**Ricardo Alonso and Niko Matouschek**

# Optimal delegation

## Article (Accepted version) (Refereed)

**Original citation:**
Alonso, Ricardo and Matouschek, Niko (2008) *Optimal delegation*. Review of Economic Studies, 75 (1). pp. 259-293. ISSN 0034-6527 DOI: 10.1111/j.1467-937X.2007.00471.x

# Optimal Delegation*

Ricardo Alonso
University of Southern California

Niko Matouschek
Northwestern University

**Abstract**

We analyze the design of decision rules by a principal who faces an informed but biased agent and who is unable to commit to contingent transfers. The contracting problem reduces to a delegation problem in which the principal commits to a set of decisions from which the agent chooses his preferred one. We characterize the optimal delegation set and perform comparative statics on the principal's willingness to delegate and the agent's discretion. We also provide conditions for interval delegation to be optimal and show that they are satisfied when the agent's preferences are sufficiently aligned. Finally, we apply our results to the regulation of a privately informed monopolist and to the design of legislatives rules.

Keywords: delegation, decision rights

JEL Classification: D82, L23

*We are grateful to the Editor Andrea Prat and to three anonymous referees for their comments and suggestions. We would also like to thank David Austen-Smith, Simon Board, Jim Dana, Mike Fishman, Joshua Gans, Tom Hubbard, Mark Satterthwaite, Scott Schaefer, Kathy Spier, Asher Wolinsky and, especially, Marco Ottaviani for very helpful discussions. Finally, we thank participants at various seminars and conferences for their comments. All errors are our own.

# 1 Introduction

Organizations are governed by rules. A key function of these rules is to limit the agency costs that principals incur when they involve better informed but biased agents in the decision making. The decision rules that organizations employ for this purpose take many different forms. Some rules allow the agents to make any decisions but give the principals the right to overrule or otherwise adjust them. When the U.S. House of Representatives delegates the drafting of bills to standing committees, for instance, it employs a variety of legislative rules that specify if and how the committees' proposals can be amended. Other rules simply specify what decisions agents are and are not allowed to make and give them the right to freely choose among the permissible decisions. Regulators who delegate pricing decisions, for instance, often specify price caps below which the regulated firms can set any prices. In this paper we analyze the optimal design of such rules. In particular, we analyze the design of decision rules by a principal who faces an informed but biased agent and who is unable to commit to contingent monetary transfers.

To illustrate the issues we are interested in, consider the classic problem of regulating a monopolist who is privately informed about his costs (Baron and Myerson 1982). In contrast to the standard literature, however, suppose that transfers between the regulator and the monopolist are ruled out by law.[1] In this setting, how should the regulator decide on the monopolist's price? Should she simply impose the welfare maximizing price given the expected costs or should she involve the monopolist in the price setting process? In the latter case, what form should the monopolist's involvement take? Should he be allowed to set any price below a price cap, for instance, or should the regulator opt for some other decision making process? In short, what is the best decision rule the regulator can adopt?

The regulation problem illustrates the three main features of our model: (i.) an organization that consists of a principal and an agent must make a decision. The principal and the agent have different preferences over the possible decisions and their preferred decisions depend on the state of the world. In the regulation problem, the regulator and the monopolist disagree on the price that should be charged and their respective preferred prices vary with

[1] In many countries and industries, regulators are legally forbidden to make transfers to firms (Laffont and Tirole 1990, 1993; Armstrong and Sappington 2004; see also our discussion below).

the monopolist's costs. (ii.) 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. The regulator, for instance, has the legal right to set the price but the monopolist is better informed about his costs. (iii.) as mentioned before, contingent transfers between the principal and the agent are not feasible, for instance, because of legal constraints as in the regulation problem.

In this set-up the principal can commit to any deterministic decision rule.[2] The problem of finding the optimal decision rule reduces to a *delegation problem* (Holmström 1977, 1984). In this problem the principal simply decides on a set of decisions and the agent is then allowed to make any decision from this *delegation set.* In other words, the principal merely needs to decide what decisions the agent should and should not be allowed to make.

The first step in addressing this problem is to decide whether the principal should delegate any decision rights to the agent. In other words, the principal needs to decide whether she should let the agent choose between at least two alternatives or whether she should simply impose her best uninformed decision. In our setting the principal benefits from delegation if and only if the principal and the agent are *minimally aligned*, a condition that depends on the preferences and the agent's information. As one would expect, the principal is more willing to delegate to a more aligned agent. More surprisingly, in many cases the principal's willingness to delegate does not depend on the agent's informational advantage. In many specifications of the regulation problem, for example, the regulator benefits from giving the monopolist some discretion over pricing if and only if the expected marginal costs are above a threshold that depends only on the demand characteristics. An increase in the monopolist's informational advantage, as captured by a mean preserving spread of the cost distribution, therefore does not affect the regulator's willingness to give the monopolist some discretion over pricing.

Once the principal has determined that she can benefit from delegation, she must decide how much discretion the agent should have. In other words, she needs to decide which decisions the agent should be allowed to make and which should be ruled out. At first glance one might think that the principal simply rules out those decisions for which the agent is very biased. A key insight of our analysis is that this is not the case. In particular, whether a

[2] The analysis therefore encompases decisions rules that can be implemented through cheap talk (Crawford and Sobel 1982; Dessein 2002) and veto-delegation (Marino 2006; Mylovanov 2006), among many others.

decision should be ruled out depends not just on the level of the bias but also on its slope. To see this, suppose that for some states the slope of the agent's bias is not very big, that is, his preferred decisions are flat relative to the principal's. If this agent were allowed to make any decision, his decisions would, from the principal's perspective, be too insensitive to changes in the state. The principal may then rule out intermediate decisions to encourage more state-sensitive decision making and she may do so even if the agent is locally very aligned. This insight allows us to derive a full characterization of optimal delegation sets which is the central result in the paper.

We then build on the characterization result to gain additional insights into the design of decision rules. In practice principals often engage in *interval delegation*, that is, they let their agents make any decision from a single interval (Holmström 1977, 1984). To understand why this is so, we provide simple conditions for interval delegation to be optimal. These conditions imply that interval delegation is optimal whenever the agent is sufficiently aligned. This suggests that the apparent widespread use of interval delegation may be due to the ability of organizations to carefully screen their employees or to use incentive schemes that closely align their interests. Applied to the regulation problem, our results on interval delegation show that optimal regulation without transfers often takes a remarkably simple form. In particular, for any linear or constant elasticity demand curve and any unimodal cost distribution a welfare maximizing regulator cannot do better than to either give the monopolist no discretion at all or to impose a price cap below which the monopolist can choose his preferred price. This result is consistent with the widespread use of price cap regulation in the United Kingdom and the United States (Armstrong and Sappington 2004).

It may seem intuitive that a principal gives more discretion to a more aligned agent and to one with a bigger informational advantage. Indeed, in the political economy literature on delegation these comparative statics have been shown to hold in a number of models and have become known as the *Ally Principle* and the *Uncertainty Principle* respectively.[3] Also, Holmström (1977, 1984) has shown that if delegation sets are required to take the form of a single interval, then the Ally Principle holds under general conditions. In contrast, we show that these principles do not hold when delegation sets can take any form. In other words,

[3] See, for instance, Huber and Shipan (2006).

when the principal's ability to delegate is unrestricted, she may give less discretion to an agent with a bigger informational advantage or, indeed, to one who is less biased.

Our analysis can be used to gain new insights into well-known economic problems. We have already illustrated this for the problem of regulating an informed monopolist. Another example is the optimal design of legislative rules. As mentioned above, the U.S. House of Representatives uses a wide variety of such rules to regulate the relationship between the legislature and its standing committees. A large literature in political economy argues that these rules are designed to motivate the acquisition and transmission of information by the committees (Gilligan and Krehbiel 1987, 1989; Krishna and Morgan 2001).[4] The standard approach taken in this literature is to compare the performance of specific rules that are observed in practice. In contrast, our analysis can be used to determine the optimal among all possible rules. In particular, we show that the optimal legislative rule is closely related to, but different from, those considered in the literature.

A key feature of our set up is that the principal cannot commit to contingent transfers. We focus on such situations since we believe that they are widespread in practice and not yet sufficiently well understood in theory.[5] There are various reasons for why a principal may not be able to commit to contingent transfers. In some situations such transfers are simply ruled out by law. As mentioned above, for instance, there are many countries and industries in which regulators are legally forbidden to make transfers to firms.[6]

The inability to observe the agents decision and the associated payoffs are another reason for why a principal may be unable to commit to contingent transfers. In such a situation our analysis applies if the principal can restrict the decisions the agent can make through technological constraints. For instance, parents cannot observe what TV shows their children watch in their absence and are therefore unable to reward or punish them depending on their decisions. However, if they own a Tivo recorder, or a similar device, they can restrict the

[4] The literature almost exclusively assumes that the legislature is not able to use transfers to motivate the committees (a notable exception is Baron 2000).

[5] Most of the literature on mechanism design, and its applications to organizational structure (for an overview see Mookherjee 2006), assumes that the principal is able to commit to contingent transfers. In contrast, most of the literature on delegation rules out contingent transfers (see, for instance, Holmström 1977, 1984; Aghion and Tirole 1997; Dessein 2002; Szalay 2005).

[6] This fact has been explained by regulatory failures due to collusion (Laffont and Tirole 1990) and commitment problems (Laffont and Tirole 1993). See also Armstrong and Sappington (2004).

channels their children can watch while they are gone. Similarly, in larger organizations it can be prohibitively expensive for principals to observe all the decisions their agents make and available performance measures may depend on so many decisions, as well as external factors, that they cannot be used to influence the decision making of individual agents. In such organizations, however, the principals may still be able to rule out certain decisions altogether, such as making long-distance phone calls, using the internet etc.[7]

It should also be noted that in some situations the extra benefit of committing to contingent transfers, in addition to limiting an agent's discretion, can be quite limited, as we show in an example below. Moreover, in some cases, such as the regulation problem, optimal delegation takes a very simple form that does not depend on the details of the information structure. In contrast, optimal transfers rules are in general sensitive to changes in the information structure and they can be computationally difficult to derive. This suggests that in some situations principals may find it more economical to simply restrict the agent's discretion than to commit to a potentially complicated complete contract.

The rest of the paper is structured as follows. In the next section we briefly discuss the related literature and we then present the model in Section 3. In Section 4 we show that the problem of finding the optimal decision rule reduces to the delegation problem. The solution to this problem is characterized in Sections 5 and 6. In Section 7 we then discuss applications of our analysis before revisiting three of our main assumptions in Section 8. Finally, we conclude in Section 9. All proofs are in the appendices.

## 2 Literature Review

Our paper builds on, and borrows from, Holmström (1977, 1984). He was the first to pose the general class of delegation problems to which our model belongs. After posing the general problem, he provides conditions under which a solution exists. He also considers a series of examples in which he characterizes optimal delegation sets assuming, for the most part, that

[7] Even when a principal is able to commit to contingent transfers, she may choose not to do so due to multi-tasking considerations (Holmström and Milgrom 1991). Suppose, for instance, that in addition to making a decision, the agent has to put effort into two tasks, 'A' and 'B.' Suppose also that there is only one performance measure which depends on the agent's decision and the effort he puts into task A. The principal may then refrain from tying the agent's pay to the available performance measure since doing so might induce him to substitute effort from task B to task A (for an adverse selecion version of this multi-tasking argument see Armstrong 1995).

they consist of single intervals.[8] The main contribution of our paper is to provide a general characterization of the solution to Holmström's delegation problem when no restrictions are made on the set of feasible delegation sets and the principal's preferences take a generalized quadratic form.

Melumad and Shibano (1991) were the first to fully characterize the solution to the delegation problem in the absence of restrictions on the feasible delegation sets. They do so in a model with quadratic loss functions in which the preferred decisions are linear functions of the state and the state is uniformly distributed. We differ from Melumad and Shibano (1991) in that we allow for general distributions and more general utility functions. Allowing for more generality enables us to derive new insights into delegation. It also allows us to show which salient features of optimal delegation sets are robust to perturbations in the economic environment. Finally, it makes our results applicable to a large class of economic problems.

Another related paper is Green (1982) who considers a version of the delegation problem with general preferences and a finite number of states and feasible decisions. He shows that the search for the optimal decision rule reduces to a linear programming problem and that the principal may benefit from randomizing over decisions. In a setting with a continuum of states and decisions, Kovac and Mylovanov (2006) also show that the optimal mechanism may not be deterministic. We discuss stochastic mechanisms in Section 8.

Most of the literature that builds directly on Holmström (1977, 1984) focuses on characterizing the optimal delegation set. In contrast, Mylovanov (2006) concentrates on how to implement the optimal delegation set and provides conditions under which this can be done through veto delegation.

A number of recent papers have analyzed versions of the delegation problem in which the agent's bias does not vary with the state. Dessein (2002) considers a model in which a principal can either let an agent make any decision or engage in cheap talk communication and then make the decision herself.[9] Martimort and Semenov (2006b) provide a sufficient condition for interval delegation to be optimal. In contrast to these static models Alonso and Matouschek (2007) develop an infinitely repeated delegation game to endogenize the commitment power

[8] For a particular example he shows that a single interval is optimal among all compact delegation sets (see page 44 in Holmström 1977).

[9] See also Harris and Raviv (2005) and Marino and Matsusaka (2005).

of the principal. Ottaviani (2000) and Krishna and Morgan (2006) also analyze the basic delegation problem with a constant bias but, in contrast to the previous papers, allow for contingent monetary transfers.[10]

Variants of the delegation problem have also received attention in the political economy literature. A large literature uses this basic set up to analyze the delegation of public policy making from elected politicians to bureaucrats (Epstein and O'Halloran 1999; Huber and Shipan 2006) and from legislatures to their standing committees (Gilligan and Krehbiel 1987, 1989; Krishna and Morgan 2001). Martimort and Semenov (2006a) consider a delegation problem with multiple agents to analyze the organization of lobbying by interest groups.

Our paper is also related to the literature on cheap talk (Crawford and Sobel 1982). The key difference with this literature is that we allow the principal to commit to a decision rule while this is ruled out in cheap talk models. Finally, our model is related to the mechanism design approach to organizational structure (for an overview see Mookherjee 2006). In contrast to our paper, this literature assumes that the principal can commit to contingent transfers.

## 3 The Model

An organization that consists of a principal and an agent has to make a decision. The principal has the legal right to make the decision but only the agent has the information necessary to make the 'right' decision. The principal is unable to make contingent transfers and must decide on the decision making rules that the organization should adopt.

*Preferences:* The principal's and the agent's utilities depend on the implemented decision and on the state of the world. The decision is represented by $y \in Y$, where the set of admissible decisions $Y$ is a large compact interval of $\mathbb{R}$, and the state is denoted by $\theta \in \Theta = [\theta_1, \theta_2] \subset \mathbb{R}$. 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 $u_P(y, \theta) = -r(\theta)(y - y_P(\theta))^2$, where $y_P(\theta)$ is continuous in $\theta \in \Theta$ and $r(\theta)$ is a continuously differentiable and strictly positive function of the state $\theta$. The assumption that the principal's preferences are quadratic allows us to fully characterize the solution to this model. We discuss more general utility functions in Section 8. The agent has a von

[10] On delegation with transfers see also Filipi and Singh (2006).

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. Thus, given the state $\theta$, the principal's preferred decision is $y_P(\theta)$ and the agent's is $y_A(\theta)$. Below we often work with the inverse of the agent's preferred decisions which we denote by $\theta_A(y) \equiv y_A^{-1}(y)$. Note that the ranges of the principal's and the agent's preferred decisions, $Y_P = \{y \in Y : y_P(\theta) = y\}$ and $Y_A = \{y \in Y : y_A(\theta) = y\}$ respectively, are intervals of the form $Y_A = \left[\underline{d}_A, \overline{d}_A\right]$ and $Y_P = \left[\underline{d}_P, \overline{d}_P\right]$. Note also that the specified utility functions allow for an arbitrary continuous divergence of preferences $(y_A(\theta) - y_P(\theta))$. Moreover, they allow for variable degrees of risk aversion of the principal with respect to the decision $y$ for each realization of $\theta$, as characterized by the function $r(\theta)$.

*Information:* The agent knows the state $\theta$ but the principal does not. Her prior beliefs over its realization are represented by the cumulative distribution function $F(\theta)$. The corresponding probability density function $f(\theta)$ is absolutely continuous and strictly positive for all $\theta \in \Theta$.

*Contracts:* The principal can offer the agent any mechanism $(M, h)$, where $M$ is a message space and $h : M \to X$ is a decision rule that maps the messages into a set of allocations $X$, to be defined momentarily. Note that we restrict attention to deterministic mechanisms where, after receiving a message $m \in M$, the principal makes a particular decision $h(m)$ with certainty. Also, as discussed in the introduction, we rule out contingent transfers. We discuss stochastic mechanisms and contingent transfers in Section 8. Finally, 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 guarantee his participation.[11] The set of feasible mechanisms is therefore restricted to those in which the set of allocations $X$ is the set of admissible decisions $Y$.

*Timing:* The principal selects a mechanism $(M, h)$. The agent then observes the state and sends a message $m \in M$ to the principal who chooses a decision according to the decision rule $h$. Finally, payoffs are realized and the game ends.

*Transformations:* Without loss of generality we can restrict attention to the case in which the principal's preferences take the simple quadratic form $u_P(y,\theta) = -(y - y_P(\theta))^2$ and the agent's preferred decisions are a linear function of the state of the form $y_A(\theta) = \alpha + \beta\theta$, where

[11] This assumption is not crucial. See the working paper version of this paper (Section 8 in Alonso and Matouschek 2005).

$\beta > 0$. Transforming the model in this way simply requires making appropriate adjustments to the density and the principal's preferred decisions.[12] In what follows we work with this transformed version of the general model.

*Measuring the Preference Divergence:* We make use of three different measures of the agent's and the principal's preference divergence. First, for a given state $\theta$ the agent's *bias* $b(\theta) \equiv y_A(\theta) - y_P(\theta)$ measures the difference between the agent's and the principal's preferred decisions. Second, for a given state $\theta$ the *backward bias*

$$T(\theta) \equiv F(\theta) \left[ y_A(\theta) - \mathrm{E}\left[ y_P(z) \,| z \leq \theta \right] \right] \tag{1}$$

measures the difference between the agent's preferred decision in state $\theta$ and the principal's preferred decision if she believes that the state is smaller than $\theta$, weighted by the probability $F(\theta)$ that the state is indeed smaller than $\theta$. Similarly, for a given state the *forward bias* is given by

$$S(\theta) \equiv (1 - F(\theta)) \left[ y_A(\theta) - \mathrm{E}\left[ y_P(z) \,| z \geq \theta \right] \right] \tag{2}$$

and measures the difference between the agent's preferred decision in state $\theta$ and the principal's preferred decision if she believes that the state is bigger than $\theta$, weighted by the probability $(1 - F(\theta))$ that the state is indeed bigger than $\theta$. The backward bias $T(\theta)$ and the forward bias $S(\theta)$ are key in determining the solution to the contracting problem.

*Regulation Example:* It is useful to introduce a simple application that we will use throughout the analysis to illustrate and interpret our results. For this purpose, consider the regulation of an informed monopolist with a linear demand curve.[13] In particular, suppose that the 'principal' is a welfare maximizing regulator and the 'agent' is a profit maximizing monopolist. The monopolist can produce $q \geq 0$ units of a good at costs $\theta q$ and he faces a linear inverse demand function $y = A - Bq$, where $y$ is the price. Marginal costs $\theta$ are drawn from a unimodal distribution with support $[0, 1]$ and the maximum willingness to pay is higher than the highest cost, i.e. $A > 1$. Profits can then be expressed as a linear function of $(y - y_A(\theta))^2$, where

[12] Suppose that $u_P(y, \widehat{\theta}) = -r(\widehat{\theta})(y - y_P(\widehat{\theta}))^2$ and that $y_A(\widehat{\theta})$ is non-linear and let the density be denoted by $\widehat{f}(\widehat{\theta})$. First, to restrict attention to $u_P(y, \widehat{\theta}) = -(y - y_P(\widehat{\theta}))^2$, we simply need to redefine the density function as $f(\widehat{\theta}) = r(\widehat{\theta})\widehat{f}(\widehat{\theta}) / \int_0^1 r(t)\widehat{f}(t)\mathrm{d}t$. Second, to further restrict attention to a linear $y_A(\cdot)$, the state needs to be redefined as $\theta \equiv \left( y_A(\widehat{\theta}) - \alpha \right) / \beta$.

[13] We further generalize this application in Section 7.

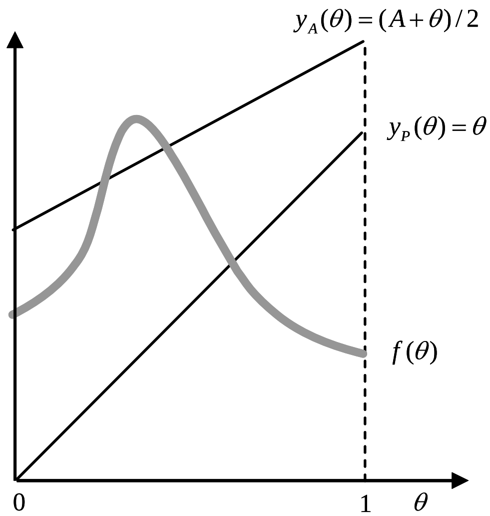


Figure 1: The Linear Regulation Model

$y_A(\theta) = (A+\theta)/2$ is the profit maximizing price, and welfare can be expressed as a linear function of $(y - y_P(\theta))^2$, where $y_P(\theta) = \theta$ is the price that maximizes welfare. This *linear regulation model* is illustrated in Figure 1. In the context of this model, the key question is what rules the regulator should adopt to determine the monopolist's price.

# 4 The Delegation Problem

The contracting problem that the principal solves is to choose a deterministic mechanism that maximizes her expected utility. In this section we show that this problem can be stated in two equivalent ways, as a direct mechanism design problem and as a *delegation problem* in which the principal offers the agent a compact set of decisions from which he then chooses his preferred one.

In our search for the optimal deterministic mechanism we can restrict attention to direct deterministic mechanisms in which the agent truthfully reports the state. This fact is reminiscent of the Revelation Principle and proven formally in the next lemma.

**Lemma 1.** *The principal's contracting problem can be stated as*

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

*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 into decisions.*

The outcome functions that satisfy the incentive compatibility constraint take a simple form, as shown in the next lemma. This lemma is the adaptation to our setting of Proposition 1 in Melumad and Shibano (1991). To be able to state it, let $X^-(\widehat{\theta}) \equiv \lim_{\theta\to\widehat{\theta}^-} X(\theta)$ and $X^+(\widehat{\theta}) \equiv \lim_{\theta\to\widehat{\theta}^+} X(\theta)$.[14]

**Lemma 2.** *An incentive compatible outcome function $X(\theta)$ must satisfy the following: (i.) $X(\theta)$ is weakly increasing, (ii.) if $X(\theta)$ is strictly increasing and continuous in $(\theta_1,\theta_2)$, 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: (iii.a.) $u_A(X^-(\theta),\widehat{\theta}) = u_A(X^+(\theta),\widehat{\theta})$, (iii.b.) $X(\theta) = X^-(\widehat{\theta})$ for $\theta \in [\max\{0,\theta_A(X^-(\widehat{\theta}))\},\widehat{\theta})$ and $X(\theta) = X^+(\widehat{\theta})$ for $\theta \in (\widehat{\theta},\min\{1,\theta_A(X^+(\widehat{\theta}))\}]$ and (iii.c.) $X(\widehat{\theta}) \in \{X^-(\widehat{\theta}), X^+(\widehat{\theta}))\}$.*

An illustration of this lemma is provided in Figure 2. It can be seen that the outcome function is weakly increasing and consists of flat segments as well as strictly increasing segments in which the agent's preferred decisions are 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, 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.

The direct mechanism design problem (3) is equivalent to the delegation problem in which the principal offers the agent a *delegation set*, i.e. a set of decisions from which he chooses his preferred one (Holmström 1977, 1984). Essentially, this is the case since the same decisions are implemented whether the principal commits to an outcome function $X(\theta)$ or lets the agent choose among the decisions $D = \{y : y = X(\theta), \theta \in \Theta\}$ that are in the range of $X(\theta)$. To state this result formally, let $X_D$ denote the set of outcome functions with range $D$, i.e. $X_D = \{X(\theta) : D \text{ is the range of } X(\theta)\}$.

[14] Note that $X^-(\widehat{\theta})$ and $X^+(\widehat{\theta})$ are well defined since any implementable $X(\cdot)$ is monotone.

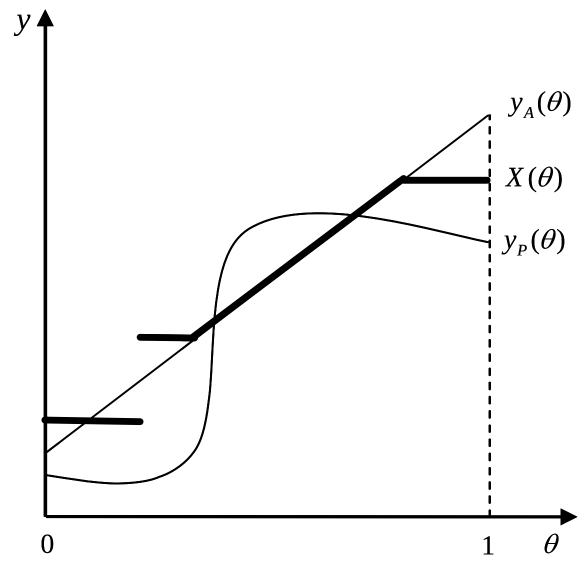


Figure 2: An Incentive Compatible Outcome Function

**Lemma 3.** *The principal's contracting problem can be stated as*

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

*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 $Y$.*[15]

It follows from Theorem 1 in Holmström (1984) that a solution to the delegation problem exists, as stated in the next lemma.

**Lemma 4.** *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 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. For the remainder of this paper, 'optimal delegation sets' refers to optimal minimal delegation sets unless otherwise stated.

[15] 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.

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 $D_2 \subset D_1$.

# 5 The Value of Delegation

When is delegation valuable? In other words, when does the principal benefit from letting the agent choose between at least two alternatives? In this section we address this question and show that delegation is valuable if and only if the principal and the agent are *minimally aligned*, a condition on the players' preferences and the agent's information that we formally define below. We also analyze how the value of delegation is affected by changes in the agent's preferences and his informational advantage. In the Section 6 we then investigate what decisions the principal should delegate when delegation is indeed valuable.

## 5.1 When Is Delegation Valuable?

A principal who does not delegate any decision rights simply implements her best uninformed decision $y_P^* \equiv \mathrm{E}(y_P(\theta))$. Delegation is therefore valuable if the principal can improve on implementing $y_P^*$ by letting the agent choose between at least two alternatives. To understand when this is the case, we first need to introduce the concept of a *minimally aligned* principal and agent. For this purpose consider Figure 3 which illustrates the principal's preferred decisions for different beliefs. In particular, $\mathrm{E}(y_P(z)\,|z \geq \theta\,)$ gives her preferred decision if she believes that the state is above $\theta$ while $\mathrm{E}(y_P(z)\,|z \leq \theta\,)$ gives her preferred decision if she believes that it is below $\theta$. The principal's best uninformed decision $y_P^*$ is a convex combination of $\mathrm{E}(y_P(z)\,|z \leq \theta\,)$ and $\mathrm{E}(y_P(z)\,|z \geq \theta\,)$ and therefore lies in-between these two curves. We say that the principal and the agent are *minimally aligned* if there exists a state $\theta^* \in (0,1)$ such that $\mathrm{E}(y_P(z)\,|z \leq \theta^*\,) < y_A(\theta^*) < \mathrm{E}(y_P(z)\,|z \geq \theta^*\,)$. Note that this is equivalent to requiring that there exists a state $\theta^* \in (0,1)$ such that $T(\theta^*) > 0$ and $S(\theta^*) < 0$, where $T(\cdot)$ and $S(\cdot)$ are, respectively, the backward bias defined in (1) and the forward bias defined in (2).

Delegation is valuable if the principal and the agent are minimally aligned. To see this, suppose that the principal and the agent are minimally aligned and let $\theta^* \in (0,1)$ denote a state such that $T(\theta^*) > 0$ and $S(\theta^*) < 0$. Suppose also that $y_A(\theta^*) > y_P^*$ and define

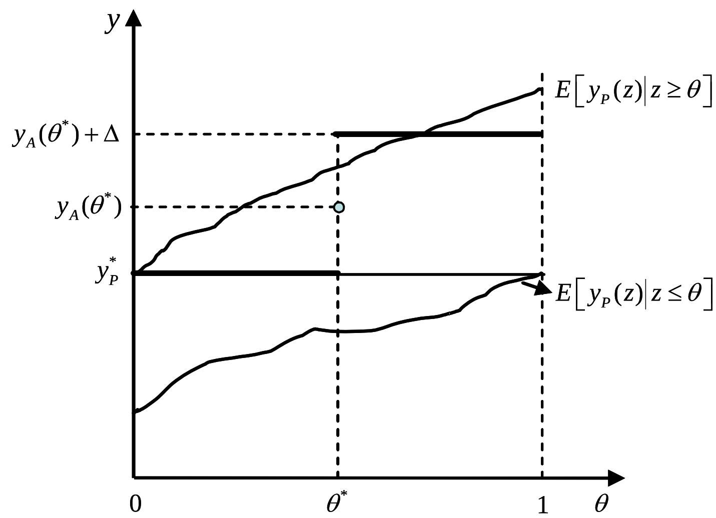


Figure 3: The Value of Delegation

$\Delta \equiv y_A(\theta^*) - y_P^*$.[16] The principal can then improve on implementing $y_P^*$ by letting the agent choose between $y_A(\theta^*) - \Delta = y_P^*$ and $y_A(\theta^*) + \Delta$. Faced with the choice between these two decisions the agent implements $y_P^*$ if $\theta \leq \theta^*$ and $y_A(\theta^*) + \Delta$ if $\theta > \theta^*$. The key observation, which is illustrated in Figure 3, is that $y_A(\theta^*) + \Delta$ is closer to the decision $\mathrm{E}(y_P(z)\,|z \geq \theta^*)$ that maximizes the principal's expected utility if $\theta > \theta^*$ than $y_P^*$ is. This of course implies that the principal prefers the delegation set $\{y_P^*, y_A(\theta^*) + \Delta\}$ to always making decision $y_P^*$.

Delegation is therefore valuable if the principal and the agent are minimally aligned. It turns out that the reverse is also true: if the principal and the agent are not minimally aligned, then delegation is not valuable. The following lemma is the key step in proving this result. To state the lemma, we define the *value of delegation* $V$ as the increase in the principal's expected utility when, instead of making $y_P^*$, she employs the optimal delegation set. Formally, $V \equiv \mathrm{E}_\theta[u_P(X_{D^*}(\theta), \theta)] - \mathrm{E}_\theta[u_P(y_P^*, \theta)]$, where $D^*$ solves the delegation problem (4).

**Lemma 5.** *Let $X = \{X(\theta) \in X_D, D \in N\}$ be the set of incentive compatible deterministic outcome functions. Then the value of delegation is given by*

$$V = \max_{X(\theta) \in X} -(y_P^* - X(1))^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) \qquad (5)$$

[16]The argument for $y_A(\theta^*) < y_P^*$ is analogous.

*or, equivalently,*

$$V = \max_{X(\theta)\in X} -(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta). \tag{6}$$

Since the first term on the right hand sides of (5) and (6) are weakly negative, it follows immediately that if, for all $\theta \in (0,1)$, $T(\theta) < 0$ and $S(\theta) > 0$, then $V = 0$. In other words, if the principal and the agent are not minimally aligned, then delegation is not valuable. We can therefore establish the following proposition.

**Proposition 1.** *Delegation is valuable if and only if the principal and the agent are minimally aligned, i.e. if and only if there exists a state $\theta \in (0,1)$ such that $T(\theta) > 0$ and $S(\theta) < 0$.*

The proposition shows that, in general, the principal's willingness to delegate depends on the players' preferences and the agent's information.[17] In the next sub-section we discuss how changes in the preferences and the information structure affect the principal's willingness to delegate. Before we do so, however, it is illustrative to apply our results to the linear regulation model that was introduced in Section 3.

*Regulation Example:* Should the regulator give the monopolist any discretion over its price? It is evident that the monopolist always wants to set a higher price than the regulator and thus $T(\theta) > 0$ for all $\theta \in (0,1)$. It is also straightforward to verify that $S(\theta) \geq 0$ for all $\theta \in (0,1)$ if and only if $A/2 \geq \mathrm{E}(\theta)$, where $A$ is the intercept of the demand curve. The following result then follows from Proposition 1.

**Result 1.** *Delegation is valuable if and only if $A/2 < \mathrm{E}(\theta)$.*

Note that the bias $b(\theta) = y_A(\theta) - y_P(\theta) = (A-\theta)/2$ is increasing in $A$. We therefore have the intuitive result that the regulator gives the monopolist some discretion over pricing if and only if the monopolist's preferences are sufficiently aligned. The effect of an increase in the monopolist's informational advantage, however, may at first be less intuitive. In particular, the result shows that mean preserving spreads have no effect on the regulator's willingness to delegate. Essentially, the regulator can benefit from delegation if and only if there exists a state in which the monopolist's preferred price $y_A(\theta)$ coincides with the principal's best

[17] We say that the principal is *willing to delegate* if and only if delegation is valuable, i.e. $V > 0$.

uninformed price $y_P^*$. Since $y_P^* = \mathrm{E}(\theta)$ this condition only depends on the mean and not on other properties of the cost distribution.

### 5.2 Comparative Statics on the Value of Delegation

How does the value of delegation depend on the agent's preferences and his informational advantage? The next corollary shows that, for a given information structure, the principal is more likely to delegate to a more aligned agent. To state the corollary, we introduce the concept of an agent whose preferences are *uniformly closer* to the principal's than those of another agent: suppose there are two agents, $A$ and $\overline{A}$, with preferred decisions $y_A(\cdot)$ and $y_{\overline{A}}(\cdot)$ respectively. We say that $\overline{A}$'s preferences are *uniformly closer* to the principal's than $A$'s if, for each $\theta \in \Theta$, $y_{\overline{A}}(\cdot)$ lies between $y_A(\cdot)$ and $y_P(\cdot)$.

**Corollary 1.** *Suppose that agent $\overline{A}$'s preferences are uniformly closer to the principal's than agent $A$'s. Then, if delegation to $A$ is valuable, delegation to $\overline{A}$ is also valuable.*

In general, little can be said about the effect of changes in the agent's informational advantage on the value of delegation. This is the case since such changes have an ambiguous effect on the backward and forward biases. However, we can still shed light on this comparative static with the help of examples. We have already seen that in the linear regulation model the agent's informational advantage has no effect on the principal's willingness to delegate. In Section 7 we show that this is also the case in another applications of our model. What these examples have in common is that the principal's preferred decisions are increasing in the state. For this case we can establish the following general result.

**Corollary 2.** *Suppose that $y_P(\theta)$ is weakly increasing for all $\theta \in (0,1)$. Then delegation is valuable if $y_P^* \in Y_A^\circ$.*

In this class of models the principal is therefore willing to delegate if her best uninformed decision lies in the range of the agent's preferred decisions. Changes in the information structure then only matter to the extent that they affect $y_P^* = \mathrm{E}(y_P(\theta))$. Once the principal has determined that she can benefit from delegation, she of course still needs to figure out exactly what decision rights she should delegate. We turn to this question next.

# 6 Optimal Delegation

What decisions should the principal delegate when delegation is valuable? In other words, what is the solution to the delegation problem (4)? The difficulty in solving this problem is the need to optimize over sets which precludes us from using standard optimization techniques. To characterize the solution, we 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. Roughly speaking, the next sub-section shows that the principal benefits from reducing the agent's discretion by removing intermediate decisions if the principal's preferred decisions are sufficiently steep relative to the agent's. We then use this insight to provide a general characterization of optimal delegation sets in Section 6.2 and to provide optimality conditions for interval delegation in Section 6.3. Finally we conclude this section by discussing comparative statics on the agent's discretion in Section 6.4.

## 6.1 Changing the Agent's Discretion

We first analyze the effects of adding a single intermediate decision to, and removing such a decision from, a delegation set. After having considered such a discrete change in the agent's discretion, we then analyze the effects of adding and removing decision intervals.

*Adding and Removing Discrete Decisions:* Consider a delegation set $D^+$ that contains three consecutive decisions $y_1 < y_2 < y_3$, all of which are within the range of the agent's preferred decisions $Y_A$. Removing the intermediate decision $y_2$ from $D^+$ yields a delegation set $D^- \equiv D^+ \setminus y_2$ that gives the agent strictly less discretion. To compare the decision making under the two delegation sets, consider Figure 4 in which the solid step function $X_{D^+}(\theta)$ describes the agent's decision making under $D^+$ and the dashed step function $X_{D^-}(\theta)$ describes his decision making under $D^-$. Recall that $\theta_A(y) \equiv y_A^{-1}(y)$ and note that the discontinuities occur at the states $r \equiv \theta_A((y_1+y_2)/2)$, $s \equiv \theta_A((y_1+y_3)/2)$ and $t \equiv \theta_A((y_2+y_3)/2)$ at which the agent is indifferent between decisions $y_1$ and $y_2$, $y_1$ and $y_3$, and $y_2$ and $y_3$ respectively. It can be seen that decision making only differs between the two delegation sets for the intermediate states $\theta \in [r, t]$. In particular, removing the intermediate decision forces the agent to make more extreme decisions: instead of implementing the intermediate decision for all intermediate states, he implements the low decision for sufficiently low states and the high decision for

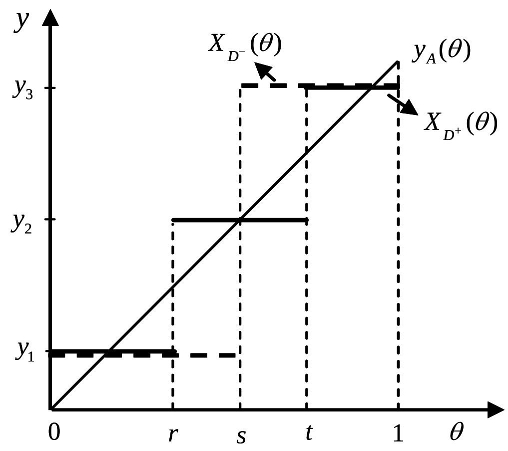


Figure 4: The Effect of a Change in Discretion on Decision Making

sufficiently high states. Which delegation set the principal prefers therefore depends simply on whether, over the relevant interval, the principal's preferred decisions are better approximated by a flat function or by a step function. This suggests that if the principal's preferred decisions are sufficiently steep relative to the agent's, then she prefers to reduce the agent's discretion. The formal analysis to which we turn next makes this intuition precise.

Let $\Delta U \equiv \mathrm{E}\left(u_P(y,\theta)\,|D^+\right) - \mathrm{E}\left(u_P(y,\theta)\,|D^-\right)$. Using the definitions of $r$, $s$ and $t$ introduced above, it follows that $\Delta U = -2\left[(y_3 - y_1)\,T(s) - (y_2 - y_1)T(r) - (y_3 - y_2)T(t)\right]$. We can further simplify $\Delta U$ 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 $\lambda \in (0,1)$. Substituting into the above expression for $\Delta U$ gives $\Delta U = -2\,(y_3 - y_1)\left[T(s) - \lambda T(r) - (1-\lambda)T(t)\right]$. Since $s = \lambda r + (1-\lambda)t$, it follows that if the backward bias $T(\theta)$ is strictly concave over the interval $\left[\theta_A\left(y_1\right), \theta_A\left(y_3\right)\right]$ then $\Delta U < 0$. Note that this sufficient condition is independent of $y_2$. Thus, under this condition, the principal strictly benefits from removing *any* $y_2 \in (y_1, y_3)$ from the delegation set $D^+$. Similarly, if the backward bias $T(\theta)$ is strictly convex over the interval $\left[\theta_A\left(y_1\right), \theta_A\left(y_3\right)\right]$ then $\Delta U > 0$ and if it is linear then $\Delta U = 0$.

The agent's discretion therefore depends crucially on the curvature of the backward bias. Before turning to the intuition for this condition, we summarize the above analysis in the following lemma.

**Lemma 6.** *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* $D^- \equiv D^+ \setminus y_2$ *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 backward bias* $T(\theta)$ *is strictly concave over the interval* $[\theta_A(y_1), \theta_A(y_3)]$ *and it does not change the principal's expected utility if* $T(\theta)$ *is linear over the interval* $[\theta_A(y_1), \theta_A(y_3)]$.

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

We can now relate this lemma to our previous intuition. The lemma shows that the principal prefers to reduce the agent's discretion if the backward bias is convex or, equivalently, if $b'(\theta) < -[\beta + b(\theta)f'(\theta)/f(\theta)]$, where $\beta = y'_A(\theta)$. Thus, as anticipated, the principal gives the agent less discretion if her preferred decisions are sufficiently steep relative to the agent's. Note that she may prefer to restrict the agent's discretion even if his bias is close to zero. Essentially, when the principal's preferred decisions are relatively steep, she wants to force the agent's decision making to be more sensitive to the state and she can ensure this, albeit in a coarse manner, by removing the intermediate decision. Moreover, she may want to do so even if the agent is locally very aligned.

Exactly how steep $y_P(\theta)$ has to be depends on the agent's private information. To focus on the agent's information, suppose that $b(\theta) = b > 0$, in which case the backward bias is convex if $f'(\theta)/f(\theta) < -\beta/b$. Thus, the principal can only benefit from a reduction in the agent's discretion if the density is decreasing. To understand this, note that if the agent has a positive bias $b > 0$, the principal benefits from reducing the agent's discretion in low states but not in high states. A decreasing density then ensures that low states are sufficiently more likely than high states so that the principal benefits overall. In Figure 5, for instance, removing decision $y_2$ makes the principal better off in states $\theta \in [r, s]$ but it makes her worse off in states $\theta \in [s, t]$. Overall the principal then benefits if the probability of being in interval $[r, s]$ is sufficiently high relative to the probability of being in the interval $[s, t]$.

*Adding and Removing Decision Intervals:* We have just seen that if $T(\theta)$ is convex over the interval $[\theta_A(y_1), \theta_A(y_3)]$, then the principal benefits from adding any decision $y_2 \in (y_1, y_3)$ to

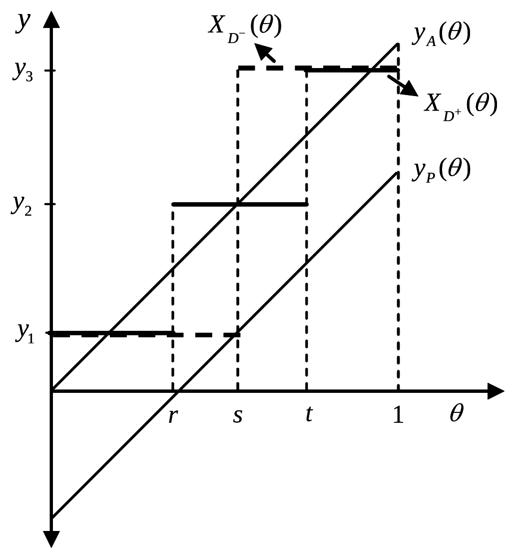


Figure 5: The Costs and Benefits of Removing a Decision

the delegation set $D^-$. This suggests that the principal would benefit even more from adding any number of decisions or, in fact, a continuum $(y_1, y_3)$ of decisions, to the delegation set. The following lemma considers the effect of adding decision intervals to, and removing such intervals from, a delegation set and shows that this is indeed the case.

**Lemma 7.** *Let $D^+$ be a delegation set that contains an interval $[y_1, y_3] \subset Y_A$ and let $D^- \equiv D^+ \setminus (y_1, y_3)$. Then*

*(i.)* *removing decisions $(y_1, y_3)$ from $D^+$ strictly increases the principal's expected utility if $T(\theta)$ is strictly concave over the interval $[\theta_A(y_1), \theta_A(y_3)]$ and it does not change the principal's expected utility if $T(\theta)$ is linear over the interval $[\theta_A(y_1), \theta_A(y_3)]$.*

*(ii.)* *adding decisions $(y_1, y_3)$ to $D^-$ strictly increases the principal's expected utility if $T(\theta)$ is strictly convex over the interval $[\theta_A(y_1), \theta_A(y_3)]$ and it does not change the principal's expected utility if $T(\theta)$ is linear over the interval $[\theta_A(y_1), \theta_A(y_3)]$.*

Having derived these lemmas, we can now turn to the analysis of optimal delegation sets.

## 6.2 Characterizing the Optimal Delegation Set

In this section we build on the above analysis to characterize the optimal delegation set. Essentially, we define a partition $\{y_1, ..., y_i, ...\}$ of $Y$ such that for each interval $[y_i, y_{i+1}]$, the backward bias $T(\theta)$ is either convex, concave or linear for all $\theta \in [\theta_A(y_i), \theta_A(y_{i+1})]$. We then use Lemmas 6 and 7 to characterize the number of decisions that the optimal delegation set can contain in each interval. For instance, if $T(\theta)$ is concave for all $\theta \in [\theta_A(y_i), \theta_A(y_{i+1})]$, the optimal delegation set can contain at most two decisions within $[y_i, y_{i+1}]$. The optimal delegation set can then be obtained by comparing the principal's expected utility under the different possible combinations of decisions in the different intervals.

The following proposition states the main characterization result. To understand the proposition, recall that the ranges of the principal's and the agent's preferred decisions are denoted by $Y_P = [\underline{d}_P, \overline{d}_P]$ and $Y_A = [\underline{d}_A, \overline{d}_A]$ respectively.

**Proposition 2.** *Let $D^*$ be an optimal delegation set and let $y_1, y_2 \in Y_A \cap Y_P$. Then,*

*(i.)* *if $T(\theta)$ is strictly convex for all $\theta \in [\theta_A(y_1), \theta_A(y_2)]$, 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.*

*(ii.)* *if $T(\theta)$ is strictly concave for all $\theta \in [\theta_A(y_1), \theta_A(y_2)]$, then $D^* \cap [y_1, y_2]$ contains at most two decisions.*

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

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

As anticipated, the first three parts of the proposition characterize the number of decisions that the optimal delegation set can contain within intervals for which the backward bias is either convex, concave or linear. Together, these parts of the proposition characterize the optimal delegation set in the region of the decision space in which the ranges of the principal's and the agent's preferred decisions intersect, i.e. in $Y_A \cap Y_P$. Part (iv.) completes the characterization by showing that the optimal delegation set can contain at most one decision above and one decision below $Y_A \cap Y_P$.

The characterization result enables us to generically reduce the delegation problem to a finite dimensional problem that can be solved with standard techniques.[18] As an illustration

[18] Specifically, we can reduce the delegation problem to a finite dimensional problem as long as $T'(\theta)$ has a finite number of extrema.

of how this result can be applied, we now return to the linear regulation model.

*Regulation Example:* In this model $T''(\theta) = \frac{1}{2}(A - \theta) f'(\theta)$. The backward bias is therefore convex if $f'(\theta) > 0$ and it is concave if $f'(\theta) < 0$. Part (i.) of Proposition 2 then implies that the optimal delegation set contains either no price, one price or a price interval within $[A/2, \min[(A + \theta_m)/2, 1]]$, where $\theta_m$ is the mode of the distribution, and part (ii.) implies that it contains at most two prices within $[\min[(A + \theta_m)/2, 1], 1]$. Moreover, part (iv.) implies that the optimal delegation set contains at most one price below $y = A/2$ and one above $y = 1$. The next result shows which one of the different possible combinations is optimal.

**Result 2.** *Suppose that delegation is valuable, i.e.* $A/2 < \mathrm{E}(\theta)$. *Then,* $D^* = \left[A/2, \left(A + \overline{\theta}\right)/2\right]$, *where* $\overline{\theta} \in (0, 1)$ *solves* $y_A(\overline{\theta}) = \mathrm{E}(\theta \,|\theta \geq \overline{\theta})$. *Moreover, the regulator gives the monopolist more discretion, the more aligned the monopolist's preferences, i.e. the smaller* $A$.

In the case of linear demand curves, regulation without transfers therefore often takes a remarkably simple form. In particular, for any unimodal distribution, the regulator cannot do better than to engage in price cap regulation. Essentially, faced with a monopolist who prefers higher prices than herself, the regulator simply imposes a price cap and lets the monopolist choose any price below this threshold. The result also confirms the standard intuition that an agent gets more discretion the more aligned he is with the principal. In Section 6.4, however, we will see that this comparative static does not hold in general.

## 6.3 Interval Delegation

In practice organizations often allow their agents to make any decision from a single interval, that is, they engage in interval delegation (Holmström 1977, 1984). To understand why this may be so, we now investigate when interval delegation is optimal.

Formally, the principal engages in *interval delegation* if the optimal delegation set consists of a single, non-degenerate interval $[\underline{y}, \overline{y}]$, where $\underline{y} < \overline{y}$. A particular type of interval delegation is *threshold delegation*, in which case the interval lies strictly within the range of the agent, i.e. $\underline{y} > \min Y_A$ and $\overline{y} < \max Y_A$. The next proposition provides conditions for threshold delegation to be optimal.[19]

[19] The conditions under which other types of interval delegation, such as complete delegation, are optimal are closely related to those stated in Proposition 3 and we therefore relegate them to the appendix (see Propositions 6 and 7).

**Proposition 3.** *Threshold delegation is optimal if and only if there exist two states* $\underline{\theta}, \overline{\theta} \in (0,1)$ *and* $\overline{\theta} > \underline{\theta}$ *such that (i.)* $S(\overline{\theta}) = 0$ *and* $S(\theta) \geq 0$ *for* $\theta > \overline{\theta}$*, (ii.)* $T(\underline{\theta}) = 0$ *and* $T(\theta) \leq 0$ *for* $\theta < \underline{\theta}$ *and (iii.)* $T(\theta)$ *is convex for all* $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$*.*

For an intuition, consider first condition (i.). By requiring that the forward bias is weakly positive for all $\theta \geq \overline{\theta}$, this condition ensures that for any $\widehat{\theta} \geq \overline{\theta}$ 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 any $\widehat{\theta} \leq \underline{\theta}$ 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 backward bias is convex between $\underline{\theta}$ and $\overline{\theta}$, condition (iii.) ensures that it is optimal for the principal to include all decisions between the two thresholds in the delegation set.

To understand why interval delegation may be widespread in organizations, consider the next proposition.

**Proposition 4.** *Suppose the preferred decisions* $y_P(\theta)$ *and* $y_A(\theta)$ *are strictly increasing and twice continuously differentiable 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 the principal 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* $\lambda \geq \overline{\lambda}$*.*

The proposition shows that interval delegation tends to be optimal when the agent's preferences are sufficiently aligned with those of the principal. This suggests that the apparent widespread use of interval delegation in organizations may be due to their ability to carefully screen their agents or to employ incentive schemes that sufficiently align their interests.

## 6.4 Comparative Statics on the Agent's Discretion

It seems intuitive that a principal gives more discretion to a less biased agent, that is, to an agent whose preferences are uniformly closer to hers. In fact, Holmström (1977, 1984) shows that this comparative static holds under general conditions if delegation sets are required to

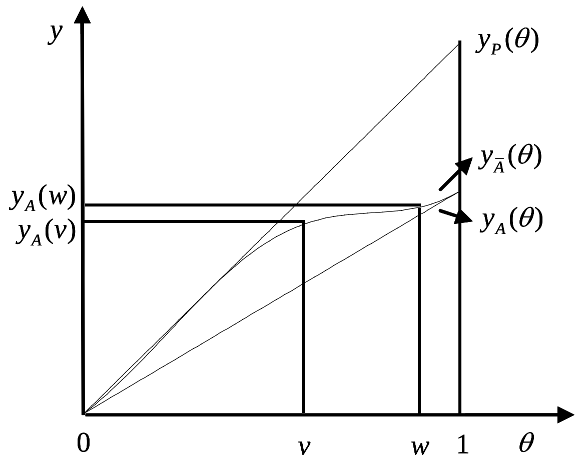


Figure 6: A Less Biased Agent May Have Less Discretion

take the form of a single interval. It turns out, however, that it does not hold if the principal's ability to delegate is unrestricted. In particular, this can be the case if the less biased agent is locally very unresponsive to changes in the state. Suppose, for instance, that the principal's preferred decisions are given by $y_P(\theta) = \theta$ and that $\theta$ is uniformly distributed on $[0, 1]$. Suppose also that there is an agent $A$ with preferred decisions $y_A(\theta) = \beta\theta$, where $\beta \in (1/2, 17/20)$. The principal cannot do better than to let this agent make any decision. Consider now another agent, agent $\overline{A}$, with the preferred decisions $y_{\overline{A}}(\theta) = \frac{1+\beta}{2}\theta\left(1 + \frac{1-\beta}{1+\beta}\sin\frac{3}{2}\pi\theta\right)$ which are illustrated in Figure 6. It can be seen that while this agent is uniformly closer to the principal than agent $A$ is, he is also much less responsive to changes in the state if $\theta \in [v, w]$. To encourage more state-sensitive decision making, the principal restricts agent $\overline{A}$'s discretion by ruling out the intermediate decisions $(y_A(v), y_A(w))$.[20] This confirms our previous insight that the agent's discretion depends not only on the level of his bias but also, and importantly, on its slope.

It may also seem intuitive that a principal gives more discretion to an agent with a bigger informational advantage. However, in general, this is not the case either.[21] Technically this

[20] The fact that agent $A$ is allowed to make any decision follows from Proposition 7. The result that agent $\overline{A}$'s gets strictly less discretion follows from the fact that if $\beta \in (1/2, 17/20)$ then there exist $v, w \in (0, 1)$ such that the backward bias is concave.

[21] Holmström (1977, 1984) shows that changes in the agent's informational advantage have an ambiguous effect on optimal interval delegation sets.

is so since changes in the agent's informational advantage have an ambiguous effect on the backward and forward biases. Situations in which the principal gives less discretion to an agent with a bigger informational advantage arise naturally in settings in which it is optimal for the principal to encourage more state-sensitive decision making by ruling out intermediate decisions. It is easy to show that ruling out such decisions may be more costly for the principal the bigger the agent's informational advantage, and that, as a result, she may rule out fewer decisions. For an illustration we refer to Section 7, where we provide a simple example in which the effect of changes in the agent's informational advantage on his discretion depend only on the slope of his bias.

# 7 Applications

Our analysis can be used to gain new insights into well-known economic problems. We have already shown that our framework can be applied to a version of the classic regulation problem in Baron and Myerson (1982) in which transfers are ruled out by law. In the linear regulation model, we restricted attention to linear demand curves and showed that for any unimodal distribution the regulator cannot do better than to engage in price cap regulation. This result can also been shown for constant elasticity demand curves (see Appendix B).[22] Thus, for a large class of distributions and for common demand functions, optimal regulation without transfers takes a remarkably simple form. Moreover, it is widely observed in practice.[23]

As discussed in the introduction, our analysis can also be used to gain new insights into the design of legislative rules. In the remainder of this section we apply our analysis to a version of the standard model in the literature and show that the optimal legislative rules are similar to, but different from, the rules that the literature has focused on.[24]

*The Model:* A legislature and a committee care about an outcome $x \in \mathbb{R}$. The utility of the legislature is given by $u_P(\cdot) = -x^2$ and that of the committee is given by $u_A(\cdot) = -(x-b)^2$, where $b \in \mathbb{R}$. The outcome $x$ is determined by an outcome function $x(y, \theta) = y/\theta - 1$ that

[22] In the case of constant elasticity demand curves, profits and welfare are no longer quadratic functions of the price. However, we can reasonably approximate them as quadratic functions using Taylor series approximations.

[23] For the use of price cap regulation in practice see, for instance, Armstrong and Sappington (2004).

[24] We follow the strand of the literature that focuses on homogenous committees and on information transmission rather than the incentives to acquire information.

depends on the policy $y \in Y \subset \mathbb{R}$ and the state $\theta \in [-a, a]$, where $a > 0$.[25] The committee knows the state $\theta$ but the legislature does not. The prior of the legislature is given by a truncated normal distribution with zero mean and support $[-a, a]$.[26] The legislature can employ any legislative rule, that is, it can commit to any deterministic decision rule. Among the many rules it could commit to is the *open rule* under which the committee proposes a bill $\widehat{y} \in Y$ and the legislature is free to choose any policy $y \in Y$. Another example is the *closed rule* under which the committee proposes a bill $\widehat{y} \in Y$ and the legislature can either choose $\widehat{y}$ or an exogenously given default policy $y_0 \in Y$. Most of the existing literature has focused on the relative performance of open and closed rules. In contrast, we are interested in the optimal among all possible rules. The timing is as follows: first the legislature chooses the legislative rule, then the committee proposes a bill and finally the legislature chooses a policy.

*Optimal Legislative Rules:* In a given state $\theta$ the legislature's preferred policy is $y_P(\theta) = \theta$ and the committee's is $y_A(\theta) = \beta\theta$, where $\beta \equiv (1 + b)$. To ensure that $y_A(\theta)$ is increasing we now assume that $\beta > 0$. Moreover, to avoid having to discuss a large number of different cases, we assume that the support is 'sufficiently large,' in the sense that $\beta a \geq 2\sigma^2 [f(0) - f(a)]$. The model is illustrated in Figure 7. The first result shows that delegation is always valuable, independent of the committee's informational advantage.

**Result 3.** *Delegation is always valuable.*

Essentially, the legislature and the committee agree that a positive policy should be implemented when the state is positive and a negative policy should be implemented when the state is negative. The legislature can then always elicit the sign of the state by offering a binary delegation set $\{-\overline{y}, \overline{y}\}$, where $\overline{y} = \mathrm{E}[\theta \,|\theta \geq 0]$.

To investigate which decisions the legislature should delegate, it is useful to distinguish between three types of committees: *responsive committees* for which $\beta > 1$, *moderate committees* for which $1 \geq \beta \geq 1/2$, and *unresponsive committees* for which $1/2 > \beta > 0$. The next proposition focuses on responsive committees.

[25] In the standard model in Gilligan and Krehbiel (1987) and Krishna and Morgan (2001) the outcome function is assumed to be linear, i.e. $x(y, \theta) = y - \theta$. We focus on multiplicative outcome functions since they yield additional insights into delegation.

[26] In the standard model in Gilligan and Krehbiel (1987) and Krishna and Morgan (2001) it is assumed that the state is uniformly distributed. The normal distribution allows us to perform comparative statics on the agent's informational advantage in a straightforward manner.

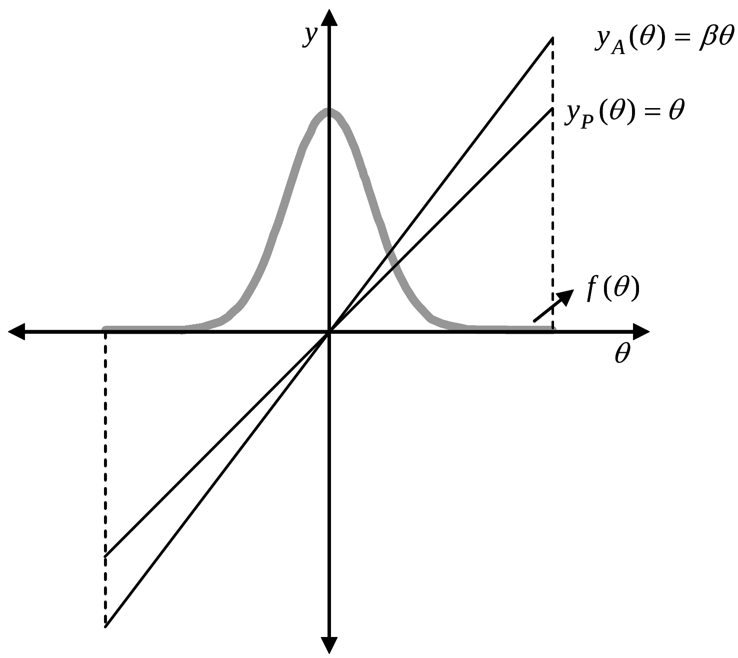


Figure 7: Decision Making with Multiplicative Outcome Functions

**Result 4.** *Suppose the committee is responsive, i.e. $\beta > 1$. Then threshold delegation is optimal and the optimal delegation set takes the form $[-\overline{y}, \overline{y}]$, where $\overline{y} = \mathrm{E}\left[y_P(\theta)\,|\theta \geq \theta_A(\overline{y})\right]$. The committee's discretion is increasing in its informational advantage and decreasing in its bias.*

A responsive committee always prefers more extreme policies than the legislature, i.e. $|y_A(\theta)| > |y_P(\theta)|$. Moreover, the discrepancy between the preferred policies $|y_A(\theta) - y_P(\theta)|$ is increasing in $|\theta|$. It is then intuitive that the legislature restricts the committee's discretion by imposing an upper and a lower bound on the delegation set. Moreover, since the committee's preferred decisions are sufficiently steep relative to the legislature's, the committee is allowed to make any decision within these bounds.

Moderate and unresponsive committees always prefer less extreme policies than the legislature. For this reason the legislature does not impose a binding upper or lower threshold on the delegation set. Indeed, in the case of a moderate committee, the legislature does not put any restrictions on the committee, as shown in the next result.

**Result 5.** *Suppose the committee is moderate, i.e. $1 \geq \beta \geq 1/2$. Then it is optimal to let the* *committee make any decision.*

In contrast, in the case of an unresponsive committee the legislature does limit the committee's discretion, as shown next.[27]

**Result 6.** *Suppose the committee is unresponsive, i.e.* $1/2 > \beta > 0$. *Then the optimal delegation set contains all policies* $Y$ *except those in an interval* $(-\overline{y}, \overline{y})$, *where* $\overline{y} = \mathrm{E}\left[y_P(\theta)\,|0 \leq \theta \leq \theta_A(\overline{y})\right]$. *The committee's discretion is decreasing in its informational advantage and its bias.*

To understand why the legislature bans the policies $(-\overline{y}, \overline{y})$, note that the preferred policies of an unresponsive committee are, by definition, much less sensitive to changes in the state than the legislature's. Essentially, by removing the intermediate decisions $(-\overline{y}, \overline{y})$ the legislature forces the committee's decision making to be more state-sensitive than it would be under complete delegation. Figure 8 shows that doing so hurts the legislature in the states $[-v, v]$ but makes it better off in states $[-w, -v] \cup [v, w]$. An increase in $\sigma^2$ makes it relatively less likely that the state lies in $[-v, v]$ than in $[-w, -v] \cup [v, w]$. As a result, the legislature bans more decisions when the informational advantage of an unresponsive committee increases.

The open rule cannot implement the optimal delegation set for any type of committee. This is so since optimal delegation requires a continuum of decisions to be implemented which is not possible in a cheap talk equilibrium. Similarly, the closed rule can only implement the optimal delegation set in the knife-edge cases.[28] However, the optimal delegation set can be implemented through a *modified closed rule* which is identical to the closed rule but allows the legislature to set the appropriate default policy (see Theorem 1 in Mylovanov 2006).

# 8 Revisiting Key Assumptions

To conclude the formal analysis we now return to three key assumptions and discuss the implications of relaxing them.

[27] Note that the optimal delegation set described in this result is not minimal, i.e. it includes decisions that are never selected in equilibrium. We focus on this optimal delegation set to ensure that the comparative statics are unambiguous.

[28] In particular, the closed rule can implement the optimal delegation set for responsive and for unresponsive committee's only if the default policy happens to be equal to either $\overline{y}$ or $-\overline{y}$. In the case of a moderate committee, the default policy has to be sufficiently unattractive to the legislature to ensure that it never wants to implement it.

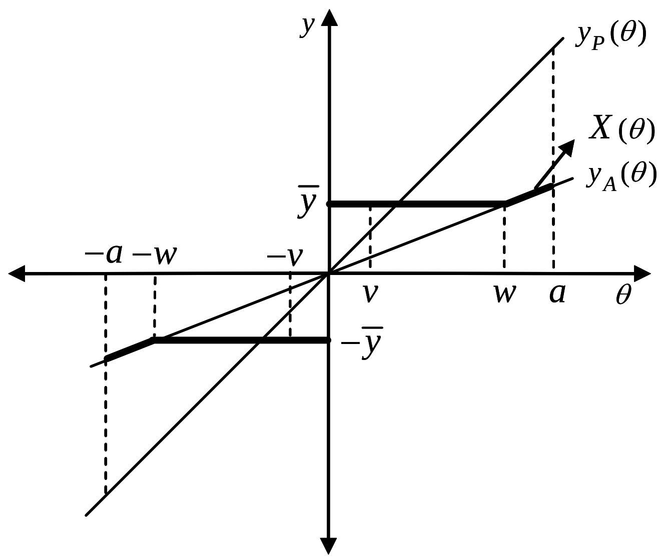


Figure 8: Decision Making with an Unresponsive Committee

### 8.1 Contingent Transfers

It is evident that in our setting the principal could always do weakly better by committing to contingent transfers. To get a sense for how much better she would be able to do, we now compare the performances of optimal delegation and complete contracting in the leading example of Crawford and Sobel (1982). We leave a general analysis of the returns from contracting for future research.[29]

In the leading example of Crawford and Sobel (1982), $\theta$ is uniformly distributed on $[0,1]$, $u_P(y,\theta) = -(y-\theta)^2$ and $u_A(y,\theta) = -(y-\theta-b)^2$, where $b > 0$. The optimal delegation set then takes the form $D^* = [b, 1-b]$ if $b \leq 1/2$ and $D^* = \{1/2\}$ if $b > 1/2$. To characterize the complete contract, we draw on Krishna and Morgan (2006) who provide such a characterization under the assumption that the agent is liquidity constrained.[30] We consider two 'no contract' benchmarks: under *principal control* the principal is free to implement any decision without interacting with the agent and thus implements $y_P^* = 1/2$. Similarly, under *agent control* the agent is free to make any decision and therefore implements $y_A(\theta) = \theta + b$.

Figure 9 illustrates the percentage increase in the principal's expected utility of moving

[29] See also Krishna and Morgan (2006).

[30] In particular, they require transfers to the agent to be positive.

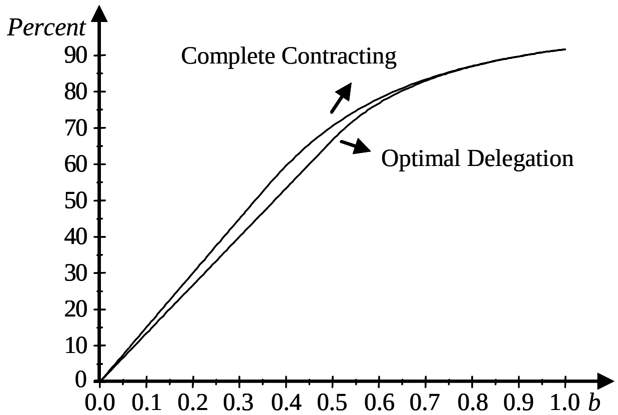


Figure 9: The Returns from Contracting: Agent Control

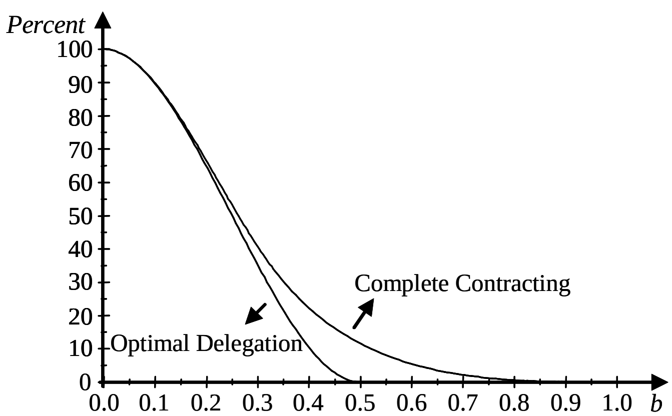


Figure 10: The Returns from Contracting: Principal Control

from agent control to optimal delegation and complete contracting respectively. The figure shows that there are strongly decreasing returns from contracting. In particular, when the difference in the relative performances is the greatest, namely at $b \simeq 0.4$, optimal delegation increases the principal's expected utility by about 50% while complete contracting increases it by about 56%. Moreover, for small and for large biases, the performances of the two types of contracts are almost indistinguishable. Figure 10 shows that this is also true if the benchmark is principal control. While in this case the maximum difference in the relative performances is larger, it still does not surpass 15 percentage points.

In some cases, therefore, the benefits of using contingent transfers, in addition to restricting the agent's discretion, are quite limited. It should also be noted that optimal transfer rules tend to be highly non-linear and sensitive to changes in the informational structure. In contrast, in some cases, such as the regulation model, optimal delegation takes a very simple form that is robust to changes in the informational structure. To the extent that deriving and implementing sophisticated transfer rules requires resources, this suggests that it may sometimes be more economical to simply restrict an agent's discretion than to design a potentially complicated complete contract.

### 8.2 General Utility Functions

In Section 6 we characterized the solution to the delegation problem (4) by analyzing the effects of adding decisions to, and removing them from, a delegation set. When the principal's utility function is quadratic, this optimization technique delivers very simple optimality conditions. Also, for such preferences the principal benefits from removing any number of intermediate decisions whenever the sufficient condition for removing a single intermediate decision is satisfied. These features greatly simplify the characterization of the optimal delegation set. The key problem of allowing for more general preferences is that, in general, these features do not hold for non-quadratic preferences. It should be noted, however, that while the analysis is more complicated, in principle, the same optimization technique can be used to obtain a characterization of the optimal delegation set when the principal's utility function takes a more general form.

Suppose, for instance, that the principal's utility $u_P(y, \theta)$ is strictly concave in $y$ and that $y_A(\theta) = \theta$. The rest of the model is as in Section 3. Using the same approach as we did in the analysis above, we can then establish a sufficient condition for interval delegation to be optimal.

**Proposition 5.** *Let* $R(\theta) \equiv \frac{\partial u_P(\theta,\theta)}{\partial y} f(\theta)$ *and* $y_P(\theta) \equiv \arg\max_y u_P(y, \theta)$.. *If* $y_P(\theta)$ *is (weakly) increasing and* $R(\theta)$ *is (weakly) decreasing in* $[0, 1]$ *and there exists* $\overline{\theta} \in (0, 1)$ *such that* $R(\overline{\theta}) = 0$ *then interval delegation is optimal.*

To understand how this condition relates to the sufficient conditions derived above, note that if the principal's preferences are quadratic, then $R(\theta) = -2b(\theta)f(\theta)$ and $T'(\theta) = F(\theta) - R(\theta)/2$. Thus, if $R(\theta)$ is weakly decreasing, then the backward bias $T(\theta)$ is convex.

### 8.3 Stochastic Mechanisms

Our model follows most of the literature on mechanism design in restricting attention to deterministic mechanisms. In general, this is not without loss of generality. In a recent paper, however, Kovac and Mylovanov (2006) consider a version of our model with quadratic preferences and provide a sufficient condition for the optimal mechanism to be deterministic. When their condition is satisfied, the backward bias is everywhere convex and thus interval delegation is optimal. In contrast, when interval delegation is not optimal, it is possible to construct examples in which the optimal mechanism is stochastic.

Suppose, for instance, that $\theta$ is distributed on $[-1, 1]$ according to a symmetric probability density function $f(\theta)$, $u_A(y, \theta) = -(y - \beta\theta)^2$ and that $u_P(y, \theta) = -(y - \theta)^2$. Further suppose that $\mathrm{E}[\theta \,|\theta \geq 0] > 2\beta$. The optimal delegation set then consists of only two decisions, $-\overline{y}$ and $\overline{y}$, where $\overline{y} = \mathrm{E}[\theta \,|\theta \geq 0]$. Although the agent's bias is very small for $\theta$ close to zero, the principal rules out the intermediate decisions $(-\overline{y}, \overline{y})$ to encourage more state-sensitive decision making. The principal can improve on this deterministic mechanism by letting the agent choose between $\overline{y}$, $-\overline{y}$ and a zero-mean lottery over two decisions $-z$ and $z$, where $z < \overline{y}$. The agent will then prefer the lottery to either $\overline{y}$ or $-\overline{y}$ whenever $\theta$ is close to zero.[31]

We leave a full characterization of optimal stochastic mechanisms for future research. It should be noted, however, that even if a stochastic mechanism is optimal, enforcing such a mechanism may be impossible.[32]

## 9 Conclusions

Rules are a pervasive feature of organizations. In this paper we studied the optimal design of decision rules in organizations in which there is a misalignment in the interests between the managers who have the legal right to make decisions and those who possess the relevant information. Our focus has been on situations in which it is impossible to commit to contingent monetary transfer schemes. We believe that such situations are widespread in practice and

---

[31] In particular, the agent chooses the lottery for $|\theta| \leq \overline{\theta} \equiv \left(\overline{y}^2 - z^2\right) / (2\beta\overline{y})$. The increment in the principal's expected utility when offering this stochastic mechanism is given by $\Delta U = 4\overline{y}\left(F(\overline{\theta}) - 1/2\right)\left(\beta\overline{\theta} - \mathrm{E}\left[\theta \left|0 \leq \theta \leq \overline{\theta}\right.\right]\right)$. For a fixed $\beta$, there exist distributions with sufficient probability mass at zero such that $\Delta U > 0$.

[32] Enforcement is a standard concern with stochastic mechanisms. See, for instance, page 67 in Laffont and Martimort (2002).

not yet sufficiently well understood in theory.

The central result in the paper is the characterization of optimal decision rules. In general, such rules can take many different forms. In practice, however, decision rules are often very simple and merely specify an interval of decisions from which the agent is allowed to choose his preferred one. We showed that such simple decision rules are optimal when the agent is sufficiently aligned with the principal. When this is not the case, optimal decision rules may contain gaps, that is, they may allow the agent to make high or low decisions but not intermediate ones. Such gaps may be optimal since they can be used to induce more state-sensitive decision making by an otherwise unresponsive agent. When such gaps are optimal, some of the perceived wisdoms about decision making in organizations do no longer hold. For instance, the principal may then give less discretion to a more aligned agent or to one with a bigger informational advantage. In other words, the Ally Principle and the Uncertainty Principle no longer hold.

Our theoretical results can be used to gain new insights into well-known economic problems such as the regulation of a monopolist who is privately informed about his costs. This problem was first analyzed in Baron and Myerson (1982) in a setting in which the regulator is able to make contingent transfers to the monopolist. While there are industries in which such transfers are indeed feasible, there are others in which they are ruled out by law. Our analysis can be applied to study optimal regulation in such industries. In this context we showed that for a large class of distributions and for common demand functions, optimal regulation without transfers takes a remarkably simple form which, moreover, is often observed in practice. In particular, the regulator cannot do better than to either give the monopolist no discretion at all or to impose a price cap and let him choose any price below this threshold.

Some of the decision rules that are optimal in our setting are therefore widely observed in practice. There are other rules, however, which are pervasive in organizations and which our model cannot rationalize. For instance, organizations often engage in *management by exception*, that is, they make the identity of the decision maker contingent on the state of the world. Existing research has emphasized the importance of ability differences between players in explaining management by exception.[33] Such ability differences are absent in our

[33] See Athey, Gans, Schaefer, and Stern (1994) and Garicano (2000).

model. While this assumption seems appropriate in some situations, such as setting a price for a regulated monopolist, it is not in others. It would be interesting to extend our analysis by allowing for ability differences.

Our analysis also assumes that decision rules do not affect the incentives of the agent to acquire information. Again this assumption is appropriate in some situations. The incentives of a regulated firm to learn its own production costs, for instance, are largely independent of the regulatory environment it faces. In other situations, however, the rules that an agent faces have a first order effect on his incentives to acquire information. A judge who knows that he has to impose a sentence of exactly five years, for instance, has a smaller incentive to learn about the specifics of case than a judge who is allowed to impose any sentence. In such a situation rules must then be designed, not just to elicit information from an agent, but also to motivate him to acquire information in the first place.[34] This issue, and others, awaits future research.

[34] Aghion and Tirole (1997) and Szalay (2005) analyze the effect of delegation on the incentives to acquire information but they largely abstract from the need to elicit information.

# 10 Appendix A

This appendix contains the proofs of all propositions and lemmas as well as the statements and proofs of Propositions 6 and 7. For some of the proofs it is convenient to define $\widetilde{S}(y) \equiv S(\theta_A(y))$ and $\widetilde{T}(y) \equiv T(\theta_A(y))$. Thus $\widetilde{S}(\cdot)$ and $\widetilde{T}(\cdot)$ represent the forward bias and backward bias as functions of the decision instead of the state.

**Proof of Lemma 1**: The proof is similar to the arguments presented in Strausz (2003). Consider an arbitrary deterministic mechanism and an equilibrium of that mechanism that implements outcome function $\sigma(\theta)$. 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 $u(\theta) = \mathrm{E}_{\sigma(\theta)}[u_P(m,\theta)]$ be the interim expected utility of the principal for type $\theta$. For each $\theta$ define $X(\theta) \in$supp $\sigma(\theta)$ such that $u_P(X(\theta),\theta) \geq \mathrm{E}_{\sigma(\theta))}[u_P(m,\theta)]$. Note that $X(\theta)$ is non-decreasing and hence Borel-measurable. Therefore the direct deterministic mechanism $(\Theta, X)$ is well defined, incentive compatible and satisfies $\mathrm{E}_\theta[u_P(s(\theta),\theta)] \geq \mathrm{E}_\theta\left[\mathrm{E}_{\sigma(\theta)}[u_P(m,\theta)]\right]$. Thus we can reduce the search for an optimal deterministic mechanism to the set of direct deterministic mechanisms that are incentive compatible. ■

**Proof of Lemma 2**: Follows immediately from Proposition 1 in Melumad and Shibano (1991). ■

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

Let $\theta, \widetilde{\theta} \in W$ and $\theta \neq \widetilde{\theta}$. It follows from single peakedness and symmetry of $u_A(y,\theta)$ w.r.t. to $y$ that $X(\theta) \subset \{s_\theta, s'_\theta\}$, where $s_\theta < y_A(\theta) < s'_\theta$ . Since $y_A(\theta)$ is strictly increasing, 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 $\theta \in W$ the number $s'_\theta - s_\theta > 0$. Define the sets $A_n = \{\theta \in W : \frac{1}{n} > s'_\theta - s_\theta \geq \frac{1}{1+n}\}$, $n \in \mathbb{N}$ and $A_0 = \{\theta \in W : s'_\theta - s_\theta \geq 1\}$. Note that each $A_n$ is a finite set. Since $W = \cup_{i=0}^{\infty} A_i$, $S$ is countable. ■

**Proof of Lemma 4**: Follows immediately from Theorem 1 in Holmström (1984). ■

**Proof of Lemma 5:** Since, following Lemma 2, 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 the delegation set $\{y_P^*\}$

$$\begin{aligned}\Delta(X(\theta)) &= \int_0^1 \left((y_P^* - y_P(\theta))^2 - (X(\theta) - y_P(\theta))^2\right)\mathrm{d}F(\theta) \\ &= -\left(y_P^*\right)^2 + \int_0^1 2X(\theta)y_P(\theta)\mathrm{d}F(\theta) - \int_0^1 X^2(\theta)\mathrm{d}F(\theta). \qquad (7)\end{aligned}$$

First, integrating by parts the second term of the right hand side of (7) we have

$$\int_0^1 2X(\theta)y_P(\theta)\mathrm{d}F(\theta) = 2X(1)y_P^* - \int_0^1 \left[\int_0^\theta 2y_P(z)\mathrm{d}F(z)\right]\mathrm{d}X(\theta).$$

Note that for every state $\theta$, incentive compatibility of the agent implies that $X^-(\theta) + X^+(\theta) = 2y_A(\theta)$, which implies that

$$(X^+(\theta))^2 - (X^-(\theta))^2 = 2y_A(\theta)\left(X^+(\theta) - X^-(\theta)\right). \qquad (8)$$

We can then integrate by parts the third term of the right hand side of (7) 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

$$\begin{aligned}\Delta(X(\theta)) &= -\left(y_P^*\right)^2 + 2X(1)y_P^* - X^2(1) \\ &\quad + \int_0^1 2y_A(\theta)F(\theta)\mathrm{d}X(\theta) - \int_0^1\int_0^\theta 2y_P(z)\mathrm{d}F(z)\mathrm{d}X(\theta) \\ &= -\left(y_P^* - X(1)\right)^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta). \qquad (9)\end{aligned}$$

Alternatively we can integrate by parts the second term on the right hand side of (7) to obtain $\int_0^1 2X(\theta)y_P(\theta)\mathrm{d}F(\theta) = 2X(0)y_P^* + \int_0^1 \left[\int_\theta^1 2y_P(z)\mathrm{d}F(z)\right]\mathrm{d}X(\theta)$. By application of (8) 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

$$\begin{aligned}\Delta(X(\theta)) &= -(y_P^*)^2 + 2X(0)y_P^* - X^2(0) - \int_0^1 2y_A(\theta)(1-F(\theta))\mathrm{d}X(\theta) \\ &\quad + \int_0^1 \left[\int_\theta^1 2y_P(z)\mathrm{d}F(z)\right]\mathrm{d}X(\theta) \\ &= -(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta). \qquad (10)\end{aligned}$$

Therefore,

$$\begin{aligned}V &= \max_{X(\theta)\in X} \Delta(X(\theta)) = \max_{X(\theta)\in X} -(y_P^* - X(1))^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) \\ &= \max_{X(\theta)\in \widehat{X}_D} -(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta). \quad \blacksquare\end{aligned}$$

**Proof of Proposition 1:** *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 $\theta^*$ 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 the delegation set $\{y_P^*\}$ 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 (5) $-(y_P^* - X(1))^2 + 2\int_0^1 T(\theta)\mathrm{d}X(\theta) \leq 0$. 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 (10) $-(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta) \leq 0$. This implies that $V = \max_{X(\theta)\in\widehat{X}_D} -(y_P^* - X(0))^2 - 2\int_0^1 S(\theta)\mathrm{d}X(\theta) = 0$. $\blacksquare$

**Proof of Lemma 6**: Follows immediately from the discussion in the text. ■

**Proof of Lemma 7**: 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 $D(\bar{t}) = \overline{D}$. 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

$$\begin{aligned}\Delta(t) &= \int_{\theta_A(y_2-t)}^{\theta_A(y_2)} (y_2 - t - y_P(\theta))^2 - (y_A(\theta) - y_P(\theta))^2\,\mathrm{d}F(\theta) \\ &+ \int_{\theta_A(y_2)}^{\theta_A(y_2+t)} (y_2 + t - y_P(\theta))^2 - (y_A(\theta) - y_P(\theta))^2\,\mathrm{d}F(\theta).\end{aligned} \tag{11}$$

Differentiating this expression gives $\Delta'(t) = 2\left[T(\theta_A(y_2+t)) + T(\theta_A(y_2-t)) - 2T(\theta_A(y_2))\right]$. Thus, if $T(\theta)$ is strictly concave in $[\theta_A(y_2 - \bar{t}), \theta_A(y_2 + \bar{t})]$ we have that $\Delta'(t) \leq 0$ for $t \leq \bar{t}$ and with strict inequality if $T(\theta)$ is strictly concave. Thus, in this case $\Delta(t) = \int_0^{\bar{t}} \Delta'(t)\mathrm{d}t \leq 0$ for $t \in (0, \bar{t}]$ and with strict inequality if $T(\theta)$ is strictly concave in $[\theta_A(y_2 - \bar{t}), \theta_A(y_2 + \bar{t})]$. It follows that, if $T(\theta)$ is concave in $[\theta_A(y_2 - \bar{t}), \theta_A(y_2 + \bar{t})]$, then (11) is negative for all $t \in [0, \bar{t}]$ and takes its minimum value for $t = \bar{t}$. This proves part (i.).

Conversely, if $T(\theta)$ is convex in $[\theta_A(y_2 - \bar{t}), \theta_A(y_2 + \bar{t})]$ we have that $\Delta'(t) \geq 0$ for $t \leq \bar{t}$ and with strict inequality if $T(\theta)$ is strictly convex. Since $\Delta(0) = 0$ this implies that $\Delta(t) = \int_0^{t} \Delta'(t)\mathrm{d}t \geq 0$ for $t \in (0, \bar{t}]$ and with strict inequality if $T(\theta)$ is strictly convex in $[\theta_A(y_2 - \bar{t}), \theta_A(y_2 + \bar{t})]$. Hence, if $T(\theta)$ is strictly convex in $[\theta_A(y_2 - \bar{t}), \theta_A(y_2 + \bar{t})]$, then (11) is positive for all $t \in [0, \bar{t}]$. This proves part (ii.). ■

**Proof of Proposition 2:** *Part (i.)*: 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 $\Delta((v-u)/2)$, where $\Delta(\cdot)$ is defined in (11) and the difference is evaluated at $(v+u)/2$. Since $T(\theta)$ is strictly convex in $[\theta_A(u), \theta_A(v)]$ by Lemma 7.ii. $\Delta((v-u)/2) > 0$. Thus, $D^*$ cannot be optimal.

*Part (ii.)*: We establish this claim in two steps. We first show that if $T(\theta)$ is strictly concave in $[\theta_A(y_1), \theta_A(y_2)]$ an optimal delegation set cannot contain any non-degenerate interval.

Second, we establish that the principal strictly prefers a delegation set with only two decisions in $[y_1, y_2]$ to one with more than 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 $\Delta((v-u)/2)$, where $\Delta(\cdot)$ is defined in (11) and the difference is evaluated at $(v+u)/2$. Since $T(\theta)$ is strictly concave in $[\theta_A(u), \theta_A(v)]$ by Lemma 6.i., $\Delta((v-u)/2) > 0$ 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.[35]

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 6.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 $(\theta_A(\underline{a}_i), \theta_A(\overline{a}_i))$. 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 $D^*$

[35] 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)).

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 $i$, $|\mathrm{E}(u_P(y,\theta)\,|D_i\,) - \mathrm{E}(u_P(y,\theta)\,|D^*\,)| \le \left|\max_{y\in Y} y - \min_{y\in Y} y\right| \Pr\left[\theta \in \left(\cup_{j=1}^{i} B_i\right)^c\right]$ 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 (iii.)*: 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 will show that if $D^*$ is optimal then $\widehat{D}$ is also optimal. Suppose on the contrary that $\mathrm{E}(u_P(y,\theta)\left|\widehat{D}\,\right.) < \mathrm{E}(u_P(y,\theta)\,|D^*\,)$. Then there must exist an interval $[u,v]$, $u \neq v$ such that $\mathrm{E}(u_P(y,\theta)\left|\widehat{D}\,\right.) < \mathrm{E}(u_P(y,\theta)\left|\widehat{D}_1\right.)$, where $\widehat{D}_1 = \widehat{D} \cap (u,v)^c$, or, equivalently $\Delta\left((v-u)/2\right) < 0$. However, since $T(\theta)$ is linear in $[\theta_A(u), \theta_A(v)]$ we have by Lemma 7 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.

*Part (iv.)*: 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 $S_{D^*}^c$ 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 Prob$[\theta \in S_{D^*}] > 0$ then $\mathrm{E}\left[u_P(y,\theta) \middle| \widetilde{D}\right] > \mathrm{E}\left[u_P(y,\theta) \,|D^*\right]$ contradicting the assumed optimality of $D^*$. ■

**Proof of Proposition 3**: *Necessity*: Suppose that $D^* = \left[\underline{y}, \overline{y}\right]$ is a minimal optimal delegation set and let $\underline{\theta} = \theta_A(\underline{y})$ and $\overline{\theta} = \theta_A(\overline{y})$. From Proposition 2 it is then necessary that $T(\theta)$ is convex for $y \in \left[\underline{\theta}, \overline{\theta}\right]$. 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

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

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

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 $\widetilde{\theta} < \underline{\theta}$, 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 the outcome functions 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 (9) 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.) delegation set $\{y_P^*\}$ 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.) Delegation set $\{y_P^*\}$ is not optimal*: Note that, since $T(\underline{\theta}) = 0$ and $T(\theta) \leq 0$ for $\theta < \underline{\theta}$, $T(\theta)$ is (weakly) increasing at $\theta = \underline{\theta}$ . Convexity of $T(\theta)$ in $\left[\underline{\theta}, \overline{\theta}\right]$ implies that $T(\theta) > 0$ for

$\theta \in (\underline{\theta}, \overline{\theta})$. A similar argument applied to $S(\theta)$ establishes that $S(\theta) < 0$ for $\theta \in (\underline{\theta}, \overline{\theta})$. By Proposition 3 it follows that $\{y_P^*\}$ cannot be optimal.

*(ii.) $D^*$ is empty outside of* $[\underline{y}, \overline{y}]$: First, since $\widetilde{T}(y) \leq 0$ for $y < \underline{y}$ , by (5) 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 (6) an optimal delegation set must have at most one decision above $\overline{y}$.

Second, we establish that $D^* \cap [\underline{y}, \overline{y}] \neq \varnothing$, i.e. any optimal delegation set must contain at least one decision in $[\underline{y}, \overline{y}]$. Suppose not, i.e. $D^* \cap [\underline{y}, \overline{y}] = \varnothing$. Then, since delegation is valuable $D^*$ must contain exactly two decisions. We will show that the optimal two-decision delegation set necessarily has at least one decision in $[\underline{y}, \overline{y}]$ thus reaching a contradiction. To this end, define $\widetilde{S}(y) \equiv S(\theta_A(y))$ and $\widetilde{T}(y) \equiv T(\theta_A(y))$ and 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 $y^* \in (\underline{y}, \overline{y})$. If $D^* \cap [\underline{y}, \overline{y}] = \varnothing$ 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 $[\underline{y}, \overline{y}]$. Given that both $\widetilde{T}(y)$ and $\widetilde{S}(y)$ are increasing, the last two inequalities imply that $y^* > (\overline{y} + \underline{y})/2$ and $y^* < (\overline{y} + \underline{y})/2$ which leads to a contradiction. Therefore it must be that $D^* \cap [\underline{y}, \overline{y}] \neq \varnothing$.

Third, we prove that, since any optimal delegation set $D^*$ must contain at least one decision in $[\underline{y}, \overline{y}]$, if there are decisions allowed by the principal in $D^* \cap [\underline{y}, \overline{y}]^c$ 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 $D^* \cap [\underline{y}, \overline{y}]^c = \varnothing$. 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 $D^* \cap [\underline{y}, \overline{y}]$. If $(y_1 + y_2)/2 > \overline{y}$ by (6) 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

$$\Delta U = 2\left[\widetilde{T}(y_1 + \frac{\epsilon}{2})\epsilon + \widetilde{T}(\frac{y_1 + y_2}{2} + \frac{\epsilon}{2})[y_2 - y_1 - \epsilon] - \widetilde{T}(\frac{y_1 + y_2}{2})(y_2 - y_1)\right] > 0$$

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 $\{y_P^*\}$ 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 show 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 $V = -\left(y_P^* - y_2\right)^2 + 2\int_{y_1}^{y_2} \widetilde{T}(y)\mathrm{d}y = -\left(y_P^* - y_1\right)^2 - 2\int_{y_1}^{y_2} \widetilde{S}(y)\mathrm{d}y$, 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:** Let $\widehat{F}(y,\lambda) = F(\theta_A(y,\lambda))$ and $\widehat{f}(y,\lambda) = f(\theta_A(y,\lambda))$. It suffices to show that there exists a $\overline{\lambda} \in (0,1)$ such that for $\lambda > \overline{\lambda}$, $\widetilde{T}(y,\lambda) = \widehat{F}(y,\lambda)y - \int_0^{\theta_A(y,\lambda)} y_P(\theta)\mathrm{d}F(\theta)$ 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.

Let $Y_A(\lambda)$ be the range of preferred decisions of an agent with preferred decisions $y_A(\theta,\lambda)$. For $y \in Y_A(\lambda)$ define $r(y,\lambda) \equiv \frac{\partial}{\partial y}\left[\theta_A(y,\lambda)\right]\widehat{f}(y,\lambda) = \widehat{f}(y,\lambda)/\left[1-\lambda+\lambda y_P'(\theta_A(y,\lambda)\right]$. 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, $\frac{\partial^2}{\partial y^2}\widetilde{T}(y,\lambda)$ satisfies

$$\begin{aligned} & \left[1-\lambda+\lambda y_P'(\theta_A(y,\lambda)\right]\frac{\partial^2}{\partial y^2}\widetilde{T}(y,\lambda) \\ = \quad & \widehat{f}(y,\lambda) + (1-\lambda)\left[1-\lambda+\lambda y_P'(\theta_A(y,\lambda)\right]\frac{\partial}{\partial y}\left[\left(y - y_P(\theta_A(y,\lambda))\right)r(y,\lambda)\right]. \end{aligned} \tag{12}$$

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

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

is uniformly bounded in $\Omega$, 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 (12) 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 Proposition 5:** Letting $D^*$ be an optimal delegation set we will show that (i.) $D^*$ is non-empty in $[0,1]$, (ii.) $D^*$ is empty in $[0,1]^c$ and (iii.) $D^*$ is an interval. We first note that from $R(0) > 0$ and $R(1) < 0$ it follows that $y_P(0) > 0$ and $y_P(1) < 1$ and, since $y_P(\theta)$ is (weakly) increasing $Y_P \subset [0,1]$ where $Y_P$ is the image of $y_P(\theta)$.

*(i.)* $D^* \cap [0,1] \neq \emptyset$. Since the only potential case that satisfies $D^* \cap [0,1] = \emptyset$ is when the optimal delegation set consists of two decisions outside of $[0,1]$, it suffices to show that the two decision optimal delegation set always has one decision in $[0,1]$. For the purpose of this proof we will define $P(y,\theta) = \int_0^\theta \frac{\partial}{\partial y} u_P(y,s)\,\mathrm{d}F(s)$ and $Q(y,\theta) = \int_\theta^1 \frac{\partial}{\partial y} u_P(y,s)\,\mathrm{d}F(s)$. Note that if $\theta_1 < \theta_2$ are the optimal decisions then they must satisfy

$$P(\theta_1, \frac{\theta_1+\theta_2}{2}) = Q(\theta_2, \frac{\theta_1+\theta_2}{2}) \equiv W.$$

Concavity of $u_P(y,\theta)$ implies that $P(y,\theta)$ and $Q(y,\theta)$ are both decreasing in $y$ while the condition $Y_P \subset [0,1]$ results in $P(0,\theta) > 0 > P(1,\theta)$ and $Q(0,\theta) > 0 > Q(1,\theta)$. Now suppose that $W > 0$. Then it must be that $0 < \theta_1 < \theta_2 < 1$ and thus $\theta_2 \in (0,1)$. Equivalently if we suppose that $W < 0$ we obtain $\theta_1 \in (0,1)$. In either case the optimal two-decision delegation set contains at least one point in $[0,1]$ and thus $D^* \cap [0,1] \neq \emptyset$.

*(ii.)* $D^* \cap [0,1]^c = \emptyset$. Suppose on the contrary that $\theta_2 \in D^*$ with $\theta_2 > 1$ (the proof for the case that $\theta_2 < 0$ follows a similar argument), and let $\overline{\theta}$ and $\theta_1$ be such that $R\left(\overline{\theta}\right) = 0$ and $\theta_1 = \max D^* \cap [0,1]$.

*(ii-a.)* $\frac{\theta_1+\theta_2}{2} > \overline{\theta}$. Suppose first that $\frac{\theta_1+\theta_2}{2} > \overline{\theta}$. We will consider two sub-cases depending on whether $\theta_1 < \overline{\theta}$ or $\theta_1 > \overline{\theta}$.

If $\theta_1 < \overline{\theta}$ the principal can increase her expected utility by offering the delegation set $\widehat{D} = (D^* - \{\theta_2\}) \cup [\theta_1, \theta_1 + \epsilon] \cup \{\theta_2 - \epsilon\}$. The change in expected utility is then given by

$$\begin{aligned}\Delta U &= \int_{\theta_1}^{\theta_1+\epsilon} \int_{\theta_1}^{\theta} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\mathrm{d}F(\theta) + \int_{\theta_1+\epsilon}^{\frac{\theta_1+\theta_2}{2}} \int_{\theta_1}^{\theta_1+\epsilon} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\mathrm{d}F(\theta) \\ &\quad - \int_{\frac{\theta_1+\theta_2}{2}}^{1} \int_{\theta_2+\epsilon}^{\theta_2} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\mathrm{d}F(\theta).\end{aligned}$$

We can compute $\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon}$ to obtain

$$\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon} = \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \frac{\partial}{\partial y} u_P(\theta_1,\theta)\,\mathrm{d}F(\theta) - \int_{\frac{\theta_1+\theta_2}{2}}^{1} \frac{\partial}{\partial y} u_P(\theta_2,\theta)\,\mathrm{d}s\mathrm{d}F(\theta) > 0.$$

Since $R(\theta_1) > 0$ and $y_P(\theta)$ is increasing we must have $\frac{\partial}{\partial y} u_P(\theta_1,\theta) > 0$ for $\theta > \theta_1$ and since $R(1) < 0$ and $y_P(\theta) < 1$ we must have $\frac{\partial}{\partial y} u_P(\theta_2,\theta) < 0$ for $\theta > \frac{\theta_1+\theta_2}{2}$. This implies that $\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon} > 0$ and thus for sufficiently small $\epsilon$ the expected utility of the principal increases when offering the delegation set $\widehat{D}$. and thus $D^*$ cannot be optimal.

Now we turn to the case that $\theta_1 > \overline{\theta}$. Note that optimality of $D^*$ requires that $\theta_2$ satisfies the first order condition

$$\int_{\frac{\theta_1+\theta_2}{2}}^{1} \frac{\partial}{\partial y} u_P(\theta_2, \theta)\, \mathrm{d}F(\theta) = \frac{1}{2} \int_{\theta_1}^{\theta_2} \frac{\partial}{\partial y} u_P\left(s, \frac{\theta_1+\theta_2}{2}\right) f(\frac{\theta_1+\theta_2}{2})\mathrm{d}s. \tag{13}$$

Now consider the delegation set $\widehat{D} = D^* \cup [\theta_1, \theta_1 + \epsilon]$ with $\epsilon$ sufficiently small. The change in the expected utility of the principal is given by

$$\begin{aligned}\Delta U &= \int_{\theta_1}^{\theta_1+\epsilon} \int_{\theta_1}^{\theta} \frac{\partial}{\partial y} u_P(s, \theta)\, \mathrm{d}s\mathrm{d}F(\theta) + \int_{\theta_1+\epsilon}^{\frac{\theta_1+\theta_2}{2}} \int_{\theta_1}^{\theta_1+\epsilon} \frac{\partial}{\partial y} u_P(s, \theta)\, \mathrm{d}s\mathrm{d}F(\theta) \\ &\quad - \int_{\frac{\theta_1+\theta_2}{2}}^{\frac{\theta_1+\theta_2+\epsilon}{2}} \int_{\theta_1+\epsilon}^{\theta_2} \frac{\partial}{\partial y} u_P(s, \theta)\, \mathrm{d}s\mathrm{d}F(\theta).\end{aligned}$$

We can compute $\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon}$ to obtain

$$\begin{aligned}\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon} &= \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \frac{\partial}{\partial y} u_P(\theta_1, \theta)\, \mathrm{d}F(\theta) - \frac{1}{2} \int_{\theta_1}^{\theta_2} \frac{\partial}{\partial y} u_P\left(s, \frac{\theta_1+\theta_2}{2}\right) f(\frac{\theta_1+\theta_2}{2})\mathrm{d}s = \\ &= \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \frac{\partial}{\partial y} u_P(\theta_1, \theta)\, \mathrm{d}F(\theta) - \int_{\frac{\theta_1+\theta_2}{2}}^{1} \frac{\partial}{\partial y} u_P(\theta_2, \theta)\, \mathrm{d}F(\theta)\end{aligned}$$

where the last equality follows from application of (13). By a similar argument we can compute $\Delta U$ when the principal offers the delegation set $\widehat{D} = (D^* \cap [\theta_1, \theta_1 - \epsilon]^c) \cup \{\theta_1 - \epsilon\}$. As a result, optimality of $D^*$ would imply that $\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon} = 0$ and thus

$$\int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \frac{\partial}{\partial y} u_P(\theta_1, \theta)\, \mathrm{d}F(\theta) = \int_{\frac{\theta_1+\theta_2}{2}}^{1} \frac{\partial}{\partial y} u_P(\theta_2, \theta)\, \mathrm{d}F(\theta) \equiv W. \tag{14}$$

Let $\widetilde{\theta}$ be such that $Q(\widetilde{\theta}, \theta_1) = \int_{\theta_1}^{1} \frac{\partial}{\partial y} u_P\left(\widetilde{\theta}, \theta\right) \mathrm{d}F(\theta) = 0$. Then (14) implies that $Q(\theta_1, \theta_1) > 0 > Q(\theta_2, \theta_1)$ and the principal can improve her expected utility by offering the delegation set $\widehat{D} = (D^* \cap \{\theta_2\}^c) \cup \left\{\widetilde{\theta}\right\}$ and thus $D^*$ is not optimal.

*(ii-b.)* $\frac{\theta_1+\theta_2}{2} < \overline{\theta}$. We now turn to the case that $\frac{\theta_1+\theta_2}{2} < \overline{\theta}$. We will show that the principal is made better off by allowing all decisions in $[\theta_1, \theta_1 + \epsilon]$ for $\epsilon > 0$ sufficiently small.

Again, optimality of $D^*$ requires that $\theta_2$ satisfies the first order condition

$$\int_{\frac{\theta_1+\theta_2}{2}}^{1} \frac{\partial}{\partial y} u_P(\theta_2, \theta)\, \mathrm{d}F(\theta) = \frac{1}{2} \int_{\theta_1}^{\theta_2} \frac{\partial}{\partial y} u_P\left(s, \frac{\theta_1+\theta_2}{2}\right) f(\frac{\theta_1+\theta_2}{2})\mathrm{d}s. \tag{15}$$

Now consider the delegation set $\widehat{D} = D^* \cup [\theta_1, \theta_1 + \epsilon]$ with $\epsilon$ sufficiently small. The change in the expected utility of the principal is given by

$$\begin{aligned} \Delta U &= \int_{\theta_1}^{\theta_1+\epsilon} \int_{\theta_1}^{\theta} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\mathrm{d}F(\theta) + \int_{\theta_1+\epsilon}^{\frac{\theta_1+\theta_2}{2}} \int_{\theta_1}^{\theta_1+\epsilon} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\mathrm{d}F(\theta) \quad (16) \\ &- \int_{\frac{\theta_1+\theta_2}{2}}^{\frac{\theta_1+\theta_2+\epsilon}{2}} \int_{\theta_1+\epsilon}^{\theta_2} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\mathrm{d}F(\theta). \end{aligned}$$

We can compute $\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon}$ to obtain

$$\begin{aligned} \lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon} &= \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \frac{\partial}{\partial y} u_P(\theta_1,\theta)\,\mathrm{d}F(\theta) - \frac{1}{2}\int_{\theta_1}^{\theta_2} \frac{\partial}{\partial y} u_P\left(s, \frac{\theta_1+\theta_2}{2}\right) f(\frac{\theta_1+\theta_2}{2})\mathrm{d}s = (17) \\ &= \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \frac{\partial}{\partial y} u_P(\theta_1,\theta)\,\mathrm{d}F(\theta) - \int_{\frac{\theta_1+\theta_2}{2}}^{1} \frac{\partial}{\partial y} u_P(\theta_2,\theta)\,\mathrm{d}F(\theta) \end{aligned}$$

where the last equality follows from application of (15). Since $u_P(y,\theta)$ is strictly concave in $y$ we have that

$$\begin{aligned} \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \frac{\partial}{\partial y} u_P(\theta_1,\theta)\,\mathrm{d}F(\theta) &\geq \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} R(s)\,\mathrm{d}s \geq R\left(\frac{\theta_1+\theta_2}{2}\right)\frac{\theta_2-\theta_1}{2} \quad (18) \\ \int_{\frac{\theta_1+\theta_2}{2}}^{1} \frac{\partial}{\partial y} u_P(\theta_2,\theta)\,\mathrm{d}F(\theta) &\leq \int_{\frac{\theta_1+\theta_2}{2}}^{1} R(s)\,\mathrm{d}s \leq R\left(\frac{\theta_1+\theta_2}{2}\right)\left(1-\frac{\theta_1+\theta_2}{2}\right), \quad (19) \end{aligned}$$

where in each case the first inequality follows from the concavity of $u_P(y,\theta)$ w.r.t $y$ and the second from $R(\theta)$ being weakly decreasing. Therefore combining (18) and (19) in (17) we obtain

$$\lim_{\epsilon\to 0} \frac{\Delta U}{\epsilon} \geq R\left(\frac{\theta_1+\theta_2}{2}\right)(\theta_2 - 1) > 0.$$

Therefore there exists $\epsilon > 0$ such that $\Delta U > 0$ and $D^*$ cannot be optimal.

*(iii.) $D^*$ is connected.* Suppose that $D^*$ is not connected, i.e. there exist two decisions $\theta_1, \theta_2 \in D^*$ with $\theta_1 \leq \theta_2$, such that $D^* \cap [\theta_1, \theta_2] = \varnothing$. Now consider a new delegation set $\widehat{D} = D^* \cup [\theta_1, \theta_2]$ The change in the expected utility of the principal $\Delta U$ when offering $\widehat{D}$ instead of $D^*$ is given by

$$\begin{aligned} \Delta U &= \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} [u_P(\theta,\theta) - u_P(\theta_1,\theta)]\,\mathrm{d}F(\theta) + \int_{\frac{\theta_1+\theta_2}{2}}^{\theta_2} [u_P(\theta,\theta) - u_P(\theta_2,\theta)]\,\mathrm{d}F(\theta) = \\ &= \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \left(\int_{\theta_1}^{\theta} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\right)\mathrm{d}F(\theta) - \int_{\frac{\theta_1+\theta_2}{2}}^{\theta_2} \left(\int_{\theta}^{\theta_2} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s\right)\mathrm{d}F(\theta). \quad (20) \end{aligned}$$

Since $u_P(y,\theta)$ is strictly concave in $y$ we have that

$$\int_{\theta_1}^{\theta} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s \quad > \quad \frac{\partial}{\partial y} u_P(\theta,\theta)(\theta-\theta_1), \tag{21}$$

$$\int_{\theta}^{\theta_2} \frac{\partial}{\partial y} u_P(s,\theta)\,\mathrm{d}s \quad < \quad \frac{\partial}{\partial y} u_P(\theta,\theta)(\theta_2-\theta). \tag{22}$$

Then using (21) and (22) in (20) we have that

$$\begin{aligned}
\Delta U \quad &> \quad \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} R(\theta)(\theta-\theta_1)\,\mathrm{d}\theta - \int_{\frac{\theta_1+\theta_2}{2}}^{\theta_2} R(\theta)(\theta_2-\theta)\,\mathrm{d}\theta = \\
&= \quad \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \left( \int_{\theta}^{\frac{\theta_1+\theta_2}{2}} R(s)\,\mathrm{d}s \right) \mathrm{d}\theta - \int_{\frac{\theta_1+\theta_2}{2}}^{\theta_2} \left( \int_{\frac{\theta_1+\theta_2}{2}}^{\theta} R(s)\,\mathrm{d}s \right) \mathrm{d}\theta \geq \\
&\geq \quad R\left(\frac{\theta_1+\theta_2}{2}\right) \left[ \int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \int_{\theta}^{\frac{\theta_1+\theta_2}{2}} \mathrm{d}s\mathrm{d}\theta - \int_{\frac{\theta_1+\theta_2}{2}}^{\theta_2} \int_{\frac{\theta_1+\theta_2}{2}}^{\theta} \mathrm{d}s\mathrm{d}\theta \right] = 0,
\end{aligned}$$

where the first equality follows by integrating by parts each integrand, and the last inequality follows from the assumption that $R(\theta)$ is weakly decreasing. Thus we have that $\Delta U > 0$ and reach the contradiction that $D^*$ is not an optimal delegation set.

As a result of steps (i.)-(iii.) $D^*$ must be an interval and interval delegation is optimal. ■

To state the next proposition, we define *upper-threshold delegation* as a delegation scheme in which $D = [\underline{d}_A, \overline{y}]$, where $\overline{y} \in (\underline{d}_A, \overline{d}_A)$, and *lower-threshold delegation* as a delegation scheme in which $D = [\underline{y}, \overline{d}_A]$, where $\underline{y} \in (\underline{d}_A, \overline{d}_A)$.

**Proposition 6.** *Upper-threshold delegation is optimal if and only if there exist a state* $\overline{\theta} \in (0,1)$ *such that*

*(i.)* $S(\overline{\theta}) = 0$, $S(\theta) \geq 0$ *for* $\theta > \overline{\theta}$ *and* $T(\theta) \geq 0$ *for* $\theta \leq \overline{\theta}$ *and*

*(ii.)* $T(\theta)$ *is convex for all* $\theta \in [0, \overline{\theta}]$.

*Lower-threshold delegation is optimal if and only if there exist a state* $\underline{\theta}$ *such that*

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

*(ii.)* $T(\theta)$ *is convex for all* $\theta \in [\underline{\theta}, 1]$.

**Proof:** 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 $D^* = \left[\underline{d}_A, y_A\left(\overline{\theta}\right)\right]$ is a minimal optimal delegation set then by Proposition 2, $T(\theta)$ is convex for $\theta \in \left[0, \overline{\theta}\right]$. Since the principal cannot improve by adding decisions above $y_A\left(\overline{\theta}\right)$ we must have $S(\theta) \geq 0$ for $\theta > \overline{\theta}$. Furthermore, optimality of $y_A\left(\overline{\theta}\right)$ requires that $S(\overline{\theta}) = 0$. Finally, if $T(\theta) < 0$ for some $\theta \in \left(0, \overline{\theta}\right)$ then convexity of $T(\theta)$ and the fact that $T(0) = 0$ implies that $T(\theta) < 0$ for $\theta \in \left(0, \overline{\theta}\right)$. This leads to a contradiction since the principal could increase her expected utility by an amount $-2\int_0^{\theta} T(x)\mathrm{d}x > 0$ by banning all decisions $[\underline{d}_A, y_A(\theta))$. Therefore we must have $T(\theta) \geq 0$ for $\theta \in \left[0, \overline{\theta}\right]$.

*Sufficiency:* We establish sufficiency by proving that (i.) the delegation set $\{y_P^*\}$ is not optimal, (ii.) an optimal delegation set has no decisions above $y_A\left(\overline{\theta}\right)$, and (iii.) $D^*$ is an interval and $D^* = \left[\underline{d}_A, y_A\left(\overline{\theta}\right)\right]$.

*(i.) The delegation set $\{y_P^*\}$is not optimal*: Note that, since $S(\overline{\theta}) = 0$ and $S(\theta)$ is concave in $\left[0, \overline{\theta}\right]$ (since $T(\theta)$ is concave in $\left[0, \overline{\theta}\right]$), it follows that $S(\theta) < 0$ for $\theta \in \left(0, \overline{\theta}\right)$. Since by assumption $T(\theta) > 0$ in the same region, Proposition 1 implies that the delegation set $\{y_P^*\}$ cannot be optimal.

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

*(iii.) $D^*$ is an interval and $D^* = \left[\underline{d}_A, y_A\left(\overline{\theta}\right)\right]$.* The value of delegation from offering the agent an interval $[y_A(\theta_1), y_A(\theta_2)]$ is given by $V = -\left(y_P^* - y_A(\theta_2)\right)^2 + 2\int_{\theta_1}^{\theta_2} T(\theta)\mathrm{d}\theta = -\left(y_P^* - y_A(\theta_1)\right)^2 - 2\int_{\theta_1}^{\theta_2} S(\theta)\mathrm{d}\theta$. Since $T(\theta) \geq 0$, we must have $\theta_1 = 0$. Also by differentiating w.r.t. $\theta_2$ we must have $S(\theta_2) = 0$ which implies $\theta_2 = 1$. ■

To state the next proposition, we define *complete delegation* as a delegation scheme in which $D = [\underline{d}_A, \overline{d}_A]$.

**Proposition 7.** *Complete delegation is optimal if and only if*

- *(i.)* $Y_A \subseteq Y_P$,
- *(ii.)* $T(\theta)$ *and* $S(\theta)$ *are increasing and* $T(\theta)$ *is convex for* $\theta \in [0, 1]$ *and*
- *(iii.)* $y_P^* \in Y_A^{\circ}$.

**Proof :** *Necessity*: Suppose complete delegation is optimal, i.e. $D^* = Y_A$. Now, let $S = Y_A \cap Y_P^c$ . By Proposition 2-iv $D^*$ 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 $T(\theta)$ is increasing and convex and $S(\theta)$ 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. $\Delta\left((v-u)/2\right) > 0$ (where the increment is computed at $(v+u)/2$). By the proof of Lemma 7 this implies that $T(\theta)$. Next we show $T'(0) \geq 0$ which, coupled with convexity of $T(\theta)$, entails that $T(\theta)$ is increasing in $\theta$. Suppose not, i.e. $T'(0) < 0$. Then $T(\theta) < 0$ in a region $(0, \widetilde{\theta})$. Consider now the delegation set $\widehat{D} = Y_A \cap [\underline{d}_A, y_A\left(\widetilde{\theta}\right))^c$. The difference in expected utility from $\widehat{D}$ and $D^*$ can be expressed as $\Delta U \equiv \mathrm{E}(u_P(y,\theta) \,\Big|\, \widehat{D}) - \mathrm{E}(u_P(y,\theta)\,|D^*) = -\int_0^{\widetilde{\theta}} T(\theta)\mathrm{d}\theta > 0$ which implies that $D^*$ is not optimal. Therefore it must be that $T'(0) \geq 0$ and, consequently, $T(\theta)$ is increasing. A similar argument shows that $S(\theta)$ is concave, $S(\theta) \leq 0$, and, since $S(1) = 0$, $S(\theta)$ is increasing.

Finally, since $T(0)+S(0) = S(0) < 0$ and $T(1)+S(1) = T(1) > 0$, continuity of $T(\theta)+S(\theta)$ implies that for some $\theta' \in (0,1)$ $T(\theta') + S(\theta') = y_A\left(\theta'\right) - y_P^* = 0$ . This establishes that $y_P^* \in \left(\underline{d}_A, \overline{d}_A\right)$.

*Sufficiency:* Note first that, under the conditions of the proposition, the delegation set $\{y_P^*\}$ can never be optimal. Indeed, since $y_P^* \in \left(\underline{d}_A, \overline{d}_A\right)$, $T(0)+S(0) = y_A\left(0\right) - y_P^* < 0$. Since $T(0) = 0$ this implies that $S(0) < 0$. By continuity and monotonicity of $T(\theta)$ there exists $\theta'$ such that $T(\theta') > 0$ and $S(\theta') < 0$. It follows from Proposition 1, that the delegation set $\{y_P^*\}$ can never be optimal.

Next we show that a two-decision delegation set cannot be optimal. To this end, recall that $\widetilde{S}(y) \equiv S(\theta_A(y))$ and $\widetilde{T}(y) \equiv T(\theta_A(y))$ and 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 $y^* + d^* \in \left(\underline{d}_A, \overline{d}_A\right)$. The case that $y^* - d^* \in \left(\underline{d}_A, \overline{d}_A\right)$ 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 $\Delta U = 2\left[\widetilde{T}(\overline{y} + \frac{\epsilon}{2})\epsilon + \widetilde{T}(\frac{\widehat{y}+\overline{y}}{2} + \frac{\epsilon}{2})\left[\widehat{y} - \overline{y} - \epsilon\right] - \widetilde{T}(\frac{\widehat{y}+\overline{y}}{2})\left(\widehat{y} - \overline{y}\right)\right] > 0$. We therefore reach a contradiction implying that adding projects above and below $Y_A$ can never be optimal. Therefore complete delegation is optimal. ■

# 11 Appendix B

This appendix includes the proofs of all results and the analysis of the complete contract.

**Proof of Result 1**: The forward bias is given by $S(\theta) = (1 - F(\theta))(A + \theta)/2 - \int_\theta^1 x \mathrm{d}F(x)$. Differentiating gives $S'(\theta) = (1 - F(\theta) - (A - \theta) f(\theta))/2$ and $S''(\theta) = -(A - \theta) f'(\theta)/2$. Thus we have that (i.) $S(0) \geq 0$ if and only if $A \geq 2\mathrm{E}(\theta)$, (ii.) $S(1) = 0$ and $S'(1) < 0$ and (iii.) $S''(\theta) \leq 0$ for all $\theta \in [0, \theta_m]$ and $S''(\theta) > 0$ for all $\theta \in (\theta_m, 1]$ where $\theta_m$ is the mode of the distribution. It follows from these facts that if $A \geq 2\mathrm{E}(\theta)$, then $S(\theta) > 0$ for all $\theta \in (0,1)$ and if $A < 2\mathrm{E}(\theta)$, then there exists a $\overline{\theta} \in (0, \theta_m)$ such that $S(\overline{\theta}) = 0$, $S(\theta) < 0$ for all $\theta \in \left[0, \overline{\theta}\right)$ and $S(\theta) > 0$ for all $\theta \in \left(\overline{\theta}, 1\right)$. Since, as explained in the text, $T(\theta) > 0$ for all $\theta \in (0,1)$, the result then follows from Proposition 1. ■

**Proof of Result 2**: To economize on the calculations, we apply Proposition 6. From the proof of Result 1 we know that if $A < 2\mathrm{E}(\theta)$, then there exists a $\overline{\theta} \in (0,1)$ such that $S(\overline{\theta}) = 0$ and $S(\theta) > 0$ for all $\theta \in \left(\overline{\theta}, 1\right)$. Since $T(\theta) \geq 0$ for all $\theta \in [0,1]$, condition (i.) in Proposition 6 is satisfied. We also know from the proof of Result 1 that $S''(\theta) < 0$ for all $\theta \in \left[0, \overline{\theta}\right]$. Since $T''(\theta) = -S''(\theta)$ this implies that condition (ii.) in Proposition 6 is also satisfied. Finally, implicitly differentiating $S(\overline{\theta}) = 0$ gives $\mathrm{d}\overline{\theta}/\mathrm{d}A = -2S'(\overline{\theta})/(1 - F(\overline{\theta})) < 0$. ■

**Proof of Result 3:** The backward and forward biases are given by $T(\theta) = F(\theta)\beta\theta + \sigma^2 [f(\theta) - f(-a)]$ and $S(\theta) = (1 - F(\theta))\beta\theta - \sigma^2 [f(\theta) - f(-a)]$. Note that $T(0) = \sigma^2 [f(0) - f(a)] > 0$ and $S(0) = -\,T(0) < 0$. The result then follows from Proposition 1. ■

**Proof of Result 4:** 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.) $\mathrm{E}[s \,|s \geq \theta] > \theta \;\; \forall \theta \in (-a, a)$, (b.) $0 < \frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}[s \,|s \leq \theta] < 1$ and $0 < \frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}[s \,|s \geq \theta] < 1 \; \forall \theta \in (-a, a)$, (c.) $\frac{\mathrm{d}}{\mathrm{d}\sigma}\mathrm{E}[s \,|s \geq \theta] \geq 0$.

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 $T(\theta)$ is convex in $[\underline{\theta}, \overline{\theta}]$. Note that for $\theta \in (-a, a)$, $T(\theta) = 0$ if and only if $R(\theta) \equiv \beta\theta - \mathrm{E}[s \,|s \leq \theta] = 0$. Since $R'(\theta) = \beta - \frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}[s \,|s \leq \theta] > 0$, $R(\underline{\theta}) = 0$ 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 non-degenerate interval $[\underline{\theta}, \overline{\theta}] \subset [-a, a]$. Furthermore, since $T(\theta) = -S(-\theta)$ we necessarily have that $\overline{\theta} = -\underline{\theta}$.

Since $T''(\theta) = \left[2\beta - 1 - (\beta-1)\theta^2/\sigma^2\right] f(\theta)$ 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 $T'(-a) = f(-a)a\,[1-\beta] < 0$ and $T'(\underline{\theta}) = F(\underline{\theta})[\beta - \frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{E}\,[s\,|s \leq \theta]\big|_{\theta=\underline{\theta}}] > 0$ 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 $T(\theta)$ is convex in $[\underline{\theta}, \overline{\theta}]$.

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

**Proof of Result 5:** Note that $Y_A \subset Y_P$. We have shown above that if $1/2 \leq \beta < 1$, then $T(\theta)$ is convex and $S(\theta)$ is concave in $[-a, a]$. Since $T'(-a) = f(-a)a\,[1-\beta] > 0$ and $S'(a) = -f(a)a\,[\beta - 1] > 0$, we must then have $T'(\theta) > 0$ and $S'(\theta) > 0$ $\forall\theta \in [-a, a]$. The result then follows from Proposition 7 (in Appendix A). ■

**Proof of Result 6:** For this proof it is useful to define $\widetilde{S}(y) \equiv S(\theta_A(y))$ and $\widetilde{T}(y) \equiv T(\theta_A(y))$. We first partition the decision space $Y$ in five regions $\Omega_i$ $i \in \{1, ..., 5\}$. $\Omega_1$ and $\Omega_5$ denote the region below and the region above the range of preferred decisions of the agent $[-\beta a, \beta a]$. In the text it has been shown that if $\beta < 1/2$ then $\widetilde{T}(y)$ is always concave in a neighborhood of zero. Given the assumption that $\beta a \geq 2\sigma^2\,[f(0) - f(a)]$, there exists $\widetilde{y} \in (0, \beta a)$ such that $\widetilde{T}''(\pm\widetilde{y}) = 0$, with $\widetilde{T}(y)$ convex in the regions above $\widetilde{y}$ and below $-\widetilde{y}$. Let $\Omega_2$ and $\Omega_4$ be the regions where $\widetilde{T}(y)$ is convex, namely $\Omega_2 = [-\beta a, -\widetilde{y}]$ and $\Omega_4 = [\widetilde{y}, \beta a]$. 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 $\beta a \geq 2\sigma^2\,[f(0) - f(a)]$ 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 4.

*(i.)* $D^* \cap (\Omega_1 \cup \Omega_2) \neq \varnothing$ *and* $D^* \cap (\Omega_4 \cup \Omega_5) \neq \varnothing$: Note that for $0 < \beta < 1/2$ both effective biases have a constant sign, i.e. $\widetilde{T}(y) > 0$ and $\widetilde{S}(y) < 0$ for $y \in (-\beta a, \beta a)$. 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 $\Omega_2$, contradicting the assumed optimality of $D^*$. Consider adding the decision $y_2 - 2\epsilon$

to $D^*$, where $\epsilon$ is such that $(y_1+y_2)/2-\epsilon$ belongs to $\Omega_2$. The increment of the expected utility of the principal is $\Delta U = 2\left[\widetilde{T}(y_2-\epsilon)2\epsilon + \widetilde{T}(\frac{y_1+y_2}{2}-\epsilon)\left[y_2-y_1-2\epsilon\right] - \widetilde{T}(\frac{y_1+y_2}{2})(y_2-y_1)\right]$. 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 $\Omega_2$ 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 $\Omega_3$ (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 $\Omega_2$ 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 $\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\left(\frac{y_2+\underline{y}_3}{2}\right)\left(\frac{\underline{y}_3-y_2}{2}\right) = \int_{-\beta a}^{\frac{y_2+\underline{y}_3}{2}}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\mathrm{d}y$. Since $\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}$ is strictly increasing and positive in $\Omega_2$ we have that $\int_{-\beta a}^{\frac{y_2+\underline{y}_3}{2}}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\mathrm{d}y < \frac{\mathrm{d}\widetilde{T}}{\mathrm{d}y}\left(\frac{y_2+\underline{y}_3}{2}\right)\left(\frac{y_2+\underline{y}_3}{2}-(-\beta a)\right)$ implying that $\left(\frac{\underline{y}_3-y_2}{2}\right) < \left(\frac{y_2+\underline{y}_3}{2}-(-\beta a)\right)$ and $y_2 > -\beta a$. 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 $\Omega_3$, 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 $q = \mathrm{E}\left[s\,|s\geq 0\right] = 2\sigma^2\left[f(0)-f(a)\right]$. Therefore whenever $\mathrm{E}\left[s\,|s\geq 0\right] < \beta a$ 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 $D^* = [-\beta a, \underline{y}]\cup[\overline{y}, \beta a]$. 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 $\overline{y} = \sigma^2\frac{f(0)-f(\overline{y}/\beta)}{F(\overline{y}/\beta)-1/2}$. ■

## 11.1 Regulation with Constant Elasticity Demand Curves

Reconsider the linear regulation model but suppose now that demand is characterized by constant elasticity demand curves for which $y = (\gamma_0/q)^{1/\gamma_1}$ if $q \geq \varepsilon$ and $y = (\gamma_0/\varepsilon)^{1/\gamma_1}$ if $q < \varepsilon$, where $\varepsilon > 0$ is arbitrarily small, $\gamma_0 > 0$ and $\gamma_1 > 1$.[36] The support of the marginal costs is given by $[\theta_l, \theta_h]$, where $\theta_h > \theta_l > 0$. Using Taylor series approximations we can express profits and welfare, respectively, as $\Pi(y) = \Pi(y_A(\theta)) + \frac{1}{2}\Pi''(y_A(\theta))(y - y_A(\theta))^2$ and $W(y) = W(y_P(\theta)) + \frac{1}{2}W''(y_P(\theta))(y - y_P(\theta))^2$, where $y_A(\theta) = \theta\gamma_1/(\gamma_1 - 1)$ and $y_P(\theta) = \theta$ are the profit and welfare maximizing prices. We can now establish the following result

**Result 7.** *Delegation is valuable if and only if $y_A(\theta_l) < \mathrm{E}(\theta)$. When delegation is valuable, then $D^* = \left[y_A(\theta_l), y_A(\overline{\theta})\right]$, where $y_A(\overline{\theta}) = \mathrm{E}(\theta \,|\theta \geq \overline{\theta})$. Moreover, the regulator gives the monopolist more discretion, the more elastic the demand.*

**Proof of Result 7:** The term $W''(\cdot)$ that pre-multiplies the quadratic term $(p - p_P(\theta))^2$ in the approximated welfare function depends on $\theta$. We therefore need to transform the cumulative density function as described in Section 3. Thus, let $G(\theta) = \int_{\underline{c}}^{\theta} W''(t)f(t)dt / \int_{\underline{c}}^{\overline{\theta}} W''(t)f(t)\mathrm{d}t$ and note that $g(\theta) = K\theta^{-(\gamma_1+1)}f(\theta)$ and $g'(\theta) = K\theta^{-(\gamma_1+2)}(\theta f'(\theta) - (\gamma_1 + 1)f(\theta))$, where $K \equiv -\gamma_0\gamma_1 / \int_{\underline{c}}^{\overline{\theta}} W''(t)f(t)\mathrm{d}t > 0$. The backward bias is then given by $T(\theta) = G(\theta)\left(\gamma_1\frac{\theta}{\gamma_1 - 1}\right) - \int_{\underline{c}}^{\theta} xg(x)\mathrm{d}x$ and the forward bias is given by $S(\theta) = (1 - G(\theta))\left(\gamma_1\frac{\theta}{\gamma_1 - 1}\right) - \int_{\theta}^{\overline{\theta}} xg(x)\mathrm{d}x$. Differentiation gives $T''(\theta) = \frac{K}{\gamma_1 - 1}\theta^{-\gamma_1}f'(\theta)$, $S'(\theta) = (1 - G(\theta))\left(\frac{\gamma_1}{\gamma_1 - 1}\right) - \frac{\theta}{\gamma_1 - 1}g(\theta)$ and $S''(\theta) = -\frac{K}{\gamma_1 - 1}\theta^{-\gamma_1}f'(\theta)$.

The backward bias is strictly positive for all $\theta \in (\theta_l, \theta_h)$ and it is convex for $\theta \in [\theta_l, \theta_m)$, where $\theta_m$ is the mode of the distribution. Consider next the forward bias. Note first that $S(\theta) > 0$ for all $\theta \in [\theta_m, \theta_h)$. This follows from the facts that $S(\theta_h) = 0$, $S'(\theta_h) < 0$ and $S''(\theta) > 0$ for all $\theta \in (\theta_m, \theta_h]$. Note next that $S''(\theta) < 0$ for all $\theta \in [\theta_l, \theta_m)$. Since $S(\theta_m) > 0$ we find that $S(\theta) > 0$ for all $\theta \in (\theta_l, \theta_h)$ if and only if $S(\theta_l) \geq 0$, i.e. if and only if $y_A(\theta_l) \geq \mathrm{E}(\theta)$. It then follows from Proposition 2 that the principal cannot do better than to set $y = \mathrm{E}(\theta)$. Moreover, if $S(\theta_l) < 0$, then there exists a unique $\overline{\theta} \in (\theta_l, \theta_m)$ for which $S(\overline{\theta}) = 0$. It then follows from Propositions 6 that the regulator cannot do better than to let the monopolist set any price below $y_A(\overline{\theta}) = \mathrm{E}(\theta \,|\theta \geq \overline{\theta})$. Implicitly differentiating $S(\overline{\theta}) = 0$

[36] We assume $\epsilon > 0$ since welfare is not well-defined for $\epsilon = 0$. In particular, if $y = (\gamma_0/\varepsilon)^{1/\gamma_1}$ for $q \geq 0$, then welfare is infinite for any $q > 0$.

shows that $\mathrm{d}\overline{\theta}/\mathrm{d}\gamma_1 > 0$. ■

### 11.2 Contingent Transfers

For the characterization of the complete contract (under limited liability) we draw on Krishna and Morgan (2006).[37] Suppose first that $b \leq 1/3$. Then the optimal decision rule $y(\theta)$ and the optimal transfers $t(\theta)$ are given by

$$y(\theta) = \begin{cases} \frac{3}{2}\theta + \frac{1}{2}b & \text{if } \theta \leq b \\ \theta + b & \text{if } b \leq \theta \leq 1 - 2b \\ 1 - b & \text{if } 1 - 2b \leq \theta \end{cases} \quad \text{and} \quad t(\theta) = \begin{cases} \frac{3}{4}\left(b^2 - \theta\left(2b - \theta\right)\right) & \text{if } \theta \leq b \\ 0 & \text{if } \theta > b. \end{cases}$$

The principal's expected utility is then given by $\mathrm{E}(u_P(y(\theta), \theta) - t(\theta)) = -\frac{1}{2}b^2\left(2 - 3b\right)$. Suppose next that $1/3 \leq b \leq 1$. Then

$$y(\theta) = \begin{cases} \frac{3}{2}\theta + \frac{1}{2}b & \text{if } \theta \leq z \\ \frac{3}{2}z + \frac{1}{2}b & \text{if } z \leq \theta \leq 1 \end{cases} \quad \text{and} \quad t(\theta) = \begin{cases} \frac{3}{4}\left(\theta^2 - 2b\theta + z(2b - z)\right) & \text{if } \theta \leq z \\ 0 & \text{if } z \leq \theta \leq 1 \end{cases}$$

where $z = \left(\frac{1}{2} - \frac{1}{6}\sqrt{3}\sqrt{4b - 1}\right)$. The principal's expected utility is then given by $\mathrm{E}(u_P(y(\theta), \theta) - t(\theta)) = -\frac{1}{24}\left(-\left(4b - 1\right)\sqrt{3\left(4b - 1\right)} + 6b\left(b + 1\right) - 1\right)$. The principal's expected utility under optimal delegation, agent control and principal control can also easily be worked out. Full numerical details for the comparisons in the text are available from the authors.

[37] In particular, see Appendix C in Krishna and Morgan (2006).

## References


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

Alonso, R., and N. Matouschek (2005): "Optimal Delegation," *Kellogg School of Management, mimeo.*

——— (2007): "Relational Delegation," *The RAND Journal of Economics*, forthcoming.

Armstrong, M. (1995): "Delegating Decision-Making to an Agent with Unknown Preferences," *University College London, mimeo.*

Armstrong, M., and D. Sappington (2004): "Recent Developments in the Theory of Regulation," *Handbook of Industrial Organization*, 3.

Athey, S., J. Gans, S. Schaefer, and S. Stern (1994): "The Allocation of Decisions in Organizations," *Stanford Graduate School of Business Research Paper No. 1322*.

Baron, D. (2000): "Legislative Organization with Informational Committees," *American Journal of Political Science*, 44(3), 485–505.

Baron, D., and R. Myerson (1982): "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50(4), 911–930.

Crawford, V., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451.

Dessein, W. (2002): "Authority and Communication in Organizations," *Review of Economic Studies*, 69(4), 811–838.

Epstein, D., and S. O'Halloran (1999): *Delegating Powers: A Transaction Cost Politics Approach to Policy Making Under Separate Powers*. Cambridge University Press.

Filipi, A., and R. Singh (2006): "Incentives, Reputation and the Allocation of Authority," *Harvard Business School, mimeo.*

Garicano, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108(5), 874–904.

Gilligan, T., and K. Krehbiel (1987): "Collective Decision-making and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures," *Journal of Law, Economics, & Organization*, 3(2), 287–335.

——— (1989): "Asymmetric Information and Legislative Rules with a Heterogeneous Committee," *American Journal of Political Science*, 33(2), 459–490.

Green, J. (1982): "Statistical Decision Theory Requiring Incentives for Information Transfer," in *The Economics of Information and Uncertainty*, ed. by J. McCall. The University of Chicago Press, Chicago and London.

Harris, M., and A. Raviv (2005): "Allocation of Decision-Making Authority," *Review of Finance*, 9(3), 353–383.

Holmström, B. (1977): "On Incentives and Control in Organizations," Ph.D. thesis, Stanford University.

——— (1984): "On the Theory of Delegation," in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom. North-Holland, New York.

Holmström, B., and P. Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics & Organization*, 7(1), 24–51.

Huber, J., and C. Shipan (2006): "Politics, Delegation, and Bureaucracy," in *The Handbook for Political Economy*, ed. by B. Weingast, and D. Wittman. Oxford University Press, Oxford.

Kovac, E., and T. Mylovanov (2006): "Stochastic Mechanisms in Settings without Monetary Transfers: Regular Case," *University of Bonn, mimeo.*

Krishna, V., and J. Morgan (2001): "Asymmetric Information and Legislative Rules: Some Amendments," *American Political Science Review*, 95(02), 435–452.

——— (2006): "Contracting for Information under Imperfect Commitment," *Penn State University and University of California at Berkeley, mimeo.*

LAFFONT, J.-J., AND D. MARTIMORT (2002): *The Theoy of Incentives: The Principal Agent Problem.* Princeton Univeristy Press, Princeton.

LAFFONT, J.-J., AND J. TIROLE (1990): "The Politics of Government Decision Making: Regulatory Institutions," *Journal of Law, Economics, & Organization*, 6(1), 1–31.

——— (1993): *A Theory of Incentives in Procurement and Regulation.* Massachusetts Institute of Technology Press, Cambridge, MA.

MARINO, A. (2006): "Does Delegation Always Dominate Veto in Organizational Games of Strategic Communication?," *University of Southern California, mimeo.*

MARINO, A., AND J. MATSUSAKA (2005): "Decision Processes, Agency Problems, and Information: An Economic Analysis of Capital Budgeting Procedures," *Review of Financial Studies*, 18(1), 301–325.

MARTIMORT, D., AND A. SEMENOV (2006a): "Competition vs. Collusion in Legislative Politics: A Mechanism Design Approach," *University of Toulouse, mimeo.*

——— (2006b): "Continuity in Mechanism Design without Transfers," *Economics Letters*, 93(2), 182–189.

MELUMAD, N., AND T. SHIBANO (1991): "Communication in Settings with No Transfers," *The RAND Journal of Economics*, 22(2), 173–198.

MOOKHERJEE, D. (2006): "Decentralization, Hierarchies and Incentives: A Mechanism Design Perspective," *Journal of Economic Literature*, 44, 367–390.

MYLOVANOV, T. (2006): "Veto-Based Delegation," *University of Bonn, mimeo.*

OTTAVIANI, M. (2000): "The Economics of Advice," *University College London, mimeo.*

RUDIN, W. (1987): *Real and Complex Analysis.* McGraw-Hill, Inc. New York, NY, USA.

Strausz, R. (2003): "Deterministic Mechanisms and the Revelation Principle," *Economics Letters*, 79(3), 333–337.

Szalay, D. (2005): "The Economics of Clear Advice and Extreme Options," *Review of Economic Studies*, 72(4), 1173–1198.