Review of Economic Studies (2005) **72**, 1173–1198 0034-6527/05/00461173$02.00

The Economics of Clear Advice and Extreme Options

DEZSÖ SZALAY
HEC-University of Lausanne and FAME

First version received June 2000; *final version accepted November* 2004 (*Eds.*)

I study a principal–agent model in which the agent collects information and then chooses a verifiable action. I show that the principal can find it desirable to constrain the agent's action set even though there is no disagreement about the ranking of actions *ex post*. The elimination or penalization of "intermediate" actions, which are optimal when information is poor, improves incentives for information collection. I characterize optimal action sets when the agent is infinitely risk averse with respect to income shocks and optimal incentive schemes when the agent is risk neutral.

Give me a one-handed economist! All my economists say: "On the one hand . . . , but on the other hand . . . " [Harry Truman]

1. INTRODUCTION

Precise information is essential for good decision-making. Yet, many decision makers (henceforth, principals) have no time to acquire expertise themselves. Therefore, they rely on experts (henceforth, agents) to acquire expertise for them and communicate their information to them. Examples abound; *e.g.* politicians and CEOs have boards of advisers, faculties delegate the screening of job candidates to recruiting committees, and judges rule on behalf of societies. Delegation of expertise opens the door to at least two incentive problems. An agent can spend too little effort on information acquisition and can try to mislead the principal. How can a principal overcome these incentive problems?

In principle the answer is simple: the principal must reward the agent if the facts prove that the agent's advice was correct. But it is surprising how far away reality is from this ideal. For example, judges' pay does not depend on the quality of their verdicts. Quite the opposite, judges' salaries do not even depend on their verdicts as such. To explain how expertise is delegated in practice it is important to depart from the ideal world of perfect monetary contracts. But, if not by monetary contracts, how else can a principal guarantee high-quality decisions?

I develop a stylized model that allows me to answer this question for a class of decision problems. I begin my analysis with an abstract formulation of a communication game in which the principal commits to some choice rule before the agent acquires information and communicates it to the principal. However, since I abstract from contingent money payments, there is essentially no value to centralized decision-making. The principal can do just as well by delegating the right of choice to the agent and constraining this right appropriately (see Holmström (1978), Green and Stokey (1981), Melumad and Shibano (1991)). The question is now how to design the constraints on the agent's discretion.

On the one hand, the agent's freedom of choice influences his incentives to acquire information. On the other hand, the agent's freedom of choice also influences how the agent uses the information he has acquired. To focus on the first aspect, I assume that the principal has selected an agent who shares her opinion. If the agent's information were exogenous, this assumption would eliminate all incentive problems and unlimited discretion would be optimal. However, with endogenous information the *ex ante* incentive problem remains, because the agent cares for the private rather than the social value of information. Is it possible to mitigate this incentive problem by constraining the agent's freedom of choice? I show that the answer to this question is yes.

First, I show that the principal can improve the agent's incentives for information acquisition by *limiting* his freedom of choice. If the principal removes the compromising choices and leaves the agent with only extreme options, then the agent's utility is reduced both when he makes a decision based on prior information and with more precise information. But the reduction of the agent's expected utility is greater when he decides with prior information. Heuristically, the agent can depart into the wrong direction relative to the status quo if he does not know the true state of the world. Put differently, the utility gain when precise information becomes available is larger and the agent finds it relatively more attractive to acquire information.

Second, I show that it can be optimal to remove the compromising choices from the agent's choice set. On the one hand, limited freedom of choice can lead to suboptimal decision-making *ex post*. On the other hand, it increases incentives for information acquisition *ex ante*. The beneficial *ex ante* effect is dominant when the agent cares a lot for good decision-making relative to the costs of acquiring information. In this case, I say that *ex ante* interests are relatively well aligned. In my model a good alignment of *ex ante* interests implies that the increase in the amount of effort spent on information acquisition due to a reduction in freedom of choice is large. At the same time, the probability is small that any of the removed choices is optimal *ex post*.

Finally, I perform a comparative statics analysis of the optimal discretion offered to the agent. I show that the optimal discretion is non-monotonic in the *ex ante* alignment of interests. If the agent cares very little for the right course of action, the costs of additional incentives for effort are too large relative to the benefits thereof; if the agent cares very much for the right course of action, there is no need to provide additional incentives. In between, however, it can be optimal to limit his discretion.

My model squares well with a number of real-life situations of delegated expertise. To return to my earlier example, judges in bench trials (where judges collect evidence and apply the law) must sometimes choose between the extreme options of either acquitting or convicting. In educational institutions, the recruiting committees of some faculties are exposed to extreme options. Economics department recruiting committees can choose from the three options {reject, interview, fly-out without interview} for each candidate. In contrast, political science department committees cannot interview and must choose either to reject or to fly-out. Competition authorities must sometimes take a stand on whether or not a proposed merger is detrimental to consumers' welfare. "On the one hand . . . , but . . . " is not an admissible answer. The advantage is that these authorities think harder before they take a position. Some firms commit to cultures of activism. Such innovation biases can be interpreted as optimal incentive mechanisms: for example, division managers who are in charge of screening projects relative to the status quo have better incentives to screen among potential innovations than under a more flexible policy. Finally, my theory can explain active portfolio management. In Szalay (2004*b*) an active portfolio manager promises to not track the index. To ensure the credibility of his promise the manager accepts to pay some large fine for index tracking. The more precise the manager's information the longer (shorter) his optimal long (short) position when he receives good (bad) news. Consequently, the manager wants to track the index less often if he acquires more precise information, and active portfolio management can be explained as an incentive instrument that provides incentives for information acquisition.

My theory combines several views of delegation that have so far been studied separately. Holmström (1978, 1984) and Armstrong (1994) study constrained delegation when the agent's information is exogenous. When the principal and the agent have different preferences *ex post*, it is optimal to *exclude the extreme options*. On the other hand, Aghion and Tirole (1997) take the agent's information as endogenous, but abandon the contractibility of actions of the previous theories. They show that delegating unconstrained authority improves incentives for information acquisition. I assume that actions are contractible *and* that information is endogenous. My theory can explain why and when it is optimal to *remove the intermediate options*, which cannot be explained by either one of these views of delegation alone.

Tirole (1999) also discusses constrained delegation with endogenous information. Tirole develops the complete contracting version of the Aghion and Tirole (1997) paper to show that the notion of authority is not an artefact of incomplete contracts. He discusses delegation under which the principal keeps "gatekeeping counterpower" but does not analyse the optimal use of this power.

Monetary contracts are absent in most of these theories, with the exception of a chapter in Holmström (1978) and the second part of my paper. This feature distinguishes the theories from the literature on contracts with endogenous information structures (see Demski and Sappington (1987) for an early contribution, Laffont and Martimort (2002, Chapter 9.8) and the references cited there for an overview, Bergemann and Valimaki (2002) and Szalay (2004*a*) for more recent contributions).

My paper is related to a growing literature that combines frictions on monetary contracting and endogenous information to explain features of real-world institutions of decision-making. Dewatripont and Tirole (1999) explain "advocates" as specialized information collectors; Prendergast (1993) explains why workers should "second guess" their bosses' opinions. Most closely related in this literature is Li (2001), which explains excessive conservatism in committee decision-making. Li analyses a trade-off similar to the one in my paper, but emphasizes a different aspect of it. In his problem of committee decision-making, committee members are tempted to free-ride on others' efforts in fact-finding. He shows that a committee that is biased for hiring (rejecting) should commit to do so only if there is strong evidence in favour of (against) the candidate. The reason is that a biased rule for decisions between given options increases the marginal value of information *ex ante*.

I should emphasize that my theory and the literature I have just discussed neglect some aspects that may be important in practice, notably implicit incentives and limited commitment. An agent who has career concerns has an additional incentive to acquire information. If these implicit incentives are strong enough, there might be no need to provide still more incentives through limited discretion. In some real-life situations, it might also be possible to reconsider the decision rule after the information acquisition stage. Allowing for such reconsiderations would alter the analysis substantially. My model applies only to situations in which such a possibility is absent in practice, *e.g.* to judicial decision-making, because the law cannot be changed in every trial, to the analysis of firm policies that are deeply routed in corporate cultures, or to situations in which reputation concerns keep the principal from reneging on freedom of choice.[1]

The remainder of this article is organized as follows. Before I dive into the analysis of my model I explain the logic of my main results by means of a simple example in Section 2. Section 3 introduces the model. In Section 4 I analyse the case of an agent who is infinitely risk averse with respect to income shocks and study freedom of action as an incentive device. In Section 5 I study the case where the agent is risk neutral with respect to money payments.

1. For related models with career concerns, see, *e.g.* Prendergast and Stole (1996) or Holmström and Ricart-i-Costa (1986). For a model of communication without commitment, see Crawford and Sobel (1982), for a comparison of models with exogenous information with and without commitment, see Dessein (2002) and de Garidel-Thoron and Ottaviani (2000), and for a simultaneous consideration of cheap talk and reputational concerns, see Ottaviani and Sorensen (2000).

In Section 6 I discuss extensions of my analysis in various other directions. All proofs are relegated to the Appendix.

2. A TOY MODEL

Suppose that there are three possible states of nature, H, M, L, with prior probabilities $\Pr(H) = \frac{3}{8}$, $\Pr(M) = \frac{1}{4}$, and $\Pr(L) = \frac{3}{8}$, and that there are three possible actions, ℓ, m, r. There are two people, a principal ("she") and an agent ("he"). If the state of nature is η and the action x is chosen, each of the two parties receives the pay-off $u(\eta, x)$ indicated in the following table.

state/action	ℓ	m	r
H	0	−4	−16
M	−4	0	−4
L	−16	−4	0

For example, if the action r is chosen when the state is M, the principal and the agent receive a joint pay-off equal to −8.

The two parties are agreed on the optimal action in each state. The action ℓ is optimal in the state H, m is optimal in the state M, and r is optimal in the state L, providing each of them with a pay-off of zero. However, this scheme requires that the state be known when the action is taken. If the state is not known and information is given by the prior probabilities $\Pr(H) = \frac{3}{8}$, $\Pr(M) = \frac{1}{4}$, and $\Pr(L) = \frac{3}{8}$, both parties prefer the action m to be taken. In this case, they obtain an expected pay-off equal to −3 each.

Suppose now that the agent can learn the state by spending resources $c > 0$. In a first-best world, the agent would acquire this information if and only if the cost c is less than the aggregate benefit 6 that the principal and the agent together obtain from the improvement in action choices. In fact, if c is less than the benefit 3 that the agent himself obtains, this first-best outcome is achieved as a matter of course. However, if $c > 3$, the agent will acquire the information only if the principal provides him with an additional incentive to do so.

If actions are verifiable, one way to provide an additional incentive for information acquisition is to have a contractual clause prohibiting or penalizing the action m. If m is off limits and the state is not known, then both ℓ and r are optimal. Both actions provide the agent as well as the principal with a pay-off equal to $\frac{3}{8}(-16) + \frac{1}{4}(-4) = -7$. In contrast, when the state η is known, the elimination of m has no effect on the optimal action for $\eta = H$ or $\eta = L$. But removing m from the set of feasible choices changes the optimal action from m to either ℓ or r if $\eta = M$. The resulting pay-off in state M is $u(M, \ell) = u(M, r) = -4$, and the *ex ante* expectation of $u(\eta, x(\eta))$ is equal to −1.

When the principal eliminates m from the set of admissible actions, the agent's information acquisition depends on the comparison of the expected pay-off −7, which he gets from the action ℓ or r in the absence of information, and the difference $-1 - c$ between the expected pay-off from making the optimal choice of ℓ or r in each state and the cost of information acquisition. He is willing to acquire information whenever c is less than $-1 - (-7) = 6$. For c between 3 and 6, the prohibition of the action m thus induces an acquisition of information that would otherwise not take place. The reason is that the prohibition of the action m reduces the agent's expected pay-off in the absence of information, when m is optimal with probability 1, rather more than it reduces his pay-off with information, when m is optimal only with probability $\frac{1}{4}$.

The improvement in information acquisition incentives comes at a cost in *ex post* efficiency. In the state M, there is a welfare loss, so that the *ex ante* expected aggregate welfare is equal to $-2 - c$. For $c < 4$, this is still better than the *ex ante* expected aggregate welfare of −6 that is achieved without information. The social benefits of having the information exceed the sum of the information acquisition costs and efficiency losses arising from the prohibition of m. Therefore, for $3 < c < 4$, the two parties will find it mutually beneficial to conclude a contract which prohibits m and at the same time provides the agent with a side payment as compensation for part of the cost c.

This simple analysis applies directly to the example of judicial decision-making when there is clear evidence on the nature of a crime, but the question is whether the defendant did or did not commit this crime. The judge must spend time and effort to go through the arguments of the prosecutor and the defence. Without going through the case, the judge is unbiased. The three states correspond to the nature of the evidence. It can incriminate or exonerate the defendant, or it can be insufficient for either purpose. But there are only two verdicts available to the judge. Under current law the judge has no compromising choice available on finding that the evidence is insufficient, but has to decide "*in dubio pro reo*" and acquit the defendant. The example shows why these extreme options are optimal.

I proceed as follows. After explaining my model in the next section, I will study the optimal specification of action sets in Section 4. Subsequently, in Section 5, I will show that the basic idea is robust to the introduction of risk neutrality and a sensitivity of the agent to monetary rewards. In this case, intermediate actions are not prohibited outright, but are penalized by a reduction of the agent's monetary reward.

3. THE MODEL

I consider a principal who derives utility from a decision x according to

$$V(x, \eta) = K - \pi(x, \eta) \tag{1}$$

and an agent who derives utility from the same decision according to

$$U(x, \eta, \alpha) = \alpha(K - \pi(x, \eta)). \tag{2}$$

η is a parameter and $\pi(x, \eta)$ is a quadratic loss function

$$\pi(x, \eta) = \frac{1}{2}(x - \eta)^2. \tag{3}$$

The parameter η is the realization of a random variable $\tilde{\eta}$, that takes values in an interval $\left[\underline{\eta}, \bar{\eta}\right]$. The distribution of $\tilde{\eta}$ has a density $f(\eta)$ with full support. The expected value of $\tilde{\eta}$ is μ and $\sigma > 0$ is its standard deviation. The parameter $\alpha > 0$ measures the relative value of information to the agent. The set of *a priori* feasible actions, x, is $\left[\underline{\eta}, \bar{\eta}\right]$. I denote by $W(x, \eta)$ the joint utility that the principal and the agent derive from the decision x, *i.e.*

$$W(x, \eta) = U(x, \eta, \alpha) + V(x, \eta) = (1 + \alpha)(K - \pi(x, \eta)). \tag{4}$$

At the time of contracting, both the agent and the principal know $f(\cdot)$, but neither of them knows the realization of $\tilde{\eta}$. The agent can perform an experiment on η after the contract is written but before the action x must be chosen. If the agent exerts effort e, the experiment is a success with probability e and is unsuccessful with probability $(1 - e)$. The agent observes whether the outcome is a success or a failure. A successful experiment reveals the realization of $\tilde{\eta}$ to the agent. If the experiment fails, the agent does not acquire additional information. In other words, the agent's effort is success-enhancing in the sense of Green and Stokey (1981). The cost of an experiment that succeeds with probability e is $g(e)$, where $g(e)$ satisfies the Inada conditions:[2]

2. Throughout the paper, subscripts denote derivatives of functions.

$g_e(e) > 0\ \forall e > 0$, $g_{ee}(e) > 0\ \forall e$, $g_e(e)|_{e=0} = 0$, $\lim_{e\to 1} g_e(e) = \infty$. The agent's choice of e is not observable to the principal. If the experiment is successful, the realization of $\tilde{\eta}$ is the agent's private information, and the principal does not even observe whether the experiment is a success or not. Hence, the agent's information is soft and can be forged.

The principal commits to a mechanism before the agent performs the experiment. She specifies a message space M such that the agent can send messages $m \in M$ to the principal once he has observed the outcome of the experiment. Moreover, she commits to taking the action $x(m)$ and paying a transfer $t(m, \eta)$ to the agent if his message is m and the realized state is η. I assume that the principal must commit to this mechanism. In particular, she cannot renege on the choice scheme $x(m)$ once the agent has acquired information.

The principal is risk neutral with respect to income shocks, *i.e.* her overall utility is $V(x, \eta) - t(m, \eta)$. The agent's overall utility is $U(x, \eta, \alpha) + \upsilon(t(m, \eta)) - g(e)$. The agent is infinitely risk averse with respect to income risk. More specifically, I assume that $E\upsilon(t(m, \eta))$ is equal to the smallest realization of the transfer.[3] Infinite risk aversion implies that the agent's effort choice does not respond to monetary incentives. As a result, the agent receives a constant payment, t. The principal sets this payment so as to equalize the agent's utility from participating to the agent's outside wage, which I normalize to zero.

Consider now the nature of the incentive problem. It is obvious from (1) and (2) that there is no conflict of interest with respect to the choice of x *ex post*. However, *ex ante*, the principal and the agent disagree on the choice of e. To see this, suppose that the principal uses information efficiently *ex post* and chooses $x = \eta$ when information η is available and $x = \mu$ when only prior information is available. The marginal value of information to the agent is equal to the expected incremental utility he obtains when information η becomes available, $\frac{\alpha\sigma^2}{2}$. The agent exerts effort until the marginal cost of effort, $g_e(e)$, equals the private marginal value of information. By contrast, the social value of information includes the value of information to the principal, $\frac{\sigma^2}{2}$. Consequently, the agent exerts too little effort from a social perspective if the principal uses information efficiently *ex post*.

4. FREEDOM OF CHOICE AS AN INCENTIVE DEVICE

4.1. The contracting problem

By the revelation principle I can think of a contract as specifying a direct, incentive-compatible mechanism. The agent communicates his information to the principal and is given incentives to be truthful. Given that the principal cannot use monetary transfers, she cannot do better than give the agent the right to choose an action x out of a closed subset of $\left[\underline{\eta}, \bar{\eta}\right]$.[4] I let $\mathbf{\Gamma}$ denote the set of closed subsets of $\left[\underline{\eta}, \bar{\eta}\right]$ with typical element Γ. From this perspective, the principal's problem is to choose a closed set Γ and a constant payment t, with the interpretation that the agent is free to choose $x \in \Gamma$. I let $x(\eta)$ denote his choice of action if the experiment succeeds and $x(\phi)$ denote

3. Infinite risk aversion is the standard assumption for justifying non-monetary contracting (see Aghion and Tirole, 1997, p. 6). The specific functional form I assume here renders the agent's utility function *in equilibrium* linear in transfers, which allows me to discuss the agent's participation decision in a simple way. I relax this assumption in Section 4.

4. Holmström (1978, 1984) and Green and Stokey (1981) first observed this. I use the result in the spirit of Melumad and Shibano (1991).

his choice of action if the experiment fails. The principal's problem is[5]

$$\max_{\Gamma\in\mathbf{\Gamma},t} eEV(x(\eta),\eta) + (1-e)EV(x(\phi),\eta) - t \tag{5}$$

$$s.t.$$

$$x(\eta) \in \arg\max_{x\in\Gamma} U(x,\eta,\alpha) \quad \forall\eta \tag{6}$$

$$x(\phi) \in \arg\max_{x\in\Gamma} EU(x,\eta,\alpha) \tag{7}$$

$$EU(x(\eta),\eta,\alpha) - EU(x(\phi),\eta,\alpha) = g_e(e) \tag{8}$$

$$eEU(x(\eta),\eta,\alpha) + (1-e)EU(x(\phi),\eta,\alpha) - g(e) + t \geq 0. \tag{9}$$

The first term in (5) represents the principal's expected pay-off conditional on a successful information acquisition experiment weighted by the probability of success, e. The principal takes into account (6), *i.e.* that the agent will choose the alternative he most prefers for each realization of $\tilde{\eta}$ subject to the restriction Γ. The second term is the analogue for the case where the agent chooses x after an unsuccessful experiment. In this case the incentive compatibility condition on the agent's choice of x is (7). Equation (8) is the incentive compatibility condition for the agent's effort choice. The condition states that the agent's level of effort equates the private marginal value of information to the marginal cost. Equation (9) states that the agent must be willing to go along with the principal's contract proposal. Since the principal has quasi-linear utility and unlimited wealth and the agent's utility from money income is equal to the smallest payment, (9) is binding at the optimum and the principal maximizes expected social surplus subject to incentive compatibility of the agent's choices, $x(\phi)$, $x(\eta)$, and e.

4.2. *A characterization of optimal contracts*

In this section I show that it is optimal to have the agent choose any action he likes, with the possible exception of actions in a symmetric interval around the prior optimal one, μ. Let Γ^* denote an optimal contract. Then:

Proposition 1. *The principal's problem has a solution.* $\Gamma^* \in \mathbf{\Gamma}^D$ *where*

$$\mathbf{\Gamma}^D = \left\{\Gamma \in \mathbf{\Gamma} : \Gamma = \left[\underline{\eta}, \overline{\eta}\right] \setminus (\mu-\delta, \mu+\delta) \text{ for all } \delta \in \left[0, \min\left\{\mu-\underline{\eta}, \overline{\eta}-\mu\right\}\right]\right\}.$$

If the principal wants to increase the agent's choice of effort, she must, by (8), increase the marginal value of information to the agent, $EU(x(\eta),\eta,\alpha) - EU(x(\phi),\eta,\alpha)$. She cannot increase $EU(x(\eta),\eta,\alpha)$ relative to the case in which the agent has unlimited discretion, because the agent's expected utility conditional on a successful experiment is maximized if information is used efficiently. Therefore, she must reduce the agent's expected utility conditional on experiment failure, *i.e.* punish him if the experiment fails. Since $EU(x(\phi),\eta,\alpha)$ is decreasing in the Euclidean distance between the agent's restricted preferred act, $x(\phi)$, and his unrestricted preferred act, μ, the principal can reduce $EU(x(\phi),\eta,\alpha)$ by forcing the agent to depart from μ by, say, Δ. The choice sets she can offer to enforce this departure from μ in an incentive-compatible way must include at least one action with distance Δ to μ and must not include any actions closer to μ than this one.

Contracts in $\mathbf{\Gamma}^D$ are optimal in this set of incentive-compatible contracts because they minimize the cost of inflicting a given punishment on the agent. Prohibiting the agent from choosing actions with more than distance Δ to μ does not influence the punishment if the experiment fails, but it does prohibit him from choosing the *ex post* efficient action for some

5. Throughout the paper, E denotes the expectation operator.

realizations of η if the experiment succeeds. Consequently, the principal removes exclusively a convex set, including μ, from the agent's choice set.

Finally, the bounds of the set of feasible acts, $\underline{\eta}$ and $\overline{\eta}$, are contained in an optimal choice set. Removing an extreme policy reduces the sensitivity of $x(\eta)$ to the agent's information and reduces his incentive to acquire information. In consequence, the set of removed actions is also symmetric around μ.

To prove existence of an optimal solution I show that Holmström's (1978) existence proof can be used in the present context as well. These arguments show that Γ^* exists and that $\Gamma^* \in \mathbf{\Gamma}^D$.

At the optimum, the principal excludes a convex, symmetric set from the agent's choice set regardless of the distribution of $\tilde{\eta}$. The convexity result contrasts with Holmström's (1978) analysis of delegation contracts with preference divergence and given information in which convex prohibitions are optimal only if the distribution of $\tilde{\eta}$ is uniform. The difference is due to the simplicity of my agency problem *ex post* as well as the simplicity of the information acquisition technology. My symmetry result is due to the fact that pay-off functions depend only on the distance between the restricted and the unrestricted choice of x. In a more general analysis with more general utility functions, the optimal prohibition might well be asymmetric. But as long as the information acquisition technology is success-enhancing, the prohibited set will still be convex.

4.3. *The economics of extreme options*

I now show that the prohibited set can be nonempty. By Proposition 1, the principal's problem is to choose the bounds, $\mu \pm \Delta$, of a symmetric interval around the mean. Because of this symmetry I can write expected losses conditional on the agent's information and conditional on incentive compatibility of his choices under contracts of the optimal structure as

$$E\pi\ (x(\eta), \eta) = \frac{1}{2}\int_{\mu}^{\mu+\Delta} (\mu + \Delta - \eta)^2 dF(\eta) + \frac{1}{2}\int_{\mu-\Delta}^{\mu} (\mu - \Delta - \eta)^2 dF(\eta)$$
$$and \tag{10}$$
$$E\pi\ (x(\phi), \eta) = \frac{1}{2}(\sigma^2 + \Delta^2).$$

To simplify notation I let $E\pi\ (\Delta, \eta) := E\pi\ (x(\eta), \eta)$ and $E\pi\ (\Delta, \phi) := E\pi\ (x(\phi), \eta)$. I use analogous notation for the agent's, the principal's, and joint expected utility, respectively. From (8), the agent's incentive-compatible effort choice is

$$e\,(\Delta, \alpha) = g_e^{-1}\,[\alpha\,(E\pi\ (\Delta, \phi) - E\pi\ (\Delta, \eta))] \tag{11}$$

where $g_e^{-1}(\cdot)$ exists because $g(e)$ is strictly convex.

Proposition 2. *$e\,(\Delta, \alpha)$ is strictly increasing in Δ for $\Delta > 0$.*

Forcing the agent to depart from the prior optimal choice decreases expected conditional utilities conditional on both failure and success, because the principal cannot observe the outcome of the experiment. But the reduction of the level of expected utility conditional on failure is always greater than the reduction of the level of expected utility conditional on success. Conditional on an unsuccessful experiment, the agent would like to choose the prior optimal action with probability 1 and must deviate by Δ from this action. Conditional on a successful experiment, the agent's choice of x is affected only in the event that $\eta \in (\mu - \Delta, \mu + \Delta)$. Moreover, the induced deviation from the agent's unrestricted preferred action is almost surely smaller than Δ. Therefore, the principal can use the agent's inability to choose freely to provide incentives.[6]

Using (11) I can write the principal's problem, (5) s.t. (6)–(9), as the unconstrained maximization problem

$$\max_{\Delta} P(\Delta, \alpha) = (1 - e(\Delta, \alpha)) EW(\Delta, \phi, \alpha) + e(\Delta, \alpha) EW(\Delta, \eta, \alpha) - g(e(\Delta, \alpha)). \quad (12)$$

The principal's expected utility is equal to expected joint welfare net of costs of information acquisition. Heuristically, the principal faces the following trade-off. On the one hand, an increase in Δ increases (decreases) the probability that the experiment succeeds (fails), but on the other hand it decreases expected welfare conditional both on failure and success of the experiment and it increases the cost of information acquisition. To understand this trade-off analytically, I take the derivative of $P(\Delta, \alpha)$ with respect to Δ, and use (10) and (11) to write

$$\begin{aligned} P_{\Delta}(\Delta, \alpha) = & \{-(1 - e(\Delta, \alpha))(1 + \alpha) E\pi_{\Delta}(\Delta, \phi) - e(\Delta, \alpha)(1 + \alpha) E\pi_{\Delta}(\Delta, \eta)\} \\ & + \left[\frac{g_e(e(\Delta, \alpha))}{g_{ee}(e(\Delta, \alpha))} (E\pi_{\Delta}(\Delta, \phi) - E\pi_{\Delta}(\Delta, \eta)) \right]. \end{aligned} \quad (13)$$

The expression in {} is the marginal cost of a small increase in Δ, the expected marginal increase in *ex post* welfare losses due to removing some more actions around μ. The expression in [] is the marginal benefit of a small increase in Δ, the marginal increase in the principal's expected utility due to a small increase in the agent's effort. The increase in effort increases (reduces) the probability that the relatively higher (lower) expected utility from decision-making with precise (prior) information is realized. For ease of expression, I will call these terms the marginal costs and benefits of Δ, respectively.

Observe that the marginal benefit of Δ is inversely proportional to the curvature of the cost of effort function. The reason is that the agent is the more susceptible to incentives the less his marginal cost of effort increases with effort.

Before I solve problem (12) I address the question of when it is optimal to increase Δ away from zero. To answer this question I use (10) to evaluate the marginal costs and benefits of Δ at $\Delta = 0$. Both expressions are equal to zero at $\Delta = 0$, so $P_{\Delta}(\Delta, \alpha)|_{\Delta=0} = 0$. Therefore, $P(\Delta, \alpha)$ is increasing in Δ for small but positive Δ if and only if $P(\Delta, \alpha)$ is convex around the origin.

Proposition 3. *Let Δ^* denote a solution to problem* (12). *Δ^* is strictly positive,* i.e. *the principal benefits from introducing small action restrictions, if $P_{\Delta\Delta}(\Delta, \alpha)|_{\Delta=0} > 0$ or equivalently if*

$$\left. \left(-(1 - e(\Delta, \alpha))(1 + \alpha) + \frac{g_e(e(\Delta, \alpha))}{g_{ee}(e(\Delta, \alpha))} \right) \right|_{\Delta=0} > 0. \quad (14)$$

Specifically, let the cost of effort be $\hat{g}(e) = \beta g(e)$ where $\beta > 0$ and suppose $g(e)$ satisfies (a) $\lim_{e \to 1} \frac{\partial}{\partial e} \left(\frac{g_e(e)}{g_{ee}(e)} \right) < -(1 + \alpha)$ and (b) $\frac{g_e(e)}{g_{ee}(e)}$ concave in e. Then, for any strictly positive values of the parameters α and σ, there exists a unique $\beta' > 0$ such that (14) *is satisfied iff $\beta \in (0, \beta')$. Conditions (a) and (b) are satisfied by,* e.g.

$$g(e) = ((1 - e) \ln(1 - e) + e).$$

6. Again, it is interesting to contrast my result with Holmström's. He observed that giving more freedom of action to the agent in the sense of giving him more discretion to choose among extreme options provides the agent with more of an incentive to acquire information (see Holmström, 1978, p. 98). The same is true for my problem (see Szalay, 2000). However, in contrast to the case with conflicting interests, there is no reason to prohibit the agent from choosing extreme actions in the first place.

To obtain condition (14), I differentiate (13) with respect to Δ, and note that all first-order effects are zero around $\Delta = 0$ and that $E\pi_{\Delta\Delta}(\Delta, \phi)|_{\Delta=0} = 1$ and $E\pi_{\Delta\Delta}(\Delta, \eta)|_{\Delta=0} = 0$. The only nonvanishing effects of a small increase of the restricted area around μ, say by $d\Delta$, are related to the increase in the marginal loss in the case where the experiment fails: the expected increase in the marginal welfare cost, $(1 - e(\Delta, \alpha))(1 + \alpha)d\Delta$, and the increase in the marginal value of successful experimenting to the principal, $\frac{g_e(e(\Delta,\alpha))}{g_{ee}(e(\Delta,\alpha))}d\Delta$. It is optimal to introduce action restrictions if the former effect is small and the latter is large. Economically, this requires that the agent be *well motivated already in the absence of restrictions* and that he also be *susceptible to additional incentives*.

To demonstrate that condition (14) is nonvacuous, I introduce the parameter β, which allows me to do comparative statics on the agent's effort choice on the L.H.S. of condition (14). If β is close to zero this choice of effort is close to one. Because of the Inada assumption this implies that the agent is not at all susceptible to additional incentives. Formally, $g_{ee}(e)$ goes to infinity faster than $g_e(e)$ as e approaches one. On the other hand, the expected increase in the marginal welfare cost of increasing Δ is also zero when e is close to one. Hence, the L.H.S. of condition (14) goes to zero as e gets close to one (β gets close to zero). If condition (a) is satisfied, the agent's susceptibility to incentives improves faster with a decrease in e (an increase in β) than the probability of experiment failure increases with a decrease in e. Condition (b) is quite natural for Inada functions. Taken together, conditions (a) and (b) imply that (14) is satisfied if the agent is well, but not extremely well, motivated.

4.4. The optimal size of restrictions

I now investigate how large optimal restrictions should be. This analysis is complicated by the fact that problem (12) is not necessarily everywhere concave in Δ. Therefore, I introduce two simplifications to obtain analytical results. I investigate local conditions and study the smallest local maximizer of the principal's pay-off function. In addition I drop either the Inada assumption or the agent's individual rationality constraint. These simplifications allow me to obtain comparative statics results of optimal restrictions with respect to the *ex ante* alignment of interests. For a subset of parameter values, I can show that problem (12) has a unique solution, which coincides with the smallest local maximizer. Finally, I derive explicit solutions for a specific distribution of types to quantify the size of optimal restrictions without any of these simplifications.

4.4.1. Alignment of *ex ante* interests and the size of restrictions. If $g_{eee}(e) \geq 0$, then the difference between the first-best and the second-best level of effort at $\Delta = 0$ is decreasing in α and I can interpret the parameter α as a measure of the *ex ante* alignment of interests. I begin with a heuristic discussion of the effects of such an alignment of interests on the marginal cost and the marginal benefit of Δ. The effect on the marginal cost of Δ is ambiguous. On the one hand, it increases $e(\Delta, \alpha)$ and shifts more weight towards the relatively lower marginal loss conditional on a successful experiment. On the other hand, it increases marginal welfare losses both conditional on failure and on success of the experiment. The effect on the marginal benefit of Δ is non-monotonic in α. The agent is reluctant to exert any effort when α is small however large Δ is. For α very large $e(\Delta, \alpha)$ is close to one. In this case the agent is reluctant to increase his effort choice when Δ is increased because marginal costs of effort rise very fast. For intermediate values of α, however, the agent is susceptible to incentives.

The interplay of these effects is too complex. However, the picture becomes clear if I drop the Inada condition and assume that the marginal costs of effort are bounded for all e. Let $\underline{\Delta}$ denote the smallest local maximizer of the principal's pay-off function in the following sense:

in the case where the set $\{\Delta : P_{\Delta}(\Delta, \alpha) = 0 \text{ and } P_{\Delta\Delta}(\Delta, \alpha) < 0\}$ is nonempty, $\underline{\Delta}$ is the smallest element in this set; otherwise $\underline{\Delta}$ is the smallest element in the set $\{\Delta : e(\Delta, \alpha) = 1\}$.

Proposition 4. *Suppose that* $\lim_{e\to 1} g_e(e) < \infty$ *and let* $\frac{g_e(e)}{g_{ee}(e)}$ *be nondecreasing in* e.[7] *Then, there exist* α', α'', *and* α''' *(defined explicitly in the Appendix) satisfying* $\alpha''' > \alpha'' > \alpha' > 0$, *such that:*

(i) $\underline{\Delta}$ *is equal to zero for* $\alpha \leq \alpha'$, *increasing in* α *for* $\alpha' < \alpha \leq \alpha''$, *decreasing in* α *for* $\alpha'' < \alpha < \alpha'''$, *and equal to zero for* $\alpha \geq \alpha'''$.
(ii) $e(\underline{\Delta}, \alpha)$ *is nondecreasing in* α *for all* α. $e(\underline{\Delta}, \alpha)$ *is equal to 1 for all* $\alpha \geq \alpha''$.
(iii) *If* $\alpha \geq \alpha''$ *then* $\underline{\Delta} = \Delta^*$.

For $\alpha < \alpha'$ the principal's pay-off function is concave and decreasing in Δ for Δ small but positive, and small action restrictions are unattractive. The marginal cost of Δ is relatively high at $\Delta = 0$, because the experiment fails with a high probability and the marginal benefit of Δ is relatively small, because the agent responds badly to incentives. An increase in α improves this cost–benefit comparison for restrictions of infinitesimal size to the point that it eventually becomes optimal to introduce action restrictions. Formally, $P(\Delta, \alpha)$ is strictly convex around the origin for $\alpha > \alpha'$. Likewise, in this range an increase in α increases the marginal benefit of Δ by more than it increases the marginal cost of Δ, so that larger restrictions are optimal. Since marginal costs of effort are bounded, corner solutions are possible and the optimal contract induces the agent to experiment successfully with probability 1 at $\alpha = \alpha''$. Even better motivated agents will also be induced to succeed with probability 1, but the principal can implement this effort level with smaller-size restrictions. Very well-motivated agents succeed with probability 1 even in the absence of restrictions. Finally, for $\alpha \geq \alpha''$ the principal's pay-off is increasing in Δ until $\underline{\Delta}$. Hence, the corner solution is the unique optimal solution to the contracting problem.

I obtain a second tractable case if I drop the agent's participation constraint. In this case I can reimpose the more appealing Inada assumption.

Suppose the agent is protected by limited liability and that K is sufficiently larger than $\frac{\sigma^2}{2}$. In this case, the agent is willing to participate even if he receives the lowest possible wage, $t = 0$. The principal's pay-off in this case is $P(\Delta, \alpha) = (1 - e(\Delta, \alpha))EV(\Delta, \phi) + e(\Delta, \alpha)EV(\Delta, \eta)$. It is easy to verify that the set $\{\Delta : P_{\Delta}(\Delta, \alpha) = 0 \text{ and } P_{\Delta\Delta}(\Delta, \alpha) < 0\}$ is nonempty. Let $\underline{\Delta}$ denote the smallest element of this set. It satisfies the first-order condition

$$\left.\begin{pmatrix} -(1 - e(\Delta, \alpha))\, E\pi_{\Delta}(\Delta, \phi) - e(\Delta, \alpha)\, E\pi_{\Delta}(\Delta, \eta) \\ + \frac{g_e(e(\Delta, \alpha))}{g_{ee}(e(\Delta, \alpha))}\left(E\pi_{\Delta}(\Delta, \phi) - E\pi_{\Delta}(\Delta, \eta)\right) \end{pmatrix}\right|_{\Delta = \underline{\Delta}} = 0.$$

Proposition 5. *Suppose that the agent is protected by limited liability. Suppose also that* K *is sufficiently larger than* $\frac{\sigma^2}{2}$, *so that the agent's IR constraint is nonbinding. Suppose further that* $g(e)$ *satisfies* $\lim_{e\to 1} \frac{\partial}{\partial e}\left(\frac{g_e(e)}{g_{ee}(e)}\right) < -1$ *and* $\frac{g_e(e)}{g_{ee}(e)}$ *concave in* e. *Then, the qualitative features of Proposition* 4 *remain intact except that* $e(\underline{\Delta}, \alpha)$ *is bounded away from one.*

7. The property "$\frac{g_e(e)}{g_{ee}(e)}$ is nondecreasing in e" is possessed by convex power functions, e^n, exponentials, $\exp(e)$, and products of these two classes. Sums of (two) power functions $e^n + e^l$ display the property provided that $|n - l|$ is small. As an example consider a quadratic cost function and let $\frac{\sigma^2}{2} = 1$. One can easily verify that (14) holds if $1 > \alpha > \frac{\sqrt{5}-1}{2}$.

For the now-familiar reasons, the principal does not use action restrictions when α is small. They become attractive when α is large enough, provided that a technical condition analogous to the one in Proposition 3, is satisfied. For relatively low levels of the implemented effort level, the agent's susceptibility to incentives, $\frac{g_e(e(\underline{\Delta},\alpha))}{g_{ee}(e(\underline{\Delta},\alpha))}$, improves with increase in effort. Hence, an increase in α increases the marginal benefit of Δ. On the other hand, it decreases the marginal cost of Δ, because it increases the probability that the relatively small marginal loss after a successful experiment is realized. To balance these effects, $\underline{\Delta}$ is increased. Eventually, a further increase in α makes the agent less susceptible to incentives, that is, moves $e\left(\underline{\Delta},\alpha\right)$ into the range where $\frac{g_e(e)}{g_{ee}(e)}$ is decreasing in effort. To balance the resulting reduction in the marginal benefit of Δ, $\underline{\Delta}$ is decreased again. Finally, the principal's pay-off function is concave in Δ beyond $\underline{\Delta}$ when a further increase of effort beyond $e\left(\underline{\Delta},\alpha\right)$ decreases the agent's susceptibility to incentives. Therefore $\underline{\Delta} = \Delta^*$ for large enough α, *i.e.* I pick up the global maximizer of the principal's problem.

4.4.2. The case for the very extreme options. Although a comparative statics analysis on the alignment of interests is not possible in the full model, I can of course calculate the size of the optimal restriction if I assume a specific cost of effort function, say the one in Proposition 3, and a specific distribution of types. Suppose that η is distributed on $[-1, 1]$, for the sake of the argument, potentially with mass points $\frac{1-p}{2}$ at the extremes. For $p < 1$ ($p = 1$) let the distribution of η have uniform density $f(\eta) = \frac{p}{2}$ on $(-1, 1)$ ($[-1, 1]$). Suppose the parameters in the example take values $\alpha = 1$ and $\beta = \frac{1}{16}$. $\underline{\Delta}$ satisfies the first-order condition

$$\left.\left(\begin{array}{c} -(1-e(\Delta,\alpha))(1+\alpha)E\pi_\Delta(\Delta,\phi) - e(\Delta,\alpha)(1+\alpha)E\pi_\Delta(\Delta,\eta) \\ +\frac{g_e(e(\Delta,\alpha))}{g_{ee}(e(\Delta,\alpha))}\left(E\pi_\Delta(\Delta,\phi) - E\pi_\Delta(\Delta,\eta)\right) \end{array}\right)\right|_{\Delta=\underline{\Delta}} = 0. \quad (15)$$

If $p = 1$, the optimal restriction, calculated from (15) rules out roughly 4·5% of the decision space. In contrast, if a lot of mass is concentrated at the bounds, say for $p = 0{\cdot}01$, the optimal contract rules out about 17% of the decision space. Finally, there is $\underline{p} > 0$ such that for $p \leq \underline{p}$, it is optimal to rule out everything but the most extreme choices, *i.e.* $\Delta^* = 1$.

This example shows that the optimal restriction is relatively small when all type realizations are equally likely. But the optimal restriction can get extremely large when intermediate choices are much less likely than the extreme ones.

5. MONETARY SANCTIONS AND MOTIVATION

I now turn to the opposite extreme where the agent is risk neutral with respect to income shocks. Now the results depend crucially on whether performance and messages are contractible or whether only messages are contractible. In the former case, which Osband (1989) has studied, it is possible to implement the first-best. The solution involves a transfer that reflects the principal's pay-off, *i.e.* a transfer scheme $t\left(\hat{\eta},\eta\right) = \bar{t} - \pi\left(x\left(\hat{\eta}\right),\eta\right)$. Confronted with this transfer the agent has the correct incentive to acquire information as well as to communicate information truthfully and the first-best is achieved. The result is due to the agent having unlimited wealth (and being risk neutral with respect to money payments).

I study the case where performance is noncontractible.

By the revelation principle, I can restrict attention to direct mechanisms that give the agent the incentive to tell the truth. The principal's problem is a straightforward extension of problem (5) s.t. (6)–(9), with the additional complication that transfers may depend on the agent's recommendation. I let $\{x(\eta), t(\eta)\}$ denote the contract tuple offered to the agent in the case where he announces that the experiment was a success and that the realization of $\tilde{\eta}$ was η. $\{x(\phi), t(\phi)\}$ denotes the contract offered to the agent in the case where he announces that the experiment failed. Formally, the principal's problem is

$$\max_{\substack{x(\eta),t(\eta),\\ x(\phi),t(\phi),e}} (1-e)E\left[V\left(x\left(\phi\right),\eta\right)-t\left(\phi\right)\right]+eE\left[V\left(x(\eta),\eta\right)-t(\eta)\right] \tag{16}$$

$$s.t.$$

$$\begin{aligned} \forall\eta: \quad U\left(x(\eta),\eta,\alpha\right)+t\left(\eta\right) &\geq U\left(x(\phi),\eta,\alpha\right)+t(\phi) \\ \forall\eta: \quad U\left(x(\eta),\eta,\alpha\right)+t\left(\eta\right) &\geq U\left(x\left(\hat{\eta}\right),\eta,\alpha\right)+t\left(\hat{\eta}\right) \quad \forall\hat{\eta} \end{aligned} \tag{17}$$

$$E\left[U\left(x(\phi),\eta,\alpha\right)+t\left(\phi\right)\right] \geq E\left[U\left(x\left(\hat{\eta}\right),\eta,\alpha\right)+t\left(\hat{\eta}\right)\right] \quad \forall\hat{\eta} \tag{18}$$

$$E\left[U\left(x(\eta),\eta,\alpha\right)+t\left(\eta\right)\right]-E\left[U\left(x(\phi),\eta,\alpha\right)+t(\phi)\right]=g_e(e) \tag{19}$$

$$(1-e)E\left[U\left(x(\phi),\eta,\alpha\right)+t\left(\phi\right)\right]+eE\left[U\left(x(\eta),\eta,\alpha\right)+t(\eta)\right]\geq g(e). \tag{20}$$

The principal gives the agent incentives to truthfully announce the failure of the experiment (18), and to truthfully announce the true realization of η if he knows it, rather than any other realization of $\tilde{\eta}$ or that the experiment failed (17). The agent's effort choice is determined by the marginal value of information that the contract provides (19), and the agent is willing to participate (20). Maximization is performed—as is usual—with respect to piecewise C^1 functions $x(\eta)$ and $t(\eta)$.

It is easy to see that no contract with transfers depending exclusively on reports can implement the first-best:

Lemma 1. *There exists no contract that is* ex post *and* ex ante *efficient.*

Ex post efficiency requires that the agent receives the same transfer for all values of η that he announces, if he announces that the experiment was a success. To see this, consider local incentive compatibility of the agent's announcement conditional on a successful experiment, *i.e.* $\alpha(x(\hat{\eta})-\eta)\frac{\partial x(\hat{\eta})}{\partial\hat{\eta}} = \frac{\partial t(\hat{\eta})}{\partial\hat{\eta}}$ *a.e. Ex post* efficiency requires that the L.H.S. of this equation be zero. In consequence, the transfer scheme must be flat where it is differentiable. By global incentive compatibility, it must be continuous. Otherwise some types would have an incentive to misrepresent their types. Because $x(\phi) = x(\mu) = \mu$ in an *ex post* efficient contract, the only remaining possibility for increasing the agent's incentive to exert effort is to set $t(\phi)$ sufficiently smaller than $t(\mu)$. But if the same decision x is taken for reports ϕ and μ, the agent can always claim to be the type that would receive the higher transfer. But these arguments imply that the transfer scheme cannot be used at all to give the agent more of an incentive to exert effort. Consequently, the first-best cannot be achieved by any contract. Conversely, giving extra incentives for information acquisition implies a departure from the *ex post* efficient use of information.[8]

By quasi-linearity of utilities the agent's individual rationality constraint is binding at the optimum and the principal maximizes joint surplus subject to incentive compatibility of the

8. As a referee pointed out, first-best would be implementable by a Crémer and McLean (1988) type mechanism if the principal also acquired information with a certain probability (which I rule out). Then, the principal could give the agent a very low transfer if he announced some type $\hat{\eta} = \eta_1$ while in fact his experiment was a failure, but the principal's succeeded and revealed that $\eta = \eta_2 \neq \eta_1$. In consequence, $\tau(\phi)$ can be decreased relative to $\tau(\eta)$ for all η, while $x(\eta) = \eta$ is implemented for all η without giving the agent an incentive to misrepresent his type.

agent's choices. The solution procedure essentially parallels the one presented in Section 3. To avoid duplication, I present the solution directly.[9]

Proposition 6. *The principal implements an* ex post *inefficient contract. Suppose that* $\frac{\partial f(\eta)}{\partial \eta} \geq 0$ *for* $\eta \leq \mu$ *and that* $\frac{\partial f(\eta)}{\partial \eta} \leq 0$ *for* $\eta > \mu$ *and that* $g_{eee}(e) \geq 0$ $\forall e$. *Then, the optimal contract has the following properties:*

(a) *The action* $x(\phi)$ *is* ex post *efficient. The function* $x(\eta)$ *is strictly increasing with a discontinuity at the mean. Formally,*

$$x(\phi) = \mu$$
$$x(\eta) = \eta - \gamma(\alpha)\frac{F(\eta)}{f(\eta)} \text{ for } \eta \leq \mu$$
$$x(\eta) = \eta + \gamma(\alpha)\frac{1 - F(\eta)}{f(\eta)} \text{ for } \eta > \mu.$$

(b) *The transfer scheme is decreasing in* η *for* $\eta \leq \mu$, *increasing in* η *for* $\eta > \mu$, *and displays a discontinuity at the mean. Formally,*

$$t(\phi) = T - \alpha\gamma(\alpha)\int_{\underline{\eta}}^{\mu}\frac{F(z)}{f(z)}dz$$
$$t(\eta) = T + \frac{\alpha}{2}\left(\gamma(\alpha)\frac{F(\eta)}{f(\eta)}\right)^2 - \alpha\gamma(\alpha)\int_{\underline{\eta}}^{\eta}\frac{F(z)}{f(z)}dz \qquad \text{for } \eta \leq \mu$$
$$t(\eta) = T + \frac{\alpha}{2}\left(\gamma(\alpha)\frac{1 - F(\eta)}{f(\eta)}\right)^2 + \alpha\gamma(\alpha)\int_{\underline{\eta}}^{\eta}\frac{1_{z>\mu} - F(z)}{f(z)}dz \text{ for } \eta > \mu$$

where

$$\gamma(\alpha) = \left.\left(\frac{\frac{g_e(e)}{\alpha} - \frac{\sigma^2}{2}}{\int_{\underline{\eta}}^{\bar{\eta}}\frac{(1_{\eta>\mu} - F(\eta))^2}{f(\eta)}d\eta}\right)\right|_{e=e^*}.$$

(c) e^* *is uniquely defined by the equation*

$$\gamma(\alpha) = \left.\left(\frac{\frac{\alpha+1}{2}\sigma^2 - g_e(e)}{\frac{1+\alpha}{2}\left(\frac{g_e(e)}{\alpha} - \frac{\sigma^2}{2}\right) + e\frac{1+\alpha}{\alpha}g_{ee}(e)}\right)\right|_{e=e^*}.$$

For α *smaller than some critical value (defined in the Appendix),* e^* *is an optimal effort level in the local sense. For* α *large,* e^* *is the unique optimal effort level in the global sense.*

Figure 1 illustrates the optimal contract.

The principal now has two incentive instruments to hand for making experiment failure relatively unattractive to the agent. The principal can increase either the wedge between the

9. 1_{Φ} is equal to 1 if the statement in Φ is true, and 0 if not. T is the optimal indirect utility level given to type $\underline{\eta}$, $u^*\left(\underline{\eta}\right)$. It is derived from the binding IR constraint. $T = g\left(e^*\right) - \alpha K - \left(1 - e^*\right)\frac{\alpha\sigma^2}{2} + \alpha\gamma(\alpha)\left\{\left(1 - e^*\right)\int_{\underline{\eta}}^{\mu}\frac{F(\eta)}{f(\eta)}d\eta + e^*\int_{\underline{\eta}}^{\bar{\eta}}\frac{\left(1_{\eta>\mu} - F(\eta)\right)^2}{f(\eta)}d\eta\right\}$.

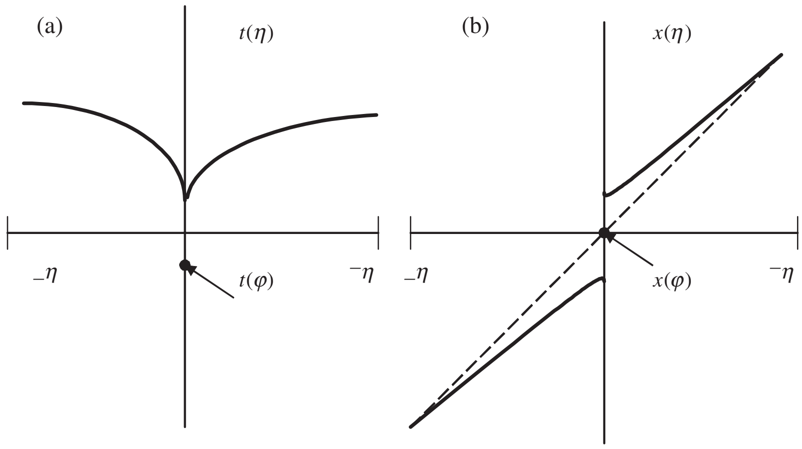

FIGURE 1

Diagram (a) shows the transfer scheme. The agent is punished with a low transfer for a failure of the experiment and the transfer is the higher the farther his type is from the mean. Diagram (b) depicts the optimal choice function $x(\cdot)$. If the experiment fails, the *ex post* efficient act is implemented. If the experiment succeeds, $x(\eta)$ is closer to the bounds than the *ex post* efficient choice for all η except at the bounds of the support, where the *ex post* efficient action is implemented.

agent's expected utility from decision-making conditional on a successful and an unsuccessful experiment, respectively, or/and the analogous wedge in his expected income. Since the two instruments are substitutes it is not necessary to implement an inefficient choice of action if the experiment fails in order to give the agent an added incentive to acquire information. The principal can simply reduce $t(\phi)$.

The principal indeed prefers to punish the agent financially rather than by implementing an inefficient action if the experiment fails. In other words, an optimal contract involves $x(\phi) = \mu$. The reason is the following. *Ex ante* the principal expects to receive message ϕ with the discrete probability $1 - e$, whereas any other message has likelihood $ef(\eta)$. Therefore, the *ex ante* cost of departing from μ if the message is ϕ is infinitely higher than the expected cost of departing from *ex post* efficiency for any η. In consequence, it is always optimal to implement the efficient choice of action conditional on experiment failure, and use the transfer scheme instead of the decision schedule to make experiment failure relatively unattractive to the agent.

It is always optimal to give the agent more of an incentive to acquire information. The reason is precisely that the optimal contract involves no efficiency loss if the experiment fails. Therefore, the *ex post* cost of introducing distortions is smaller by an order of magnitude than the *ex ante* gain that arises through the beneficial effect on the agent's incentive to acquire more information.

Consider now the form of the decision schedule $x(\cdot)$. If the agent is punished with a low transfer contingent on experiment failure, then the truth-telling constraint of type ϕ implies that $x(\eta)$ must be bounded away from *ex post* efficient choices for statements close to μ. In this way, type ϕ is confronted with a choice between the *ex post* efficient action and a low transfer if he tells the truth, and an inefficient choice of action together with a relatively higher transfer if he claims to be type μ or a type close to but different from μ. By local incentive compatibility of the contract, the distortion spreads out over the whole support of the distribution, with the bounds being an exception. The size of the distortion at a given type η reflects a trade-off between increasing the joint *ex post* loss and increasing the agent's marginal information rent.

Under the standard conditions on hazard rates, $\frac{\partial}{\partial \eta}\frac{F(\eta)}{f(\eta)} \leq 1$ for $\eta < \mu$ and $\frac{\partial}{\partial \eta}\frac{1-F(\eta)}{f(\eta)} \geq -1$ for $\eta > \mu$, which are equivalent to the stated condition on the density, the contract implementing a locally optimal effort level is strictly monotonic. Under a relatively mild additional condition on $g(\cdot)$, the principal's objective function is strictly concave in effort for α large enough. Consequently, the contract I characterize implements the unique optimal effort level in the global sense for α large enough. Unfortunately, however, the solution is too involved to allow for a comparative statics analysis.

The second-best contract no longer allows for the interpretation that the agent is forced to depart from the prior optimal choice. If he announces that he has not acquired new information, then the principal follows the efficient course of action. But like the case without transfers, the principal goes too far in following the agent's advice for some reports: it is optimal to implement $x(\eta) > \eta$ for $\eta > \mu$ and $x(\eta) < \eta$ for $\eta \leq \mu$, except at the bounds of the support. Some actions are never taken in equilibrium, as the contract is discontinuous at the mean. However, the set of actions that are not taken in equilibrium is no longer convex but rather consists of two convex sets, separated by the isolated point, μ. Finally, there is no bunching in the contract.[10]

6. EXTENSIONS

It is possible to generalize the model with infinite risk aversion with respect to money income in a number of directions.

Consider alternative location experiments and let the agent have a normally distributed prior about the true state of the world, η, and have him acquire a normally distributed signal, whose deterministic precision is increasing in effort. Some of my results are robust to this extension, some are not. In particular, the incentives for information acquisition for an agent who has been forbidden to choose actions in a set $[\mu - \Delta, \mu + \Delta]$ are increasing in Δ. To see why, observe that the agent would like to choose x *ex post* equal to (or as close as possible to) his conditional expected value of η. For normal location experiments, the conditional expectation is a weighted average of the prior mean and the signal. The higher its precision, the higher the weight on the signal. From the *ex ante* perspective, the conditional expectation function is itself a normally distributed random variable. The higher the agent's effort level, the higher its variance. On the one hand, the signal is more precise, which tends, *ceteris paribus*, to decrease the variance in the *ex ante* distribution. On the other hand, the agent puts more weight on the signal relative to the prior in forming the conditional expectation, which tends to increase the variance in the *ex ante* distribution. For normal location experiments the second effect dominates over the first. Therefore, the higher the agent's effort level, the less likely it is that the agent wants to choose an action close to the prior optimal one. In consequence, a higher effort helps the agent to avoid the punishment and the principal can use clear advice as an incentive device. Moreover, the principal may benefit from using the incentive instrument. It proves very difficult to characterize the optimal choice set for normal location experiments. The reason is that the agent's effort no longer shifts mass away from a mass point, but rather from a whole set around

10. This result is interesting from a technical perspective, because Lewis and Sappington (1993) have a result that takes the opposite direction. The reason for this discrepancy is that in contrast to theirs, my agent is benevolent. As it is here in my paper, in their procurement problem, the possibility of ignorance induces a strong desire for implementing the efficient course of action in the case where the agent confesses ignorance and the principal is tempted to increase the quantity bought towards the efficient amount. However, this desire runs *counter* to the desire to limit the agent's informational rents, which induces the principal to buy less than would be efficient. In my problem, there is no conflict between these goals because the optimal choice is distorted downwards for $\eta < \mu$ and upwards for $\eta > \mu$. Consequently, the efficient choice at the mean does not conflict with monotonicity in the contract and there is no bunching under the usual conditions.

the prior optimal action. The possibilities of providing incentives are much richer than for success-enhancing technologies.

Consider the role of identical preferences *ex post*. Two points are important here. First, the formulation is more general than in the case of exogenously given identical objectives. My application to portfolio management (Szalay, 2004*b*) makes it clear that a congruence of *ex post* objectives arises naturally if the agent does not care for performance per se, but receives a payment that is linear in performance. Second, one can allow for diverging objectives. In Szalay (2000), I consider the case in which the agent's bias is not known to the principal, but the principal knows that the agent's preferred action is correlated with hers. If the correlation is large enough, the principal benefits from removing actions around the prior optimal action from the agent's choice set. If the correlation is low, the agent's information is of no use to the principal, because the agent would not use it in the principal's interest. Therefore, it is not optimal to provide extra incentives by banning the prior optimal action and actions close to it. But the principal benefits from constraining the agent's right to choose among extreme actions in this case. The optimal extreme bounds on the agent's choice set display the Aghion–Tirole trade-off of initiative vs. loss of control.

Consider the role of the continuity of the action space. In Szalay (2002) I consider a model in which the agent faces the discrete choice of whether to innovate or not. The manager chooses between a risky and a safe option, *e.g.* an uncertain innovation and the safe status quo. The manager screens, by exerting effort, innovation paths. Imposing an innovation bias may increase *ex ante* expected pay-offs, although the manager has to gamble sometimes, because he puts in more effort to screen the innovation paths.

Finally, one may wonder about in what sense the results of the model with monetary transfers depend on the information acquisition technology. For instance, one can model information acquisition in a tractable way by assuming that the agent acquires a deterministic signal s that is correlated with the state η. The higher the agent's effort is, the more informative the signal is about the state. For concreteness, consider the multiplicative linear environment studied by Ottaviani and Sorensen (2001). The signal conditional on the state and the agent's effort is distributed according to the density $f(s\,|\eta, e) = \frac{1+es\eta}{2}$, with η and s each in $[-1, 1]$. Assume that the marginals are uniform on $[-1, 1]$. Then, one can show:[11] (i) it is optimal to depart from *ex post* efficient use of information; (ii) the optimal contract displays exaggeration: for statements $\hat{\eta} > 0$ one has $x\left(\hat{\eta}\right) > E[\eta\,|s, e]$, for statements $\hat{\eta} < 0$ one has $x\left(\hat{\eta}\right) < E[\eta\,|s, e]$. I conclude that the exaggeration feature is a robust finding of my model. The discontinuity of the contract that I obtain in the present paper is due to the success-enhancing information acquisition technology.

7. CONCLUSION

Previous theories of delegation show that excluding extreme options may be optimal to correct for differences in opinions (Holmström (1978, 1984), Armstrong (1994)), or that when information collection is important (Aghion and Tirole, 1997), increasing freedom of choice may be optimal even when there are differences in opinions. However, these theories cannot explain why compromising choices should be excluded.

I show that excluding compromising choices increases incentives for information acquisition and may increase the *ex ante* expected utility from decision-making. I also discuss a number of applications.

11. Results are available from the author upon request.

I do not discuss multi-agent situations. One interesting possibility would be to study rules of arbitration, as analysed, *e.g.* by Gibbons (1988). Gibbons (1988) studies arbitration from the perspective of providing incentives to the disputing parties to make reasonable offers. My analysis suggests that rules of arbitration may also be important for providing incentives for effort to the arbitrator. Eliminating the possibility of compromise may be good for the arbitrator's incentive to find out which of two opposing positions is closer to the truth. I believe that this is an interesting avenue for research and leave it to future work.

APPENDIX

Proof of Proposition 1. I use a two-step sequential maximization procedure to find the solution to the principal's problem. In the first step, I take as given that the principal wants to implement an $x(\phi)$ such that $|x(\phi)-\mu|=\Delta$. In the second step, I search for the optimal Δ within the set of maximizers to the Step 1 problem. In the present proof, I discuss the existence of a solution to the overall problem (Steps 1 and 2) and characterize the solution to the Step 1 problem. I characterize the solution to the Step 2 problem in the subsequent propositions.

By quasi-linearity of pay-offs and unlimited wealth the principal maximizes joint pay-off subject to incentive compatibility of the agent's choices. The principal's problem, given that $|x(\phi)-\mu|=\Delta$, is

$$\begin{gathered}\max_{\Gamma\in\mathbf{\Gamma}} EW(x(\phi),\eta,\alpha)+e\,(EW(x(\eta),\eta,\alpha)-EW(x(\phi),\eta,\alpha))-g(e)\\ s.t.\\ x(\eta)\in\arg\max_{x\in\Gamma}U(x,\eta,\alpha)\quad\forall\eta\\ x(\phi)\in\arg\max_{x\in\Gamma}EU(x,\eta,\alpha)\\ |x(\phi)-\mu|=\Delta\\ EU(x(\eta),\eta,\alpha)-EU(x(\phi),\eta,\alpha)=g_e(e).\end{gathered}$$

A convenient, and equivalent, way of writing this problem is to replace its constraints 2 and 3 with a restriction on the admissible set of choice sets. The choice set offered to the agent must be in the following set of sets:

$$\mathbf{\Gamma^{\Delta}}=\left\{\Gamma\in\mathbf{\Gamma}:\Gamma\text{ is a closed subset of }\Gamma^{\Delta}\text{ and }\mu+(-)\,\Delta\notin\Gamma\text{ implies }\mu-(+)\,\Delta\in\Gamma\right\}$$

where

$$\Gamma^{\Delta}=\begin{cases}\left[\underline{\eta},\overline{\eta}\right]\setminus(\mu-\Delta,\mu+\Delta) & for\ \Delta\in\left[0,\min\left\{\mu-\underline{\eta},\overline{\eta}-\mu\right\}\right]\\ [\mu+\Delta,\overline{\eta}] & for\ \Delta>\min\left\{\mu-\underline{\eta},\overline{\eta}-\mu\right\}\quad if\ \overline{\eta}-\mu>\mu-\underline{\eta}\\ \left[\underline{\eta},\mu-\Delta\right] & for\ \Delta>\min\left\{\mu-\underline{\eta},\overline{\eta}-\mu\right\}\quad if\ \overline{\eta}-\mu<\mu-\underline{\eta}.\end{cases}$$

To see the equivalence of the two formulations, note that $\Gamma\in\mathbf{\Gamma^{\Delta}}$ is a necessary condition on a choice set that implements an $x(\phi)$ with distance Δ from μ. The reason is that the agent chooses $x(\phi)$ as close as possible to μ if the experiment fails, because $EU(x(\phi),\eta,\alpha)$ is decreasing in $|x(\phi)-\mu|$. It is obvious that $\Gamma\in\mathbf{\Gamma^{\Delta}}$ is also sufficient for implementing an $x(\phi)$ with distance Δ from μ. Hence, $\Gamma\in\mathbf{\Gamma^{\Delta}}$ is necessary and sufficient for constraints 2 and 3 in the program above.

To obtain a further simplification I use the incentive compatibility of the choices $x(\phi)$, $x(\eta)$, and e to rewrite the principal's objective function. Using $(EW(x(\eta),\eta,\alpha)-EW(x(\phi),\eta,\alpha))=\frac{1+\alpha}{\alpha}(EU(x(\eta),\eta,\alpha)-EU(x(\phi),\eta,\alpha))$, I find that the Step 1 problem has the equivalent representation:

$$\begin{gathered}\max_{\Gamma\in\mathbf{\Gamma^{\Delta}}} EW(x(\phi),\eta,\alpha)+e\frac{1+\alpha}{\alpha}g_e(e)-g(e)\\ s.t.\\ x(\eta)\in\arg\max_{x\in\Gamma}U(x,\eta,\alpha)\quad\forall\eta\\ e=g_e^{-1}[EU(x(\eta),\eta,\alpha)-EU(x(\phi),\eta,\alpha)]\end{gathered}$$

where $x(\phi)$ is taken as fixed. Observe that the objective function is strictly increasing in e and that e is strictly increasing in $EU(x(\eta),\eta,\alpha)$ for given $x(\phi)$. It follows that the principal's problem in Step 1 has a solution iff the problem

$$\begin{gathered}\max_{\Gamma\in\mathbf{\Gamma^{\Delta}}} EU(x(\eta),\eta,\alpha)\\ s.t.\ x(\eta)\in\arg\max_{x\in\Gamma}U(x,\eta,\alpha)\ \forall\eta\end{gathered}$$

has a solution. This problem corresponds to Holmström's (1978) problem for the case in which the principal and the agent have utility function $U(\cdot)$ and the set of admissible choice sets is $\mathbf{\Gamma^{\Delta}}$. As he showed, a solution exists. Trivially, the solution is the set Γ^{Δ}. Also trivially, it is unique. To see this, suppose it were not unique and that there were two closed sets that solve this problem. Since the two sets would achieve the same value of the objective, they could differ only on measure 0 sets. By full support of $f(\eta)$, measure 0 sets are isolated points. But then at least one of the sets would have to be open, *i.e.* would not be an admissible choice set.

Next, I prove that optimal choice set must satisfy $|x(\phi) - \mu| < \min\left\{\mu - \underline{\eta}, \overline{\eta} - \mu\right\}$. I prove this result for the case $\mu - \underline{\eta} < \overline{\eta} - \mu$. The proof of the reverse case is analogous and omitted. Consider the incremental loss when $x(\eta)$ and $x(\phi)$ are chosen according to (6) and (7), respectively, and $|x(\phi) - \mu| = \Delta$. By straightforward algebra,

$$\begin{aligned} E\pi\,(x(\phi), \eta) - E\pi\,(x\,(\eta)\,, \eta) \quad &= \quad \tfrac{1}{2}(\sigma^2 + \Delta^2) - \tfrac{1}{2}\textstyle\int_{\mu}^{\mu+\Delta}(\mu + \Delta - \eta)^2 dF(\eta) \\ &\qquad - \tfrac{1}{2}\textstyle\int_{\mu-\Delta}^{\mu}(\mu - \Delta - \eta)^2 dF\,(\eta) \qquad if\ \Delta < \mu - \underline{\eta} \\ &= \quad \tfrac{1}{2}(\sigma^2 + \Delta^2) - \tfrac{1}{2}\textstyle\int_{\underline{\eta}}^{\mu+\Delta}(\mu + \Delta - \eta)^2 dF(\eta) \qquad else. \end{aligned}$$

By straightforward calculus one finds that $(\sigma^2 + \Delta^2) - \int_{\underline{\eta}}^{\mu+\Delta}(\mu + \Delta - \eta)^2 dF\,(\eta)$ is decreasing convex in Δ (with slope 0 at $\Delta = \overline{\eta} - \mu$). Moreover, one can show that

$$\begin{aligned} &\left(\Delta^2 - \int_{\mu}^{\mu+\Delta}(\mu + \Delta - \eta)^2 dF(\eta) - \int_{\mu-\Delta}^{\mu}(\mu - \Delta - \eta)^2 dF(\eta)\right)\Bigg|_{\Delta=\mu-\underline{\eta}} \\ &> \left(\Delta^2 - \int_{\underline{\eta}}^{\mu+\Delta}(\mu + \Delta - \eta)^2 dF(\eta)\right)\Bigg|_{\Delta=\mu-\underline{\eta}}. \end{aligned}$$

Hence, the maximal implementable effort level under a contract with $|x(\phi) - \mu| = \min\{\mu - \underline{\eta}, \overline{\eta} - \mu\}$ is higher than the maximal effort level that is implementable by any contract with $|x(\phi) - \mu| > \min\{\mu - \underline{\eta}, \overline{\eta} - \mu\}$. Since $EW(x(\phi), \eta, \alpha)$ is decreasing in Δ, this proves that any $\Delta > \min\{\mu - \underline{\eta}, \overline{\eta} - \mu\}$ is suboptimal. Hence $\Delta \in \left[0, \min\left\{\mu - \underline{\eta}, \overline{\eta} - \mu\right\}\right]$. I show in the text that the principal's pay-off when x and e satisfy incentive compatibility and the contract is Γ^{Δ} is a continuous function of Δ. Since a continuous function on a compact domain attains a maximum, a solution to problem (5) s.t. (6)–(9) exists. ‖

Proof of Proposition 2. From (10), the difference between expected losses conditional on experiment failure and conditional on success is

$$E\pi\,(\Delta, \phi) - E\pi\,(\Delta, \eta) = \frac{(\sigma^2 + \Delta^2)}{2} - \frac{\int_{\mu}^{\mu+\Delta}(\mu + \Delta - \eta)^2 dF(\eta) + \int_{\mu-\Delta}^{\mu}(\mu - \Delta - \eta)^2 dF(\eta)}{2}. \tag{A.1}$$

Differentiating (11), using Leibniz's rule and the inverse function theorem I find that

$$e_{\Delta}\,(\Delta, \alpha) = \frac{\alpha}{g_{ee}(e)}\left\{\Delta - \int_{\mu}^{\mu+\Delta}(\mu + \Delta - \eta) dF(\eta) + \int_{\mu-\Delta}^{\mu}(\mu - \Delta - \eta) dF(\eta)\right\}. \tag{A.2}$$

Hence, $e_{\Delta}\,(\Delta, \alpha) > 0$ iff $\Delta - \int_{\mu}^{\mu+\Delta}(\mu + \Delta - \eta) dF\,(\eta) + \int_{\mu-\Delta}^{\mu}(\mu - \Delta - \eta) dF\,(\eta) > 0$. Let now $\mu + Z^{+}\,(\Delta) := E\,[\eta|\,\eta \in [\mu, \mu + \Delta]]$ and $\mu - Z^{-}\,(\Delta) := E\,[\eta|\,\eta \in [\mu - \Delta, \mu]]$. Using these definitions I can write

$$\begin{aligned} &\Delta - \int_{\mu}^{\mu+\Delta}(\mu + \Delta - \eta) dF\,(\eta) + \int_{\mu-\Delta}^{\mu}(\mu - \Delta - \eta) dF(\eta) \\ &= \Delta - (F\,(\mu + \Delta) - F\,(\mu - \Delta))\,\Delta + (F\,(\mu + \Delta) - F(\mu))\,Z^{+}\,(\Delta) + (F(\mu) - F\,(\mu - \Delta))\,Z^{-}\,(\Delta) \\ &\geq (F\,(\mu + \Delta) - F(\mu))\,Z^{+}\,(\Delta) + (F(\mu) - F\,(\mu - \Delta))\,Z^{-}\,(\Delta) \\ &\geq 0 \end{aligned}$$

where the last inequality is strict if $\Delta > 0$, because $Z^{+}\,(\Delta)$ and $Z^{-}\,(\Delta)$ are strictly positive by construction. ‖

Proof of Proposition 3. Using (11), the principal's pay-off function is

$$P\,(\Delta, \alpha) = EW\,(\Delta, \phi, \alpha) + e\,(\Delta, \alpha)\,\frac{1 + \alpha}{\alpha} g_e\,(e\,(\Delta, \alpha)) - g\,(e\,(\Delta, \alpha))\,.$$

Differentiating $P(\Delta, \alpha)$ with respect to Δ, using (A.1), (A.2), and the envelope theorem, I obtain

$$P_{\Delta}(\Delta, \alpha) = \left(e(\cdot)(1+\alpha) + \frac{g_e(e(\cdot))}{g_{ee}(e(\cdot))}\right)\left\{\begin{array}{c}\Delta - \int_{\mu}^{\mu+\Delta}(\mu+\Delta-\eta)dF(\eta) \\ +\int_{\mu-\Delta}^{\mu}(\mu-\Delta-\eta)dF(\eta)\end{array}\right\} - \Delta(1+\alpha).$$

Observe that $P_{\Delta}(\Delta, \alpha)|_{\Delta=0} = 0$; $\Delta = 0$ is a stationary point. Differentiating a second time I obtain

$$P_{\Delta\Delta}(\Delta, \alpha) = -1(1+\alpha) + \left(e(\cdot)(1+\alpha) + \frac{g_e(e(\cdot))}{g_{ee}(e(\cdot))}\right)\left\{1 - \int_{\mu}^{\mu+\Delta} dF(\eta) - \int_{\mu-\Delta}^{\mu} dF(\eta)\right\}$$
$$+ \frac{\partial}{\partial e}\left(e(\cdot)(1+\alpha) + \frac{g_e(e(\cdot))}{g_{ee}(e(\cdot))}\right)\frac{\alpha\left\{\Delta - \int_{\mu}^{\mu+\Delta}(\mu+\Delta-\eta)dF(\eta) + \int_{\mu-\Delta}^{\mu}(\mu-\Delta-\eta)dF(\eta)\right\}^2}{g_{ee}(e(\cdot))}.$$

At $\Delta = 0$ I have $P_{\Delta\Delta}(\Delta, \alpha)|_{\Delta=0} = -1(1+\alpha) + \left(e(\Delta, \alpha)(1+\alpha) + \frac{g_e(e(\Delta,\alpha))}{g_{ee}(e(\Delta,\alpha))}\right)\Big|_{\Delta=0}$. Hence, iff the condition in the proposition is satisfied, $P_{\Delta\Delta}(\Delta, \alpha)|_{\Delta=0} > 0$ and $\Delta = 0$ is a local minimum.

I now consider the example with $e\left(\Delta, \frac{\alpha}{\beta}\right) = g_e^{-1}\left[\frac{\alpha}{\beta}(E\pi(\Delta, \phi) - E\pi(\Delta, \eta))\right]$. The parameter β serves to do comparative statics on the effort level in (14) leaving everything else unchanged: $e\left(\Delta, \frac{\alpha}{\beta}\right) = g_e^{-1}\left[\frac{\alpha}{\beta}(E\pi(\Delta, \phi) - E\pi(\Delta, \eta))\right]$ is decreasing in β and $\frac{\hat{g}_e(e)}{\hat{g}_{ee}(e)} = \frac{g_e(e)}{g_{ee}(e)}$. Let $y(e) := \frac{g_e(e)}{g_{ee}(e)} - (1+\alpha)(1-e)$, the expression on the L.H.S. of (14). By assumption, $y(e)$ is concave in e. Also, $y(e)|_{e=0} = -(1+\alpha)$. I now show that $\lim_{e\to 1} y(e) = 0$: since $\lim_{e\to 1} g_e(e) = \infty$ and $g_{ee}(e) > 0\ \forall e$, $g_e(e)$ must be convex for e close to one; hence, for e close to one, $g_{ee}(e) \geq e g_{ee}(e) > g_e(e)$; hence $\lim_{e\to 1}\frac{g_e(e)}{g_{ee}(e)} = 0$ and hence $\lim_{e\to 1} y(e) = 0$. Economically, as e goes to one, the need to provide incentives diminishes, because the agent exerts the socially optimal level of effort.

I now write $y\left(e\left(\Delta, \frac{\alpha}{\beta}\right)\right)$ to take care of the dependence of e on β. By straightforward calculus,

$$y_{\beta}\left(e\left(\Delta, \frac{\alpha}{\beta}\right)\right) = \left((2+\alpha) - \frac{g_e\left(e\left(\Delta, \frac{\alpha}{\beta}\right)\right) g_{eee}\left(e\left(\Delta, \frac{\alpha}{\beta}\right)\right)}{\left(g_{ee}\left(e\left(\Delta, \frac{\alpha}{\beta}\right)\right)\right)^2}\right) e_{\beta}\left(\Delta, \frac{\alpha}{\beta}\right).$$

Since $\lim_{\beta\to 0} e\left(\Delta, \frac{\alpha}{\beta}\right) = 1$ and $\frac{\partial e\left(\Delta, \frac{\alpha}{\beta}\right)}{\partial\beta} < 0\ \forall\beta$, $\lim_{e\to 1}\frac{g_e(e) g_{eee}(e)}{(g_{ee}(e))^2} > 2+\alpha$ implies that $\lim_{\beta\to 0}\frac{\partial y\left(e\left(\Delta, \frac{\alpha}{\beta}\right)\right)}{\partial\beta} > 0$. Hence $y\left(e\left(\Delta, \frac{\alpha}{\beta}\right)\right) > 0$ for β positive but small. Since $y(e)$ is concave in e and $e\left(\Delta, \frac{\alpha}{\beta}\right)$ monotonic in β, there is a unique β' such that $\frac{g_e\left(e\left(\Delta, \frac{\alpha}{\beta'}\right)\right)}{g_{ee}\left(e\left(\Delta, \frac{\alpha}{\beta'}\right)\right)}\Bigg|_{\Delta=0} = (1+\alpha)\left(1 - e\left(\Delta, \frac{\alpha}{\beta'}\right)\right)\Big|_{\Delta=0}$. Hence (14) is satisfied iff $\beta \in (0, \beta')$. ‖

Proof of Proposition 4. First, I prove that the solution to the principal's problem is characterized by the first-order condition for α small. Second, I show that the size of the optimal restriction is nondecreasing in α if the solution is characterized by the first-order condition. Third, I show that the solution to the principal's problem is a corner solution for α large enough. Finally, I show that the size of the optimal restriction for large enough α (in a corner solution) is nonincreasing in α.

The principal's problem has an interior solution for α small. From (15), an interior solution satisfies

$$\left(\left(e(\Delta, \alpha) + \frac{1}{1+\alpha}\frac{g_e(e(\Delta, \alpha))}{g_{ee}(e(\Delta, \alpha))}\right)(E\pi_{\Delta}(\Delta, \phi) - E\pi_{\Delta}(\Delta, \eta)) - \Delta\right)\Bigg|_{\Delta=\underline{\Delta}} = 0.$$

I define $S(\Delta, \alpha) = e(\Delta, \alpha) + \frac{1}{1+\alpha}\frac{g_e(e(\Delta,\alpha))}{g_{ee}(e(\Delta,\alpha))}$ and let $S(0, \alpha) = \left(e(\Delta, \alpha) + \frac{1}{1+\alpha}\frac{g_e(e(\Delta,\alpha))}{g_{ee}(e(\Delta,\alpha))}\right)\Big|_{\Delta=0}$. Clearly, $S(0, 0) = 0$. Since $S(0, \alpha)$ is continuous in α, $\underline{\Delta}$ is characterized by the first-order condition for small but positive α.

Using $e_{\alpha}(\Delta, \alpha) = \frac{1}{\alpha}\frac{g_e(e(\Delta,\alpha))}{g_{ee}(e(\Delta,\alpha))}$ I find

$$S_{\alpha}(\Delta, \alpha) = \frac{\partial}{\partial\alpha}\left(e(\Delta, \alpha) + \frac{1}{1+\alpha}\frac{g_e(e(\Delta, \alpha))}{g_{ee}(e(\Delta, \alpha))}\right)$$
$$= \left(\frac{g_e(e(\Delta, \alpha))}{g_{ee}(e(\Delta, \alpha))}\left(\frac{1}{\alpha} - \frac{1}{(1+\alpha)^2}\right) + \frac{1}{1+\alpha}\frac{\partial}{\partial e}\left(\frac{g_e(e(\Delta, \alpha))}{g_{ee}(e(\Delta, \alpha))}\right)e_{\alpha}(\Delta, \alpha)\right).$$

It follows from monotonicity of $\frac{g_e(e)}{g_{ee}(e)}$ that $S_\alpha(\Delta,\alpha) > 0$ for all Δ. In turn, this result has two implications. First, since (14) is satisfied iff $S(0,\alpha) > 1$, it implies that there is α' defined by $S(0,\alpha') = 1$ such that (14) is satisfied if $\alpha > \alpha'$. Second, by the implicit function theorem, if $\underline{\Delta}$ satisfies the first-order and the second-order condition, then it is nondecreasing in α.

It follows further that $e(\underline{\Delta},\alpha)$ is increasing in α when $\underline{\Delta}$ is characterized by the first-order condition. Since $(g_e(e))|_{e=1} < \infty$, $e(\underline{\Delta},\alpha)$ tends to one as α increases. Define $\alpha'' = \min\{\alpha : e(\underline{\Delta},\alpha) = 1\}$. From $S_\alpha(\Delta,\alpha) > 0$ for all Δ, it follows that $\underline{\Delta}$ cannot satisfy the first-order condition for $\alpha > \alpha''$. Consequently, $e(\underline{\Delta},\alpha) = 1$ for all $\alpha \geq \alpha''$. The optimal solution is the smallest restriction that implements an effort level equal to one. Since $e(\Delta,\alpha)$ is increasing in α, $\underline{\Delta}$ is decreasing in α. α''' is defined as the smallest α such that $e(0,\alpha) = 1$. Finally, $\underline{\Delta} = \Delta^*$ for $\alpha \geq \alpha''$ because $P(\Delta,\alpha)$ is increasing in Δ until $\underline{\Delta}$. ‖

Proof of Proposition 5. In this case

$$P(\Delta,\alpha) = EV(\Delta,\alpha) + \frac{e(\Delta,\alpha)}{\alpha} g_e(e(\Delta,\alpha))$$

and

$$P_\Delta(\Delta,\alpha) = \left(e(\cdot) + \frac{g_e(e(\cdot))}{g_{ee}(e(\cdot))}\right)(E\pi_\Delta(\Delta,\phi) - E\pi_\Delta(\Delta,\eta)) - \Delta. \tag{A.3}$$

From the condition $P_{\Delta\Delta}(\Delta,\alpha)|_{\Delta=0}$ I find that $y(e) = \frac{g_e(e)}{g_{ee}(e)} - (1-e)$. If $\lim_{e\to 1} \frac{g_e(e)g_{eee}(e)}{(g_{ee}(e))^2} > 2$, then $\lim_{e\to 1} \frac{\partial y(e)}{\partial e} < 0$. Since $y(e)$ is concave in e and $e(\Delta,\alpha)$ monotonic in α, there is a unique α' such that $\left.\frac{g_e(e(\Delta,\alpha'))}{g_{ee}(e(\Delta,\alpha'))}\right|_{\Delta=0} = (1-e(\Delta,\alpha'))|_{\Delta=0}$. Hence (14) is satisfied iff $\alpha > \alpha'$.

I now consider the comparative statics of the optimal solution. $\underline{\Delta}$ satisfies both the first-order and the second-order condition for an optimum, *i.e.* $P_\Delta(\Delta,\alpha)|_{\Delta=\underline{\Delta}} = 0$ and $P_{\Delta\Delta}(\Delta,\alpha)|_{\Delta=\underline{\Delta}} < 0$. In what follows I use the notation $P_\Delta(\underline{\Delta},\alpha) = P_\Delta(\Delta,\alpha)|_{\Delta=\underline{\Delta}}$. By the implicit function theorem,

$$\frac{\partial\underline{\Delta}}{\partial\alpha} = \frac{P_{\Delta\alpha}(\underline{\Delta},\alpha)}{-P_{\Delta\Delta}(\underline{\Delta},\alpha)}. \tag{A.4}$$

Clearly, $-P_{\Delta\Delta}(\underline{\Delta},\alpha) > 0$. Consider now the numerator:

$$P_{\Delta\alpha}(\underline{\Delta},\alpha) \propto \frac{\partial}{\partial e}\left(e(\underline{\Delta},\alpha) + \frac{g_e(e(\underline{\Delta},\alpha))}{g_{ee}(e(\underline{\Delta},\alpha))}\right)\frac{\partial e(\underline{\Delta},\alpha)}{\partial\alpha}$$

where $\propto$ means "proportional to". Since $e + \frac{g_e(e)}{g_{ee}(e)}$ is concave in e, $e + \frac{g_e(e)}{g_{ee}(e)}$ is maximized for e'' that satisfies $2 = \left.\frac{g_e(e)g_{eee}(e)}{(g_{ee}(e))^2}\right|_{e=e''}$ and therefore $\frac{\partial}{\partial e}\left(e + \frac{g_e(e)}{g_{ee}(e)}\right) \lesseqgtr \Leftrightarrow e \gtreqless e''$. Thus, I need to show that $\frac{de(\underline{\Delta},\alpha)}{d\alpha} > 0$. By straightforward differentiation of (11) and some manipulations, using (A.4), I obtain

$$\frac{de}{d\alpha} = \frac{\partial e(\underline{\Delta},\alpha)}{\partial\alpha} + \frac{\partial e(\underline{\Delta},\alpha)}{\partial\Delta}\frac{\partial\underline{\Delta}}{\partial\alpha} > 0$$
$$\Leftrightarrow 1 - \left(e(\underline{\Delta},\alpha) + \frac{g_e(e(\underline{\Delta},\alpha))}{g_{ee}(e(\underline{\Delta},\alpha))}\right)(E\pi_{\Delta\Delta}(\Delta,\phi) - E\pi_{\Delta\Delta}(\Delta,\eta)) > 0$$
$$\Leftrightarrow 1 - \frac{\Delta(E\pi_{\Delta\Delta}(\Delta,\phi) - E\pi_{\Delta\Delta}(\Delta,\eta))}{E\pi_\Delta(\Delta,\phi) - E\pi_\Delta(\Delta,\eta)} > 0.$$

The last equivalence follows from using the condition $P_\Delta(\Delta,\alpha)|_{\Delta=\underline{\Delta}} = 0$ to substitute the term $\frac{\Delta}{E\pi_\Delta(\Delta,\phi)-E\pi_\Delta(\Delta,\eta)}$ for the term $\left(e(\underline{\Delta},\alpha) + \frac{g_e(e(\underline{\Delta},\alpha))}{g_{ee}(e(\underline{\Delta},\alpha))}\right)$. Since $(E\pi_\Delta(\Delta,\phi) - E\pi_\Delta(\Delta,\eta))|_{\Delta=0} = 0$ and $(E\pi_\Delta(\Delta,\phi) - E\pi_\Delta(\Delta,\eta))$ is increasing concave for all $\Delta > 0$, the result follows. Therefore, there is a unique α'', associated with e'' by the equation $e'' = e(\underline{\Delta},\alpha'')$, such that $\frac{\partial\underline{\Delta}}{\partial\alpha} > 0$ for $\alpha \in (\alpha',\alpha'']$ and $\frac{\partial\underline{\Delta}}{\partial\alpha} < 0$ for $\alpha > \alpha''$.

Finally, I show that $\underline{\Delta} = \Delta^*$ for α large enough. To see this, consider the second derivative of $P(\Delta,\alpha)$ with respect to Δ:

$$P_{\Delta\Delta}(\Delta,\alpha) = \underbrace{-1 + \left(e(\cdot) + \frac{g_e(e(\cdot))}{g_{ee}(e(\cdot))}\right)(E\pi_{\Delta\Delta}(\Delta,\phi) - E\pi_{\Delta\Delta}(\Delta,\eta))}_{=X(\Delta,\alpha)}$$
$$\underbrace{+\frac{\partial}{\partial e}\left(e(\cdot) + \frac{g_e(e(\cdot))}{g_{ee}(e(\cdot))}\right)\frac{\alpha\,(E\pi_{\Delta}(\Delta,\phi) - E\pi_{\Delta}(\Delta,\eta))^2}{g_{ee}(e(\cdot))}}_{=Y(\Delta,\alpha)}.$$

Since the agent's effort is increasing in Δ for all Δ it follows that $Y(\Delta,\alpha) \leq 0$ for all $\Delta \geq \underline{\Delta}$ if $\alpha \geq \alpha''$. Consider next $X(\Delta,\alpha)$. Using the first-order condition, $P_{\Delta}(\Delta,\alpha)\big|_{\Delta=\underline{\Delta}} = 0$, and the fact that $E\pi_{\Delta}(\Delta,\phi) - E\pi_{\Delta}(\Delta,\eta)$ is increasing concave in Δ and satisfies $(E\pi_{\Delta}(\Delta,\phi) - E\pi_{\Delta}(\Delta,\eta))|_{\Delta=0} = 0$ I observe that $X\left(\underline{\Delta},\alpha\right) < 0\ \forall\alpha$. Moreover, $X(\Delta,\alpha)$ is decreasing in Δ for $\Delta \geq \underline{\Delta}$ if $\alpha \geq \alpha''$, because $\frac{\partial}{\partial e}\left(e(\cdot) + \frac{g_e(e(\cdot))}{g_{ee}(e(\cdot))}\right) \leq 0$ for $e \geq e''$ and $E\pi_{\Delta}(\Delta,\phi) - E\pi_{\Delta}(\Delta,\eta)$ is increasing concave in Δ. Hence, $P_{\Delta\Delta}(\Delta,\alpha) < 0$ for all $\Delta \geq \underline{\Delta}$. ‖

Proof of Proposition 6. The proof has two parts. In the first part I derive an equivalent, more tractable formulation of the principal's problem, which I solve in the second part.[12] ‖

Part I: "Equivalent formulation"

Lemma I. *Problem* (16) *s.t.* (17), (18), (19), *and* (20) *is solved if and only if the following problem is solved:*

$$\max_{\substack{x(\eta)\\ x(\phi),e}} (1+\alpha)K + e\left\{\int_{\underline{\eta}}^{\bar{\eta}}\left(-\frac{\alpha+1}{2}(x(\eta)-\eta)^2\right)dF(\eta)\right\} - (1-e)\frac{\alpha+1}{2}\left((x(\phi)-\mu)^2+\sigma^2\right)$$
$$-g(e) + \lambda\left\{\int_{\underline{\eta}}^{\bar{\eta}}\alpha\,(x(\eta)-\eta)\left(1_{\eta>\mu} - F(\eta)\right)d\eta + \frac{\alpha\sigma^2}{2} - g_e(e)\right\} \tag{M}$$
$$s.t.\ \frac{\partial x(\eta)}{\partial\eta} \geq 0;\ x(\eta) \leq x(\phi)\forall\eta<\mu;\ x(\eta) \geq x(\phi)\forall\eta>\mu.$$

Proof of Lemma I. Since (20) is binding, the objective function is the expected joint surplus net of the cost of information acquisition. Let $u(\eta)$ denote the indirect utility of type η. I show that the truth-telling constraints are equivalent to requiring that

$$u(\eta) = u\left(\underline{\eta}\right) + \int_{\underline{\eta}}^{\eta}\alpha\,(x(z)-z)\,dz \tag{A.5}$$

$$Eu(\phi) = u(\mu) - \alpha\frac{\sigma^2}{2} \tag{A.6}$$

and in addition, $\frac{\partial x(\eta)}{\partial\eta} \geq 0$, $x(\eta) \leq x(\phi)$ for $\eta < \mu$, and $x(\eta) \geq x(\phi)$ for $\eta > \mu$. The effort incentive constraint follows then using (A.5), (A.6) and an integration by parts. I proceed in five steps. For ease of reference I call the incentive constraint in the first line of (17) (17a) and the incentive constraint in the second line (17b).

(i) Equation (A.5) and $\frac{\partial x(\eta)}{\partial\eta} \geq 0$ are necessary and sufficient for (17b): (standard). The first-order condition, $\frac{\partial}{\partial\hat{\eta}}\left(U\left(x\left(\hat{\eta}\right),\eta,\alpha\right) + t\left(\hat{\eta}\right)\right)\big|_{\hat{\eta}=\eta} = 0$, is necessary for an optimal report. Taking this in combination with the envelope theorem, one has $\frac{\partial u(\eta)}{\partial\eta} = \alpha\,(x(\eta)-\eta)$. Equation (A.5) results from integrating this condition. If $\frac{\partial x(\eta)}{\partial\eta} \geq 0$, then the local first-order condition is also sufficient for (17b).

12. The development of the solution concept builds on Lewis and Sappington (1993) and Crémer, Khalil and Rochet (1998). Both the Lewis and Sappington and the Crémer *et al*. studies examine versions of a procurement problem. In Lewis and Sappington, the adverse selection problem is complicated by the possibility of the agent's ignorance. In their model, they take the probability of ignorance as exogenous. In Crémer *et al*., information acquisition is endogenous. Information acquisition involves a discrete cost and $e \in \{0, 1\}$, so in equilibrium, the agent is either completely informed or completely ignorant. In my analysis, the adverse selection problem is complicated by the simultaneous presence of both problems.

(ii) Necessity of (A.6): First, I take (18) at $\hat{\eta} = \mu$, rearrange it, and integrate out to get

$$\frac{\alpha}{2}(x(\mu)-\mu)^2 + \frac{\alpha\sigma^2}{2} - \frac{\alpha}{2}(x(\phi)-\mu)^2 - \frac{\alpha\sigma^2}{2} \geq \tau(\mu) - \tau(\phi).$$

Second, I take (17a) for type $\eta = \mu$ and rearrange it to get

$$\tau(\mu) - \tau(\phi) \geq \frac{\alpha}{2}(x(\mu)-\mu)^2 - \frac{\alpha}{2}(x(\phi)-\mu)^2.$$

The two inequalities can hold simultaneously only if

$$\tau(\mu) - \tau(\phi) = \frac{\alpha}{2}(x(\mu)-\mu)^2 - \frac{\alpha}{2}(x(\phi)-\mu)^2. \tag{A.7}$$

Hence

$$-\int \frac{\alpha}{2}(x(\phi)-\eta)^2 dF(\eta) + \tau(\phi) + \frac{\alpha}{2}(x(\mu)-\mu)^2 - \tau(\mu) = Eu(\phi) - u(\mu) = -\alpha\frac{\sigma^2}{2}.$$

(iii) Necessity of $x(\eta) \leq x(\phi)$ for $\eta < \mu$, and $x(\eta) \geq x(\phi)$ for $\eta > \mu$: First, I consider (18) for every $\hat{\eta}$. After integrating out and rearranging, this condition is equivalent to

$$-\frac{\alpha}{2}x(\phi)^2 + \frac{\alpha}{2}x(\hat{\eta})^2 + \alpha\mu\left(x(\phi) - x(\hat{\eta})\right) \geq \tau(\hat{\eta}) - \tau(\phi)\ \forall\hat{\eta}.$$

Second, I consider (17a) and rearrange it to get

$$-\frac{\alpha}{2}x(\eta)^2 + \frac{\alpha}{2}x(\phi)^2 + \alpha\eta\left(x(\eta) - x(\phi)\right) \geq \tau(\phi) - \tau(\eta)\ \forall\eta.$$

I take $\hat{\eta} = \eta$, add the two inequalities and obtain

$$(\mu - \eta)\left(x(\phi) - x(\eta)\right) \geq 0.$$

(iv) Sufficiency: Type η's utility from report ϕ is $U(x(\phi), \eta, \alpha) + t(\phi) = -\frac{\alpha}{2}(x(\phi)-\mu)^2 + \tau(\phi) - \alpha(x(\phi)-\mu)(\mu-\eta) - \frac{\alpha}{2}(\eta-\mu)^2$. Using (A.5) and (A.6), truth-telling is better than deviating to reporting type ϕ when

$$\int_{\underline{\eta}}^{\eta} \alpha(x(z)-z)\,dz \geq \int_{\underline{\eta}}^{\mu} \alpha(x(z)-z)\,dz - \alpha(x(\phi)-\mu)(\mu-\eta) - \frac{\alpha}{2}(\eta-\mu)^2.$$

Integrating out and simplifying, this inequality is equivalent to

$$\int_{\mu}^{\eta} \alpha\left[x(z) - x(\phi)\right]dz \geq 0.$$

Hence, if $x(\eta) > x(\phi)$ for $\eta > \mu$ and $x(\eta) < x(\phi)$ for $\eta < \mu$ and $\frac{\partial x(\eta)}{\partial \eta} \geq 0$, then the local incentive constraint implies that reporting ϕ is suboptimal. Likewise, using (A.5) and completing the square appropriately, type ϕ's expected utility from reporting type $\hat{\eta}$ is $EU\left(x(\hat{\eta}), \eta, \alpha\right) + t(\hat{\eta}) = u(\hat{\eta}) - \frac{\alpha\sigma^2}{2} - \alpha\left(x(\hat{\eta}) - \hat{\eta}\right)\left(\hat{\eta} - \mu\right) - \frac{\alpha}{2}\left(\hat{\eta}-\mu\right)^2$. Reporting ϕ is better than reporting any $\hat{\eta}$ if

$$\int_{\underline{\eta}}^{\mu} \alpha(x(z)-z)\,dz - \frac{\alpha\sigma^2}{2} \geq \int_{\underline{\eta}}^{\hat{\eta}} \alpha(x(z)-z)\,dz - \frac{\alpha\sigma^2}{2} - \alpha\left(x(\hat{\eta}) - \hat{\eta}\right)\left(\hat{\eta}-\mu\right) - \frac{\alpha}{2}\left(\hat{\eta}-\mu\right)^2.$$

This condition is equivalent to

$$\int_{\hat{\eta}}^{\mu} \alpha\left[x(z) - x(\hat{\eta})\right]dz \geq 0.$$

Hence, $\frac{\partial x(\eta)}{\partial \eta} \geq 0$ makes the local incentive constraint of type ϕ sufficient for truth-telling to be globally optimal.

(v) To get the effort constraint, observe that $g_e(e) = \int_{\underline{\eta}}^{\overline{\eta}} u(\eta)\,dF(\eta) - Eu(\phi)$. Using (A.5) and (A.6) I find

$$\int_{\underline{\eta}}^{\overline{\eta}} u(\eta)dF(\eta) - Eu(\phi) = \int_{\underline{\eta}}^{\overline{\eta}}\int_{\underline{\eta}}^{\eta} \alpha(x(z)-z)\,dz\,dF(\eta) - \int_{\underline{\eta}}^{\mu}\int_{\underline{\eta}}^{\eta} \alpha(x(z)-z)\,dz + \alpha\frac{\sigma^2}{2}.$$

Integrating by parts gives the result. ‖

Part II: "Solution of Problem (M)"

The solution uses a three-step sequential maximization procedure. In Step 1 (Problem M1) I take as given that the principal implements some $x(\phi) = \bar{x}$ and $e = \bar{e}$ and solve for a constrained optimal function $x(\eta\,|\bar{x}, \bar{e})$, that implements $\bar{x}$ and $\bar{e}$ optimally. In Step 2 (Problem M2), I take account of the solution to the Step 1 problem and treat $x(\phi)$ as a choice variable, thus obtaining a constrained function $x(\eta\,|x(\phi), \bar{e})$ and an optimal $x(\phi\,|\bar{e})$. In Step 3 (Problem M3), I make e endogenous, which delivers the optimal function $x(\eta)$, an optimal choice of $x(\phi)$, and the optimal effort level that the principal implements.

Problems M1–M3 are *tractable* only if $\bar{e}$, the effort level that shall be implemented, is not excessively high. If $\bar{e}$ is too high, it causes bunching problems. Anticipating the results that I obtain below, I must restrict attention to $\bar{e} \leq \min\left\{e^{fb}(\alpha), e^{\max}(\alpha)\right\}$ where $e^{fb}(\alpha) := g_e^{-1}\left(\frac{1+\alpha}{2}\sigma^2\right)$ and $e^{\max}(\alpha) := g_e^{-1}\left(\frac{\alpha\sigma^2}{2} + \alpha\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu} - F(\eta)\right)^2}{f(\eta)}\right)d\eta\right)$. Since any *optimal* effort level is necessarily smaller than $e^{fb}(\alpha)$, $e^{\max}(\alpha) - e^{fb}(\alpha)$ is monotonic in α, and $e^{\max}(\alpha) > e^{fb}(\alpha)$ for large α, I will obtain the global optimum for α large, but will obtain a local optimum for α small.

For reasons of space I omit stating Problems M1 and M2. I present the solution to Problem M2 in the following lemma:

Lemma II. *Suppose that $\bar{e} \leq \min\left\{e^{\max}, e^{fb}\right\}$. Then the optimal contract that implements the given effort level $\bar{e}$ (the solution to Problem M2) takes the following form:*

$$\begin{aligned} x(\eta\,|\bar{e}) &= \eta - \frac{\lambda(\bar{e})}{\bar{e}}\frac{\alpha}{1+\alpha}\frac{F(\eta)}{f(\eta)} \quad \textit{for } \eta < \mu \\ x(\phi\,|\bar{e}) &= \mu \qquad \textit{for } \eta = \mu \\ x(\eta\,|\bar{e}) &= \eta + \frac{\lambda(\bar{e})}{\bar{e}}\frac{\alpha}{1+\alpha}\frac{1-F(\eta)}{f(\eta)} \quad \textit{for } \eta > \mu. \end{aligned} \tag{A.8}$$

where $\frac{\lambda(\bar{e})}{\bar{e}}$ is uniquely defined by the incentive constraint on effort:

$$\frac{\lambda(\bar{e})}{\bar{e}} = \frac{g_e(\bar{e}) - \frac{\alpha\sigma^2}{2}}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\alpha^2}{1+\alpha}\frac{\left(1_{\eta>\mu} - F(\eta)\right)^2}{f(\eta)}\right)d\eta} \tag{A.9}$$

$\frac{\lambda(\bar{e})}{\bar{e}}$ is positive only if $\bar{e} > g_e^{-1}\left(\frac{\alpha\sigma^2}{2}\right)$. The larger $\bar{e}$, the larger the departure from the ex post *efficient scheme at any η,* i.e. *$\frac{\lambda(\bar{e})}{\bar{e}}$ is increasing in $\bar{e}$.*

For the complete proof of Lemma II see Szalay (2002). Heuristically, the argument is as follows: $x(\phi\,|\bar{e}) = \mu$ is optimal because this choice occurs with *ex ante* probability mass $1 - e$; all other choices have *ex ante* probability 0; therefore the optimal contract avoids distortions for type ϕ; with $x(\phi) = \mu$ and taking $e = \bar{e}$, pointwise maximization with respect to $x(\eta)$ in Problem M gives (A.8); pointwise maximization is justified: $e^{\max}$ is defined such that $\frac{\lambda(\bar{e})}{\bar{e}}\frac{\alpha}{1+\alpha}\Big|_{e=e^{\max}} = 1$; together with the regularity conditions on $F(\eta)$ this ensures that $\frac{\partial x(\eta|\bar{x},\bar{e})}{\partial\eta} \geq 0$; hence, all monotonicity constraints in Problem M are satisfied by (A.8); to obtain (A.9) note that by definition $\lambda(\bar{e})\left\{\int_{\underline{\eta}}^{\bar{\eta}}\alpha\left((\eta\,|\bar{x}, \bar{e}) - \eta\right)\left(1_{\eta>\mu} - F(\eta)\right)d\eta + \frac{\alpha\sigma^2}{2} - g_e(e)\right\} \equiv 0\ \forall\bar{e}$; substitute (A.8) into this identity and solve for $\frac{\lambda(\bar{e})}{\bar{e}}$.

Problem M3 is obtained by substituting (A.8) into Problem M. The objective function is

$$\begin{aligned} P(e, \cdot) \quad = \quad & (1+\alpha)K - (1-e)\left(\frac{\alpha+1}{2}\right)\sigma^2 \\ & -e\left\{\int_{\underline{\eta}}^{\bar{\eta}}\frac{\alpha+1}{2}\left(\frac{\alpha}{\alpha+1}\frac{\lambda(e)}{e}\frac{1_{\eta>\mu} - F(\eta)}{f(\eta)}\right)^2 dF(\eta)\right\} - g(e) \\ & +\lambda(e)\left\{\int_{\underline{\eta}}^{\bar{\eta}}\alpha\left(x(\eta\,|e) - \eta\right)\left(1_{\eta>\mu} - F(\eta)\right)d\eta + \frac{\alpha\sigma^2}{2} - g_e(e)\right\}. \end{aligned}$$

The objective function is well defined for $e \leq \min\left\{e^{fb}, e^{\max}\right\}$. The principal's problem is

$$\max_{e\leq\min\{e^{fb}, e^{\max}\}} P(e, \cdot). \tag{M3}$$

Lemma III. *At the optimum the effort constraint is strictly binding. That is, let e^* denote an optimal effort choice. Then $\frac{\lambda(e^*)}{e^*} > 0$. Problem $(M3)$ has an interior solution for all α. For α sufficiently large the solution to $(M3)$ is the unique optimal effort level.*

Proof of Lemma III. From the identity $\lambda(e)\left\{\int_{\underline{\eta}}^{\bar{\eta}} \alpha\,(x(\eta)-\eta)\left(1_{\eta>\mu}-F(\eta)\right)d\eta + \frac{\alpha\sigma^2}{2} - g_e(e)\right\} \equiv 0\ \forall e$ and using (A.8), with $\bar{e}$ replaced by e, to substitute for $x(\eta)$, I solve for $\frac{\lambda(e)}{e}$:

$$\frac{\lambda(e)}{e} = \frac{g_e(e) - \frac{\alpha\sigma^2}{2}}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\alpha^2}{1+\alpha}\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta}. \tag{A.10}$$

Let $\hat{e} = g_e^{-1}\left(\frac{\alpha\sigma^2}{2}\right)$. I find that $\left.\left(\frac{\lambda(e)}{e}\right)\right|_{e=\hat{e}} = 0$. By straightforward calculus,

$$P_e(e,\cdot) = \frac{\alpha+1}{2}\sigma^2 - \int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\alpha^2}{2(1+\alpha)}\left(\frac{\lambda(e)}{e}\frac{1_{\eta>\mu}-F(\eta)}{f(\eta)}\right)^2\right)dF(\eta) - g_e(e)$$
$$- e\left\{\int_{\underline{\eta}}^{\bar{\eta}}\frac{\alpha^2}{2(1+\alpha)}\left(\frac{1_{\eta>\mu}-F(\eta)}{f(\eta)}\right)^2 dF(\eta)\right\}\frac{2\lambda(e)}{e}\frac{\partial}{\partial e}\left[\frac{\lambda(e)}{e}\right].$$

Observe that $P_e(e,\cdot)|_{e=\hat{e}} = \frac{\alpha+1}{2}\sigma^2 - g_e(e)|_{e=\hat{e}} = \frac{\sigma^2}{2} > 0$. Hence, it is optimal to introduce *ex post* distortions.

I now prove that problem (M3) has an interior solution for all α. An optimal effort level satisfies the first- and second-order conditions

$$\left.\left(\frac{\alpha+1}{2}\sigma^2 - g_e(e) - \frac{\frac{1+\alpha}{2}\left(\frac{g_e(e)}{\alpha}-\frac{\sigma^2}{2}\right)^2}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta} - e\frac{\frac{1+\alpha}{\alpha}\left(\frac{g_e(e)}{\alpha}-\frac{\sigma^2}{2}\right)}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta}g_{ee}(e)\right)\right|_{e=e^*} = 0$$

$$\left.\left(-g_{ee}(e) - \frac{\frac{1+\alpha}{\alpha}\left(\frac{g_e(e)}{\alpha}-\frac{\sigma^2}{2}\right)}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta}\left(2g_{ee}(e) + e\frac{g_{ee}(e)^2}{\alpha} + eg_{eee}(e)\right)\right)\right|_{e=e^*} < 0.$$

Note that $g_{eee}(e) \geq 0$ implies that the first-order condition is sufficient for a local optimum. I need to show that satisfaction of the first-order condition is consistent with $e^* \leq e^{\max}$ for all α. To see this, rearrange the first-order condition to get

$$\left.\left(\frac{\left(\frac{g_e(e)}{\alpha}-\frac{\sigma^2}{2}\right)}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta}\right)\right|_{e=e*} = \left.\left(\frac{\frac{\alpha+1}{2}\sigma^2 - g_e(e)}{\frac{1+\alpha}{2}\left(\frac{g_e(e)}{\alpha}-\frac{\sigma^2}{2}\right) + e\frac{1+\alpha}{\alpha}g_{ee}(e)}\right)\right|_{e=e*}.$$

If $g_{eee}(e) \geq 0$ then $e\frac{1+\alpha}{\alpha}g_{ee}(e) \geq \frac{1+\alpha}{\alpha}g_e(e) > \frac{\alpha+1}{2}\sigma^2$. Therefore, the value of the expression on the R.H.S. is smaller than one and $\gamma(\alpha) = \left.\left(\frac{\alpha}{1+\alpha}\frac{\lambda(e)}{e}\right)\right|_{e=e^*} < 1$. Thus, for all α, the stated contract is locally optimal. Finally, for α large, the contract is the unique globally optimal contract. $e \leq e^{fb}$ is equivalent to $g_e(e) \leq (1+\alpha)\frac{\sigma^2}{2}$. Therefore also $\frac{\left(\frac{g_e(e)}{\alpha}-\frac{\sigma^2}{2}\right)}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta} \leq \frac{\left(\frac{\alpha+1}{2\alpha}\sigma^2-\frac{\sigma^2}{2}\right)}{\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta} = \frac{\sigma^2}{2\alpha\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta}$. Hence, $\gamma(\alpha)$ is bounded above by one for all e smaller or equal to the first-best effort level and $\alpha \geq \frac{\sigma^2}{2\int_{\underline{\eta}}^{\bar{\eta}}\left(\frac{\left(1_{\eta>\mu}-F(\eta)\right)^2}{f(\eta)}\right)d\eta}$. Consequently, e^* is the unique global maximizer of Problem (M3). ∥

Acknowledgements. The paper is a revised version of a chapter, entitled "Optimal Delegation", of my 2001 doctoral thesis written at the University of Mannheim. An earlier version circulated under the title "On the One-handed Economist". I owe special thanks to Martin Hellwig for many discussions and comments. Three anonymous referees and Mark Armstrong, the editor, provided excellent comments that greatly improved the paper. I also would like to thank Dirk Bergemann, Jacques Crémer, Wouter Dessein, Mathias Dewatripont, Christian Ewerhart, Anke Kessler, Christian Laux, Nicolas Melissas, Benny Moldovanu, John Moore, Georg Nöldeke, Marco Ottaviani, Daniel Probst, Klaus Schmidt, Elu von Thadden, Thomas von Ungern-Sternberg, and seminar participants on various occasions. Finally, many thanks to Sandra Sizer for her suggestions that greatly improved the expression and style in my paper. Financial support from the Swiss National Science Foundation is gratefully acknowledged. All remaining errors are my own.

REFERENCES

AGHION, P. and TIROLE, J. (1997), "Formal and Real Authority", *Journal of Political Economy*, **105**, 1–29.

ARMSTRONG, M. (1994), "Delegation and Discretion" (Discussion Papers in Economics and Econometrics No. 9421, University of Southampton).

BERGEMANN, D. and VALIMAKI, J. (2002), "Information Acquisition and Efficient Mechanism Design", *Econometrica*, **70**, 1007–1033.

CRAWFORD, V. and SOBEL, J. (1982), "Strategic Information Transmission", *Econometrica*, **50**, 1431–1451.

CRÉMER, J., KHALIL, F. and ROCHET, J.-C. (1998), "Contracts and Productive Information Gathering", *Games and Economic Behavior*, **25** (2), 174–193.

CRÉMER J. and MCLEAN (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions", *Econometrica*, **56**, 1247–1257.

DEMSKI, J. and SAPPINGTON, D. (1987), "Delegated Expertise", *Journal of Accounting Research*, **25**, 68–89.

DESSEIN, W. (2002), "Authority and Communication in Organizations", *Review of Economic Studies*, **69**, 811–838.

DEWATRIPONT, M. and TIROLE, J. (1999), "Advocates", *Journal of Political Economy*, **107**, 1–39.

DE GARIDEL-THORON, T. and OTTAVIANI, M. (2000), "The Economics of Advice" (Mimeo UCL).

GIBBONS, R. (1988), "Learning in Equilibrium Models of Arbitration", *American Economic Review*, **78**, 896–912.

GREEN, J. and STOKEY, N. (1981), "The Value of Information in the Delegation Problem" (Discussion Paper, Harvard University).

HOLMSTRÖM, B. (1978), "On Incentives and Control in Organizations" (Ph.D. Dissertation, Stanford University).

HOLMSTRÖM, B. (1984), "On the Theory of Delegation", in M. Boyer and R. Kihlstrom (eds.) *Bayesian Models in Economic Theory* (Amsterdam: Elsevier Science Publishers B.V.).

HOLMSTRÖM, B. and RICART-i-COSTA, J. (1986), "Managerial Incentives and Capital Management", *Quarterly Journal of Economics*, **101**, 835–860.

LAFFONT, J.-J. and MARTIMORT, D. (2002) *The Principal–Agent Model* (Princeton and Oxford: Princeton University Press).

LEWIS, T. and SAPPINGTON, D. (1993), "Ignorance in Agency Problems", *Journal of Economic Theory*, **61**, 169–183.

LI, H. (2001), "A Theory of Conservatism", *Journal of Political Economy*, **109**, 617–636.

MELUMAD, N. and SHIBANO, T. (1991), "Communication in Settings with No Transfers", *Rand Journal of Economics*, **22**, 173–198.

OSBAND, K. (1989), "Optimal Forecasting Incentives", *Journal of Political Economy*, **97**, 1091–1112.

OTTAVIANI, M. and SORENSEN, P. (2000), "Professional Advice" (Mimeo UCL).

OTTAVIANI, M. and SORENSEN, P. (2001), "Professional Advice: The Theory of Reputational Cheap Talk" (Mimeo LBS).

PRENDERGAST, C. (1993), "A Theory of Yes Man", *American Economic Review*, **83**, 757–770.

PRENDERGAST, C. and STOLE, L. (1996), "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning", *Journal of Political Economy*, **104**, 1005–1134.

SZALAY, D. (2000), "Optimal Delegation" (Working Paper, University of Mannheim).

SZALAY, D. (2002), "On the One-Handed Economist" (Working Paper, University of Mannheim).

SZALAY, D. (2004*a*), "Contracts with Endogenous Information" (Working Paper, University of Lausanne).

SZALAY, D. (2004*b*), "On Information Acquisition in Delegated Portfolio Management" (Working Paper, University of Lausanne).

TIROLE, J. (1999), "Incomplete Contracts: Where Do We Stand?", *Econometrica*, **67**, 741–781.