DISCUSSION PAPER SERIES

No. 5289

OPTIMAL DELEGATION

Ricardo Alonso and Niko Matouschek

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

OPTIMAL DELEGATION

Ricardo Alonso, Kellogg Graduate School of Management, Northwestern University and CEPR
Niko Matouschek, Kellogg Graduate School of Management, Northwestern University and CEPR

Discussion Paper No. 5289
October 2005

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 5289

October 2005

ABSTRACT

Optimal Delegation*

We analyse the optimal delegation of decision rights by a uninformed principal to an informed but biased agent. When the principal cannot use message-contingent transfers, she offers the agent a set of decisions from which he can choose his preferred one. We fully characterize the optimal delegation set for general distributions of the state space and preferences with arbitrary continuous state-dependent biases. We also provide necessary and sufficient conditions for particular delegation sets to be optimal. Finally, we show that the optimal delegation set takes the form of a single interval if the agent's preferences are sufficiently similar to the principal's.

JEL Classification: D82 and L23
Keywords: decision rights, delegation and mechanism design

Ricardo Alonso
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60201
USA
Email: r-alonso@kellogg.northwestern.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=161747

Niko Matouschek
Management and Strategy
Kellogg School of Management
Northwestern University
2001 Sheridan Rd
Evanston, IL 60208
USA
Tel: (1 847) 491 4166
Fax: (1 847) 467 1777
Email:
n-matouschek@kellogg.northwestern.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=144676

* We are very grateful to Jim Dana, Joshua Gans, Marco Ottaviani, Scott Schaefer and Kathy Spier for their feedback. We also would like to thank seminar participants at the University of Chicago's GSB, Harvard-MIT and the Kellogg School of Management as well as participants of the 2005 ALEA Annual Meetings, the IDEI Conference in Tribute to Jean-Jacques Laffont, the 2005 CEPR European Summer Symposium in Economic Theory and the 2005 Stony Brook Workshop on Law and Economics for their comments and suggestions. All errors are our own.

Submitted 21 September 2005

1 Introduction

The delegation of decision rights to subordinates is a pervasive feature of organizations. Such delegation, however, can take many different forms. CEOs who delegate investment decisions to their division managers, for instance, employ a variety of different capital budgeting rules that specify what decisions the division managers are allowed to make. As a result, some division managers can make essentially any decision whereas others face various constraints, such as upper limits on the size of the investments they are permitted to make.[3] Similarly, until recently the United States Congress employed a variety of mandatory sentencing guidelines that specified precisely what sentences judges could impose on an offender who had been found guilty of a particular crime.[4]

What decisions should a subordinate be allowed to make? A key trade-off that a principal faces when answering this question is between the agency cost of biased decision making and the benefit of utilizing the agent's information (Holmström 1977, 1984; Jensen and Meckling 1992). In designing optimal capital budgeting rules, for instance, CEOs must balance the need to control the investment decisions made by potentially biased division managers with the desire to make these decisions sensitive to the managers' information. Similarly, while proponents of federal sentencing guidelines argue that they are necessary to limit the influence of the political and moral beliefs of individual judges, opponents contend that they restrict judges' abilities to set sentences that fit the specifics of the crime. In this paper we study the optimal delegation of decision rights by a principal who faces this information-control trade-off. In particular, we investigate the optimal rules that an uninformed principal puts in place to constrain the decisions that a better informed but biased agent is permitted to make.

Our model has three main features. First, there is a principal and an agent who have different preferences over a decision that has to be made. The decision they each prefer,

[3] A large number of studies have described the capital budgeting rules that firms employ. See, in particular, Bower (1970).

[4] The 1984 Sentencing Reform Act introduced mandatory federal sentencing guidelines. In January 2005 the US Supreme Court ruled that these guidelines can only be advisory and not mandatory. See, for instance, "Supreme Court Transforms use of Sentence Guidelines," in *The New York Times*, 13 January 2005.

and how large a loss they incur if a different decision is made, depends on the state of the world. Second, there is a mismatch between authority and information: while the principal has the legal right to make the decision, only the agent is informed about the state of the world. Third, the principal cannot use message-contingent transfers to elicit information from the agent.

In this setting, the principal specifies a delegation set, i.e. a compact set of decisions from which the agent can choose his preferred one. In the sentencing case, for instance, the delegation set could be any compact subset of the set of feasible decisions. The Congress could, for example, specify that an offender found guilty of manslaughter can be sentenced to either three years, anywhere between five to ten years or thirteen years. The judge can then choose his preferred sentence from this delegation set. Similarly, when delegating an investment decision to a division manager, a CEO could offer any compact subset of the set of feasible decisions as a delegation set.

The main contribution of this paper is to fully characterize the optimal delegation set that maximizes the principal's expected utility for general distributions and utility functions with arbitrary continuous state-dependent biases. We also provide necessary and sufficient conditions for the optimality of particular delegation sets, such as centralization – in which case the only decision the agent is allowed to choose is the principal's preferred decision given her prior – and interval delegation – in which case the agent can make any decision in-between two thresholds. Moreover, we show that interval delegation is optimal as long as the agent's preferences are 'sufficiently similar' to the principal's. Finally, we present a normal-linear example and investigate the effect of changes in the economic environment on the agent's optimal degree of discretion.

Our paper also makes a methodological contribution. In the delegation problem that we investigate the principal optimizes over delegation *sets.* This precludes us from using standard optimization techniques in solving the problem. Instead we characterize the solution by investigating the effect on the principal's expected utility of adding decisions to, and removing decisions from, a delegation set. To our knowledge this is the first paper in the economics literature that uses this method to solve such an optimization problem.

The key departure of our setup from the standard mechanism design literature is the assumption that message-contingent transfers are not feasible. This assumption is common in the literature on delegation (see, for instance, Holmström 1977, 1984; Melumad and Shibano 1991; Aghion and Tirole 1997; Dessein 2002; Szalay 2004; Martimort and Semenov 2005)[5] and can be justified formally by assuming that the agent is infinitely risk averse to income shocks. It can also be justified from a positive perspective by noting that in many settings message-contingent transfers do not get used. For instance, the salaries of judges are not directly linked to the sentences they impose. The aim of this paper is to investigate the optimal delegation of decision rights taking as given that message-contingent transfers are not feasible.

The model has a number of applications. It could, for example, be applied to the before-mentioned federal sentencing guidelines (Shavell 2005) and capital budgeting rules (Harris and Raviv 1996, 1998; Marino and Matsusaka 2004). Other applications include the delegation of a pricing policy by the government to a regulatory agency (Armstrong 1995), the interaction between congressional committees and the Congress (Gilligan and Krehbiel 1987; Krishna and Morgan 2001) and the interaction between lobby groups and policy makers (Martimort and Semenov 2005). In all these situations the use of contingent monetary transfers is limited. Thus, a model in which such transfers are ruled out may provide a useful benchmark to investigate such situations.

The papers most closely related to ours are Holmström (1977, 1984). He considers a setting that is very similar to ours and proves the existence of an optimal delegation set. He does not, however, characterize the optimum. Instead he restricts the set of feasible delegation sets to intervals and characterizes optimal interval delegation sets.[6] Melumad and Shibano (1991) do solve for the optimal delegation set but restrict attention to the uniform distribution and particular preferences. Armstrong (1995) considers a model similar to Holmström (1977, 1984) and allows for uncertainty over the agent's preferences. Like Holmström (1977, 1984) he restricts the set of feasible delegation sets to intervals. Dessein (2002) allows the principal to commit to a particular type of delega-

[5] In our terminology 'delegation' is equivalent to 'communication with commitment' which is the focus of Melumad and Shibano (1991) and Martimort and Semenov (2005).

[6] The only exception is on page 44 of Holmström (1977) where he considers a particular example and argues that for this example an interval delegation set is the optimal among all compact sets.

tion, namely complete delegation, and shows that for a large number of distributions the principal does better when she commits to complete delegation than when she engages in cheap talk communication with the agent. Alonso and Matouschek (2004) develop an infinitely repeated version of the basic delegation game to endogenize the commitment power of the principal. They restrict attention to quadratic preferences with constant biases and, for any discount rate, provide sufficient conditions for particular delegation sets to be optimal. In a recent paper, Martimort and Semenov (2005) restrict attention to similar preferences but allow for multiple agents. In this setting they provide a sufficient condition for a particular delegation set to be optimal.

Our paper is also related to Ottaviani (2000) and Krishna and Morgan (2004). They consider similar set ups as the before-mentioned papers but, in contrast to these papers, they allow for contingent monetary transfers. Finally, our paper is related to the large literature on cheap talk that followed Crawford and Sobel (1982). The key difference between this literature and our model is our assumption that the principal can commit to any delegation set or, equivalently, any decision rule. In contrast, cheap talk models are characterized by the principal's lack of commitment power.

In the next section we describe the model. In Section 3 we then state the delegation problem and show that a solution exists. The core of the analysis is contained in Sections 4 - 6: in Section 4 we characterize the optimal delegation set and in Sections 5 and 6 we provide necessary and sufficient conditions for particular delegation sets to be optimal. In Section 7 we present an example to illustrate the results from the previous sections and to perform comparative statics. We extend the model by allowing for wage payments in Section 8 and we conclude in Section 9.

2 The Model

There is a principal and an agent who have to make a decision. The decision is represented by $y \in Y$ and the set of admissible decisions Y is a large compact interval of $\mathbb{R}$. The utilities of the principal and the agent depend on the decision y and the state of the world θ, where θ takes values over the compact interval $\Theta = [\theta_1, \theta_2] \subset \mathbb{R}$. For most of the paper we assume, without loss of generality, that $\Theta = [0, 1]$.

The principal has a von Neumann-Morgenstern utility function that takes the generalized quadratic form $v_P(y,\theta) = -\alpha(\theta)(y - y_P(\theta))^2$, where $y_P(\theta)$ is continuous in $\theta \in \Theta$ and $\alpha(\theta)$ is a continuously differentiable and strictly positive function of the state θ. Her agent has a von Neumann-Morgenstern utility function given by $u_A(y,\theta) = v_A(y - y_A(\theta), \theta)$, where $y_A(\theta)$ is continuously differentiable and strictly increasing in $\theta \in \Theta$ and, for each $\theta \in \Theta$, the function $v_A(\cdot, \theta)$ is single peaked and symmetric around zero.[7] Thus, given the state of the world θ, the principal's preferred decision is $y_P(\theta)$ and the agent's preferred decision is $y_A(\theta)$. The divergence of the preferences between the principal and the agent is then given by the agent's bias $b(\theta) \equiv y_A(\theta) - y_P(\theta)$. Note that the specified utility functions allow for an arbitrary continuous divergence of preferences. Also, they allow for variable degrees of risk aversion of the principal with respect to the decision y for each realization of θ, as characterized by the function $\alpha(\theta)$. We denote the ranges of the principal's and the agent's preferred decisions by $Y_P = \{y \in Y : y_P(\theta) = y\}$ and $Y_A = \{y \in Y : y_A(\theta) = y\}$. Note that Y_P and Y_A are compact intervals of the form $Y_A = \left[\underline{d}_A, \overline{d}_A\right]$ and $Y_P = \left[\underline{d}_P, \overline{d}_P\right]$.

We assume that the agent is informed about the state of the world θ and that the principal is not. Her prior beliefs over its realization are represented by the cumulative distribution function $\widetilde{F}(\theta)$. The corresponding probability density function $\widetilde{f}(\theta)$ is absolutely continuous and strictly positive for all $\theta \in \Theta$.

We denote a mechanism by (M, h), where M is a message space and $h : M \to X$ is a decision rule that maps the messages into a set of allocations X. We restrict attention to *deterministic* mechanisms where, after receiving a message $m \in M$, the principal makes a particular decision $h(m)$ with certainty and does not randomize. The key assumption that we make and that distinguishes our analysis from the standard mechanism design literature is that message-contingent transfers between the agent and the principal are not feasible. Also, we assume that the participation of the agent in any mechanism (M, h) is assured so that the principal does not have to pay the agent any wages to ensure his participation.[8] The set of feasible mechanisms is therefore restricted to those

[7] As a special case, this specification of the agent's utility function includes the generalized quadratic function $v_A(y,\theta) = -\beta(\theta)(y - y_A(\theta))^2$, where $\beta(\theta)$ is continuous and strictly positive for $\theta \in \Theta$.

[8] This assumption will be relaxed in Section 8 where we discuss optimal delegation with wages.

in which the set of allocations X is the set of admissible decisions Y.

The timing is as follows. The principal selects a mechanism (M, h) after which the principal and the agent play the mechanism. The agent observes the state of the world and sends a message $m \in M$ to the principal who then chooses a decision according to the decision rule h. Payoffs are then realized and the game ends.

3 The Delegation Problem

In this section we first show that the principal's problem – choosing a deterministic mechanism to maximize her expected utility subject to the agent's incentive compatibility constraint – is equivalent to the delegation problem in which the principal offers the agent a set of decisions from which he can choose his preferred one. We then show that the delegation problem always has a solution.

We refer to any mapping $X : \Theta \to \Delta Y$ as an *outcome function* and say that it is *implementable* if there exists a deterministic mechanism (M, h) and an equilibrium strategy for the agent $\sigma : \Theta \to \Delta M$ such that $X(\theta) = h(\sigma(\theta))$ for all $\theta \in \Theta$. Furthermore, we say that an outcome function $X(\theta)$ is *truthfully implementable* if there exists a direct deterministic mechanism (Θ, h) in which $\sigma(\theta) = \theta$ is an equilibrium strategy for the agent and $X(\theta) = h(\sigma(\theta))$ for all $\theta \in \Theta$.

The principal's problem is to choose an implementable outcome function $X(\theta)$ that maximizes her expected utility. Thus, the principal solves

$$\max_{X(\cdot)} \mathrm{E}_{\theta} \left[\mathrm{E}_{X(\theta)} \left[v_P(x(\theta)), \theta) \right] \right] \tag{1}$$

subject to

$$u_A(x(\theta), \theta) \geq u_A(x, \theta) \;\; \forall \; x(\theta) \in \mathrm{supp} X(\theta), \; x \in \cup_{\theta \in \Theta} \mathrm{supp} X(\theta), \;\; \theta \in \Theta.$$

Note that the agent's equilibrium strategy σ is independent of the principal's prior $\widetilde{F}(\theta)$. Furthermore, if we define the adjusted cumulative density function $F(\theta) \equiv \int_0^{\theta} \alpha(t) \widetilde{f}(t) \mathrm{d}t / \int_0^1 \alpha(t) \widetilde{f}(t) \mathrm{d}t$ and its probability density function $f(\theta) \equiv F'(\theta)$, $f(\theta)$ satisfies the assumptions in Section 2. These observations allow us to reformulate the principal's problem by replacing the generalized quadratic utility functions $v_P(y,\theta) = -\alpha(\theta)(y - y_P(\theta))^2$ with the simple quadratic utility function $u_P(y,\theta) \equiv -(y - y_P(\theta))^2$ and the cumulative density function $\widetilde{F}(\theta)$ with the adjusted cumulative density function $F(\theta)$. The principal's problem can therefore be restated as

$$\max_{X(\theta)} \mathrm{E}_\theta \left[\mathrm{E}_{X(\theta)} \left[u_P(x(\theta)),\theta) \right] | F(\theta) \right] \tag{2}$$

subject to

$$u_A(x(\theta),\theta) \geq u_A(x,\theta) \;\; \forall \; x(\theta) \in \mathrm{supp}X(\theta), \; x \in \cup_{\theta\in\Theta}\mathrm{supp}X(\theta), \;\; \theta \in \Theta.$$

For the remainder of the paper all expectations are taken using the adjusted cumulative density function $F(\theta)$ unless explicitly stated otherwise.

Next we show that to solve the principal's problem (2) we can restrict attention to *deterministic* incentive compatible outcome functions $X(\theta) : \Theta \to Y$ that can be truthfully implemented. In particular, the next lemma shows that for every deterministic indirect mechanism and an equilibrium strategy of the agent for that mechanism there exists a truthful direct *deterministic* mechanism that gives the principal weakly higher expected utility.[9]

LEMMA 1 (Truthful Direct Deterministic Mechanisms) *Let $S = (M, h)$, with $h : M \to Y$, be a deterministic mechanism, and let $\sigma : \Theta \to \Delta M$ be an equilibrium of S. Then there exists a direct truth-telling deterministic mechanism $S' : \Theta \to Y$, with agent's equilibrium strategy $s'(\theta) = \theta$, such that the principal obtains (weakly) higher expected utility, i.e.* $\mathrm{E}_\theta \left[u_P(s'(\theta),\theta) \right] \geq \mathrm{E}_\theta \left[\mathrm{E}_{\sigma(\theta)} \left[u_P(m,\theta) \right] \right]$.

Proof: See appendix. ■

The original problem (2) can therefore be restated without loss of generality as

$$\max_{X(\theta)} \mathrm{E}_\theta \left[u_P(X(\theta),\theta) \right] \tag{3}$$

[9] Note that this result is not implied by the Revelation Principle. The Revelation Principle states that any outcome function obtained with a mechanism can also be implemented with a truthful direct mechanism. It does not follow that any outcome function obtained with a *deterministic* mechanism can also be implemented with a truthful direct *deterministic* mechanism.

subject to the incentive compatibility constraint

$$u_A(X(\theta),\theta) \geq u_A(X(\theta'),\theta) \ \ \forall\ \theta,\theta' \in \Theta,$$

where $X(\theta) : \Theta \to Y$ is an outcome function that maps states of the world into decisions. It turns out that the outcome functions that satisfy the incentive compatibility constraint take a simple form. This is shown in the next lemma which is the adaptation to our setting of Proposition 1 in Melumad and Shibano (1991). To be able to state the lemma, let $X^-(\widehat{\theta}) \equiv \lim_{\theta \to \widehat{\theta}^-} X(\theta)$ and $X^+(\widehat{\theta}) \equiv \lim_{\theta \to \widehat{\theta}^+} X(\theta)$.

LEMMA 2 (Melumad and Shibano 1991) *An incentive compatible $X(\theta)$ must satisfy the following: i. $X(\theta)$ is weakly increasing, ii. if $X(\theta)$ is strictly increasing and continuous in (θ_1,θ_2), then $X(\theta) = y_A(\theta)$ for $\theta \in (\theta_1,\theta_2)$, iii. if $X(\theta)$ is discontinuous at $\widehat{\theta}$, then the discontinuity must be a jump discontinuity that satisfies:*

a. $u_A(X^-(\theta),\widehat{\theta}) = u_A(X^+(\theta),\widehat{\theta})$,

b. $X(\theta) = X^-(\widehat{\theta})$ *for* $\theta \in [\max\left\{0, y_A^{-1}(X^-(\widehat{\theta}))\right\}, \widehat{\theta})$,

$X(\theta) = X^+(\widehat{\theta})$ *for* $\theta \in (\widehat{\theta}, \min\left\{1, y_A^{-1}(X^+(\widehat{\theta}))\right\}]$ *and*

c. $X(\widehat{\theta}) \in \left\{X^-(\widehat{\theta}), X^+(\widehat{\theta}))\right\}$.

Proof: See proof of Proposition 1 in Melumad and Shibano (1991). ■

A graphical illustration of the lemma is provided in Figure 1. It can be seen that the outcome function is weakly increasing and consists of, first, flat segments in which the decision is not sensitive to the agent's message and, second, strictly increasing segments in which the agent's preferred decision is implemented. Also, if the outcome function is discontinuous, then it must be symmetric around the agent's preferred decision at the point of discontinuity. Finally, part iii-b. implies that there must be flat segments to the left and the right of the discontinuity point.

For the remainder of the analysis we denote by $X(\theta)$ any incentive compatible outcome function that satisfies the conditions in Lemma 2. In spite of $X(\theta)$ taking a relatively simple form, we will often find it more convenient to work with an alternative formulation of the principal's problem. In this alternative problem, which we refer to as the *delegation problem*, the principal offers the agent a set of decisions from which the agent can choose his preferred one. We show next that the delegation problem is equivalent to the direct mechanism problem (3).

To do so, denote the range of an incentive compatible outcome function $X(\theta)$ by D, i.e. $D = \{y : y = X(\theta), \forall \theta \in \Theta\}$. Essentially, D is the set of decisions from which the agent can choose when the principal offers the mechanism $(\Theta, X(\theta))$. Note that different incentive compatible outcome functions might have the same range D. We denote by X_D the set of outcome functions with range D, i.e. $X_D = \{X(\theta) : D \text{ is the range of } X(\theta)\}$. The next lemma shows that the expected utilities of the principal and, respectively, the agent are the same for all $X(\theta) \in X_D$.

LEMMA 3 (Utility Equivalence) *For any* $D \subseteq Y$*, let* $X_D = \{X(\theta) : D \text{ is the range of } X(\theta)\}$*. Then* $\mathrm{E}_\theta\left[u_i(X'(\theta), \theta)\right] = \mathrm{E}_\theta\left[u_i(X''(\theta), \theta)\right] \ \forall \ X'(\theta), X''(\theta) \in X_D$ and $i = A, P$.

Proof: See appendix. ■

It follows from Lemma 3 that instead of choosing an outcome function $X(\theta)$ we can think of the principal as choosing a *delegation set* D from which the agent can select his preferred decision. In other words, the principal's contracting problem (3) is equivalent to the delegation problem

$$\max_{D \in N} \mathrm{E}_\theta\left[u_P(y^*(\theta), \theta)\right] \tag{4}$$

subject to

$$y^*(\theta) \in X_D(\theta) \equiv \arg\max_{y \in D} u_A(y, \theta),$$

where N is the collection of compact sets of the decision space Y.[10]

Having posed the delegation problem, we show next that it has a solution. This follows immediately from Lemma 3 and Theorem 1 in Holmström (1984).

LEMMA 4 (Existence) *The delegation problem (4) has a solution.*

In general the solution to (4) will not be unique. For example, different optimal delegation sets can be created by adjoining to an optimal delegation set decisions that

[10]Note that if for an arbitrary delegation set (not necessarily compact) a solution to the agent's problem exists, then the range of decisions implemented by the agent is compact. Hence, there is no loss of generality in restricting delegation sets to being compact.

are never chosen by the agent. We therefore focus on optimal *minimal* delegation sets, defined as a solution D to (4) such that all decisions $y \in D$ are chosen, i.e. there exists a state θ such that $y \in X_D(\theta)$. For the remainder of this paper, 'optimal delegation sets' refers to optimal minimal delegation sets unless otherwise stated.

Finally, we partially order delegation sets by how much discretion they bestow on the agent. Specifically, we say that a delegation set D_1 gives the agent more discretion than a delegation set $D_2 \neq D_1$ if and only if $D_2 \subset D_1$.

4 Optimal Delegation

We have shown above that the principal's contracting problem can be stated in two equivalent ways, either as the direct mechanism design problem (3) or as the delegation problem (4). The need to determine the optimal number of jump discontinuities prevents us from applying standard optimal control techniques to solve (3). In this section we therefore focus on (4) to derive a characterization of the principal's contracting problem. The key difficulty in solving the delegation problem (4) is the need to optimize over sets which precludes us from using standard optimization methods. To characterize the solution, we instead investigate the effect on the principal's expected utility of changing the agent's discretion by adding decisions to, and removing decisions from, a delegation set. In the next sub-section we derive a basic condition that determines whether the principal benefits from an increase or a reduction in the agent's discretion. In Sub-section 4.2 we then use this condition to provide a general characterization of optimal delegation sets and provide an example to illustrate how our result can be applied.

4.1 Changing the Agent's Discretion

To investigate the effect of changes in the agent's discretion on the principal's expected utility, consider a delegation set D that contains three consecutive decisions $y_1 < y_2 < y_3$ within the range of the agent's preferred decisions. Consider also a delegation set $\widehat{D}$ that is identical to D but does not include the intermediate decision y_2. Note that, given our definition of discretion, D gives the agent more discretion than $\widehat{D}$. An example of the two delegation sets is given in Figures 2a and 2b.

The figures illustrate that different decisions get implemented under the two delegation sets if and only if $\theta \in [r, t]$, where $r \equiv y_A^{-1}((y_1 + y_2)/2)$ is the state of the world in which the agent is indifferent between decisions y_1 and y_2 and $t \equiv y_A^{-1}((y_2+y_3)/2)$ is the state of the world in which he is indifferent between y_2 and y_3. In particular, under D the agent implements only one decision, y_2, for all $\theta \in [r, t]$ while under $\widehat{D}$ he implements two decisions, y_1 for $\theta \in [r, s]$ and y_3 for $\theta \in [s, t]$, where $s \equiv y_A^{-1}((y_1 + y_3)/2)$. Thus, in those states of the world in which the two delegation sets actually induce different decisions, decision making is *less* sensitive to changes in the state of the world if the agent has *more* discretion. This suggests the somewhat counter-intuitive result that the principal prefers to increase the agent's discretion by adding an intermediate decision to a delegation set if she wants to induce decision making that is locally less sensitive to changes in the state of the world. To derive this result formally, let the difference in the principal's expected utility be given by $\Delta U \equiv \mathrm{E}\left(u_P(y,\theta)\,|D\right) - \mathrm{E}\left(u_P(y,\theta)\,\middle|\widehat{D}\right)$. Using the definitions of r, t and s introduced above, we can then write

$$\Delta U = \int_r^s (y_1 - y_P(\theta))^2 \,\mathrm{d}F(\theta) + \int_s^t (y_3 - y_P(\theta))^2 \,\mathrm{d}F(\theta) - \int_r^t (y_2 - y_P(\theta))^2 \,\mathrm{d}F(\theta).$$

This expression in turn can be restated as

$$\Delta U = -2\left[(y_3 - y_1)\,T(s) - (y_2 - y_1)T(r) - (y_3 - y_2)T(t)\right],$$

where $T(\theta) \equiv F(\theta)\left[y_A(\theta) - \mathrm{E}\left[y_P(s)\,|s \leq \theta\right]\right]$ is the *effective backward bias.* This concept is key for all the results in this paper. To give it an economic interpretation, consider an arbitrary state θ and suppose that the principal knows that the state of the world is below θ. Given this information, the principal's preferred decision is $\mathrm{E}\left[y_P(s)\,|s \leq \theta\right]$. Thus, $T(\theta)$ measures the difference between the agent's preferred decision at θ and the principal's preferred decision conditional on the state of the world being smaller than θ, weighted by the probability $F(\theta)$ that the state of the world is indeed smaller than θ.

Next we can further simplify ΔU by expressing y_2 as a convex combination of y_1 and y_3, i.e. $y_2 = (1 - \lambda)y_1 + \lambda y_3$ for $\lambda \in (0, 1)$. Then,

$$\Delta U = -2(y_3 - y_1)\left[\widetilde{T}\left(\frac{y_1 + y_3}{2}\right) - \lambda\widetilde{T}\left(\frac{y_1 + y_2}{2}\right) - (1 - \lambda)\widetilde{T}\left(\frac{y_2 + y_3}{2}\right)\right], \quad (5)$$

where $\widetilde{T}(y) \equiv T(y_A^{-1}(y))$ is the effective backward bias as a function of decisions rather than states of the world. Since $(y_1 + y_3) = \lambda (y_1 + y_2) + (1 - \lambda)(y_2 + y_3)$, it follows from the above expression that $\Delta U > 0$ if the effective backward bias $\widetilde{T}(y)$ is strictly convex over the interval $[y_1, y_3]$, that $\Delta U < 0$ if $\widetilde{T}(y)$ is strictly concave over $[y_1, y_3]$ and that $\Delta U = 0$ if $\widetilde{T}(y)$ is linear over $[y_1, y_3]$. This allows us to establish the following lemma.

LEMMA 5 (Adding and Removing Discrete Decisions) *Let D be a delegation set which contains three consecutive decisions $y_1 < y_2 < y_3$ that are within the range of the agent's preferred decisions, i.e. $y_1, y_2, y_3 \in Y_A$. Let $\widehat{D} = D \setminus y_2$ be a delegation set derived from D by excluding the decision y_2. Then*

i. *removing decision y_2 from D strictly increases the principal's expected utility if the agent's effective backward bias $\widetilde{T}(y)$ is strictly concave over the interval $[y_1, y_3]$ and it does not change the principal's expected utility if $\widetilde{T}(y)$ is linear over the interval $[y_1, y_3]$.*

ii. *adding any decision $y_2 \in (y_1, y_3)$ to $\widehat{D}$ strictly increases the principal's expected utility if $\widetilde{T}(y)$ is strictly convex over the interval $[y_1, y_3]$ and it does not change the principal's expected utility if $\widetilde{T}(y)$ is linear over the interval $[y_1, y_3]$.*

Proof: Follows immediately from the discussion in the text. ■

To get an intuition for this lemma, consider a particular intermediate decision y_2 that is the average of y_1 and y_3, i.e. $y_2 = (y_1 + y_3)/2$. Furthermore, let

$$R_i(r, s, t) \equiv \int_s^t y_i(\theta)\mathrm{d}F(\theta) - \int_r^s y_i(\theta)\mathrm{d}F(\theta) \quad \text{for } i = A, P \tag{6}$$

be the the agent's and the principal's respective *average responsiveness to changes in the state of the world* over the partition $\{r, s, t\}$. Thus, for instance, the principal is said to be more responsive to changes in the state of the world the higher her average preferred decision over the interval $[s, t]$ relative to her average preferred decision over the interval $[r, s]$.

The above analysis implies that the principal is better off adding decision y_2 to delegation set $\widehat{D}$ if

$$\widetilde{T}\left(\frac{y_1 + y_3}{2}\right) - \lambda \widetilde{T}\left(\frac{y_1 + y_2}{2}\right) - (1 - \lambda)\widetilde{T}\left(\frac{y_2 + y_3}{2}\right) \tag{7}$$

is negative and she is better off removing y_2 from D if this expression is positive. Using (6) we can rewrite (7) as

$$\frac{1}{2}\left[-Z + R_P(r,s,t) - R_A(r,s,t)\right], \tag{8}$$

where $Z \equiv \int_r^s y_A(\theta) - (y_1 + y_2)/2 \mathrm{d}F(\theta) - \int_s^t y_A(\theta) - (y_2 + y_3)/2 \mathrm{d}F(\theta) > 0$. The inequality is implied by $y_A(r) = (y_1 + y_2)/2$ and $y_A(s) = (y_2 + y_3)/2$. This expression implies that the principal benefits from increasing the agent's discretion by adding an intermediate decision if she is relatively unresponsive to changes in the states of the world, in the sense that $R_P(r,s,t) < R_A(r,s,t)$. Essentially, in this case the principal's preferred decision function $y_P(\theta)$ is relatively flat and is therefore better approximated by the flat outcome function induced by D then by the increasing outcome function induced by $\widehat{D}$. This confirms our earlier intuition that the principal prefers to increase the agent's discretion if she wants to induce locally less state-sensitive decision making.

It also follows from (8) that the principal prefers to reduce the agent's discretion by removing an intermediate decision only if she is relatively responsive to changes in the state of the world, i.e. only if $R_P(r,s,t) > R_A(r,s,t)$. Essentially, in this case the principal's preferred decision function is relatively steep and is therefore better approximated by the increasing outcome function induced by $\widehat{D}$ than by the flat outcome function induced by D.

Finally, (8) implies that the principal gains from giving the agent more discretion by adding an intermediate decision if their preferences are sufficiently aligned. Specifically, if the agent's bias is zero for all $\theta \in [r,t]$, then $R_P(r,s,t) = R_A(r,s,t)$ and thus $\Delta U > 0$.

So far we have considered the effect of adding a single discrete decision to, and removing a single discrete decision from, a delegation set. We have seen that if $\widetilde{T}(y)$ is convex over the interval $[y_1, y_3]$, then the principal benefits from adding a decision $y_2 \in (y_1, y_3)$ to the delegation set $\widehat{D}$. Clearly, when $\widetilde{T}(y)$ is convex over the entire interval $[y_1, y_3]$, it is also convex over any interval that lies within $[y_1, y_3]$. This implies that the principal would benefit not just from including one additional decision y_2 but from including any number of decisions. In fact, it suggests that the principal would benefit the most if she added a continuum of decisions (y_1, y_3) to the delegation set.

The following lemma considers the effect of adding decision intervals to, and removing decision intervals from, a delegation set and shows that this is indeed the case.

LEMMA 6 (Adding and Removing Decision Intervals) *Let D be a delegation set that contains an interval $[y_1, y_3] \subset Y_A$ and let $\overline{D} = D \setminus (y_1, y_3)$. Then*

i. removing decisions (y_1, y_3) from D strictly increases the principal's expected utility if $\widetilde{T}(y)$ is strictly concave over the interval $[y_1, y_3]$ and it does not change the principal's expected utility if $\widetilde{T}(y)$ is linear over the interval $[y_1, y_3]$.

ii. adding decisions (y_1, y_3) to $\overline{D}$ strictly increases the principal's expected utility if $\widetilde{T}(y)$ is strictly convex over the interval $[y_1, y_3]$ and it does not change the principal's expected utility if $\widetilde{T}(y)$ is linear over the interval $[y_1, y_3]$.

Proof: See appendix. ■

Lemmas 5 and 6 show that the effect on the principal's expected utility of adding decisions to, and removing decisions from, a delegation set depend crucially on the curvature of the effective bias. In the next subsection we use this insight to characterize optimal delegation sets.

4.2 Characterizing Optimal Delegation Sets

To characterize the optimal delegation set, we partition the decision space Y into different intervals and provide a characterization for each interval.

Suppose first that the ranges of the agent's and the principal's preferred decisions do not intersect, i.e. $Y_A \cap Y_P = \varnothing$. In this case we partition the decision space Y into two overlapping intervals $Y \cap (\infty, \max\{\underline{d}_A, \underline{d}_P\}]$ and $Y \cap [\min\{\overline{d}_A, \overline{d}_P\}, \infty)$. The next proposition then shows that the optimal delegation set contains at most one decision in each interval. Suppose next that the ranges of the agent's and the principal's preferred decisions do intersect, i.e. $Y_A \cap Y_P \neq \varnothing$. The first step in the partitioning then is to divide the decision space into i. an interval above $Y_A \cap Y_P$, ii. an interval equal to $Y_A \cap Y_P$ and iii. an interval below $Y_A \cap Y_P$. The next step is to further divide $Y_A \cap Y_P$ into i. intervals in which the effective bias $\widetilde{T}(y)$ is strictly concave, ii. intervals in which it is linear, and iii. intervals in which it is strictly convex. The next proposition shows that the optimal delegation set contains at most one decision above and one decision below $Y_A \cap Y_P$ and that the number of decisions offered within $Y_A \cap Y_P$ depends on the curvature of the effective bias $\widetilde{T}(y)$ over the different intervals.

PROPOSITION 1 (Characterization) *Let D^* be an optimal delegation set and let $y_1, y_2 \in Y_A \cap Y_P$. Then,*

i. $D^* \cap [\min\{\overline{d}_A, \overline{d}_P\}, \infty)$ *contains at most one decision and* $D^* \cap (\infty, \max\{\underline{d}_A, \underline{d}_P\}]$ *contains at most one decision.*

ii. *if $\widetilde{T}(y)$ is strictly convex for all $y \in [y_1, y_2]$, then $D^* \cap [y_1, y_2]$ is a connected set, i.e. it contains either no decision, one decision, or an interval of decisions.*

iii. *if $\widetilde{T}(y)$ is strictly concave for all $y \in [y_1, y_2]$, then $D^* \cap [y_1, y_2]$ contains at most two decisions.*

iv. *if $\widetilde{T}(y)$ is linear for all $y \in [y_1, y_2]$, then there exists a delegation set $D^{*'} \subseteq D^*$ such that a. the principal is indifferent between $D^{*'}$ and D^* and b. $D^{*'} \cap [y_1, y_2]$ is a connected set.*

Proof: See appendix. ■

Part i. characterizes the optimal delegation set outside of the intersection of the agent's and the principal's ranges and shows that it contains at most one decision above- and at most one decision below $Y_A \cap Y_P$. To get an intuition for this result, suppose the principal offers two decisions above Y_A (the argument when the decisions are below Y_A is similar). The agent then always prefers the smaller decision to the larger one. As a result the larger decision never gets chosen and can therefore not be part of an optimal minimal delegation set. Next, suppose that the principal offers two decisions above Y_P (the argument when the decisions are below Y_P is similar). The principal can then increase her expected utility by removing the larger decision from the delegation set. Whenever the agent would have chosen the larger decision he will then choose the smaller one. The agent's decisions in all other states are not affected by the removal of the largest decision and thus the principal's expected payoff must be higher.

Parts ii. - iv. characterize the optimal delegation set within the intersection of the agent's and the principal's ranges and show that, in this region, it depends critically on the curvature of the agent's effective bias $\widetilde{T}(y)$. In particular, if $\widetilde{T}(y)$ is strictly convex, then the optimal delegation set contains either no decision, one decision or an interval within $Y_A \cap Y_P$. If, instead, $\widetilde{T}(y)$ is strictly concave, then the optimal delegation set contains at most two decisions within $Y_A \cap Y_P$. It is evident that these two results are strongly driven by Lemmas 5 and 6. To understand the characteristics of the optimal delegation set when $\widetilde{T}(y)$ is linear, recall from Lemmas 5 and 6 that in this case the principal is indifferent between delegation sets that offer two decisions and delegation sets that offer any compact set within these two decisions. Thus, when $\widetilde{T}(y)$ is linear over an interval within $Y_A \cap Y_P$, there always exists an optimal delegation set that contains either no decision, one decision, or an interval of decisions.

The characterization result enables us to generically reduce the delegation problem to a finite dimensional problem that can be solved with standard techniques.[11] As an illustration of how this result can be applied, consider the model analyzed in Melumad and Shibano (1991). Their model is a special case of ours in which $u_A(y, \theta) = -(y-\theta)^2$ and $u_P(y, \theta) = -(y - (n_P + m_P\theta))^2$, where $n_P, m_P \in \mathbb{R}$ and $F(\theta) = \theta$. This model is illustrated in Figure 3. We can use Proposition 1 to reproduce their characterization result in a straightforward manner. In particular, it follows from Proposition 1 i. that the optimal delegation set can contain at most one decision above $\min\{1, n_P + m_P\}$ and one decision below $\max\{0, n_P\}$. Moreover, it follows from Proposition 1 ii. - iv. that the characteristics of the optimal delegation set within $[\max\{0, n_P\}, \min\{1, n_P + m_P\}]$ depend on the curvature of the effective bias which, in this example, is given by $\widetilde{T}''(y) = (2 - m_P)$. Thus, if $m_P < 2$, then the optimal delegation set contains either no decision, one decision or an interval within $[\max\{0, n_P\}, \min\{1, n_P + m_P\}]$. If, instead, $m_P > 2$, then the optimal delegation set contains at most two decisions within $[\max\{0, n_P\}, \min\{1, n_P + m_P\}]$.

These results reduce the search for the optimal delegation set to a combinatorial problem. Suppose, for instance, that $m_P < 2$. Then one has to compare the principal's expected utility under the different possible combinations, i.e. only one decision above

[11] Specifically, we can reduce the delegation problem to a finite dimensional problem as long as $\widetilde{T}'(y)$ has a finite number of extrema.

$\min\{1, n_P + m_P\}$ and none below, only one decision below $\max\{0, n_P\}$ and none above etc. Doing so shows that the optimal delegation set consists of a single interval. Using the same technique one can also show that the optimal delegation set consists of only two decisions if $m_P > 2$.

The results of this example are consistent with our basic intuition that the principal increases the agent's discretion by adding intermediate decisions if she wants to induce less state-sensitive decision making. In particular, when $m_P < 2$ the principal is relatively unresponsive to changes in the state of the world. She then benefits from increasing the agent's discretion by adding intermediate decisions since doing so makes his decision making locally less state-sensitive. In contrast, when $m_P > 2$ the principal is relatively responsive to changes in the state of the world. In this case she benefits from reducing the agent's discretion by removing intermediate decisions since this makes his decision making locally more state-sensitive.

5 Centralization

The characterization result in the previous section suggests that optimal delegation sets can take many different forms. In practice, however, some delegation sets appear to be more common than others. Capital budgeting rules and sentencing guidelines, for instance, often take the form of a single interval.[12] Also, in the authors' experience principals often simply impose their preferred decision and forgo their agents' information. In other words, they offer delegation sets that consist of a single decision, namely their preferred decision given their priors. In this section and the next we consider such delegation sets and provide necessary and sufficient conditions for them to be optimal. Specifically, in this section we focus on *centralization*, i.e. a delegation set that contains only the principal's preferred decision given her prior, and in the next section we focus on *interval delegation*, i.e. delegation sets that consist of a single interval.

We define the *value of delegation* V as the difference between the expected utility of the principal when she offers the optimal delegation set D^* and the expected utility

[12] In the context of capital budgeting rules see, in particular, Bower (1970).

derived from the best uninformed decision y_P^*, $V \equiv \mathrm{E}_\theta\left[u_P(X_{D^*}(\theta),\theta)\right] - \mathrm{E}_\theta\left[u_P(y_P^*,\theta)\right]$. If $V = 0$, then centralization is optimal since the principal cannot do better than to impose her preferred decision. If, instead, $V > 0$, then there is value to delegation, i.e. the optimal delegation set contains at least two decisions. Also, we define the *effective forward bias* as $S(\theta) \equiv (1 - F(\theta))\left[y_A(\theta) - \mathrm{E}\left[y_P(s)\,|s \geq \theta\right]\right]$. This concept is closely related to the effective backward bias $T(\theta)$ introduced above and has a similar interpretation: consider a state θ and suppose that the principal knows that the state of the world is above θ. Given this information, the principal's preferred decision is $\mathrm{E}\left[y_P(s)\,|s \geq \theta\right]$. Thus, $S(\theta)$ measures the difference between the agent's preferred decision at θ and the principal's preferred decision conditional on the state of the world being higher than θ, weighted by the probability $(1 - F(\theta))$ that the state is indeed higher than θ.

We can now establish the following lemma that provides a useful characterization of the value of delegation V as a function of the effective forward and backward biases.

LEMMA 7 (Value of Delegation) *Let $X = \{X(\theta) \in X_D, D \in N\}$ be the set of incentive compatible deterministic outcome functions. Then,*

$$V = \max_{X(\theta)\in X} -\left(y_P^* - X(1)\right)^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) \tag{9}$$

$$= \max_{X(\theta)\in X} -\left(y_P^* - X(0)\right)^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta). \tag{10}$$

Proof: See appendix. ■

This lemma allows us to establish the following proposition which gives necessary and sufficient conditions for centralization to be optimal.

PROPOSITION 2 (Centralization) *Centralization is optimal, i.e. $V = 0$, if and only if there does not exist a $\theta \in (0,1)$ such that $T(\theta) > 0$ and $S(\theta) < 0$.*

Proof: See appendix. ■

The sufficiency part of the proposition follows immediately from Lemma 7: if $T(\theta) \leq 0$ for all $\theta \in (0,1)$, then it follows from (9) that $V = 0$. Similarly, it follows from (10) that $V = 0$ if $S(\theta) \geq 0$ for all $\theta \in (0,1)$. To get an intuition for the necessity part of the proposition, suppose that there exists a $\widehat{\theta} \in (0,1)$ such that $T(\widehat{\theta}) > 0$ and

$S(\widehat{\theta}) < 0$. We can then construct a two-decision delegation set that the principal prefers to centralization. To see this, note first that the principal's preferred uninformed decision y_P^* lies strictly between $\overline{y}_1 \equiv \mathrm{E}(y_P(\theta) \,\big|\, \theta \leq \widehat{\theta})$ and $\overline{y}_2 \equiv \mathrm{E}(y_P(\theta) \,\big|\, \theta > \widehat{\theta})$. Moreover, if $T(\widehat{\theta}) > 0$ and $S(\widehat{\theta}) < 0$, then the agent's preferred decision in state $\widehat{\theta}$, $y_A(\widehat{\theta})$, also lies strictly between $\overline{y}_1$ and $\overline{y}_2$, i.e. $\overline{y}_1 < y_A(\widehat{\theta}) < \overline{y}_2$. Suppose now that $y_A(\widehat{\theta}) > y_P^*$ and let $h \equiv y_A(\widehat{\theta}) - y_P^*$ (the argument for $y_A(\widehat{\theta}) \leq y_P^*$ is similar). We can then construct a delegation set D that implements $y_1 \equiv y_P^*$ for $\theta \in [0, \widehat{\theta})$ and $y_2 \equiv y_P(\widehat{\theta}) + h$ for $\theta \in [\widehat{\theta}, 1]$. An example of such a delegation set is illustrated in Figure 4. Note that the decisions implemented under D differ from the one implemented under centralization only if $\theta \in [\widehat{\theta}, 1]$. The principal therefore prefers D to centralization if and only if

$$\int_{\widehat{\theta}}^{1} (y_P^* - y_P(\theta))^2 - (y_2 - y_P(\theta))^2 \mathrm{d}F(\theta) > 0.$$

The LHS of this inequality is equal to $(1 - F(\widehat{\theta}))(y_2 - y_P^*)\left((\overline{y}_2 - y_P^*) - (y_2 - \overline{y}_2)\right)$. Thus, the principal prefers D to centralization if the distance between decision y_2 and her preferred decision for $\theta \in [\widehat{\theta}, 1]$, $\overline{y}_2$, is less than the distance between y_P^* and $\overline{y}_2$. That this must always be the case can be seen in Figure 4. It can also be shown formally by noting that $|y_2 - \overline{y}_2| = \left|y_A(\widehat{\theta}) + h - \overline{y}_2\right| < \left|y_A(\widehat{\theta}) - h - \overline{y}_2\right| = |\overline{y}_2 - y_P^*|$, where the inequality follows from $y_A(\widehat{\theta}) < \overline{y}_2$. Thus, if $T(\widehat{\theta}) > 0$ and $S(\widehat{\theta}) < 0$, then there exists a two-decision delegation set that dominates centralization and, thus, centralization cannot be optimal.

The proposition provides simple conditions under which centralization is optimal. These conditions are satisfied in a number of examples, two of which are described in the following corollary.

COROLLARY *Centralization is optimal if*

i. *(Incongruence) the principal's preferred decision $y_P(\theta)$ is weakly decreasing for all $\theta \in (0, 1)$.*

ii. *(Disjoint Ranges) the ranges of the principal's and the agent's preferred decisions are disjoint, i.e. $Y_A \cap Y_P = \varnothing$.*

The proposition and the corollary show that, in the absence of contingent monetary transfers, it can be optimal for the principal to forgo the information that her agent possesses and to impose her preferred decision given her prior. Essentially, agency costs can be sufficiently high so that the principal is better off making an ignorant but unbiased decision than trying to elicit information from the agent by giving him some discretion. This suggests that the rigid and much bemoaned bureaucratic rules that many firms impose on their employees may simply be the firms' optimal responses to the agency problems they face.

Note also that when centralization is optimal, there is no value to the agent's information. Since in our model the agent does not engage in any productive activities, this implies that there is no value to employing the agent in the first place. Thus, when the conditions of the proposition are satisfied, observed agency relationships must be due to either the principal's ability to use contingent transfers or to productive activities of the agent that are outside of our model.

Finally, while the proposition shows that in some cases there is no value to the agent, it also shows that in other cases the agent's information is valuable to the principal. One such case is described in the following corollary.

COROLLARY *Centralization is not optimal if $y_P(\theta)$ is weakly increasing for all $\theta \in (0,1)$ and $y_P^* \in Y_A^\circ$.*

As an example, the conditions in the corollary are satisfied in the standard constant bias case (in which $y_A(\theta) = \theta + b$ and $y_P(\theta) = \theta$) for any cumulative density function as long as the bias b satisfies $b \leq \mathrm{E}(\theta) \leq 1 + b$. Thus, in this case, there is always value to offering the agent at least two decisions.

6 Interval Delegation

Next we consider interval delegation, i.e. delegation sets that consist of a single interval $[\underline{y}, \overline{y}]$, where $\underline{y} < \overline{y}$. We first provide necessary and sufficient conditions for interval delegation to be optimal and then show that these conditions are satisfied when the agent's preferences are 'sufficiently similar' to the principal's.

We distinguish between four types of interval delegation sets. Under *threshold delegation* the interval lies strictly within the range of the agent, i.e. $\underline{y} > \underline{d}_A \equiv \min Y_A$ and

$\overline{y} < \overline{d}_A \equiv \max Y_A$. In this sense, both thresholds are binding for the agent. In contrast, under *upper-threshold delegation* the agent only faces a binding upper threshold, i.e. $\underline{y} = \underline{d}_A$ and $\overline{y} < \overline{d}_A$, and under *lower-threshold delegation* he only faces a binding lower threshold, i.e. $\underline{y} > \underline{d}_A$ and $\overline{y} = \overline{d}_A$. Finally, under *complete delegation* the agent does not face any binding thresholds, i.e. $\underline{y} = \underline{d}_A$ and $\overline{y} = \overline{d}_A$.

The next proposition provides conditions for threshold delegation to be optimal.

PROPOSITION 3 (Threshold Delegation) *Threshold delegation is optimal if and only if there exist two decisions* $\underline{y}, \overline{y} \in (\underline{d}_A, \overline{d}_A)$ and $\overline{y} > \underline{y}$ *such that*

i. $\widetilde{T}(\underline{y}) = 0$ *and* $\widetilde{T}(y) \leq 0$ *for* $y < \underline{y}$,

ii. $\widetilde{S}(\overline{y}) = 0$ *and* $\widetilde{S}(y) \geq 0$ *for* $y > \overline{y}$ *and*

iii. $\widetilde{T}(y)$ *is convex for all* $y \in \left[\underline{y}, \overline{y}\right]$.

Proof: See appendix. ■

To get an intuition for this proposition, consider first condition i. By requiring that the effective forward bias is weakly positive for all $y \geq \overline{y}$, this condition ensures that for all $\widehat{\theta} \geq y_A^{-1}(\overline{y})$ the agent's preferred decision $y_A(\widehat{\theta})$ is larger than the principal's preferred decision given that $\theta \geq \widehat{\theta}$, $\mathrm{E}\left[y_P(\theta) \middle| \theta \geq \widehat{\theta}\right]$. In this sense the agent has an incentive to make decisions that are too large from the principal's perspective, leading her to impose an upper threshold. The intuition for condition ii. is similar: it implies that for all $\widehat{\theta} \leq y_A^{-1}(\underline{y})$ the agent's preferred decision $y_A(\widehat{\theta})$ is smaller than the principal's preferred decision given that $\theta \leq \widehat{\theta}$, $\mathrm{E}\left[y_P(\theta) \middle| \theta \leq \widehat{\theta}\right]$. The principal therefore imposes a lower threshold to prevent the agent from making decisions that are too small from her perspective. Finally, by requiring that the effective backward bias is convex between $\underline{y}$ and $\overline{y}$, condition iii. ensures that it is optimal for the principal to include all decisions between the two thresholds in the delegation set.

Next we turn to upper- and lower-threshold delegation. These types of delegation sets are closely related to threshold delegation and, as the next proposition shows, the conditions under which they are optimal are similar to those in Proposition 3.

PROPOSITION 4 (Upper- and Lower-Threshold Delegation) *Upper-threshold delegation is optimal if and only if there exist a decision* $\overline{y} \in (\underline{d}_A, \overline{d}_A)$ *such that* *i.* $\widetilde{S}(\overline{y}) = 0$*,* $\widetilde{S}(y) \geq 0$ *for* $y > \overline{y}$ *and* $\widetilde{T}(y) \geq 0$ *for* $y \leq \overline{y}$ *and*

ii. $\widetilde{T}(y)$ *is convex for all* $y \in \left[\underline{y}, \overline{y}\right]$ *.*

Lower-threshold delegation is optimal if and only if there exist a decision $\underline{y} \in (\underline{d}_A, \overline{d}_A)$ *such that*

i. $\widetilde{T}(\underline{y}) = 0$*,* $\widetilde{T}(y) \leq 0$ *for* $y < \underline{y}$*,* $\widetilde{S}(y) \leq 0$ *for* $y \geq \underline{y}$ *and*

ii. $\widetilde{T}(y)$ *is convex for all* $y \in \left[\underline{y}, \overline{y}\right]$ *.*

Proof: See appendix. ■

Finally, we can turn to complete delegation.

PROPOSITION 5 (Complete Delegation) *Complete delegation is optimal if and only if*

i. $Y_A \subseteq Y_P$*,*

ii. $\widetilde{T}(y)$ *and* $\widetilde{S}(y)$ *are increasing and* $\widetilde{T}(y)$ *is convex for* $y \in Y_A$ *and*

iii. $y_P^* \in \left(\underline{d}_A, \overline{d}_A\right)$ *.*

Proof: See appendix. ■

Under complete delegation the principal finds it optimal not to restrict the discretion granted to the agent. Therefore it is necessarily the case that the range of the agent's preferred decisions Y_A is contained in the range of the principal's preferred decisions Y_P. If this were not the case, the principal could improve upon complete delegation by excluding all decisions outside Y_P, thus setting an interior upper bound on the delegation set. This is captured by condition i. Conditions ii. and iii. together guarantee that there is value to delegation, i.e. that centralization is never optimal. Finally, by requiring that the effective backward bias is convex for $y \in Y_A$, condition ii. ensures that the principal does not profit from reducing the agent's discretion in Y_A.

As argued above, interval delegation appears to be widespread in organizations, suggesting that the conditions under which it is optimal are often satisfied in practice. In this context it is interesting to note that in our model interval delegation is optimal if the agent's preferences are 'sufficiently similar' to the principal's. This result is stated formally in the following proposition.

PROPOSITION 6 (Interval Delegation) *Consider any strictly increasing and twice continuously differentiable functions* $y_P(\theta)$ *and* $y_A(\theta)$ *and let* $y_A(\theta, \lambda) = (1 - \lambda)y_A(\theta) +$

$\lambda y_P(\theta)$, where $\lambda \in [0,1]$. Also, let D^* be the optimal delegation set that a principal with preferred decisions $y_P(\theta)$ offers to an agent with preferred decisions $y_A(\theta,\lambda)$. Then there exists a $\overline{\lambda} \in (0,1)$ such that D^* takes the form of a single interval for all $\lambda \geq \overline{\lambda}$.

Proof: See appendix. ■

Note that this proposition requires $y_P(\theta)$ to be strictly increasing which is more restrictive than what we assume in the rest of the analysis. The reason for this assumption is the need to ensure that the agent's preferred decision $y_A(\theta,\lambda)$ is strictly increasing. In light of this proposition, the apparent widespread use of interval delegation in organizations is consistent with our model as long as organizations are able to screen their agents and only hire those with sufficiently aligned preferences.

7 The Normal-Linear Example

In this section we investigate an example that is similar to that considered in Holmström (1977, 1984).[13] Our main aim is to illustrate the results that we derived above and to investigate how changes in the economic environment affect the agent's discretion.

Suppose that the state θ takes values in $\Theta = [-a, a]$ with $a > 0$ and that the principal's and agent's utility functions are respectively given by $u_P(y,\theta) = -(y-\theta)^2$ and $u_A(y,\theta) = -(y - m_A\theta)^2$, where $m_A > 0$. The principal's prior beliefs over θ are distributed according to a zero-mean truncated normal distribution with a probability density function $f(\theta) = K \exp\left(-\theta^2/2\sigma^2\right)/\sqrt{2\pi\sigma^2}$, where $K \equiv 1/\Pr\left[|Z| \leq a/\sigma\right]$ and $Z \sim N(0,1)$. To avoid having to discuss a large number of different cases, we assume that the support is sufficiently large so that the principal's preferred decision conditional on the state being positive belongs to the range of the agent's preferred decisions. Formally this is ensured by assuming that $m_A a \geq 2\sigma^2\left[f(0) - f(a)\right]$.[14] The normal-linear example is illustrated in Figure 5.[15]

[13] In contrast to our analysis of this example, Holmström (1977, 1984) restricts attention to interval delegation sets.

[14] When this condition is violated, the principal finds it optimal to include decisions that are outside the agent's range of preferred decisions. Full details of the cases when $m_A a < \sigma^2\left[f(0) - f(-a)\right]$ are available from the authors.

[15] This example differs from the main example in Holmström (1977, 1984) in two ways. First, we

In this example it is never optimal for the principal to centralize. This is the case since, for any value of m_A, the principal can always improve upon centralization by offering two decisions $\{-y, y\}$ with $y = \mathrm{E}[\theta \,|\theta \geq 0]$. Essentially, such a scheme allows her to elicit whether the state of the world is positive or negative and to implement a different decision in each case. We can therefore state the first result:

RESULT 1 (Centralization) *Centralization is never optimal.*

Proof: The effective biases are given by $T(\theta) = F(\theta) m_A \theta + \sigma^2 [f(\theta) - f(-a)]$ and $S(\theta) = (1 - F(\theta)) m_A \theta - \sigma^2 [f(\theta) - f(-a)]$. Note that $T(0) = \sigma^2 [f(0) - f(a)] > 0$ and $S(0) = -T(0) < 0$. The result then follows from Proposition 2. ■

To characterize the optimal delegation set, we need to investigate the curvature of the effective backward bias $\widetilde{T}''(y) = [2m_A - 1 - (m_A - 1)(y/m_A)^2/\sigma^2] f(y/m_A)/m_A^2$. Let $\widetilde{y}$ be implicitly defined by $\widetilde{T}''(\widetilde{y}) = 0$ and define $t \equiv \min\{|\widetilde{y}|, \overline{d}_A\}$. It then follows that i. if $m_A > 1$, then $\widetilde{T}''(y) \geq 0$ if and only if $y \in [-t, t]$, ii. if $1/2 \leq m_A \leq 1$, then $\widetilde{T}''(y) \geq 0$ for all $y \in Y_A$ and iii. if $m_A < 1/2$, then $\widetilde{T}''(y) \leq 0$ if and only if $y \in [-t, t]$. Next we analyze these three cases in turn.

The Extremist Agent ($m_A > 1$): The next proposition characterizes the optimal delegation set and provides comparative statics for this case.

RESULT 2 (Extremist Agent) *Suppose the agent is an extremist, i.e.* $m_A > 1$. *Then threshold delegation is optimal. The optimal delegation set is a symmetric interval* $[-\overline{y}, \overline{y}]$, *where*

$$\overline{y} = \sigma^2 \frac{f(\overline{y}/m_A) - f(-a)}{1 - F(\overline{y}/m_A)}.$$

The agent's discretion increases i. if the agent's informational advantage increases, i.e. if σ^2 *increases, or ii. if the agent's preferences become more aligned with the principal's, i.e. if* m_A *decreases.*

Proof: See appendix. ■

---

assume that the principal's prior beliefs are summarized by a truncated normal distribution whereas he assumes they are summarized by a normal distribution with infinite support. Second, he restricts attention to $m_A > 1$ while we do not.

In this case the agent always prefers more extreme decisions than the principal, i.e. he prefers larger decisions for positive realizations of the state of the world and smaller decisions for negative realizations. Moreover, the discrepancy between the principal's and the agent's preferred decision increases as $|\theta|$ increases. It is then intuitive that the principal restricts the agent's discretion by imposing an upper and a lower bound on the delegation set. However, within these bounds the principal allows the agent to make any decision. She does so since the agent is relatively responsive to changes in the states of the world. As argued above, in such a case the principal finds it optimal to reduce the sensitivity of the agent's decision making by including all intermediate decisions in the delegation set. The proposition also shows that the agent's discretion, as measured by $\overline{y}$, is increasing in his informational advantage and in the extent to which his preferences are aligned with the principal's.

The Moderate Agent ($1/2 \leq m_A \leq 1$): In contrast to the previous case, the agent now prefers smaller decisions than the principal when the state of the world is positive and larger decisions otherwise. It is then intuitive that the principal does not impose an upper or a lower threshold on the delegation set. Indeed, as the next proposition shows she does not put any restrictions on the agent.

RESULT 3 (Moderate Agent) *Suppose the agent is a moderate, i.e.* $1/2 \leq m_A \leq 1$. *Then complete delegation is optimal.*

Proof: See appendix. ■

The Passive Agent ($0 < m_A < 1/2$): Just as in the previous case, the agent prefers smaller decisions than the principal when the state of the world is positive and larger decisions otherwise and this discrepancy grows as $|\theta|$ increases. As a result, it can not be optimal for the principal to impose an upper or a lower bound on the delegation set. Note next that the principal now faces an agent who is much less responsive to changes in the states than she is. We know from our analysis above that in such a situation the principal may want to make the agent's decision making more state-sensitive by reducing his discretion and, in particular, by excluding intermediate decisions. The next proposition shows that the principal does indeed find it optimal to rule out intermediate decisions, including the optimal centralized decision.

RESULT 4 (Passive Agent) *Suppose the agent is passive, i.e.* $0 < m_A < 1/2$. *Then the optimal delegation set contains all decisions* Y *except those in an interval* $(-\overline{y}, \overline{y})$, *where*

$$\overline{y} = \sigma^2 \frac{f(0) - f(\overline{y}/m_A)}{F(\overline{y}/m_A) - 1/2} > 0.$$

The agent's discretion increases i. if the agent's informational advantage decreases, i.e. if σ^2 *decreases, or ii. if the agent's preferences become more aligned with the principal's, i.e. if* m_A *increases.*

Proof: See appendix. ■

Note that the optimal delegation set described in this result is not minimal, i.e. it includes decisions that are never selected in equilibrium.[16] We focus on this optimal delegation set to ensure that the comparative statics are unambiguous. In this context, note that, in contrast to the extremist agent case, an increase in the agent's informational advantage now leads to a reduction in his discretion. This illustrates the fact that, in general, changes in the agent's informational advantage have an ambiguous effect on the agent's discretion. To understand why this is the case in this example, recall the extremist agent case in which $m_A > 1$. In this case the principal sets an upper bound $\overline{y}$ on the agent's admissible decisions such that $\overline{y} = \mathrm{E}\left[y_P(\theta)\,|\theta \geq \overline{y}/m_A\right]$. Therefore the maximum admissible decision is the same as the decision that the principal would optimally select conditional on the event $\{\theta \geq \overline{y}/m_A\}$. A reduction in σ^2 unambiguously decreases $\mathrm{E}\left[y_P(s)\,|s \geq \overline{y}/m_A\right]$, leading to a reduction in the agent's discretion. In the passive agent case, however, the principal bans all decisions in $(-\overline{y}, \overline{y})$, where $\overline{y} = \mathrm{E}\left[y_P(\theta)\,|0 \leq \theta \leq \overline{y}/m_A\right]$. Therefore the smallest positive decision the agent is allowed to make is the same decision that the principal would optimally select conditional on the event $\{0 \leq \theta \leq \overline{y}/m_A\}$. Again, a reduction in σ^2 unambiguously decreases $\mathrm{E}\left[y_P(\theta)\,|0 \leq \theta \leq \overline{y}/m_A\right]$. In this case, however, this reduction corresponds to an *increase* in the discretion of the agent.

[16] The optimal minimal delegation set would be $Y_A \backslash (-\overline{y}, \overline{y})$.

8 Optimal Delegation with Wages

So far we have ruled out wage payments and have assumed that the agent's participation constraint is not binding. In this section we relax both these assumptions while maintaining our assumption that message-contingent transfers are not feasible.

Specifically, we change the model presented in Section 2 in two ways. First, the von Neuman-Morgenstern utility functions for the principal and the agent are now given by $v_P(y, \theta) = K_P - \alpha(\theta)(y - y_P(\theta))^2 - w$ and $v_A(y, \theta) = K_A - \beta(\theta)(y - y_A(\theta))^2 + w$, where w is the wage payment from the principal to the agent and where $K_P \geq 0$ and $K_A \geq 0$ are the utilities that the principal and the agent realize simply by forming a relationship. To rule out the uninteresting case in which the principal does not hire the agent, we assume that K_A and K_P are always sufficiently large for an employment relationship to be formed. Second, we now assume that the principal makes the agent a take-it-or-leave-it offer that consists of a delegation set D and a wage w before the agent learns the state of the world. If the agent rejects the offer, he realizes his reservation utility which we normalize to zero.[17] If he accepts the offer, the agent is costlessly informed of the state of the world θ and optimally selects a decision from the delegation set D.

The principal's contracting problem then is to $\max_{w,D\in N} \mathrm{E}_\theta \left[v_P(y^*(\theta), \theta) \middle| \widetilde{F}(\theta) \right] - w$ subject to $y^*(\theta) \in X_D(\theta) \equiv \arg\max_{y\in D} v_A(y, \theta)$ and $\mathrm{E}_\theta \left[v_A(y^*(\theta), \theta) \right] + w \geq 0$, where the first constraint is the incentive constraint and the second the participation constraint. Similarly to Section 3 we can simplify the utility functions provided we adjust the cumulative density function appropriately. For this purpose let $u_P(y, \theta) = K_P - (y - y_P(\theta))^2 - w$ and $u_A(y, \theta) = K_A - (y - y_A(\theta))^2 + w$. Since the utility functions are quasi-linear, the participation constraint is binding. Thus we can substitute for the wage in the objective function to obtain the reduced contracting problem

$$\max_{D\in N} \mathrm{E}_\theta \left[u_P(y^*(\theta), \theta) + u_A(y^*(\theta), \theta) \right] \tag{11}$$

[17] This implies that the agent is unaffected by the principal's ulterior decision. This is a natural assumption if the agent, upon rejecting the principal's offer, is not influenced by the posterior evolution of the organization.

subject to

$$y^*(\theta) \in X_D(\theta) \equiv \arg\max_{y \in D} u_A(y, \theta),$$

where the expectations are taken using the adjusted cumulative density function $F(\theta)$. Note that, in contrast to the delegation problem without wages (4), the principal now chooses the delegation set that maximizes joint expected utility rather than her individual expected utility. In spite of this differences, however, the solutions to (4) and (11) are closely related.

To see this, let $\widetilde{D}^*$ denote a delegation set that solves the delegation problem without wages (4) when the principal's parameter of risk aversion is $\widetilde{\alpha}(\theta) \equiv \alpha(\theta) + \beta(\theta)$ and her preferred decision is given by the following convex combination of $y_P(\theta)$ and $y_A(\theta)$: $\widetilde{y}_P(\theta) \equiv (\alpha(\theta) y_P(\theta) + \beta(\theta) y_A(\theta)) \,/\, (\alpha(\theta) + \beta(\theta))$. We can then state the following proposition.

PROPOSITION 7 (Delegation with Wages) *In the model with wages, $\widetilde{D}^*$ is an optimal delegation set.*

Proof: First note that since w and K_A are independent of the agent's decision y, the incentive constraints in (4) and (11) are equivalent.

Consider next the objective functions in (4) and (11). Expanding $u_P(y, \theta) + u_A(y, \theta)$ gives $u_P(y, \theta) + u_A(y, \theta) = -\widetilde{\alpha}(\theta) \left(y - \widetilde{y}_P(\theta)\right)^2 + J$, where J is independent of y. It follows that any solution to (4) when the principal's parameter of risk aversion is given by $\widetilde{\alpha}(\theta)$ and her preferred decision is given by $\widetilde{y}_P(\theta)$ is also a solution to (11). ■

We can thus use Proposition 1 to characterize the optimal delegation set in a model with wages. Also, by adjusting the principal's preferences in the way described above, we can use Propositions 2 - 5 to provide necessary and sufficient conditions for centralization and interval delegation. Finally, we can adjust Proposition 6 to show that once we allow for wages, interval delegation remains to be optimal when the agent's preferences are sufficiently similar to the principal's.

9 Conclusions

In this paper we have investigated the optimal delegation of decision rights by an uninformed principal to an informed but biased agent. We have shown that the characteristics of an optimal delegation set depend crucially on the agent's effective bias. First, the value of this bias determines whether there is any value of delegation, i.e. whether it is in the principal's interest to offer the agent a delegation set that contains more than one decision. Second, the curvature of the effective bias determines the qualitative nature of the optimal delegation set. In particular, it determines whether or not it is in the principal's interest to change the agent's discretion by adding intermediate decisions to, or removing such decisions from, a delegation set. In this context a key, and somewhat counter-intuitive, insight is that the principal finds it optimal to *increase* the agent's discretion by adding an intermediate decision to a delegation set, if she wants to induce decision making that is locally *less sensitive* to changes in the world. Finally, together the value of the effective bias and its curvature determine the form of the optimal delegation set. We have shown that depending on the value and the curvature of the effective bias it can be optimal for the principal to engage in interval delegation, i.e. to offer a delegation set that consists of a single interval. Moreover, we have shown that interval delegation is optimal if the agent's preferences are sufficiently similar to the principal's.

Our paper also makes a methodological contribution. In the delegation problem that we investigate the principal optimizes over delegation sets which precludes us from using standard optimization techniques. Instead we characterize the solution by investigating the effect on the principal's expected utility of adding decisions to, and removing decisions from, a delegation set. To our knowledge this is the first paper in the economics literature that uses this method to solve such an optimization problem.

In our analysis we have focused on the role that delegation plays in eliciting information and have abstracted from its role in motivating agents to acquire information in the first place (Aghion and Tirole 1997; Szalay 2004). We believe that the optimal design of delegation sets when both roles are important is an interesting research topic that has not been fully investigated in the literature. We leave this topic for future research.

10 Appendix

Proof of Lemma 1: The proof is similar to the arguments presented in Strausz (2003). First we establish that for $\theta' > \theta$, $\inf\{y : y \in \text{supp } \sigma(\theta')\} \geq \sup\{y : y \in \text{supp } \sigma(\theta)\}$. In other words, for any deterministic mechanism offered by the principal and any equilibrium of that mechanism, the agent always induces higher decisions for higher realizations of the state of the world. Suppose on the contrary that for $\theta' > \theta$ there exists $y(\theta') < y(\theta)$, where $y(\theta') \in$supp $\sigma(\theta')$ and $y(\theta) \in$supp $\sigma(\theta)$. From single-peakedness and symmetry of $u_A(y, \theta)$ it has to be the case then that $y_A(\theta) \geq (y(\theta') + y(\theta))/2$ and $y_A(\theta') \leq (y(\theta') + y(\theta))/2$. Since $y_A(\theta)$ is strictly increasing this leads to a contradiction. Therefore we have that $y(\theta') \geq y(\theta)$ for $y(\theta') \in$supp $\sigma(\theta')$ and $y(\theta) \in$supp $\sigma(\theta)$.

Next let $u(\theta) = \text{E}_{\sigma(\theta)}[u_P(m, \theta)]$ be the interim expected utility of the principal for type θ. For each θ define $s(\theta) \in$supp $\sigma(\theta)$ such that $u_P(s(\theta), \theta) \geq \text{E}_{\sigma(\theta))}[u_P(m, \theta)]$. Note that $s(\theta)$ is non-decreasing and hence Borel-measurable. Therefore the direct deterministic mechanism $S' : \Theta \to Y$ such that $s'(\theta) = s(\theta)$ is well defined, incentive compatible, and satisfies $\text{E}_\theta[u_P(s'(\theta), \theta)] \geq \text{E}_\theta\left[\text{E}_{\sigma(\theta)}[u_P(m, \theta)]\right]$. ■

Proof of Lemma 3: Let $D \subset Y$ be a compact set and define $S = \{\theta \in \Theta : X^{'}(\theta) \neq X''(\theta); X^{'}, X'' \in X_D\}$. The set S contains all states where outcome functions in X_D differ. This in turn implies that for the agent at $\theta \in S$, $y_A(\theta) \notin D$. We will prove that Prob$[\theta \in S] = 0$ which then establishes that $\text{E}_\theta[u_i(X'(\theta), \theta)] = \text{E}_\theta[u_i(X''(\theta), \theta)]$ for all $X', X'' \in X_D$ and $i = A, P$. To prove that Prob$[\theta \in S] = 0$ we show that the set S is countable. Since F is absolutely continuous it follows that Prob$[\theta \in S] = 0$.

Let $\theta \in S$. By single peakedness and symmetry of $u_A(y, \theta)$ w.r.t. to y, then $X(\theta) \subset \{s_\theta, s'_\theta\}$ and $s_\theta < y_A(\theta) < s'_\theta$. Since $y_A(\theta)$ is strictly increasing, for $\widetilde{\theta} \in S$, $\theta \neq \widetilde{\theta}$ it must be that $\{x : x = X(\theta), X \in X_D\} \neq \left\{x : x = X(\widetilde{\theta}), X \in X_D\right\}$. Next associate with each $\theta \in S$ the number $s'_\theta - s_\theta > 0$. Define the sets $A_n = \{\theta \in S : \frac{1}{n} > s'_\theta - s_\theta \geq \frac{1}{1+n}\}$, $n \in \mathbb{N}$ and $A_0 = \{\theta \in S : s'_\theta - s_\theta \geq 1\}$. Note that each A_n is a finite set. Since $S = \cup_{i=0}^{\infty} A_i$, S is countable. ■

Proof of Lemma 6: Let $y_2 = (y_1 + y_3)/2$ and consider the collection of delegation sets $D(t) = D \setminus (y_2 - t, y_2 + t)$ for $t \in [0, \bar{t}]$, where $\bar{t} \equiv (y_3 - y_1)/2$. Note that $D(0) = D$ and

$D(\overline{t}) = \overline{D}$. The difference in the principal's expected utility from these two delegation sets $\Delta(t) \equiv \mathrm{E}(u_P(y,\theta)\,|D) - \mathrm{E}(u_P(y,\theta)\,|D(t))$ is given by

$$\Delta(t) = \int_{y_A^{-1}(y_2-t)}^{y_A^{-1}(y_2)} (y_2 - t - y_P(\theta))^2 - (y_A(\theta) - y_P(\theta))^2 \,\mathrm{d}F(\theta) + \int_{y_A^{-1}(y_2)}^{y_A^{-1}(y_2+t)} (y_2 + t - y_P(\theta))^2 - (y_A(\theta) - y_P(\theta))^2 \,\mathrm{d}F(\theta). \quad (12)$$

Differentiating this expression gives $\Delta'(t) = 2\left[\widetilde{T}(y_2+t) + \widetilde{T}(y_2-t) - 2\widetilde{T}(y_2)\right]$. Thus, if $\widetilde{T}(y)$ is strictly concave in $[y_2 - \overline{t}, y_2 + \overline{t}]$ we have that $\Delta'(t) \leq 0$ for $t \leq \overline{t}$ and with strict inequality if $\widetilde{T}(y)$ is strictly concave. Thus, in this case $\Delta(t) = \int_0^{\overline{t}} \Delta'(t)\mathrm{d}t' \leq 0$ for $t \in (0, \overline{t}]$ and with strict inequality if $\widetilde{T}(y)$ is strictly concave in $[y_2 - \overline{t}, y_2 + \overline{t}]$. It follows that, if $\widetilde{T}(y)$ is concave in $[y_2 - \overline{t}, y_2 + \overline{t}]$, then (12) is negative for all $t \in [0, \overline{t}]$ and takes its minimum value for $t = \overline{t}$. This proves part i.

Conversely, if $\widetilde{T}(y)$ is convex in $[y_2 - \overline{t}, y_2 + \overline{t}]$ we have that $\Delta'(t) \geq 0$ for $t \leq \overline{t}$ and with strict inequality if $\widetilde{T}(y)$ is strictly convex. Since $\Delta(0) = 0$ this implies that $\Delta(t) = \int_0^{t} \Delta'(t)\mathrm{d}t' \geq 0$ for $t \in (0, \overline{t}]$ and with strict inequality if $\widetilde{T}(y)$ is strictly convex in $[y_2 - \overline{t}, y_2 + \overline{t}]$. Hence, if $\widetilde{T}(y)$ is strictly convex in $[y_2 - \overline{t}, y_2 + \overline{t}]$, then (12) is positive for all $t \in [0, \overline{t}]$. This proves part ii. ■

Proof of Proposition 1: *Part i.-* We first show that D^* can contain at most one point above and one point below the range of preferred decisions of the agent Y_A. Suppose that $D^* \cap [\overline{d}_A, \infty)$ and $D^* \cap (-\infty, \underline{d}_A]$ are non empty and let $\overline{c}_A = \min D^* \cap [\overline{d}_A, \infty)$ and $\underline{c}_A = \max D^* \cap (-\infty, \underline{d}_A]$. Single peakedness of $u_A(y,\theta)$ w.r.t. y implies that $\overline{c}_A$ and $\underline{c}_A$ are strictly preferred by the agent to all other points in $D^* \cap [\overline{d}_A, \infty)$ and $D^* \cap (-\infty, \underline{d}_A]$. Minimality of D^* implies that $D^* \cap [\overline{d}_A, \infty) = \{\overline{c}_A\}$ and $D^* \cap (-\infty, \underline{d}_A] = \{\underline{c}_A\}$.

Now we establish that D^* can contain at most one point above and one point below the range of preferred decisions of the principal Y_P. Suppose the set $D^* \cap (-\infty, \underline{d}_P]$ contains more than one point. The corresponding analysis for the set $D^* \cap [\overline{d}_P, \infty)$ is entirely analogous. Let $\underline{c}_P = \max D^* \cap (-\infty, \underline{d}_P]$ be the highest decision in D^* (weakly) below the principal's range of preferred decisions and define the set of states $S_{D^*} = \{\theta \,|\underline{c}_P > \arg\max_{y \in D^*} u_P(y, \theta)\}$ where the agent selects a decision strictly below $\underline{c}_P$. Now consider the alternative delegation set $\widetilde{D} = (D^* \cap [\underline{d}_P, \infty)) \cup \{\underline{c}_P\}$ which is obtained by replacing all decisions in D^* below $\underline{d}_P$ with the single decision $\underline{c}_P$. For states in $S^c_{D^*}$ the agent's optimal choice remains unchanged under $\widetilde{D}$ while for states in S_{D^*} the agent will optimally select $\underline{c}_P$ from the delegation set $\widetilde{D}$. Strict concavity of the principal's utility w.r.t. y implies that $u_P(\underline{c}_P, \theta) > u_P(y, \theta)$ for $\theta \in S_{D^*}, y \in D^* \cap (-\infty, \underline{d}_P)$. If $\text{Prob}[\theta \in S_{D^*}] > 0$ then $\text{E}\left[u_P(y, \theta) \middle| \widetilde{D}\right] > \text{E}\left[u_P(y, \theta) \,| D^*\right]$ contradicting the assumed optimality of D^*.

Part ii.- Suppose on the contrary that $D^* \cap [y_1, y_2]$ is not a connected set. Since $D^* \cap [y_1, y_2]$ is closed there exist two points $u, v \in D^* \cap [y_1, y_2]$, $u \neq v$, such that the interval (u, v) does not contain any points of D^*. Consider the alternative (compact) delegation set $\widehat{D} = D^* \cup [u, v]$. The difference in expected utility to the principal under $\widehat{D}$ and D^* is given by $\Delta\left((v - u)/2\right)$ (where the difference is evaluated at $(v + u)/2$). Since $\widetilde{T}(y)$ is strictly convex in $[u, v]$ by Lemma 6.ii $\Delta\left((v - u)/2\right) > 0$. Thus, D^* cannot be optimal.

Part iii.- We establish this claim in two steps. We first show that if $\widetilde{T}(y)$ is strictly concave in $[y_1, y_2]$ an optimal delegation set cannot contain any non-degenerate interval. Second, we establish that any delegation set with more than two decisions in $[y_1, y_2]$ is strictly dominated (from the principal's perspective) by a delegation set with only two decisions in $[y_1, y_2]$.

Suppose first that $D^* \cap [y_1, y_2]$ contains a closed interval $[u, v]$. Let $\widehat{D} = D^* \cap (u, v)^C$ be an alternative delegation set where all decisions in (u, v) are prohibited by the principal. The difference in the principal's expected utility under D^* and $\widehat{D}$ is given by $\Delta\left((v - u)/2\right)$ (where the difference is evaluated at $(v + u)/2$). Since $\widetilde{T}(y)$ is strictly concave in $[u, v]$ by Lemma 6.i., $\Delta\left((v - u)/2\right) > 0$ contradicting the assumed optimality of D^*. Thus, D^* cannot contain any non-degenerate interval in $[y_1, y_2]$.

Consider now the case in which $D^* \cap [y_1, y_2]$ contains more than two points but does not contain any non-degenerate interval. We will distinguish two cases: a. there exist three decisions $\widehat{y}_1 < \widehat{y}_2 < \widehat{y}_3$ with $\widehat{y}_1, \widehat{y}_2, \widehat{y}_3 \in D^* \cap [y_1, y_2]$ that are consecutive in the sense that $(\widehat{y}_1,\widehat{y}_2)\cap D^* = (\widehat{y}_2,\widehat{y}_3)\cap D^* = \varnothing$, b. $D^*\cap[y_1,y_2]$ does not contain three consecutive decisions[18].

a. *Three consecutive decisions* $\widehat{y}_1 < \widehat{y}_2 < \widehat{y}_3$. Suppose that there are three consecutive decisions $\widehat{y}_1 < \widehat{y}_2 < \widehat{y}_3$. We now propose an alternative delegation set $\widehat{D}$ which coincides with D^* except for the decision $\widehat{y}_2$ which is banned by the principal. Then, letting $\Delta U \equiv \mathrm{E}(u_P(y,\theta)\,|D^*) - \mathrm{E}(u_P(y,\theta)\,\Big|\widehat{D})$ be the difference in the expected utility of the principal from D^* and $\widehat{D}$ by Lemma 5.i we have that $\Delta U < 0$ so that $\mathrm{E}(u_P(y,\theta)\,\Big|\widehat{D}) > \mathrm{E}(u_P(y,\theta)\,|D^*)$. Therefore the delegation set D^* with three consecutive decisions $\widehat{y}_1,\widehat{y}_2,\widehat{y}_3$ in $D^*\cap[y_1,y_2]$ cannot be optimal.

b. *No three consecutive decisions.* Let $\overline{s} = \max D^*\cap[y_1,y_2]$ and $\underline{s} = \min D^*\cap[y_1,y_2]$ be the highest and lowest decisions in the range $[y_1,y_2]$ allowed in D^*. Note that the complement in $[y_1,y_2]$ of D^* is an open set whose intersection with $[\underline{s},\overline{s}]$ can be described as the union of a countable collection of pairwise disjoint intervals A_i ,$i\geq 1$, of the form $A_i = (\underline{a}_i,\overline{a}_i)$. For convenience define B_i, $i\geq 1$, the set of states in which the agent's preferred decision lies in $\left(y_A^{-1}(\underline{a}_i)\,,y_A^{-1}(\overline{a}_i)\right)$. We now construct a sequence of delegation sets D_i such that the expected utility of the principal $\mathrm{E}(u_P(y,\theta)\,|D_i)$ converges to $\mathrm{E}(u_P(y,\theta)\,|D^*)$. Define $D_0 = \{\underline{s},\overline{s}\}$ and $D_i = D_{i-1}\cup\,\{\underline{a}_i,\overline{a}_i\}$ for $i\geq 1$. Next note that the agent's optimal response under D_i and D^* coincide in the set $\cup_{j=1}^{i}B_j$, and that $\lim_{i\to\infty}\Pr\left[\theta\in\left(\cup_{j=1}^{i}B_i\right)^c\right] = 0$. Therefore we have that for each i, $|\mathrm{E}(u_P(y,\theta)\,|D_i) - \mathrm{E}(u_P(y,\theta)\,|D^*)| \leq \left|\max_{y\in Y} y - \min_{y\in Y} y\right| \Pr\left[\theta\in\left(\cup_{j=1}^{i}B_i\right)^c\right]$ which implies that $\mathrm{E}(u_P(y,\theta)\,|D_i) \to \mathrm{E}(u_P(y,\theta)\,|D^*)$ as $i\to\infty$.

By the previous proof for three consecutive decisions we know that $\mathrm{E}(u_P(y,\theta)\,|D_{i-1}) > \mathrm{E}(u_P(y,\theta)\,|D_i)$. Thus $\mathrm{E}(u_P(y,\theta)\,|D_0) > \mathrm{E}(u_P(y,\theta)\,|D^*)$ and D^* cannot be optimal.

Part iv.- Let $\overline{s} = \max D^*\cap[y_1,y_2]$ and $\underline{s} = \min D^*\cap[y_1,y_2]$ be the highest and lowest decisions in the range $[y_1,y_2]$ allowed in D^*. Consider the alternative delegation set $\widehat{D} = [\underline{s},\overline{s}]\cup D^*$ where the principal offers the entire interval $[\underline{s},\overline{s}]$ to the agent. We

[18] We note that the existence in $D^*\cap[y_1,y_2]$ of three decisions that are consecutive is equivalent to the existence of an isolated point of $D^*\cap[y_1,y_2]$ different from its extremal points (i.e. different from $\max D^*\cap[y_1,y_2]$ and $\min D^*\cap[y_1,y_2]$). There are however compact sets that are nowhere dense (therefore do not contain any nondegenerate interval) but have no isolated points, i.e. all its points are accumulation points. An example of such a set would be the *Cantor ternary set* (see e.g. Rudin 1987).

will show that if D^* is optimal then $\widehat{D}$ is also optimal. Suppose on the contrary that $\mathrm{E}(u_P(y,\theta)\,\Big|\,\widehat{D}) < \mathrm{E}(u_P(y,\theta)\,|D^*)$. Then there must exist an interval $[u,v]$, $u \neq v$ such that $\mathrm{E}(u_P(y,\theta)\,\Big|\,\widehat{D}) < \mathrm{E}(u_P(y,\theta)\,\Big|\,\widehat{D}_1)$, where $\widehat{D}_1 = \widehat{D} \cap (u,v)^c$, or, equivalently $\Delta\left((v-u)/2\right) < 0$. However, since $\widetilde{T}(y)$ is linear in $[u,v]$ we have by Lemma 6 that $\Delta\left((v-u)/2\right) = 0$ reaching a contradiction. Therefore if D^* is optimal then $\widehat{D}$, which follows by substituting the set of decisions $D^* \cap [y_1, y_2]$ by its convex hull $[\underline{s}, \overline{s}]$, is also optimal. ■

Proof of Lemma 7: Since for all $X(\theta) \in X$, $X(\theta)$ is (weakly) monotonic and bounded and $T(\theta)$ is continuous, the Riemann-Stieltjes integral $\int_0^1 T(\theta)\mathrm{d}X(\theta)$ is well-defined and finite. Now consider a given $X(\theta)$ and compute the difference $\Delta(X(\theta)) = \mathrm{E}_\theta\left[u_P(X(\theta),\theta)\right] - \mathrm{E}_\theta\left[u_P(y_P^*,\theta)\right]$ between the expected utility of the principal under $X(\theta)$ and under centralization

$$\begin{aligned}\Delta(X(\theta)) &= \int_0^1 \left((y_P^* - y_P(\theta))^2 - (X(\theta) - y_P(\theta))^2\right)\mathrm{d}F(\theta) \\ &= -\left(y_P^*\right)^2 + \int_0^1 2X(\theta)y_P(\theta)\mathrm{d}F(\theta) - \int_0^1 X^2(\theta)\mathrm{d}F(\theta). \qquad (13)\end{aligned}$$

First, integrating by parts the second term of the right hand side of (13) we have $\int_0^1 2X(\theta)y_P(\theta)\mathrm{d}F(\theta) = 2X(1)y_P^* - \int_0^1\left[\int_0^\theta 2y_P(s)\mathrm{d}F(s)\right]\mathrm{d}X(\theta)$. Note that for every state θ, incentive compatibility of the agent implies that $X^-(\theta) + X^+(\theta) = 2y_A(\theta)$, which implies that

$$(X^+(\theta))^2 - (X^-(\theta))^2 = 2y_A(\theta)\left(X^+(\theta) - X^-(\theta)\right). \qquad (14)$$

We can then integrate by parts the third term of the right hand side of (13) to find that $\int_0^1 X^2(\theta)\mathrm{d}F(\theta) = X^2(1) - \int_0^1 2y_A(\theta)F(\theta)\mathrm{d}X(\theta)$. Rearranging terms it follows that

$$\begin{aligned}\Delta(X(\theta)) &= -\left(y_P^*\right)^2 + 2X(1)y_P^* - X^2(1) \\ &\quad + \int_0^1 2y_A(\theta)F(\theta)\mathrm{d}X(\theta) - \int_0^1\int_0^\theta 2y_P(s)\mathrm{d}F(s)\mathrm{d}X(\theta) \\ &= -\left(y_P^* - X(1)\right)^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) \qquad (15)\end{aligned}$$

Alternatively we can integrate by parts the second term on the right hand side of (13) to obtain $\int_0^1 2X(\theta)y_P(\theta)\mathrm{d}F(\theta) = 2X(0)y_P^* + \int_0^1\left[\int_\theta^1 2y_P(s)\mathrm{d}F(s)\right]\mathrm{d}X(\theta)$. By application of (14) and integration by parts we have $\int_0^1 X^2(\theta)\mathrm{d}F(\theta) = X^2(0) + \int_0^1 2y_A(\theta)(1 - F(\theta))\mathrm{d}X(\theta)$. Rearranging terms we finally obtain

$$\begin{aligned}\Delta(X(\theta)) &= -(y_P^*)^2 + 2X(0)y_P^* - X^2(0) - \int_0^1 2y_A(\theta)(1-F(\theta))\mathrm{d}X(\theta) \\ &\quad + \int_0^1 \left[\int_\theta^1 2y_P(s)\mathrm{d}F(s)\right]\mathrm{d}X(\theta) \\ &= -(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta). \end{aligned} \tag{16}$$

Therefore, $V = \max_{X(\theta)\in X} \Delta(X(\theta)) = \max_{X(\theta)\in X} -(y_P^* - X(1))^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) = \max_{X(\theta)\in \widehat{X}_D} -(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta)$. ■

Proof of Proposition 2: *Necessity*: We will prove the contra-positive, i.e. if there exists $\theta \in (0,1)$ such that $T(\theta) > 0$ and $S(\theta) < 0$ then $V > 0$. Let $\theta^* \in (0,1)$ be such that $S(\theta^*) < 0 < T(\theta^*)$ and let $y = y_A(\theta^*)$. Consider the delegation set D comprised of only two decisions such that at θ^* the agent is indifferent between the two decisions, i.e. $D = \{y - d, y + d\}$ with $d > 0$. The difference in the principal's expected utility from D and centralization is given by $\Delta(X(\theta)) \equiv -(y_P^* - X(1))^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) = -(y_P^* - (y+d))^2 + 4T(\theta^*)d$. Selecting $d = T(\theta^*) - S(\theta^*) > 0$ we have $\Delta(X(\theta)) = -(2T(\theta^*))^2 + 4T(\theta^*)[T(\theta^*) - S(\theta^*)] = -4T(\theta^*)S(\theta^*) > 0$. Therefore $V > 0$.

Sufficiency: Note that the condition in the proposition is equivalent to requiring that for all $\theta \in (0,1)$, $T(\theta) \leq 0$ or $S(\theta) \geq 0$. Suppose first that for all $\theta \in (0,1)$ $T(\theta) \leq 0$. Then, for all $X(\theta) \in X$ we have that $\int_0^1 T(\theta)\mathrm{d}X(\theta) \leq 0$ and from (9) $-(y_P^* - X(1))^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) \leq 0$. Therefore $V = \max_{X(\theta)\in \widehat{X}_D} -(y_P^* - X(1))^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) = 0$.

Next consider the case that $S(\theta) \geq 0$ for all $\theta \in (0,1)$. Then, for all $X(\theta) \in X$ we have that $\int_0^1 S(\theta)\mathrm{d}X(\theta) \geq 0$ and from (16) $-(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta) \leq 0$. This implies that $V = \max_{X(\theta)\in \widehat{X}_D} -(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta) = 0$. ■

Proof of Proposition 3: *Necessity*: Suppose that $D^* = [\underline{y}, \overline{y}]$ is a minimal optimal delegation set and let $\underline{\theta} = y_A^{-1}(\underline{y})$ and $\overline{\theta} = y_A^{-1}(\overline{y})$. From Proposition 1 it is then necessary that $\widetilde{T}(y)$ is convex for $y \in [\underline{y}, \overline{y}]$. The expected utility of the principal under D^* is given by

$$U_P = -\int_0^{\underline{\theta}} [y_A(\underline{\theta}) - y_P(\theta)]^2 \mathrm{d}F(\theta) - \int_{\underline{\theta}}^{\overline{\theta}} [y_A(\theta) - y_P(\theta)]^2 \mathrm{d}F(\theta) - \int_{\overline{\theta}}^1 \left[y_A(\overline{\theta}) - y_P(\theta)\right]^2 \mathrm{d}F(\theta)$$

Optimality of D^* requires $\underline{\theta}$ and $\overline{\theta}$ to satisfy the first order conditions

$$\begin{aligned}\frac{\partial U_P}{\partial \underline{\theta}} &= -2y_A'(\underline{\theta}) \int_0^{\underline{\theta}} \left[y_A(\underline{\theta}) - y_P(\theta)\right] \mathrm{d}F(\theta) = 0 \\ \frac{\partial U_P}{\partial \overline{\theta}} &= -2y_A'(\overline{\theta}) \int_{\overline{\theta}}^{1} \left[y_A(\overline{\theta}) - y_P(\theta)\right] \mathrm{d}F(\theta) = 0\end{aligned}$$

Since, by assumption, $y_A'(\theta) > 0$ we must have $y_A(\overline{\theta}) = \mathrm{E}\left[y_P(s)\,|s \geq \overline{\theta}\right]$ and $y_A(\underline{\theta}) = \mathrm{E}\left[y_P(s)\,|s \leq \underline{\theta}\right]$.

Finally, since $D^* = \left[\underline{y}, \overline{y}\right]$ is optimal, U_P cannot increase if the principal adds decisions below $\underline{y}$ and above $\overline{y}$ to the delegation set D^*. First consider adding decisions below $\underline{y}$ and, for each $\widetilde{\theta} < \overline{\theta}$ consider adding the decision y to D^* such that the agent at state $\widetilde{\theta}$ is indifferent between the lower bound $\underline{y}$ and the new decision y, i.e. $y_A(\widetilde{\theta}) = (\underline{y} + y)/2$. Let $X_{D^*}(\theta)$ and $X_{D^* \cup \{y\}}(\theta)$ be an outcome function associated with D^* and $D^* \cup \{y\}$, respectively. These two functions only differ for $\theta \leq \widetilde{\theta}$, where $X_{D^* \cup \{y\}}(\theta)$ selects the new decision y and has a jump discontinuity at $\widetilde{\theta}$ of magnitude $\underline{y} - y$. Using the representation (15) we have that the increment in the expected utility of the principal by adding a new decision y is $\Delta U_P = 2T(\widetilde{\theta})\left(\underline{y} - y\right)$. Optimality of D^* implies that $\Delta U_P \leq 0$ and therefore $T(\widetilde{\theta}) \leq 0$. A similar reasoning shows that adding a decision y above $\overline{y}$ leads to a variation in the principal's expected utility $\Delta U_P = -2S(\widetilde{\theta})\left(\overline{y} - y\right)$, where $\widetilde{\theta} \geq \overline{\theta}$ is such that $y_A(\widetilde{\theta}) = \left(\overline{y} + y\right)/2$. Optimality of D^* implies that $\Delta U_P \leq 0$ and therefore $S(\widetilde{\theta}) \geq 0$.

Sufficiency: We establish sufficiency by proving that i. centralization is not optimal, ii. an optimal delegation set has no decisions above $\overline{y}$ and no decisions below $\underline{y}$, and iii. D^* is an interval and $D^* = \left[\underline{y}, \overline{y}\right]$.

i. Centralization is not optimal: Note that, since $\widetilde{T}(\underline{y}) = 0$ and $\widetilde{T}(y) \leq 0$ for $y < \underline{y}$, $\widetilde{T}(y)$ is (weakly) increasing at $y = \underline{y}$. Convexity of $\widetilde{T}(y)$ in $\left[\underline{y}, \overline{y}\right]$ implies that $\widetilde{T}(y) > 0$ for $y \in \left(\underline{y}, \overline{y}\right)$. A similar argument applied to $\widetilde{S}(y)$ establishes that $\widetilde{S}(y) < 0$ for $y \in \left(\underline{y}, \overline{y}\right)$. By Proposition 3 it follows that centralization cannot be optimal.

ii. D^* is empty outside of $\left[\underline{y}, \overline{y}\right]$: First, since $\widetilde{T}(y) \leq 0$ for $y < \underline{y}$, by (9) it follows that an optimal delegation set must have at most one decision below $\underline{y}$. By a similar argument, from $\widetilde{S}(y) \geq 0$ for $y > \overline{y}$ and representation (10) an optimal delegation set must have at most one decision above $\overline{y}$.

Second, we establish that $D^* \cap \left[\underline{y}, \overline{y}\right] \neq \varnothing$, i.e. any optimal delegation set must contain at least one decision in $\left[\underline{y}, \overline{y}\right]$. Suppose not, i.e. $D^* \cap \left[\underline{y}, \overline{y}\right] = \varnothing$. Then, since centralization is not optimal and from the previous paragraph D^* must contain exactly two decisions. We will show that the optimal two-decision delegation set necessarily has at least one decision in $\left[\underline{y}, \overline{y}\right]$ thus reaching a contradiction. Let $D_{y^*} = \{y^* - d^*, y^* + d^*\}$ be an optimal two-decision delegation set. Then we must have $d^* = \widetilde{T}(y^*) - \widetilde{S}(y^*)$ and $\widetilde{T}(y^*) > 0, \widetilde{S}(y^*) < 0$, which requires that $y^* \in \left(\underline{y}, \overline{y}\right)$. If $D^* \cap \left[\underline{y}, \overline{y}\right] = \varnothing$ it must be that $y^* + d^* > \overline{y}$ and $y^* - d^* < \underline{y}$ which implies that $2\widetilde{T}(y^*) > \widetilde{T}(\overline{y}) = \widetilde{T}(\overline{y}) + \widetilde{T}(\underline{y}) > 2\widetilde{T}(\frac{\overline{y}+\underline{y}}{2})$ and $2\widetilde{S}(y^*) < \widetilde{S}(\underline{y}) = \widetilde{S}(\overline{y}) + \widetilde{S}(\underline{y}) < 2\widetilde{T}(\frac{\overline{y}+\underline{y}}{2})$, where in each case the last inequality follows, respectively, from the convexity of $\widetilde{T}(y)$ and the concavity of $\widetilde{S}(y)$ in $\left[\underline{y}, \overline{y}\right]$. Given that both $\widetilde{T}(y)$ and $\widetilde{S}(y)$ are increasing, the last two inequalities imply that $y^* > \left(\overline{y} + \underline{y}\right)/2$ and $y^* < \left(\overline{y} + \underline{y}\right)/2$ which leads to a contradiction. Therefore it must be that $D^* \cap \left[\underline{y}, \overline{y}\right] \neq \varnothing$.

Third, we prove that, since any optimal delegation set D^* must contain at least one decision in $\left[\underline{y}, \overline{y}\right]$, if there are decisions allowed by the principal in $D^* \cap \left[\underline{y}, \overline{y}\right]^c$ she can always increase her expected utility by either banning these decisions or appropriately increasing the discretion of her agent. This will contradict the assumed optimality of D^* and hence prove that $D^* \cap \left[\underline{y}, \overline{y}\right]^c = \varnothing$. We will only explicit show that $D^* \cap (\overline{y}, \infty) = \varnothing$ since the analysis required to prove $D^* \cap (-\infty, \underline{y}) = \varnothing$ is entirely analogous.

Suppose that $D^* \cap (\overline{y}, \infty) = \{y_2\}$ and let y_1 be the highest decision allowed to the agent in $D^* \cap \left[\underline{y}, \overline{y}\right]$. If $(y_1 + y_2)/2 > \overline{y}$ by (10) we see that the principal could obtain a higher expected utility by removing the decision y_2 from D^*. Now suppose that $(y_1 + y_2)/2 \leq \overline{y}$, which implies that $y_1 < \overline{y}$. Consider the delegation set $D^* \cup \{y_1 + \epsilon\}$ where $0 < \epsilon < 2\overline{y} - y_1 - y_2$. The increment of the principal's expected utility is $\Delta U = 2\left[\widetilde{T}(y_1 + \frac{\epsilon}{2})\epsilon + \widetilde{T}(\frac{y_1+y_2}{2} + \frac{\epsilon}{2})\left[y_2 - y_1 - \epsilon\right] - \widetilde{T}(\frac{y_1+y_2}{2})\left(y_2 - y_1\right)\right] > 0$ since $\widetilde{T}(y)$ is convex in $\left[\underline{y}, \overline{y}\right]$. We see that in both cases D^* cannot be optimal. Therefore we must have $D^* \cap (\overline{y}, \infty) = \varnothing$.

iii. D^* is an interval and $D^* = \left[\underline{y}, \overline{y}\right]$. Since $D^* \cap \left[\underline{y}, \overline{y}\right]^c = \varnothing$ and centralization is not optimal, by Proposition 1 D^* must be a (non-degenerate) interval contained in $\left[\underline{y}, \overline{y}\right]$. Therefore threshold delegation is optimal. We will now prove that indeed $D^* = \left[\underline{y}, \overline{y}\right]$.

Since the value of delegation from offering the agent an interval $[y_1, y_2]$ is given by $V = -(y_P^* - y_2)^2 + 2\int_{y_1}^{y_2} \widetilde{T}(y)\mathrm{d}y = -(y_P^* - y_1)^2 - 2\int_{y_1}^{y_2} \widetilde{S}(y)\mathrm{d}y$, by differentiating this expressions w.r.t. y_1 and y_2, respectively, we have that for an optimal $[y_1, y_2]$, $\widetilde{T}(y_1) = 0$ and $\widetilde{S}(y_2) = 0$. Therefore $D^* = \left[\underline{y}, \overline{y}\right]$. ■

Proof of Proposition 4: We will present a proof for the case of upper-threshold delegation since the case of lower-threshold delegation can be treated in a similar manner. Furthermore, we will explicitly include those steps that differ from the proof of Proposition 3**,** referring to this proof for all remaining details.

Necessity: If $D^* = [\underline{d}_A, \overline{y}]$ is a minimal optimal delegation set then by Proposition 1, $\widetilde{T}(y)$ is convex for $y \in [\underline{d}_A, \overline{y}]$. Since the principal cannot improve by adding decisions above $\overline{y}$ we must have $S(y) \geq 0$ for $y > \overline{y}$. Furthermore, optimality of $\overline{y}$ requires that $S(\overline{y}) \geq 0$. Finally, if $\widetilde{T}(\widetilde{y}) < 0$ for some $\widetilde{y} \in (\underline{d}_A, \overline{y})$ then convexity of $\widetilde{T}(y)$ and the fact that $\widetilde{T}(\underline{d}_A) = 0$ implies that $\widetilde{T}(y) < 0$ for $\in (\underline{d}_A, \widetilde{y})$. This leads to a contradiction since the principal could increase her expected utility by an amount $-2\int_{\underline{d}_A}^{\widetilde{y}} \widetilde{T}(y)\mathrm{d}y > 0$ by banning all decisions $[\underline{d}_A, \widetilde{y})$. Therefore we must have $\widetilde{T}(y) \geq 0$ for $y \in [\underline{d}_A, \overline{y}]$.

Sufficiency: We establish sufficiency by proving that i. centralization is not optimal, ii. an optimal delegation set has no decisions above $\overline{y}$, and iii. D^* is an interval and $D^* = [\underline{d}_A, \overline{y}]$.

i. Centralization is not optimal: Note that, since $\widetilde{S}(\overline{y}) = 0$ and $\widetilde{S}(y)$ is concave in $[\underline{d}_A, \overline{y}]$ (since $\widetilde{T}(y)$ is assumed convex in $[\underline{d}_A, \overline{y}]$), it follows that $\widetilde{S}(y) < 0$ for $y \in (\underline{d}_A, \overline{y})$. Since by assumption $\widetilde{T}(y) > 0$ in the same region, Proposition 1 implies that centralization cannot be optimal.

ii. D^* is empty above $\overline{y}$: This can be established by the same proof presented in Proposition 3 and is thus omitted.

iii. D^* is an interval and $D^* = [\underline{d}_A, \overline{y}]$. The value of delegation from offering the agent an interval $[y_1, y_2]$ is given by $V = -(y_P^* - y_2)^2 + 2\int_{y_1}^{y_2} \widetilde{T}(y)\mathrm{d}y = -(y_P^* - y_1)^2 - 2\int_{y_1}^{y_2} \widetilde{S}(y)\mathrm{d}y$. Since $\widetilde{T}(y) \geq 0$, we must have $y_1 = \underline{d}_A$. Also by differentiating w.r.t. y_2 we must have $\widetilde{S}(y_2) = 0$ which implies $y_2 = \overline{y}$. ■

Proof of Proposition 5: *Necessity*: Suppose complete delegation is optimal, i.e. $D^* = Y_A$. Now, let $S = Y_A \cap Y_P^c$. By Proposition 1-i D^* contains at most two points (one above and one below) outside the range of the principal Y_P. If $S \neq \varnothing$, S contains an open interval and hence $D^* \neq Y_A$, and we reach a contradiction. Therefore it must be that $S = \varnothing$ which implies that $Y_A \subseteq Y_P$.

Next we show that $\widetilde{T}(y)$ is increasing and convex and $\widetilde{S}(y)$ is increasing and concave. Convexity follows by noticing that for each $[u, v] \subset Y_A$ the delegation set $\widehat{D} = Y_A \cap [u, v]^c$ cannot improve upon Y_A, i.e. $\Delta((v-u)/2) > 0$ (where the increment is computed at $(v+u)/2$). By the proof of Lemma 6 this implies that $\widetilde{T}(y)$ is convex for $y \in Y_A$. Next we show $\widetilde{T}'(\underline{d}_A) \geq 0$ which, coupled with convexity of $\widetilde{T}(y)$, entails that $\widetilde{T}(y)$ is increasing for $y \in Y_A$. Suppose not, i.e. $\widetilde{T}'(\underline{d}_A) < 0$. Then $\widetilde{T}(y) < 0$ in a region $(\underline{d}_A, v)$. Consider now the delegation set $\widehat{D} = Y_A \cap [\underline{d}_A, v)^c$. The difference in expected utility from $\widehat{D}$ and D^* can be expressed as $\Delta U \equiv \mathrm{E}(u_P(y,\theta) \,\big|\, \widehat{D}) - \mathrm{E}(u_P(y,\theta)\,|D^*) = -\int_{\underline{d}_A}^{v} \widetilde{T}(y)\mathrm{d}y > 0$ which implies that D^* is not optimal. Therefore it must be that $\widetilde{T}'(\underline{d}_A) \geq 0$ and, consequently, $\widetilde{T}(y)$ is increasing. A similar argument shows that $\widetilde{S}(y)$ is concave, $\widetilde{S}(y) \leq 0$ for $y \in Y_A$, and, since $\widetilde{S}(\overline{d}_A) = 0$, $\widetilde{S}(y)$ is increasing.

Finally, since $\widetilde{T}(\underline{d}_A) + \widetilde{S}(\underline{d}_A) = \widetilde{S}(\underline{d}_A) < 0$ and $\widetilde{T}(\overline{d}_A) + \widetilde{S}(\overline{d}_A) = \widetilde{T}(\overline{d}_A) > 0$, continuity of $\widetilde{T}(y) + \widetilde{S}(y)$ implies that for some $y' \in (\underline{d}_A, \overline{d}_A)$ $\widetilde{T}(y') + \widetilde{S}(y') = y' - y_P^* = 0$. This establishes that $y_P^* \in (\underline{d}_A, \overline{d}_A)$.

Sufficiency: Note first that, under the conditions of the proposition, centralization can never be optimal. Indeed, since $y_P^* \in (\underline{d}_A, \overline{d}_A)$, $\widetilde{T}(\underline{d}_A) + \widetilde{S}(\underline{d}_A) = \underline{d}_A - y_P^* < 0$. Since $\widetilde{T}(\underline{d}_A) = 0$ this implies that $\widetilde{S}(\underline{d}_A) < 0$. By continuity and monotonicity of $\widetilde{T}(y)$ there exists a decision y such that $\widetilde{T}(y) > 0$ and $\widetilde{S}(y) < 0$. It follows from Proposition 2, that centralization is not optimal.

Next we show that a two-decision delegation set cannot be optimal. Let $D_{y^*} = \{y^* - d^*, y^* + d^*\}$ be the optimal two-decision delegation set. Then we must have $\widetilde{T}(y^*) > 0, \widetilde{S}(y^*) < 0$ and $d^* = \widetilde{T}(y^*) - \widetilde{S}(y^*)$. We first show that at least one of the decisions in D_{y^*} belongs to the range of the agent Y_A. Suppose not. Then it must be that $y^* + d^* > \overline{d}_A$ and $y^* - d^* < \underline{d}_A$ which, given the convexity of $\widetilde{T}(y)$ and the concavity of $\widetilde{S}(y)$ implies that $2\widetilde{T}(y^*) > \widetilde{T}(\overline{d}_A) = \widetilde{T}(\overline{d}_A) + \widetilde{T}(\underline{d}_A) > 2\widetilde{T}(\frac{\overline{d}_A + \underline{d}_A}{2})$ and $2\widetilde{S}(y^*) < \widetilde{S}(\underline{d}_A) = \widetilde{S}(\overline{d}_A) + \widetilde{S}(\underline{d}_A) < 2\widetilde{T}(\frac{\overline{d}_A + \underline{d}_A}{2})$. Given that both $\widetilde{T}(y)$ and $\widetilde{S}(y)$ are increasing, the last two inequalities imply that $y^* > \frac{\overline{d}_A + \underline{d}_A}{2}$ and $y^* < \frac{\overline{d}_A + \underline{d}_A}{2}$ which leads to a contradiction. To prove that D_{y^*} is not optimal, suppose that $y^* + d^* \in (\underline{d}_A, \overline{d}_A)$. The case that $y^* - d^* \in (\underline{d}_A, \overline{d}_A)$ can be treated similarly. Consider increasing the discretion of the agent by adding the decision $y^* + d^* - \varepsilon$, i.e. offering the delegation set $D = D_{y^*} \cup \{y^* - d^* - \varepsilon\}$. The increment in the expected utility from offering D is given by

$$\Delta U = 4d^* \left[\widetilde{T}(y^* + d - \frac{\epsilon}{2}) \frac{\epsilon}{2d^*} + \widetilde{T}(y^* - \frac{\epsilon}{2}) \left[1 - \frac{\epsilon}{2d^*} \right] - 2\widetilde{T}(y^*) \right].$$

Since $\widetilde{T}(y)$ is strictly convex, $\Delta U > 0$ and two-decision delegation cannot be optimal. Applying the same logic one can show that three-decision delegation also cannot be optimal. Since $\widetilde{T}(y)$ is strictly convex for $y \in Y_A \cap Y_P = Y_A$ and delegation with one, two or three decisions is never optimal, the optimal delegation set must consist of an interval in Y_A. Moreover, since $\widetilde{T}(y) \geq 0$ and $\widetilde{S}(y) \leq 0$, the optimal interval D must indeed be $D = Y_A$.

To complete the proof, we show that the optimal delegation set has no decisions outside the range of the agent. Indeed, suppose that a decision $\widehat{y}$ above $\overline{d}_A$ is added to the delegation set D such that $\widehat{y}$ is selected with positive probability. The case for decisions below $\underline{d}_A$ is entirely analogous. Let $\overline{y} < \overline{d}_A$ be the highest decision allowed in the range of the agent Y_A. Consider now the delegation set $D \cup \{\overline{y} + \epsilon\}$, where $0 < \epsilon < 2\overline{d}_A - \widehat{y} - \overline{y}$. Since $\widetilde{T}(y)$ is convex, the increment of the principal's expected utility is $\Delta U = 2 \left[\widetilde{T}(\overline{y} + \frac{\epsilon}{2})\epsilon + \widetilde{T}(\frac{\widehat{y}+\overline{y}}{2} + \frac{\epsilon}{2}) [\widehat{y} - \overline{y} - \epsilon] - \widetilde{T}(\frac{\widehat{y}+\overline{y}}{2}) (\widehat{y} - \overline{y}) \right] > 0$. We therefore reach a contradiction implying that adding projects above and below Y_A can never be optimal. Therefore complete delegation is optimal. ■

Proof of Proposition 6: Let $\widehat{F}(y, \lambda) = F(y_A^{-1}(y, \lambda))$ and $\widehat{f}(y, \lambda) = f(y_A^{-1}(y, \lambda))$. It suffices to prove that there exists a $\overline{\lambda} \in (0, 1)$ such that for $\lambda > \overline{\lambda}$, $\widetilde{T}(y, \lambda) = \widehat{F}(y, \lambda)y - \int_0^{y_A^{-1}(y,\lambda)} y_P(\theta) \mathrm{d}F(\theta)$ is strictly convex for all $y \in Y_A(\lambda)$. If this condition is satisfied then for each $\lambda > \overline{\lambda}$ interval delegation is optimal.

For $y \in Y_A(\lambda)$ define $r(y, \lambda) \equiv \frac{\partial}{\partial y} \left[y_A^{-1}(y, \lambda) \right] \widehat{f}(y, \lambda) = \widehat{f}(y, \lambda) / \left[1 - \lambda + \lambda y_P'(y_A^{-1}(y, \lambda)) \right]$. Given our conditions on $y_A(\theta)$, $y_P(\theta)$ and $f(\theta)$, $r(y, \lambda)$ is continuously differentiable in the compact set $\Omega = \{(y, \lambda) : y \in Y_A(\lambda), \lambda \in [0, 1]\}$. By successive differentiation we obtain that

$$\begin{aligned}&\left[1-\lambda+\lambda y_P'(y_A^{-1}(y,\lambda)\right]\frac{\partial^2}{\partial y^2}\widetilde{T}(y,\lambda) \\ = \quad &\widehat{f}(y,\lambda)+(1-\lambda)\left[1-\lambda+\lambda y_P'(y_A^{-1}(y,\lambda)\right]\frac{\partial}{\partial y}\left[b(y)r(y,\lambda)\right].\end{aligned} \tag{17}$$

From the assumption $\min f(\theta)=\underline{f}>0$ and the fact that the term

$$\left|\left[1-\lambda+\lambda y_P'(y_A^{-1}(y,\lambda)\right]\frac{\partial}{\partial y}\left[b(y)r(y,\lambda)\right]\right|$$

is uniformly bounded in Ω, say by M, we infer the existence of a $\overline{\lambda}\in(0,1)$ such that $\underline{f}+(1-\overline{\lambda})M>0$. Thus, $\forall\lambda>\overline{\lambda}$ the RHS of (17) is strictly positive and thus $\frac{\partial^2}{\partial y^2}\widetilde{T}(y,\lambda)>0$ $\forall\lambda>\overline{\lambda}$, $y\in Y_A(\lambda)$. This establishes the convexity of $\widetilde{T}(y,\lambda)$ $\forall\lambda>\overline{\lambda}$. ■

Proof of Result 2: In this proof we will make use of the following facts which are adaptations to our setting of similar results proven in Holmström 1984: a. $\mathrm{E}\left[s\,|s\geq\theta\right]>\theta$ $\forall\theta\in(-a,a)$, b. $0<\frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}\left[s\,|s\leq\theta\right]<1$ and $0<\frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}\left[s\,|s\geq\theta\right]<1$ $\forall\theta\in(-a,a)$, c. $\frac{\mathrm{d}}{\mathrm{d}\sigma}$ $\mathrm{E}\left[s\,|s\geq\theta\right]\geq0$.

We first note that $T(\theta)$ and $S(\theta)$ take both positive and negative values in $(-a,a)$. To apply Proposition 3 we will show that $T(\underline{\theta})=0$ and $S(\overline{\theta})=0$ have unique solutions in $(-a,a)$ and that $\widetilde{T}(y)$ is convex in $[m_A\underline{\theta},\,m_A\overline{\theta}]$. Note that for $\theta\in(-a,a)$, $T(\theta)=0$ if and only if $R(\theta)\equiv m_A\theta-\mathrm{E}\left[s\,|s\leq\theta\right]=0$. Since $R'(\theta)=m_A-\frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}\left[s\,|s\leq\theta\right]>0$, $R(\underline{\theta})=0$ has a unique solution $\underline{\theta}\in(-a,a)$. The same rationale applies to $S(\overline{\theta})=0$. Since it must be that $\underline{\theta}<0<\overline{\theta}$, these two values define a nondegenerate interval $[\underline{\theta},\overline{\theta}]\subset[-a,a]$. Furthermore, since $T(\theta)=-S(-\theta)$ we necessarily have that $\overline{\theta}=-\underline{\theta}$.

Since $T''(\theta)=\left[2m_A-1-(m_A-1)\theta^2/\sigma^2\right]f(\theta)$ there is a value $\widehat{\theta}$ such that $T''(\theta)$ is strictly convex for $\theta\in\Omega_c=\left[-\widehat{\theta},\widehat{\theta}\right]$. Since $T'(-a)=f(-a)a\left[1-m_A\right]<0$ and $T'(\underline{\theta})=F(\underline{\theta})\left[m_A-\frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}\left[s\,|s\leq\theta\right]\big|_{\theta=\underline{\theta}}\right]>0$ it must be that $T''(\underline{\theta})>0$ and therefore $\underline{\theta}\in\Omega_c$ and $-\underline{\theta}=\overline{\theta}\in\Omega_c$. This proves that $\widetilde{T}(y)$ is convex in $[\underline{\theta},\overline{\theta}]$.

$S(\overline{\theta})=0$ can be rewritten as $m_A\overline{\theta}=\mathrm{E}\left[s\,\middle|s\geq\overline{\theta}\right]$. Totally differentiating and using results b. and c. above we see that $\mathrm{d}\overline{\theta}/\mathrm{d}m_A<0$ and $\mathrm{d}\overline{\theta}/\mathrm{d}\sigma>0$. ■

Proof of Result 3: Note that $Y_A\subset Y_P$. We have shown above that if $1/2\leq m_A<1$, then $\widetilde{T}(y)$ is convex and $\widetilde{S}(y)$ is concave in $[-m_Aa,m_Aa]$. Since $T'(-a)=$

$f(-a)a\,[1-m_A] > 0$ and $S'(a) = -f(a)a\,[m_A - 1] > 0$, we must then have $\widetilde{T}'(y) > 0$ and $\widetilde{S}'(y) > 0$ $\forall y \in [-m_A a, m_A a]$. The result then follows from Proposition 5. ■

Proof of Result 4: We first partition the decision space Y in five regions Ω_i $i \in \{1, ..., 5\}$. Ω_1 and Ω_5 denote the region below and the region above the range of preferred decisions of the agent $[-m_A a, m_A a]$. In the text it has been shown that if $m_A < 1/2$ then $\widetilde{T}(y)$ is always concave in a neighborhood of zero. Given the assumption that $m_A a \geq 2\sigma^2\,[f(0) - f(a)]$, there exists $\widetilde{y} \in (0, m_A a)$ such that $\widetilde{T}''(\pm\widetilde{y}) = 0$, with $\widetilde{T}(y)$ convex in the regions above $\widetilde{y}$ and below $-\widetilde{y}$. Let Ω_2 and Ω_4 be the regions where $\widetilde{T}(y)$ is convex, namely $\Omega_2 = [-m_A a, -\widetilde{y}]$ and $\Omega_4 = [\widetilde{y}, m_A a]$. Finally let $\Omega_3 = (-\widetilde{y}, \widetilde{y})$ be the region where $\widetilde{T}(y)$ is concave. We will establish this result by proving that if D^* is an optimal *minimal* delegation set then i. $D^* \cap (\Omega_1 \cup \Omega_2) \neq \varnothing$ and $D^* \cap (\Omega_4 \cup \Omega_5) \neq \varnothing$, ii. if $D^* \cap \Omega_2$ is non-empty, then $D^* \cap \Omega_1$ is empty, and if $D^* \cap \Omega_4$ is non-empty, then $D^* \cap \Omega_5$ is empty, iii. $D^* \cap \Omega_3 = \varnothing$, iv. if $m_A a \geq 2\sigma^2\,[f(0) - f(a)]$ then $D^* \cap \Omega_1 = D^* \cap \Omega_5 = \varnothing$. These steps show that an optimal delegation set D^* consists of one interval for positive decisions and one interval for negative decisions. We conclude by characterizing the bounds of both intervals. The comparative statics follow immediately by the same considerations of Result 2.

i. $D^* \cap (\Omega_1 \cup \Omega_2) \neq \varnothing$ and $D^* \cap (\Omega_4 \cup \Omega_5) \neq \varnothing$: Note that for $0 < m_A < 1/2$ both effective biases have a constant sign, i.e. $\widetilde{T}(y) > 0$ and $\widetilde{S}(y) < 0$ for $y \in (-m_A a, m_A a)$. This implies that $D^* \cap (\Omega_1 \cup \Omega_2) \neq \varnothing$ and $D^* \cap (\Omega_4 \cup \Omega_5) \neq \varnothing$.

ii. $D^* \cap \Omega_2 \neq \varnothing \Rightarrow D^* \cap \Omega_1 = \varnothing$: Suppose on the contrary that $y_1 \in D^* \cap \Omega_1$ and let y_2 be the smallest decision in $D^* \cap \Omega_2$. We will show that the principal can improve by adding decisions in Ω_2, contradicting the assumed optimality of D^*. Consider adding the decision $y_2 - 2\epsilon$ to D^*, where ϵ is such that $(y_1 + y_2)/2 - \epsilon$ belongs to Ω_2. The increment of the expected utility of the principal is $\Delta U = 2\left[\widetilde{T}(y_2 - \epsilon)2\epsilon + \widetilde{T}(\frac{y_1+y_2}{2} - \epsilon)\,[y_2 - y_1 - 2\epsilon] - \widetilde{T}(\frac{y_1+y_2}{2})\,(y_2 - y_1)\right]$. Letting $\widetilde{T}(\frac{y_1+y_2}{2} - \epsilon) = \widetilde{T}(\frac{y_1+y_2}{2}) - \epsilon\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}(\xi)$ with $\xi \in \left(\frac{y_1+y_2}{2} - \epsilon, \frac{y_1+y_2}{2}\right)$ we have that $\Delta U = 4\epsilon\int_{\frac{y_1+y_2}{2}}^{y_2-\epsilon}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\mathrm{d}y - 4\epsilon\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}(\xi)\left(\frac{y_2-y_1-2\epsilon}{2}\right)$. Since $\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}$ is strictly increasing in Ω_2 we must have $\Delta U > 0$. Replacing $\widetilde{T}(y)$ with $\widetilde{S}(y)$ in the preceding argument allows also to establish that if $D^* \cap \Omega_4 \neq \varnothing$, then $D^* \cap \Omega_5 = \varnothing$.

iii. $D^* \cap \Omega_3 = \varnothing$: Suppose on the contrary that $D^* \cap \Omega_3 \neq \varnothing$. Let $\underline{y}_3 \leq \overline{y}_3$ be the two decisions allowed by the principal in Ω_3 (where we could have $\underline{y}_3 = \overline{y}_3$) and y_2 and y_4 be the highest decision in $D^* \cap (\Omega_1 \cup \Omega_2)$ and the lowest decision in $D^* \cap (\Omega_4 \cup \Omega_5)$, respectively. Suppose first that $y_2 \in \Omega_2$ and that $\left(y_2 + \underline{y}_3\right)/2 < -\widetilde{y}$. By offering instead the delegation set $D^* \cup \{y_2 + 2\epsilon\}$ such that $\left(y_2 + \underline{y}_3\right)/2 + \epsilon < -\widetilde{y}$, the principal can increment his expected utility by $\Delta U = 4\epsilon \left[\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}(\xi)\left(\frac{\underline{y}_3 - y_2 - 2\epsilon}{2}\right) - \int_{y_2+\epsilon}^{\frac{y_2+\underline{y}_3}{2}} \frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\mathrm{d}y\right]$, where $\xi \in \left(\frac{y_2+\underline{y}_3}{2}, \frac{y_2+\underline{y}_3}{2} + \epsilon\right) \subset \Omega_2$. Since $\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}$ is strictly increasing in Ω_2 we must have $\Delta U > 0$ and thus we reach a contradiction. Suppose now that $y_2 \in \Omega_1$ and that $\left(y_2 + \underline{y}_3\right)/2 < -\widetilde{y}$. Then the first order condition on y_2 requires that $-\widetilde{T}(\frac{y_2+\underline{y}_3}{2}) + \frac{1}{2}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\left(\frac{y_2+\underline{y}_3}{2}\right)\left(\underline{y}_3 - y_2\right) = 0$ which can be written as $\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\left(\frac{y_2+\underline{y}_3}{2}\right)\left(\frac{\underline{y}_3 - y_2}{2}\right) = \int_{-m_A a}^{\frac{y_2+\underline{y}_3}{2}} \frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\mathrm{d}y$. Since $\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}$ is strictly increasing and positive in Ω_2 we have that $\int_{-m_A a}^{\frac{y_2+\underline{y}_3}{2}} \frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\mathrm{d}y < \frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\left(\frac{y_2+\underline{y}_3}{2}\right)\left(\frac{y_2+\underline{y}_3}{2} - (-m_A a)\right)$ implying that $\left(\frac{\underline{y}_3 - y_2}{2}\right) < \left(\frac{y_2+\underline{y}_3}{2} - (-m_A a)\right)$ and $y_2 > -m_A a$. This contradicts the fact that $y_2 \in \Omega_1$.

The same argument would reach a contradiction if we had assumed that $\left(y_4 + \underline{y}_3\right)/2 > \widetilde{y}$. But if we assume that $\left(y_2 + \underline{y}_3\right)/2 \geq -\widetilde{y}$ and $\left(y_4 + \underline{y}_3\right)/2 \leq \widetilde{y}$ then, from the concavity of $\widetilde{T}(y)$ in Ω_3, the principal can improve by banning both decisions $\underline{y}_3$ and $\overline{y}_3$. In both cases we reach a contradiction and thus conclude that $D^* \cap \Omega_3 = \varnothing$.

iv. If $D^* \cap \Omega_1 \neq \varnothing$ and $D^* \cap \Omega_5 \neq \varnothing$, the optimal delegation set consists of two decisions, both outside the range of the agent. The optimal two-decision delegation set is given by $\{-q, q\}$ where $q = \mathrm{E}[s \,|s \geq 0] = 2\sigma^2 [f(0) - f(a)]$. Therefore whenever $\mathrm{E}[s\,|s \geq 0] < m_A a$ any optimal minimal delegation set must have only decisions in the regions where $\widetilde{T}(y)$ is convex.

v. Optimal upper and lower bounds $\underline{y}$ and $\overline{y}$: Let $D^* = [-m_A a, \underline{y}] \cup [\overline{y}, m_A a]$. The first order conditions on $\underline{y}, \overline{y}$ are given by $\widetilde{T}(\frac{\underline{y}+\overline{y}}{2}) - \widetilde{T}(\underline{y}) = \frac{1}{2}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}(\frac{\underline{y}+\overline{y}}{2})\left(\overline{y} - \underline{y}\right)$ and $\widetilde{T}(\overline{y}) - \widetilde{T}(\frac{\underline{y}+\overline{y}}{2}) = \frac{1}{2}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}(\frac{\underline{y}+\overline{y}}{2})\left(\overline{y} - \underline{y}\right)$. Given the symmetry in the model we have $\underline{y} = -\overline{y}$ and $\widetilde{T}(\overline{y}) - \widetilde{T}(0) = \frac{1}{2}\overline{y}$ or $\overline{y} = \sigma^2 \frac{f(0) - f(\overline{y}/m_A)}{F(\overline{y}/m_A) - 1/2}$. ■

References

[1] Aghion, P. and J. Tirole (1997), "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105, 1-29.

[2] Alonso, R. and N. Matouschek (2004), "Relational Delegation," CEPR Discussion Paper No. 4870.

[3] Armstrong, M. (1995), "Delegating Decision-Making to an Agent with Unknown Preferences," mimeo, UCL.

[4] Bower, J. (1970), *Managing the Resource Allocation Process: A Study of Corporate Planning and Investment*, Boston, Division of Research, Graduate School of Business Administration, Harvard.

[5] Crawford, V. and J. Sobel (1982), "Strategic Information Transmission," *Econometrica*, 50, 1431-1452.

[6] Dessein, W. (2002), "Authority and Communication in Organizations," *Review of Economic Studies*, 69, 811-838.

[7] Gilligan, T. and K. Krehbiel (1987), "Collective Decision-Making and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures," *Journal of Law, Economics, and Organization*, 3, 287-335.

[8] Harris, M. and A. Raviv (1996), "The Capital Budgeting Process, Incentives and Information," Journal of Finance, 51, 1139-1174.

[9] Harris, M. and A. Raviv (1998), "Capital Budgeting and Delegation," *Journal of Financial Economics,* 50, 259-289.

[10] Holmström, B. (1977), "On Incentives and Control in Organizations," Ph.D. Dissertation, Stanford.

[11] ____ (1984), "On the Theory of Delegation," in M. Boyer and R. Kihlstrom (eds.) *Bayesian Models in Economic Theory*, New York: North-Holland.

[12] Jensen, M. and W. Meckling (1992), "Specific and General Knowledge and Organizational Structure," in L. Werin and H. Wijkander (eds.) *Contract Economics*, Oxford: Blackwell.

[13] Krishna, V. and J. Morgan (2001), "Asymmetric information and Legislative Rules: Some Amendments," *American Political Science Review*, 95, 435-452.

[14] ____ and ____ (2004), "Contracting for Information under Imperfect Commitment," mimeo, Penn State and UC Berkeley.

[15] Marino, A. and J. Matsusaka (2004), "Decision Processes, Agency Problems, and Information: An Economic Analysis of Capital Budgeting Procedures," *Review of Financial Studies*, forthcoming.

[16] Martimort, D. and A. Semenov (2005), "Communication by Multiple Interest Groups," mimeo, University of Toulouse.

[17] Melumad, N. and T. Shibano (1991), "Communication in Settings with No Transfers," *Rand Journal of Economics*, 22, 173-198.

[18] Ottaviani, M. (2000), "Economics of Advice," mimeo, London Business School.

[19] Rudin, W. (1987), *Real and Complex Analysis*, McGraw-Hill, Third Edition.

[20] Shavell, S. (2005), "Optimal Discretion in the Application of Rules," Discussion Paper No. 509, John M. Olin Center for Law, Economics, and Business.

[21] Strausz, R. (2003), "Deterministic Mechanisms and the Revelation Principle," *Economics Letters,* 79, 333-37.

[22] Szalay, D. (2004), "The Economics of Clear Advice and Extreme Options," *Review of Economic Studies,* forthcoming.

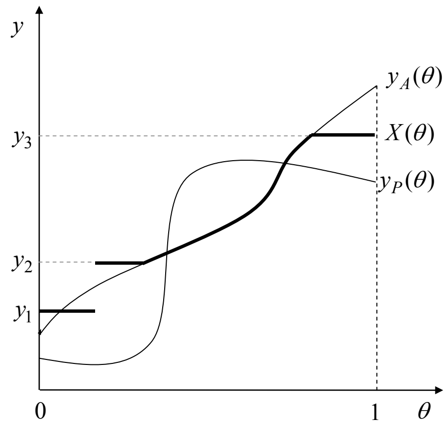

Figure 1: An Incentive Compatible Outcome Function

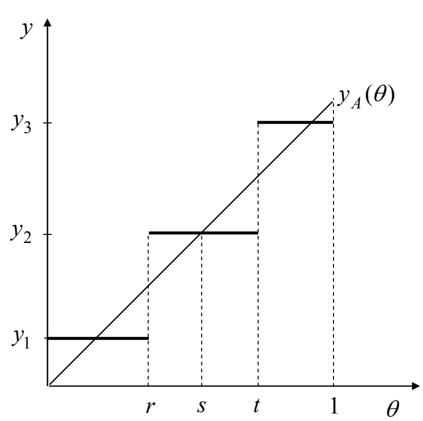

Figure 2a: Delegation set D

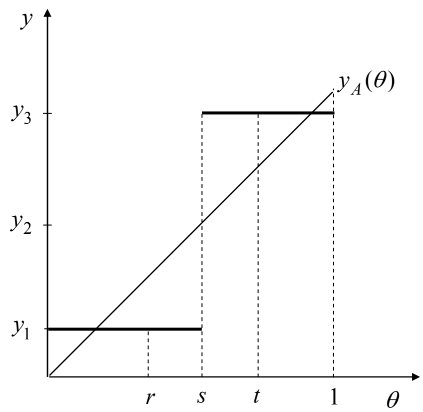

Figure 2b: Delegation set $\hat{D}$

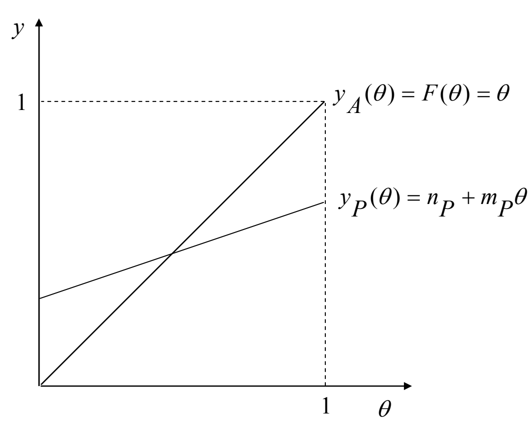

Figure 3: Melumad and Shibano (1991)

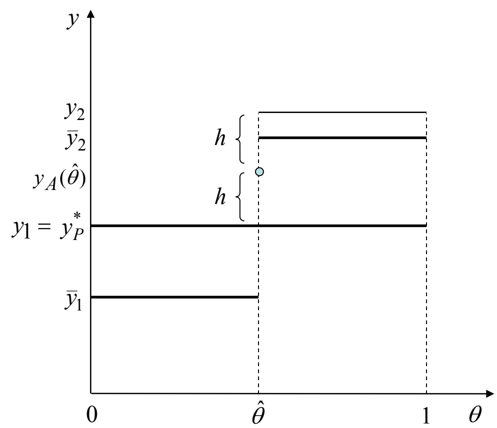

Figure 4: Centralization

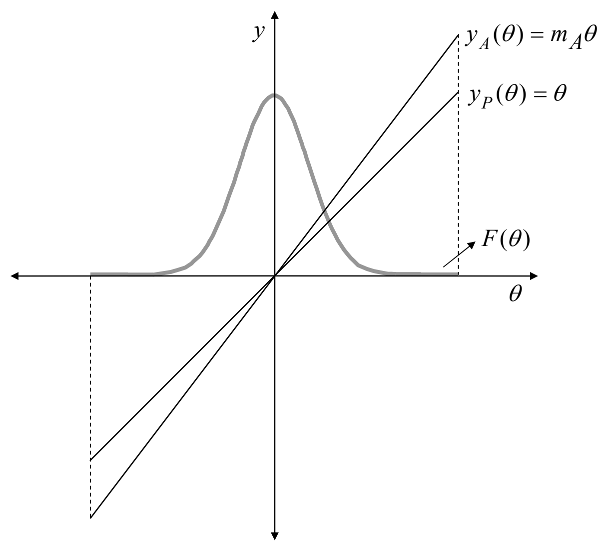

Figure 5: The Normal-Linear Example