

# Optimal Sequential Delegation

Eugen Kováč\*      Daniel Krähmer†

July 6, 2012

## Abstract

The paper extends the optimal delegation framework pioneered by Holmström (1977, 1984) to a dynamic environment where, at the outset, the agent privately knows his ability to interpret decision relevant private information received later on. We show that any mechanism can be implemented by a menu of delegation sets and characterize the optimal menu for the uniform quadratic case. Sequential delegation is strictly better than static delegation whenever the conflict of interest is small, the agent's ability is sufficiently dispersed, and the likelihood of facing a high ability agent is sufficiently small.

*Keywords:* optimal delegation, sequential screening, dynamic mechanism design, non-transferable utility.

*JEL Codes:* D02, D20, D82, D86.

---

\*Universität Bonn, Department of Economics, Adenauer Allee 24–42, D-53113 Bonn (Germany); E-mail: [eugen.kovac@uni-bonn.de](mailto:eugen.kovac@uni-bonn.de); URL: [www.uni-bonn.de/~kovac](http://www.uni-bonn.de/~kovac).

†Universität Bonn, Department of Economics, Adenauer Allee 24–42, D-53113 Bonn (Germany), E-mail: [kraehmer@hcm.uni-bonn.de](mailto:kraehmer@hcm.uni-bonn.de). D. Krähmer gratefully acknowledges financial support by the DFG (German Science Foundation) under SFB/TR-15.

# 1 Introduction

Decision makers frequently lack decision relevant information and are reliant on informed agents who have conflicting preferences over the appropriate course of action. When firm headquarters need to make strategic business decisions, division managers possess superior knowledge about how these decisions affect the firm as a whole but prefer firm strategies that benefit only their own division. Regulators seek to make regulatory decisions based on privately known firm characteristics but, instead of the firm's profits, want to maximize social welfare. Financial advisors have better information than investors about portfolio-investor match but are often remunerated based on commissions, biasing them towards recommending certain funds.

The problem of a principal who is dependent on decision relevant information known to a biased agent has been extensively studied in the literature (see the review below). The literature typically focuses on static situations in which the agent has *perfect* information about the decision relevant aspects of the world already at the beginning of the relation. In this paper, we depart from this assumption and investigate dynamic environments where the agent receives decision relevant private information only during the course of the relation, and at the outset, has private information about the quality of the information he will obtain later on.

Take a firm, for example, a car maker, who considers introducing a new model line. The decision has to be made years before the actual market conditions realize that determine the ultimate profitability of the new model. While division managers responsible for developing the new car cannot know future conditions perfectly, they will be better able than headquarters to predict the success of the launch. Moreover, they will frequently have additional private information about the accuracy of their prediction, for example, they may have private contacts to competitors and, knowing about the rival's plans, be better able to predict profitability than if they do not have this information.

A similar information structure can be also identified in the regulation example. To provide planning certainty for firms (and due to other administrative constraints), an industry regulator typically has to irrevocably set its policy over a longer duration of time. A firm consulted by the regulator at the outset has thus only imprecise information, say about future production costs, but, depending on the privately known strength of the firm's market research division may be more or less able to predict those future costs.

Also, when an investor and a financial advisor engage in business, the advisor learns about the former's needs (e.g. the type of desired diversification) only over time, yet most advisors are specialized in particular areas (emerging markets, technology markets, etc.), and so, unknown to the customer, may be more or less able to tie their advice to the investor's actual needs.

In such a situation, how should the principal structure the decision making process to arrive at an optimal decision? In this paper, we address this question within a dynamic extension of Holmström's (1977, 1984) seminal mechanism design framework with non-transferable utility where the principal can commit to a decision contingent on the information provided by agent.

In the first step of our analysis, we extend Holmström's insight that any mechanism can be implemented by delegating decision making to the agent. Instead of enacting the decision herself, the principal offers the agent a delegation set of permissible decisions. To account for the sequential revelation of information, in our setup any mechanism can be implemented by a *menu* of delegation sets where, first, the agent chooses a delegation set from the menu, and then, after having obtained new information, he makes a decision within the initially chosen set.

Moreover, only menus are implementable which are *incentive compatible*, that is, the menu contains one delegation set for each realization of the agent's ex ante information, and the agent has an incentive to select the delegation set corresponding to his ex ante information. A special menu is the degenerate menu which consists only of identical delegation sets. We refer to such a menu as static because, in this case, the sequential nature of information arrival is irrelevant, and the ultimate decision only depends on the agent's final information. Whenever the menu is not static, we refer to it as sequential.

Relative to static delegation, the principal can implement a larger set of decisions under sequential delegation, because this allows her to adapt the delegation set to the agent's ex ante information. This is possible because at the ex ante stage, the agent chooses the delegation set with those decisions that he *believes* he is most likely to favor ex post, after observing his signal realization. Since the agent holds private ex ante beliefs about the distribution of his ex post signal, the agent can be induced to choose different delegation sets for different ex ante beliefs.

It is hence clear that the optimal sequential delegation menu is at least weakly better for the principal than the optimal static delegation set. In this paper we ask when sequential delegation is *strictly* superior to static delegation, and if so, how the optimal sequential delegation menu looks like.

We address this question in the familiar setting with quadratic utility, constant bias, and uniformly distributed state. We assume that the agent is biased towards higher decisions than the principal. Unlike in the standard setting, the agent in our setup privately observes an imperfect signal about the state during the course of the relation, but has private information at the outset about the precision of the signal he will receive. We interpret the signal's precision as the agent's forecasting ability and consider the simplest case that the agent's forecasting ability can be of two possible types: either high

or low. The difference between the high and the low ability, *the ability difference*, is a measure of the agent's ex ante private information.

We identify three factors that determine whether sequential delegation is more profitable than static delegation: (i) the ability difference, (ii) the size of the conflict of preferences (measured by the agent's bias), and (iii) the distribution of the agent's types. First, we show that if the ability difference is sufficiently small compared to the agent's bias, static delegation is optimal irrespective of the distribution of the agent's types. On the other hand, for sufficiently large ability difference (relative to the agent's bias), sequential delegation becomes optimal irrespective of the distribution of the agent's types. Finally, if the ability difference is in an intermediate range, static delegation is optimal whenever the high type is sufficiently likely. Moreover, we show that sequential delegation, whenever optimal, involves offering moderate decisions to the low ability type and more extreme decisions to the high ability type. In fact, the delegation set for the low type is an interval, while the delegation set for the high type is disconnected with moderate decisions being excluded.

To provide intuition for these results, it is useful to consider the benchmark case in which the agent's ability is publicly known so that his only private information is the signal realization he observes. The design of an optimal delegation set is then determined by the interplay of two forces: by including more decisions, the principal makes better use of the agent's private information (*information effect*), but at the same time gives up control over the decision (*loss-of-control effect*). The loss-of-control effect makes the principal averse to delegate large decisions, because due to his positive bias, the agent tends, from the principal's perspective, towards to excessively large decisions. Thus the larger the agent's bias, the smaller the decisions the principal delegates to the agent. Moreover, the information effect is stronger for the high than for the low ability agent because the former possesses more precise information. Therefore, the principal would like to delegate more decisions, represented by a larger delegation set, to the agent with the high forecasting ability.

When the agent's ability is his private information, offering a larger delegation set to the high ability type is not incentive compatible, as nothing prevents the low type agent from mimicking the high type. Incentive compatibility is trivially restored under a static delegation set that does not condition on the agent's ability. We show that, compared to the case with publicly observable ability, under optimal static delegation, the principal delegates too many decisions to the low type and too few to the high type.

Alternatively, the principal may restore incentive compatibility by offering the high type a set of relatively extreme, including large, decisions but excluding some moderate decisions. In fact, the low ability type agent is aware that his final information is less

precise and anticipates that he will likely favor moderate decisions after having observed the signal. Therefore, the low ability agent will refrain from taking a delegation set that offers only extreme yet no moderate decisions. Reversely, the high type agent, who anticipates to observe more dispersed signal realizations, considers it less likely that his optimal decisions are moderate. Therefore, such an agent will prefer a delegation set with only extreme to one with only moderate decisions. In other words, offering a delegation menu with a set consisting of extreme decisions and one consisting of moderate decisions allows the principal to screen the agent's ability type.

The key question is when such a sequential delegation menu is better for the principal than the optimal static delegation set. On one hand, sequential delegation makes better use of the high ability type's information, because it allows the final decision to be better adapted to the more dispersed signal of the high ability agent. On the other hand, disallowing moderate decisions is also costly if the signal indicates that a moderate decision should be taken. The three factors identified earlier determine when the benefits outweigh the costs of sequential delegation. First, if the ability difference is small compared to the bias, offering the high ability agent more extreme decisions has relatively few benefits (as his informational advantage is small) but high costs (because due to the large bias, the loss-of-control effect makes offering larger decisions relatively costly). Second, the reverse logic applies when the ability difference is large compared to the bias, and, in that case, sequential delegation is optimal.

Third, for intermediate values of the ability difference, the profitability of sequential delegation also depends on the prior distribution of the agent's types. To understand this, observe that when the high ability type is very likely, the optimal static delegation set is approximately the same as the optimal delegation set when the agent's ability is publicly known to be high. Therefore, already under static delegation, relatively large decisions are offered to the agent, and the marginal benefit of offering even larger decisions is relatively low. By the same token, if the high ability type is very unlikely, the optimal static delegation set is not much different from the optimal delegation set when the agent's ability is publicly known to be low. In particular, in this case, under static delegation only moderate decisions are offered. Therefore, offering the high type more extreme decisions has relatively large benefits due to the information effect, and thus sequential delegation becomes optimal.

### *Related Literature*

Our paper brings together two strands of the literature: the literature on optimal delegation and the literature on sequential screening. Starting with the seminal work of Holmström (1977, 1984) the optimal delegation literature asks how a principal should optimally delegate decision making to an informed, but biased agent. Alonso and Ma-

touscheck (2008) provide conditions that characterize the optimal delegation set in general environments, including characterizations for when simple interval delegation is optimal.<sup>1</sup> Our main contribution to this literature is that, instead of a static environment with a fully informed agent, we consider a dynamic environment where the agent’s information arrives sequentially over time. Our result that sequential delegation requires the principal offer delegation sets with only extreme options is similar to the insight by Szalay (2005) that offering extreme options is optimal to provide incentives for an initially uninformed agent to endogenously acquire information. Instead of providing incentives for information acquisition, in our setup offering only extreme options provides incentives for the agent to reveal information about his *ex ante* expertise. Our paper draws heavily on the insights in Kováč and Mylovanov (2009) who provide a handy and intuitive representation of the principal’s expected utility as a weighted average of the agent’s utility. In particular, we exploit this representation to derive the optimal static delegation set in our setup. This is not trivial since, for some parameters, our setup violates the conditions for which Alonso and Matouscheck (2008) or Kováč and Mylovanov (2009) characterize the solution to the static delegation problem.<sup>2</sup>

Secondly, our paper is related to the literature on sequential screening in the spirit of Courty and Li (2000) who study a dynamic price discrimination problem where the agent, while knowing a private signal at the time of contracting, learns his actual willingness to pay at the time of consumption only over time.<sup>3</sup> In terms of information structure and timing, our setup is essentially the same as that in Courty and Li (2000). But, the key difference is that we consider parties without transferable utility so that contingent monetary transfers are not available. Courty and Li (2000) establish that under standard regularity conditions the optimal contract always elicits the agent’s information sequentially and strictly improves over the optimal static screening contract. In contrast, in our setup both static and sequential delegation can be optimal, depending on parameters. In this sense, the availability of money as a screening instruments works in favor of, but is not necessary, for the optimality of sequential screening.<sup>4</sup>

This paper is organized as follows.

---

<sup>1</sup>See also Mylovanov (2009) who studies optimal delegation when the principal has veto power *ex post*. Martimort and Semenov (2006) investigate when the optimal mechanism is continuous.

<sup>2</sup>The optimal delegation literature assumes that the principal can contractually constrain the agent’s discretion, that is, decisions are contractible. A large literature studies the problem with non-contractible decisions and asks how the unconstrained delegation of authority to an informed and biased agent compares to other modes of organization, such as communication. See e.g. Riordan and Sappington (1987), Dessein (2002), Krähmer (2006), Bester (2009), Goltsman, Hörner, Pavlov, and Squintani (2009) to name only a few.

<sup>3</sup>See also Baron and Besanko (1984), Battaglini (2005), Esö and Szentes (2007a, 2007b), Dai, Lewis, and Lopomo (2006), Krähmer and Strausz (2008, 2011), Pavan, Segal, and Toikka (2008), Inderst and Hoffmann (2009), Nocke, Peitz, and Rosar (2011), Inderst and Peitz (2011).

<sup>4</sup>On a related note, Krähmer and Strausz (2012) show that in the standard sequential screening model with money, a sequential contract is never optimal if no *ex post* losses can be imposed on the agent.

Section 2 introduces the setup. Section 3 derives the principal’s mechanism design problem, and Section 4 shows that the principal’s problem can be equivalently stated as a sequential delegation problem. Section 5 describes the principal’s fundamental trade-off when designing delegation sets, and Section 6 characterizes two benchmark cases: when the agent’s ability is publicly known and when only static delegation is feasible. Section 7 fully characterizes the solution to the sequential delegation problem. Section 8 concludes. Proofs of all propositions and lemmas are relegated to the Appendix.

## 2 Model

A principal (she) has to take a decision  $x \in \mathbb{R}$ . Her payoff from the decision depends on a state of the world given by the random variable  $\tilde{\theta}$  which we assume to be uniformly distributed on the unit interval. The principal does not know the true state and consults an agent (he) for advice. When the principal approaches the agent, also the agent does not know the true state, but before the decision is taken, he privately observes an imperfect signal  $\tilde{s}$  about the true state. In addition, the agent privately knows his degree of expertise or his forecasting ability which is captured by the quality his signal. We assume that  $\tilde{s}$  coincides with the true state with probability  $p \in [0, 1]$ , whereas it is pure noise with probability  $1 - p$ . By pure noise we mean a random variable that is identically distributed as, yet stochastically independent from, the true state. The agent does not know whether the realization of the signal is true or noise.<sup>5</sup> Thus, the “precision”  $p$  is a measure for the agent’s ability. We refer to  $p$  as the agent’s ability type, and assume that the type can be either “high” or “low”:  $p \in \{h, \ell\}$ ,  $h > \ell$ . The probability with which the agent’s type is  $p$  is commonly known by the parties and denoted by  $\mu_p \in [0, 1]$ .

The parties have state-dependent preferences and disagree about the ideal action to be taken. We adopt the familiar quadratic utility representation: if action  $x$  is taken in state  $\theta$ , the principal’s utility is  $-(x - (\theta - b))^2$ , and the agent’s utility is  $-(x - \theta)^2$ , where the “bias”  $b > 0$  measures the conflict of interest between the parties.<sup>6,7</sup> To avoid

---

<sup>5</sup>This signal structure is sometimes referred to as the “replacement noise” model.

<sup>6</sup>This setting is the leading example in Crawford and Sobel (1982) and has also been extensively used in the literature. See, for example, Alonso, Dessein, and Matouschek (2008), Dessein (2002), Blume, Board, and Kawamura (2008), Goltsmann et al. (2010), Kartik, Ottaviani, and Squintani (2007), Ottaviani and Squintani (2006).

<sup>7</sup>The literature frequently ascribes the bias to the agent in the sense that the principal’s ideal action is the state  $\theta$ , and the agent’s ideal action is  $\theta + b$ . As will become clear below, our formulation is more convenient for our purposes.

uninteresting case distinctions, we assume in the whole paper that<sup>8</sup>

$$b \leq \ell/2. \quad (1)$$

At the outset, the principal can *commit* to a mechanism which specifies an action to be taken after the agent has observed the signal. The principal's objective is to design a mechanism which maximizes her expected utility.

Before we describe the principal's design problem in detail, we argue that in our setup with quadratic preferences, the expected state conditional on the signal is a sufficient statistics for the signal. In what follows, it will be more convenient to work with this sufficient statistics. Let

$$\tilde{\omega} \equiv E[\tilde{\theta} \mid \tilde{s}] = p\tilde{s} + (1-p) \cdot 1/2 \quad (2)$$

be the expected state conditional on the signal  $\tilde{s}$ . Observe that since the state is uniformly distributed, the random variable  $\tilde{\omega}$  is uniformly distributed on the support  $[\underline{\omega}_p, \bar{\omega}_p]$ , where the support endpoints are given by

$$\underline{\omega}_p = 1/2 - p/2, \quad \bar{\omega}_p = 1/2 + p/2.$$

We denote by  $f_p$  the density of  $\tilde{\omega}$  and  $F_p$  the corresponding cumulative distribution function.

Since preferences are quadratic, a party's expected utility, conditional on the signal, is of the mean-variance form: for  $z \in \{0, b\}$ ,

$$E[-(x - (\tilde{\theta} - z))^2 \mid s] = -(x - (\omega - z))^2 + Var(\tilde{\theta} \mid s),$$

where  $Var(\tilde{\theta} \mid s)$  is the conditional variance of the state, conditional on signal realization  $s$ . Because the conditional variance enters utility only as a constant, a party's marginal utility from an action depends only on the expected state  $\omega$ , yet not on  $s$ . We can thus normalize utility by subtracting the variance. Therefore, our model is isomorphic to the model where the agent first privately observes his ex ante type  $p$  and then privately observes the expected state  $\omega$  as specified in (2), and where the principal's and the agent's utilities are respectively given by  $-(x - (\omega - b))^2$ , and  $-(x - \omega)^2$ .

The following timing summarizes the description of the model.

0. The principal commits to a mechanism.

---

<sup>8</sup>This assumption will guarantee that when the agent's ability is publicly known, the optimal delegation set is a non-degenerate interval. See Lemma 4 for more details. In the terminology of Alonso and Matouscheck (2008), assumption (1) means that delegation is valuable when facing the low type.



1. The agent privately observes his ability type  $p \in \{h, \ell\}$ .
2. The agent privately observes the expected state  $\omega \in [\underline{\omega}_p, \bar{\omega}_p]$ .
3. An action is implemented according to the terms of the mechanism.

We have chosen to study the specific environment with uniformly distributed state, quadratic preferences, constant bias, and two ability types, because it is tractable and allows us to fully solve the model for all parameter constellations. The main question of the paper is to what extent the principal can exploit the fact that the agent receives his information sequentially. In particular, we ask when the principal can do strictly better by using a “sequential” mechanism, which conditions on the agent’s ability type, than by using the optimal “static” mechanism, which does not condition on ability type. To address this question, it is necessary to be able to solve for the optimal “static” mechanism. As will become clear below, this is already demanding in our setup and is likely to be even more so in more general environments. Throughout the analysis we will explicitly point out where we make use of our specific assumptions and explain the difficulties in going beyond them.

### 3 The principal’s problem

The principal’s objective is to design a mechanism which maximizes her expected utility. In this section, we describe the principal’s problem formally. Since the agent has private information, the action implemented by the mechanism optimally depends on communication by the agent to the principal. By the revelation principle for sequential games (Myerson, 1986), the optimal mechanism is in the class of direct and incentive compatible mechanisms. A direct mechanism requires the agent to report his private information as soon as it has arrived. Formally, a *direct* mechanism

$$M \equiv \{\xi_{\hat{p}}(\hat{\omega}) \in \mathbb{R} \mid \hat{p} \in \{h, \ell\}, \hat{\omega} \in [\underline{\omega}_p, \bar{\omega}_p]\}$$

requires the agent to first submit a report  $\hat{p}$  about his ability and then a report  $\hat{\omega}$  about the expected state and then implements the action  $\xi_{\hat{p}}(\hat{\omega})$ .<sup>9</sup>

If the agent’s ex post type is  $\omega$  and his period 1 report was  $\hat{p}$ , then his utility from reporting  $\hat{\omega}$  in period 2 is  $-(\xi_{\hat{p}}(\hat{\omega}) - \omega)^2$ . The mechanism is *incentive compatible in period 2* if it gives the agent an incentive to report the expected state truthfully, conditional on

---

<sup>9</sup>We restrict attention to deterministic mechanisms. For an investigation of stochastic mechanisms in the static delegation problem, see Kováč and Mylovanov (2009).

having reported his ability truthfully. That is, if for all  $p \in \{h, \ell\}$  and  $\omega, \hat{\omega} \in \Omega_p$ ,

$$-(\xi_p(\omega) - \omega)^2 \geq -(\xi_p(\hat{\omega}) - \omega)^2. \quad (3)$$

Observe that the revelation principle does not require truth-telling off the equilibrium path, i.e. after a lie in period 1.<sup>10</sup>

Next, we turn to incentive compatibility in period 1. If the agent's ability is  $p$ , the expected state is uniformly distributed on  $[\underline{\omega}_p, \bar{\omega}_p]$ . Thus, the agent's expected utility from reporting  $\hat{p}$  when his true ability is  $p$  is.

$$\int_{\underline{\omega}_p}^{\bar{\omega}_p} \max_{\hat{\omega}} [-(\xi_{\hat{p}}(\hat{\omega}) - \omega)^2] \frac{1}{p} d\omega.$$

Again, because truth-telling is not required after a lie in period 2, the agent may find it optimal to lie again in period 2 after a lie in period 1. Thus the “max” operator under the integral.

The mechanism is *incentive compatible in period 1* if it gives the agent an incentive to report his type truthfully in period 1. Observe that if the agent tells the truth in period 1, then second period incentive compatibility guarantees truth-telling in period 2. Hence, the mechanism is incentive compatible in period 1 if for all  $p, \hat{p}$ ,

$$\int_{\underline{\omega}_p}^{\bar{\omega}_p} -(\xi_p(\omega) - \omega)^2 \frac{1}{p} d\omega \geq \int_{\underline{\omega}_p}^{\bar{\omega}_p} \max_{\hat{\omega}} -(\xi_{\hat{p}}(\hat{\omega}) - \omega)^2 \frac{1}{p} d\omega. \quad (4)$$

The principal's (conditional) expected utility under an incentive compatible mechanism  $M$  when the agent's ability is  $p$  is

$$V_p(M) \equiv \int_{\underline{\omega}_p}^{\bar{\omega}_p} -(\xi_p(\omega) - (\omega - b))^2 \frac{1}{p} d\omega,$$

and her ex ante expected utility is  $V(M) \equiv \mu_h V_h(M) + \mu_\ell V_\ell(M)$ . The principal's problem, referred to as  $\mathcal{M}$ , can therefore be stated as follows:

$$\mathcal{M} : \quad \max_M V(M) \quad \text{s.t.} \quad (3), (4).$$

---

<sup>10</sup>The reason is that incentive compatibility in period 1 implies that a lie in period 1 is a zero probability event and, thus, what happens afterwards does not affect the principal's utility. On the other hand, allowing for lying off the path may increase the set of implementable outcomes (see Myerson, 1986). See also Krämer and Strausz (2008) for an elaboration of this point in the context of a principal agent problem with money.

## 4 Sequential Delegation

It is well-known that in the static analogue to our problem, when the agent knows the true state at the outset, the principal's problem can be equivalently stated as a "delegation problem" where instead of requiring a report by the agent and implementing the action herself, the principal offers the agent a set of actions from which he can freely choose.

In this section, we argue that also the mechanism design problem  $\mathcal{M}$  can be equivalently stated as a *sequential* delegation problem where the principal offers the agent a *menu*  $(D_h, D_\ell)$  of delegation sets, and the agent chooses a delegation set from the menu in period 1 and picks an action from the chosen delegation set in period 2.

To see that the problem  $\mathcal{M}$  is equivalent to a sequential delegation problem, consider a menu of delegation sets and observe that, if the agent has chosen  $D_{\hat{p}}$  in period 1, then he chooses in period 2 the action<sup>11,12</sup>

$$x_{\hat{p}}(\omega) \equiv \arg \max_{x' \in D_{\hat{p}}} -(x' - \omega)^2. \quad (A)$$

We call a menu  $(D_h, D_\ell)$  of delegation sets *incentive compatible* if the agent of any type  $p \in \{h, \ell\}$  chooses the delegation set  $D_p$  from the menu, i.e. if

$$-\int_{\underline{\omega}_p}^{\bar{\omega}_p} (x_p(\omega) - \omega)^2 \frac{1}{p} d\omega \geq -\int_{\underline{\omega}_p}^{\bar{\omega}_p} (x_{\hat{p}}(\omega) - \omega)^2 \frac{1}{p} d\omega \quad \text{for all } \hat{p}. \quad (IC_p)$$

We now argue that any outcome that can be implemented by a direct, incentive compatible mechanism  $M$  can also be implemented by an incentive compatible menu of delegation sets, and vice versa. Indeed, for an incentive compatible mechanism  $M$ , define the menu of delegation sets by the set of all possible actions that can arise under the mechanism:  $D_p = \{\xi_p(\omega) \mid \omega \in [\underline{\omega}_p, \bar{\omega}_p]\}$ ,  $p \in \{h, \ell\}$ . Then it follows by (3) that the choice function defined in (A) satisfies  $x_p(\omega) = \xi_p(\omega)$ , and (4) implies that the menu  $(D_h, D_\ell)$  is incentive compatible and, thus, implements the same outcome as the direct mechanism  $M$ . Reversely, given an incentive compatible menu  $(D_h, D_\ell)$  of delegation sets, define the mechanism  $M$  by the choice function given in (A):  $\xi_p(\omega) \equiv x_p(\omega)$ . Then the mechanism is (trivially) incentive compatible in period 2, and  $(IC_p)$  implies that  $M$  is also incentive compatible in period 1.

Therefore, we can state the principal's problem  $\mathcal{M}$  equivalently as the following se-

---

<sup>11</sup>Without loss of generality, we may assume that delegation sets are closed and bounded. In that case the maximizer is well defined. See footnote 15 in Alonso and Matouscheck (2008).

<sup>12</sup>This notation is somewhat inaccurate. Strictly speaking, the agent's optimal action is a function of the delegation set.

quential delegation problem:

$$\mathcal{D} : \quad \max_{(D_h, D_\ell)} \sum_{p \in \{h, \ell\}} \mu_p \int_{\underline{\omega}_p}^{\bar{\omega}_p} -(x_p(\omega) - (\omega - b))^2 \frac{1}{p} d\omega \quad \text{s.t.} \quad (A), (IC_\ell), (IC_h).$$

In general, the solution to  $\mathcal{D}$  will not be unique. This is so because one can always add redundant actions to the solution which in no state would be chosen by the agent. We therefore restrict attention to minimal optimal delegation menus in the sense that any action in the delegation set is chosen in some state and that unchosen actions are removed.<sup>13,14</sup>

The minimal optimal delegation menu has the property that in the most extreme states, the most extreme actions are chosen, and that there is at most one action outside the support of the state. To state this formally, we introduce the following notation for the minimal and maximal actions in a delegation set  $D_p$  (for  $p \in \{h, \ell\}$ ) which we shall use throughout the paper:<sup>15</sup>

$$\underline{x}_p = \min D_p, \quad \bar{x}_p = \max D_p.$$

**Lemma 1.** *Consider a minimal optimal delegation menu  $(D_h, D_\ell)$  and let  $p \in \{h, \ell\}$ . Then  $x_p(\underline{\omega}_p) = \underline{x}_p$  and  $x_p(\bar{\omega}_p) = \bar{x}_p$ . Moreover, each of the sets  $D_p \cap (-\infty, \underline{\omega}_p)$  and  $D_p \cap (\bar{\omega}_p, +\infty)$  contains at most one action. If so, then it is the action  $\underline{x}_p$  and  $\bar{x}_p$ , respectively.*

The proof of the lemma is straightforward and is, thus, omitted.

## 5 Information and loss-of-control effect

In this section, we explain the fundamental trade-off the principal faces when designing a delegation set, conditional on facing a given ability type, and capture this trade-off formally.

At a fundamental level, granting the agent more discretion involves the trade-off of making better use of the agent's information versus suffering a loss of control. Take a delegation set  $D_p$  and consider the effect of adding a single action  $\bar{x}_p + \varepsilon$  to  $D_p$  (where  $\varepsilon > 0$  is small). On the one hand, for states higher than  $\bar{x}_p + \varepsilon/2 + b$ , the agent's decision

<sup>13</sup>Alonso and Matouscheck (2008) proceed in the same fashion.

<sup>14</sup>Observe that removing these actions does not upset incentive compatibility. In fact, it relaxes incentive compatibility, as removing actions from  $D_{\hat{p}}$  only reduces the agent type  $p$ 's incentives to pick the delegation set  $D_{\hat{p}}$ .

<sup>15</sup>Note that both the minimum and maximum exist, as  $D_p$  is closed and bounded.

is now closer to the principal's ideal decision (the *information effect*). On the other hand, more freedom of action is costly, because for states in  $[\bar{x}_p + \varepsilon/2, \bar{x}_p + \varepsilon/2 + b]$ , the agent's decision moves farther away from the principal's ideal decision which equals state minus bias (the *loss-of-control effect*).

To capture the information and loss-of-control effect formally, we transform the principal's objective by exploiting the second period incentive compatibility constraint (A). Let  $u_p(\omega) = -(x_p(\omega) - \omega)^2$  denote the agent's utility in state  $\omega$  when having picked  $D_p$ . The constraint (A) pins down the agent's utility and, in fact is equivalent, to<sup>16</sup>

$$u_p(\omega_2) - u_p(\omega_1) = \int_{\omega_1}^{\omega_2} 2(x_p(\omega) - \omega) d\omega \quad \forall \omega_1, \omega_2. \quad (5)$$

Therefore, under an incentive compatible delegation menu, the principal's expected utility, conditional on facing type  $p$  (multiplied through by  $p$ ), can be written as

$$\begin{aligned} p \cdot V_p &= \int_{\underline{\omega}_p}^{\bar{\omega}_p} -(x_p(\omega) - \omega)^2 d\omega - b \int_{\underline{\omega}_p}^{\bar{\omega}_p} 2(x_p(\omega) - \omega) d\omega - b^2 \\ &= \underbrace{bu_p(\underline{\omega}_p) + \int_{\underline{\omega}_p}^{\bar{\omega}_p} u_p(\omega) d\omega}_{\text{information effect}} - \underbrace{bu_p(\bar{\omega}_p) - b^2}_{\text{loss-of-control effect}}. \end{aligned} \quad (6)$$

The first and the second term in (6) say that the principal's utility goes up with the agent's utility in the states  $\omega \in [\underline{\omega}_p, \bar{\omega}_p]$ . However, according to the third term, the principal's utility goes down the better off the agent in the highest state  $\bar{\omega}_p$ . Thus, the principal's objective is a weighted average of the agent's utility across states, where the agent's utility in the lowest state receives weight  $b$  and the agent's utility in the highest state receives the *negative* weight  $-b$ .<sup>17</sup>

The decomposition (6) cleanly disentangles the principal's trade-off. The information effect corresponds to the agent's utility gain in states  $[\underline{\omega}_p, \bar{\omega}_p]$ , and the loss-of-control effect corresponds to the agent's utility gain in the state  $\bar{\omega}_p$  (which lowers the principal's utility).

In what follows, we provide an auxiliary lemma which provides conditions how the principal's expected payoff changes when we add a single action to the delegation set. This will allow us below to check the (sub)optimality of delegation sets by evaluating those changes. To state the lemma, we introduce the following critical value which will

<sup>16</sup>This is formally shown in Lemma 1 of Kováč and Mylovanov (2009). Intuitively, because  $x_p(\omega)$  is a maximizer of  $-(x - \omega)^2$ , the envelope theorem implies that  $u'_p(\omega) = 2(x_p(\omega) - \omega)$ .

<sup>17</sup>A similar decomposition, but with weight on a different state, is used by Kováč and Mylovanov (2009).

appear repeatedly throughout the rest of the paper:

$$\beta_p^0 \equiv \bar{\omega}_p - 2b.$$

**Lemma 2.** *Consider an arbitrary menu  $(D_h, D_\ell)$  of delegation sets.<sup>18</sup> Let  $p \in \{h, \ell\}$  and consider an action  $y \notin D_p$  which is non-redundant in the sense that, when added to the delegation set, it would be chosen in some state, i.e.,*

$$\underline{\omega}_p - |\underline{\omega}_p - \underline{x}_p| < y < \bar{\omega}_p + |\bar{\omega}_p - \bar{x}_p|. \quad (7)$$

- (i) *Let  $y < \bar{x}_p$ . Then adding the action  $y$  to  $D_p$  improves the principal's expected utility if  $\frac{1}{2}(y + \bar{x}_p) \leq \bar{\omega}_p$ .*
- (ii) *Let  $y > \bar{x}_p$ . Then adding the action  $y$  to  $D_p$  improves the principal's expected utility, if and only if  $\frac{1}{2}(y + \bar{x}_p) < \beta_p^0$ .*

To understand the lemma, observe first that adding an action that is redundant in the sense that it is not chosen in any state, will leave the principal's expected payoff unchanged. Therefore, we focus only on non-redundant actions as given by condition (7).

Given non-redundancy, part (i) of the lemma describes when it is beneficial to add an action that is smaller than the maximal action. The condition  $\frac{1}{2}(y + \bar{x}_p) < \bar{\omega}_p$  simply means that the new action  $y$  is not chosen by the highest type  $\bar{\omega}_p$ . That this condition is sufficient for an improvement can be readily seen from (6): Since the highest agent type  $\bar{\omega}_p$  does not choose the new action  $y$ , the loss-of-control effect is not affected by including  $y$ . Since the new action is non-redundant, at least one agent type chooses the new action, which implies that adding the action strictly improves the utility of a positive mass of agent types, and thus leads to a strict improvement of the objective via the information effect.

Part (ii) characterizes when it is beneficial to add a new action to the delegation set that is larger than the maximal action. Adding such an action increases both the information effect (as the utility of all agent types in  $[\frac{1}{2}(y + \bar{x}_p), \bar{\omega}_p)$  goes up) and the loss-of-control effect (because the highest type  $\bar{\omega}_p$  chooses the new action). The interplay of these two effects is determined by the location of the new action  $y$ . If  $y$  is close to  $\bar{x}_p$ , the information effect is relatively large, and the the loss-of-control effect is relatively small. When  $\frac{1}{2}(y + \bar{x}_p) = \beta_p^0$ , the two effects are in balance.

Lemma 2 implies the following lemma.

---

<sup>18</sup>The menu of delegation sets is not required to be incentive compatible here.

**Lemma 3.** *Consider an arbitrary menu  $(D_h, D_\ell)$  of delegation sets and  $p \in \{h, \ell\}$ . Let  $\beta_p^0 < y < \bar{x}_p \leq \bar{\omega}_p$ . Then replacing  $D_p$  by  $D'_p = (D_p \cap [-\infty, y]) \cup \{y\}$  improves the principal's expected utility.*

The set  $D'_p$  in Lemma 3 is obtained by chopping off a piece of the upper end of the original delegation set  $D_p$  and keeping the action  $y$  at which the set was chopped off. Lemma 3 says that modifying the delegation set in this way is beneficial if the new maximal action  $y$  lies above  $\beta_p^0$ .

## 6 Benchmarks

In this section, we discuss two benchmark cases that will play an important role in the subsequent analysis. First, we consider the principal's problem when the agent's ability type is publicly known. Second, we consider the optimal "static" delegation set which does not depend on the agent's ability.

### 6.1 Publicly known ability

When the agent's ability is publicly known, the constraint  $(IC_p)$  in  $\mathcal{D}$  is redundant, and the sequential delegation problem reduces to Holmström's (1977, 1984) classic static delegation problem with uniformly distributed state  $\omega$ . It is well-known that in this case, the optimal delegation set is an interval with upper endpoint  $\beta_p^0$ .<sup>19</sup> Note that by assumption (1),  $\underline{\omega}_p < \beta_p^0$ . We summarize this result in the following lemma.

**Lemma 4.** *If the agent's ability  $p$  is publicly known, the solution to the delegation problem is given by  $D_p^0 = [\underline{\omega}_p, \beta_p^0]$ .*

We briefly reiterate the reason for the result. The fact that interval delegation is optimal, can be inferred from (6). Indeed, in state  $\bar{\omega}_p$ , the agent chooses the highest action available.<sup>20</sup> Thus, allowing the agent to pick all smaller actions does not affect the agent's utility in the highest state (and so does not lower the principal's utility), but (weakly) improves the agent's utility in the other states  $[\underline{\omega}_p, \bar{\omega}_p)$  (and so also improves the principal's utility). Extending the interval by including additional higher actions improves the principal's utility via the information effect but lowers the principal's utility via the loss-of-control effect. At the point  $\beta_p^0$ , the two effects are in balance (see Lemma 2, (ii)).

<sup>19</sup>See, for example, Holmström (1977) or Kováč and Mylovanov (2009).

<sup>20</sup>It is not hard to see that offering a delegation set that includes actions larger than  $\bar{\omega}_p$  cannot be optimal.

## 6.2 Optimal static delegation

We refer to a menu of delegation sets as *static* if it can be implemented with a single delegation set  $D$ .<sup>21</sup> In terms of the mechanisms design perspective, this means that both pieces of agent's private information are elicited at the same time.

The principal's utility under a static delegation set is

$$V = \int_{\underline{\omega}_h}^{\bar{\omega}_h} -(x(\omega) - (\omega - b))^2 dF(\omega),$$

where  $x(\omega) = \arg \max_{x \in D} -(x - \omega)^2$  denotes the choice function of the agent under the delegation set  $D$ , and

$$F(\omega) = \mu_h F_h(\omega) + \mu_\ell F_\ell(\omega)$$

is the average distribution function of the expected state, averaged across the two ability types.

Analogously to (6), the principal's utility can be expressed as

$$\begin{aligned} V &= \frac{\mu_h}{h} bu(\underline{\omega}_h) + \frac{\mu_\ell}{\ell} bu(\underline{\omega}_\ell) + \int_{\underline{\omega}_h}^{\bar{\omega}_h} u(\omega) dF(\omega) \\ &\quad - \frac{\mu_h}{h} bu(\bar{\omega}_h) - \frac{\mu_\ell}{\ell} bu(\bar{\omega}_\ell) \\ &\quad - b^2, \end{aligned} \tag{8}$$

where  $u(\omega) = -(x(\omega) - \omega)^2$  is the agent's utility. The key difference to the benchmark case with publicly known types is that, because the delegation set can no longer condition on the agent's ability type, now the loss-of-control effect corresponds to the agent's utility not only in one state but in two states, as reflected in the second line of (8): conditional on facing the  $h$ -type, the principal's utility goes down in the state  $\omega_h$ , and conditional on facing the  $\ell$ -type, the principal's utility goes down in the state  $\omega_\ell$ . Therefore, the simple argument that establishes optimality of interval delegation in the benchmark case no longer works. In fact, consider a delegation set with a maximal action larger than  $\bar{\omega}_\ell$ . Then it would never be beneficial to include all actions below this maximal action but to insert a small gap around the action  $\bar{\omega}_\ell$ , because the agent's utility in state  $\bar{\omega}_\ell$  enters the principal's utility with a negative weight  $-\mu_\ell/\ell \cdot b$ .

Yet, the argument from the benchmark case can be extended as follows. For any delegation set containing two actions which are both below the action  $\bar{\omega}_\ell$ , (8) implies that the principal's utility (weakly) goes up when all actions between the two actions are included in the delegation set. This follows from that fact that, conditional on each type,

<sup>21</sup>For example, the menu of intervals  $D_h = [\bar{\omega}_h, x]$  and  $D_\ell = [\bar{\omega}_\ell, x]$  can be implemented by the single set  $D = [\bar{\omega}_h, x]$ .



adding these actions only improves the information effect yet leaves the loss-of-control effect unchanged. Similarly, for any delegation set containing two actions which are both inside  $(\bar{\omega}_\ell, \bar{\omega}_h)$ , the principal's utility (weakly) goes up when all actions between the two actions are included in the delegation set. Therefore, the optimal delegation set consists of an interval  $[\underline{\omega}_h, x]$ , where  $x < \bar{\omega}_\ell$ , and a set  $D'$  which is either an interval or a single point (within  $(\bar{\omega}_\ell, \bar{\omega}_h)$ ), or empty. The next lemma makes this reasoning more precise.

**Lemma 5.** *The optimal static delegation set is given by the union of an interval  $[\underline{\omega}_h, \bar{\omega}_\ell - d]$  and an additional set  $D'$ , where  $d \in [0, \bar{\omega}_h - \bar{\omega}_\ell]$  and*

$$D' = \emptyset, \quad \text{or} \quad D' = \{\bar{\omega}_\ell + d\}, \quad \text{or} \quad D' = [\bar{\omega}_\ell + d, \bar{x}] \quad (9)$$

for some  $\bar{x} \in [\bar{\omega}_\ell + d, \bar{\omega}_h]$ .

The lemma says that if the optimal static delegation set is not an interval, then it has a symmetric gap around the upper endpoint  $\bar{\omega}_\ell$  of the low type's support. The symmetry of the gap is a consequence of the information effect. If the gap was not symmetric, say with a larger gap below than above the type  $\bar{\omega}_\ell$ , then the type  $\bar{\omega}_\ell$  would choose the action at the upper end of the gap. Thus, closing the gap slightly from below would not affect this type's action and thus would not affect the loss-of-control effect. Yet this would increase improve the principal's expected utility by the information effect.

To identify which of the possible three shapes the optimal static delegation set takes, we use Lemma 2, (ii) which allows us to compare the principal's utility from an interval and from an interval plus a point. It follows from Lemma 5 and Lemma 2, (ii) that interval delegation is optimal when  $\beta_h^0 \leq \bar{\omega}_\ell$ . To see this, observe first that if  $\beta_h^0 \leq \bar{\omega}_\ell$ , the optimal delegation set cannot be the union of two intervals because by Lemma 3 it would be beneficial to lower the upper endpoint of the second interval which is located in  $(\bar{\omega}_\ell, \bar{\omega}_h]$ . Hence, the optimal delegation set is either an interval or an interval plus a point. In the latter case, because the delegation set has a symmetric gap around  $\bar{\omega}_\ell$ , the average of the two largest actions in the delegation set,  $\bar{\omega}_\ell - d$  and  $\bar{\omega}_\ell + d$ , is exactly  $\bar{\omega}_\ell$ . Hence, Lemma 2, (ii) implies that since  $\beta_h^0 \leq \bar{\omega}_\ell$ , it would be better to remove the largest action from the delegation set. These considerations are summarized in Proposition 1. Observe that the condition  $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$  stated in the proposition is equivalent to the condition  $\beta_h^0 \leq \bar{\omega}_\ell$ .

**Proposition 1.** *Let  $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ . Then the optimal static delegation set is an interval of the form*

$$D^{st} = [\underline{\omega}_h, \beta^{st}], \quad \text{with } \beta^{st} \in [\beta_\ell^0, \beta_h^0].$$

Moreover, the upper endpoint  $\beta^{st}$  monotonically increases in  $\mu_h$  with  $\beta^{st} = \beta_\ell^0$  for  $\mu_h = 0$  and  $\beta^{st} = \beta_h^0$  for  $\mu_h = 1$ .

The comparative statics properties of  $\beta^{st}$  are driven by the fact that the principal's objective is the weighted average of her objective if she knew the ability type. Therefore, the optimal upper endpoint is between the optimal upper endpoints of the benchmark cases with known ability types and converges to the benchmark cases as the principal's uncertainty about the agent's ability diminishes.

Next, we consider the case that  $\beta_h^0 > \bar{\omega}_\ell$  (which is equivalent to  $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ ). In this case, the optimal static delegation set consists of an interval plus a point or of two intervals. This is stated in Proposition 2.

**Proposition 2.** *Let  $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$  and  $\mu_h \in (0, 1)$ . Then the optimal static delegation set consists of the disjoint union of an interval and one other set which is either an interval or a single point, and has a symmetric gap around  $\bar{\omega}_\ell$ , i.e., there is a  $d \in [0, \bar{\omega}_h - \bar{\omega}_\ell]$  so that*

$$D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup D',$$

where either  $D' = [\bar{\omega}_\ell + d, \beta_h^0]$  and  $\beta_h^0 > \bar{\omega}_\ell + d$ , or  $D' = \{\bar{\omega}_\ell + d\}$  and  $\beta_h^0 \leq \bar{\omega}_\ell + d$ .

The intuition is analogous to Proposition 1. By Lemma 2, (ii), interval delegation cannot be optimal, because in the case  $\beta_h^0 > \bar{\omega}_\ell$  it would always pay to include the additional action  $\bar{\omega}_\ell + d$  in the delegation set. Whether the optimal delegation set is the union of an interval and a point or of two intervals, can now be determined by maximizing the principal's expected utility with respect to  $d$ . (We omit the somewhat tedious calculation.)

*Remark 1.* It is useful to relate our results on optimal static delegation to the literature. The static problem has, in some generality, been studied by Alonso and Matouscheck (2008) and by Kováč and Mylovanov (2009). In particular, Kováč and Mylovanov (2009) show that the solution is interval delegation if the density associated with the distribution of the state is absolutely continuous, and a “regularity” condition holds. However, we cannot directly apply Kováč and Mylovanov (2009)'s result, because in our case,  $f$  is not absolutely continuous, as it has jumps at the points  $\omega = \underline{\omega}_\ell$  and  $\omega = \bar{\omega}_\ell$ ; thus,  $f$  is only piece-wise absolutely continuous. More seriously, our environment is not regular for all biases. This is the main reason why we consider a setup with two ability types. A major obstacle to go beyond this case is the difficulty to characterize the optimal static delegation set.<sup>22</sup>

---

<sup>22</sup>More precisely, regularity means that the function  $g(\omega) \equiv 1 - F(\omega) - bf(\omega)$  is decreasing at the point  $\omega$  whenever  $0 \leq g(\omega) \leq 1$ . For small bias, the function  $g$  jumps upward to a value larger than 0 at the point  $\bar{\omega}_\ell$ , thus violating regularity. Moreover, the more ability types  $p$  one allows for, the “more” irregular  $g$  becomes. (That  $f$  is only piece-wise continuous causes minor problems only.) This is so since  $g$  makes an upward jump at each point  $\bar{\omega}_p$ . Note that this observation remains true if the state  $\theta$  is not uniformly distributed. One could, of course, allow for more ability types and impose assumptions (e.g. on  $b$ ) that guarantee regularity. However, our analysis below suggests that the comparison between static

## 7 Optimal sequential delegation

In this section, we derive the optimal sequential delegation menu when the agent's ability type is his private information. We say that *static delegation is optimal*, if there exists a static delegation menu that solves the principal's problem. Otherwise, the agent's ability type is explicitly screened, and we say that *sequential delegation is optimal*.

Our approach to solving the principal's problem is to first consider a relaxed problem where we ignore one incentive constraint. We then show that the optimal delegation menu for the relaxed problem must be in a class of simple delegation menus which are pinned down by three parameters. In a second step, we establish that a solution to the relaxed problem automatically satisfies the other incentive constraint and thus also solves the original problem. These two steps will then allow us to determine the optimal delegation menu by optimizing over menus in the class mentioned above.

The benchmark case with publicly known ability types as described in Lemma 4 suggests that we can ignore the high type's incentive constraint. To see this, observe that in the benchmark case, the high type's delegation set is larger than that of the low type:  $D_h^0 = [\underline{\omega}_h, \beta_h^0] \supset [\underline{\omega}_\ell, \beta_\ell^0] = D_\ell^0$ . Thus, the low ability type would have an incentive to pretend to be of high ability. This intuitively suggests that with privately known ability types, the low type's incentive constraint is binding at the optimum.

Thus, we consider the relaxed delegation problem

$$\mathcal{D}^R : \quad \max_{(D_h, D_\ell)} \sum_{p \in \{h, \ell\}} \mu_p \int_{\underline{\omega}_p}^{\bar{\omega}_p} -(x_p(\omega) - (\omega - b))^2 \frac{1}{p} d\omega \quad \text{s.t.} \quad (A), (IC_\ell).$$

where we ignore the  $(IC_h)$  constraint. We denote a solution to  $\mathcal{D}^R$  by  $(D_h^*, D_\ell^*)$ .

We shall now formally confirm the intuition that  $(IC_h)$  is binding, and moreover that the low ability type is offered an interval at the optimum of the relaxed problem.

**Lemma 6.** *At the solution to the relaxed problem  $\mathcal{D}^R$ ,  $(IC_\ell)$  is binding and  $D_\ell^*$  is an interval of the form*

$$D_\ell^* = [\underline{\omega}_\ell, \beta_\ell^*], \quad \text{where } \beta_\ell^* \in [\beta_\ell^0, \bar{\omega}_\ell].$$

The fact that the delegation set for the low type is an interval is a consequence of the effects stated in Lemma 2. Suppose for example that  $D_\ell^*$  contained gaps. Then filling the gaps (adding the actions inside) would clearly improve the  $\ell$ -type's utility from picking the modified delegation set and thus relax the incentive constraint. At the same time, filling gaps would improve the principal's expected utility by Lemma 2, (i).

---

and sequential delegation becomes interesting exactly in irregular cases. Therefore, excluding irregular cases would be a loss.

We shall now exploit the fact that  $(IC_\ell)$  is binding to derive restrictions on the optimal delegation set  $D_h^*$ . Towards this end, we first provide a representation of the optimal delegation menu of the relaxed problem. The following lemma says that the high type's delegation is a union of two connected sets.

**Lemma 7.** *Without loss of generality, the high type's delegation set at the optimum of the relaxed problem  $\mathcal{D}^R$  is a union of two connected sets:  $D_h^* = D_{h1} \cup D_{h2}$ , where:*

- (i)  $D_{h1} = \{\gamma_1\}$  with  $\gamma_1 \leq \underline{\omega}_h$ , or  $D_{h1} = [\underline{\omega}_h, \gamma_1]$  with  $\gamma_1 > \underline{\omega}_h$ .
- (ii)  $D_{h2} = \{\gamma_2\}$ , or  $D_{h2} = [\gamma_2, \beta_h^0]$  with  $\bar{\omega}_\ell \leq \gamma_2 < \beta_h^0$ .
- (iii)  $\gamma_1 \leq \beta_\ell^* \leq \gamma_2$  and  $\frac{1}{2}(\beta_\ell^* + \gamma_2) \leq \bar{\omega}_\ell$ .

Part (i) says that  $D_{h1}$  is a singleton or an interval. The first case occurs only if  $D_h^*$  contains an action which is smaller than the lower endpoint of the support. Otherwise, the second case occurs. Part (ii) says that also  $D_{h2}$  is either a singleton or an interval. Importantly, it can be an interval only if  $\bar{\omega}_\ell < \beta_h^0$ , or equivalently  $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ . This is precisely the case when interval delegation is not optimal in the static case. In the other case,  $D_{h2}$  is a singleton, and thus  $D_h^*$  may be represented as a union of an interval and an action or as a two-action set.

Part (iii) restricts the location of the sets  $D_{h1}$  and  $D_{h2}$ . The first inequality says that the two actions  $\gamma_1, \gamma_2$  cannot be jointly larger or smaller than the highest action of the  $\ell$ -type's delegation set,  $\beta_\ell^*$ . The second inequality says that in state  $\bar{\omega}_\ell$ , the agent (at least weakly) prefers the action  $\gamma_2$  over the action  $\beta_\ell^*$ .

Let us emphasize that the optimal static delegation sets from Propositions 1 and 2 are included in the class of menus described in Lemma 7: Interval delegation occurs if  $\gamma_1 = \gamma_2 = \beta_\ell^*$  and  $D_{h2}$  is a singleton. Moreover, the static delegation set from Proposition 2 occurs if  $\gamma_1 = \beta_\ell^*$  and  $\gamma_2 = \bar{\omega}_\ell + d$  with  $d = \bar{\omega}_\ell - \beta_\ell^*$ .

Finally, note that Lemma 7 does not claim that  $D_h^*$  is of the specified form in *every* optimal delegation menu. Rather, it states that for any optimal delegation menu  $(D_h^*, D_\ell^*)$  there is a delegation set  $D'_h$  of the form described in Lemma 7 such that the menu  $(D'_h, D_\ell^*)$  is optimal as well. The intuition is as follows. Given the delegation menu  $(D_h^*, D_\ell^*)$ , we construct the set  $D'_h$  so that it (weakly) improves the principal's expected utility and at the same time preserves the binding  $(IC_\ell)$ .<sup>23</sup> More specifically, let us denote  $\gamma_2$  the action that the agent chooses from  $D_h^*$  in state  $\bar{\omega}_\ell$  (if the agent is indifferent between two actions, let  $\gamma_2$  be the higher one). Then, type  $\ell$  would (weakly) prefer the interval  $[\underline{x}_h, \gamma_2]$  to  $D_h^*$ . On the other hand, type  $\ell$  would be (weakly) worse off from the two

<sup>23</sup>The construction actually applies to any delegation menu where  $(IC_\ell)$  is binding.

point set  $\{\underline{x}_h, \bar{x}_h\}$  than from  $D_\ell^*$ . Therefore, an intermediate value argument implies that there is  $\gamma_1 \in [\underline{x}_h, \gamma_2]$  such that type  $\ell$  remains indifferent between the set  $D_\ell^*$  and the set  $[\underline{x}_h, \gamma_1]$  in between the large and the small set. Let us now consider the delegation set  $D'_h$  where all actions in  $D_h^*$  below  $\gamma_2$  are replaced by  $[\underline{x}_h, \gamma_1]$ . This clearly leaves  $(IC_\ell)$  binding. The question is then how the principal's expected utility is affected by the change. While the value of  $D_h^*$  and  $D'_h$  is the same based on the *low* agent type's belief about the expected state, all that matters for the principal is the utility from  $D_h^*$  and  $D'_h$ , conditional on facing the *high* ability type. But observe that the conditional distribution of the expected state is uniform conditional on *both* the high and the low ability type. Therefore, the principal effectively evaluates  $D_h^*$  and  $D'_h$  with the same beliefs as the low agent type. Thus, if the change does not affect the agent's choices in states  $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell]$ , also the principal's expected utility is unaffected by the change.<sup>24</sup> On the other hand, if it does affect the agent's choices in states  $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell]$ , the delegation set  $D'_h$  may even improve the principal's expected utility.

Lemma 7 characterizes the shape of the optimal delegation menu for the relaxed problem. We now show that any delegation menu in the class described in Lemma 7 automatically satisfies the incentive compatibility constraint for type  $h$ . Therefore, a solution to the relaxed problem is automatically a solution to the original problem.

**Lemma 8.** *A solution  $(D_h^*, D_\ell^*)$  to the relaxed problem  $\mathcal{D}^R$  of the form as in Lemma 7, is also a solution to the original problem  $\mathcal{D}$ .*

The basic idea behind Lemma 8 is as follows. Suppose there is a sequential solution  $(D_h^*, D_\ell^*)$  to the relaxed problem that violates  $(IC_h)$ .<sup>25</sup> Then, since  $(IC_h)$  is violated, the  $h$ -type's utility would be improved when the agent was offered the static menu consisting of  $D'_h = [\underline{\omega}_h, \underline{\omega}_\ell] \cup D_\ell^* = [\underline{\omega}_h, \bar{x}_\ell]$  and  $D'_\ell = D_\ell^*$ .<sup>26</sup> In terms of the principal's utility, by (6), a modification that improves the  $h$ -type's utility improves the objective by the information effect. Moreover, since  $\bar{x}_h \geq \bar{x}_\ell$  by Lemma 7, the utility of the highest type  $\bar{\omega}_h$  is higher under  $D_h^*$  than under  $D'_h$ , and therefore, the loss-of-control effect is smaller under  $D'_h$  than under  $D_h^*$ . Consequently, if  $(IC_h)$  is violated, offering the static menu is  $(D'_h, D'_\ell)$  is an improvement in terms of both the information and the loss-of-control effect.

Lemma 7 and Lemma 8 imply that the optimal delegation menu can be found by optimizing over the class described in Lemma 7. By part (ii) of Lemma 7, the properties

<sup>24</sup>If the underlying state  $\theta$  is not uniformly distributed, a change of  $D_h^*$  that leaves the low type indifferent does not necessarily leave the principal indifferent, because the conditional distributions of the expected state, conditional on the various ability types, are not linear transformations of one another. It is an open question to what extent the previous argument extends to the non-uniform case.

<sup>25</sup>Clearly, a static solution meets  $(IC_h)$ .

<sup>26</sup>This static menu can also be implemented by offering the delegation set  $D'_h = [\underline{\omega}_h, \beta_\ell^*]$  to both types.

of the  $h$ -type's optimal delegation set  $D_h^*$  depend on whether, for the static problem, interval delegation is optimal or not. It is therefore useful to make this case distinction.

### 7.1 Large bias: $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$

We now consider the set of biases so that interval delegation is optimal in the static problem. Recall that this is the case if  $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ , or equivalently,

$$\beta_h^0 \leq \bar{\omega}_\ell. \quad (10)$$

Propositions 3 and 4 below fully characterize the solution to the principal's problem. The propositions say that sequential delegation is optimal if and only if the bias is small (relative to  $\bar{\omega}_h - \bar{\omega}_\ell$ ) and the prior probability of the high type,  $\mu_h$ , is not too large. Lemma 7 implies that in this case  $D_h^*$  is a union of an interval and an isolated point, or it is a two-action set. Otherwise, static delegation remains to be optimal.

**Proposition 3.** *Let  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$ . Then static interval delegation is optimal for the problem  $\mathcal{D}$ . The optimum is achieved at the delegation set  $D^{st} = [\underline{\omega}_h, \beta^{st}]$ .*

**Proposition 4.** *Let  $\frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell) \leq b < \bar{\omega}_h - \bar{\omega}_\ell$ , and define<sup>27</sup>*

$$\hat{b} \equiv \frac{1 + \sqrt{5}}{4} (\bar{\omega}_h - \bar{\omega}_\ell). \quad (11)$$

*Then for every  $b \in [\frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell), \hat{b})$  there is  $\hat{\mu}_h(b) \in (0, 1]$  such that:*

- (a) *If  $b \geq \hat{b}$ , or if  $b < \hat{b}$  and  $\mu \geq \hat{\mu}_h(b)$ , then static delegation is optimal for the problem  $\mathcal{D}$ . The optimum is achieved at the delegation set  $D^{st} = [\underline{\omega}_h, \beta^{st}]$ .*
- (b) *If  $b < \hat{b}$  and  $\mu < \hat{\mu}_h(b)$ , then sequential delegation is optimal for the problem  $\mathcal{D}$ .<sup>28</sup>*

To illuminate the forces behind the result, we shall identify the costs and benefits of sequential delegation. For a fixed delegation set of the low type, the principal's choice of the high type's delegation set is restricted by the incentive compatibility requirement that  $(IC_\ell)$  be binding. This pins down the agent's utility from the high type's delegation set, aggregated over the range of the low type's support  $[\underline{\omega}_\ell, \bar{\omega}_\ell]$ :

$$\ell \cdot U_\ell = \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_\ell(\omega) d\omega = \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_h(\omega) d\omega. \quad (12)$$

<sup>27</sup>Observe that  $\frac{1}{2}(1 + \sqrt{5}) \approx 1.62$  is the *Golden ratio*.

<sup>28</sup>By Lemmata 6 and 7, at the optimum,  $D_\ell$  is an interval, and  $D_h$  is a union of an interval and an isolated point, or it is a two-action set.

Therefore, by (6), the principal's expected utility, conditional on facing the  $h$ -type, can be expressed as<sup>29</sup>

$$h \cdot V_h = \ell \cdot U_\ell + \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega + bu_h(\underline{\omega}_h) - bu_h(\bar{\omega}_h) - b^2. \quad (13)$$

Again, we can identify an information and a loss-of-control effect. The information effect corresponds to the agent's utility gain in states  $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell) \cup (\bar{\omega}_\ell, \bar{\omega}_h)$ , and the loss-of-control effect corresponds to the agent's utility gain in the state  $\bar{\omega}_h$  (which lowers the principal's utility). Observe that, compared to the benchmark case with publicly known ability types, since  $U_\ell$  is fixed by incentive compatibility, the information effect is diminished, as the principal now only benefits from the agent's utility in the states  $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell) \cup (\bar{\omega}_\ell, \bar{\omega}_h)$ , but the loss-of-control effect is the same as in the benchmark case. In this sense, the cost of incentive compatibility is that it weakens the information effect with respect to the high ability type.

To understand the effects of sequential delegation more specifically, it is useful to consider marginal changes of the optimal static delegation set. Recall that for the case under consideration in this subsection ( $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ ), the optimal static delegation set is an interval  $[\underline{\omega}_h, \beta^{st}]$  by Proposition 1. Now suppose that, instead of offering the static menu  $D_h = D_\ell = [\underline{\omega}_h, \beta^{st}]$ , the principal offers the sequential delegation menu where  $D_\ell$  is maintained and the upper endpoint of  $D_h$  is slightly increased by  $d\beta^{st}$  and an appropriate gap  $dz$  is inserted in  $D_h$  to maintain incentive compatibility:

$$D'_\ell = D^{st}, \quad D'_h = [\underline{\omega}_h, \beta^{st} - dz] \cup \{\beta^{st} + d\beta^{st}\}. \quad (14)$$

We refer to this modification as *marginal sequential delegation*. Since marginal sequential delegation is feasible for the relaxed problem  $\mathcal{D}^R$ , it follows that if marginal sequential delegation is beneficial (relative to static delegation), then some sequential delegation menu is optimal for the relaxed problem, and thus, by Lemma 8 also for the original problem. We now investigate when marginal sequential delegation is beneficial.

We may use (13) to evaluate the effect of marginal sequential delegation on the principal's utility, conditional on facing the  $h$ -type. Since  $\beta^{st}$  is increased only marginally, the gap necessary to maintain incentive compatibility is also only marginal. In particular, as under static delegation, all actions  $[\underline{\omega}_h, \underline{\omega}_\ell)$  will be included in the  $h$ -type's modified delegation set. Hence, marginal sequential delegation does not affect that part of the information effect that is attributable to the agent's utility in the states  $[\underline{\omega}_h, \underline{\omega}_\ell)$  (the sec-

---

<sup>29</sup>It is important to note that this follows from the assumption that the state is uniformly distributed. The binding ( $IC_\ell$ ) uses the distribution conditional on facing type  $\ell$ . On the other hand, the above formula involves the distribution conditional on facing type  $h$ . If the types are distributed according to the uniform distribution, the two conditional distributions differ only by a constant factor.

ond and fourth term in (13)). Yet it does affect both what remains from the information effect (the third term in (13)) and the loss-of-control effect (the fifth term in (13)).

Since the upper endpoint,  $\beta^{st}$ , of the optimal static delegation set is smaller than  $\beta_h^0$  by Proposition 1, it is also smaller than  $\bar{\omega}_\ell$  by condition (16). Therefore, in the states  $[\bar{\omega}_\ell, \bar{\omega}_h]$ , the agent chooses the highest action from the  $h$ -type's delegation set both under static and under marginal sequential delegation. Together with the argument in the previous paragraph, this implies that the benefit of marginal sequential delegation can be expressed as

$$\frac{d(hV_h)}{d\beta^{st}} = \frac{d}{d\beta^{st}} \left\{ \underbrace{\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} -(\beta^{st} - \omega)^2 d\omega}_{\text{information effect}} - \underbrace{b[-(\beta^{st} - \bar{\omega}_h)^2]}_{\text{loss-of-control effect}} \right\}. \quad (15)$$

Observe that both the information and the loss-of-control effect go up through marginal sequential delegation, as the agent now obtains a higher utility in all relevant states  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h]$ . Hence, what matters is how the two effects increase *relative* to one another. Three forces drive the relative size of the information and the loss-of-control effect.

First, because the agent's utility is quadratic, the agent's utility increases by more in state  $\bar{\omega}_h$  than in any state  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h)$  when  $\beta^{st}$  is marginally increased. Yet, the *difference* in the increase declines the smaller is the action  $\beta^{st}$ . This is because the agent's loss function is concave so that the difference between the marginal loss in state  $\bar{\omega}_h$  and in some state  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h)$  diminishes as the loss becomes large. Therefore, the smaller is  $\beta^{st}$ , the more the information effect works in favor of marginal sequential delegation. Recall from Proposition 1 that  $\beta^{st}$  increases in the likelihood  $\mu_h$  of the  $h$ -type.

Second, the loss-of-control effect is evidently less pronounced, the smaller is the bias. Thus, the smaller is  $b$ , the less the loss-of-control effect works against marginal sequential delegation. Third, the information effect is more pronounced the larger is the interval  $[\bar{\omega}_\ell, \bar{\omega}_h)$ . Thus, the larger is  $\bar{\omega}_h - \bar{\omega}_\ell$ , the more the information effect works in favor of marginal sequential delegation.

We now explain how these effects play out for the cases distinguished in Propositions 3 and 4. Consider first Proposition 3, i.e.,  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$ . As said above, a marginal increase of  $\beta^{st}$  increases the agent's utility by more in state  $\bar{\omega}_h$  than in states  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h)$ . Therefore, the agent's utility increase in state  $\bar{\omega}_h$ , multiplied by  $b$ , is larger than the agent's aggregate utility increase over the range  $[\bar{\omega}_\ell, \bar{\omega}_h)$  which has measure  $\bar{\omega}_h - \bar{\omega}_\ell$  smaller than  $b$ . Thus, the loss-of-control effect dominates the information effect. This shows that, for  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$ , the principal cannot benefit from marginal sequential delegation. While this argument deals only with marginal modifications of the optimal static delegation menu, the proof



of Proposition 3 shows that, in fact, no profitable modification of the optimal static delegation menu exists when  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$ .<sup>30</sup>

Next, consider Proposition 4, where  $b < \bar{\omega}_h - \bar{\omega}_\ell$ . As argued above, the benefit of marginal sequential delegation is large when  $b$  and  $\mu_h$  are small and  $\bar{\omega}_h - \bar{\omega}_\ell$  is large. A computation shows that at the values  $b = \hat{b}$  and  $\mu_h = \hat{\mu}_h$ , the benefit of marginal sequential delegation (15) is zero. Thus, for parameter constellation (b), marginal sequential delegation improves upon static delegation, and a fortiori, some sequential delegation menu must be the solution to the principal's problem.

Likewise, for the parameter constellation under (a), marginal sequential delegation does not improve upon static delegation. In principle, therefore, a non-marginal modification of the static delegation menu could be profitable. That this is not the case, can intuitively be seen from (13) and the insight from Lemma 7 that the optimal delegation set for the  $h$ -type is an interval plus an action, or it is a two-action set. Because a marginal modification is not profitable, any increase of the upper endpoint of the  $h$ -type's delegation set increases the agent's utility in state  $\bar{\omega}_h$  (times  $b$ ) by more than the agent's aggregate utility over the states  $[\bar{\omega}_\ell, \bar{\omega}_h)$ . Hence, by (13), any potential gain from a non-marginal modification must come from the effect it has on the agent's utility in state  $\underline{\omega}_h$  and on the agent's aggregate utility over the states  $(\underline{\omega}_h, \underline{\omega}_\ell)$ . But the agent's utility in those states is already maximized under marginal sequential delegation. Hence, there is no room for a non-marginal modification to yield the principal additional utility.

This concludes our discussion of the case where the agent's bias is relatively large. Next, we turn to the case with small bias.

## 7.2 Small bias: $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$

We now consider the set of biases so that interval delegation is no longer optimal in the static problem. Recall that this is the case if  $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ , or equivalently,

$$\beta_h^0 > \bar{\omega}_\ell. \quad (16)$$

We show that in this case, sequential delegation is always strictly better than static delegation.

**Proposition 5.** *Let  $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ . Then sequential delegation is optimal for the principal's problem  $\mathcal{D}$ .*<sup>31</sup>

<sup>30</sup>Interestingly, the condition  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$  stated in Proposition 3 is exactly the regularity condition established in Kováč and Mylovanov (2009). It is an open question of whether, in general, regularity is sufficient for the optimality of static delegation.

<sup>31</sup>By Lemmata 6 and 7, at the optimum,  $D_\ell$  is an interval, and  $D_h$  is a union of an interval and an isolated point, or it is the disjoint union of two intervals, or it is a two-action set.

The intuition can again be best understood in terms of the information and the loss-of-control effect as captured by expression (13) which, for convenience, we reiterate:

$$h \cdot V_h = \ell \cdot U_\ell + \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega + bu_h(\underline{\omega}_h) - bu_h(\bar{\omega}_h) - b^2. \quad (17)$$

Recall that by Proposition 2, the optimal static delegation set is the disjoint union of an interval and one other set which is either an interval or a single point, and has a symmetric gap around  $\bar{\omega}_\ell$ . For the sake of the discussion, suppose it is the union of an interval and a single point (the argument for the other case is identical):

$$D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup \{\bar{\omega}_\ell + d\}, \quad \text{with } \bar{\omega}_\ell + d \leq \bar{\omega}_h. \quad (18)$$

Consider a modification of the optimal static delegation menu where the delegation set of the low type is the same as under static delegation except that the action  $\{\bar{\omega}_\ell + d\}$  is removed:  $D_\ell = [\underline{\omega}_h, \bar{\omega}_\ell - d]$ . Further, the delegation set of the high type is modified by chopping off a piece of the upper end of the interval and adding a small interval to the left of the upper endpoint so that  $(IC_\ell)$  remains binding:

$$D_h = [\underline{\omega}_h, \bar{\omega}_\ell - d - \eta] \cup [\bar{\omega}_\ell + d - \varepsilon, \bar{\omega}_\ell + d], \quad \eta, \varepsilon > 0. \quad (19)$$

We now argue that this modification is profitable for the principal (relative to static delegation). Since the modification is feasible for the relaxed problem  $\mathcal{D}^R$ , some sequential delegation menu is therefore optimal for the relaxed problem, and thus, by Lemma 8, also for the original problem.

To see that the modification is profitable, observe that because  $D_\ell$  and  $D^{st}$  induce the same action for the  $\ell$ -type agent, the modification does not affect the principal's utility conditional on facing the  $\ell$ -type. The effect on the principal's utility conditional on facing the  $h$ -type can be seen from (17). Because the agent's utility in state  $\bar{\omega}_h$  has not changed under the modification, the loss-of-control effect — the term  $bu_h(\bar{\omega}_h)$  in (17) — is unaffected by the modification.

Consider next the impact of the modification on the information effect — the second to fourth terms in (17). Suppose that  $\varepsilon$  (and thus  $\eta$ ) is so small that  $\bar{\omega}_\ell - d - \eta > \underline{\omega}_\ell$  (we may always find such an  $\varepsilon$ ). Then the information effect that is attributable to the agent's utility in the states  $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell]$  — the second and fourth term in (17) — is not affected by the modification. But, the information effect that is attributable to the agent's utility in the states  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h)$  — the third term in (17) — *does* increase under the modification. To see this, observe that since the modification maintains  $(IC_\ell)$ , it must be that in the states  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_\ell + d - \varepsilon]$ , the agent chooses the action  $\bar{\omega}_\ell + d - \varepsilon$  under

the modification, whereas under static delegation he chooses the worse action  $\bar{\omega}_\ell + d$ . Moreover, in the states  $\omega \in [\bar{\omega}_\ell + d - \varepsilon, \bar{\omega}_\ell + d)$ , the agent gets his ideal action  $\omega$  under the modification, whereas under static delegation he chooses the worse action  $\bar{\omega}_\ell + d$ . Finally, in the states  $\omega \in [\bar{\omega}_\ell + d, \bar{\omega}_h]$  the agent's action under the modification is the same as under static delegation. In sum, the principal's expected utility unambiguously goes up under the modification, and thus, a sequential delegation menu is strictly better than static delegation.

## 8 Conclusion

In this paper, we have provided insights into the nature of optimal delegation when the agent's private information arrives over time. While we focus on a special environment to obtain a full characterization, the intuitions developed in the paper suggest that many of our findings are likely to carry over to more general settings. In particular, it seems likely that in many environment offering delegation menus with only extreme options is necessary to screen the agent's ex ante ability. Also, whenever the principal faces a trade-off between information and loss of control, it is reasonable to conjecture that a smaller conflict of interest and a larger uncertainty about the agent's ex ante ability work in favor of sequential delegation.

The design approach adopted in our paper requires that actions are contractible. Starting with Crawford and Sobel (1982), there is a huge literature that investigates the static problem when actions are not contractible and the parties have to rely on non-binding forms of communication such as cheap talk. An interesting question for future research that can be addressed within the model of this paper is how the nature of non-binding communication is affected by the sequential arrival of new information.

## A Appendix

*Proof of Lemma 2.* (i) By assumption,  $\frac{1}{2}(y + \bar{x}_p) \leq \bar{\omega}_p$ . Therefore, adding  $y$  does not modify the agent's utility in state  $\bar{\omega}_p$ . In addition, (7) implies that some agent type  $\omega \in [\underline{\omega}_p, \bar{\omega}_p]$  chooses  $y$ . Hence, adding  $y$  strictly improves the agent's utility in some neighborhood of  $y$  when  $y > \underline{\omega}_p$ , or in some neighborhood of  $\underline{\omega}_p$  when  $y \leq \underline{\omega}_p$ . By (6), these observations imply that adding  $y$  strictly improves the principal's utility.

(ii) Let us first consider the case  $\bar{x}_p \geq \underline{\omega}_p$ . Condition (7) implies that  $\frac{1}{2}(y + \bar{x}_p) < \bar{\omega}_p$ . Thus, by (6), the principal's utility from  $D'_p = D_p \cup \{y\}$  and  $D_p$  are respectively given by

$$p \cdot V'_p = C + \int_{\bar{x}_p}^{(y+\bar{x}_p)/2} (\bar{x}_p - \omega)^2 d\omega - \int_{(y+\bar{x}_p)/2}^{\bar{\omega}_p} (y - \omega)^2 d\omega + b(y - \bar{\omega}_p)^2 - b^2, \quad (20)$$

$$p \cdot V_p = C + \int_{\bar{x}_p}^{\bar{\omega}_p} (\bar{x}_p - \omega)^2 d\omega + b(\bar{x}_p - \bar{\omega}_p)^2 - b^2, \quad (21)$$

where  $C$  is the principal's utility, conditional on the states  $[\underline{\omega}_p, \bar{x}_p]$  which is the same for both delegation sets. Thus,

$$p(V'_p - V_p) = - \int_{(y+\bar{x}_p)/2}^{\bar{\omega}_p} (y - \omega)^2 - (\bar{x}_p - \omega)^2 d\omega + b[(y - \bar{\omega}_p)^2 - (\bar{x}_p - \bar{\omega}_p)^2]. \quad (22)$$

The integral is equal to

$$- \int_{(y+\bar{x}_p)/2}^{\bar{\omega}_p} (y - \bar{x}_p)(y + \bar{x}_p - 2\omega) d\omega = (y - \bar{x}_p) \left( \frac{y + \bar{x}_p}{2} - \bar{\omega}_p \right)^2. \quad (23)$$

The second term in  $p(V'_p - V_p)$  can be written as

$$b[(y - \bar{\omega}_p)^2 - (\bar{x}_p - \bar{\omega}_p)^2] = b(y - \bar{x}_p)(y + \bar{x}_p - 2\bar{\omega}_p) = 2b(y - \bar{x}_p) \left( \frac{y + \bar{x}_p}{2} - \bar{\omega}_p \right). \quad (24)$$

Hence,

$$p(V'_p - V_p) = (y - \bar{x}_p) \left( \bar{\omega}_p - \frac{y + \bar{x}_p}{2} \right) \left( \bar{\omega}_p - \frac{y + \bar{x}_p}{2} - 2b \right). \quad (25)$$

Since the first two factors are positive by assumption, we obtain that  $p(V'_p - V_p) > 0$  is equivalent to  $\frac{1}{2}(y + \bar{x}_p) < \bar{\omega}_p - 2b = \beta_p^0$ , as desired.

Second, consider the case  $\bar{x}_p < \underline{\omega}_p < \frac{1}{2}(y + \bar{x}_p)$ . Due to non-redundancy, we have  $D_p = \{\bar{x}_p\}$  and  $D'_p = \{\bar{x}_p, y\}$ . In this case, the lower bound in the integrals in (20) and (21) is  $\bar{\omega}_p$  instead of  $\bar{x}_p$ . The formula (27) as well as the rest of the proof remain the same.

Finally, third, consider the case  $\frac{1}{2}(y + \bar{x}_p) \leq \underline{\omega}_p$ . We show that in this case, adding

$y$  improves the principal's expected utility (recall that  $\underline{\omega}_p < \beta_p^0$ ). Note that now non-redundancy implies that  $D_p = \{\bar{x}_p\}$  and  $D'_p = \{y\}$ , as in  $D'_p$  the action  $\bar{x}_p$  becomes redundant. Let us for any  $z \in \mathbb{R}$  define

$$\sigma(z) = -b(z - \underline{\omega}_p)^2 - \int_{\underline{\omega}_p}^{\bar{\omega}_p} (z - \omega)^2 d\omega + b(z - \bar{\omega}_p)^2 - b^2 \quad (26)$$

the principal's expected utility (scaled by  $p$ ) from a singleton delegation set  $\{z\}$ . Then  $p \cdot V_p = \sigma(\bar{x}_p)$  and  $p \cdot V'_p = \sigma(y)$  and a direct computation gives

$$p(V'_p - V_p) = (\bar{\omega}_p - \underline{\omega}_p)(y - \bar{x}_p)[\underline{\omega}_p + \bar{\omega}_p - (y + \bar{x}_p) - 2b]. \quad (27)$$

The first and the second brackets are clearly positive. The third bracket is also positive, as  $\underline{\omega}_p + \bar{\omega}_p - (y + \bar{x}_p) \geq \underline{\omega}_p + \bar{\omega}_p - 2\underline{\omega}_p = \bar{\omega}_p - \underline{\omega}_p = \ell > 2b$ , where the first inequality follows from  $\frac{1}{2}(y + \bar{x}_p) \leq \underline{\omega}_p$ , assumed in this case, and the second inequality follows from assumption (1). Thus, we have shown that indeed  $V'_p > V_p$ , which completes the proof.  $\square$

*Proof of Lemma 3.* Consider first the case that  $[y, \bar{x}_p] \subseteq D_p$ . Then  $D_p = D'_p \cup (y, \bar{x}_p]$  which means that  $D_p$  can be obtained by adding actions  $y'$  from the interval  $(y, \bar{x}_p]$  to  $D'_p$ . Since  $y$  is the maximal action in  $D'_p$ , and since  $y' \leq \bar{\omega}_p$ , such an action  $y'$  is non-redundant for the set  $D'_p$ . Moreover, the average of  $y = \max D'_p$  and  $y'$  is larger than  $\beta_p^0$  since  $y \geq \beta_p^0$ . Hence, Lemma 2, (ii), implies that adding  $y'$  to  $D'_p$  lowers the principal's expected utility. Therefore, the set  $D'_p$  yields the principal a higher expected utility than the set  $D_p$ , as we wanted to show.

Next, consider the case that  $[y, \bar{x}_p] \not\subseteq D_p$ . Then adding all (missing) actions in the interval  $[y, \bar{x}_p)$  causes an improvement for the principal by (6), because the action chosen by type  $\bar{\omega}_p$ , and thus his utility is not affected by the availability of the new actions. Together with the argument in the previous paragraph, this implies that the principal is better off with  $D'_p$  than with  $D_p$ , as desired.  $\square$

*Proof of Lemma 5.* Let  $D$  be an optimal delegation set. The claim follows from the following four properties:

- (a) Each of the sets  $D \cap [\underline{\omega}_h, \bar{\omega}_\ell]$  and  $D \cap [\bar{\omega}_\ell, \bar{\omega}_h]$  is connected or empty.
- (b) If both intersections in (a) are non-empty and  $x_1 = \max(D \cap [\underline{\omega}_h, \bar{\omega}_\ell])$  and  $x_2 = \min(D \cap [\bar{\omega}_\ell, \bar{\omega}_h])$ , then  $\bar{\omega}_\ell - x_1 = x_2 - \bar{\omega}_\ell$ .

(c)  $D \cap (\bar{\omega}_h, \infty)$  is empty.

(d)  $\underline{\omega}_h \in D$ .

We prove the properties by contradiction by assuming that each of the properties holds at the optimal delegation set and then constructing a delegation set which is an improvement.

(a) Suppose first  $D \cap [\underline{\omega}_h, \bar{\omega}_\ell]$  contains a gap, i.e. there are  $y_1, y_2 \in D$ ,  $y_1 < y_2$  so that  $D \cap (y_1, y_2) = \emptyset$ . Then, since  $y_2 < \bar{\omega}_\ell < \bar{\omega}_h$ , adding the set of actions  $(y_1, y_2)$  would not affect the agent's choice (and thus his utility) in states the  $\bar{\omega}_\ell$  and  $\bar{\omega}_h$ . However, it would strictly increase the agent's utility in the states  $\omega \in (y_1, y_2)$ . Therefore, by (8), adding actions from the interval  $(y_1, y_2)$  to  $D$  strictly improves the the principal's expected payoff, a contradiction to the optimality of  $D$ . The argument for the case that  $D \cap [\underline{\omega}_h, \bar{\omega}_\ell]$  contains a gap is identical.

(b) Suppose to the contrary that the differences  $\bar{\omega}_\ell - x_1$  and  $x_2 - \bar{\omega}_\ell$  are not equal. We only consider the case that  $\bar{\omega}_\ell - x_1 < x_2 - \bar{\omega}_\ell$  (the other case is analogous). Let  $x'_1 = 2\bar{\omega}_\ell - x_1$  be the mirror action of  $x_1$  mirrored at  $\bar{\omega}_\ell$ . Then  $x'_1 \notin D_p$ . Since  $x_2 \leq \bar{\omega}_h$  and by construction of  $x'_1$ , adding the action  $x'_1$  to the delegation set would not affect agent's choice (and thus his utility) in states the  $\bar{\omega}_\ell$  and  $\bar{\omega}_h$ . However, it would strictly increase the agent's utility the neighborhood of  $x'_1$ . Therefore, by (8), adding  $x'_1$  to  $D$  strictly improves the the principal's expected payoff, a contradiction to the optimality of  $D$ .

(c) Assume to the contrary that  $D \cap (\bar{\omega}_h, \infty)$  is not empty at the optimum. Then it contains exactly one action, namely  $\bar{x} = \max D$  (other actions would be redundant). Let us also denote  $x' = \max D \setminus \{\bar{x}\}$ . Then  $x' \leq \bar{\omega}_h$  and  $\bar{x}$  is chosen in state  $\bar{\omega}_h$ , i.e.,  $\frac{1}{2}(x' + \bar{x}) < \bar{\omega}_h$ . We consider two cases.

*Case 1.* Suppose first that the agent chooses action  $\bar{x}$  in state  $\bar{\omega}_\ell$ , i.e.,  $\frac{1}{2}(x' + \bar{x}) < \bar{\omega}_\ell$ . Then it would be profitable to add the action  $y = x' + \varepsilon$  for  $\varepsilon$  small enough. To see this, observe that  $y$  satisfies the non-redundancy condition (7) for  $p \in \{h, \ell\}$ , and also  $\frac{1}{2}(y + \bar{x}) < \bar{\omega}_p$ . Thus, it follows by Lemma 2, (i), that adding  $y$  improves the principal's expected utility, a contradiction to the optimality of  $D$ .

*Case 2.* Suppose next that the agent does not choose action  $\bar{x}$  in state  $\bar{\omega}_\ell$ , i.e.,  $\frac{1}{2}(x' + \bar{x}) \geq \bar{\omega}_\ell$ . We distinguish three subcases.

*Case 2.1.* If  $x' < \beta_\ell^0$ , then adding the (non-redundant) action  $y = x' + \varepsilon$  (for  $\varepsilon$  small enough) to  $D$  improves the principal's expected utility conditional on both types: for type  $h$  strictly due to part (i) of Lemma 2 (since  $y < \bar{x}$  and  $\frac{1}{2}(x' + \bar{x}) < \bar{\omega}_h$  so that also  $\frac{1}{2}(y + \bar{x}) < \bar{\omega}_h$  for  $\varepsilon$  small enough); for type  $\ell$  due to part (ii). (Here,  $x'$ , which is the

action chosen by type  $\ell$  in state  $\bar{\omega}_\ell$ , plays the role of  $\bar{x}_\ell$  in the statement of Lemma 2. We can then apply part (ii) since  $y > x'$  and  $\frac{1}{2}(y + x') < \beta_\ell^0$  for  $\varepsilon$  small enough.) This contradicts the optimality of  $D$ .

*Case 2.2.* If  $\frac{1}{2}(\bar{x} - x') < \bar{\omega}_h - \bar{\omega}_\ell$ , then let  $y = 2\bar{\omega}_h - \bar{x}$  be the mirror action of  $\bar{x}$ , mirrored at  $\bar{\omega}_h$ . We show that adding  $y$  to  $D$  improves the principal's utility conditional on both types. To see this, note that  $\frac{1}{2}(x' + \bar{x}) < \bar{\omega}_h$  implies that  $y > x'$ , and  $\frac{1}{2}(\bar{x} - x') < \bar{\omega}_h - \bar{\omega}_\ell$  implies that  $\frac{1}{2}(x' + y) > \bar{\omega}_\ell$ . Therefore,  $y$  is not chosen by the agent in the state  $\bar{\omega}_\ell$  and adding  $y$  does not influence the principal's utility conditional on type  $\ell$ . In addition, by (8), it improves the principal's utility conditional on type  $h$ , because it does not affect  $u_h(\bar{\omega}_h)$ , but improves the agent's utility in lower states. This again contradicts the optimality of  $D$ .

*Case 2.3.* Assume that  $x' \geq \beta_\ell^0$  and  $\frac{1}{2}(\bar{x} - x') \geq \bar{\omega}_h - \bar{\omega}_\ell$ . We show that removing the action  $\bar{x}$  improves the principal's utility. First, removing  $\bar{x}$  has no effect on the principal's utility conditional on type  $\ell$ , since  $\bar{x}$  is not chosen in the state  $\bar{\omega}_\ell$  by assumption. Moreover, it holds

$$\frac{y + \bar{x}}{2} = y + \frac{\bar{x} - y}{2} \geq \beta_\ell^0 + \bar{\omega}_h - \bar{\omega}_\ell = \beta_h^0. \quad (28)$$

Therefore, removing the action  $\bar{x}$  improves the principal's utility conditional on type  $h$  due to Lemma 2, (ii). (Observe that if the above inequality holds as an equality, then the principal is indifferent between removing and not removing  $\bar{x}$ . In this case, it is therefore without loss of generality that property (c) holds.)

(d) Suppose to the contrary that  $\underline{\omega}_h \notin D$ . Then  $D \cap [\underline{\omega}_h, \bar{\omega}_\ell]$  is empty, as otherwise adding the action  $\underline{\omega}_h$  would improve the principal's expected utility conditional on both types by Lemma 2, (i). Moreover, by part (c),  $D$  cannot contain an action larger than  $\bar{\omega}_h$ . Therefore,  $D$  contains actions in  $(\bar{\omega}_\ell, \bar{\omega}_h]$  or in  $(-\infty, \underline{\omega}_h)$ . In the former case, similar arguments as in part (b) imply that the principal's utility could be improved by adding the action  $y = 2\bar{\omega}_\ell - (\min(D \cap (\bar{\omega}_\ell, \bar{\omega}_h]))$ . Note that  $y \in [\underline{\omega}_h, \bar{\omega}_\ell]$ , contradicting the assumption that  $D \cap [\underline{\omega}_h, \bar{\omega}_\ell]$  is empty. Moreover, it cannot be optimal to offer actions only in  $(-\infty, \underline{\omega}_h)$  due to our assumption that  $b < \ell/2$ .

This completes the proof. □

*Proof of Proposition 1.* We want to show that the optimal static delegation set is an interval  $D^{st} = [\underline{\omega}_h, \bar{x}]$  with  $\beta_\ell^0 \leq x \leq \beta_h^0$ . To the contrary, suppose it is not an interval. By Lemma 5, we then have

$$D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup D', \quad (29)$$

with  $D' = \{\bar{\omega}_\ell + d\}$ , or  $D' = [\bar{\omega}_\ell + d, \bar{x}]$ . Now,  $D' = [\bar{\omega}_\ell + d, \bar{x}]$  cannot be optimal, because the fact that  $\beta_h^0 \leq \bar{\omega}_\ell < \bar{x}$  implies that lowering  $x$  slightly would be an improvement by Lemma 3 (conditional on both,  $h$  and  $\ell$ ). Therefore,  $D' = \{\bar{\omega}_\ell + d\}$ . But also this cannot be (strictly) optimal, because by Lemma 2, (ii), it would be profitable to remove the action  $\bar{\omega}_\ell + d$  from  $D^{st}$  conditional on both types (observe that  $\frac{1}{2}(\bar{\omega}_\ell - d + \bar{\omega}_\ell + d) = \bar{\omega}_\ell \geq \beta_h^0 \geq \beta_\ell^0$ ).

Hence, we have shown that  $D^{st}$  is an interval. We conclude the proof by demonstrating the stated properties of the upper endpoint  $\beta^{st}$  of the optimal interval. Let  $V_p(\bar{x})$  be the principal's expected utility from a general interval  $[\underline{\omega}_h, \bar{x}]$ , conditional on facing type  $p$ . Then, the principal's expected utility from interval delegation when he does not know the ex ante type is

$$V(\bar{x}) = \mu_h V_h(\bar{x}) + \mu_\ell V_\ell(\bar{x}). \quad (30)$$

From the benchmark case with publicly ability types, we know that

$$V_p'(\bar{x}) \begin{matrix} \leq \\ \geq \end{matrix} 0 \quad \Leftrightarrow \quad \bar{x} \begin{matrix} \geq \\ \leq \end{matrix} \beta_p^0. \quad (31)$$

Therefore,  $V'(\bar{x}) > 0$  for all  $\bar{x} < \beta_\ell^0$ , and  $V'(\bar{x}) < 0$  for all  $\bar{x} > \beta_\ell^0$ . Hence, the optimal upper endpoint  $\beta^{st}$  is in between  $\beta_\ell^0$  and  $\beta_h^0$ .

Moreover, differentiating the first order condition  $V'(\beta^{st}) = 0$  with respect to  $\mu_h$  yields

$$\frac{d}{d\mu_h} \beta^{st} = V_h'(\bar{\beta}^{st}) - V_\ell'(\beta^{st}), \quad (32)$$

which is strictly positive by (31) since  $\beta^{st} \in (\beta_\ell^0, \beta_h^0)$ . Furthermore,  $V$  converges to  $V_h$  when  $\mu_h$  converges to 0. Thus, also  $\beta^{st}$  converges to  $\beta_h^0$ . Similarly,  $\beta^{st}$  converges to  $\beta_\ell^0$  as  $\mu_\ell$  converges to 1. This completes the proof.  $\square$

*Proof of Proposition 2.* We want to show that the optimal static delegation set,  $D^{st}$ , is the union of an interval  $[\underline{\omega}_h, \bar{\omega}_\ell - d]$  and a set  $D'$  with either  $D' = [\bar{\omega}_\ell + d, \beta_h^0]$  and  $\beta_h^0 > \bar{\omega}_\ell + d$ , or  $D' = \{\bar{\omega}_\ell + d\}$  and  $\beta_h^0 \leq \bar{\omega}_\ell + d$ .

Indeed, suppose, contrary to the claim, that  $D^{st}$  is not of the postulated form. Then by Lemma 5, it is an interval  $[\underline{\omega}_h, \bar{\omega}_\ell - d]$ . But Lemma 2, (ii), implies that the principal could improve by offering the set  $[\underline{\omega}_h, \bar{\omega}_\ell - d] \cup \{\bar{\omega}_\ell + d\}$ . (Adding the action  $\bar{\omega}_\ell + d$  does not affect the choices of type  $\ell$ . however, since  $\beta_h^0 > \bar{\omega}_\ell$  by assumption, Lemma 2, (ii) implies that adding the action  $\bar{\omega}_\ell + d$  improves the principal's utility conditional on type  $h$ .) Therefore, by Lemma 5, the optimal set is the union of two intervals, or an interval and a point.



Consider the first case  $D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup [\bar{\omega}_\ell + d, x]$ . Then, if  $x \neq \beta_h^0$ , then the principal's expected profit could be improved either by adding an action  $y = x + \varepsilon$  close to  $x$  (when  $x < \beta_h^0$ ) or by slightly decreasing  $x$  (when  $x > \beta_h^0$ ). The modifications would not affect the choices by type  $\ell$ , but improve the principal's utility conditional on type  $h$  by Lemma 2, (i), in the first case and by Lemma 3 in the second case. This shows that when the optimal delegation set is the union of two intervals, the second interval is  $D' = [\bar{\omega}_\ell + d, \beta_h^0]$ .

Consider next the case  $D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup \{\bar{\omega}_\ell + d\}$ . Then, if  $\bar{\omega}_\ell + d < \beta_h^0$ , then, adding an action slightly higher than  $\bar{\omega}_\ell + d$  would improve the the principal's expected profit conditional on type  $h$  by Lemma 2, (i), but leave the choices of type  $\ell$  unaffected. This shows that when the optimal delegation set is the union of an interval and a point  $\bar{\omega}_\ell + d$ , we have  $\bar{\omega}_\ell + d \geq \beta_h^0$ . This completes the proof.  $\square$

*Proof of Lemma 6.* To show the claim, we use the following auxiliary lemma proven below.

**Lemma 9.** *Consider an arbitrary menu  $(D_h, D_\ell)$  of delegation sets and  $p \in \{h, \ell\}$ . If  $D_p \neq [\underline{\omega}_p, z]$  for all  $z \geq \beta_p^0$ , then for every  $\varepsilon > 0$  there exists  $y \notin D_p$  such that replacing  $D_p$  by the delegation set  $D'_p = D_p \cup \{y\}$  improves the principal's expected utility and improves the agent's expected utility by no more than  $\varepsilon$ :*

$$0 < V'_p - V_p, \quad \text{and} \quad 0 \leq U'_p - U_p < \varepsilon. \quad (33)$$

To establish Lemma 6, observe first that at the optimum of the relaxed problem,  $D_\ell^* = [\underline{\omega}_\ell, \bar{x}_\ell^*]$  with  $\bar{x}_\ell^* \geq \beta_\ell^0$ . Otherwise, we can consider a modification of  $D_\ell^*$  as specified in Lemma 9. Such a modification preserves  $(IC_\ell)$ , as it increases  $U_\ell^*$ , but also improves  $V_\ell^*$ . This contradicts the optimality of  $D_\ell^*$ .

Second, we show that  $(IC_\ell)$  is binding at the optimum. To the contrary, suppose  $(IC_\ell)$  is slack at the optimum. We derive a contradiction in three steps.

*Step 1.* We show that  $D_h^* = [\underline{\omega}_h, \bar{x}_h^*]$ , where  $\bar{x}_h^* \geq \beta_h^0$ . Otherwise, consider a modification of  $D_h^*$  as specified in Lemma 9. Because such a modification improves the  $\ell$ -type agent's expected utility only marginally and thus preserves the slack  $(IC_\ell)$ . Moreover, it also improves  $V_h^*$ , which is a contradiction to the optimality of  $D_h^*$ .

*Step 2.* We show that  $\bar{x}_h = \beta_h^0$ , i.e.,  $D_h^* = [\underline{\omega}_h, \beta_h^0]$ . Assume to the contrary that  $\bar{x}_h > \beta_h^0$ . Non-redundancy requires that  $\bar{x}_p \leq \bar{\omega}_p$ . Let us replace the delegation set  $D_h^*$  by the set  $[\underline{\omega}_h, \bar{x}_h^* - \varepsilon]$  where  $\varepsilon > 0$  is sufficiently small (so that  $\bar{x}_h^* - \varepsilon > \beta_h^0$ ). This modification preserves  $(IC_\ell)$ , as it lowers the  $\ell$ -type's expected utility from choosing  $D_h$ .

Moreover, by Lemma 3, it improves the principal's expected utility, which contradicts the optimality of  $D_h^*$ .

*Step 3.* We finally argue that  $(IC_\ell)$  cannot be slack. Indeed, as we have already shown,  $D_\ell^*$  is an interval. Since  $D_h^* = [\underline{\omega}_h, \beta_h^0]$ , the slack  $(IC_\ell)$  implies that  $\bar{x}_\ell^* > \beta_h^0 > \beta_\ell^0$ . Then, replacing  $D_\ell^*$  by  $[\underline{\omega}_h, \bar{x}_\ell^* - \varepsilon]$  again preserves the slack  $(IC_\ell)$ , as the modification is only marginal. Moreover, by Lemma 3, the principal would benefit from the change, which is a contradiction to the assumption that  $D_\ell^*$  is optimal.  $\square$

*Proof of Lemma 9.* Adding an action always (weakly) improves the agent's utility. It is thus sufficient to specify an action  $y$  such that the improvement is marginal and that it also increases the principal's expected utility.

Assume first that  $D_p$  is not connected. Then  $D_p$  contains a gap, i.e., there are  $y_1, y_2 \in D_p$ ,  $y_1 < y_2$ , such that  $D_p \cap (y_1, y_2) = \emptyset$ . Then we can add the action  $y = y_1 + \varepsilon'$  to  $D_p$  (for  $\varepsilon' > 0$  small). Since  $y_1$  is not chosen in state  $\bar{\omega}_p$ , we can find  $\varepsilon'$  sufficiently small so that  $\frac{1}{2}(y + \bar{x}_p) < \bar{\omega}_p$ . Hence, by Lemma 2, (i), adding  $y$  also improves the principal's utility.

From now on, let us assume that  $D_p$  is connected, i.e.,  $D_p = [\underline{x}_p, \bar{x}_p]$ , or  $D_p$  is a singleton (in which case  $\underline{x}_p = \bar{x}_p$ ). If  $D_p \neq [\underline{\omega}_p, z]$  for all  $z \geq \beta_p^0$ , then we have four possible cases.

*Case 1.* Let  $\underline{x}_p < \underline{\omega}_p$ . Non-redundancy requires that  $D_p$  is a singleton with  $\underline{x}_p = \bar{x}_p < \underline{\omega}_p$ . Then we can add the action  $y = \underline{x}_p + \varepsilon'$  for  $\varepsilon' > 0$  small enough (so that  $y < \underline{\omega}_p$ ). This modification improves the agent's utility only marginally. It also improves the principal's expected utility due to Lemma 2, (ii), as  $\frac{1}{2}(y + \bar{x}_p) < \underline{\omega}_p < \beta_p^0$ .

*Case 2.* Let  $\bar{\omega}_p < \bar{x}_p$ . Non-redundancy again requires that  $D_p$  is a singleton with  $\underline{x}_p = \bar{x}_p > \bar{\omega}_p$ . Then we can add the action  $y = \bar{x}_p - \varepsilon'$  for  $\varepsilon' > 0$  small enough. This is again a marginal modification with  $D'_p = \{y\}$  being a singleton. In order to compare the principal's expected utility we use formula (27) from the Proof of Lemma 2, (ii). Now the first bracket is positive, whereas the second bracket is negative. The third bracket is negative as well, because  $\underline{\omega}_p + \bar{\omega}_p - (y + \bar{x}_p) < \underline{\omega}_p + \bar{\omega}_p - 2\bar{\omega}_p < 0 < 2b$ . Thus, adding  $y$  to  $D_p$  indeed improves the principal's expected utility.

*Case 3.* Let  $\underline{\omega}_p < \underline{x}_p$  and  $\bar{x}_p \leq \bar{\omega}_p$ . Now we can add the action  $y = \underline{x}_p - \varepsilon'$  for  $\varepsilon' > 0$  small enough (so that  $y > \underline{\omega}_p$ ). This change improves the agent's utility only marginally and is profitable for the principal, due to Lemma 2, (i).

*Case 4.* Let  $\underline{\omega}_p = \underline{x}_p$  and  $\bar{x}_p < \beta_p^0$ . In this case, we can add the action  $y = \bar{x}_p + \varepsilon'$  for  $\varepsilon' > 0$  small enough (so that  $y < \beta_p^0$ ). This change again improves the agent's utility

only marginally, and it improves the principal's expected utility due to Lemma 2, (ii), as  $\frac{1}{2}(y + \bar{x}_p) < \beta_p^0$ .  $\square$

*Proof of Lemma 7.* We start by deriving necessary conditions, (a)–(e), for  $D_h^*$ . Before doing so, recall our definition that  $\gamma_2 = x_h(\bar{\omega}_\ell)$  is the agent's choice from  $D_h^*$  in state  $\bar{\omega}_\ell$  (if in state  $\bar{\omega}_\ell$  the agent is indifferent between two actions, let  $\gamma_2$  be the higher one).

(a) Each of the intersections  $D_h^* \cap (-\infty, \underline{\omega}_\ell]$  and  $D_h^* \cap [\bar{\omega}_\ell, \infty)$  is connected or empty. If, to the contrary, one of them is not connected, it contains a gap, i.e., there are  $y_1, y_2 \in D_h^*$ ,  $y_1 < y_2$  such that  $D_h^* \cap (y_1, y_2) = \emptyset$ . Then we can add the action  $y = y_1 + \varepsilon$  to  $D_h^*$  (for  $\varepsilon > 0$  small). Since  $y_1$  is not chosen in state  $\bar{\omega}_h$ , we can find  $\varepsilon$  sufficiently small so that  $\frac{1}{2}(y + \bar{x}_h) < \bar{\omega}_h$ . Hence, by Lemma 2, (i), adding  $y$  also improves the principal's utility. Moreover, adding  $y$  also preserves  $(IC_\ell)$ , as it does not modify the agent's utility in states  $[\underline{\omega}_\ell, \bar{\omega}_\ell]$ . This contradicts the optimality of  $D_h^*$ .

(b)  $\underline{x}_h \leq \underline{\omega}_h$ . Assume the opposite. Then, consider adding the (non-redundant) action  $y = 2\underline{\omega}_h - \underline{x}_h + \varepsilon$  to  $D_h^*$ .<sup>32</sup> For  $\varepsilon > 0$  small enough, action  $y$  is chosen in the state  $\underline{\omega}_h$  but not in the state  $\omega_\ell$ . Thus, adding  $y$  does not influence the  $\ell$ -type's utility from the modified  $D_h^*$ , and so it preserves  $(IC_\ell)$ . Moreover, this operation improves the principal's utility from type  $h$ , according to Lemma 2, (i), a contradiction to the optimality of  $D_h^*$ .

(c) If  $\gamma_2 < \bar{x}_h$ , then  $\bar{\omega}_\ell \leq \gamma_2$ . Assume the opposite. Then by definition of  $\gamma_2$ , the action  $y = 2\bar{\omega}_\ell - \gamma_2$  (the mirror image of  $\gamma_2$  with respect to  $\bar{\omega}_\ell$ ) does not belong to  $D_h^*$ . Then  $y < \bar{x}_h$  and  $\frac{1}{2}(y + \bar{x}_h) \leq \bar{\omega}_h$ . Thus, adding  $y$  to  $D_h^*$  preserves  $(IC_\ell)$  and also improves the principal's expected utility by Lemma 2, (i).

(d) If  $\gamma_2 < \bar{x}_h$ , then  $\bar{x}_h = \beta_h^0$ . Otherwise, the principal's expected utility can be improved by one of the following modifications. If  $\bar{x}_h < \beta_h^0$ , we add the action  $y = \bar{x}_h + \varepsilon$  to  $D_h^*$ , where  $\varepsilon > 0$  is small enough. If  $\bar{x}_h \in (\beta_h^0, \bar{\omega}_h]$ , we reduce  $\bar{x}_h$  marginally, i.e., replace  $D_h^*$  by the delegation set  $(D_h^* \cap (-\infty, \bar{x}_h - \varepsilon]) \cup \{\bar{x}_h - \varepsilon\}$ , where  $\varepsilon > 0$  is small enough. If  $\bar{x}_h > \bar{\omega}_h$ , then  $\bar{x}_h$  is an isolated point of  $D_h^*$  by Lemma 1, and we remove the action  $\bar{x}_h$ . In all cases, because  $\bar{\omega}_\ell \leq \gamma_2 < \bar{x}_h$ , the change does not affect the agent's choices in states  $\omega \in [\underline{\omega}_\ell, \bar{\omega}_\ell]$  and, thus, it keeps  $(IC_\ell)$  binding. Moreover, it improves the principal's expected payoff: due to Lemma 2, (ii) for the first change; due to Lemma 3 (because  $\bar{x}_h - \varepsilon > z \geq \beta_h^0$ ) for the second change; and, finally, due to Lemma 2, (ii) for the third change. This contradicts the optimality of  $D_h^*$ .

We can now prove part (i) and (ii) of Lemma 7. We show that the principal weakly benefits by replacing  $D_h^*$  by the set  $D'_h = D_{h1} \cup D_{h2}$ , where the sets  $D_{h1}$  and  $D_{h2}$  are specified as follows: From (a), (c), and (d), we infer that if  $\gamma_2 < \bar{x}_h$ , then  $\bar{\omega}_\ell < \beta_h^0$  and

<sup>32</sup>Note that  $2\underline{\omega}_h - \underline{x}_h$  is the mirror action to  $\underline{x}_h$ , mirrored at  $\underline{\omega}_h$ .

$D_h^* \cap [\bar{\omega}_\ell, \infty) = [\gamma_2, \beta_h^0]$ . In that case, let us set  $D_{h2} = [\gamma_2, \beta_h^0]$ . Otherwise, if  $\gamma_2 = \bar{x}_h$ , we set  $D_{h2} = \{\gamma_2\}$ .

Consider now the set  $[\underline{x}_h, \gamma_1] \cup D_{h2}$  as constructed in the discussion of Lemma 7. (If  $\gamma_1 = \underline{x}_h$ , we identify  $[\underline{x}_h, \gamma_1]$  with  $\{\underline{x}_h\}$ .) The construction ensures that type  $\ell$  is indifferent between the set  $D_\ell^*$  and  $[\underline{x}_h, \gamma_1] \cup D_{h2}$ , so replacing  $D_h^*$  by  $[\underline{x}_h, \gamma_1] \cup D_{h2}$  keeps  $(IC_\ell)$  binding. Now it may be that  $[\underline{x}_h, \gamma_1] \cup D_{h2}$  contains redundant actions. If  $\gamma_1 \leq \underline{\omega}_h$  then all actions below  $\gamma_1$  are redundant, and we set  $D_{h1} = \{\gamma_1\}$ ; if  $\gamma_1 > \underline{\omega}_h$ , all actions below  $\underline{\omega}_h$  are redundant, and we set  $D_{h1} = [\underline{\omega}_h, \gamma_1]$ .

We now show that the principal's utility, conditional on facing the  $h$ -type, is weakly higher under  $D'_h = D_{h1} \cup D_{h2}$  than under  $D_h^*$ , implying that  $D'_h$  is optimal, too. For this we use the formula (6) for  $p = h$ , where we split the integral into three parts.

$$h \cdot V_h = \underbrace{bu_h(\underline{\omega}_h)}_{T_1} + \underbrace{\int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega}_{T_2} + \underbrace{\int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_h(\omega) d\omega}_{T_3} + \underbrace{\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega}_{T_4} - \underbrace{bu_h(\bar{\omega}_h)}_{T_5} - b^2. \quad (34)$$

In what follows we indicate all variables pertaining to  $D'_h$  (resp  $D_h^*$ ) with a prime (resp. an asterisk) and compare the principal's utilities from  $D'_h$  and  $D_h^*$  using (34).

First, observe that  $T'_3 = T_3^*$  because  $(IC_\ell)$  is binding for both  $D_h^*$  and  $D'_h$ .

Second, we argue that  $T'_4 = T_4^*$  and  $T'_5 = T_5^*$ . Indeed, this follows directly from the above construction, because we kept the agent's utility on the states  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h]$  unchanged.

Third, we show that  $T'_1 \geq T_1^*$  and  $T'_2 \geq T_2^*$ . It follows from (a) and (b) that  $D_h^* \cap (-\infty, \underline{\omega}_\ell]$  is either a singleton  $\{\underline{x}_\ell\}$ , or it is an interval of the form  $[\underline{\omega}_\ell, z]$ . In both cases, this intersection is a subset of  $[\underline{x}_\ell, \gamma_1]$ , which then yields a (weakly) higher utility to the agent. Moreover, this utility remains unchanged when omitting the redundant actions, as described above.

Summing up, we have shown that the delegation set  $D'_h$  yields at least as large utility to the principal's as the set  $D_h^*$ , and this establishes part (i) and (ii).

Finally, we proof part (iii) by deriving the inequalities for  $\gamma_1$  and  $\gamma_2$ . First, we show that  $\gamma_1 \leq \beta_\ell^*$ . Assume the opposite. Then in every state  $\omega \in (\beta_\ell^*, \bar{\omega}_\ell]$ , the agent is strictly better off under the delegation set  $D'_h$  than under  $D_\ell^*$ , as the former has an action, namely  $\min\{\omega, \gamma_1\}$ , that is closer to the agent's ideal action. Moreover, in states  $\omega \in [\underline{\omega}_\ell, \beta_\ell^*]$ , the agent chooses his ideal action  $\omega$  under both  $D'_h$  and  $D_\ell^*$ . Thus, the  $\ell$ -type agent strictly prefers the delegation set  $D'_h$ , which contradicts  $(IC_\ell)$ .

Second, we show that  $\beta_\ell^* \leq \gamma_2$ . Otherwise, we have  $\gamma_2 < \beta_\ell^* \leq \bar{\omega}_\ell$  and thus  $D_{h2} = \{\gamma_2\}$ . Then in every state  $\omega \in (\gamma_2, \bar{\omega}_\ell]$ , the agent is strictly better off under the delegation set

$D_\ell^*$  than under  $D'_h$ , as the former has an action, namely  $\min\{\omega, \beta_h^*\}$ , that is closer to the agent's ideal action. Moreover, in states  $\omega \in [\underline{\omega}_\ell, \gamma_2]$ , the agent weakly better off under  $D_\ell^*$ , where he chooses his ideal action  $\omega$ . Thus, the  $\ell$ -type agent strictly prefers the delegation set  $D_h^*$ , which contradicts  $(IC_\ell)$ .

Third, we show that  $\frac{1}{2}(\beta_\ell^* + \gamma_2) \leq \bar{\omega}_\ell$ . The inequality is trivial when  $\gamma_2 \leq \bar{\omega}_\ell$ . Let us thus assume that  $\gamma_2 > \bar{\omega}_\ell$  and  $\frac{1}{2}(\beta_\ell^* + \gamma_2) > \bar{\omega}_\ell$ . By definition, in state  $\bar{\omega}_\ell$ , the agent chooses action  $\gamma_2$  from  $D'_h$ , whereas he chooses action  $\beta_\ell^*$  from  $D_\ell^*$ . Since  $\beta_\ell^* \leq \bar{\omega}_\ell < \gamma_2$  and  $\frac{1}{2}(\beta_\ell^* + \gamma_2) > \bar{\omega}_\ell$ , the action  $\beta_\ell^*$  is strictly better than  $\gamma_2$  in state  $\bar{\omega}_\ell$ . Therefore, the agent is strictly better off under the delegation set  $D_\ell^*$  than under  $D'_h$  in state  $\bar{\omega}_\ell$ , and by continuity also in some neighborhood around  $\bar{\omega}_\ell$ . Moreover, in all other states in  $[\underline{\omega}_\ell, \bar{\omega}_\ell]$  the agent is (weakly) better off under  $D_\ell^*$  than under  $D'_h$  since  $\gamma_1 \leq \beta_\ell^*$ . Thus, the  $\ell$ -type agent strictly prefers the delegation set  $D_\ell^*$  to  $D'_h$ , which contradicts the assumption that  $(IC_\ell)$  is binding. This completes the proof.  $\square$

*Proof of Lemma 8.* Let  $(D_h^*, D_\ell^*)$  be a solution to the relaxed problem. The statement is trivial when the delegation menu is static. Let us thus consider only delegation menus that are sequential. Assume, to the contrary, that  $(IC_h)$  is violated. We show that the principal can increase her expected utility by offering both types the (static) delegation menu  $(D'_h, D'_\ell)$  such that  $D'_\ell = D_\ell$  and  $D'_h = [\underline{\omega}_h, \bar{x}_\ell]$ .

We obtain from the violated  $(IC_h)$  and binding  $(IC_\ell)$  that

$$\begin{aligned} & \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega + \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_h(\omega) d\omega + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega \\ &= \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega + \ell \cdot U_\ell + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega \\ &< - \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} (\underline{\omega}_\ell - \omega)^2 d\omega + \ell \cdot U_\ell - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} (\bar{x}_\ell - \omega)^2 d\omega, \end{aligned} \quad (35)$$

where the equality in the second line follows since  $(IC_\ell)$  is binding, and the third expression is the  $h$ -type's expected utility from  $D_\ell$ .

Now, as the delegation sets  $D_\ell$  and  $D'_\ell$  are identical, they also yield the same principal's expected utility. Let us then compare the principal's expected utility conditional on facing the  $h$ -type. By formula (6) we obtain

$$\begin{aligned} h \cdot V_h &= \ell \cdot U_\ell + \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega + bu_h(\underline{\omega}_h) - bu_h(\bar{\omega}_h) - b^2 \\ &< \ell \cdot U_\ell - \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} (\underline{\omega}_\ell - \omega)^2 d\omega - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} (\bar{x}_\ell - \omega)^2 d\omega + bu_h(\underline{\omega}_h) + b(\bar{x}_h - \bar{\omega}_h)^2 - b^2, \end{aligned} \quad (36)$$

where the inequality follows from (35) and where we have used that the largest type  $\bar{\omega}_h$  picks the largest action  $\bar{x}_h$  by Lemma 1. On the other hand, for the set  $D'_h$  we obtain

$$h \cdot V'_h = \ell \cdot U_\ell - 0 - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} (\bar{x}_\ell - \omega)^2 d\omega + 0 + b(\bar{x}_\ell - \bar{\omega}_h)^2 - b^2. \quad (37)$$

Comparing the expressions (36) and (37) term by term, we can clearly see that the latter is higher. The only inequality that is not trivial is  $b(\bar{x}_h - \bar{\omega}_h)^2 < b(\bar{x}_\ell - \bar{\omega}_h)^2$ , which is equivalent to  $\frac{1}{2}(\bar{x}_\ell + \bar{x}_h) < \bar{\omega}_h$ . This clearly holds when  $\bar{x}_h \leq \bar{\omega}_h$ . On the other hand, when  $\bar{x}_h > \bar{\omega}_h$ , then  $\bar{x}_h$  is an isolated point of  $D_h^*$  and  $D_{h2}$  as specified in Lemma 7 is a singleton. However, according to Lemma 7, (iii) we have  $\frac{1}{2}(\bar{x}_\ell + \bar{x}_h) \leq \bar{\omega}_\ell < \bar{\omega}_h$ . Thus we indeed obtain that  $V'_h > V_h$ . This proves our statement and completes the proof of the lemma.  $\square$

*Proof of Proposition 3.* Let  $(D_h^*, D_\ell^*)$  be the optimal delegation menu in the form as specified in Lemmas 6 and 7. Let us assume to the contrary that this menu is a sequential menu and is (strictly) better for the principal than any static menu.

The main idea of the proof is the comparison of this delegation menu with a static delegation set  $D' = [\underline{\omega}_h, \beta_\ell^*] = [\underline{\omega}_h, \underline{\omega}_\ell] \cup D_\ell^*$  being offered to both types. Under this menu, the  $h$ -type is also offered the  $\ell$ -type's delegation set which is extended by states that are never chosen by the  $\ell$ -type. We show that for  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$ , the delegation set  $D'$  improves the principal's expected utility.

When analyzing the differences between the principal's expected payoffs, it will turn out to be useful to define the function

$$\kappa(\omega) = (\beta_\ell^* - \omega)^2 - (\beta_h^* - \omega)^2 = (\beta_h^* - \beta_\ell^*)[2\omega - (\beta_\ell^* + \beta_h^*)]. \quad (38)$$

Now consider the menu  $(D', D')$ . Observe that the states by which  $D_\ell$  is extended do not alter the choice function of the type  $\ell$ . Thus  $D'$  yields the same expected utility to the principal as  $D_\ell$  conditional on facing the  $\ell$ -type:

$$V'_\ell = V_\ell. \quad (39)$$

For the comparison of the principal's expected utility conditional on facing the  $h$ -type, we use formula (6) and observe that the integral on the interval  $[\underline{\omega}_\ell, \bar{\omega}_\ell]$  is actually equal to type  $\ell$ 's expected utility (multiplied by factor  $\ell$ ) from the delegation sets  $D_\ell$  and  $D_h$ ,

respectively. These expected utilities are equal due to the binding ( $IC_\ell$ ). Thus,

$$\begin{aligned}
h \cdot V_h &= \ell \cdot U_\ell + \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega - \int_{\underline{\omega}_\ell}^{\bar{\omega}_h} (\beta_h^* - \omega)^2 d\omega + bu_h(\underline{\omega}_h) + b(\beta_h^* - \bar{\omega}_h)^2 - b^2, \\
h \cdot V_h' &= \ell \cdot U_\ell - \int_{\underline{\omega}_\ell}^{\bar{\omega}_h} (\beta_\ell^* - \omega)^2 d\omega + b(\beta_\ell^* - \bar{\omega}_h)^2 - b^2, \\
h(V_h - V_h') &= -b\kappa(\bar{\omega}_h) + bu_h(\underline{\omega}_h) + \int_{\underline{\omega}_\ell}^{\bar{\omega}_h} \kappa(\omega) d\omega + \int_{\underline{\omega}_\ell}^{\underline{\omega}_h} u_h(\omega) d\omega.
\end{aligned} \tag{40}$$

In the following, we show that  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$  implies  $V_h' \geq V_h$ , which contradicts sequential delegation being optimal.

First recall that according to Lemma 7 we have  $\beta_\ell^* + \beta_h^* \leq 2\bar{\omega}_\ell < 2\bar{\omega}_h$  and thus  $\kappa(\bar{\omega}_h) > 0$ . We then obtain

$$\begin{aligned}
h(V_h - V_h') &\leq -b\kappa(\bar{\omega}_h) + \int_{\underline{\omega}_\ell}^{\bar{\omega}_h} \kappa(\omega) d\omega \leq -b\kappa(\bar{\omega}_h) + \int_{\underline{\omega}_\ell}^{\bar{\omega}_h} \kappa(\bar{\omega}_h) d\omega \\
&= -[b - (\bar{\omega}_h - \bar{\omega}_\ell)]\kappa(\bar{\omega}_h) \leq 0.
\end{aligned} \tag{41}$$

The first inequality was obtained by neglecting the non-positive terms  $\int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega) d\omega$  and  $bu_h(\underline{\omega}_h)$  in  $V_h$ . The second inequality follows from the fact that  $\kappa(\omega)$  is increasing in  $\omega$ . The third inequality follows from the assumption  $b \geq \bar{\omega}_h - \bar{\omega}_\ell$  and from  $\kappa(\bar{\omega}_h) > 0$ . This completes the proof.  $\square$

*Proof of Proposition 4.* We begin with part (b): We show that the marginal sequential delegation menu (14) is strictly better than static delegation whenever  $b < \hat{b}$  and  $\mu_h < \hat{\mu}_h(b)$ , where  $\hat{\mu}_h(b)$  is determined below. As argued in the text following the statement of the proposition, this is the case if expression (15) satisfies:

$$\frac{d(hV_h)}{d\beta^{st}} > 0 \quad \Leftrightarrow \quad b < \hat{b} \quad \text{and} \quad \mu_h < \hat{\mu}_h(b). \tag{42}$$

Indeed, we have

$$\begin{aligned}
\frac{d(hV_h)}{d\beta^{st}} &= \frac{d}{d\beta^{st}} \left\{ \int_{\underline{\omega}_\ell}^{\bar{\omega}_h} -(\beta^{st} - \omega)^2 d\omega - b[-(\beta^{st} - \bar{\omega}_h)^2] \right\}. \\
&= -2 \int_{\underline{\omega}_\ell}^{\bar{\omega}_h} (\beta^{st} - \omega) d\omega + 2b(\beta^{st} - \bar{\omega}_h) \\
&= -(\beta^{st} - \bar{\omega}_h)^2 + (\beta^{st} - \bar{\omega}_\ell)^2 + 2b(\beta^{st} - \bar{\omega}_h) \\
&= [2\beta^{st} - 2\bar{\omega}_h + (\bar{\omega}_h - \bar{\omega}_\ell)] [-(\bar{\omega}_h - \bar{\omega}_\ell)] - 2b(\bar{\omega}_h - \beta^{st}) \\
&= 2(\bar{\omega}_h - \beta^{st})(\bar{\omega}_h - \bar{\omega}_\ell - b) - (\bar{\omega}_h - \bar{\omega}_\ell)^2.
\end{aligned} \tag{43}$$

Hence, since  $\bar{\omega}_h - \bar{\omega}_\ell - b > 0$  by assumption, we obtain

$$\frac{d(hV_h)}{d\beta^{st}} > 0 \quad \Leftrightarrow \quad \beta^{st} < \bar{\omega}_h - \frac{(\bar{\omega}_h - \bar{\omega}_\ell)^2}{2(\bar{\omega}_h - \bar{\omega}_\ell - b)} \equiv \xi_h. \quad (44)$$

Now, a computation (provided below) shows that

$$\beta_\ell^0 < \xi_h \quad \Leftrightarrow \quad b < \hat{b}. \quad (45)$$

By Proposition 1,  $\beta^{st}$  increases monotonically from  $\beta_\ell^0$  to  $\beta_h^0$  as  $\mu_h$  goes from 0 to 1. Thus, for every  $b < \hat{b}$  there is a unique  $\hat{\mu}_h(b) \in (0, 1]$  so that  $\beta^{st} < \xi_h$  if and only if  $\mu_h < \hat{\mu}_h(b)$ . Together with (44) and (45), this establishes (42).

To complete the proof of part (b), we show (45). Let  $a = \bar{\omega}_h - \bar{\omega}_\ell$ . Observe that  $a - b > 0$  by assumption. Recall that  $\beta_\ell^0 = \bar{\omega}_\ell - 2b$ . Thus, by definition of  $\xi_h$ :

$$\begin{aligned} \beta_\ell^0 < \xi_h &\Leftrightarrow -2b < a - \frac{a^2}{2(a-b)} \Leftrightarrow 4b^2 - 2ab - a^2 < 0 \\ &\Leftrightarrow b > \frac{1}{4}(1 - \sqrt{5})a \quad \text{and} \quad b < \frac{1}{4}(1 + \sqrt{5})a = \hat{b}. \end{aligned} \quad (46)$$

The left inequality in the last line is always satisfied since  $\frac{1}{2}a > \frac{1}{4}(1 - \sqrt{5})a$  and since, by the maintained assumption in this subsection,  $b \geq \frac{1}{2}a$ . Therefore,  $\beta_\ell^0 < \xi_h$  if and only if  $b < \hat{b}$ , as desired.

Proof of part (a): By Lemma 6 and 7, the optimal delegation menu is in the set

$$\begin{aligned} \mathcal{R} \equiv \{ &(D_h, D_\ell) \mid D_h = [\min\{\omega_h, z_h\}, z_h] \cup \{\bar{x}_h\}, \quad D_\ell = [\underline{\omega}_\ell, \bar{x}_\ell], \\ &\beta_\ell^0 \leq \bar{x}_\ell \leq \bar{x}_h, \quad \frac{1}{2}(\bar{x}_\ell + \bar{x}_h) \leq \bar{\omega}_\ell, \\ &(IC_\ell) \text{ binding} \} \end{aligned} \quad (47)$$

Due to the fact that  $(IC_\ell)$  is binding, any menu in  $\mathcal{R}$  is pinned down by the points  $\bar{x}_h$  and  $\bar{x}_\ell$ , and we denote the associated expected utility for the principal by

$$V(\bar{x}_h, \bar{x}_\ell). \quad (48)$$

Observe that static interval delegation corresponds to the menu in  $\mathcal{R}$  for which  $\bar{x}_h = \bar{x}_\ell$ .



By (6) and (13),

$$\begin{aligned}
V(\bar{x}_h, \bar{x}_\ell) &= \frac{\mu_h}{h} hV_h + \frac{\mu_\ell}{\ell} \ell V_\ell \\
&= \frac{\mu_h}{h} \left[ \ell \cdot U_\ell + \int_{\underline{\omega}_h}^{\bar{\omega}_\ell} u_h(\omega) d\omega + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega + bu_h(\underline{\omega}_h) - bu_h(\bar{\omega}_h) \right] \\
&\quad + \frac{\mu_\ell}{\ell} \left[ \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_\ell(\omega) d\omega + bu_\ell(\underline{\omega}_\ell) - bu_\ell(\bar{\omega}_\ell) \right]. \tag{49}
\end{aligned}$$

Since  $u_h(\omega) \leq 0$  for all  $\omega$ , the second and the fourth term in the first square brackets are negative, and we obtain that

$$\begin{aligned}
\Psi(\bar{x}_h, \bar{x}_\ell) &\equiv \frac{\mu_h}{h} \left[ \ell \cdot U_\ell + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega) d\omega - bu_h(\bar{\omega}_h) \right] \\
&\quad + \frac{\mu_\ell}{\ell} \left[ \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_\ell(\omega) d\omega + bu_\ell(\underline{\omega}_\ell) - bu_\ell(\bar{\omega}_\ell) \right] \tag{50}
\end{aligned}$$

is an upper bound on  $V$  which is actually attained if  $\bar{x}_h = \bar{x}_\ell$ , i.e.:

$$V(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_h, \bar{x}_\ell) \quad \text{for all } \bar{x}_\ell \leq \bar{x}_h, \quad \text{and} \quad V(x, x) = \Psi(x, x). \tag{51}$$

We will show that  $\Psi$  is maximized at the point  $(\beta^{st}, \beta^{st})$ . Since  $\Psi$  is an upper bound on  $V$  that coincides with  $V$  for  $\bar{x}_h = \bar{x}_\ell$ , this will imply that also  $V$  is maximized at the point  $(\beta^{st}, \beta^{st})$  so that static delegation is optimal. We begin by computing  $\Psi$  explicitly. As argued in the proof of Lemma 7, for all  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h]$ , we have  $x_h(\omega) = \bar{x}_h \leq \bar{\omega}_h$ . Recall, moreover, that  $\ell \cdot U_\ell = \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_\ell(\omega) d\omega$ . This, together with the fact that  $D_\ell = [\underline{\omega}_\ell, \bar{x}_\ell]$  is an interval, implies that

$$\begin{aligned}
\Psi(\bar{x}_h, \bar{x}_\ell) &= \frac{\mu_h}{h} \left[ - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} (\bar{x}_h - \omega)^2 d\omega + b(\bar{x}_h - \bar{\omega}_h)^2 \right] \\
&\quad - \left( \frac{\mu_h}{h} + \frac{\mu_\ell}{\ell} \right) \int_{\bar{x}_\ell}^{\bar{\omega}_\ell} (\bar{x}_\ell - \omega)^2 d\omega + \frac{\mu_\ell}{\ell} b(\bar{x}_\ell - \bar{\omega}_\ell)^2 - b^2. \tag{52}
\end{aligned}$$

We will be interested in the partial derivative of  $\Psi$ . Observe that the partial derivative of  $\Psi$  with respect to  $\bar{x}_h$  is the same as (a multiple of) the derivative of  $hV_h$  with respect to  $\beta^{st}$  in (43). Thus, we can deduce as above:

$$\frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_h} > 0 \quad \Leftrightarrow \quad \bar{x}_h < \xi_h. \tag{53}$$

With these preliminaries, we can now prove the claim.

We consider first the case  $b \geq \hat{b}$ . To show the optimality of static delegation, we show

that when choosing a delegation menu  $(D_h, D_\ell)$  in  $\mathcal{R}$ , it is optimal to choose  $\bar{x}_\ell = \bar{x}_h$ , that is,  $V(\bar{x}_h, \bar{x}_\ell) \leq V(\bar{x}_\ell, \bar{x}_\ell)$  for all  $\bar{x}_\ell \leq \bar{x}_h$ . Since  $V(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_h, \bar{x}_\ell)$  and  $V(\bar{x}_\ell, \bar{x}_\ell) = \Psi(\bar{x}_\ell, \bar{x}_\ell)$ , it is sufficient to show that

$$\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_\ell, \bar{x}_\ell) \quad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \quad (54)$$

Indeed, from part (b),  $\xi_h \leq \beta_\ell^0 \Leftrightarrow b \geq \hat{b}$ . Therefore,  $b \geq \hat{b}$  together with (53) and the fact that  $\beta_\ell^0 \leq \bar{x}_\ell$  implies that

$$\frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_h} \leq 0 \quad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \quad (55)$$

We can deduce that  $\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_\ell, \bar{x}_\ell)$ , establishing (54) as desired.

Next, we show that static delegation is optimal if  $b < \hat{b}$  and  $\mu_h \geq \hat{\mu}_h(b)$ . Because for all  $\bar{x}_\ell \leq \bar{x}_h$ , we have  $V(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_h, \bar{x}_\ell)$ , it is sufficient for the optimality of static delegation that

$$\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\beta^{st}, \beta^{st}) = V(\beta^{st}, \beta^{st}) \quad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \quad (56)$$

We distinguish three cases. Suppose first that  $\bar{x}_\ell \geq \xi_h$ . Then equation (53) yields that  $\partial \Psi / \partial \bar{x}_h \leq 0$  for all  $\bar{x}_h \geq \bar{x}_\ell$ . Thus,  $\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_\ell, \bar{x}_\ell) \leq \Psi(\beta^{st}, \beta^{st})$ .

Second, suppose that  $\bar{x}_h > \xi_h > \bar{x}_\ell$ . Then equation (53) yields that  $\partial \Psi / \partial \bar{x}_h \leq 0$  for all  $\bar{x}_h \geq \xi_h$ . Thus,  $\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\xi_h, \bar{x}_\ell)$ . Thus, it cannot be true that  $\bar{x}_h > \xi_h$  at the optimum.

Third, suppose that  $\bar{x}_h \leq \xi_h$ . By definition,  $\mu_h \geq \hat{\mu}_h(b)$  implies that  $\xi_h \leq \beta^{st}$ . Accordingly,

$$\bar{x}_\ell \leq \bar{x}_h \leq \beta^{st}, \quad (57)$$

and, by (53),

$$\frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_h} \leq 0. \quad (58)$$

Since  $\beta^{st}$  satisfies the first order condition for optimal interval delegation, we also have

$$0 = \frac{dV(\bar{x}, \bar{x})}{d\bar{x}} \Big|_{\bar{x}=\beta^{st}} = \frac{d\Psi(\bar{x}, \bar{x})}{d\bar{x}} \Big|_{\bar{x}=\beta^{st}} = \frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_h} + \frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_\ell}. \quad (59)$$

The two previous properties imply

$$\frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_\ell} \geq 0. \quad (60)$$

Now, by definition of  $\Psi$ , we can compute

$$\begin{aligned} \frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_\ell} &= -\left(\frac{\mu_h}{h} + \frac{\mu_\ell}{\ell}\right) \cdot 2 \cdot \int_{\bar{x}_\ell}^{\bar{\omega}_\ell} (\bar{x}_\ell - \omega) d\omega + \frac{\mu_\ell}{\ell} 2b(\bar{x}_\ell - \bar{\omega}_\ell) \\ &= \left(\frac{\mu_h}{h} + \frac{\mu_\ell}{\ell}\right) (\bar{x}_\ell - \bar{\omega}_\ell)^2 - \frac{\mu_\ell}{\ell} 2b(\bar{\omega}_\ell - \bar{x}_\ell) \\ &= (\bar{\omega}_\ell - \bar{x}_\ell) \left[ \left(\frac{\mu_h}{h} + \frac{\mu_\ell}{\ell}\right) (\bar{\omega}_\ell - \bar{x}_\ell) - \frac{\mu_\ell}{\ell} 2b \right]. \end{aligned} \quad (61)$$

Note that this expression does not depend on  $\bar{x}_h$ . Now, since  $\beta^{st} \leq \bar{\omega}_\ell$ , (60) implies that the square bracket in the last line is non-negative when evaluated at  $\bar{x}_\ell = \beta^{st}$ . It follows that the square bracket in the last line is a fortiori non-negative when evaluated at  $\bar{x}_\ell \leq \beta^{st}$ . Thus, we get

$$\frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_\ell} \geq 0 \quad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \quad (62)$$

Hence,  $\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_h, \bar{x}_h) \leq \Psi(\beta^{st}, \beta^{st})$ . This establishes (56) and completes the proof.  $\square$

*Proof of Proposition 5.* We only consider the case where the optimal static delegation set is the union of an interval plus a point. (The argument for the other case is identical.) Let  $D_\ell = [\underline{\omega}_\ell, \bar{\omega}_\ell - d]$ , and consider the marginal modification defined in (19),

$$D_h(\varepsilon) = [\underline{\omega}_h, \bar{\omega}_\ell - d - \eta] \cup [\bar{\omega}_\ell + d - \varepsilon, \bar{\omega}_\ell + d], \quad \eta, \varepsilon > 0, \quad (63)$$

so that  $(IC_\ell)$  is binding. Observe that for  $\varepsilon = 0$ , the delegation menu  $(D_h(0), D_\ell)$  essentially reduces to the optimal static menu. We show that the principal's expected utility goes up in  $\varepsilon$  by showing that the derivative of the principal's utility with respect to  $\varepsilon$  is strictly positive when evaluated at  $\varepsilon = 0$ . Indeed, conditional on facing the  $\ell$ -type, the principal's utility does not depend on  $\varepsilon$ . Moreover, conditional on facing the  $h$ -type, the principal's utility is given by (17). For  $\varepsilon$  small enough,  $\bar{\omega}_\ell - d - \eta$  is larger than  $\underline{\omega}_\ell$ . Because  $(IC_\ell)$  is binding and because the  $\ell$ -type is indifferent between  $D_\ell$  and  $D_h(0)$ , it must be that the action  $\bar{\omega}_\ell + d - \varepsilon$  is closer to  $\bar{\omega}_\ell$  than is the action  $\bar{\omega}_\ell - d - \eta$ . Hence, the agent chooses action  $\bar{\omega}_\ell + d - \varepsilon$  in the states  $\omega \in [\bar{\omega}_\ell, \bar{\omega}_\ell + d - \varepsilon]$ . Because of this and

since  $\bar{\omega}_\ell + d \leq \bar{\omega}_h$ , (17) writes

$$\begin{aligned}
h \cdot V_h &= \ell \cdot U_\ell - \int_{\bar{\omega}_\ell}^{\bar{\omega}_\ell + d - \varepsilon} (\bar{\omega}_\ell + d - \varepsilon - \omega)^2 d\omega - \int_{\bar{\omega}_\ell + d}^{\bar{\omega}_h} (\bar{\omega}_\ell + d - \omega)^2 d\omega \\
&\quad + b(\bar{\omega}_\ell + d - \omega_h)^2 - b^2.
\end{aligned} \tag{64}$$

The derivative with respect to  $\varepsilon$  is

$$2 \int_{\bar{\omega}_\ell}^{\bar{\omega}_\ell + d - \varepsilon} (\bar{\omega}_\ell + d - \varepsilon - \omega) d\omega = (\bar{\omega}_\ell + d - \varepsilon - \bar{\omega}_\ell)^2, \tag{65}$$

which is strictly positive for  $\varepsilon = 0$ . This establishes the claim.  $\square$

## References

- [1] Alonso, R. and N. Matouschek (2008). Optimal Delegation. *Review of Economic Studies* 75 (1), 259–293.
- [2] Alonso, R., W. Dessein, and N. Matouschek (2008). When does coordination require centralization? *American Economic Review* 98 (1), 145–179.
- [3] Baron, D. and D. Besanko (1984). Regulation and Information in a Continuing Relationship. *Information Economics and Policy* 1, 267–302.
- [4] Battaglini, M. (2005). Long-Term Contracting with Markovian Consumers. *American Economic Review* 95, 637–658.
- [5] Bester, H. (2009). Externalities, Communication and the Allocation of Decision Rights. *Economic Theory* 41, 269–296.
- [6] Blume, A. O. Board, K. Kawamura (2008). Noisy talk. *Theoretical Economics* 2, 395–440.
- [7] Courty, P. and H. Li (2000). Sequential Screening. *Review of Economic Studies* 67, 697–717.
- [8] Crawford, V. P., and J. Sobel (1982). Strategic Information Transmission. *Econometrica* 50, 1431–1451.
- [9] Dai, C., T. R. Lewis, and G. Lopomo (2006). Delegating Management to Experts. *RAND Journal of Economics* 37, 503–520.
- [10] Dessein, W. (2002). Authority and Communication in Organizations. *Review of Economic Studies* 69, 2002, 811–838.
- [11] Esö, P. and B. Szentes (2007a). The Price of Advise. *RAND Journal of Economics* 38, 863–880.
- [12] Esö, P. and B. Szentes (2007b). Optimal Information Disclosure in Auctions and the Handicap Auction. *Review of Economic Studies* 74, 705–731.
- [13] Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009). Mediation, Arbitration, and Negotiation. *Journal of Economic Theory* 144, 1397–1420.
- [14] Holmström, B. (1977). On Incentives and Control in Organizations. Ph.D. thesis, Stanford University.

- [15] Holmström, B. (1984). On the Theory of Delegation. In: *Bayesian Models in Economic Theory*. Ed. by M. Boyer, and R. Kihlstrom. North-Holland, New York.
- [16] Hoffmann, F. and R. Inderst (2009). Price Discrimination and the Provision of Information. mimeo.
- [17] Inderst, R. and M. Peitz (2011). Selling Service Plans to Differentially Informed Customers. mimeo.
- [18] Kartik, N., M. Ottaviani, and F. Squintani (2007). Credulity, lies, and costly talk. *Journal of Economic Theory* 134 (1), 93–116
- [19] Kováč, E. and T. Mylovanov (2009). Stochastic Mechanisms in Settings without Monetary Transfers: The Regular Case. *Journal of Economic Theory* 144 (4), 1373–1395.
- [20] Krähmer, D. (2006). Message-Contingent Delegation. *Journal of Economic Behavior and Organization* 60, 490–506.
- [21] Krähmer, D. and R. Strausz (2008). Ex Post Private Information and Monopolistic Screening. *The B.E. Journal of Theoretical Economics, Topics* 8 (1), Article 25.
- [22] Krähmer, D. and R. Strausz (2011). Optimal Procurement Contracts with Pre-Project Planning. *Review of Economic Studies* 78, 1015–1041.
- [23] Krähmer, D. and R. Strausz (2012). The Benefits of Sequential Screening. mimeo.
- [24] Martimort, D, and A. Semenov (2006). Continuity in Mechanism Design without Transfers. *Economics Letters* 93, 182–189.
- [25] Myerson, R. (1986). Multistage Games with Communication. *Econometrica* 54, 323–358.
- [26] Mylovanov, T. (2008). Veto-based Delegation. *Journal of Economic Theory* 138, 297–307.
- [27] Nocke, V., M. Peitz, and F. Rosar (2011). Advance-purchase discounts as a price discrimination device. *Journal of Economic Theory* 146, 141–162.
- [28] Ottaviani, M. and F. Squintani (2006). Naive audience and communication bias. *International Journal of Game Theory* 35 (1), 129–150.
- [29] Pavan, A., I. Segal, and J. Toikka (2008). Dynamic Mechanism Design: Revenue Equivalence, Profit Maximization and Information Disclosure. mimeo.

- [30] Riordan, M. H. and D. E. Sappington (1987). Information, Incentives and Organizational Mode. *Quarterly Journal of Economics* 102, 243–264.
- [31] Szalay, D. (2005). The Economics of Clear Advice and Extreme Options. *Review of Economic Studies* 72, 1173–1198.