Optimal Delegation[*]

Dezső Szalay
University of Mannheim

January 26, 2000

Abstract

How should a principal delegate a task to an agent? This paper studies the choice of an agent's discretion as a contracting problem. We show that the agent's freedom of action can be used as an effective incentive device: the agent's initiative is determined by the discretion he has in decision making. Due to this incentive effect the relationship between the severity of the conflict of principal's and agent's interests and the agent's optimal discretion in decision making is potentially non-monotonic: it may be optimal to curtail a subordinate's authority over decision making even if there are no conflicting interests concerning that decision.

Our theory provides a rationale for commonly observed phenomena such as "demanding clear statements" from advisors or "imposing an innovation bias" on an organizational structure.

Key words: authority, discretion, initiative

1 Introduction

Standard agency theory assumes that the agent's action space is given. In some contexts, however, it is of interest to study how much freedom of action an agent should be given. This is a key issue for the theory of delegation.

The problem of delegation has previously been studied by Holmström (1984) and Armstrong (1994) and from a different perspective by Aghion and Tirole (1997). Holmström and Armstrong address the question of how much discretion an agent should be given if principal and agent do not share the same objective. The principal's optimal choice reflects the following tradeoff: on the one hand the agent has some superior information, which can be used to improve the

[*]Acknowledgements: Special thanks are due to Martin Hellwig for ongoing discussions and advice. I would also like to thank Sascha Haller, Christian Laux and Nicolas Melissas as well as the participants of the Mannheim workshop in economic theory for very useful discussions and comments. Financial Support of the Swiss National Science Foundation is gratefully acknowledged. All remaining errors are my own. email: Szalay@pool.uni-mannheim.de

quality of decision making. Giving the agent more discretion may therefore allow the principal to make use of the agent's information. On the other hand the principal faces the risk that the agent may abuse the freedom of action this provides. In this literature on the optimal "discretion" of agents, the information structure is taken as given.

Aghion and Tirole (1997) in contrast study the problem of delegation in a context where the contracting parties' information structures are endogenous. Their main point is that the transfer of authority to an agent will encourage the agent's initiative. Consequently when deciding on the optimal allocation of authority the principal's decision will re°ect the tradeo® between loss of control and increased initiative of the agent. However, Aghion and Tirole address the problem of delegation as a binary choice problem between di®erent institutions, i.e. the principal has no choice on how to delegate the decision as he has in Holmström's or Armstrong's paper.[1]

The present paper studies a hybrid of these two distinct approaches to the allocation of authority (or optimal choice of discretion). In our model, the agent's information arises endogenously from costly e®orts of learning as in Aghion and Tirole (1997). We focus on a particular situation where giving the right to decide to the agent is optimal. In addition we assume that the principal has the right - as in the literature on discretion - to exclude certain projects from the agent's choice set, i.e., the principal has "veto" or "gatekeeping counterpower" (Tirole (1999)). We ask: how should the principal optimally use this veto power? or equivalently: how much discretion should the agent optimally be given in this context?

Our analysis highlights another incentive e®ect of discretion. The agent's freedom of action itself can be used as an incentive device. The most interesting result of the paper is that due to these incentive e®ects of discretion the relationship between the severity of the con°ict of interest and the principal's optimal choice of discretion is not necessarily monotonic. Somewhat surprisingly the present paper shows that in the very case where objectives are shared (read: preferences are identical) the principal may want to restrict the agent's freedom of action! This stands in sharp contrast with the results of Holmström (1984) and Armstrong (1994) as well as commonly held wisdom.

Though counterintuitive at ¯rst sight the economic rationale for the result is easy to explain heuristically. The timing of events and the information structure is as follows: at the time of contracting, agent and principal have symmetric information. After the contract is signed, the agent learns about the nature of the

[1] Tirole (1999) develops the "complete contracting" version of the Aghion and Tirole (1997) paper. He discusses the institution of giving authority to the agent with the principal having "gatekeeping counterpower". In this institutional setting the principal is given the right to exclude certain alternatives a priori. However, his point is to show that the concept of authority is not an artefact of modelling the problem as an incomplete contracting problem. To illustrate this idea, he shows that we can ¯nd an institution in an incomplete contracting context that implements the same outcome as the optimal complete contract. Giving authority to the agent with the principal having gatekeeping counterpower is an example of such an institution. This point is distinct from our question of how to use gatekeeping counterpower optimally.

optimal choices from his own as well as the principal's perspective. Thereafter he picks his preferred alternative out of those that are allowed in the contract. We assume that the e®ort he spends to acquire information is not observable and hence noncontractible.

This situation gives rise to a number of tradeo®s at the time of contract design. The nature of the tradeo® depends on whether the principal expects there to be con°icting interests once the agent has acquired his information or not. If the principal has to expect such con°icts then his optimal choice of contract re°ects the tradeo® of improved quality of decision making and increased initiative by the agent on the one hand versus the danger of opportunistic behavior by the agent (i.e. loss of control) on the other hand. If the principal expects no such con°icts concerning project choices, the nature of this tradeo® is di®erent, because there is no loss of control when the agent gets more discretion. On the contrary: interfering with the choices of an agent who always acts in the principal's best interest can only decrease the quality of decision making. However, limiting the agent's discretion (or curtailing his authority) appropriately will elicit more initiative from the agent. The nature of the tradeo® is thus completely reversed.

Limiting the agent's discretion e®ectively serves as a punishment scheme. Maybe somewhat surprisingly this punishment scheme will always (almost surely) hit the agent harder when he is ignorant than when he is well informed. He therefore has a strong incentive to avoid being ignorant. The really surprising thing, however, is that the principal might actually bene¯t from such a device in the very case where agent and principal have identical preferences!

Our theory might help explain some real world phenomena: politicians, when they approach economists for their advice, want to hear clear statements about whether a policy is good or bad. Thus, the answer, "it depends" is excluded from the agent's choice set, although it is often the best answer that can be given (and it is certainly the best answer for an economist, if he has no speci¯c information). As a consequence the economist - provided that he has a preference for correct answers - will spend more time thinking about the correct solution.[2] Judges or members of a jury can e®ectively only choose between guilty and non guilty. The judgement "don't know" is implicitly excluded from the choice set, because the guiltyness of an accused must be proven beyond reasonable doubt. Consequently the judge and the jurors have a strong incentive to think hard about guiltyness or nonguiltyness of the accused.[3] As a last example, consider a corporate culture with a strong bias for innovations. The ¯nding in the paper might give some rationale for the "we do it di®erently" doctrine. If headquarters follows the policy "whatever we do we will not stick to the status quo" division managers will have a strong incentive to become informed, about what change exactly should be implemented.

[2] However, there are other factors which also come to mind in this particular situation. Politicians also like to have someone to blame if things go wrong.

[3] We do not claim that society's rationale for this particular aspect of our legal system was the one given in our proposition 1 below. But it is fair to say that it is a side e®ect of the procedure.

The remainder of the paper is structured as follows. Section 2 introduces the model. Section 3 treats the case of complete alignement of interests. Section 4 offers a more general treatment of the problem, allowing for conflicts of interests. Section 5 concludes.

2 The model

We consider an agency problem in which the payoffs of both the principal and the agent depend on an action x chosen by the agent as well as on two parameters, η and φ; according to the specification

$$U(x;\eta) = k - \frac{A}{2}(x - \eta)^2 \tag{1}$$

and

$$\pi(x;\varphi) = K - \frac{1}{2}(x - \varphi)^2 \tag{2}$$

where U is the payoff of the agent and π the payoff of the principal. A is a measure of the agent's relative distaste for risk (read: relative to the principal). We will henceforth interpret the action x as the choice of a particular "project". The parameters η and φ are assumed to be realizations of random variables $\tilde{\eta}$ and $\tilde{\varphi}$ with joint distribution function $F_{\eta;\varphi}$. More precisely

Assumption 1: $(\eta;\varphi)$ have identical, symmetric marginals $f_\eta = f_\varphi$:

At the time of contracting both agent and principal know f but neither of them knows the realization of $\tilde{\eta}$ and $\tilde{\varphi}$:

Between the time the contract is written and the choice of the action x the agent tries to get informed. By exerting effort e the agent is informed about the realizations of both $\tilde{\eta}$ and $\tilde{\varphi}$ with probability e - uninformed with probability $1 - e$ - and bears costs of effort $g(e)$; where $g(e)$ is convex and satisfies the Inada conditions: $g'(e) > 0$; $g''(e) > 0$; $\lim_{e \to 0} g'(e) = 0$; $\lim_{e \to 1} g'(e) = \infty$: The agent's choice of e is not observable to the principle. Moreover, the principal does not observe whether the agent successfully learned the true realizations of $\tilde{\eta}$ and $\tilde{\varphi}$ or not.

The agent chooses the action x according to his information[4]. We study a situation where the principal cannot use monetary contracting schemes:

Assumption 2: The agent does not react to monetary incentives.

In light of assumption 2 there is no way to make the reward of the agent dependent on the realized payoff of the principal, π; even if π is perfectly observable

[4] In contrast to Aghion and Tirole (1997) we assume that the agent directly chooses a project. There is thus no communication of information about the realizations of $\tilde{\eta}$ and $\tilde{\varphi}$:

and verifiable.[5] Therefore, the principal can only contract on the observable decision x in a nonmonetary sense:

Contracts: the principal restricts the choice of x to an admissable set Γ.

The situation is thus one where the agent cares "much more" for his private benefit than he cares for other sources of income. Assumption 2 is admittedly very restrictive but it serves the purpose of emphasizing the impact of the agent's discretion on his initiative. There is, however, an alternative way to justify our approach. Instead of assumption 2 one can impose the following set of assumptions: suppose U and π are noncontractible. In addition to the private benefits, the agent derives utility $V(w)$ from financial income. The agent is infinitely risk averse with respect to income shocks. The principal is risk neutral with respect to financial income.

In this situation, the principal can write monetary contracts on the decision x. However, the optimal nonmonetary contract and the optimal monetary decision based contract will - under our assumptions - be equivalent in the following sense: they will both implement the same outcome in terms of effort and project choices. (Payoffs will differ of course.)[6][7]

The nature of conflicts: Under assumption 1 there is no conflict of interest concerning project choice ex ante: both principal and agent agree that $x \overset{!}{=} \mu$ is the best choice. However, ex post - i.e. conditional a specific realization $(\eta; \varphi)$ of $(\tilde{\eta}; \tilde{\varphi})$ - the principal and his agent do not necessarily agree what the best project is: the agent's preferred alternative is $x \overset{!}{=} \eta$ while the choice $x \overset{!}{=} \varphi$ is in the principal's best interest. Since marginals are identical, the correlation coefficient ρ between η and φ is a sufficient statistic for the expected extent of this conflict of interest concerning project choice ex post.

2.1 First Best

The first best corresponds to a situation where all variables can be written into an enforcable contract, i.e. everything is observable and verifiable. Such a contract will directly specify a particular choice of e and a choice of x contingent on the agent's information.

To provide a benchmark we will characterize this first best for the case of perfectly aligned interests concerning project choice, i.e $\eta \equiv \varphi$. The contract will thus specify $x = \eta$ for the case where the agent knows the realization of $\tilde{\eta}$ and $x = \mu$ for the case where the agent does not know the true realization of $\tilde{\eta}$.

[5] Recall that the agent's information on realizations of $\tilde{\eta}$ and $\tilde{\varphi}$ is not observable to the principal.

[6] This is shown in Appendix B part (i).

[7] This discussion shows in particular that our approach of nonmonetary contracting on decisions, i.e. the choice of discretion, can equivalently (in well defined environments) be understood as an optimal choice of monetary decision based rewards. Monetary decision based rewards have first been studied by Dewatripont and Tirole (1999) in a different context: they show that it may be better to have different agents advocating the pros and cons, respectively, of a certain matter rather than having one agent searching for the truth directly.

The first best effort level is defined as the effort level that maximizes joint surplus.[8] To derive this we must first calculate the value of information. From (2) the principal's payoff conditional on the agent knowing η is K; while the agent will enjoy utility k (from (1)). Conditional on the agent being ignorant the principal's payoff is: $K - \frac{1}{2}\int(\mu - \eta)^2 dF_\eta = K - \frac{1}{2}\sigma^2$; while the agent will then have utility $k - \frac{A}{2}\sigma^2$. The optimal choice of e solves

$$\max_e K + k - \frac{(1+A)\sigma^2}{2} + e\frac{(1+A)\sigma^2}{2} - g(e)$$

and therefore satisfies

$$\frac{(1+A)\sigma^2}{2} \overset{!}{=} g'(e)$$

I.e. the agent spends effort to learn until the marginal cost of effort is equal to the marginally avoided disutility of risk for the principal and himself.

2.2 Discretion as a Contracting Problem

The agent's information and his effort choice are not observable to the principal. In the absence of monetary compensation schemes, the principal will choose the admissable set Γ in such a way as to maximize his expected payoff. More formally the principal's maximization problem is:

$$\max_{x'';x'(.);\Gamma;e} eE\pi(x'(\eta);\tilde{\eta};\tilde{\mu}) + (1-e)E\pi(x'';\tilde{\mu}) \tag{3}$$

s.t.

$$\begin{aligned} x'(\eta) &\in \arg\max_{x\in\Gamma} U(x;\eta) \quad \forall\eta \\ x'' &\in \arg\max_{x\in\Gamma} EU(x;\tilde{\eta}) \end{aligned} \tag{4}$$

$$EU(x'(\eta);\tilde{\eta}) - EU(x'';\tilde{\eta}) = g'(e) \tag{5}$$

$$\begin{aligned} &eEU(x'(\eta);\tilde{\eta}) + \\ &(1-e)EU(x'';\tilde{\eta}) - g(e) \geq 0; \end{aligned} \tag{6}$$

The first term inside the brackets in (3) represents the principal's expected payoff when he knows that an agent who knows the realizations of $\tilde{\eta}$ and $\tilde{\mu}$ (henceforth an informed agent) chooses his most preferred alternative (given the restriction Γ). The second term represents the analogue for an ignorant agent's choice. (4) are the incentive compatibility conditions on the choice of alternative of the

[8] The present model is one of private benefits, which makes the interpretation of first best a bit difficult. It should be noted that our results do not depend on this definition of first best. The only important thing is that there is some underinvestment at all.

informed agent (henceforth $x'(.)$) and on the choice of an ignorant agent (x'' henceforth), respectively. (5) is the IC condition for the agent's effort choice. Finally (6) is the agent's individual rationality constraint, which is assumed to be nonbinding for all Γ.[9]

Notation: The following shortcuts will henceforth be used for the agent's and the principal's expected payoffs under contract Γ (anticipating incentive compatible choices according to his information and (4)):

$$E\pi'_\Gamma := E\pi(x'(\eta);\lambda;\mu); \quad E\pi''_\Gamma := E\pi(x'';\mu) \tag{7}$$
$$EU'_\Gamma := EU(x'(\eta);\lambda); \quad EU''_\Gamma := EU(x'';\lambda)$$

The tradeoff the principal faces - for instance if the correlation between η and φ is relatively low - can be seen from the incentive compatibility conditions on the agent's project choice (4) and on his effort choice (5). Through its impact on $x'(.)$ and x'', Γ directly affects the principal's payoff. Increasing Γ (to be made precise below) increases the danger of misbehavior of selfish agents. On the other hand there is also an indirect impact of Γ on the principal's payoff through the agent's effort choice (via (5)). The agent's effort choice is an increasing function of the wedge between the expected utility levels across the uninformed and informed state, respectively. The levels themselves are increasing in the degree of discretion or, put the other way, decreasing in the stringency of the restrictions, the agent faces.[10] To induce higher effort the agent has either to be rewarded in case he is informed, i.e he has to be given greater discretion for the choice of alternatives, or he has to be punished in case he is ignorant.
The problem is not completely standard. The difference stems from the fact that the principal controls the degree of freedom the agent has when he takes his decision: the principal maximizes with respect to sets which in turn are the side constraints on the agent's choices. To get a handle on the problem we impose the following restriction on the principal's choices:

Assumption 3: The principal restricts himself to use only one interval control and one interval prohibition at a time.

To clarify what we have in mind here, think of a worker in charge of buying goods for his company. His principal restricts his choices with the following rules: "buy at least y_l but not more than y^h items of article xy." Formally $\Gamma = \mathrm{dom}\, f(\eta) \setminus [-\infty; y_l); (y^h; \infty]$. This will be called an interval control. $\Gamma = \mathrm{dom}\, f(\eta) \setminus (y_l; y^h)$ shall be called an interval prohibition. Under this restriction the agent is allowed to buy either less than y_l or more than y^h items of the good, but nothing in between. In contrast $\Gamma = \mathrm{dom}\, f(\eta)$ means "do whatever

[9] Without monetary transfers this assumption is obviously needed for a meaningful analysis. We will discuss the impact of this assumption on our results below in detail.

[10] This will be made precise below. For the moment it suffices to note, that the value of a constrained maximization problem cannot be higher than the value of an unconstrained problem.

you deem right". Any collection of open sets can be described as a combination of controls and prohibitions. But for the moment, we simply assume that the principal uses only one prohibition and one control at the same time. We thus do not allow him to forbid any collection of open sets. However, it will be shown below, that this is a result rather than an assumption!

3 Perfectly Aligned Interests: The Degree of Freedom as an Incentive Device

If the principal and his subordinate have perfectly aligned interests, there is obviously no need for safeguards against opportunistic behavior ex post. Therefore full freedom of action would be optimal if it were not for the problem of effort choice. To see this formally, note that both $E\pi'_\Gamma$ and $E\pi''_\Gamma$ are maximized at $\Gamma = \text{dom } f(\theta)$ and any $\Gamma' \neq \Gamma$ must introduce deviations from first best choices ex post. It is immediate that the principal will not find it optimal to use interval controls.[11] One is tempted to conclude that the agent's decision space should be completely unrestricted. However, this is not generally true.
Although the incentive problem concerning project choice vanishes if ρ is equal to 1; the underinvestment problem due to moral hazard in effort choice does not. Relative to the social optimum the agent exerts too little effort, because he does not value the improvement in the principal's utility from an informed decision. The principal may therefore want to restrict the agent's choice set so as to increase his effort choice. The questions are then whether this is feasible and rational for the principal. Consider first feasibility: assume that the principal uses an interval prohibition of the form

$$\hat{\Gamma} = \text{dom } f \setminus (\mu - \epsilon, \mu + \epsilon) \tag{8}$$

for some specific $\epsilon > 0$. I.e. the principal allows everything but choices within a symmetric interval around the mean.

The agent's best response to this rule is

$$\begin{aligned} x'(\hat{\Gamma}, \theta) &= \begin{cases} \theta & \text{for } \theta \in \text{dom } f \setminus (\mu - \epsilon, \mu + \epsilon) \\ \mu - \epsilon & \text{for } \theta \in [\mu - \epsilon, \mu) \\ \mu + \epsilon & \text{for } \theta \in [\mu, \mu + \epsilon] \end{cases} \\ x''(\hat{\Gamma}) &= \{\mu + \epsilon \end{aligned} \tag{4'}$$

i.e. he minimizes the (expected) deviation from the unrestricted best choice, which would be $x = \theta$ if the agent is informed and $x = \mu$ if the agent is ignorant.[12] The agent's expected utility levels under $\hat{\Gamma}$; conditional on being

[11] A formal proof of this statement follows in proposition 3.

[12] We can pick $x'' = \mu + \epsilon$ without loss of generality since both principal and agent are indifferent between $\mu - \epsilon$ and $\mu + \epsilon$.

informed (EU'_ε)[13] and, respectively, on being ignorant (EU''_ε) are then (by symmetry)

$$EU'_\varepsilon = k - A \int_{\mu}^{\mu+\varepsilon} (\mu + \varepsilon - \eta)^2 dF_\eta \tag{9}$$

$$EU''_\varepsilon = k - \frac{A}{2} \int (\mu + \varepsilon - \eta)^2 dF_\eta = k - \frac{A}{2}(\sigma^2 + \varepsilon^2)$$

From (5) and (9) the agent's incentive compatible effort choice is

$$e(\varepsilon) = h\left[A \frac{(\sigma^2 + \varepsilon^2) - 2\int_{\mu}^{\mu+\varepsilon} (\mu + \varepsilon - \eta)^2 dF_\eta}{2}\right] \tag{5'}$$

where $h := g'^{-1}(.)$ (which exists by $g''(e) > 0$). It turns out that $\hat{\Gamma}$ has a very clearcut influence on the effort choice of the agent:

Lemma 1 e is increasing in ε.

Proof: see in the appendix.

The intuition is very simple. Since the agent's effort choice must be incentive compatible, we know from (5) that it is determined by the marginal gains the agent gets if he learns about (η, φ). This wedge between expected utility, conditional on being informed and conditional on being uninformed, $EU'_\varepsilon - EU''_\varepsilon$, can be increased by two ways. The principal can reward the agent if he is informed or he can punish him if he is not. Since, relative to Γ, it is not possible to reward the agent more the principal punishes the agent[14]. This punishment is effective because its impact on the agent's utility is more severe when the subordinate is ignorant: conditional on the agent being informed, the optimal choice will only by coincidence, i.e. small probability, lie in the excluded interval, while conditional on being ignorant the agent always wants to choose the excluded alternative $x = \mu$. It should be noted that this result holds generically , i.e. does not depend on any distributional assumptions as long as there is some uncertainty at all.

A positive impact of ε on e is a necessary condition for the optimality of a $\hat{\Gamma}$-type control, where $\hat{\Gamma}$ is given by (8), but it is of course not sufficient. Analogously to (9) one can derive the principal's expected utility conditional on the fact that his agent has learned the realizations of $\tilde{\eta}$ and $\tilde{\varphi}$ (that the agent is ignorant of the realizations, respectively)

$$E\pi'_\varepsilon = K - \int_{\mu}^{\mu+\varepsilon} (\mu + \varepsilon - \eta)^2 dF_\eta \tag{10}$$

$$E\pi''_\varepsilon = K - \frac{1}{2}(\sigma^2 + \varepsilon^2)$$

[13]With a slight abuse of notation we write EU'_ε for $EU'_{dom f \setminus (\mu-\varepsilon;\mu+\varepsilon)}$.

[14]Note that the principal thereby decreases both EU' and EU'' but the difference $EU' - EU''$ increases.

where $E\pi'_\varepsilon$ and $E\pi''_\varepsilon$ have been defined in (7). Plugging[15] (5') and (10) into (3), the principal's maximization problem can be restated as an unconstrained problem:

$$\max_{\varepsilon} e(\varepsilon)E\pi'_\varepsilon + (1 - e(\varepsilon))E\pi''_\varepsilon \qquad (11)$$

So if $\varepsilon^* > 0$, it satisfies the first order condition

$$\frac{\partial E\pi''_\varepsilon}{\partial \varepsilon} + e\frac{\partial}{\partial \varepsilon}[E(\pi'_\varepsilon - \pi''_\varepsilon)] + \frac{\partial e}{\partial \varepsilon}E(\pi'_\varepsilon - \pi''_\varepsilon) \overset{!}{=} 0 \qquad (12)$$

Obviously, the principal will use a $\hat{\Gamma}$-type control if and only if the derivative of his payoff with respect to ε is positive at $\varepsilon = 0$. Assume thus:

Assumption 4a: $(1 - e(\varepsilon = 0)) < \frac{A\sigma^2}{g''(e(\varepsilon=0))2}$.
Assumption 4b: $\frac{g'''(e)}{g''(e)^2} \geq \frac{4}{A\sigma^2}, \forall e.$[16]

Remark 1: Assumption 4a is not yet a condition on the primitives of the model. Yet, it contains the economic rationale of the arguments. Proposition 2 below translates the results into conditions on the parameters of the model.

Then:

Proposition 1 (I) the principal will optimally set $\varepsilon^* > 0$ if A4a holds.
(II) If in addition A4b holds, then there exists exactly one value ε^*, $\varepsilon^* > 0$, which solves (12).
(III) $\hat{\Gamma} = dom\ f \setminus (\mu - \varepsilon^*, \mu + \varepsilon^*)$ is the unique optimal interval prohibition.
(IV) $\hat{\Gamma}$ is the unique optimal policy, i.e. for $\rho = 1$ imposing A3 is without loss of generality.

Proof: see in the appendix.

The crucial question is whether the increase in the agent's effort supply is large enough to justify the ex post utility losses that the principal inflicts on himself. To give an intuitive explanation for the condition in A4a, $1 - e(0) < \frac{A}{g''(e(0))}\frac{\sigma^2}{2}$, consider ε very small but positive: with probability (approximately) $1 - e(0)$ the agent is ignorant and the principal just shoots himself in the foot, by setting $\varepsilon > 0$. The marginal impact of increasing ε on $E\pi''_\varepsilon$ is then equal to $-\varepsilon$. (With

[15] Observe that the direct cost of using the prohibition is identical for the principal and the agent. Thus, although the rule has favorable incentive effects, the principal will - once the agent's effort is sunk (but of course before the agent takes his decision concerning projects) - always regret having prohibited some choices as he cuts on expectation into his own flesh by using them.

[16] A4a is a condition on the first derivative of the principal's problem, A4b is one on the second derivative. A detailed discussion of A4a follows right after the statement of proposition 1. A4b is a sufficient condition guaranteeing that $\frac{\partial^2 e}{\partial \varepsilon^2} < 0$ for large enough ε, implying that the problem gets concave eventually.

probability $e(0)$ there is also a negative impact on $E\pi'_\varepsilon$, but this effect vanishes as ε is chosen small enough.) This must be compared with the incremental risk avoidance, i.e. additional effort supplied due to a marginal increase in the wedge $E(U'_\varepsilon - U''_\varepsilon)$ times the impact of avoided risk on the principal's utility: $\frac{A}{g''(e(0))}\frac{\sigma^2}{2}$. (It is shown in the appendix that the impact of increasing ε on $E(U'_\varepsilon - U''_\varepsilon)$ becomes indistinguishable from ε for ε very small.) If the net impact of these effects is positive the principal will use the rule $\hat{\Gamma}$.

Restricting attention to symmetric rules like $\hat{\Gamma}$ only is without loss of generality. The important characteristic of $\hat{\Gamma}$ is that it extracts maximum initiative from the agent at a given cost. Any nonsymmetric restriction, for instance, allows the agent to escape the punishment partially: if he is ignorant anyway he does not really care in which direction he has to deviate from the optimal choice. This depends of course on the symmetry of both the payoff function and the underlying distribution of random variables. Furthermore, $\hat{\Gamma}$ excludes the optimal choice under ignorance. This gives the agent a strong incentive to avoid being ignorant, i.e. to learn, since the expected marginal disutility of such a policy is always larger when he ends up without knowing the true realizations of the random variables.

When interests are perfectly aligned, it is not necessary to impose A3: the principal will voluntarily choose the policy $\hat{\Gamma}$ even if we allow him to choose from any possible combination of collections of open sets. The reason being that $\hat{\Gamma}$ is cheapest rule that increases the agent's effort supply. Any other rule either decreases the agent's effort supply and the expected payoff from a well informed decision (i.e. $E\pi'_\varepsilon$) or if it increases the agent's effort supply it will be more expensive than $\hat{\Gamma}$.

Remark 2: The comparison so far involves only direct cost and benefits, i.e. the private costs of additional effort are borne entirely by the agent. Since the principal cannot use any monetary transfers, there is simply no way to compensate the agent. But since the IR-condition is nonbinding initially, this causes no problems. Moreover, it is straightforward to extend the argument to the case of a binding IR-condition when monetary contracts are admitted. (see Appendix B (ii))

The next proposition first (part (i)) states conditions under which we are more likely to observe the use of such interval prohibitions and second (part (ii)), given that we observe them, how large they should be. For the derivation of result (iii)[17] and the analysis of the remaining part of the paper it is necessary to impose more structure on the distribution of (η, φ):

Assumption 1': (η, φ) are bivariate normal with joint density $f_{\eta,\varphi}$ and identical marginals $f_\eta = f_\varphi$, i.e. $\eta \sim N(\mu, \sigma^2)$, $\varphi \sim N(\mu, \sigma^2)$.

[17] Note however, that (i) and (ii) do not make use of this assumption!

Proposition 2 Under A1, A4a and A4b:
(i) the principal is more likely to set $\epsilon^* > 0$ if the agent's distaste for risk is weak (A small) and/or the underlying risk is small (σ^2 small) and
(ii) the optimal size of the prohibition is the larger the smaller A is.
Under A1':
(iii) The impact of σ on ϵ^* is ambiguous.

Proof: see in the appendix.

Intuitively, the principal should use prohibitions as an incentive device if they are not too costly and have a large and positive impact on the agent's initiative. A priori, the e®ects due to increases in the agent's distaste for risk or the underlying risk itself are not clear. This ambiguity is due to the fact, that these factors in°uence both how the agent reacts to changes in his freedom of action as well as how large his e®ort supply already is when he is given full freedom of action. While the level e®ect, $\frac{\partial e}{\partial A}$ is always positive, the growth or change e®ect, $\frac{\partial^2 e}{\partial \epsilon \partial A}$ is negative under A4b. Moreover, under A4b, the second e®ect dominates the ¯rst. Thus, the negative impact on the incrementally avoided risk due to the introduction of a prohibition is always dominant. It must be emphasized, however, that these e®ects are so clearcut if and only if we are willing to impose a lower bound (more speci¯cally $\frac{4}{A\sigma^2}$) on $\frac{1}{g''}\frac{g'''}{g''}$; the product of $\frac{1}{g''(e)}$ and its derivative, i.e. the curvature of the marginal cost function.[18]
The economic rationale for these e®ects is then very straightforward. There is little reason to use a costly incentive device if the agent's motivation is already high enough and/or he does not respond appropriately to prohibitions.
The reason that the prohibited interval be the smaller the more the agent hates risk, i.e. $\frac{\partial \epsilon}{\partial A} < 0$; is exactly the same. This, however, is not true for a change in the underlying risk, which has an ambiguous e®ect on the optimal size of the prohibited interval. A detailed discussion is relegated to the appendix, so we discuss here only the e®ect which is of ambiguous sign: it is not clear how the agent's responsiveness to incentive schemes is a®ected by changes in the underlying risk, i.e. the sign of $\frac{\partial^2 e}{\partial \epsilon \partial \sigma}$ is not clear: There is a negative impact, because a higher risk causes the agent to supply a higher level of e®ort at full freedom of action: c.p. it is harder to motivate the agent if he already works quite hard. But now there is also a positive impact because any given incentive scheme will have a more pronounced e®ect on the agent's e®ort supply. This is because the wedge between his utility levels is an increasing function of underlying risk. It is not clear which of these two e®ect dominates.

4 Con°icting Interests

So far only the border line case of perfect alignement of interests has been discussed. In this special case there was no need for the principal to protect himself

[18] Unfortunately, I have no good intuition for the third derivative of a cost function. However, A4b was imposed to guarantee that $\frac{\partial^2 e}{\partial \epsilon^2} < 0$ if the size of the prohibition gets large enough:

from ex-post-opportunistic behavior of his agent, since the only deviation from first best was the agent's effort choice. In general, i.e. when $\tilde{\eta}$ and $\tilde{\varphi}$ do not co-vary one for one with each other, both the agent's effort choice and his project choice will not coincide with the first best choices. Concerning the project choice, the worst case scenario for the principal is given by a realization-tuple $\eta; \varphi$ of $\tilde{\eta}; \tilde{\varphi}$ where η and φ are extremely far apart. As a consequence, the agent when left unrestricted to choose whatever he deems right, might then choose an alternative which gives the principal a very low private benefit. In other words we introduce divergence of interests ex post. The principal can majorize the expected impact on his private benefits of this divergence of interests ex post if he forbids extreme alternatives a priori.
We will say that by setting some specific $\lambda = \lambda'$ the principal allows all choices of $x \in [\mu - \lambda'; \mu + \lambda']$; i.e. forbids all choices of $x \in (\mu + \lambda'; \infty]$ and $x \in [-\infty; \mu - \lambda')$; combined with the ϵ-controls from above, we have $\hat{\Gamma}' = \text{dom}\, \{\eta \setminus [-\infty; \mu - \lambda)(\mu - \epsilon; \mu + \epsilon)(\mu + \lambda; \infty]$: The agent's best response to this restriction is

$$
\begin{aligned}
x'(\hat{\Gamma}') &= \begin{cases} \mu - \lambda & \text{for } \eta \in [-\infty; \mu - \lambda) \\ \eta & \text{"} \quad [\mu - \lambda; \mu - \epsilon] \\ \mu - \epsilon & \text{"} \quad (\mu - \epsilon; \mu] \\ \mu + \epsilon & \text{"} \quad (\mu; \mu + \epsilon) \\ \eta & \text{"} \quad [\mu + \epsilon; \mu + \lambda] \\ \mu + \lambda & \text{"} \quad (\mu + \lambda; \infty] \end{cases} \qquad (4") \\
x'' &= \{\mu + \epsilon
\end{aligned}
$$

giving him expected utility levels

$$
\begin{aligned}
EU'_{\lambda,\epsilon} &= k - A \left\{ \int_{\mu+\lambda}^{\infty} (\mu + \lambda - \eta)^2 dF_\eta + \int_{\mu}^{\mu+\epsilon} (\mu + \epsilon - \eta)^2 dF_\eta \right\} \qquad (13) \\
EU''_{\lambda,\epsilon} &= k - A \frac{\sigma^2 + \epsilon^2}{2}
\end{aligned}
$$

where we use the fact, that - by symmetry - we only have to consider the upper half of the support of η: As for the principal, his expected utility from informed decisions of his agent changes to: $E\pi'_{\lambda,\epsilon} = K -$

$$
\left\{ \begin{array}{c} \iint_{\mu}^{\mu+\epsilon} (\mu + \epsilon - \varphi)^2 dF_\eta dF_{\varphi|\eta} \\ + \iint_{\mu+\epsilon}^{\mu+\lambda} (\eta - \varphi)^2 dF_\eta dF_{\varphi|\eta} + \iint_{\mu+\lambda}^{\infty} (\mu + \lambda - \varphi)^2 dF_\eta dF_{\varphi|\eta} \end{array} \right\} \qquad (14)
$$

while $E\pi''_{\lambda,\epsilon}$ is unchanged relative to $E\pi''_{\epsilon}$ by the introduction of λ-controls. ($E\pi'_{\lambda,\epsilon} := E\pi'(x'(\eta); \tilde{\eta}; \tilde{\varphi})$ and $E\pi''_{\lambda,\epsilon} := E\pi''(x''; \tilde{\varphi})$ where $x'(:)$ and x'' obey now (4")). The benefit of forbidding extreme alternatives a priori for the principal is the protection against extreme conflicts. The restriction effectively majorizes the utility loss of the principal in the worst case. The flipside of the coin is of course that the principal might end up forbidding a choice that he would actually want to have implemented ex post and that the agent will respond with less effort.

Formally, the principal now solves

$$\max_{\varepsilon,\lambda} e(\varepsilon,\lambda)E\pi'_{\varepsilon,\lambda} + (1 - e(\varepsilon,\lambda))E\pi''_{\varepsilon,\lambda} \tag{15}$$

where $e(\varepsilon,\lambda)$ now combines the information contained in (5') and (13), while $E\pi'_{\varepsilon,\lambda}$ is given by (14). This program yields the first order conditions

$$\frac{\partial}{\partial\varepsilon}E\pi''_{\varepsilon} + e\frac{\partial}{\partial\varepsilon}\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon})\right] + \frac{\partial e}{\partial\varepsilon}E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon}) \overset{!}{=} 0 \tag{FOC$_\varepsilon$}$$

$$e\frac{\partial}{\partial\lambda}E\pi'_{\varepsilon,\lambda} + \frac{\partial e}{\partial\lambda}E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda}) \overset{!}{=} 0 \tag{FOC$_\lambda$}$$

Proposition 3 (ia) An optimal solution to problem (15) exists. (ib) The solution is unique (a.e.).
Value of Delegation:
(ii) The principal is better off if he contracts with agents whose interests are more in line with his own, i.e. $\frac{\partial}{\partial\rho}E\pi'_{\varepsilon,\lambda} > 0$.
Optimal Use of Interval Prohibitions:
(iii) Under A4a and A4b it is optimal to set $\varepsilon > 0$ if ρ is sufficiently close to 1, $\frac{\partial\varepsilon^*}{\partial\rho} \geq 0$.
Optimal Use of Interval Controls:
(iv) If $\rho \leq 0$ it is optimal to set $\lambda^* = 0$.
(v) If $0 < \rho < 1$ the principal will choose the boundary of the interval control such that $\lambda^* > \bar{\lambda}$, where $\bar{\lambda} = \arg\max E\pi'_{\varepsilon,\lambda}$. If $\rho = 1$, $\lambda^* = \bar{\lambda}$.
(vi) λ^* increases in ρ.

Proof of part (i)-(iii) and (vi): see in the appendix.
Proofs of parts (iv) and (v): To make delegation worthwhile at all, the principal must benefit if his agent is well informed. This requires that $E\left[\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda}\right] \geq 0$, which in turn affords $\rho \geq 0$. The principal can only loose if he gives discretion to an agent who's interests are opposed to his own. To see part (v), consider the equation $e\frac{\partial E\pi'}{\partial\lambda} + \frac{\partial e}{\partial\lambda}E\left[\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda}\right] \overset{!}{=} 0$. The agent reacts positively on increased freedom of action ($\frac{\partial e}{\partial\lambda} > 0$). If a solution exists, we must therefore necessarily have $\frac{\partial E\pi'}{\partial\lambda} < 0$, (since $E\left[\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda}\right] \geq 0$, too). An important though purely technical result is established in Lemma 2, which has been relegated to the appendix: $E\pi'_{\varepsilon,\lambda}$ is a quasiconcave function of λ. This implies, for $0 < \rho < 1$, together with the arguments above that $\lambda^* > \bar{\lambda}$. That $\rho = 1 \Rightarrow \lambda^* = \bar{\lambda}$ is shown in the appendix. Essentially this is due to the fact that for $\rho < 1$, the principal can, by decreasing λ, increase $E\pi'_{\varepsilon,\lambda}$ a bit without reducing effort supply a lot. This proves the claims.

Interval prohibitons serve to motivate the agent to supply high effort. Interval controls are designed to protect the principal from opportunistic behavior of his agent. Forbidding the choice which is best under ignorance is a very costly

incentive instrument. It will only be used if the benefit of doing so is sufficiently large. However, the principal will not benefit greatly if an agent, whose preferred choices are always "far away" from the principal's preferred choices, puts in more effort. Therefore the principal will contend himself with using safeguards only. The principal can never gain anything by delegating to an agent who is against him on average ($\rho \leq 0$) - no matter how he delegates the task. As long as the principal does not find it optimal to give his agent full discretion, he will give a freedom of action, which is larger than would be optimal from an ex post perspective only. By sacrificing some of the expected benefits from an informed decision of his agent, he induces a higher effort choice of him. Agents whose interests are closer aligned with those of the principal get more freedom. It should be noted that this result is due to the assumption of identical marginals. If the prior means of principal and agent differ, agents with higher ρ need not be better for the principal nor need they be given a larger degree of freedom. However the result that a principal can induce higher effort by the use of interval-prohibitions - i.e. by forbidding the agent's most preferred alternative when ignorant - applies in this case too. Note also that the cost of using these prohibitions decreases as the prior means move apart. In contrast to the present situation the principal does not punish himself ex post when he makes the ignorant life for his agent unpleasant. Apart from these considerations it should be possible to eliminate differences in opinions ex ante trough some kind of screening procedure.
An important robustness consideration concerns the policies the principal is allowed to choose from. So far we have simply assumed that the principal uses only prohibitions and controls as specified in Assumption 3. But this is actually a result, rather than an assumption:

Proposition 4 Under A1', if σ is large, $\hat{¡}^0$ is the unique optimal policy.

Proof: see in the appendix.

While the formal analysis is relegated to the appendix, an intuitive sketch shall be given here. We have already shown, that in the special case of $\rho = 1$; $\hat{¡}$ was the unique optimal solution. This implies that we cannot improve upon the interval prohibition as specified there, i.e. either $\varepsilon = 0$ or excluding a symmetric interval $(\mu - \varepsilon; \mu + \varepsilon)$ from the choice set of the agent is part of an optimal policy. But in stead of interval controls only we should allow the principal to exclude any collection of open sets from the agent's choice set. As a result, we now must consider three candidates for optimal policies: in addition to prohibiting $(\mu - \varepsilon; \mu + \varepsilon)$ use (i) interval controls only, or (ii) a countable but possibly infinite number of prohibitions of the type $(\mu + \gamma_i - t_i; \mu + \gamma_i + t_i)$ with $\mu + \gamma_{h+1} - t_{h+1} > \mu + \gamma_h + t_h$ or (iii) a combination of these two policy types (i) and (ii). Note carefully that restricting attention to symmetric policies located around some specific points $\mu + \gamma_i$ is without loss of generality, since we can also take care of asymmetric policies by simply redefining $\mu + \gamma_i$ appropriately. However, the concept can only apply to finite values, since otherwise we cannot

¯nd the "middle" of a prohibited interval. Thus interval controls are not the limit of these policies as t ! 1!

It is economically intuitive and proven formally in the appendix that prohibiting two intervals ($^1 + °_1$ ¡ t_1; $^1 + °_1 + t_1$) and ($^1 + °_2$ ¡ t_2; $^1 + °_2 + t_2$) with $^1 + °_2$ ¡ $t_2 > {}^1 + °_1 + t_1$ is a suboptimal policy, since we leave the agent the possibility of reacting in the wrong direction. But consequently, if it is optimal to exclude some speci¯c alternatives beyond a certain threshold $^1 + {}_¸$; then it must be optimal to exclude all alternatives beyond the threshold[19]. Hence interval controls are optimal.

5 Conclusions

The principal uses two types of instruments to control the agent's behavior. Agents which are known to have interests very much out of line with those of the principal have to be controlled with instruments restricting their misbehavior. Their choice set will be a narrow set around the optimal choice under ignorance. Obviously those agents will also deliver little e®ort, because they will hardly be able to use their improved knowledge in their own best interest. Agents with interests which parallel those of the Principal may be subject to another kind of restriction on their choice set. Since the principal bene¯ts very much from their increased e®orts he could try to make life unpleasent for them if they have to decide under ignorance. Such an e®ort enhancing control must therefore forbid the choice which is optimal under ignorance. The surprising thing is that a principal might ¯nd it optimal to use such instruments even if the agent is a perfect clone of himself!

The bottom line of our argument is that curtailing the discretion or limiting the authority of agents may have as favorable incentive e®ects as increasing their discretion or giving them more authority. Moreover, the impact on the expected payo® of the principal might well be positive. While Aghion and Tirole (1997) have found that delegating authority -thus increasing discretion- might well be a good thing, even if there are con°icting interests, the present paper points at the mirror image of this result: curtailing discretion might well be a good thing even if there are no con°icts of interests! As a consequence, there need not be a monotonic relationship between the severity of the con°ict of interest and the restrictiveness of the optimal decision space. An agent with better aligned interests need not be given a higher degree of freedom - even if he would always (on expectation) use this additional freedom in the principal's interest!

6 References

² Aghion, P. and Tirole, J. (1997) "Formal and Real Authority" Journal of Political Economy 105(1), 1-29

[19] Reallowing some selected alternatives within the interval control is e®ectively nothing else than (iii):

- Armstrong, Mark (1994) "Delegation and Discretion" Discussion Papers in Economics and Econometrics No. 9421, University of Southampton
- Dewatripont, M. and Tirole, J. (1999) "Advocates" Journal of Political Economy 107(1), 1-39
- Holmström, B. "On the Theory of Delegation" in Boyer, M. and Kihlstrom, R. eds. (1984) "Bayesian Models in Economic Theory" Elsevier Science Publishers B.V.
- Johnson, N. and Kotz, S. (1970) "Distributions in Statistics-Continous Univariate Distributions", Vol. 2, New York: Wiley
- Tirole, J. (1999) "Incomplete Contracts: Where Do We Stand?" Econometrica 67(4), 741-781

7 Appendix A

We consider controls of the type $\Gamma = \text{dom } f_\theta \setminus [-\infty; \mu - \lambda] \cup [\mu - \varepsilon; \mu + \varepsilon] \cup [\mu + \lambda; \infty]$. Γ effectively truncates the distribution of θ. In view of the quadratic payoff functions, the mathematical analysis will involve first and second moments of the resulting truncated distribution. We will use the following notation troughout:

Definitions and Notation:

- $dF_\theta = f(\theta)d\theta$; $\quad dF_{\varphi|\theta} = f(\varphi|\theta)d\varphi$
- $\Delta F_\varepsilon := F(\mu) - F(\mu - \varepsilon) = F(\mu + \varepsilon) - F(\mu)$
- $dG_\varepsilon := \frac{dF_\theta}{\Delta F_\varepsilon}$
- $\delta := \mu - \int_{\mu-\varepsilon}^{\mu} \theta \, dG_\varepsilon = \int_{\mu}^{\mu+\varepsilon} \theta \, dG_\varepsilon - \mu$
- $F_{\mu-\lambda} := 1 - F(\mu + \lambda) = F(\mu - \lambda)$
- $dG_\lambda := \frac{f(\theta)}{F_{\mu-\lambda}} d\theta$
- $\nu := \int_{\mu+\lambda}^{\infty} \theta \, dG_\lambda - \mu = \mu - \int_{-\infty}^{\mu-\lambda} \theta \, dG_\lambda$

$\frac{dG_\varepsilon}{d\theta}$ is the density of θ, conditional on the fact that $\theta \in [\mu - \varepsilon; \mu]$. $\mu - \delta$ is the mean of θ, conditional on $\theta \in [\mu - \varepsilon; \mu]$. Likewise, $\frac{dG_\lambda}{d\theta}$ is the probability density of θ, conditional on the fact that $\theta \in [\mu + \lambda; \infty]$. $\nu + \mu$ is the conditional mean of θ, conditional on the fact that $\theta \in [\mu + \lambda; \infty]$.

Proof of Lemma 1 ($\frac{\partial e}{\partial \varepsilon} > 0$):. We know that the agent's effort choice satisfies

$$\begin{aligned} g' &\stackrel{!}{=} EU(x'(\theta); \tilde{\theta}) - EU(x''; \tilde{\theta}) \\ &\stackrel{!}{=} \frac{A}{2}(\sigma^2 + \varepsilon^2) - A \int_{\mu}^{\mu+\varepsilon} (\mu + \varepsilon - \theta)^2 dF_\theta \end{aligned}$$

Hence

$$e = h\left(\frac{A}{2}(\sigma^2+\varepsilon^2) - A\int_{\mu}^{\mu+\varepsilon}(\mu+\varepsilon-\theta)^2 dF_\theta\right)$$

where $h := g'^{-1}$ exists by strict convexity of g:

$$\frac{\partial e}{\partial \varepsilon} = \frac{1}{g''(e)}\frac{\partial}{\partial \varepsilon}\left[\frac{A}{2}(\sigma^2+\varepsilon^2) - A\int_{\mu}^{\mu+\varepsilon}(\mu+\varepsilon-\theta)^2 dF_\theta\right]$$

By Leibniz's rule we have $\frac{\partial}{\partial \varepsilon}\left[\frac{A}{2}(\sigma^2+\varepsilon^2) - A\int_{\mu}^{\mu+\varepsilon}(\mu+\varepsilon-\theta)^2 dF_\theta\right]$

$$\begin{aligned} &= -A\left\{\int_{\mu}^{\mu+\varepsilon} 2(\mu+\varepsilon-\theta)f(\theta)d\theta - \varepsilon\right\} \\ &= A[\varepsilon - 2\cdot F_\varepsilon f_\varepsilon - \delta] \end{aligned}$$

Hence

$$\frac{\partial e}{\partial \varepsilon} > 0 \iff \varepsilon - 2\cdot F_\varepsilon f_\varepsilon - \delta > 0$$

Observe now that this will always be true except in pathological cases: We have $\delta \leq \varepsilon$ because a mean is a convex combination. $2\cdot F_\varepsilon \leq 1$ because at most half of the mass can lie in the upper half of the distribution. Thus only in the case where $\delta = 0$ and ε is set equal to the upper bound of the distribution effort will not increase. Observe however that $\delta = 0$ means that the distribution has point mass around μ; i.e. is degenerate on μ: We can thus safely ignore this case. ■

Proof of Proposition 1:. For the sake of clarity, we first state all derivatives needed for the evaluation of (12):

$$\frac{\partial e}{\partial \varepsilon} = \frac{A[\varepsilon - 2\cdot F_\varepsilon f_\varepsilon - \delta]}{g''(e)}$$

$$\frac{\partial^2 e}{\partial \varepsilon^2} = \frac{A(1-2\cdot F_\varepsilon)}{g''(e)} - \frac{A^2[\varepsilon - 2\cdot F_\varepsilon f_\varepsilon - \delta]^2}{g''(e)}\frac{g'''(e)}{g''(e)^2}$$

$$\frac{\partial}{\partial \varepsilon}E(\pi'_\varepsilon - \pi''_\varepsilon) = [\varepsilon - 2\cdot F_\varepsilon f_\varepsilon - \delta]$$

$$\frac{\partial^2}{\partial \varepsilon^2}E(\pi'_\varepsilon - \pi''_\varepsilon) = 1 - 2\cdot F_\varepsilon$$

(I) The principal will benefit from setting $\varepsilon > 0$ if

$$\lim_{\varepsilon\to 0}\frac{\partial E\pi''_\varepsilon}{\partial \varepsilon} + e\frac{\partial}{\partial \varepsilon}E(\pi'_\varepsilon - \pi''_\varepsilon) + \frac{\partial e}{\partial \varepsilon}E(\pi'_\varepsilon - \pi''_\varepsilon) > 0:$$

More specifically if

$$\lim_{\epsilon\to 0} -\epsilon + e[\epsilon - 2\cdot F_\epsilon f\epsilon - \delta g] + \frac{A[\epsilon - 2\cdot F_\epsilon f\epsilon - \delta g]}{g''(e)} E(\pi_\epsilon' - \pi_\epsilon'') > 0$$

or

$$\lim_{\epsilon\to 0}\left[\left(e(\epsilon) + A\frac{(\sigma^2+\epsilon^2) - 2\epsilon\int_\mu^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_\eta}{2g''(e(\epsilon))}\right)\times\frac{\epsilon - 2\cdot F_\epsilon f\epsilon - \delta g}{\epsilon}\right] > 1$$

By lemma (1) $\epsilon - 2(\epsilon - \delta)\cdot F_\epsilon > 0$. By l'Hôpital $\lim_{\epsilon\to 0}\frac{\epsilon - 2\cdot F_\epsilon f\epsilon - \delta g}{\epsilon} = 1$. $\lim_{\epsilon\to 0} (\sigma^2+\epsilon^2) - 2\epsilon\int_\mu^{\mu+\epsilon}(\mu+\epsilon-\eta)^2 dF_\eta = \sigma^2$. Thus the principal can benefit from ϵ-controls if $e(\epsilon=0) + \frac{A\sigma^2}{2g''(e(\epsilon=0))}$ is larger than 1, which proves the first claim.

(II)Given that it pays to increase ϵ from zero, we must now show that there exists one and only one finite ϵ which solves the first order condition.

Existence: If an $\epsilon^* > 0$ solving the principal's problem exists, it is the solution of the first order necessary condition:

$$\frac{\partial E\pi_\epsilon''}{\partial\epsilon} + e\frac{\partial}{\partial\epsilon}[E(\pi_\epsilon' - \pi_\epsilon'')] + \frac{\partial e}{\partial\epsilon}E(\pi_\epsilon' - \pi_\epsilon'') \overset{!}{=} 0 \qquad \text{(FOC)}$$

or

$$f\epsilon - 2\cdot F_\epsilon(\epsilon - \delta)g\left(e + \frac{AE(\pi_\epsilon' - \pi_\epsilon'')}{g''}\right) \overset{!}{=} \epsilon \qquad (16)$$

and satisfies the second order condition:

$$\begin{aligned} 0 > \; & \frac{\partial^2\pi''}{\partial\epsilon^2} + e\frac{\partial^2}{\partial\epsilon^2}[E(\pi_\epsilon' - \pi_\epsilon'')] + \\ & 2\frac{\partial e}{\partial\epsilon}\frac{\partial}{\partial\epsilon}[E(\pi_\epsilon' - \pi_\epsilon'')] + \frac{\partial^2 e}{\partial\epsilon^2}E(\pi_\epsilon' - \pi_\epsilon'') \end{aligned} \qquad \text{(SOC)}$$

or

$$\begin{aligned} 0 > \; & -1 + \left(e + \frac{AE(\pi_\epsilon' - \pi_\epsilon'')}{g''}\right)(1 - 2\cdot F_\epsilon) \\ & + \frac{A(\epsilon - 2\cdot F_\epsilon(\epsilon-\delta))^2}{g''}\left\{2 - \frac{Ag'''E(\pi_\epsilon' - \pi_\epsilon'')}{(g'')^2}\right\} \end{aligned} \qquad (17)$$

Define the left hand side of (16) as

$z := e f\epsilon - 2\cdot F_\epsilon(\epsilon - \delta)g + \frac{A(\epsilon - 2\cdot F_\epsilon(\epsilon-\delta))}{g''}E(\pi_\epsilon' - \pi_\epsilon'')$.

Consequently $z(0) = 0; \lim_{\epsilon\to 0}\frac{\partial z}{\partial\epsilon} = \lim_{\epsilon\to 0}\frac{z}{\epsilon} > 1$ by l'Hôpital and A4a. Existence requires that $\frac{\partial z}{\partial\epsilon}$ eventually be smaller than one, the problem thus eventually being concave. The problem is not globally concave. To see this, consider the second derivative of the principal's payoff function:

$$\frac{\partial^2\pi_\epsilon''}{\partial\epsilon^2} + e\frac{\partial^2}{\partial\epsilon^2}E(\pi_\epsilon' - \pi_\epsilon'') + 2\frac{\partial e}{\partial\epsilon}\frac{\partial}{\partial\epsilon}E(\pi_\epsilon' - \pi_\epsilon'') + \frac{\partial^2 e}{\partial\epsilon^2}E(\pi_\epsilon' - \pi_\epsilon'').$$

$\frac{\partial^2}{\partial\varepsilon^2}E(\pi_2' - \pi_2'') = 1 - 2\cdot F > 0$; $\frac{\partial e}{\partial\varepsilon}\frac{\partial}{\partial\varepsilon}E(\pi_2' - \pi_2'') > 0$. Finally $\frac{\partial^2 e}{\partial\varepsilon^2} = \frac{A}{g''}(1 - 2\cdot F) - \frac{A^2 f_\varepsilon - 2\cdot F_\varepsilon(\varepsilon - \delta)g^2}{g''}\frac{g'''}{g''^2}$. Now because $\lim_{\varepsilon\to 0}\frac{\partial z}{\partial\varepsilon} > 1$ we must also have $\lim_{\varepsilon\to 0}\frac{\partial^2 z}{\partial\varepsilon^2} > 1$; implying that the problem be convex for small ε. Let then $\frac{g'''}{g''^2}$ be bounded below and larger than $\frac{4}{A\sigma^2}$. Then the expression in the second line of (17): $\frac{A f_\varepsilon - 2\cdot F_\varepsilon(\varepsilon-\delta)g^2}{g''}\times\left[2 - \frac{Ag'''E(\pi_2' - \pi_2'')}{g''^2}\right]$ is smaller than zero. The expression in the first line is positive upon assumption for small ε. However, as ε is increased its influence must vanish: $\frac{\partial}{\partial\varepsilon}\left[(1 - 2\cdot F_\varepsilon)\left(e + \frac{AE(\pi_2' - \pi_2'')}{g''}\right)\right] =$

$$-2f_{1+\varepsilon}\left(e + \frac{AE(\pi_2' - \pi_2'')}{g''}\right) + \left\{2 - Ag'''\frac{E(\pi_2' - \pi_2'')}{g''^2}\right\}\times \frac{(1 - 2\cdot F_\varepsilon)A(\varepsilon - 2\cdot F_\varepsilon(\varepsilon - \delta))}{g''}$$

< 0. Therefore as ε goes out of bounds we must have $\frac{\partial^2 z}{\partial\varepsilon^2} < 0$. Consequently z and the 45°degree line must cross eventually, which proves existence of ε^*.

Uniqueness: Finally we prove that there is a unique value ε^*, which solves the first order condition. Since the first order condition implies at ε^* that

$$\left(e + A\frac{E(\pi_2' - \pi_2'')}{g''}\right)(1 - 2\cdot F_\varepsilon) - 1 = -\left(e + A\frac{E(\pi_2' - \pi_2'')}{g''}\right)2\cdot F_\varepsilon\frac{\delta}{2} < 0$$

and we know that $\frac{\partial}{\partial\varepsilon}\left(\left(e + A\frac{E(\pi_2' - \pi_2'')}{g''}\right)(1 - 2\cdot F_\varepsilon)\right) < 0$, z must be concave for all $\varepsilon > \varepsilon^*$. (It was only the first line in (17) that caused problems). Therefore (16) has exactly one solution.

(III) $\hat{\Gamma}$ is a symmetric restriction around the mean. Therefore we must consider candidates Γ which allow $x = \mu$ or/and are nonsymmetric. Consider thus any Γ with $\mu \in \Gamma$. The agent's choice under ignorance and his as well as the principal's payoff in this case are left unchanged. But if $\Gamma \not\subset \hat{\Gamma}$ there is a negative impact on the choices and the payoffs π' and U' in case the agent is informed. Therefore Γ decreases the wedge between U' and U'' and the agent supplies less effort. These policies are therefore dominated and $\mu \notin \Gamma$.

Consider then nonsymmetric restrictions $\Gamma = \text{dom } f \setminus [\mu - \varepsilon_1; \mu + \varepsilon_2]$. let $\varepsilon_1 < \varepsilon_2$ (w.l.o.g.). Consider the interval $[\mu + \varepsilon_1; \mu + \varepsilon_2]$. In case the agent is not informed there is no payoff relevant effect since the agent prefers to choose $x'' = \mu - \varepsilon_1$. In case he is informed the effect of excluding $[\mu + \varepsilon_1; \mu + \varepsilon_2]$ from the choice set of the agent is to reduce the payoff of the agent and the principal, and to reduce the difference of payoffs across the informed and uninformed state for the agent. Therefore the overall effect must be negative. This proves the claim. ■

Proof of Proposition 2. (i) We want to know whether an increase in A or

σ makes $\lim_{\varepsilon \to 0}\left(e + \frac{A E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e(\varepsilon))}\right) > 1$ more likely. Thus $\frac{\partial}{\partial A}\left(e + A\frac{E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e(\varepsilon))}\right) =$

$$\frac{E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e)}\left(2 - Ag'''(e)\frac{E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e)^2}\right) < 0 \Leftarrow \frac{g'''(e)}{g''(e)^2} > \frac{4}{A\sigma^2}; \forall e$$

The argument of the effect of σ^2 is exactly the same.
(ii) Consider the effect of an increase in A on the optimal size of the prohibited interval. Define the function

$$H(A; \varepsilon(A)) := \frac{\partial E\pi''}{\partial \varepsilon} + e\frac{\partial}{\partial \varepsilon}[E(\pi'_\varepsilon - \pi''_\varepsilon)] + \frac{\partial e}{\partial \varepsilon}E(\pi'_\varepsilon - \pi''_\varepsilon)$$

By the implicit function theorem: $\frac{\partial \varepsilon^*}{\partial A} = -\frac{\frac{\partial H}{\partial A}}{\frac{\partial H}{\partial \varepsilon}}$: The denominator is the second order condition, thus negative. $\frac{\partial H}{\partial A} = \frac{\partial e}{\partial A}\frac{\partial}{\partial \varepsilon}[E(\pi'_\varepsilon - \pi''_\varepsilon)] + \frac{\partial^2 e}{\partial \varepsilon \partial A}E(\pi'_\varepsilon - \pi''_\varepsilon) =$

$$\frac{E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''}\frac{\partial}{\partial \varepsilon}[E(\pi'_\varepsilon - \pi''_\varepsilon)] + \frac{(\varepsilon - 2 \cdot F_\varepsilon(\varepsilon - \delta))E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e)}\left(g'''(e)A\frac{E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e)^2} - 1\right)$$

Hence: $-\frac{\partial H}{\partial A} = \frac{(\varepsilon - 2 \cdot F_\varepsilon(\varepsilon - \delta))E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e)}\left(g'''(e)A\frac{E(\pi'_\varepsilon - \pi''_\varepsilon)}{g''(e)^2} - 2\right) > 0$ by assumption A4b. Hence $\frac{\partial \varepsilon^*}{\partial A} < 0$:
(iii) Next, redefine H as $H(\sigma; \varepsilon(\sigma))$. Then $\text{sign}\frac{\partial \varepsilon^*}{\partial \sigma} = \text{sign}\frac{\partial H}{\partial \sigma}$:

$$\frac{\partial H}{\partial \sigma} = \frac{\partial e}{\partial \sigma}\frac{\partial}{\partial \varepsilon}[E(\pi'_\varepsilon - \pi''_\varepsilon)] + e\frac{\partial^2}{\partial \varepsilon \partial \sigma}[E(\pi'_\varepsilon - \pi''_\varepsilon)] + \frac{\partial^2 e}{\partial \varepsilon \partial \sigma}E(\pi'_\varepsilon - \pi''_\varepsilon) + \frac{\partial e}{\partial \varepsilon}\frac{\partial}{\partial \sigma}[E(\pi'_\varepsilon - \pi''_\varepsilon)]$$

Consider all terms in order:

$$\frac{\partial e}{\partial \sigma} = \frac{A}{g''(e)}\left(\sigma + \int_\mu^{\mu+\varepsilon}(\mu + \varepsilon - \theta)^2\left[\frac{1}{\sigma} - \frac{(\theta - \mu)^2}{\sigma^3}\right]dF_\theta\right)$$

which is larger than zero for $\varepsilon < \sigma$ (which is natural, because nobody would ever want to prohibit an interval containing two thirds of the mass!)

$$\frac{\partial^2}{\partial \varepsilon \partial \sigma}[E(\pi'_\varepsilon - \pi''_\varepsilon)] = 2\int_\mu^{\mu+\varepsilon}(\mu + \varepsilon - \theta)\left[\frac{1}{\sigma} - \frac{(\theta - \mu)^2}{\sigma^3}\right]dF_\theta > 0:$$

$$\frac{\partial^2 e}{\partial \varepsilon \partial \sigma} = \frac{2A}{g''}\int_\mu^{\mu+\varepsilon}(\mu + \varepsilon - \theta)\left[\frac{1}{\sigma} - \frac{(\theta - \mu)^2}{\sigma^3}\right] - \frac{A(\varepsilon - 2 \cdot F(\varepsilon - \delta))}{g''(e)}\frac{Ag'''(e)}{g''(e)^2} \times \left(\sigma + \int_\mu^{\mu+\varepsilon}(\mu + \varepsilon - \theta)^2\left[\frac{1}{\sigma} - \frac{(\theta - \mu)^2}{\sigma^3}\right]dF_\theta\right):$$

The sign of this expression is ambigous. Finally $\frac{\partial}{\partial\sigma}[E(\pi'_2 - \pi''_2)] =$

$$\left(\sigma + \int_{\mu}^{\mu+\epsilon} (\mu+\epsilon-\theta)^2 \left[\frac{1}{\sigma} - \frac{(\theta-\mu)^2}{\sigma^3}\right] dF_\theta\right) > 0.$$

Thus, the overall effect is of no clear sign. ∎

General Derivation of Payoffs: The expressions in the text are not directly accessible to analysis. For the sake of completeness, the statistical derivation of a more workable form is provided here:
We consider $\Gamma = \text{dom}\{\theta \setminus [-1;\mu-\lambda); (\mu-\epsilon;\mu+\epsilon); (\mu+\lambda;1]\}$ with $\lambda \geq \epsilon$. The principal allows all choices within the intervals $[\mu+\epsilon;\mu+\lambda]; [\mu-\lambda;\mu-\epsilon]$. By symmetry total utility losses due to these instruments are double the losses in the upper half of the support of θ.
Derivation of utility losses due to prohibitions and controls: (1) controls: $EU'_\lambda = k - A\int_{\mu+\lambda}^{1}(\mu+\lambda-\theta)^2 dF_\theta$:

$$\begin{aligned} & AF_{\mu-\lambda}\int_{\mu+\lambda}^{1}(\mu+\lambda-\theta)^2 dG_\lambda \\ = \; & AF_{\mu-\lambda}\left\{\int_{\mu+\lambda}^{1}\theta^2 - (\mu+\kappa)^2 dG_\lambda + (\lambda-\kappa)^2\right\} \\ = \; & AF_{\mu-\lambda}\left\{Var(\theta \mid \theta\in[\mu+\lambda;1]) + (\lambda-\kappa)^2\right\} \end{aligned}$$

$$\begin{aligned} & A\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\theta)^2 f(\theta)d\theta \\ = \; & A\cdot F_\epsilon\left\{Var(\theta \mid \theta\in[\mu;\mu+\epsilon]) + (\epsilon-\delta)^2\right\} \end{aligned}$$

Derivation of $E\pi'_{2;\lambda}$: $E\pi'_{2;\lambda} = K -$

$$\left[\int_{\mu}^{\mu+\epsilon}\int(\mu+\epsilon-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta + \int_{\mu+\epsilon}^{\mu+\lambda}\int(\theta-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta + \int_{\mu+\lambda}^{1}\int(\mu+\lambda-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta\right] \tag{18}$$

Proceed term by term: $\int_{\mu+\epsilon}^{\mu+\lambda}\int(\theta-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta =$

$$(1-\rho)\sigma^2 - \int_{\mu}^{\mu+\epsilon}\int(\theta-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta - \int_{\mu+\lambda}^{1}\int(\theta-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta \tag{19}$$

since $2\int\int_{\mu}^{1}(\theta-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta = \int\int(\theta-\varphi)^2 dF_{\varphi|\theta}\,dF_\theta = 2(1-\rho)\sigma^2$.

Consider next $\int\int_{\mu+\lambda}^{1}(\mu+\lambda-\phi)^2 dF_{\phi|\eta}\, dF_\eta$:

$$
\begin{aligned}
&= \int\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta+\eta-\phi)^2 dF_{\phi|\eta}\, dF_\eta \\
&= \int\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)^2 dF_{\phi|\eta}\, dF_\eta + \int\int_{\mu+\lambda}^{1} (\eta-\phi)^2 dF_{\phi|\eta}\, dF_\eta + \\
&\quad 2\int\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)(\eta-\phi) dF_{\phi|\eta}\, dF_\eta
\end{aligned}
\tag{20}
$$

The term $\int\int_{\mu}^{\mu+\epsilon}(\mu+\epsilon-\phi)^2 dF_{\phi|\eta}\, dF_\eta$ can be expanded the very same way. Making use of (19) and (20), we can simplify (18) to yield:

$$
\begin{aligned}
&(1-\rho)\sigma^2 + \\
&\int\int_{\mu}^{\mu+\epsilon} (\mu+\epsilon-\eta)^2 dF_{\phi|\eta}\, dF_\eta + 2\int\int_{\mu}^{\mu+\epsilon} (\mu+\epsilon-\eta)(\eta-\phi) dF_{\phi|\eta}\, dF_\eta + \\
&\int\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)^2 dF_{\phi|\eta}\, dF_\eta + 2\int\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)(\eta-\phi) dF_{\phi|\eta}\, dF_\eta
\end{aligned}
$$

Consider now the term $\int\int_{\mu+\lambda}^{1}(\mu+\lambda-\eta)^2 dF_{\phi|\eta}\, dF_\eta$. Because we can integrate out over ϕ (since $\int dF_{\phi|\eta} = 1$) the term equals $\int_{\mu+\lambda}^{1}(\mu+\lambda-\eta)^2 dF_\eta$, which has been shown to equal $F_{1-\lambda}\left\{Var(\eta|\eta \in [\mu+\lambda, 1]) + (\lambda-\kappa)^2\right\}$. Hence

$$
\begin{aligned}
&\int\int_{\mu}^{\mu+\epsilon} (\mu+\epsilon-\eta)^2 dF_{\phi|\eta}\, dF_\eta + \int\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)^2 dF_{\phi|\eta}\, dF_\eta \\
= \;& \Delta F_\epsilon \left\{Var(\eta|\eta \in [\mu, \mu+\epsilon]) + (\epsilon-\delta)^2\right\} + \\
& F_{1-\lambda}\left\{Var(\eta|\eta \in [\mu+\lambda, 1]) + (\lambda-\psi)^2\right\}
\end{aligned}
$$

To evaluate the remaining terms, we make use of both the fact that the marginals are identical and that they are normal. As a consequence, conditional means take the form $\int \phi\, dF_{\phi|\eta} = (1-\rho)\mu + \rho\eta$. Consider then the term

$$
\begin{aligned}
&2\int\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)(\eta-\phi) dF_{\phi|\eta}\, dF_\eta \\
= \;& 2\int dF_{\phi|\eta} \int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)\eta\, dF_\eta - 2\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)\int \phi\, dF_{\phi|\eta}\, dF_\eta \\
= \;& 2F_{1-\lambda}\left\{\int_{\mu+\lambda}^{1} (\mu+\lambda-\eta)\eta - (\mu+\lambda-\eta)[(1-\rho)\mu+\rho\eta]\, dG_\lambda\right\} \\
= \;& -2F_{1-\lambda}(1-\rho)\left\{Var(\eta|\eta \geq \mu+\lambda) + \kappa^2 - \lambda\kappa\right\}
\end{aligned}
$$

Do the same for the analogous term involving ϵ and add all up to get: $E\pi^0_{a;\lambda,\kappa} =$ K

$$
-\left\{
\begin{array}{c}
(1-\rho)\sigma^2 + (2\rho-1)F_{1-\lambda}Var_\lambda + (2\rho-1)\Delta F_\epsilon Var_\epsilon + \\
F_{1-\lambda}(\kappa-\lambda)^2 + \Delta F_\epsilon(\epsilon-\delta)^2 - 2(1-\rho)F_{1-\lambda}\kappa(\kappa-\lambda) + \\
2(1-\rho)\Delta F_\epsilon \delta(\epsilon-\delta)
\end{array}
\right\}
\tag{21}
$$

The following lemma contains a very useful technical result. Because it is not by itself economically interesting, it has been relegated to the Appendix.

Lemma 2: $E\pi'_{2,\lambda}$ is a singlepeaked and quasiconcave function of λ. For $1 > \rho > 0$; it has an interior maximum, say at $\lambda = \bar{\lambda}(\rho)$ is concave for $\lambda \leq \tilde{\lambda}(\rho)$ and convex for $\lambda > \tilde{\lambda}(\rho)$:

Proof of Lemma 2. Straightforward differentiation of (21) shows that the derivative of $E\pi'_{2,\lambda}$ with respect to λ equals:

$$\frac{\partial E\pi'_{2,\lambda}}{\partial\lambda} = 2F_{\mu-\lambda}(\rho\nu - \lambda)$$

We want to show that for $\rho < 1$; there exists a $\bar{\lambda}$, such that $\frac{\partial\pi'_{2,\lambda}(\rho)}{\partial\lambda} < 0; \forall\lambda \geq \bar{\lambda}$: This can only be the case if there exists a λ; such that $\rho\nu < \lambda; \forall\lambda \geq \lambda$: Consider the equation $\rho\nu = \lambda$ or

$$\rho\sigma\frac{\phi(\frac{\lambda}{\sigma})}{1-\Phi(\frac{\lambda}{\sigma})} = \lambda$$

where ϕ and Φ are the p.d.f. and c.d.f., respectively, of the standard normal and we make use of the fact that $\mu + \nu$ is the first moment of a truncated normal distribution.[20] At $\lambda = 0$ the right hand side of the equation is zero obviously while the left hand side is strictly positive. The right hand side is increasing in λ with slope 1: The derivative of the left hand side is

$$\begin{aligned} & -\rho\frac{\frac{\lambda}{\sigma}\phi(\frac{\lambda}{\sigma})}{1-\Phi(\frac{\lambda}{\sigma})} + \frac{\rho\phi(\frac{\lambda}{\sigma})^2}{(1-\Phi(\frac{\lambda}{\sigma}))^2} \\ = & \frac{f_{\mu+\lambda}}{(1-F_{\mu+\lambda})}\rho(\nu-\lambda) \end{aligned}$$

thus increasing in λ: Start with the case $\rho = 1$: for $\rho = 1$ the equation cannot have a solution because $\nu \geq \lambda; \forall\lambda$: We can also state that $\nu' \leq 1; \forall\lambda$: To see this consider the change in the slope of the left hand side, i.e. ν'':

$$\begin{aligned} \frac{\partial}{\partial\lambda}\left(\frac{f}{1-F}\rho(\nu-\lambda)\right) = & \rho\frac{\partial}{\partial\lambda}\left(\frac{f}{1-F}\right)(\nu-\lambda) + \rho\frac{f}{1-F} \times \\ & \left(\frac{f}{1-F}(\nu-\lambda) - 1\right) \end{aligned}$$

It is well known that $\frac{\partial}{\partial\lambda}\left\{\frac{f}{1-F}\right\} > 0$ for a normal distribution. Thus if $\frac{f}{1-F}(\nu - \lambda) > 1$ the slope of $\rho\nu$ will increase more and more. Therefore $\nu'(\lambda) > 1$; for any finite $\tilde{\lambda}$; implies that $\nu'(\lambda) > 1$ for any $\lambda \geq \tilde{\lambda}$: But eventually this implies

[20]see Johnson and Kotz (1970) p. 81

$\lim_{\lambda\to\infty}(\pi-\lambda) > 0.$
Now

$$\pi-\lambda = \sigma\frac{\phi(\frac{\lambda}{\sigma})}{1-\Phi(\frac{\lambda}{\sigma})}-\lambda$$

Using l'Hôpital twice we see that

$$\lim_{\lambda\to\infty}\frac{\sigma\phi(\frac{\lambda}{\sigma})}{\lambda(1-\Phi(\frac{\lambda}{\sigma}))}=1 \qquad (22)$$

and therefore

$$\lim_{\lambda\to\infty}\pi-\lambda=0$$

(22) says $\lim_{\lambda\to\infty}\frac{\pi}{\lambda}=1.$ But then also $\lim_{\lambda\to\infty}\pi'=1$; again by l'Hôpital. This establishes that $\pi'(\lambda)\le 1, \forall\lambda.$
Decrease now ρ from 1. This has two effects: it shifts the function $\rho\pi$ down and it decreases its slope to $\rho\pi'$. Since $\pi'\le 1, \forall\lambda$, $\rho\pi'<1$ for $\rho<1, \forall\lambda$. Thus, the equation has exactly one solution, $\lambda(\rho)$; for $\rho<1$ and $\rho\pi-\lambda$ is negative for all $\lambda > \lambda^1(\rho)$. This finally establishes that $E\Pi'(\lambda)$ is a singlepeaked, quasiconcave function. Consider now $\frac{\partial^2 E\Pi'_{\varepsilon,\lambda}}{\partial\lambda^2} = 2((1-\rho)f_{\mu+\lambda}\lambda - F_{\mu-\lambda})$. This is positive if $\frac{f}{1-F_{\mu+\lambda}} > \frac{1}{(1-\rho)\lambda}$. Since the left hand side is increasing in λ, the right hand side is decreasing in λ, we will -if $\rho<1$- always find a value $\tilde{\lambda}$ such that $E\Pi'_{\varepsilon,\lambda}$ is convex for all $\lambda>\tilde{\lambda}$. This proves the claim. ∎

Proof of Proposition 3. For the sake of clarity, we first state all derivates. The results follow from straightforward differentiation of (21) and $e = h\left[A\frac{(\sigma^2+\varepsilon^2)}{2} - \int_{\mu}^{\mu+\varepsilon}(\mu+\varepsilon-\theta)^2 - \int_{\mu+\lambda}^{\infty}(\mu+\lambda-\theta)^2\right]$.
Derivatives with respect to ε:

$$\frac{\partial}{\partial\varepsilon}E\Pi'_{\varepsilon,\lambda} = 2\cdot F_\varepsilon(\rho\delta-\varepsilon)$$

$$\frac{\partial^2}{\partial\varepsilon^2}E\Pi'_{\varepsilon,\lambda} = -2\cdot F_\varepsilon - 2(1-\rho)f_{\mu+\varepsilon}\varepsilon$$

$$\frac{\partial e}{\partial\varepsilon} = \frac{A[\varepsilon-2\cdot F_\varepsilon f\varepsilon-\delta g]}{g''(e)}$$

$$\frac{\partial^2 e}{\partial\varepsilon^2} = \frac{A(1-2\cdot F_\varepsilon)}{g''(e)} - \frac{A^2[\varepsilon-2\cdot F_\varepsilon f\varepsilon-\delta g]^2}{g''(e)}\frac{g'''(e)}{g''(e)^2}$$

Derivatives with respect to λ:

$$\frac{\partial E\pi'_{\varepsilon,\lambda}}{\partial \lambda} = 2F_{1-\lambda}(\rho\nu - \lambda)$$

$$\frac{\partial^2}{\partial \lambda^2} E\pi'_{\varepsilon,\lambda} = 2\left((1-\rho)f_{1+\lambda}\lambda - F_{1-\lambda}\right).$$

$$\frac{\partial e}{\partial \lambda} = \frac{2AF_{1-\lambda}(\nu - \lambda)}{g''(e)}$$

$$\frac{\partial^2 e}{\partial \lambda^2} = -\frac{g'''}{g''}\left(\frac{2A}{g''}F_{1-\lambda}(\nu-\lambda)\right)^2 - \frac{2A}{g''}F_{1-\lambda} < 0$$

(ia) Existence of a solution follows from Weierstrass's Theorem. Before we proove uniqueness a.e. (since it uses the same arguments) let us first proove (ii). The change in the principal's payoff due to a change in ρ is given by

$$\left(\frac{\partial}{\partial \varepsilon}E\pi''_{\varepsilon} + e\frac{\partial}{\partial \varepsilon}\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right] + \frac{\partial e}{\partial \varepsilon}E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right)\frac{\partial \varepsilon}{\partial \rho} +$$
$$\left(e\frac{\partial}{\partial \lambda}E\pi'_{\varepsilon,\lambda} + \frac{\partial e}{\partial \lambda}E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right)\frac{\partial \lambda}{\partial \rho} +$$
$$e\frac{\partial}{\partial \rho}\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right]$$

The proof of proposition 1 showed that the solution for ε is an interior solution for $\rho = 1$. By continuity the same is true for ρ close to 1. Let us further assume for the moment that the solution for λ is an interior solution too. Then, by the envelope theorem, the first two lines are equal to zero. The final step is to show that $\frac{\partial}{\partial \rho}\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right] > 0, \forall \lambda, \varepsilon \geq 0, \lambda \geq \varepsilon$: $\frac{\partial}{\partial \rho}\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right] =$

$$\sigma^2 - 2F_{1-\lambda}Var_\lambda - 2\cdot F_\varepsilon Var_\varepsilon - 2F_{1-\lambda}\nu(\nu - \lambda) + 2\cdot F_\varepsilon\delta(\varepsilon - \delta)$$

As $\frac{\partial^2}{\partial \varepsilon \partial \rho}\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right] = 2\cdot F_\varepsilon \delta > 0$ and $\frac{\partial^2}{\partial \lambda \partial \rho}\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right] = 2(1 - F_{1+\lambda})v > 0$; the expression is minimized at $\varepsilon = \lambda = 0$: $\frac{\partial}{\partial \rho}\min\left[E(\pi'_{\varepsilon,\lambda} - \pi''_{\varepsilon,\lambda})\right] =$

$$\sigma^2 - \int_{\mu}^{\infty} \eta^2 - (\mu + \nu)^2 \frac{dF_\eta}{\frac{1}{2}} - \nu^2$$
$$= \sigma^2 - \left(\sigma^2 - \nu^2\right) - \nu^2 = 0.$$

Assume then that the solutions are corner solutions. Since at $\varepsilon = 0$ all derivatives with respect to ε are exactly equal to zero, the same argument as above applies. As for $\lambda = \lambda^{max}$, both $\frac{\partial E\pi'_{\varepsilon,\lambda}}{\partial \lambda}$ and $\frac{\partial e}{\partial \lambda}$ tend to zero as $\lambda \to \infty$. Then again, the only term we need consider is $\frac{\partial}{\partial\rho}\left[E(\pi'_{\varepsilon;\lambda} - \pi''_{\varepsilon;\lambda})\right]$. This proves the claim.

(ib) Consider now uniqueness. In the case of ε, uniqueness was proven in proposition 1 for the case of $\rho = 1$. By continuity the same arguments hold also true for ρ close to 1. Moreover, it will be shown below that $\varepsilon > 0$ if and only if ρ is close to 1. Consider then uniqueness in the case of λ: since $\nu \geq \lambda$, $\frac{\partial e}{\partial\lambda} \geq 0$. Furthermore we know from lemma 2 that $\frac{\partial\pi'}{\partial\lambda}$ is positive for λ small and $\rho > 0$. Hence, the principal's objective must either reach an interior maximum or it must increase till $\lambda = \lambda^{max}$. In the latter case the solution is obviously unique. In the former, we have at least one interior solution and a sufficient condition for uniqueness is that the problem be globally concave. Consider the "first" interior extremum, i.e. the smallest value λ^*, satisfying $e\frac{\partial\pi'}{\partial\lambda} + \frac{\partial e}{\partial\lambda}(\pi' - \pi'') \overset{!}{=} 0$. Since the first extremum is a max, the second order condition holds locally:

$$\frac{\partial^2 e}{\partial\lambda^2}(\pi' - \pi'') + 2\frac{\partial e}{\partial\lambda}\frac{\partial E\pi'}{\partial\lambda} + e\frac{\partial^2 E\pi'}{\partial\lambda^2} < 0 \tag{23}$$

However, the problem need not be globally concave. From Lemma 2 we know that $\frac{\partial\pi'}{\partial\lambda} < 0$ for all $\lambda > \lambda^1(\rho)$. Hence, the first and the second term in (23) are negative as $\frac{\partial\pi'}{\partial\lambda} < 0$; $\frac{\partial e}{\partial\lambda} > 0$; and

$$\frac{\partial^2 e}{\partial\lambda^2} = -\frac{g'''}{g''}\left[\frac{2A}{g''}F_{\mu-\lambda}(\nu - \lambda)\right]^2 - \frac{2A}{g''}F_{\mu-\lambda} < 0$$

However, if $\rho < 1$, $\frac{\partial^2 E\pi'}{\partial\lambda^2} = 2((1-\rho)f_{\mu+\lambda}\lambda - F_{\mu-\lambda})$ is larger than zero for λ sufficiently large. It is thus possible - and we have found no way to exclude this possibility in general - that there are several interior solutions. However, it is shown below that this is not an issue for often used effort cost functions. (see below). Assume nevertheless for the sake of the argument that there are several local extrema. In this case the principal's payoff must be evaluated at all these values of λ satisfying the first and second order condition.

In this case we can only show uniquness a.e.: let $\lambda_1^*(\rho)$ and $\lambda_2^*(\rho)$ be two values of λ which both maximize the principal's payoff. This cannot be the case except at a finite number of disconnected points in $[0; 1]$; i.e. values of ρ. To see this, consider the change in the principal's payoff at the two interior solutions as ρ increases. By the argument in (ii) this is equal to

$$\frac{\partial E(\pi'_{\varepsilon;\lambda} - \pi''_{\varepsilon;\lambda})}{\partial\rho} > 0$$

By the fact that $\frac{\partial^2 E(\pi'_{\varepsilon;\lambda} - \pi''_{\varepsilon;\lambda})}{\partial\rho\partial\lambda} > 0$ (see above) the change in expected payoff for the principal is always higher at higher values of ρ. Therefore this implies that $\lambda(\rho + d\rho)$ gives the principal a higher payoff than $\lambda(\rho + d\rho)$. This proves the claim.

(iii) The first order condition in this case is given by

$$-\varepsilon + e\{\varepsilon - 2\cdot F_{\varepsilon}(\varepsilon - \rho\delta)\} + \frac{\partial e}{\partial \varepsilon}\times E(\pi'_{2;\lambda} - \pi''_{2;\lambda}) \stackrel{!}{=} 0 \qquad (24)$$

while the second order condition is

$$-1 + e(1 - 2\cdot F_{\varepsilon} - 2(1-\rho) f_{1+\varepsilon}\varepsilon) + 2\frac{\partial e}{\partial \varepsilon}\{\varepsilon - 2\cdot F_{\varepsilon}(\varepsilon - \rho\delta)\} + \frac{\partial^2 e}{\partial \varepsilon^2} E(\pi'_{2;\lambda} - \pi''_{2;\lambda}) \qquad (25)$$

< 0. Define

$$H(\varepsilon;\varepsilon(\rho)) := \frac{\partial \pi''}{\partial \varepsilon} + e\frac{\partial}{\partial \varepsilon}\left[E(\pi'_{2;\lambda} - \pi''_{2;\lambda})\right] + \frac{\partial e}{\partial \varepsilon} E(\pi'_{2;\lambda} - \pi''_{2;\lambda}).$$

By the implicit function theorem $\frac{\partial \varepsilon}{\partial \rho} = \frac{-\frac{\partial H(\varepsilon)}{\partial \rho}}{\frac{\partial H(\varepsilon)}{\partial \varepsilon}}$. The denominator is the second order condition. It was proved that the solution is a maximum for $\rho = 1$. By continuity this is also true for ρ close to 1. Thus, the denominator is negative. Therefore $\operatorname{sign}\left[\frac{\partial \varepsilon}{\partial \rho}\right] = \operatorname{sign}\left[\frac{\partial H(\varepsilon)}{\partial \rho}\right]$.

$$\frac{\partial H(\varepsilon)}{\partial \rho} = e\frac{\partial^2}{\partial \varepsilon \partial \rho}\left[E(\pi'_{2;\lambda} - \pi''_{2;\lambda})\right] + \frac{\partial e}{\partial \varepsilon}\frac{\partial}{\partial \rho}\left[E(\pi'_{2;\lambda} - \pi''_{2;\lambda})\right].$$

$$\frac{\partial^2}{\partial \varepsilon \partial \rho}\left[E(\pi'_{2;\lambda} - \pi''_{2;\lambda})\right] = e2\cdot F_{\varepsilon}\delta > 0.$$

$\left[\frac{\partial}{\partial \rho} E(\pi'_{2;\lambda} - \pi''_{2;\lambda})\right] \geq 0, \forall \lambda, \varepsilon \geq 0, \lambda \geq \varepsilon$ by (ii) above. This proves the claim.

(v) λ^* is finite for $\rho < 1$: $e\frac{\partial E\pi'}{\partial \lambda} + \frac{\partial e}{\partial \lambda} E(\pi'_{2;\lambda} - \pi''_{2;\lambda}) = 0$. e and $E(\pi'_{2;\lambda} - \pi''_{2;\lambda})$ both attend fixed values for $\lambda \to \infty$. Consider $\lim_{\lambda\to\infty}\left(\frac{\frac{\partial E\pi'}{\partial \lambda}}{\frac{\partial e}{\partial \lambda}}\right)$:

$$\lim_{\lambda\to\infty} \frac{2(1 - F_{1+\lambda})(\rho\kappa - \lambda)}{\frac{2A}{g(e)''}(1 - F_{1+\lambda})(\kappa - \lambda)} = -\infty$$

Thus, the increase in $E\pi'_{2;\lambda}$ overcompensates the decrease in e.

(vi) To see that $\frac{d\lambda}{d\rho} > 0$, consider first local changes. Define the function

$$H(\rho;\lambda(\rho)) := e\frac{\partial \pi'}{\partial \lambda} + \frac{\partial e}{\partial \lambda} E(\pi'_{2;\lambda} - \pi''_{2;\lambda}).$$

By the familiar argument: $\frac{\partial\lambda}{\partial\rho} = -\frac{\frac{\partial H}{\partial\rho}}{\frac{\partial H}{\partial\lambda}}$. Because the second order condition holds locally at a maximum, the denominator is negative and sign $\left[\frac{\partial\lambda}{\partial\rho}\right]$ = sign $\left[\frac{\partial H}{\partial\rho}\right]$.

$$\frac{\partial H}{\partial\rho} = \left[\frac{\partial e}{\partial\lambda}\frac{\partial E\pi'_{2;\lambda}}{\partial\rho} + e\frac{\partial^2 E\pi'_{2;\lambda}}{\partial\lambda\partial\rho}\right]$$

$\frac{\partial E\pi'_{2;\lambda}}{\partial\rho} > 0$ (all $\varepsilon;\ \lambda \geq 0;\ \lambda \geq \varepsilon$) by the argument in (ii). There it has also been established that $\frac{\partial^2 E\pi'_{2;\lambda}}{\partial\lambda\partial\rho} = 2(1 - F_{\mu+\lambda})\nu > 0$. Hence $\frac{\partial\lambda^*}{\partial\rho} > 0$.
Assume again that the principal's problem has two interior solutions. Even in this case we must have $\frac{d\lambda}{d\rho} \geq 0$ by the same argument as above, because $\frac{\partial^2 E(\pi'_{2;\lambda} - \pi''_{2;\lambda})}{\partial\rho\partial\lambda} > 0$ any discrete jumps will be rightwards. ■

An Example to Proposition 3: The problem has only one local extremum in the relevant range if we restrict attention to the family of effort cost functions $g = c\frac{e^i}{i};\ i \geq 3$.[21]
In this case we have

$$e = \left[\frac{A}{c}\left\{\frac{\sigma^2}{2} - (1 - F_{\mu+\lambda})(Var_\lambda + (\nu - \lambda)^2)\right\}\right]^{\frac{1}{i-1}}$$

$$\frac{\partial e}{\partial\lambda} = \frac{1}{i-1}e^{\frac{1}{i-1}-1}\frac{2A}{c}(1 - F_{\mu+\lambda})(\nu - \lambda)$$

The first order condition can then be expressed as

$$z(\lambda) \overset{!}{=} -\frac{(\nu - \lambda)}{(\rho\nu - \lambda)} \tag{26}$$

where

$$z(\lambda) = (i-1)\frac{\frac{\sigma^2}{2} - (1 - F_{\mu+\lambda})(Var_\lambda + (\nu - \lambda)^2)}{E(\pi'_{2;\lambda} - \pi''_{2;\lambda})}$$

The solution to the problem must satisfy $\frac{\partial E\pi'_{2;\lambda}}{\partial\lambda} < 0$ and $E(\pi'_{2;\lambda} - \pi''_{2;\lambda}) \geq 0$ thus both the left hand side and the right hand side of (26) are nonnegative. Let $\bar\lambda$ solve $\rho\nu - \lambda = 0$ and $\hat\lambda := \min[\lambda^{max}; \grave\lambda]$, where $\grave\lambda$ solves $E(\pi'_{2;\lambda} - \pi''_{2;\lambda}) = 0$. Now for $\rho > 0;\ \grave\lambda > \bar\lambda$. Next observe that $\lim_{\lambda\to\bar\lambda} -\frac{(\nu-\lambda)}{(\rho\nu-\lambda)} = \infty$; while $\infty > z(\bar\lambda) > 0$; $\lim_{\lambda\to\hat\lambda} z(\lambda) = \infty$ iff $\hat\lambda = \grave\lambda$; while $\infty > -\frac{(\nu(\grave\lambda)-\grave\lambda)}{(\rho\nu(\grave\lambda)-\grave\lambda)} > 0$. As $\frac{(\nu-\lambda)(\rho\nu'-1)-(\nu'-1)(\rho\nu-\lambda)}{(\rho\nu-\lambda)^2} < 0$ and
$\frac{\partial z}{\partial\lambda} = (i-1)\frac{\frac{1}{A}\frac{\partial(U'-U'')}{\partial\lambda}E(\pi'_{2;\lambda} - \pi''_{2;\lambda}) - \frac{1}{A}(U'-U'')\frac{\partial E(\pi'_{2;\lambda} - \pi''_{2;\lambda})}{\partial\lambda}}{(E(\pi'_{2;\lambda} - \pi''_{2;\lambda}))^2} > 0$ (26) has exactly

[21] Instead of the Inada condition $\lim_{e\to\infty} g' = \infty$ we ensure $e^* < \infty$ by choice of c.

one solution. If $\hat{\lambda} = \lambda^{max}$ then we must let λ go to infinity: $\lim_{\lambda\to\infty} z(\lambda) = \frac{i-1}{2\rho-1}$; while $\lim_{\lambda\to\infty} \frac{\gamma-\lambda}{\lambda-\rho\gamma} = 0$. Observe that $\frac{i-1}{2\rho-1} > 0$ if $\rho > \frac{1}{2}$. But then there always exists exactly one solution $\lambda < \lambda^{max}$ that solves (26) as long as $\rho < 1$. In particular this also shows that $\lambda = \lambda^{max}$ if and only if $\rho = 1$. This example shows that everything works well, i.e. the product $e(\lambda)E(\pi'_{2;\lambda} - \pi''_{2;\lambda})$ can achieve only one extremum in the relevant range under quite mild assumptions about effort costs.

Proof of Proposition 4. Consider the candidate policy $\Gamma_t = \text{dom } f(\eta)\setminus(\mu+\gamma-t;\mu+\gamma+t)$. By the familiar arguments $E\pi'_t = K - (1-\rho)\sigma^2 -$

$$\begin{pmatrix} \int_{\mu+\gamma-t}^{\mu+\gamma}(\mu+\gamma-t-\eta)^2 dF_\eta + \int_{\mu+\gamma}^{\mu+\gamma+t}(\mu+\gamma+t-\eta)^2 dF_\eta \\ +2(1-\rho)\int_{\mu+\gamma-t}^{\mu+\gamma}(\mu+\gamma-t-\eta)(\eta-\mu)dF_\eta \\ +2(1-\rho)\int_{\mu+\gamma}^{\mu+\gamma+t}(\mu+\gamma+t-\eta)(\eta-\mu)dF_\eta \end{pmatrix}$$

Taking derivatives with respect to t: $\frac{\partial}{\partial t}E\pi'_t =$

$$(-1)\left(\begin{array}{l} (-2)\int_{\mu+\gamma-t}^{\mu+\gamma}(\mu+\gamma-t-\eta)dF_\eta - 2(1-\rho)\int_{\mu+\gamma-t}^{\mu+\gamma}(\eta-\mu)dF_\eta \\ +2\int_{\mu+\gamma}^{\mu+\gamma+t}(\mu+\gamma+t-\eta)dF_\eta + 2(1-\rho)\int_{\mu+\gamma}^{\mu+\gamma+t}(\eta-\mu)dF_\eta \end{array}\right) \tag{27}$$

We must show that the term inside the brackets is nonnegative. If ρ is equal to 1, the term inside the brackets becomes $(-2)\int_{\mu+\gamma-t}^{\mu+\gamma}(\mu+\gamma-t-\eta)dF_\eta + 2\int_{\mu+\gamma}^{\mu+\gamma+t}(\mu+\gamma+t-\eta)dF_\eta$, which can never be negative: in the first integrand $(\mu+\gamma-t-\eta)$, all realizations of $\tilde{\eta}$ are larger than $\mu+\gamma-t$, in the second all realizations are smaller than $\mu+\gamma+t$. (Since marginals are identical, this is also a proof for the case where distributions are not normal. If $\rho < 1$, this is no longer true, since we have used the fact that conditional means are linear.) Making use of the fact that the distribution is normal the term inside the brackets is equal to

$$\begin{array}{l} -2(\gamma-t)\left(F_{\mu+\gamma} - F_{\mu+\gamma-t}\right) + 2\rho\sigma\left(\phi\left(\frac{\gamma-t}{\sigma}\right) - \phi\left(\frac{\gamma}{\sigma}\right)\right) \\ +2(\gamma+t)\left(F_{\mu+\gamma+t} - F_{\mu+\gamma}\right) - 2\rho\sigma\left(\phi\left(\frac{\gamma}{\sigma}\right) - \phi\left(\frac{\gamma+t}{\sigma}\right)\right) \end{array} \tag{28}$$

We must show that this is positive. The first line in (28) is positive iff

$$(\gamma-t) < \rho\sigma\frac{\left(\phi\left(\frac{\gamma-t}{\sigma}\right) - \phi\left(\frac{\gamma}{\sigma}\right)\right)}{(F_{\mu+\gamma} - F_{\mu+\gamma-t})} = \rho E\left(\eta \,\middle|\, \tilde{\eta} \in [\frac{\gamma-t}{\sigma};\frac{\gamma}{\sigma}]\right)$$

hence iff ρ is large enough. The second line in (28) is positive iff

$$(\gamma+t) > \rho E\left(\eta \,\middle|\, \tilde{\eta} \in [\frac{\gamma}{\sigma};\frac{\gamma+t}{\sigma}]\right)$$

hence for all values of ρ. In particular these two things imply that $\frac{\partial}{\partial t}E\pi_t^0 < 0$ for ρ close to 1.
If ρ is equal to zero, we must only consider

$$-2(\gamma - t)(F_{\mu+\gamma} - F_{\mu+\gamma-t}) + 2(\gamma + t)(F_{\mu+\gamma+t} - F_{\mu+\gamma})$$

This term will be positive if σ is large enough. $(\gamma + t) - (\gamma - t) = 2t$.
However $(F_{\mu+\gamma} - F_{\mu+\gamma-t}) - (F_{\mu+\gamma+t} - F_{\mu+\gamma}) > 0$ since an interval of size t contains more mass the lower the location of the interval. Therefore we must make sure that $(F_{\mu+\gamma+t} - F_{\mu+\gamma})$ is not "much" smaller than
$(F_{\mu+\gamma} - F_{\mu+\gamma-t})$. This can be done by choosing σ large enough, since

$$\lim_{\sigma \to \infty} (F_{\mu+\gamma} - F_{\mu+\gamma-t}) - (F_{\mu+\gamma+t} - F_{\mu+\gamma}) = 0.$$

To sum up: if σ is large, we cannot find a γ such that F is "very steep" in the interval $[\mu + \gamma - t, \mu + \gamma]$ and "very flat" in the interval $[\mu + \gamma, \mu + \gamma + t]$. has no "steep" parts.
Since (28) is linear in ρ, it is either minimized at $\rho = 1$ or at $\rho = 0$, depending on

$$\operatorname{sign}\left(\phi\left(\frac{\gamma - t}{\sigma}\right) + \phi\left(\frac{\gamma + t}{\sigma}\right) - 2\phi\left(\frac{\gamma}{\sigma}\right)\right)$$

But then, provided that σ is large enough, the above arguments imply that $\frac{\partial}{\partial t}E\pi_t^0 < 0$ for all values of ρ and $t \geq 0$. This completes the proof. ∎

8 Appendix B

(i) Equivalence of nonmonetary and monetary decision based rewards:
Suppose that the principal can write monetary contracts, but that benefits are noncontractible. The agent is infinitely risk averse with respect to financial income and derives utility $V(w(x))$ from such income. V thus satisfies: $V' > 0$, $V'' < 0$. It is easy to see that - everywhere where $w()$ is differentiable - we must have $\frac{\partial w}{\partial x} = 0$ for all x that the principal might want the agent to choose: For if $\frac{\partial w}{\partial x} \neq 0$, the realized wage would depend on η, hence would be risky ex ante. Given the infinite risk aversion of the agent this must be suboptimal. Suppose that the principal does not want the agent to choose an alternative $x \in (\mu - \varepsilon, \mu + \varepsilon)$. He can guarantee this by using the following wage scheme $w_{M,N}$:=

$$\begin{aligned} w_M &= -M \text{ for } x \in (\mu - \varepsilon, \mu + \varepsilon) \\ w_N &= N \text{ otherwise} \end{aligned}$$

for M and N large enough. Provided that $N - (-M)$ is large enough, the agent will never choose an $x \in (\mu - \varepsilon, \mu + \varepsilon)$. There is thus no risk in realized monetary income, since the agent can guarantee himself a fixed income in all states. Thus, the wage scheme (w_M, w_N) can implement exactly the same outcome, in terms of e and x as the contract ¡ does when the agent does not respond to monetary incentives at all.

(ii) The case of a binding IR-constraint: Assumption 4a and Proposition 1 compare only the direct costs and bene¯ts on the principal's utility of introducing interval prohibitions. The cost of increased e®ort is born entirely by the agent. Since, if the agent does not respond to monetary incentives, there is no way of compensating the agent for increased e®ort, we had to assume that the IR-constraint was nonbinding. However, there is a straightforward extension to proposition 1, when monetary transfers have value. Assume that V is unbounded below such that, at the optimal solution of the principal's problem, the IR-constraint of the agent will be binding. An increase in e must therefore be compensated by an increase in N: By proposition 1 $eEU^0 + (1 ¡ e)EU^{00}$ will be higher under a wage scheme $w_{M;N}$ than under a °at wage scheme, w_N; say. But lets assume that the principal compensates the entire additional e®ort cost. Then introducing $w_{M;N}$ will result in an increased N; according to

$$\frac{@N}{@²} = \frac{1}{\frac{@V}{@N}} \frac{@g}{@e} \frac{@e}{@²}:$$

Assume that the marginal utility of income for the agent is at least 1. To overstate the matter, let $\frac{@V}{@N} = 1$ at the value of N that has the IR constraint exactly binding for ² = 0: In this case the principal will set ² > 0 if

$$e + \frac{A¾^2}{g^{00}(e)2} ¡ g^0(e)\frac{A}{g^{00}(e)} > 1$$

or since $g^0(e) = \frac{A¾^2}{2}$

$$e + \frac{A¾^2}{g^{00}(e)2}(1 ¡ A) > 1 \qquad \text{(A4a')}$$

This condition now compares both direct and indirect costs: with a probability 1 ¡ e the principal just shoots himself in the foot by setting ² > 0: In addition he now bears additional wage costs of $\frac{A^2¾^2}{g(e)^{00}2}$: The bene¯t is as before, the marginally avoided risk: $\frac{A¾^2}{g(e)^{00}2}$: The derivative of the left hand side of this inequality is

$$(1 ¡ A)\frac{¾^2}{g(e)^{00}2} ½ 2 ¡ \frac{A¾^2}{2}\frac{g(e)^{000}}{g(e)^{002}} ¾ < 0:$$

A4a' is harder to satisfy than A4a but again more likely to hold the smaller A: