# Optimal Sequential Delegation*

Eugen Kováč[†] Daniel Krähmer[‡]

April 11, 2013

**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 sequential menu of delegation sets where the agent first picks a delegation set and then chooses an action within this set. For the uniform–quadratic case, we characterize when sequential delegation is strictly better than static delegation and derive the optimal delegation menu. We provide sufficient conditions so that our results extend beyond the uniform distribution.

*Keywords:* optimal delegation, sequential screening, dynamic mechanism design, non-transferable utility.
*JEL Codes:* D02, D20, D82, D86.

*The authors would like to thank Ricardo Alonso, Matthias Kräkel, Tymofiy Mylovanov, Juuso Toikka, and seminar audiences in Bonn, Madrid, and Paris for helpful comments. D. Krähmer gratefully acknowledges financial support by the DFG (German Science Foundation) under SFB/TR-15.

[†]Department of Economics, University of Bonn, Adenauerallee 24–42, D-53113 Bonn (Germany); E-mail: eugen.kovac@uni-bonn.de; URL: www.uni-bonn.de/~kovac.

[‡]Department of Economics, University of Bonn, Adenauerallee 24–42, D-53113 Bonn (Germany), E-mail: kraehmer@hcm.uni-bonn.de.

# 1 Introduction

In many important settings, uninformed decision makers rely on informed, but biased experts, and, moreover, the exchange of monetary transfers is impossible. When making strategic business decisions, firm headquarters depend on the superior knowledge of division managers who care for their own division more than for the firm as a whole. Industry regulators seek to base their decisions on privately known firm characteristics, but instead of the firm's profits, want to maximize social welfare. Financial advisors have better information than investors but are often remunerated based on commissions, biasing them towards recommending certain funds. Other examples include politician-lobbyist, patient-doctor, or client-advocate relationships.

Holmström's (1977, 1984) seminal delegation principle implies that an optimal mechanism for the principal in such an environment amounts to delegating decision making to the expert by offering him a set of permissible decisions from which he can freely choose. While more discretion allows the expert to better adapt the decision to his information, the key reason to limit discretion is to constrain the expert in pursuing his partisan objectives. In many natural settings, optimal delegation takes a remarkably simple form, known as "interval delegation": the expert's discretion is only constrained from below and/or above. The desire to limit the expert's discretion can thus provide a rationale for institutions such as price or budget caps, tariff regulations, or minimum wages.

Our point of departure in this paper is the observation that in many situations the relationship between a principal and an expert is dynamic: instead of being perfectly informed at the beginning, the expert obtains relevant information often only during the course of the relationship. A central problem for the principal in such a situation is that at the outset she does not know whether the information the expert will obtain is actually useful for her needs. For example, in a relationship between headquarters and a country division, the division has better information about local demand conditions, like demographics or heterogeneity of consumers. However, when introducing a new product, the country division might a priori not know how the product will be perceived, but will only gradually learn whether the product is suitable for the local consumers.

In such a dynamic situation, how should the principal structure the decision making process to arrive at an optimal decision? This is the question we address in this paper. To study this question, we extend Holmström's delegation principle and show that, in our dynamic environment, any mechanism can be implemented by a *menu* of delegation sets where, initially, the expert chooses a delegation set from the menu, and then, after having obtained new information, he makes a decision within the initially chosen set. The contribution of the paper is to provide a characterization of the optimal menu of delegation sets for the familiar setting with quadratic utility and constant bias where we assume that the expert is biased towards higher decisions.[1] Unlike in the standard setting, the expert privately observes an imperfect signal about the state during the course of the relation, but at the outset has private information about the precision of the signal he will receive. We interpret the signal's precision as the expert's ability and consider the simplest case with two possible expert types: a "good" expert and a "bad" expert.

The good expert receives a more precise signal. This means, his posterior beliefs conditional on his signal will be more dispersed (i.e., better match the true distribution) from an ex ante perspective. Hence, if the expert's ability is publicly known, the good expert is granted more discretion since in this way he can better adapt the decision to the more dispersed beliefs. In fact, in our setting, the good expert is simply offered a larger interval of decisions. The key novel issue that arises when the expert's ability is his *private* information, is whether the principal can and wants to offer a menu of delegation sets to *screen* the expert's ability. Clearly, simply offering the menu that consists of the delegation sets from the public information case is not feasible, because a bad expert would always pretend to be a good expert so as to enjoy the good expert's larger discretion.

Effectively, when the principal screens the expert's ability, she *sequentially* delegates decision making: in the first stage, the expert chooses a delegation set, and in the second stage, he chooses an action from the set. Sequential delegation is feasible only if the menu of delegation sets is incentive compatible so that each expert type has an incentive to select the delegation set which corresponds to his true ability. A key insight of our analysis is that screening the expert's ability through sequential delegation is feasible even though, absent contingent monetary transfers, the number of screening instruments available to the principal is quite limited.[2] In fact, the only instrument available to the principal is the degree of discretion granted to the agent. In order to prevent the bad expert from pretending to be a good one, the principal needs to impose a "cost" on the good expert by excluding some of the decisions from his delegation set. In particular, we show that the optimal way to screen the expert's ability is to offer the good expert a set of extreme decisions and to bar him from making moderate decisions. At the same time, the bad expert is offered an interval of more moderate decisions much like in the case with publicly known abilities.

Intuitively, such a menu is incentive compatible because the bad expert, knowing he receives less precise final information, anticipates that he will likely favor moderate decisions after having observed the signal. Therefore, the bad expert refrains from taking a delegation set that offers only extreme yet no moderate decisions. Reversely, the good

[1]While our main analysis is conducted for the case with a uniformly distributed state, we also provide sufficient conditions so that our results hold for other than uniform distributions.

[2]It is well–known that sequential screening is feasible when contingent monetary transfers are available. See, e.g., Courty and Li (2000).

expert, who anticipates to hold more dispersed beliefs, considers it less likely that his optimal decisions are moderate. Therefore, the good expert prefers a delegation set with extreme decisions to one with moderate decisions.

While this makes clear that sequential delegation is feasible, the key question is whether it is actually optimal to engage in sequential delegation. Instead, the principal can engage in *static* delegation by offering the degenerate menu that consists of a single delegation set only. For the case with a uniformly distributed state, we fully describe when sequential delegation is strictly better than static delegation: First, if the difference between the good and bad expert's ability is sufficiently small compared to the expert's bias, then static delegation is optimal. Second, for sufficiently large ability difference (relative to the bias), sequential delegation becomes optimal. Third, if the ability difference is in an intermediate range, static delegation is optimal whenever the high type is sufficiently likely. Moreover, we show that whenever static delegation is optimal, optimal delegation takes the simple form of interval delegation.[3]

To provide intuition for these results, it is useful to consider the benchmark case in which the expert's ability is publicly known. The design of an optimal delegation set is then determined by the trade–off between utilizing the expert's information (*information effect*) and giving up control (*loss-of-control effect*). The loss-of-control effect makes the principal averse to delegate large decisions, because the expert prefers, from the principal's perspective, excessively large decisions. Thus the larger the conflict of interest (i.e., the bias), the smaller the decisions the principal delegates. Moreover, as indicated earlier, as the information effect is stronger for the good expert, the principal would offer a larger delegation set to the good expert in the benchmark case.

For the case that the agent's ability is his private information, we first derive the optimal static delegation set. We obtain the intuitive result that, compared to the case with publicly observable ability, under the optimal static delegation set too many decisions are delegated to the bad and too few to the good expert type. Relative to static delegation, sequential delegation involves the following trade-off: On the one hand, offering extreme decisions to the good expert better utilizes this type's superior information, because due to his more dispersed beliefs, the optimal decision is more likely to be an extreme one. On the other hand, disallowing moderate decisions is also costly if the signal indicates that a moderate decision should be taken. The three cases identified earlier determine when the benefits outweigh the costs of sequential delegation in the uniform case. For more general distributions, we establish sufficient conditions for the optimal sequential delegation menu to have qualitatively the same shape as in the uniform case. The condition guarantees that precluding moderate options from a delegation set hurts the bad expert more than

[3]We also show, however, that when static delegation is suboptimal, the optimal static delegation set may not be an interval.

the principal, if the latter believes to face the good expert. We also show that in this case there are always parameters so that sequential delegation is in fact optimal.

From an applied perspective, our results show both the robustness and the limits of the optimality of interval delegation in a dynamic setting. When static delegation is optimal, interval delegation survives as the optimal mechanism. This confirms the importance of simple delegation mechanisms in the form of caps or tariffs. On the other hand, our results also provide a justification for why the discretion of experts may be limited in more complex ways. For example, headquarters of multinational firms may delegate the restructuring of local business operations to national managers leaving them a choice between either cutting costs across all local activities or terminating a certain number of activities of their choosing. The flavor of our result resonates in plea bargaining procedures in the U.S. criminal law where the defendant, or the coalition between the defendant and the prosecutor (the principal), can propose an agreement to the court (the agent) in which the defendant pleads guilty and, if the court accepts the proposal, the court commits to a set of relatively mild sentences. If the court rejects the proposal, the court is free to implement extreme sentences instead. Our interpretation is that the court accepts the agreement if the case is intransparent and the court knows to receive only little conclusive evidence about the defendant's guilt, and the court rejects the agreement if it knows to obtain sufficient evidence to convict or acquit the defendant.

*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 Matouscheck (2008) and Amador and Bagwell (2012a) provide conditions that characterize the optimal delegation set in general environments, including characterizations for when simple interval delegation is optimal.[4] 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. In recent work, Semenov (2012) considers a setup with an expert who may know the state either fully or not at all. While this setup can be nested within our model (when the good expert fully learns the state and the bad expert does not learn at all), it is ruled out by assumption in our paper. Moreover, Semenov (2012) considers only static delegation and does not allow for the agent's expertise to be screened. In this sense, Semenov (2012) complements our analysis of static delegation.

[4]See also Martimort and Semenov (2006). For an application of optimal delegation to international trade agreements, see Amador and Bagwell (2012b).

Our result that sequential delegation requires the principal offer delegation sets with only extreme options is somewhat 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 screens the agent's ex ante expertise. Moreover, Szalay (2005) considers an unbiased agent which in our setup would lead to full delegation, without screening, to be optimal.

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.[5]

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.[6] 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.[7]

This paper is organized as follows. 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

[5]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, for instance, Riordan and Sappington (1987), Dessein (2002), Krähmer (2006), Mylovanov (2008), Semenov (2008), Bester (2009), Goltsman, Hörner, Pavlov, and Squintani (2009) to name only a few.

[6]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 (2012), Inderst and Hoffmann (2011), Nocke, Peitz, and Rosar (2011), Inderst and Peitz (2012).

[7]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.

only static delegation is feasible. Section 7 fully describes the solution to the sequential delegation problem for the uniform distribution. Section 8 contains the extension to non–uniform distributions of the state. Section 9 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 $\tilde{\theta}$. We assume $\tilde{\theta}$ to be uniformly distributed on the unit interval. (This assumption will be relaxed in Section 8.) 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, but stochastically independent from, the true state. The agent does not know whether the realization of the signal is true or noise.[8] Thus, $p$ is a measure for the agent's ability, and we refer to $p$ as the agent's ability type. We assume that the type can be either "high" or "low": $p \in \{h, \ell\}$, where $h > \ell$. 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.[9,10] We assume in the whole paper that

$$b \leq \ell/2. \tag{1}$$

This assumption reduces the number of case distinctions and is not substantial. It will guarantee that when the agent's ability is publicly known, the optimal delegation set is a non-degenerate interval (see Lemma 4).[11]

[8]This signal structure is sometimes referred to as the "replacement noise" model.

[9]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).

[10]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.

[11]In the terminology of Alonso and Matouscheck (2008), assumption (1) means that delegation is

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,

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

is a sufficient statistics for the signal. In what follows, it will be more convenient to work with this sufficient statistics. 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

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

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.[12]

0. The principal commits to a mechanism.

1. The agent privately observes his ability type $p \in \{h, \ell\}$.

2. The agent privately observes the expected state $\omega \in [\underline{\omega}_p, \bar{\omega}_p]$.

valuable when facing the low type. Complementary to our analysis, Semenov (2012) derives the optimal static delegation set for the case that $h = 1$ and $\ell = 0$ (where assumption (1) does not hold), but does not consider sequential delegation.

[12]The timing implicitly assumes that the principal can clearly distinguish between periods 1 and 2. This is natural in many applications and is also the standard assumption in sequential screening models with money (see, for instance, Courty and Li (2000)).

3. An action is implemented according to the terms of the mechanism.

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 with quadratic preferences, constant bias, and two ability types and is likely to be even more so in more general environments. In the case with a uniformly distributed state, we will fully characterize whether a sequential or a static mechanism is optimal in our setup. In Section 8, we go beyond the uniform distribution and provide a set of sufficient conditions so that the main insights of our analysis for the uniform case extend to more general distributions.

# 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 $M$ is a pair $(\xi_h, \xi_\ell)$ of mappings with

$$\xi_{\hat{p}} : [\underline{\omega}_{\hat{p}}, \bar{\omega}_{\hat{p}}] \to \mathbb{R}, \qquad \hat{p} \in \{h, \ell\}.$$

A direct mechanism 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 $\xi_{\hat{p}}(\hat{\omega})$.[13]

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

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

[13] We restrict attention to deterministic mechanisms. For an investigation of stochastic mechanisms in the static delegation problem, see Kováč and Mylovanov (2009).

Observe that the revelation principle does not require truth-telling off the equilibrium path, i.e., after a lie in period 1.[14]

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 $p, \hat{p}$,

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

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:

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

# 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

[14] 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 Krähmer and Strausz (2008) for an elaboration of this point in the context of a principal agent problem with money.

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[15,16]

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

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

$$-\int_{\underline{\omega}_p}^{\bar{\omega}_p} (x_p(\omega) - \omega)^2 \, \frac{1}{p} \, d\omega \geq -\int_{\underline{\omega}_p}^{\bar{\omega}_p} (x_{\hat{p}}(\omega) - \omega)^2 \, \frac{1}{p} \, d\omega \qquad \text{for all } \hat{p} \in \{h, \ell\}. \tag{$IC_p$}$$

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 $x_p(\omega) = \xi_p(\omega)$ satisfies $(A)$. Moreover, (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 sequential delegation problem:

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

[15]Without loss of generality, we may assume that delegation sets are closed and bounded. In that case the arg max is well defined. See footnote 15 in Alonso and Matouscheck (2008).

[16]This is with slight abuse of notation. Strictly speaking, the agent's optimal action is a function of the delegation set.

In general, the solution to $\mathcal{D}$ will not be unique.[17] 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.[18,19]

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 $D_p$ which we shall use throughout the paper:[20]

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

**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 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*).

[17] Analogously as in Holmström (1984), Theorem 1, it can be shown that a solution to the sequential delegation problem $\mathcal{D}$ indeed exists.

[18] Alonso and Matouscheck (2008) proceed in the same fashion.

[19] 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 agent type $p$'s incentives to pick the delegation set $D_{\hat{p}}$.

[20] Note that both the minimum and maximum exist, as $D_p$ is closed and bounded.

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[21]

$$u_p(\omega_2) - u_p(\omega_1) = \int_{\omega_1}^{\omega_2} 2(x_p(\omega) - \omega)\, d\omega \qquad \text{for all } \omega_1, \omega_2. \tag{5}$$

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

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

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 $-b$.

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 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 appear repeatedly throughout the rest of the paper:

$$\beta_p^0 \equiv \bar{\omega}_p - 2b. \tag{7}$$

**Lemma 2.** *Consider an arbitrary menu $(D_h, D_\ell)$ of delegation sets.*[22] *Let $p \in \{h, \ell\}$ and*

[21] 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 $u'_p(\omega) = 2(x_p(\omega) - \omega)$, whenever the derivative exists.

[22] The menu of delegation sets is not required to be incentive compatible here.

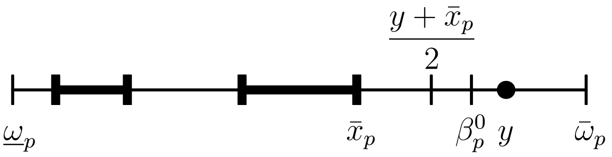


Figure 1: Illustration of Lemma 2, (ii).

*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.,*

$$\underline{\omega}_p - |\underline{\omega}_p - \underline{x}_p| < y < \bar{\omega}_p + |\bar{\omega}_p - \bar{x}_p|. \tag{8}$$

(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 (8).

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 (see Figure 1). Adding such an action increases both the information effect (as the utility of all agent types $\omega$ 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.

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

The set $\tilde{D}_p$ 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 $\beta_p^0$, as defined in (7).[23] 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.[24] 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)).

[23]See, for example, Holmström (1977) or Kováč and Mylovanov (2009).

[24]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 $D$.[25] In terms of the mechanisms design perspective, this means that both pieces of the agent's private information are elicited at the same time.

The principal's utility under a static delegation set is

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

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

$$F^{st}(\omega) = \mu_h F_h(\omega) + \mu_\ell F_\ell(\omega)$$

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

$$\begin{aligned} V = & \frac{\mu_h}{h} bu(\underline{\omega}_h) + \frac{\mu_\ell}{\ell} bu(\underline{\omega}_\ell) + \int_{\underline{\omega}_h}^{\bar{\omega}_h} u(\omega) \, dF^{st}(\omega) \\ & - \frac{\mu_h}{h} bu(\bar{\omega}_h) - \frac{\mu_\ell}{\ell} bu(\bar{\omega}_\ell) \\ & - b^2, \end{aligned} \tag{9}$$

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 (9): conditional on facing the $h$-type, the principal's utility goes down in state $\bar{\omega}_h$, and conditional on facing the $\ell$-type, the principal's utility goes down in state $\bar{\omega}_\ell$. 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 smaller than the action $\bar{\omega}_\ell$, (9) 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

[25] For example, the menu of intervals $D_h = [\underline{\omega}_h, x]$ and $D_\ell = [\underline{\omega}_\ell, x]$ can be implemented by the single set $D = [\underline{\omega}_h, x]$.

on each type, 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 the interval $(\bar{\omega}_\ell, \bar{\omega}_h)$), or empty. The next two propositions make this reasoning precise.

**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*

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

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

To see the intuition for the proposition, observe first that the condition $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ stated in the proposition is equivalent to the condition $\beta^0_h \leq \bar{\omega}_\ell$. Yet if $\beta^0_h \leq \bar{\omega}_\ell$, then, intuitively, for actions larger than $\bar{\omega}_\ell$ the loss-of-control effect dominates the information effect both conditional on facing the $\ell$- and also the $h$-type. Hence, it would be suboptimal to include actions larger than $\bar{\omega}_\ell$ in the delegation set. That the optimal delegation set is then an interval follows from the remarks preceding the proposition.

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 $\beta^0_h > \bar{\omega}_\ell$ (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 $\bar{\omega}_\ell$. More precisely, there is $d \in (0, 2b)$ so that*

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

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

To see the intuition, recall from the arguments preceding Proposition 1 that the optimal delegation set is an interval, or the union of an interval and a point or of two

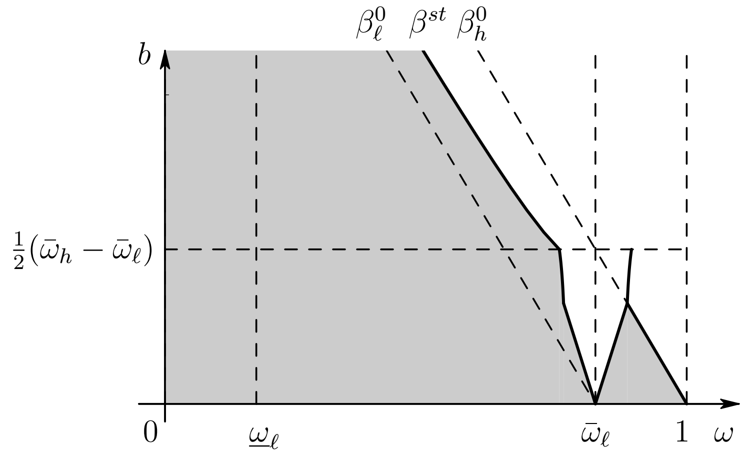


Figure 2: Optimal static delegation set (example for $h = 1$, $\ell = 0.65$, $\mu_h = 0.4$).

intervals. Observe that if interval delegation was optimal, the arguments above imply that $\bar{x} < \bar{\omega}_\ell$. In this case, we could add the action $y = \bar{\omega}_\ell + (\bar{\omega}_\ell - \bar{x})$ to the delegation set without affecting the principal's payoff conditional on facing type $\ell$. However, since $\beta_h^0 > \bar{\omega}_\ell$, we have that $\frac{1}{2}(\bar{x} + y) = \bar{\omega}_\ell < \beta_h^0$, and hence Lemma 2, (ii) implies that adding the action $y$ would improve the principal's utility conditional on facing type $h$.

Figure 2 illustrates the optimal static delegation menu (obtained numerically). Fixing the values of $h$, $\ell$, and $\mu_h$, consider some $b$ on the vertical axis and draw the corresponding horizontal line. The optimal static delegation set is then the intersection of that line with the shaded region.[26] For $b > \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$, the optimal static delegation set is an interval (Proposition 1). For lower values of $b$, it is a union of an interval and a point, whereas for even lower values it is a union of two intervals (Proposition 2).[27]

*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), by Kováč and Mylovanov (2009), and by Amador and Bagwell (2012a). These papers derive conditions for the optimality of interval delegation for the case that the density associated with the distribution of the state is continuous. We cannot directly apply these results, because in our case, $f^{st}$ is not continuous, as it has jumps at the points $\omega = \underline{\omega}_\ell$ and $\omega = \bar{\omega}_\ell$; thus, $f^{st}$ is only piece-wise continuous. In principle, the methods of Kováč and Mylovanov (2009) can be extended even if the density is not continuous as long at the environment is "regular". Yet, 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

[26]The horizontal axis depicts the state and the actions. Note that for $h = 1$, we have $\underline{\omega}_h = 0$ and $\bar{\omega}_h = 1$. The dashed diagonal lines (labeled $\beta_\ell^0$ and $\beta_h^0$) show the upper endpoints of the optimal delegation sets if the agent's type is publicly known.

[27]Note that while the delegation set is discontinuous at $b = \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$, the principal's expected utility is continuous. If $\bar{x} = \bar{\omega}_\ell - d$, then adding the action $\bar{\omega}_\ell + d$ does not influence the principal's expected payoff, as $\frac{1}{2}(\bar{\omega}_\ell - d + \bar{\omega}_\ell + d) = \bar{\omega}_\ell = \beta_h^0$ (see Lemma 2, (ii)).

to go beyond this case is the difficulty to characterize the optimal static delegation set.[28]

# 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 at the optimum, 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 (Lemmas 5 and 6). 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 (Lemma 7). 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

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

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 $(IC_\ell)$ is binding, and moreover that the low ability type is offered an interval at the optimum of the relaxed problem.

[28]More precisely, in Kováč and Mylovanov (2009), regularity means that the function $g(\omega) \equiv 1 - F^{st}(\omega) - bf^{st}(\omega)$ 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 $f^{st}$ 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 and sequential delegation becomes interesting exactly in irregular cases. Therefore, excluding irregular cases would be a loss.

**Lemma 5.** *At the solution to the relaxed problem $\mathcal{D}^R$, $(IC_\ell)$ is binding and $D_\ell^*$ is an interval of the form*

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

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).

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 6.** *Without loss of generality, the high type's delegation set at the optimum of the relaxed problem $\mathcal{D}^R$ is a union of at most two connected sets. If $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$, then $D_h^*$ takes one of the following two forms:*

(i) $D_h^* = \{\gamma_1\} \cup \{\gamma_2\}$, *where* $\gamma_1 \leq \underline{\omega}_h < \gamma_2$.

(ii) $D_h^* = [\underline{\omega}_h, \gamma_1] \cup \{\gamma_2\}$, *where* $\underline{\omega}_h < \gamma_1 \leq \gamma_2$.

*If $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$, then $D_h^*$ is of the form (ii) or of the form:*

(iii) $D_h^* = [\underline{\omega}_h, \gamma_1] \cup [\gamma_2, \beta_h^0]$, *where* $\underline{\omega}_h < \gamma_1 \leq \bar{\omega}_\ell \leq \gamma_2 < \beta_h^0$.

*Moreover, in all cases the following inequalities hold: $\gamma_1 \leq \bar{x}_\ell^* \leq \gamma_2$ and $\frac{1}{2}(\bar{x}_\ell^* + \gamma_2) \leq \bar{\omega}_\ell$.*

Let us emphasize that the optimal static delegation sets from Propositions 1 and 2 are included in the class of menus described in Lemma 6: Interval delegation occurs if $D_h^*$ is of the form (ii) and $\gamma_1 = \gamma_2 = \bar{x}_\ell^*$. Moreover, the static delegation set from Proposition 2 occurs if $\gamma_1 = \bar{x}_\ell^*$ and $\gamma_2 = \bar{\omega}_\ell + d$ with $d = \bar{\omega}_\ell - \bar{x}_\ell^*$.

Note also that Lemma 6 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 $\tilde{D}_h$ of the form described in Lemma 6 such that the menu $(\tilde{D}_h, D_\ell^*)$ is optimal as well.

The intuition is as follows. Given the delegation menu $(D_h^*, D_\ell^*)$, we construct the set $\tilde{D}_h$ so that it (weakly) improves the principal's expected utility and at the same time preserves the binding $(IC_\ell)$. It is easiest to illustrate the argument for the case when the largest action in $D_h^*$, action $\bar{x}_h^*$, is chosen by the agent in state $\bar{\omega}_\ell$ from $D_h^*$. In the proof

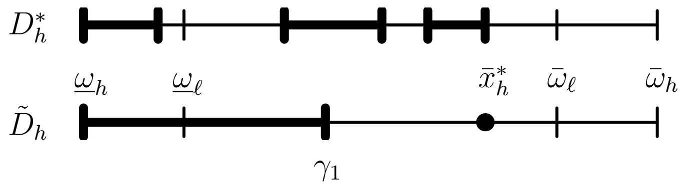


Figure 3: Construction of the set $\tilde{D}_h$ from $D_h^*$ (Lemma 6).

of the lemma, we show that optimality requires that the smallest action in $D_h^*$, action $\underline{x}_h^*$, is not larger than $\underline{\omega}_h$. For simplicity, consider now only the case $\underline{x}_h^* = \underline{\omega}_h$. Then $[\underline{\omega}_h, \bar{x}_h^*] \supseteq D_h^* \supseteq \{\underline{\omega}_h, \bar{x}_h^*\}$, and thus, type $\ell$ (weakly) prefers the interval $[\underline{\omega}_h, \bar{x}_h^*]$ over $D_h^*$ and (weakly) prefers $D_h^*$ over the two-action set $\{\underline{\omega}_h, \bar{x}_h^*\}$. Hence, since $(IC_\ell)$ is binding, type $\ell$ also prefers $[\underline{\omega}_h, \bar{x}_h^*]$ over $D_\ell^*$, and $D_\ell^*$ over $\{\underline{\omega}_h, \bar{x}_h^*\}$. Therefore, an intermediate value argument implies that there is an action $\gamma_1 \in [\underline{\omega}_h, \bar{x}_h^*]$ such that type $\ell$ remains indifferent between the set $D_\ell^*$ and the set $[\underline{\omega}_h, \gamma_1] \cup \{\bar{x}_h^*\}$. We now choose $\tilde{D}_h$ to be equal to $[\underline{\omega}_h, \gamma_1] \cup \{\bar{x}_h^*\}$ and consider the modified delegation menu $(\tilde{D}_h, D_\ell^*)$. The delegation set for the $h$-type is thus of the form (ii) as stated in Lemma 6 with $\gamma_2 = \bar{x}_h^*$ (see Figure 3). Moreover, by construction, $(IC_\ell)$ is binding under the modified delegation menu.

We now argue that the principal's expected utility (weakly) improves by the modification. While, by construction, the value from $D_h^*$ and $\tilde{D}_h$ is the same conditional on the *low* agent type's belief about the expected state, all that matters for the principal is the utility from $D_h^*$ and $\tilde{D}_h$, 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 $\tilde{D}_h$ 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. On the other hand, if it does affect the $h$-type's choices in states $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell]$, the delegation set $\tilde{D}_h$ (weakly) improves the principal's expected utility.[29]

Lemma 6 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 6 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 7.** *A solution $(D_h^*, D_\ell^*)$ to the relaxed problem $\mathcal{D}^R$ of the form as in Lemmas 5 and 6 is also a solution to the original problem $\mathcal{D}$.*

The basic idea behind Lemma 7 is as follows. Suppose there is a sequential solution

[29]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. In Section 8, we provide a sufficient condition so that the argument extends to the non-uniform case.

$(D_h^*, D_\ell^*)$ to the relaxed problem that violates $(IC_h)$.[30] Then, since $(IC_h)$ is violated, the $h$-type's utility would be improved when the agent was offered the static menu $(\tilde{D}_h, D_\ell^*)$, where $\tilde{D}_h = [\underline{\omega}_h, \underline{\omega}_\ell] \cup D_\ell^* = [\underline{\omega}_h, \bar{x}_\ell^*]$.[31] 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 6, the utility of the highest type $\bar{\omega}_h$ is higher under $D_h^*$ than under $\tilde{D}_h$, and therefore, the loss-of-control effect is smaller under $\tilde{D}_h$ than under $D_h^*$. Consequently, if $(IC_h)$ is violated, offering the static menu $(\tilde{D}_h, D_\ell^*)$ is an improvement in terms of both the information and the loss-of-control effect.

Lemma 6 and Lemma 7 imply that the optimal delegation menu can be found by optimizing over the class described in Lemma 6. According to Lemma 6, the properties of the $h$–type's optimal delegation set $D_h^*$ depend on whether, for the static problem, interval delegation is optimal or not (see Propositions 1 and 2). We now 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,

$$\beta_h^0 \leq \bar{\omega}_\ell. \tag{10}$$

Propositions 3 and 4 below 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 6 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 optimal.

**Proposition 3.** *Let $b \geq \bar{\omega}_h - \bar{\omega}_\ell$. Then static interval delegation is optimal for 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*[32]

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

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

---

[30] Clearly, a static solution meets $(IC_h)$.

[31] This menu is indeed static, as it can be implemented by offering the delegation set $\tilde{D}_h = [\underline{\omega}_h, \bar{x}_\ell^*]$ to both types.

[32] Observe that $\frac{1}{2}(1+\sqrt{5}) \approx 1.62$ is the *Golden ratio.*

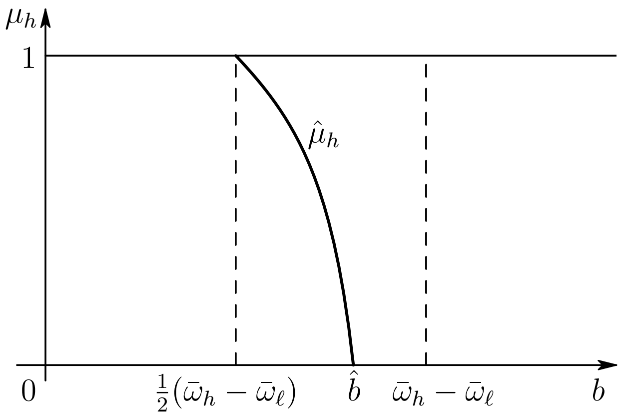


Figure 4: Illustration of Propositions 3 and 4 (for $h = 1$, $\ell = 0.65$).

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

(ii) *If $b < \hat{b}$ and $\mu < \hat{\mu}_h(b)$, then sequential delegation is optimal for problem $\mathcal{D}$.*

Figure 4 illustrates Propositions 3 and 4 graphically. In the range $b \geq \bar{\omega}_h - \bar{\omega}_\ell$, static delegation is optimal. When $\bar{\omega}_h - \bar{\omega}_\ell > b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$, sequential delegation is optimal if and only if $\mu_h$ lies below the $\hat{\mu}_h$ curve. (For the range $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ we show in the next subsection that sequential delegation is optimal.)

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 $[\underline{\omega}_\ell, \bar{\omega}_\ell]$:

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

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

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

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 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 states $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell) \cup (\bar{\omega}_\ell, \bar{\omega}_h)$,

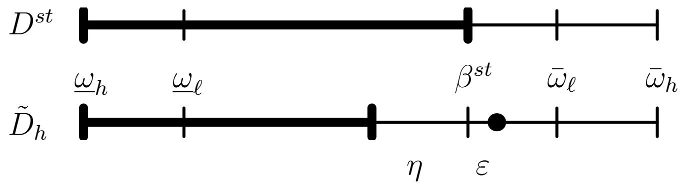


Figure 5: Marginal sequential delegation.

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 $D^{st} = [\underline{\omega}_h, \beta^{st}]$ by Proposition 1. Now suppose that, instead of offering the static menu $(D^{st}, D^{st})$, the principal offers the sequential delegation menu where the $\ell$-type's delegation set is maintained and the upper endpoint of the $h$-type's delegation set is slightly increased by $\varepsilon$ and an appropriate gap with length $\eta + \varepsilon$ (where $\eta$ depends on $\varepsilon$) is inserted in $D_h$ to maintain incentive compatibility (see Figure 5):

$$\tilde{D}_\ell = D^{st}, \quad \tilde{D}_h(\varepsilon) = [\underline{\omega}_h, \beta^{st} - \eta] \cup \{\beta^{st} + \varepsilon\}. \tag{14}$$

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 7 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 the increase of the upper endpoint by $\varepsilon$ is only marginal, 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 states $\omega \in [\underline{\omega}_h, \underline{\omega}_\ell)$ (the second 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 (10). Therefore, in states $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h]$, 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 (recall that $\eta$ is a function of $\varepsilon$):

$$h\,\frac{d\tilde{V}_h(\varepsilon)}{d\varepsilon}\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon}\left\{\underbrace{\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} -(\beta^{st}+\varepsilon-\omega)^2\,d\omega}_{\substack{\text{information}\\ \text{effect}}} - \underbrace{b[-(\beta^{st}+\varepsilon-\bar{\omega}_h)^2]}_{\substack{\text{loss-of-control}\\ \text{effect}}}\right\}\Big|_{\varepsilon=0}, \tag{15}$$

where $\tilde{V}_h(\varepsilon)$ is the principal's expected utility from the delegation set $\tilde{D}_h(\varepsilon)$ defined in (14). 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 the upper endpoint of the $h$-type's delegation set is marginally increased from $\beta^{st}$ to $\beta^{st}+\varepsilon$. Yet, the *difference* in the increase is smaller 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, marginally increasing the upper endpoint of the $h$-type's delegation 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 $b \geq \bar{\omega}_h - \bar{\omega}_\ell$.

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 the parameter constellation in part (ii) of Proposition 4, 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 in part (i) of Proposition 4, 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 6 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,

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

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 $\mathcal{D}$.*

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:

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

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):

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

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: $\tilde{D}_\ell = [\underline{\omega}_h, \bar{\omega}_\ell - d]$. Further, the delegation set of the high type is obtained from $D^{st}$ 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:

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

We now argue that this modification is profitable for the principal. 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 7, also for the original problem.

To see that the modification is profitable, observe that because $\tilde{D}_\ell$ 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 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 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 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 states $\omega \in [\bar{\omega}_\ell + d - \varepsilon,\, \bar{\omega}_\ell + d)$, the agent gets his ideal action $\omega$ under the

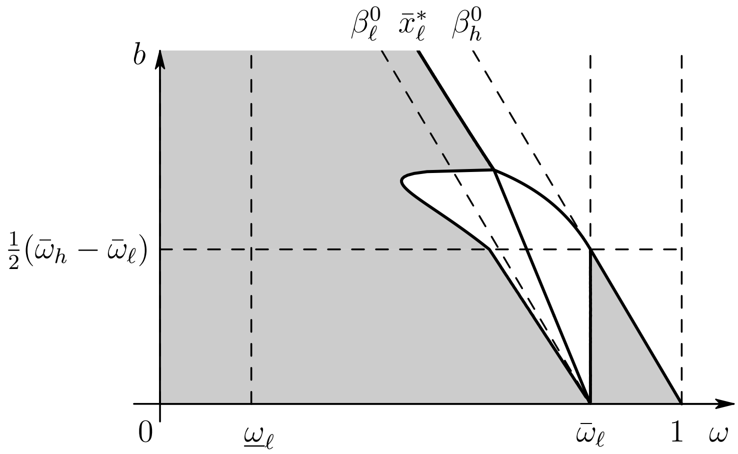


Figure 6: Optimal sequential delegation menu (example for $h = 1$, $\ell = 0.65$, $\mu = 0.4$).

modification, whereas under static delegation he chooses the worse action $\bar{\omega}_\ell + d$. Finally, in 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.

Figure 6 illustrates the optimal delegation menu (for the same values of $h$, $\ell$, and $\mu$ as in Figure 2). The intersection of a horizontal line (for some $b$) with the shaded region shows the $h$-type's delegation set $D_h^*$ at the optimal delegation menu. In addition, the figure depicts the upper endpoint $\bar{x}_\ell^*$ of the optimal delegation set $D_\ell^* = [\underline{\omega}_\ell, \bar{x}_\ell^*]$ for the $\ell$-type. For $b$ sufficiently large, static delegation is optimal. For intermediate values of $b$, the $h$-type's optimal delegation set is a union of an interval and a point, while for $b$ sufficiently small it becomes a union of two intervals.[33]

# 8 Extension beyond the uniform distribution

In this section, we provide a generalization of our results to the case that the state of the world is not uniformly distributed. Instead, we assume that $\theta$ is distributed according to some cumulative distribution function $F$ on the unit interval with full support. We denote by $f = F'$ the density, and by $\phi$ the mean of this distribution. As before, let $\tilde{\omega} = p\tilde{s} + (1-p)\phi$ be the expected state, conditional on the signal $\tilde{s}$, where $\tilde{s}$ is identically distributed as but stochastically independent from the state. Then the random variable

[33]The figure suggests that whenever $D_h^*$ is a union of two intervals, then the higher interval is of the form $[\bar{\omega}_\ell, \beta_h^0]$. Indeed, the fact that its upper endpoint is $\beta_h^0$ follows from Lemma 6, form (iii). In addition, it can be shown that if $\underline{\omega}_\ell < \gamma_1$ (at the optimum), then $\gamma_2 = \bar{\omega}_\ell$.

$\tilde{\omega}$ is distributed with cumulative distribution function and density

$$F_p(\omega) = F\Big(\phi + \frac{\omega - \phi}{p}\Big) \qquad \text{and} \qquad f_p(\omega) = \frac{1}{p} \cdot f\Big(\phi + \frac{\omega - \phi}{p}\Big) \tag{20}$$

on the support $[\underline{\omega}_p, \bar{\omega}_p] = [(1-p)\phi,\, p + (1-p)\phi]$, and has the mean $\phi$.

We can apply the same steps as in the uniform case to obtain the analogue to the decomposition (6) for the principal's expected utility conditional on facing type $p$:

$$V_p = bu_p(\underline{\omega}_p)f_p(\underline{\omega}_p) + \int_{\underline{\omega}_p}^{\bar{\omega}_p} u_p(\omega)[f_p(\omega) + bf_p'(\omega)]\, d\omega - bu_p(\bar{\omega}_p)f_p(\bar{\omega}_p) - b^2. \tag{21}$$

To highlight the essential issues, we focus our discussion on the case that interval delegation is optimal both when the expert's ability is public information and under static delegation when ability is private information (corresponding to the "large bias" case with the uniform distribution). A set of sufficient conditions for this to be the case is the following.

**A1.** *$b < \ell\phi$.*

**A2.** *The density $f$ is non-decreasing and differentiable with a bounded derivative on $[0,1]$.*

**A3.** *$\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} [1 - F_h(\omega) - bf_h(\omega)]\, d\omega < 0$.*

**A4.** *$\dfrac{\int_z^1 [1 - F(\omega)]\, d\omega}{1 - F(z)}$ is convex in $z$ on $[0,1)$.*

Assumption A1 is a generalization of our assumption (1). In the terminology of Alonso and Matouscheck (2008), it implies that delegation is valuable conditional on each type. Assumption A2 ensures that the first two terms in (21) are positive and therefore correspond to an information effect. This implies that in the case with publicly known ability $p$, inserting gaps in the delegation set is never optimal. Assumption A3 corresponds to the case where the bias is large (as in Proposition 1 and in Section 7.1).[34] Finally, Assumption A4 is an additional technical assumption.

Assumptions A1–A4 allow us to prove the following two lemmas.

**Lemma 8.** *Let Assumptions A1–A4 hold. If the expert's ability $p$ is public information, then interval delegation is optimal with optimal delegation set $D_p = [\underline{\omega}_p, \beta_p^0]$, where $\beta_p^0 \in (\underline{\omega}_p, \bar{\omega}_p)$ is the unique solution to the equation*

$$\int_{\beta}^{\bar{\omega}_p} [1 - F_p(\omega) - bf_p(\omega)]\, d\omega = 0. \tag{22}$$

*Moreover, $0 < \beta_h^0 - \beta_\ell^0 \leq \bar{\omega}_h - \bar{\omega}_\ell$.*

[34] For the uniform distribution, Assumption A3 becomes $(\bar{\omega}_h - \bar{\omega}_\ell)(\bar{\omega}_h - \bar{\omega}_\ell - 2b)/(2h) < 0$.

**Lemma 9.** *Let Assumptions A1–A4 hold. If $p$ is private information of the expert, then the optimal static delegation set is an interval $D^{st} = [\underline{\omega}_h, \beta^{st}]$ where $\beta^{st} \in (\beta_\ell^0, \beta_h^0)$.*

We next offer a sufficient condition such that our insight from Lemma 5 and 6 still remains valid that the optimal delegation menu consists of two disjoint connected sets for the high and an interval for the low type. Intuitively, for the same reason as in the uniform case, the critical incentive constraint is the one for the low ability type. Hence, we consider the relaxed problem where we ignore the high type's incentive constraint. As in the uniform case, it is then easy to see that the low type's delegation set is an interval at the optimum. Now assume, contrary to what we claim, that $D_h$ does not consist of two disjoint connected sets and contains several gaps. Then we can "merge" the several gaps in $D_h$ into a single gap and move it to the upper end so that the resulting set is an interval plus a point. By the same intermediate value argument as in the uniform case we can perform this modification so that the incentive constraint for the low type maintains to be binding.

The key question is then how this modification affects the principal's expected utility. Effectively, the modification shifts the loss due to gaps from relatively low states to higher states $\omega$. Therefore, if the modification is to be profitable, then the principal, conditional on facing the high type, needs to benefit more than the low type agent—when selecting the delegation set $D_h$—from suffering losses in relatively high than in low states. By (21), the principal weighs the loss $u_h(\omega)$ in state $\omega$ with the factor $f_h(\omega) + bf_h'(\omega)$. In contrast, the low type agent—when picking $D_h$—weighs the loss $u_h(\omega)$ in state $\omega$ with the factor $f_\ell(\omega)$. Hence, if the low type agent is indifferent between the losses suffered in the original delegation set (i.e., in relatively small states) and the modified delegation set (i.e., in relatively large states), the principal gains from the modification if $f_\ell(\omega)$ increases faster than $f_h(\omega) + bf_h'(\omega)$, that is, if

$$\frac{f_h(\omega) + bf_h'(\omega)}{f_\ell(\omega)} \quad \text{is non-increasing in } \omega. \tag{23}$$

Thus, if (23) holds, then the optimal delegation menu for the relaxed problem consists of two disjoint connected sets for the high and an interval for the low type. As it turns out, this delegation menu is also a solution to the original problem, if

$$\frac{f_h(\omega) + bf_h'(\omega)}{f_h(\omega)} \quad \text{is non-increasing in } \omega. \tag{24}$$

A sufficient condition for (23) and (24) to hold independently of $b$ is the following.[35,36]

**A5.** *$f$ is log-concave on $[0,1]$ and $f_h(\omega)/f_\ell(\omega)$ is non-increasing on $[\underline{\omega}_\ell, \bar{\omega}_\ell]$.*

We formally confirm the previous considerations in the next lemma which provides a generalization of Lemmas 6 and 7.

**Lemma 10.** *Let Assumptions A1–A5 hold. Then, at the optimal delegation menu, the $\ell$–type's delegation set is an interval $D^*_\ell = [\underline{\omega}_\ell, \bar{x}^*_\ell]$ with $\bar{x}^*_\ell \in [\beta^0_\ell, \bar{\omega}_\ell]$, and the $h$–type's delegation set is without loss of generality a union of two connected sets of the form*

(i) $D^*_h = \{\gamma_1\} \cup \{\gamma_2\}$, *where* $\gamma_1 \leq \underline{\omega}_h < \gamma_2$, *or*

(ii) $D^*_h = [\underline{\omega}_h, \gamma_1] \cup \{\gamma_2\}$, *where* $\underline{\omega}_h < \gamma_1 \leq \gamma_2$.

Lemma 10 implies that if sequential delegation is optimal, then the delegation set of the high type can be taken, without loss of generality, as (possibly degenerate) interval plus a point. We now address the question when sequential delegation is optimal. We first show that for sufficiently large bias (corresponding to the condition $b \geq \bar{\omega}_h - \bar{\omega}_\ell$ for the uniform distribution), static delegation is optimal, generalizing Proposition 3.

**Proposition 6.** *Let Assumptions A1–A5 hold. In addition, let $1 - F_h(\bar{\omega}_\ell) - bf_h(\bar{\omega}_\ell) \leq 0$. Then static interval delegation is optimal for the problem $\mathcal{D}$. The optimum is achieved at the delegation set $D^{st} = [\underline{\omega}_h, \beta^{st}]$.*

Let us now focus on the case when Proposition 6 does not apply, i.e., $1 - F_h(\bar{\omega}_\ell) - bf_h(\bar{\omega}_\ell) > 0$. Instead of providing a full characterization (as in Proposition 4), we confine ourselves to provide a sufficient condition for when marginal sequential delegation is profitable, where the delegation menu corresponding to marginal sequential delegation is defined as in (14). Of course, this then implies that sequential delegation is optimal. We begin by characterizing when it is profitable to marginally modify the static delegation menu consisting of the interval $[\underline{\omega}_h, \beta]$ with an arbitrary upper endpoint $\beta$ (instead of $\beta^{st}$).

**Lemma 11.** *Consider $\beta \in (\underline{\omega}_\ell, \bar{\omega}_\ell)$ and let us denote*

$$\Delta V_h \equiv -\frac{f_h(\beta) + bf'_h(\beta)}{f_\ell(\beta)} \cdot 2\int_\beta^{\bar{\omega}_\ell} [1 - F_\ell(\omega)]\, d\omega + 2\int_\beta^{\bar{\omega}_h} [1 - F_h(\omega) - bf_h(\omega)]\, d\omega. \quad (25)$$

[35] The log-concavity of $f$ implies also log-concavity of $f_h$. Thus, $f'_h/f_h$ and also $(f_h + bf'_h)/f_h = 1 + bf'_h/f_h$ are non-increasing. Then $(f_h + bf'_h)/f_\ell = (f_h + bf'_h)/f_h \cdot f_h/f_\ell$ is non-increasing as well. Finally, note that $zf'(z)/f(z)$ being non-decreasing is sufficient for $f_h/f_\ell$ being non-increasing for every $\ell < h$.

[36] One may wonder whether the Assumptions A1–A5 are not inconsistent with one another. It is, however, easy to check that, for example, $F(\theta) = \theta^2$ satisfies all assumptions for appropriately chosen values of $\ell, h$ and $b$ (see Appendix).

*Then, if $\Delta V_h > 0$, and only if $\Delta V_h \geq 0$, marginal sequential delegation improves the principal's expected utility compared to the static delegation set $[\underline{\omega}_h, \beta]$.*

Next, we provide a sufficient condition so that the condition for the profitability of marginal sequential delegation is satisfied at the optimal static delegation set $[0, \beta^{st}]$.

**Proposition 7.** *Consider a distribution satisfying A2 and A4. Let $0 < \ell < h \leq 1$ be such that Assumptions A1 and A3 are satisfied for some $b_0 > 0$. Then there exists an open set of parameters $(b, \mu_h)$ such that A1 and A3 are satisfied, and marginal sequential delegation improves upon static delegation.*

Proposition 7 identifies a range of parameters so that sequential delegation is optimal. If, in addition, A5 holds, then by Lemma 10, the good expert is offered a set of extreme decisions, and the bad expert is offered an interval of moderate decisions. Hence, the qualitative features of our results for the uniform case are robust and carry over to the case with more general distributions that satisfy the conditions A1–A5. Deriving the exact parametric conditions for the optimality of sequential delegation that parallel the simple conditions in the uniform case is beyond the scope of the current paper.

# 9 Conclusion

In this paper, we have provided insights into the nature of optimal delegation when the agent's private information arrives over time. Our key insight is that the principal can design a menu of delegation sets to elicit the agent's privately known expertise. In doing so, the principal offers the good expert a set of extreme decisions and the bad expert a set of moderate decisions to choose from. This provides a novel rationale for limiting the discretion of privately informed experts, next to restraining experts to pursue partisan interests or to motivate them to acquire information.

The optimal mechanism design approach adopted in our paper requires that actions are contractible and that the principal can commit to a mechanism. This approach is useful, as it characterizes the upper bound on what the principal can attain when she is constrained only by the agent's private information. An interesting avenue for future work is to relax the principal's commitment power. 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. A question 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

## A.1 Proofs for Section 5

*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, (8) 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 (8) implies that $\frac{1}{2}(y+\bar{x}_p) < \bar{\omega}_p$. Thus, by (6), the principal's utility from $\tilde{D}_p = D_p \cup \{y\}$ and $D_p$ are respectively given by

$$p \cdot \tilde{V}_p = C + \int_{\bar{x}_p}^{(y+\bar{x}_p)/2} (\bar{x}_p - \omega)^2 \, d\omega - \int_{(y+\bar{x}_p)/2}^{\bar{\omega}_p} (y-\omega)^2 \, d\omega + b(y - \bar{\omega}_p)^2 - pb^2, \tag{26}$$

$$p \cdot V_p = C + \int_{\bar{x}_p}^{\bar{\omega}_p} (\bar{x}_p - \omega)^2 \, d\omega + b(\bar{x}_p - \bar{\omega}_p)^2 - pb^2, \tag{27}$$

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,

$$p(\tilde{V}_p - V_p) = -\int_{(y+\bar{x}_p)/2}^{\bar{\omega}_p} [(y-\omega)^2 - (\bar{x}_p - \omega)^2] \, d\omega + b[(y-\bar{\omega}_p)^2 - (\bar{x}_p - \bar{\omega}_p)^2]. \tag{28}$$

The integral is equal to

$$-\int_{(y+\bar{x}_p)/2}^{\bar{\omega}_p} (y - \bar{x}_p)(y + \bar{x}_p - 2\omega) \, d\omega = (y - \bar{x}_p)\left(\frac{y+\bar{x}_p}{2} - \bar{\omega}_p\right)^2. \tag{29}$$

The second term in $p(\tilde{V}_p - V_p)$ can be written as

$$b[(y-\bar{\omega}_p)^2 - (\bar{x}_p - \bar{\omega}_p)^2] = b(y-\bar{x}_p)(y+\bar{x}_p - 2\bar{\omega}_p) = 2b(y-\bar{x}_p)\left(\frac{y+\bar{x}_p}{2} - \bar{\omega}_p\right). \tag{30}$$

Hence,

$$p(\tilde{V}_p - V_p) = (y-\bar{x}_p)\left(\bar{\omega}_p - \frac{y+\bar{x}_p}{2}\right)\left(\bar{\omega}_p - \frac{y+\bar{x}_p}{2} - 2b\right). \tag{31}$$

Since the first two factors are positive by assumption, we obtain that $p(\tilde{V}_p - V_p) > 0$ 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 $\tilde{D}_p = \{\bar{x}_p, y\}$. In this case, the lower bound in the integrals in (26) and (27) is $\bar{\omega}_p$ instead of $\bar{x}_p$. The formula (31) 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 $\tilde{D}_p = \{y\}$, as in $\tilde{D}_p$ the action $\bar{x}_p$ becomes redundant. Let us for any $z \in \mathbb{R}$ define

$$\sigma(z) = -b(z - \underline{\omega}_p)^2 - \int_{\underline{\omega}_p}^{\bar{\omega}_p} (z - \omega)^2 \, d\omega + b(z - \bar{\omega}_p)^2 - pb^2 \tag{32}$$

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

$$p(\tilde{V}_p - V_p) = (\bar{\omega}_p - \underline{\omega}_p)(y - \bar{x}_p)[\underline{\omega}_p + \bar{\omega}_p - (y + \bar{x}_p) - 2b]. \tag{33}$$

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 $\tilde{V}_p > V_p$, which completes the proof. □

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

## A.2 Proofs for Section 6

The proofs of Propositions 1 and 2 make use of the following auxiliary lemma.

**Lemma 12.** *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*

$$D' = \emptyset, \quad \text{or} \quad D' = \{\bar{\omega}_\ell + d\}, \quad \text{or} \quad D' = [\bar{\omega}_\ell + d, \bar{x}] \tag{34}$$

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

*Proof of Lemma 12.* 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 does not hold 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 the states $\bar{\omega}_\ell$ and $\bar{\omega}_h$. However, it would strictly increase the agent's utility in states $\omega \in (y_1, y_2)$. Therefore, by (9), adding actions from the interval $(y_1, y_2)$ to $D$ strictly improves the principal's expected payoff, a contradiction to the optimality of $D$. The argument for the case that $D \cap [\bar{\omega}_\ell, \bar{\omega}_h]$ 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 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 the neighborhood of $x_1'$. Therefore, by (9), 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 (8) 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., $\frac{1}{2}(x' + \bar{x}) \geq \bar{\omega}_\ell$ (i.e., $\bar{x}$ is redundant for the $\ell$-type). 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 state $\bar{\omega}_\ell$ and adding $y$ does not influence the principal's utility conditional on type $\ell$. In addition, by (9), 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 state $\bar{\omega}_\ell$ by assumption. Moreover, the two assumed inequalities imply (recall that $\beta_p^0 = \bar{\omega}_p - 2b$)

$$\frac{x' + \bar{x}}{2} = x' + \frac{\bar{x} - x'}{2} \geq \beta_\ell^0 + \bar{\omega}_h - \bar{\omega}_\ell = \beta_h^0. \tag{35}$$

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 $\beta^0_\ell \leq x \leq \beta^0_h$. To the contrary, suppose it is not an interval. By Lemma 12, we then have

$$D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup D', \tag{36}$$

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 $\beta^0_h \leq \bar{\omega}_\ell < \bar{x}$ implies that lowering $\bar{x}$ 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 $\frac{1}{2}(\bar{\omega}_\ell - d + \bar{\omega}_\ell + d) = \bar{\omega}_\ell \geq \beta^0_h \geq \beta^0_\ell$).

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 $[\underline{\omega}_h, \bar{x}]$ when he does not know the ex ante type is

$$V(\bar{x}) = \mu_h V_h(\bar{x}) + \mu_\ell V_\ell(\bar{x}). \tag{37}$$

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

$$V'_p(\bar{x}) \lesseqqgtr 0 \quad \Leftrightarrow \quad \bar{x} \gtreqqless \beta^0_p. \tag{38}$$

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

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

$$\frac{d}{d\mu_h}\beta^{st} = V'_h(\beta^{st}) - V'_\ell(\beta^{st}), \tag{39}$$

which is strictly positive by (38) since $\beta^{st} \in (\beta^0_\ell, \beta^0_h)$. Furthermore, $V$ converges to $V_h$ when $\mu_h$ converges to 0. Thus, also $\beta^{st}$ converges to $\beta^0_h$. Similarly, $\beta^{st}$ converges to $\beta^0_\ell$ as $\mu_\ell$ converges to 1. This completes the proof. □

*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 $D' = [\bar{\omega}_\ell + d, \beta^0_h]$ and $\beta^0_h > \bar{\omega}_\ell + d$, or $D' = \{\bar{\omega}_\ell + d\}$ and $\beta^0_h \leq \bar{\omega}_\ell + d$.

Indeed, suppose, contrary to the claim, that $D^{st}$ is not of the postulated form. Then by Lemma 12, 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 $\frac{1}{2}(\bar{\omega}_\ell - d + \bar{\omega}_\ell + d) = \bar{\omega}_\ell < \beta^0_h$ 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 12, the optimal set is the union of two intervals, or an interval and a point.

Consider the first case $D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup [\bar{\omega}_\ell + d, x]$. If $x \neq \beta^0_h$, then the principal's expected payoff could be improved either by adding an action $y = x + \varepsilon$ close to $x$ (when $x < \beta^0_h$) or by slightly decreasing $x$ (when $x > \beta^0_h$). 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 $D' = [\bar{\omega}_\ell + d, \beta^0_h]$.

Consider next the case $D^{st} = [\underline{\omega}_h, \bar{\omega}_\ell - d] \cup \{\bar{\omega}_\ell + d\}$. If $\bar{\omega}_\ell + d < \beta^0_h$, then, adding an action slightly higher than $\bar{\omega}_\ell + d$ would improve the principal's expected payoff 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 $\bar{\omega}_\ell + d \geq \beta^0_h$.

Finally, we show that in both cases $d < 2b$. Otherwise we can add an action $y = \bar{\omega}_\ell - d + \varepsilon$ to the delegation set. This improves the principal's expected utility conditional on both types: due to Lemma 2, (i) for the $h$-type, and due to Lemma 2, (ii) for the $\ell$-type (as actions higher than $\bar{\omega}_\ell$ are now redundant for the $\ell$-type). This completes the proof. □

## A.3 Proofs for Section 4

The proof of Lemma 5 makes use of the following auxiliary lemma.

**Lemma 13.** *Consider an arbitrary menu $(D_h, D_\ell)$ of delegation sets and $p \in \{h, \ell\}$. Assume that there is no $z \geq \beta^0_p$ such that $D_p = [\underline{\omega}_p, z]$. Then for every $\varepsilon > 0$ there* *exists $y \notin D_p$ such that replacing $D_p$ by the delegation set $\tilde{D}_p = D_p \cup \{y\}$ improves the principal's expected utility and improves the agent's expected utility by no more than $\varepsilon$:*

$$0 < \tilde{V}_p - V_p, \qquad \textit{and} \qquad 0 \leq \tilde{U}_p - U_p < \varepsilon. \tag{40}$$

*Proof of Lemma 13.* 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 there is no $z \geq \beta_p^0$ such that $D_p = [\underline{\omega}_p, z]$, then we have four possible cases.

*Case 1.* Let $\underline{x}_p < \underline{\omega}_p$. Because $D_p$ is connected, 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 the modified set $\tilde{D}_p = \{y\}$ being a singleton (after omitting the now redundant point $\bar{x}_p$). In order to compare the principal's expected utility we use formula (33) 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 $\frac{1}{2}(y + \bar{x}_p) < \beta_p^0$. □

*Proof of Lemma 5.* 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 13. 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 13. Such a modification improves the $\ell$–type agent's expected utility from picking $D_h^*$ 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., $D_h^* = [\underline{\omega}_h, \beta_h^0]$ by Step 1. Assume to the contrary that $\bar{x}_h^* > \beta_h^0$. Non-redundancy requires that $\bar{x}_h^* \leq \bar{\omega}_h$. 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 the $\ell$-type's expected utility is lower when choosing the modified $D_h^*$. 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. □

*Proof of Lemma 6.* We proceed in 5 steps.

*Step 1.* We start by deriving necessary conditions, (a)–(g), for $D_h^*$ in any optimal delegation menu. Let us denote $\gamma_2 = x_h(\bar{\omega}_\ell)$ 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) $\bar{x}_\ell^* \leq \gamma_2$. Otherwise, we have $\gamma_2 < \bar{x}_\ell^* \leq \bar{\omega}_\ell$. Now note that by the definition of $\gamma_2$, any point in $D_h^*$ higher than $\gamma_2$ does not influence the $\ell$-type's utility from choosing the delegation set $D_h^*$. Thus, 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 $\min\{\omega, \bar{x}_\ell^*\}$, that is closer to the agent's ideal action. Moreover, in states $\omega \in [\underline{\omega}_\ell, \gamma_2]$, the agent is weakly better off under $D_\ell^*$, where he chooses his ideal action $\omega$. Thus, the $\ell$-type agent strictly prefers the delegation set $D_\ell^*$ to $D_h^*$, which contradicts $(IC_\ell)$ being binding (Lemma 5).

(b) 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 $D_h^*$.

(c) $\underline{x}_h^* \leq \underline{\omega}_h$. Assume the opposite. Then, consider adding the (non-redundant) action $y = 2\underline{\omega}_h - \underline{x}_h^* + \varepsilon$ to $D_h^*$.[37] For $\varepsilon > 0$ small enough, action $y$ is chosen in state $\underline{\omega}_h$ but not in state $\underline{\omega}_\ell$. 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 $D_h^*$.

(d) If $\gamma_2 < \bar{x}_h^*$, then $\bar{\omega}_\ell \leq \gamma_2$. Assume the opposite and let us discuss two cases. First, if $\frac{1}{2}(\bar{x}_h^* - \gamma_2) \leq \bar{\omega}_h - \bar{\omega}_\ell$, consider the action $y = 2\bar{\omega}_\ell - \gamma_2$ (the mirror image of $\gamma_2$ with respect to $\bar{\omega}_\ell$). Then by definition of $\gamma_2$ we have $y \notin D_h^*$ and $y < \bar{x}_h^*$ (otherwise, $\bar{x}_h^*$ would be chosen in state $\bar{\omega}_\ell$). Moreover, it follows from the assumed inequality that $\frac{1}{2}(y + \bar{x}_h^*) \leq \bar{\omega}_h$. Thus, adding $y$ to $D_h^*$ preserves $(IC_\ell)$ and also improves the principal's expected utility by Lemma 2, (i). Second, assume that $\frac{1}{2}(\bar{x}_h^* - \gamma_2) > \bar{\omega}_h - \bar{\omega}_\ell$. Observe that the fact that $\gamma_2$ is chosen by the agent in the state $\bar{\omega}_\ell$, and $\bar{x}_h^*$ is chosen in $\bar{\omega}_h$ implies that $D_h^*$ contains no actions in the interval $(\gamma_2, \bar{x}_h^*)$. Now we show that removing $\bar{x}_h^*$ improves the principal's expected utility, while not breaking the binding $(IC_\ell)$. Indeed, removing $\bar{x}_h^*$ has no effect on the $\ell$-type's utility from $D_h^*$. Moreover, it follows from (a) and Lemma 5 that $\gamma_2 \geq \bar{x}_\ell^* \geq \beta_\ell^0$. This, together with the inequality $\frac{1}{2}(\bar{x}_h^* - \gamma_2) > \bar{\omega}_h - \bar{\omega}_\ell$ yields

$$\frac{\gamma_2 + \bar{x}_h^*}{2} = \gamma_2 + \frac{\bar{x}_h^* - \gamma_2}{2} \geq \beta_\ell^0 + \bar{\omega}_h - \bar{\omega}_\ell = \beta_h^0. \tag{41}$$

Therefore, removing the action $\bar{x}_h^*$ improves the principal's expected utility conditional on type $h$ due to Lemma 2, (ii). This contradicts the optimality of $D_h^*$.

(e) If $\gamma_2 < \bar{x}_h^*$, then $\bar{x}_h^* = \beta_h^0$. Assume the opposite. It follows from (d) that $\bar{\omega}_\ell \leq \gamma_2$ and from (b) that $\bar{x}_h^* \leq \bar{\omega}_h$ and $D_h^* \cap [\bar{\omega}_\ell, \infty) = [\gamma_2, \bar{x}_h^*]$. We show that 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 $\bar{x}_h^* > \beta_h^0$, we reduce $\bar{x}_h^*$ marginally, i.e., replace $D_h^*$ by the delegation set $\big(D_h^* \cap (-\infty, \bar{x}_h^* - \varepsilon]\big) \cup \{\bar{x}_h^* - \varepsilon\}$, where $\varepsilon > 0$ is small enough. 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, and due to Lemma 3 (because $\bar{x}_h^* - \varepsilon > \beta_h^0$) for the second change. This contradicts the optimality of $D_h^*$.

(f) If $\beta_h^0 \leq \bar{\omega}_\ell$, then $\gamma_2 = \bar{x}_h^*$. This follows directly from (d) and (e).

[37]Note that $2\underline{\omega}_h - \underline{x}_h^*$ is the mirror action to $\underline{x}_h^*$, mirrored at $\underline{\omega}_h$.

(g) If $\bar{\omega}_\ell < \beta_h^0$, then either $\beta_h^0 \le \gamma_2 = \bar{x}_h^*$ or $\bar{\omega}_\ell \le \gamma_2 < \bar{x}_h^* = \beta_h^0$. Indeed, on the one hand, if $\gamma_2 = \bar{x}$, then $\beta_h^0 \le \gamma_2$. Otherwise, we can add the action $y = 2\beta_h^0 - \bar{x}_h^* - \varepsilon$ to $D_h^*$, where $\varepsilon > 0$ is small enough. This improves the principal's expected utility by Lemma 2, (ii), because $y > \bar{x}_h^*$ and $\frac{1}{2}(y + \bar{x}_h^*) = \beta_h^0 - \frac{1}{2}\varepsilon < \beta_h^0$. On the other hand, if $\gamma_2 < \bar{x}_h^*$, the statement follows from (d) and (e).

*Step 2.* We construct a new delegation set $\tilde{D}_h = D_{h1} \cup D_{h2}$; in Step 3 we show that the principal weakly benefits by replacing $D_h^*$ by $\tilde{D}_h$. The sets $D_{h1}$ and $D_{h2}$ are specified in the following way. First, if $\gamma_2 < \bar{x}_h^*$, we infer from (b), (d), and (e) that $\bar{\omega}_\ell \le \gamma_2 < \bar{x}_h^* = \beta_h^0$ and $D_h^* \cap [\bar{\omega}_\ell, \infty) = [\gamma_2, \beta_h^0]$. In that case, let us set $D_{h2} = [\gamma_2, \beta_h^0]$. On the other hand, if $\gamma_2 = \bar{x}_h^*$, we set $D_{h2} = \{\gamma_2\}$.

Second, we construct the set $D_{h1}$ as follows: Type $\ell$ would (weakly) prefer the interval $[\underline{x}_h^*, \gamma_2]$ over $D_h^*$ and would (weakly) prefer $D_h^*$ over the two point set $\{\underline{x}_h^*, \gamma_2\}$. Hence, since $(IC_\ell)$ is binding, type $\ell$ also weakly prefers $[\underline{x}_h^*, \gamma_2]$ over $D_\ell^*$, and $D_\ell^*$ over $\{\underline{x}_h^*, \gamma_2\}$. 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 $[\underline{x}_h^*, \gamma_1] \cup \{\gamma_2\}$ in between the large and the small set. (If $\gamma_1 = \underline{x}_h^*$, we identify $[\underline{x}_h^*, \gamma_1]$ with $\{\underline{x}_h^*\}$.)

The construction ensures that type $\ell$ is also 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 $\gamma_1 \le \underline{\omega}_h$ 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 $D_{h1} = [\underline{\omega}_h, \gamma_1]$.

*Step 3.* We now show that the principal's utility, conditional on facing the $h$-type, is weakly higher under $\tilde{D}_h = D_{h1} \cup D_{h2}$ than under $D_h^*$, implying that $\tilde{D}_h$ is optimal, as well. For this we use the formula (6) for $p = h$, where we split the integral into three parts.

$$h \cdot V_h = \underbrace{b u_h(\underline{\omega}_h)}_{T_1} + \underbrace{\int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega)\, d\omega}_{T_2} + \underbrace{\int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} u_h(\omega)\, d\omega}_{T_3} + \underbrace{\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega)\, d\omega}_{T_4} - \underbrace{b u_h(\bar{\omega}_h)}_{T_5} - hb^2. \quad (42)$$

In what follows we indicate all variables pertaining to $\tilde{D}_h$ (resp $D_h^*$) with a tilde (resp. an asterisk) and compare the principal's utilities from $\tilde{D}_h$ and $D_h^*$ using (42).

First, observe that $\tilde{T}_3 = T_3^*$ because $(IC_\ell)$ is binding for both $D_h^*$ and $\tilde{D}_h$.

Second, we argue that $\tilde{T}_4 = T_4^*$ and $\tilde{T}_5 = T_5^*$. 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 $\tilde{T}_1 \geq T_1^*$ and $\tilde{T}_2 \geq T_2^*$. It follows from (b) and (c) in Step 1 that $D_h^* \cap (-\infty, \underline{\omega}_\ell]$ is either a singleton $\{\underline{x}_h^*\}$, or it is an interval of the form $[\underline{\omega}_h, z]$. In both cases, this intersection is a subset of $[\underline{x}_h^*, \gamma_1]$. Observe that by construction, the agent's utility from $\tilde{D}_h$ in the range $[\underline{\omega}_h, \underline{\omega}_\ell]$ is the same as his utility from the set of actions $[\underline{x}_h^*, \gamma_1]$. This implies that $\tilde{T}_1 \geq T_1^*$ and $\tilde{T}_2 \geq T_2^*$. Summing up, we have shown that the delegation set $\tilde{D}_h$ yields at least as large utility to the principal's as the set $D_h^*$.

*Step 4.* Let us now analyze which forms $\tilde{D}_h$ can have. If $b \geq \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$, or equivalently, $\beta_h^0 \leq \bar{\omega}_\ell$, then it follows from (f) and from construction that $D_{h2} = \{\gamma_2\}$. Thus, the set $\tilde{D}_h$ can only be of the form (i) or (ii), as specified in the lemma.

On the other hand, consider the case $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$, or equivalently, $\bar{\omega}_\ell < \beta_h^0$. First we show that $\underline{\omega}_h < \gamma_1$. This also implies that $D_h^*$ is not of the form (i). Assume, to the contrary, that $\gamma_1 \leq \underline{\omega}_h$. We show that in such a case, $(IC_\ell)$ cannot be binding. It follows from (g) that $\bar{\omega}_\ell \leq \gamma_2$. Observe that for $\gamma_1 \leq \underline{\omega}_h$ and $\bar{\omega}_\ell \leq \gamma_2$, the $\ell$-type's expected utility (scaled by $\ell$) from delegation set $\tilde{D}_h$ is increasing in $\gamma_1$ and decreasing in $\gamma_2$. Thus, it can be bounded from above by the expected utility from the delegation set $\{\underline{\omega}_h, \bar{\omega}_\ell\}$, which is

$$\int_{\underline{\omega}_\ell}^{(\underline{\omega}_h+\bar{\omega}_\ell)/2} -(\underline{\omega}_h - \omega)^2 \, d\omega + \int_{(\underline{\omega}_h+\bar{\omega}_\ell)/2}^{\bar{\omega}_\ell} -(\bar{\omega}_\ell - \omega)^2 \, d\omega < 0 + \int_0^{2b} -z^2 \, dz. \tag{43}$$

To see the inequality, note that $\frac{1}{2}(\underline{\omega}_h + \bar{\omega}_\ell) = \bar{\omega}_\ell - \frac{1}{2}(\bar{\omega}_\ell - \underline{\omega}_\ell) - \frac{1}{2}(\underline{\omega}_\ell - \underline{\omega}_h) < \bar{\omega}_\ell - \frac{1}{2}\ell - \frac{1}{2}(2b) < \bar{\omega}_\ell - 2b$, because $b < \frac{1}{2}(\bar{\omega}_h - \bar{\omega}_\ell)$ by assumption, and by applying a change of variable to the second integral. Moreover, as $\bar{x}_\ell^* \geq \beta_\ell^0 = \bar{\omega}_\ell - 2b$, we obtain for the $\ell$-type's expected utility from choosing $D_\ell^*$ that $\ell \cdot U_\ell^* \geq \int_0^{2b} -z^2 \, dz$. This together with (43) contradicts $(IC_\ell)$ being binding and establishes the inequality $\underline{\omega}_h < \gamma_1$.

Second, assume that $\tilde{D}_h$ is not of the form (ii). Then $\gamma_2 < \bar{x}_h^*$ and it follows from (g) that $\bar{\omega}_\ell \leq \gamma_2 < \bar{x}_h^* = \beta_h^0$. Moreover, clearly $\gamma_1 \leq \bar{\omega}_\ell$ and we have already shown above that $\underline{\omega}_h < \gamma_1$. It also follows from (b) that $\tilde{D}_h \cap [\bar{\omega}_\ell, \infty) = [\gamma_2, \beta_h^0]$, as specified by the form (iii).

*Step 5.* Finally, we complete the proof by deriving the inequalities for $\gamma_1$, $\gamma_2$, and $\bar{x}_\ell^*$. First, we have already shown (property (a)) that $\bar{x}_\ell^* \leq \gamma_2$. Second, we show that $\gamma_1 \leq \bar{x}_\ell^*$. Assume the opposite. Then in every state $\omega \in (\bar{x}_\ell^*, \bar{\omega}_\ell]$, the agent is strictly better off under the delegation set $\tilde{D}_h$ 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 $\omega \in [\underline{\omega}_\ell, \bar{x}_\ell^*]$, the agent chooses his ideal action $\omega$ under both $\tilde{D}_h$ and $D_\ell^*$. Thus, the $\ell$-type agent strictly prefers the delegation set $\tilde{D}_h$, which contradicts $(IC_\ell)$.

Third, we show that $\frac{1}{2}(\bar{x}_\ell^* + \gamma_2) \leq \bar{\omega}_\ell$. The inequality is trivial when $\gamma_2 \leq \bar{\omega}_\ell$. Let us thus assume, to the contrary, that $\gamma_2 > \bar{\omega}_\ell$ and $\frac{1}{2}(\bar{x}_\ell^* + \gamma_2) > \bar{\omega}_\ell$. By definition, in state

$\bar{\omega}_\ell$, the agent chooses action $\gamma_2$ from $\tilde{D}_h$, whereas he chooses action $\bar{x}_\ell^*$ from $D_\ell^*$. Since $\bar{x}_\ell^* \leq \bar{\omega}_\ell < \gamma_2$ and $\frac{1}{2}(\bar{x}_\ell^* + \gamma_2) > \bar{\omega}_\ell$, the action $\bar{x}_\ell^*$ 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 $\tilde{D}_h$ 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 $\tilde{D}_h$ since $\gamma_1 \leq \bar{x}_\ell^*$. Thus, the $\ell$-type agent strictly prefers the delegation set $D_\ell^*$ to $\tilde{D}_h$, which contradicts the assumption that $(IC_\ell)$ is binding. This completes the proof. □

*Proof of Lemma 7.* 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 $(\tilde{D}_h, D_\ell^*)$ such that $\tilde{D}_h = [\underline{\omega}_h, \bar{x}_\ell^*] = [\underline{\omega}_h, \underline{\omega}_\ell) \cup D_\ell^*$. Under this menu, the $h$-type is also offered the $\ell$-type's delegation set which is extended by actions that are never chosen by the $\ell$-type.

Under the menu $(\tilde{D}_h, D_\ell^*)$, the principal's expected utility conditional on the $\ell$-type remains unchanged. By (6), the utility conditional on the $h$-type (scaled by $h$) changes by

$$h(\tilde{V}_h - V_h^*) = b[\tilde{u}_h(\underline{\omega}_h) - u_h(\underline{\omega}_h)] - b[\tilde{u}_h(\bar{\omega}_h) - u_h(\bar{\omega}_h)] + \int_{\underline{\omega}_h}^{\bar{\omega}_h} [\tilde{u}_h(\omega) - u_h(\omega)]\, d\omega, \quad (44)$$

where $\tilde{u}_h$ denotes the $h$-type's utility under the delegation set $\tilde{D}_h$. We argue that the third term in (44) is positive and the other two are non-negative.

First, observe that $\tilde{u}_h(\underline{\omega}_h) = 0$ and $u_h(\underline{\omega}_h) \leq 0$. Thus, $\tilde{u}_h(\underline{\omega}_h) - u_h(\underline{\omega}_h) \geq 0$. Second, $-[\tilde{u}_h(\bar{\omega}_h) - u_h(\bar{\omega}_h)] = (\bar{x}_\ell^* - \bar{\omega}_h)^2 - (\bar{x}_h^* - \bar{\omega}_h)^2$, which is non-negative, if and only if $\frac{1}{2}(\bar{x}_\ell^* + \bar{x}_h^*) < \bar{\omega}_h$. 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 6 is a singleton. However, according to Lemma 6, we have $\frac{1}{2}(\bar{x}_\ell^* + \bar{x}_h^*) \leq \bar{\omega}_\ell < \bar{\omega}_h$.

Finally, because $(IC_h)$ is violated and $D_\ell^* \subseteq \tilde{D}_h$, we obtain

$$\int_{\underline{\omega}_h}^{\bar{\omega}_h} u_h(\omega)\, d\omega < \int_{\underline{\omega}_h}^{\bar{\omega}_h} u_\ell(\omega)\, d\omega \leq \int_{\underline{\omega}_h}^{\bar{\omega}_h} \tilde{u}_h(\omega)\, d\omega. \quad (45)$$

Thus, the third term in (44) is indeed positive. This shows that $\tilde{V}_h > V_h^*$, which completes the proof of the lemma. □

*Proof of Proposition 3.* Let $(D_h^*, D_\ell^*)$ be the optimal delegation menu in the form as specified in Lemmas 5 and 6. Let us assume to the contrary that this menu is a sequential menu and is (strictly) better for the principal than any static menu. We derive a contradiction by showing that the static delegation menu $(\tilde{D}_h, D_\ell^*)$, where $\tilde{D}_h = [\underline{\omega}_h, \bar{x}_\ell^*]$, weakly improves the principal's expected utility.

Now consider the menu $(\tilde{D}_h, D_\ell^*)$. The principal's expected utility conditional on facing the $\ell$-type remains unchanged. 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 $\ell$) from the delegation sets $D_\ell^*$ and $D_h^*$, respectively. These expected utilities are equal due to the binding $(IC_\ell)$. Moreover, by part (i) and (ii) and the final remark in Lemma 6, the type $h$ agent chooses action $\gamma_2 = \bar{x}_h^*$ in the states $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h]$. Thus,

$$\begin{aligned} h \cdot V_h^* &= \ell \cdot U_\ell^* + \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega)\, d\omega - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} (\bar{x}_h^* - \omega)^2\, d\omega + b u_h(\underline{\omega}_h) + b(\bar{x}_h^* - \bar{\omega}_h)^2 - hb^2, \\ h \cdot \tilde{V}_h &= \ell \cdot U_\ell^* - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} (\bar{x}_\ell^* - \omega)^2\, d\omega + b(\bar{x}_\ell^* - \bar{\omega}_h)^2 - hb^2, \\ h(\tilde{V}_h - V_h^*) &= b\kappa(\bar{\omega}_h) - b u_h(\underline{\omega}_h) - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} \kappa(\omega)\, d\omega - \int_{\underline{\omega}_\ell}^{\underline{\omega}_h} u_h(\omega)\, d\omega, \end{aligned} \tag{46}$$

where we define

$$\kappa(\omega) = (\bar{x}_\ell^* - \omega)^2 - (\bar{x}_h^* - \omega)^2 = (\bar{x}_h^* - \bar{x}_\ell^*)[2\omega - (\bar{x}_\ell^* + \bar{x}_h^*)]. \tag{47}$$

In the following, we show that $b \geq \bar{\omega}_h - \bar{\omega}_\ell$ implies $\tilde{V}_h \geq V_h^*$, which contradicts sequential delegation being optimal. First recall that by Lemma 6, we have that $\gamma_2 = \bar{x}_h^*$. Hence, by the final remark in Lemma 6, $\bar{x}_\ell^* + \bar{x}_h^* \leq 2\bar{\omega}_\ell < 2\bar{\omega}_h$ and thus $\kappa(\bar{\omega}_h) > 0$. We then obtain

$$\begin{aligned} h(\tilde{V}_h - V_h^*) &\geq b\kappa(\bar{\omega}_h) - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} \kappa(\omega)\, d\omega \geq b\kappa(\bar{\omega}_h) - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} \kappa(\bar{\omega}_h)\, d\omega \\ &= [b - (\bar{\omega}_h - \bar{\omega}_\ell)]\kappa(\bar{\omega}_h) \geq 0. \end{aligned} \tag{48}$$

The first inequality was obtained by neglecting the non-negative terms $-\int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega)\, d\omega$ and $-b u_h(\underline{\omega}_h)$. 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. □

*Proof of Proposition 4.* We begin with part (ii): 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:

$$\frac{d\tilde{V}_h(\varepsilon)}{d\varepsilon}\Big|_{\varepsilon=0} > 0 \quad \Leftrightarrow \quad b < \hat{b} \quad \textit{and} \quad \mu_h < \hat{\mu}_h(b). \tag{49}$$

Indeed, we have

$$h\,\frac{d\tilde{V}_h(\varepsilon)}{d\varepsilon}\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon}\left\{\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} -(\beta^{st}+\varepsilon-\omega)^2\,d\omega - b[-(\beta^{st}+\varepsilon-\bar{\omega}_h)^2]\right\}\Big|_{\varepsilon=0}. \tag{50}$$

$$\begin{aligned}
&= -2\int_{\bar{\omega}_\ell}^{\bar{\omega}_h}(\beta^{st}-\omega)\,d\omega + 2b(\beta^{st}-\bar{\omega}_h)\\
&= -(\beta^{st}-\bar{\omega}_h)^2 + (\beta^{st}-\bar{\omega}_\ell)^2 + 2b(\beta^{st}-\bar{\omega}_h)\\
&= [2\beta^{st}-2\bar{\omega}_h+(\bar{\omega}_h-\bar{\omega}_\ell)]\,[-(\bar{\omega}_h-\bar{\omega}_\ell)] - 2b(\bar{\omega}_h-\beta^{st})\\
&= 2(\bar{\omega}_h-\beta^{st})(\bar{\omega}_h-\bar{\omega}_\ell-b) - (\bar{\omega}_h-\bar{\omega}_\ell)^2.
\end{aligned} \tag{51}$$

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

$$\frac{d\tilde{V}_h(\varepsilon)}{d\varepsilon}\Big|_{\varepsilon=0} > 0 \quad \Leftrightarrow \quad \beta^{st} < \bar{\omega}_h - \frac{(\bar{\omega}_h-\bar{\omega}_\ell)^2}{2(\bar{\omega}_h-\bar{\omega}_\ell-b)} \equiv \xi_h. \tag{52}$$

Now, a computation (provided below) shows that

$$\beta_\ell^0 < \xi_h \quad \Leftrightarrow \quad b < \hat{b}. \tag{53}$$

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 (52) and (53), this establishes (49).

To complete the proof of part (ii), we show (53). 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 $\xi_h$:

$$\begin{aligned}
\beta_\ell^0 < \xi_h \quad &\Leftrightarrow \quad -2b < a - \frac{a^2}{2(a-b)} \quad \Leftrightarrow \quad 4b^2 - 2ab - a^2 < 0\\
&\Leftrightarrow \quad b > \tfrac{1}{4}(1-\sqrt{5})a \quad \text{and} \quad b < \tfrac{1}{4}(1+\sqrt{5})a = \hat{b}.
\end{aligned} \tag{54}$$

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 (i): By Lemma 5 and 6, the optimal delegation menu is in the set

$$\mathcal{R} \equiv \big\{(D_h, D_\ell) \mid D_h = [\min\{\underline{\omega}_h, z_h\}, z_h] \cup \{\bar{x}_h\},\ D_\ell = [\underline{\omega}_\ell, \bar{x}_\ell], \quad \beta_\ell^0 \le \bar{x}_\ell \le \bar{x}_h,\ \tfrac{1}{2}(\bar{x}_\ell + \bar{x}_h) \le \bar{\omega}_\ell, \quad (IC_\ell) \text{ binding}\big\} \tag{55}$$

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

$$V(\bar{x}_h, \bar{x}_\ell). \tag{56}$$

Observe that static interval delegation corresponds to the menu in $\mathcal{R}$ for which $\bar{x}_h = \bar{x}_\ell$. By (6) and (13), and using the fact that $(IC_\ell)$ is binding, we obtain

$$\begin{aligned} V(\bar{x}_h, \bar{x}_\ell) &= \frac{\mu_h}{h} h V_h + \frac{\mu_\ell}{\ell} \ell V_\ell \\ &= \frac{\mu_h}{h}\left[\ell \cdot U_\ell + \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega)\ d\omega + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega)\ d\omega + b u_h(\underline{\omega}_h) - b u_h(\bar{\omega}_h)\right] \\ &\quad + \frac{\mu_\ell}{\ell}\left[\ell \cdot U_\ell + b u_\ell(\underline{\omega}_\ell) - b u_\ell(\bar{\omega}_\ell)\right]. \end{aligned} \tag{57}$$

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

$$\begin{aligned} \Psi(\bar{x}_h, \bar{x}_\ell) \equiv \frac{\mu_h}{h}&\left[\ell \cdot U_\ell + \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} u_h(\omega)\ d\omega - b u_h(\bar{\omega}_h)\right] \\ &+ \frac{\mu_\ell}{\ell}\left[\ell \cdot U_\ell + b u_\ell(\underline{\omega}_\ell) - b u_\ell(\bar{\omega}_\ell)\right] \end{aligned} \tag{58}$$

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

$$V(\bar{x}_h, \bar{x}_\ell) \le \Psi(\bar{x}_h, \bar{x}_\ell) \quad \text{for all } \bar{x}_\ell \le \bar{x}_h, \quad \textit{and} \quad V(x, x) = \Psi(x, x) \quad \text{for } \underline{\omega}_\ell \le x. \tag{59}$$

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 6, for all $\omega \in [\bar{\omega}_\ell, \bar{\omega}_h]$, we have $x_h(\omega) = \bar{x}_h \le \bar{\omega}_h$. 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

$$\Psi(\bar{x}_h, \bar{x}_\ell) = \frac{\mu_h}{h}\left[-\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} (\bar{x}_h - \omega)^2 \, d\omega + b(\bar{x}_h - \bar{\omega}_h)^2\right] \\ - \left(\frac{\mu_h}{h} + \frac{\mu_\ell}{\ell}\right) \int_{\bar{x}_\ell}^{\bar{\omega}_\ell} (\bar{x}_\ell - \omega)^2 \, d\omega + \frac{\mu_\ell}{\ell} b(\bar{x}_\ell - \bar{\omega}_\ell)^2 - b^2. \tag{60}$$

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 (50). Thus, we can deduce as above:

$$\frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_h} > 0 \quad \Leftrightarrow \quad \bar{x}_h < \xi_h. \tag{61}$$

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

$$\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_\ell, \bar{x}_\ell) \qquad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \tag{62}$$

Indeed, from expression (53), $\xi_h \leq \beta_\ell^0$ if and only if $b \geq \hat{b}$. Therefore, $b \geq \hat{b}$ together with (61) and the fact that $\beta_\ell^0 \leq \bar{x}_\ell$ implies that

$$\frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_h} \leq 0 \qquad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \tag{63}$$

We can deduce that $\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_\ell, \bar{x}_\ell)$, establishing (62) 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

$$\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\beta^{st}, \beta^{st}) = V(\beta^{st}, \beta^{st}) \qquad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \tag{64}$$

We distinguish three cases. Suppose first that $\bar{x}_\ell \geq \xi_h$. Then equation (61) 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 (61) 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,

$$\bar{x}_\ell \leq \bar{x}_h \leq \beta^{st}, \tag{65}$$

and, by (61),

$$\frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_h} \leq 0. \tag{66}$$

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

$$0 = \frac{dV(\bar{x}, \bar{x})}{d\bar{x}} \Big|_{\bar{x}=\beta^{st}} = \frac{d\Psi(\bar{x}, \bar{x})}{d\bar{x}} \Big|_{\bar{x}=\beta^{st}} = \frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_h} + \frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_\ell}. \tag{67}$$

The two previous properties imply

$$\frac{\partial \Psi(\beta^{st}, \beta^{st})}{\partial \bar{x}_\ell} \geq 0. \tag{68}$$

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

$$\begin{aligned} \frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_\ell} &= -\left(\frac{\mu_h}{h} + \frac{\mu_\ell}{\ell}\right) \cdot 2 \cdot \int_{\bar{x}_\ell}^{\bar{\omega}_\ell} (\bar{x}_\ell - \omega)\, d\omega + \frac{\mu_\ell}{\ell} 2b(\bar{x}_\ell - \bar{\omega}_\ell) \\ &= \left(\frac{\mu_h}{h} + \frac{\mu_\ell}{\ell}\right)(\bar{x}_\ell - \bar{\omega}_\ell)^2 - \frac{\mu_\ell}{\ell} 2b(\bar{\omega}_\ell - \bar{x}_\ell) \\ &= (\bar{\omega}_\ell - \bar{x}_\ell)\left[\left(\frac{\mu_h}{h} + \frac{\mu_\ell}{\ell}\right)(\bar{\omega}_\ell - \bar{x}_\ell) - \frac{\mu_\ell}{\ell} 2b\right]. \end{aligned} \tag{69}$$

Note that this expression does not depend on $\bar{x}_h$. Now, since $\beta^{st} \leq \bar{\omega}_\ell$, (68) 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

$$\frac{\partial \Psi(\bar{x}_h, \bar{x}_\ell)}{\partial \bar{x}_\ell} \geq 0 \qquad \text{for all } \bar{x}_\ell \leq \bar{x}_h. \tag{70}$$

Hence, $\Psi(\bar{x}_h, \bar{x}_\ell) \leq \Psi(\bar{x}_h, \bar{x}_h) \leq \Psi(\beta^{st}, \beta^{st})$. This establishes (64) and completes the proof. □

*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 $\tilde{D}_\ell = [\underline{\omega}_\ell, \bar{\omega}_\ell - d]$, and consider the marginal modification defined in (19),

$$\tilde{D}_h(\varepsilon) = [\underline{\omega}_h, \bar{\omega}_\ell - d - \eta] \cup [\bar{\omega}_\ell + d - \varepsilon, \bar{\omega}_\ell + d], \qquad \eta, \varepsilon > 0, \tag{71}$$

so that $(IC_\ell)$ is binding. Observe that for $\varepsilon = 0$, the delegation menu $(\tilde{D}_h(0), \tilde{D}_\ell)$ essentially reduces to the optimal static menu. We now show 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 by Proposition 2, $d < 2b$, and so together with Assumption A1, we have $\bar{\omega}_\ell - d > \bar{\omega}_\ell - 2b \geq \bar{\omega}_\ell - \ell = \underline{\omega}_\ell$. Because $(IC_\ell)$ is binding and because the $\ell$-type is indifferent between $\tilde{D}_\ell$ and $\tilde{D}_h(0)$, 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 states $\omega \in [\bar{\omega}_\ell, \bar{\omega}_\ell + d - \varepsilon]$. Because of this and since, by Proposition 2, $\bar{\omega}_\ell + d < \bar{\omega}_\ell + 2b \leq \bar{\omega}_\ell + (\bar{\omega}_h - \bar{\omega}_\ell) = \bar{\omega}_h$, (17) writes

$$h \cdot \tilde{V}_h(\varepsilon) = \ell \cdot U_\ell^* - \int_{\bar{\omega}_\ell}^{\bar{\omega}_\ell + d - \varepsilon} (\bar{\omega}_\ell + d - \varepsilon - \omega)^2 \, d\omega - \int_{\bar{\omega}_\ell + d}^{\bar{\omega}_h} (\bar{\omega}_\ell + d - \omega)^2 \, d\omega \\ + b(\bar{\omega}_\ell + d - \omega_h)^2 - hb^2. \tag{72}$$

The derivative with respect to $\varepsilon$ is

$$h \, \frac{d\tilde{V}_h(\varepsilon)}{d\varepsilon} = 2 \int_{\bar{\omega}_\ell}^{\bar{\omega}_\ell + d - \varepsilon} (\bar{\omega}_\ell + d - \varepsilon - \omega) \, d\omega = (\bar{\omega}_\ell + d - \varepsilon - \bar{\omega}_\ell)^2, \tag{73}$$

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

## A.4 Proofs for Section 8

*Proof of Lemma 8.* The proof proceeds in 6 steps.

*Step 1.* First we show that equation (22) has indeed a unique solution in $(\underline{\omega}_p, \bar{\omega}_p)$. It will be more convenient to work directly with the original density $f$. Using the transformation $\omega' = \phi + (\omega - \phi)/p$ we obtain

$$\int_{\beta}^{\bar{\omega}_p} [1 - F_p(\omega) - bf_p(\omega)] \, d\omega = \int_{\phi + (\beta - \phi)/p}^{1} [p(1 - F(\omega')) - bf(\omega')] \, d\omega'. \tag{74}$$

Now define

$$\Gamma(z, p) = \int_z^1 \left[1 - F(\omega) - \frac{b}{p} f(\omega)\right] \, d\omega. \tag{75}$$

Then (74) becomes equal to $p\Gamma(\phi + (\beta - \phi)/p, p)$. Given $p \in (b/\phi, 1]$, we are now looking for $z \in (0, 1)$ that solves the equation $\Gamma(z, p) = 0$. Then $\beta = \phi + p(z - \phi)$ solves the original equation.

First note that $z = 1$ is a trivial solution to the equation, which we exclude as it does not belong to $(0,1)$. To see the existence of a solution $z \in (0,1)$, observe that $1 - F(\omega) - b/p \cdot f(\omega)$ is negative for $\omega$ close to 1, as $f$ is non–decreasing by Assumption A2. Thus, $\Gamma(z,p)$ is negative for $z$ close to 1. Now consider $z = 0$ and recall the general formula $\int_0^1 [1 - F(\omega)]\, d\omega = \phi$.[38] Then we obtain $\Gamma(0,p) = \phi - b/p$, which is positive due to A1. Continuity then implies the existence of a solution.

The see the uniqueness of the solution, observe that $\Gamma$ is strictly concave in $z$, as $F$ is increasing and $f$ is non-decreasing due to A2. Thus, it can attain the value zero in at most two points, while $z = 1$ already is one of them. Thus, there can be only one solution in the interval $(0,1)$. Let $\hat{z}(p)$ denote that solution. As indicated above, then

$$\beta_p^0 = \phi + p[\hat{z}(p) - \phi]. \tag{76}$$

Also note that as $1 - F(\omega) - b/p \cdot f(\omega)$ is decreasing, then $1 - F(z) - b/p \cdot f(z) > 0$ for $z = \hat{z}(p)$.

*Step 2.* Now we prove that $\beta_p^0$ is increasing in $p$ (assuming that $b < p\phi$). Using the *Implicit function theorem* we obtain (for simplicity we write $\hat{z}$ instead of $\hat{z}(p)$)

$$\frac{d\hat{z}(p)}{dp} = -\frac{\partial \Gamma}{\partial p} \Big/ \frac{\partial \Gamma}{\partial z} = \frac{b}{p^2} \cdot \frac{1 - F(\hat{z})}{1 - F(\hat{z}) - b/p \cdot f(\hat{z})}. \tag{77}$$

Then

$$\frac{d\beta_p^0}{dp} = (\hat{z} - \phi) + p \cdot \frac{d\hat{z}(p)}{dp} = (\hat{z} - \phi) + \frac{b}{p} \cdot \frac{1 - F(\hat{z})}{1 - F(\hat{z}) - b/p \cdot f(\hat{z})}. \tag{78}$$

Now, as $\hat{z} > 0$ we obtain

$$\begin{aligned} \phi - \hat{z} &= -\hat{z} + \int_0^1 [1 - F(\omega)]\, d\omega = -\hat{z} + \int_0^{\hat{z}} [1 - F(\omega)]\, d\omega + \int_{\hat{z}}^1 [1 - F(\omega)]\, d\omega \\ &< \int_{\hat{z}}^1 [1 - F(\omega)]\, d\omega = \frac{b}{p}[1 - F(\hat{z})], \end{aligned} \tag{79}$$

where the last equality holds, as $\hat{z}$ solves the equation $\Gamma(\hat{z},p) = 0$. Combining (78) and (79) we obtain

$$\frac{d\beta_p^0}{dp} > \frac{b}{p}[1 - F(\hat{z})] \left[ -1 + \frac{1}{1 - F(\hat{z}) - b/p \cdot f(\hat{z})} \right]. \tag{80}$$

This is positive, as $1 - F(\hat{z}) - b/p \cdot f(\hat{z}) > 0$, which we have established at the end of Step 1.

*Step 3.* We show that $\bar{\omega}_p - \beta_p^0$ is non-decreasing in $p$. For this recall that $\bar{\omega}_p =$

[38] This follows directly from the definition of the mean and integration by parts.

$\phi + p(1-\phi)$ and thus $\bar{\omega}_p - \beta_p^0 = p(1-\hat{z}(p))$. Then

$$\frac{d(\bar{\omega}_p - \beta_p^0)}{dp} = (1-\hat{z}) - p \cdot \frac{d\hat{z}(p)}{dp} = (1-\hat{z}) - \frac{b}{p} \cdot \frac{1-F(\hat{z})}{1-F(\hat{z}) - b/p \cdot f(\hat{z})}. \tag{81}$$

Now let us denote $\Phi(z) = \int_z^1 [1-F(\omega)]\, d\omega$. By definition of $\hat{z}$, we have $b/p = \Phi(\hat{z})/[1-F(\hat{z})]$. Substituting this into the above equality yields

$$\frac{d(\bar{\omega}_p - \beta_p^0)}{dp} = \Phi_1(\hat{z}), \quad \text{where} \quad \Phi_1(z) \equiv (1-z) - \frac{\Phi(z)}{1-F(z)} \cdot \frac{1}{1 - \frac{\Phi(z)}{1-F(z)} \cdot \frac{f(z)}{1-F(z)}}. \tag{82}$$

We will argue that $\Phi_1(z) \geq 0$ for all $z \in [0,1)$. Let us consider first $z \to 1^-$ (i.e., $z$ converging to 1 from below). Then $\Phi_1(z) \to 0$, since, using *L'Hospital's rule*, we have $\Phi(z)/[1-F(z)] \to 0$ and $\Phi(z)f(z)/[1-F(z)]^2 \to \frac{1}{2}$. Now we show that $\Phi_1'(z) \leq 0$. For this observe that $\Phi'(z) = -[1-F(z)]$ and

$$\frac{d}{dz}\frac{\Phi(z)}{1-F(z)} = -1 + \frac{\Phi(z)}{1-F(z)} \cdot \frac{f(z)}{1-F(z)}. \tag{83}$$

Hence,

$$\begin{aligned}
\Phi_1'(z) &= -1 - \frac{d}{dz}\Big[\frac{\Phi(z)}{1-F(z)}\Big] \cdot \frac{1}{1 - \frac{\Phi(z)}{1-F(z)} \cdot \frac{f(z)}{1-F(z)}} \\
&\quad + \frac{\Phi(z)}{1-F(z)} \cdot \frac{1}{\big[1 - \frac{\Phi(z)}{1-F(z)} \cdot \frac{f(z)}{1-F(z)}\big]^2} \cdot \frac{d}{dz}\Big[\frac{\Phi(z)}{1-F(z)} \cdot \frac{f(z)}{1-F(z)}\Big] \\
&= \frac{\Phi(z)}{1-F(z)} \cdot \frac{1}{\big[1 - \frac{\Phi(z)}{1-F(z)} \cdot \frac{f(z)}{1-F(z)}\big]^2} \cdot \frac{d}{dz}\Big[\frac{\Phi(z)}{1-F(z)} \cdot \frac{f(z)}{1-F(z)}\Big].
\end{aligned} \tag{84}$$

Now recall that due to Assumption A4, the right-hand side of (83) is non-decreasing and has thus a non-negative derivative. Thus, $\Phi_1'(z) \geq 0$, and this shows that indeed $\Phi_1(z) \geq 0$, and thus $\bar{\omega}_p - \beta_p^0$ is non-decreasing in $p$.

*Step 4.* It remains to show that with $\beta_p^0$ defined as in the lemma, the optimal delegation set is $[\underline{\omega}_p, \beta_p^0]$. This follows from the same argument that we provide in the discussion below Lemma 4 for the uniform distribution. For this argument we need to adapt the proofs of Lemmas 2 and 3. We do this in the remaining 3 steps.

*Step 5.* Let us reconsider the proof of Lemma 2. The proof is almost the same as the original proof for the uniform distribution. For the sake of brevity we don't repeat the whole proof, but rather only focus on parts that are different.

*Step 5.1.* Consider statement (ii), first case. Now we use decomposition (21) and obtain formulas similar to (26), (27), and (28) but with appropriate weights on the utilities. Consequently, formula (31) becomes

$$\tilde{V}_p - V_p = (y - \bar{x}_p) \cdot 2 \int_{(y+\bar{x}_p)/2}^{\bar{\omega}_p} [1 - F_p(\omega) - b f_p(\omega)]\, d\omega. \tag{31'}$$

This is again positive, as $\frac{1}{2}(y + \bar{x}_p) < \beta_p^0$ by assumption.

*Step 5.2.* Consider statement (ii), third case. Similarly as above, formula (33) becomes

$$\tilde{V}_p - V_p = (y - \bar{x}_p)[2\phi - (y + \bar{x}_p) - 2b]. \tag{33'}$$

This is positive, as $2\phi - (y + \bar{x}_p) \geq 2\phi - 2\underline{\omega}_p = 2p\phi > 2b$ by Assumption A1.

*Step 6.* Let us now reconsider Lemma 3. The lemma holds in the original form and, using now steps 1–5, its proof is identical to the original proof for the uniform distribution. □

*Proof of Lemma 9.* This lemma generalizes Proposition 1. For the proof we proceed in 2 steps. We first adapt the proof of Lemma 12 and then the proof of Proposition 1.

*Step 1.* Let us reconsider Lemma 12. Its proof is almost the same as the original proof for the uniform distribution. We again don't repeat the whole proof, but rather only focus on parts that are different.

In statement (c), Case 2.3, formula (35) now becomes

$$\frac{x' + \bar{x}}{2} = x' + \frac{\bar{x} - x'}{2} \geq \beta_\ell^0 + \bar{\omega}_h - \bar{\omega}_\ell \geq \beta_h^0. \tag{35'}$$

For the uniform distribution, the last inequality is an equality, which obviously holds as $\bar{\omega}_p - \beta_p^0 = 2b$. Here, we have an inequality instead, due to Lemma 8.

*Step 2.* Let us consider Proposition 1. Its proof is identical to the original proof. We would only like to point out that in the statement and on several occasions in the proof we use the inequality $\beta_\ell^0 < \beta_h^0$. This we have now established in Lemma 8. □

*Proof of Lemma 10.* In the proof we follow the structure of arguments in Section 7. We proceed in 5 steps.

*Step 1.* Let us first consider Lemma 5. This lemma holds in the original form. Its proof is also identical to the proof for the uniform distribution.

*Step 2.* Now we reconsider the auxiliary Lemma 13 that we use in the proof of Lemma 5. Lemma 13 still holds and its proof is almost the same as the original proof. It differs only in Case 2 where we refer directly to the formula (33). Instead of (33) we use the generalized formula (33') established in Step 5.1 in the proof of Lemma 8. Now, the first bracket is negative. The second bracket is negative as well, because $2\phi - (y + \bar{x}_p) < 2\phi - 2\bar{\omega}_p < 0 < 2b$. Summing up, we again obtain that $\tilde{V}_p > V_p$.

*Step 3.* We state an additional auxiliary lemma proven below. This lemma will replace the argument that due to $(IC_\ell)$ binding we could simply substitute $\ell \cdot U_\ell^*$ into the principal's expected utility when dealing with the uniform distribution.

**Lemma 14.** *Let $g_1, g_2, \lambda : [a, b] \to \mathbb{R}$ be integrable functions such that $g_1, g_2 > 0$ and $g_2(\omega)/g_1(\omega)$ is non-increasing in $\omega$. Moreover, assume there is $x \in [a, b]$ such that $\lambda(\omega) \geq 0$ for $\omega \in [a, x)$ and $\lambda(\omega) \leq 0$ for $\omega \in (x, b]$. Then the following statements hold:*

(i) *If $\int_a^b \lambda(\omega)g_1(\omega)\, d\omega \geq 0$, then $\int_a^b \lambda(\omega)g_2(\omega)\, d\omega \geq 0$.*

(ii) *If $\int_a^b \lambda(\omega)g_1(\omega)\, d\omega > 0$, then $\int_a^b \lambda(\omega)g_2(\omega)\, d\omega > 0$.*

*Step 4.* Now we establish the generalization of Lemma 6. Note that case (iii) where $D_{h2} = [\gamma_2, \beta_h^0]$ does not occur here, as Assumption A3 implies that $\beta_h^0 < \bar{\omega}_\ell$.

Consider a delegation menu $(D_h^*, D_\ell^*)$ where $D_\ell^*$ is an interval and $(IC_\ell)$ is binding. We use the same construction as in the discussion following Lemma 6: Replace $D_h^*$ by the set $\tilde{D}_h = [\underline{x}_h^*, \gamma_1] \cup \{\bar{x}_h^*\}$ so that $(IC_\ell)$ is maintained. The proof is identical to the original proof up to one point. In particular, we still obtain an analogue decomposition as in (42), with appropriate weights. However, the difference is that now the equality $\tilde{T}_3 = T_3^*$ does not need to hold. Instead, we now argue that Lemma 14 implies that $\tilde{T}_3 \geq T_3^*$ when Assumption A5 is satisfied, i.e., when $[f_h(\omega) + bf_h'(\omega)]/f_\ell(\omega)$ is non-increasing in $\omega$. To see this, let $\tilde{u}_h$ be the agent's utility under $\tilde{D}_h$ and let us set $g_1(\omega) = f_\ell(\omega)$, $g_2(\omega) = f_h(\omega) + bf_h'(\omega)$, and $\lambda(\omega) = \tilde{u}_h(\omega) - u_h(\omega)$.

By construction, there is a $x \in (\underline{\omega}_\ell, \bar{\omega}_\ell)$ such that $\tilde{u}_h(\omega) - u_h(\omega) \geq 0$ for $\omega \in [\underline{\omega}_\ell, x)$ and $\tilde{u}_h(\omega) - u_h(\omega) \leq 0$ for $\omega \in (x, \bar{\omega}_\ell]$.[39] Because $(IC_\ell)$ is maintained, we have

$$\int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} [\tilde{u}_h(\omega) - u_h(\omega)] f_\ell(\omega)\, d\omega = 0. \tag{85}$$

[39] To see this assume that $\tilde{D}_h \neq D_h^*$. As the $\ell$-type is indifferent between those two sets it cannot be that $D_h^* \cap (\gamma_1, \bar{x}_h^*)$ is empty; let $y$ be the smallest action in that intersection. We argue that $x = \frac{1}{2}(\gamma_1 + y)$ has the desirable property. First, if $\omega \leq \gamma_1$, then under $\tilde{D}_h$ the agent chooses his ideal action, and thus, $\tilde{u}_h(\omega) = 0 \geq u_h(\omega)$. If $\gamma_1 < \omega < x$, then the agent is better off when choosing action $\gamma_1 \in \tilde{D}_h$ than by choosing any action from $D_h^*$. Therefore, $\tilde{u}_h(\omega) \geq u_h(\omega)$. Finally, if $\omega > x$, then the agent prefers action $y$ to $\gamma_1$. Thus, he is better off under $D_h^*$ than under $\tilde{D}_h$, and so $\tilde{u}_h(\omega) \leq u_h(\omega)$.

By the decomposition (21), the difference in the principal's expected payoff after the modification corresponds now to

$$\tilde{T}_3 - T_3^* = \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} [\tilde{u}_h(\omega) - u_h(\omega)] \, [f_h(\omega) + b f_h'(\omega)] \, d\omega. \tag{86}$$

Lemma 14 and (85) imply that (86) is non-negative if (23) holds. The remainder of the proof is identical to the proof of Lemma 6.

*Step 5.* Finally, let us prove a generalization of Lemma 7 and show that the solution of the relaxed problem (of the form established above) also solves the original problem.

Similarly, as in the original proof, let us assume that $(IC_h)$ is violated. We again show that in such a case, the principal's expected utility can be improved by offering the static delegation menu $(\tilde{D}_h, D_\ell^*)$ where $\tilde{D}_h = [\underline{\omega}_h, \bar{x}_\ell^*]$. Under this menu, the principal's expected utility conditional on the $\ell$-type remains unchanged. Similarly, to (44), by decomposition (21) the utility conditional on the $h$-type changes by

$$\begin{aligned} \tilde{V}_h - V_h^* = {} & b[\tilde{u}_h(\underline{\omega}_h) - u_h(\underline{\omega}_h)] f_h(\underline{\omega}_h) - b[\tilde{u}_h(\bar{\omega}_h) - u_h(\bar{\omega}_h)] f_h(\bar{\omega}_h) \\ & + \int_{\underline{\omega}_h}^{\bar{\omega}_h} [\tilde{u}_h(\omega) - u_h(\omega)] \, [f_h(\omega) + b f_h'(\omega)] \, d\omega. \end{aligned} \tag{87}$$

The first and the second term are non-negative by the same argument as in the original proof. It remains to show that the integral in (87) is positive. By construction, there is an $x \in [\underline{\omega}_h, \bar{\omega}_h]$ such that $\tilde{u}_h(\omega) - u_h(\omega) \geq 0$ for $\omega \in [\underline{\omega}_h, x)$ and $\tilde{u}_h(\omega) - u_h(\omega) \leq 0$ for $\omega \in (x, \bar{\omega}_h]$.[40] Now we apply Lemma 14, where we set $g_1(\omega) = f_h(\omega)$, $g_2(\omega) = f_h(\omega) + b f_h'(\omega)$, and $\lambda(\omega) = \tilde{u}_h(\omega) - u_h(\omega)$. Then $g_1(\omega) > 0$ and $g_2(\omega) > 0$ for $\omega \in (\underline{\omega}_h, \bar{\omega}_h)$ and it follows from (23) that $g_2/g_1$ is non-increasing. Moreover, similarly to (45), we have

$$\int_{\underline{\omega}_h}^{\bar{\omega}_h} u_h(\omega) f_h(\omega) \, d\omega < \int_{\underline{\omega}_h}^{\bar{\omega}_h} u_\ell(\omega) f_h(\omega) \, d\omega \leq \int_{\underline{\omega}_h}^{\bar{\omega}_h} \tilde{u}_h(\omega) f_h(\omega) \, d\omega, \tag{88}$$

where the first inequality follows from the violated $(IC_h)$ and the second from the fact that $D_\ell^* \subseteq \tilde{D}_h$. It follows then from Lemma 14 that the last term in (87) is positive. This shows that $\tilde{V}_h > V_h^*$, which contradicts the optimality of $(D_h^*, D_\ell^*)$ and completes the proof of the lemma. □

[40]The precise argument here is similar to the argument given in footnote 39. For $D_h$ of the form as specified in Lemma 6 (see Step 4 of this proof), we now set $x = \frac{1}{2}(\bar{x}_\ell^* + \bar{x}_h^*)$.

*Proof of Lemma 14.* We have

$$\begin{aligned}\int_a^b \lambda(\omega)g_2(\omega)\,d\omega &= \int_a^b \lambda(\omega)g_1(\omega)\frac{g_2(\omega)}{g_1(\omega)}\,d\omega \\ &= \int_a^x \lambda(\omega)g_1(\omega)\frac{g_2(\omega)}{g_1(\omega)}\,d\omega + \int_x^b \lambda(\omega)g_1(\omega)\frac{g_2(\omega)}{g_1(\omega)}\,d\omega \\ &\geq \int_a^x \lambda(\omega)g_1(\omega)\frac{g_2(x)}{g_1(x)}\,d\omega + \int_x^b \lambda(\omega)g_1(\omega)\frac{g_2(x)}{g_1(x)}\,d\omega \\ &= \frac{g_2(x)}{g_1(x)}\int_a^b \lambda(\omega)g_1(\omega)\,d\omega. \end{aligned} \tag{89}$$

To appreciate the inequality in the third line, recall that $g_1 > 0$ and that $g_2/g_1$ is non-increasing. Thus,

$$\begin{aligned} &\lambda(\omega)g_1(\omega) \geq 0 \quad \text{and} \quad \frac{g_2(\omega)}{g_1(\omega)} \geq \frac{g_2(x)}{g_1(x)} \quad \text{for } \omega \in [a,x), \\ &\lambda(\omega)g_1(\omega) \leq 0 \quad \text{and} \quad \frac{g_2(\omega)}{g_1(\omega)} \leq \frac{g_2(x)}{g_1(x)} \quad \text{for } \omega \in (x,b], \end{aligned}$$

which implies the inequality in (89). Finally, as $g_2(x)/g_1(x) > 0$, both claims in the lemma follow directly from (89). □

*Proof of Proposition 6.* The proof is similar to the proof of Proposition 3. We again show that if, on the contrary, sequential delegation is optimal, then the principal can improve her expected payoff by offering the static delegation menu $(\tilde{D}_h, D_\ell^*)$, where $\tilde{D}_h = [\underline{\omega}_h, \bar{x}_\ell^*]$.

Considering $\kappa(\omega)$ defined as in (47), formula (46) needs now to be adjusted by using the proper weights:

$$\begin{aligned}\tilde{V}_h - V_h^* = b\kappa(\bar{\omega}_h)f_h(\bar{\omega}_h) - bu_h(\underline{\omega}_h)f_h(\underline{\omega}_h) - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} \kappa(\omega)[f_h(\omega) + bf_h'(\omega)]\,d\omega \\ - \int_{\underline{\omega}_h}^{\underline{\omega}_\ell} u_h(\omega)[f_h(\omega) + bf_h'(\omega)]\,d\omega \\ + \int_{\underline{\omega}_\ell}^{\bar{\omega}_\ell} [\tilde{u}_h(\omega) - u_h(\omega)]\,[f_h(\omega) + bf_h'(\omega)]\,d\omega. \end{aligned} \tag{90}$$

The second and the fourth term are clearly non-negative. Moreover, the fifth term is non-negative, which follows from the binding $(IC_\ell)$ and Lemma 14 (stated in Step 3 in the proof of Lemma 9), where we set $g_1(\omega) = f_\ell(\omega)$, $g_2(\omega) = f_h(\omega) + bf_h'(\omega)$, and $\lambda(\omega) = \tilde{u}_h(\omega) - u_h(\omega)$. We clearly have $g_1 > 0$, $g_2 > 0$ and by construction there is an

$x \in [\underline{\omega}_\ell, \bar{\omega}_\ell]$ such that $\lambda(\omega) \geq 0$ for $\omega \in [\underline{\omega}_\ell, x)$ and $\lambda(\omega) \leq 0$ for $\omega \in (x, \bar{\omega}_\ell]$. Thus,

$$\begin{aligned}\tilde{V}_h - V_h^* &\geq b\kappa(\bar{\omega}_h) f_h(\bar{\omega}_h) - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} \kappa(\omega)[f_h(\omega) + bf_h'(\omega)]\, d\omega \\ &\geq b\kappa(\bar{\omega}_h) f_h(\bar{\omega}_h) - \int_{\bar{\omega}_\ell}^{\bar{\omega}_h} \kappa(\bar{\omega}_h)[f_h(\omega) + bf_h'(\omega)]\, d\omega \\ &= \kappa(\bar{\omega}_h)[1 - F_h(\bar{\omega}_\ell) - bf_h(\bar{\omega}_\ell)] \geq 0. \end{aligned} \tag{91}$$

In the second inequality we used the fact that $\kappa(\omega)$ is increasing. The last inequality follows from the assumption $1 - F_h(\bar{\omega}_\ell) - bf_h(\bar{\omega}_\ell) > 0$ and from $\kappa(\bar{\omega}_h) > 0$, which holds by the same argument as in the proof of Proposition 3. This contradict the optimality of sequential delegation and completes the proof. □

*Proof of Lemma 11.* Consider the static delegation set $D = [\underline{\omega}_h, \beta]$ and let $\tilde{D}_h = [\underline{\omega}_h, z] \cup \{x\}$, where $z < \beta < x$, be the $h$-type's delegation set that keeps the $\ell$-type indifferent between $D$ and $\tilde{D}_h$.[41] We consider $x$ as independent variable and let $z(x)$ be such that the $\ell$-type remains indifferent. For $x$ sufficiently close to $\beta$ such $z(x)$ clearly exists, is unique, and is also close to $\beta$.

Then the $\ell$-type's expected utility from the delegation set $\tilde{D}_h$ is

$$\tilde{U}_\ell = -\int_z^{(z+x)/2} (z-\omega)^2 f_\ell(\omega)\, d\omega - \int_{(z+x)/2}^{\bar{\omega}_\ell} (x-\omega)^2 f_\ell(\omega)\, d\omega. \tag{92}$$

In order to keep the $\ell$-type indifferent, its derivative with respect to $x$ needs to be equal to 0. (Here we make use of the *Implicit function theorem.*) Thus,

$$0 = \frac{d\tilde{U}_\ell}{dx} = -z'(x) \cdot 2\int_z^{(z+x)/2} (z-\omega) f_\ell(\omega)\, d\omega - 2\int_{(z+x)/2}^{\bar{\omega}_\ell} (x-\omega) f_\ell(\omega)\, d\omega. \tag{93}$$

Now consider the principal's expected utility from $\tilde{D}_h$, using the decomposition (21):

$$\begin{aligned}\tilde{V}_h = &-\int_z^{(z+x)/2} (z-\omega)^2 [f_h(\omega) + bf_h'(\omega)]\, d\omega \\ &- \int_{(z+x)/2}^{\bar{\omega}_h} (x-\omega)^2 [f_h(\omega) + bf_h'(\omega)]\, d\omega + b(x - \bar{\omega}_h)^2 f_h(\bar{\omega}_h) - b^2. \end{aligned} \tag{94}$$

[41]Compared to the notation in Section 7.2, it is now more convenient to use $x = \beta + \varepsilon$ and $z = \beta - \eta$.

Taking the derivative with respect to $x$, we obtain

$$\frac{d\tilde{V}_h}{dx} = -z'(x) \cdot 2 \int_z^{(z+x)/2} (z-\omega)[f_h(\omega) + bf_h'(\omega)]\, d\omega$$
$$- 2 \int_{(z+x)/2}^{\bar{\omega}_h} (x-\omega)[f_h(\omega) + bf_h'(\omega)]\, d\omega + 2b(x-\bar{\omega}_h)f_h(\bar{\omega}_h). \tag{95}$$

Substituting for $z'(x)$ from (93), we obtain

$$\frac{d\tilde{V}_h}{dx} = \frac{2\int_z^{(z+x)/2}(z-\omega)[f_h(\omega) + bf_h'(\omega)]\, d\omega}{2\int_z^{(z+x)/2}(z-\omega)f_\ell(\omega)\, d\omega} \cdot 2\int_{(z+x)/2}^{\bar{\omega}_\ell} (x-\omega)f_\ell(\omega)\, d\omega$$
$$- 2 \int_{(z+x)/2}^{\bar{\omega}_h} (x-\omega)[f_h(\omega) + bf_h'(\omega)]\, d\omega + 2b(x-\bar{\omega}_h)f_h(\bar{\omega}_h). \tag{96}$$

Now, when $x \to \beta$, then $z(x) \to \beta$ as well. Thus, in the fraction, both the numerator and the denominator converge to zero. In order to evaluate the limit, we use *L'Hospital's rule* twice. Most of the terms then converge to zero and we obtain that the limit is equal to $[f_h(\beta) + bf_h'(\beta)]/f_\ell(\beta)$. Thus,

$$\left.\frac{d\tilde{V}_h}{dx}\right|_{x\to\beta} = \frac{f_h(\beta) + bf_h'(\beta)}{f_\ell(\beta)} \cdot 2\int_\beta^{\bar{\omega}_\ell} (\beta-\omega)f_\ell(\omega)\, d\omega$$
$$- 2\int_\beta^{\bar{\omega}_h} (\beta-\omega)[f_h(\omega) + bf_h'(\omega)]\, d\omega + 2b(\beta-\bar{\omega}_h)f_h(\bar{\omega}_h) \tag{97}$$
$$= -\frac{f_h(\beta) + bf_h'(\beta)}{f_\ell(\beta)} \cdot 2\int_\beta^{\bar{\omega}_\ell} [1-F_\ell(\omega)]\, d\omega$$
$$+ 2\int_\beta^{\bar{\omega}_h} [1-F_h(\omega) - bf_h(\omega)]\, d\omega, \tag{98}$$

where the second equality was obtained using integration by parts. Now the last expression is equal to $\Delta V_h$ from the Lemma, which implies the conditions for marginal sequential delegation to be profitable.[42] □

*Proof of Proposition 7.* The fact that A1 and A3 are satisfied for $b_0$ can be equivalently stated as

$$\frac{\int_{\bar{\omega}_\ell}^{\bar{\omega}_h}[1-F_h(\omega)]\, d\omega}{1-F_h(\bar{\omega}_\ell)} < b_0 < \ell\phi. \tag{99}$$

Let us now denote

$$\underline{b} = \frac{\int_{\bar{\omega}_\ell}^{\bar{\omega}_h}[1-F_h(\omega)]\, d\omega}{1-F_h(\bar{\omega}_\ell)}, \qquad \text{and} \qquad b_1 = \frac{1-F_h(\bar{\omega}_\ell)}{f_h(\bar{\omega}_\ell)}. \tag{100}$$

[42] For uniform distribution, $\Delta V_h$ simplifies to (51).

Then we have $\underline{b} < b_1$.[43] Now consider $b \in (\underline{b}, \min\{b_0, b_1\})$. Then $b$ satisfies A1 and A3, and, due to Lemma 9, the optimal static delegation set is an interval.

By definition of $\underline{b}$ and (22), we obtain in the special case for $b = \underline{b}$ that $\beta_h^0 = \bar{\omega}_\ell$. Then for $\beta$ converging to $\beta_h^0 = \bar{\omega}_\ell$ from below we have $\Delta V_h \to 0$, where the second integral in $\Delta V_h$ vanishes by definition of $\beta_h^0$ in (22). Moreover, it follows from a calculation that

$$\frac{d\Delta V_h}{d\beta}\bigg|_{\beta=\bar{\omega}_\ell} = -2[1 - F_h(\bar{\omega}_\ell) - b f_h(\bar{\omega}_\ell)], \tag{101}$$

which is negative for $b < b_1$. This means that $\Delta V_h > 0$ for some $\beta < \bar{\omega}_\ell$ and $b > \underline{b}$ such that $b$ is close to $\underline{b}$ and $\beta$ is close to $\bar{\omega}_\ell$. Consequently, $\Delta V_h > 0$ when $b > \underline{b}$ is close to $\underline{b}$ and $\mu_h$ is close to 1 (as then $\beta^{st}$ is close to $\beta_h^0$ which is close to $\bar{\omega}_\ell$). □

*Example satisfying Assumptions A1–A5.* Let us provide an example of a distribution and parameters that satisfy all assumptions. Consider the linear density $f(\omega) = 2\omega$ discussed in footnote 36. Then $F(\omega) = \omega^2$ and $\phi = \frac{2}{3}$. Assumption A2 is clearly satisfied. Assumption A1 becomes $b < \frac{2}{3}\ell$, while Assumption A3 can be rewritten as $\frac{1}{9}(h - \ell)(8h + \ell)/(5h+\ell) < b$. These *two* inequalities are consistent when $8 - 37(\ell/h) - 7(\ell/h)^2 < 0$, i.e., $\ell/h > \frac{37+3\sqrt{177}}{16} \approx 0.208$. Now, the expression in Assumption A4 becomes $\frac{1}{3}[-z+2/(1+z)]$, which is clearly convex. Finally, Assumption A5 holds as well, as $f$ is log-concave and $f_h(\omega)/f_\ell(\omega) = \ell^2/h^2 \cdot (2h+3\omega-2)/(2\ell+3\omega-2)$ is decreasing for all $\omega > \frac{2}{3}(1-\ell) = \underline{\omega}_\ell$. □

[43] To see this, observe that $\int_{\bar{\omega}_\ell}^{\bar{\omega}_h} [1 - F_h(\omega)]\, d\omega / [1 - F_h(\bar{\omega}_\ell)] < \bar{\omega}_h - \bar{\omega}_\ell < [1 - F_h(\bar{\omega}_\ell)]/f_h(\bar{\omega}_\ell)$, which follows from the monotonicity of $f_h$ and $F_h$.

# References


[1] Alonso, R. and N. Matouscheck (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] Amador, M. and K. Bagwell (2012a) The Theory of Optimal Delegation with An Application to Tariff Caps. Mimeo, Stanford University.

[4] Amador, M. and K. Bagwell (2012b) Tariff Revenue and Tariff Caps. *American Economic Review: Papers and Proceedings 102* (3), 459–465.

[5] Baron, D. and D. Besanko (1984). Regulation and Information in a Continuing Relationship. *Information Economics and Policy 1*, 267–302.

[6] Battaglini, M. (2005). Long-Term Contracting with Markovian Consumers. *American Economic Review 95*, 637–658.

[7] Bester, H. (2009). Externalities, Communication and the Allocation of Decision Rights. *Economic Theory 41*, 269–296.

[8] Blume, A., O. Board, and K. Kawamura (2008). Noisy talk. *Theoretical Economics 2*, 395–440.

[9] Courty, P. and H. Li (2000). Sequential Screening. *Review of Economic Studies 67*, 697–717.

[10] Crawford, V. P. and J. Sobel (1982). Strategic Information Transmission. *Econometrica 50*, 1431–1451.

[11] Dai, C., T. R. Lewis, and G. Lopomo (2006). Delegating Management to Experts. *RAND Journal of Economics 37*, 503-520.

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

[13] Esö, P. and B. Szentes (2007a). The Price of Advise. *RAND Journal of Economics 38*, 863–880.

[14] Esö, P. and B. Szentes (2007b). Optimal Information Disclosure in Auctions and the Handicap Auction. *Review of Economic Studies 74*, 705–731.

[15] Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009). Mediation, Arbitration, and Negotiation. *Journal of Economic Theory 144*, 1397–1420.

[16] Holmström, B. (1977). On Incentives and Control in Organizations. Ph.D. thesis, Stanford University.

[17] 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.

[18] Hoffmann, F. and R. Inderst (2011). Presale Information. *Journal of Economic Theory 146* (6), 2333–2355.

[19] Inderst, R. and M. Peitz (2012). Informing Consumers about their own Preferences. *International Journal of Industrial Organization 30*, 417-428

[20] Kartik, N., M. Ottaviani, and F. Squintani (2007). Credulity, Lies, and Costly Talk. *Journal of Economic Theory 134* (1), 93–116

[21] Kováč, E. and T. Mylovanov (2009). Stochastic Mechanisms in Settings without Monetary Transfers: The Regular Case. *Journal of Economic Theory 144* (4), 1373–1395.

[22] Krähmer, D. (2006). Message-Contingent Delegation. *Journal of Economic Behavior and Organization 60*, 490–506.

[23] 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.

[24] Krähmer, D. and R. Strausz (2011). Optimal Procurement Contracts with Pre-Project Planning. *Review of Economic Studies 78*, 1015–1041.

[25] Krähmer, D. and R. Strausz (2012). The Benefits of Sequential Screening. Mimeo.

[26] Martimort, D. and A. Semenov (2006). Continuity in Mechanism Design without Transfers. *Economics Letters 93*, 182–189.

[27] Myerson, R. (1986). Multistage Games with Communication. *Econometrica 54*, 323–358.

[28] Mylovanov, T. (2008). Veto-based Delegation. *Journal of Economic Theory 138*, 297–307.

[29] Nocke, V., M. Peitz, and F. Rosar (2011). Advance-purchase Discounts as a Price Discrimination Device. *Journal of Economic Theory 146*, 141–162.

[30] Ottaviani, M. and F. Squintani (2006). Naive Audience and Communication Bias. *International Journal of Game Theory 35* (1), 129–150.

[31] Pavan, A., I. Segal, and J. Toikka (2012). Dynamic Mechanism Design. Mimeo.

[32] Riordan, M. H. and D. E. Sappington (1987). Information, Incentives and Organizational Mode. *Quarterly Journal of Economics 102*, 243–264.

[33] Semenov, A. P. (2008). Bargaining in the appointment process, constrained delegation and the political weight of the Senate. *Public Choice 136*, 165–180.

[34] Semenov, A. P. (2012). Delegation to a Potentially Uninformed Agent. Mimeo.

[35] Szalay, D. (2005). The Economics of Clear Advice and Extreme Options. *Review of Economic Studies 72*, 1173–1198.