

Optimal Sequential Delegation*

Daniel Krähmer^{a,†} Eugen Kováč^{a,b}

March 3, 2016

Abstract

The paper extends the optimal delegation framework to a dynamic environment where the agent initially has private information merely about the distribution of the state and learns the true state only as the relation proceeds. The principal may want to elicit the agent's initial information and offers a menu of delegation sets where the agent first chooses a delegation set and subsequently an action within this set. We characterize environments under which it is optimal and under which it is not optimal to elicit the agent's initial information and characterize optimal delegation menus. In the former case, delegation sets may be disconnected and may feature gaps.

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

JEL Codes: D02, D20, D82, D86.

*We would like to thank the editor Alessandro Pavan, an associated editor, and two referees for very fruitful comments. We also thank Ricardo Alonso, Matthias Kräkel, Tymofiy Mylovanov, Mark Le Quement, Juuso Toikka, and seminar audiences in Bonn, Madrid, Mannheim, and Paris for helpful comments. D. Krähmer gratefully acknowledges financial support by the DFG (German Science Foundation) under SFB/TR-15.

^aDepartment of Economics, University of Bonn, Adenauerallee 24–42, 53113 Bonn, Germany.

^bMercator School of Management, University of Duisburg-Essen, Lotharstr. 65, 47057 Duisburg, Germany.

[†]Corresponding author.

E-mail addresses: kraehmer@hcm.uni-bonn.de (D. Krähmer), eugen.kovac@uni-due.de (E. Kováč).

1 Introduction

How much decision making discretion should be delegated to privately informed, but self-interested agents is an important question for the optimal design of firms' governance structures, hierarchies, or for the regulation of markets. Following the seminal work of Holmström [14, 15], a large body of literature studies this question in a setting where monetary transfers are infeasible, and a principal offers a set of permissible actions (a *delegation set*) to an agent who is perfectly informed about the decision relevant "state of the world". The principal has to trade off her benefit from utilizing the agent's information and her costs of giving up control. In the standard setting with one-dimensional state and action spaces, the literature has identified conditions under which an optimal delegation set takes the remarkably simple form of an interval. This amounts to imposing upper and lower thresholds on the agent's permissible actions. These thresholds depend on the distribution of the state as well as the parties' conflict of interest (*bias*).

In this paper, we study a novel issue by considering a dynamic setting where the agent learns the state only gradually while, at the outset, he privately knows the distribution of the state (to which we refer as the agent's *type*). Rather than offering the same delegation set to all types, a principal may then want to elicit (screen) the agent's type by offering a *menu of delegation sets*.¹ Facing such a menu, the agent first chooses a delegation set when he only knows his type, and then, after having observed the state, chooses an action from this set. Since monetary transfers are infeasible, the only screening instrument available is the degree of discretion provided by the delegation sets in the menu.

A classical application is the regulation of a monopolist (agent) where monetary transfers are infeasible, and the regulator (principal) determines the monopolist's discretion over prices. In this case, the state corresponds to the monopolist's marginal costs, and the regulator's objective is a weighted average of consumer surplus and profits. Our setup captures the situation that at the regulation stage, marginal costs have not yet been realized, but the monopolist has private beliefs about his future costs. In practice, offering a menu of delegation sets resembles a regulatory framework where the regulator does not impose a single regulatory plan but offers the regulated firm a choice among several plans. Such regulatory options are observed in practice in the telecommunication industry and can be interpreted as a screening instrument of the regulator.²

The contribution of our paper is to characterize environments under which it is optimal and environments under which it is not optimal for the principal to elicit the agent's

¹Considering delegation menus is in the spirit of Holmström's [14, 15] delegation principle. From an optimal mechanism design perspective, this means we restrict attention to deterministic mechanisms.

²Sappington [38, p. 234] reports the case of regional Bell Operating Companies which, in the 90s, were offered a menu of regulatory plans involving various combinations of earning sharings and price caps and, in line with our model, argues that regulatory options elicit "the regulated firm's superior knowledge of its operating environment" [38, p. 272].

type.³ In addition, we characterize optimal delegation menus. We show that for a large, precisely determined class of environments, it is not optimal to elicit the agent’s type. In this case, interval delegation remains optimal, establishing the robustness of interval delegation as an optimal mechanism even if the agent’s private information arrives sequentially. In particular, this is the case whenever the bias is sufficiently small. Crucially, however, we also show that the sequential arrival of information may call for richer forms of restricting the agent’s discretion beyond simply imposing thresholds. In these environments, the agent’s type is elicited, and optimal delegation sets may feature “gaps”, allocating discretion over “extreme” actions only.

As a key conceptual contribution of our analysis, we identify a measure that captures precisely the principal’s dynamic trade-off when designing a delegation menu. We refer to this measure as *virtual likelihood ratio*. Formally, the virtual likelihood ratio is defined as the likelihood ratio of agent type i ’s over agent type j ’s beliefs adjusted by an additional term reflecting the bias. Intuitively, consider some state and suppose that, starting from a situation in which the principal offers all actions to the agent, she inserts a tiny gap in the delegation set by removing a small set of actions around the agent’s ideal action in this state. In our setting, inserting a gap in the delegation set turns out to be costly in expected terms for both the principal and the agent. The virtual likelihood ratio in the given state is then equal to the ratio of the principal’s over the agent’s expected costs from inserting a gap in the limit as the gap gets small.⁴

Our results can be summarized as follows. First, if virtual likelihood ratios are decreasing, then, for an arbitrary (finite) number of types, *static delegation* is optimal, that is, the principal does not benefit from eliciting the agent’s type. Regardless of the shape of the virtual likelihood ratio, we also show that static delegation is optimal when the bias is small.⁵ Second, if virtual likelihood ratios are not decreasing, we provide sufficient conditions and necessary conditions for *sequential delegation* to be optimal, that is, for when the principal does benefit from eliciting the agent’s type. For tractability reasons, we establish these conditions only for the case with two agent types.⁶ We show that sequential delegation is optimal essentially if and only if the minimum of the appropriate virtual likelihood ratio is sufficiently small. Moreover, we characterize the shape of an

³We allow for environments, that is, distributions and bias functions, which essentially satisfy the conditions that, in a static setting, are necessary and sufficient for the optimal delegation set to be an interval as shown by Amador and Bagwell [3]. As explained in detail below, we allow for the same general preferences but consider slightly less general environments.

⁴The virtual likelihood ratio is a notion of a “virtual valuation” tailored to our dynamic mechanism design setting without transfers and can be seen as an analogue to what, in dynamic mechanism design with money, is sometimes referred to as informativeness measure as in Bester [4] and Courty and Li [8], or as impulse response function as in Pavan et al. [35].

⁵Even though the formal proof for this result is similar to the case with decreasing virtual likelihood ratios, the economic reason is rather different, as we explain the main text.

⁶In Section 5, we explain the difficulties with extending the analysis to more types.

optimal menu if the appropriate virtual likelihood ratio is increasing. Then an optimal delegation menu contains a delegation set that is either an interval that constrains the agent’s discretion from both above and below, or else the union of an isolated smallest action and an interval. Thus, optimal delegation sets may be disconnected, featuring gaps “at the bottom” and allocating discretion over “extreme” actions only.

The intuition for why the virtual likelihood ratio plays such a central role for our analysis can be best explained in the case with two types. In our setting, if types were publicly known, each type would be offered an interval with an upper but no lower threshold, the “low” type facing a more stringent threshold than the “high” type. The best static delegation menu, which does not discriminate between types, then corresponds to an interval whose threshold is a weighted average of these two thresholds. The basic idea to improve upon the static menu is to offer the high type a delegation set where his “discretion is shifted upwards”. This means that a more lenient threshold is imposed on the high type, while he is being banned from taking some (previously available) smaller actions. As we formally show, if the high type’s beliefs dominate the low type’s beliefs in the likelihood ratio order, then the high type benefits more, in expected terms, than the low type when the discretion provided by a delegation set is shifted upwards.⁷

Intuitively, since the high type receives a more lenient threshold when types are public, the low type’s incentive constraint will be binding at an optimal delegation menu. This means that the low type is just deterred from picking the delegation set offered to the high type, and thus his expected costs from smaller actions not being available are equal to his expected gains from larger actions being available. Hence, whether the principal benefits from shifting the high type’s discretion upwards, while keeping the low type indifferent, depends on how her *relative* expected costs from removing small actions, relative to the corresponding costs for the low type, compare to her *relative* expected gains from adding larger actions, relative to the corresponding gains for the low type.

Drawing on a representation of the principal’s payoff developed in Kováč and Mylovánov [21], we show that the relative costs of removing a tiny set of actions in the high type’s delegation set are given by the virtual likelihood ratio. Intuitively then, if the virtual likelihood ratio is increasing, the most cost effective way to elicit the agent’s type is to remove small actions in the high type’s delegation set in exchange for extending the high type’s threshold. If sequential delegation is optimal, the high type is therefore banned from choosing small actions. Reversely, if the virtual likelihood ratio is decreasing, then roughly speaking, it is best to remove actions directly below the high type’s threshold. As we show, this will imply that static delegation is optimal.

While we focus on the case where the bias is common knowledge, our insights readily

⁷Thus, the ranking of beliefs in terms of the likelihood ratio order in our setting is analogous to the “Spence-Mirrlees” single-crossing condition in a standard screening problem.

extend to the interesting special case that the bias is the agent’s private information, and the distribution of the state is commonly known. For distributions with log-concave densities it turns out that virtual likelihood ratios are decreasing. It is thus immediate from our results that the principal does not benefit from eliciting the agent’s bias. We also argue that with two types, our necessary conditions for the optimality of screening the bias are rarely satisfied in the sense that many commonly used distributions violate these conditions. Along with our results on the case with type-independent bias, this suggests the general insight that what primarily matters for sequential delegation to be optimal is the agent’s private information about the distribution, not about the bias.

Related literature

Our paper brings together the literatures on optimal delegation and on sequential screening (with money). Within the optimal delegation literature, our paper is most closely related to Amador and Bagwell [3] who, to our knowledge, present the most general sufficient and necessary conditions in the literature for interval delegation to be optimal.⁸ Next to the general setup, we draw on their sufficient conditions to prove our results on the optimality of static delegation. The basic idea is to set up a relaxed problem and to verify that the static delegation menu maximizes the Lagrangian associated to this problem. To do so, we extend the sufficiency conditions in [3] to be applicable to our Lagrangian objective.

Our result that the principal may elicit the agent’s type by allocating discretion over only extreme actions is similar to the insight by Szalay [42] that offering extreme options is optimal to induce an initially uninformed agent to acquire information. In fact, just like offering extreme options to the high type dissuades the low type from mimicking the high type in our case, so it dissuades the uninformed agent from staying uninformed in [42]. Unlike [42], we explicitly identify monotonicity and single-crossing conditions under which screening types is feasible.⁹ Moreover, [42] focuses on information acquisition costs as the main determinant of the value of inducing information acquisition for a given distribution of the state. In contrast, we identify how the benefits of screening depend on the distribution of the state and the bias. In particular, inducing information acquisition is optimal in [42] as long as information acquisition costs are small whereas in our setting, eliciting the agent’s type is never optimal for certain distributions, or, whenever the bias is sufficiently small.¹⁰

⁸Other contributions to the optimal delegation literature include Alonso and Matouscheck [1], Ivanov [18], Koessler and Martimort [19], Kolotilin et al. [20], Kováč and Mylovanov [21], Liang [28], Martimort and Semenov [30], Melumad and Shibano [31], and Mylovanov [33]. Unlike the optimal delegation literature, a large literature studies delegation when decisions are non-contractible. See, for instance, Alonso et al. [2], Bester [6], Bester and Krämer [7], Dessein [10], Goltsman et al. [13], Krämer [23], Riordan and Sappington [36], and Semenov [39].

⁹In particular, this addresses the issue of incentive compatibility in the “other direction”, that the high type does not mimic the low type. This issue, naturally, does not arise in [42].

¹⁰While Szalay [42] considers an unbiased agent, his results are robust to introducing a small bias.

Two other contributions study an optimal delegation setting in which the agent is not perfectly informed about the state. In Semenov [40], the agent only learns an imperfect signal about the state. This corresponds to a static setup where the distribution of the state features a mass-point, which induces a gap in the optimal delegation set. In contrast, in our paper a gap helps to elicit the agent’s type.¹¹ In independent work, Tanner [43] does study sequential delegation but does not allow the agent to have private information about the distribution, only about the bias. Using different arguments than [43], we arrive at the same result that eliciting the bias is not optimal, but our methods apply to a somewhat larger environment.

Within the sequential screening literature, our paper is closely related to Courty and Li [8] who study a dynamic price discrimination problem where, as in our model, the agent knows the distribution of the state (his valuation) at the time of contracting, but learns the state only afterwards.¹² Unlike in [8], monetary transfers are not feasible in our setting, and we develop a different approach to derive an optimal mechanism. Our results on the optimality of static delegation are closely related to Krähmer and Strausz [26] who establish sufficient conditions so that a static contract, which pools types, is optimal in the setting of [8] when the agent has an ex post outside option. As in [26], we consider a setting with discrete types, and to show that static delegation is optimal, we solve a relaxed problem where the incentive constraints we impose for a given type are the same as those imposed for the corresponding type in the relaxed problem of [26].¹³

The paper is organized as follows. Section 2 describes the model. Section 3 provides monotonicity and single-crossing notions. Section 4 provides sufficient conditions for the optimality of static delegation, while section 5 provides necessary conditions and sufficient conditions for the optimality of sequential delegation. Section 6 considers the case with type-dependent bias, and Section 7 concludes. All proofs are in the appendix.

2 Model

A principal (she) and an agent (he) seek to implement a contractible action $x \in \mathbb{R}$ whose payoff depends on a *state* of the world $\omega \in [\underline{\omega}, \bar{\omega}]$. There are three periods. In period 1, no party knows the true state, but the agent privately knows that the state is distributed with cdf F_i on the support $[\underline{\omega}, \bar{\omega}]$. While the agent’s *type* i is his private information, it is

¹¹An interesting question is what would happen in Semenov [40] if the agent’s information was revealed sequentially. Our analysis does not, however, directly speak to this question, since our framework does not cover the case where one agent type remains entirely ignorant.

¹²For work on dynamic mechanism design with money, see also Battaglini [5], Baron and Besanko [4], Dai et al. [9], Esö and Szentes [11, 12], Hoffmann and Inderst [16], Inderst and Peitz [17], Krähmer and Strausz [24, 25, 26, 27], Nocke et al. [34], and Pavan et al. [35].

¹³The relaxed problem is quite non-standard, as it involves global, instead of the familiar adjacent incentive constraints.

common knowledge that i is drawn from $\{1, \dots, I\}$ with probability $\mu_i > 0$. In period 2, the agent privately observes the true state ω , and in period 3, the action is implemented. We assume that F_i is twice continuously differentiable with strictly positive pdf $f_i = F'_i$.

Drawing on Amador and Bagwell [3], the agent's utility function U and the principal's utility function V are given as

$$U(\omega, x) = \omega x + a(x), \quad (1)$$

$$V(\omega, x) = \omega x + a(x) - b(\omega)x + c(\omega), \quad (2)$$

where $a : \mathbb{R} \rightarrow \mathbb{R}$ is twice differentiable and strictly concave, $b : [\underline{\omega}, \bar{\omega}] \rightarrow \mathbb{R}$ is continuously differentiable, and $c : [\underline{\omega}, \bar{\omega}] \rightarrow \mathbb{R}$ is integrable. For reasons evident in the next paragraph, we assume that $\lim_{x \rightarrow -\infty} -a'(x) < \underline{\omega}$ and $\lim_{x \rightarrow +\infty} -a'(x) > \bar{\omega}$. Let

$$x_A(\omega) \equiv \arg \max_{x \in \mathbb{R}} U(\omega, x) \quad (3)$$

be the agent's *favorite action*. By our assumptions on a , $x_A(\omega)$ is uniquely determined by the first order condition, $\omega + a'(x_A(\omega)) = 0$, and is differentiable and strictly increasing.¹⁴ We refer to the function b as the agent's *bias*, because it is a measure for the distance between the agent's and the principal's favorite actions.¹⁵ As a special case, if b does not depend on the state, we say the agent has constant bias.¹⁶

In period 1, the principal offers a mechanism to the agent. By the revelation principle for dynamic games (Myerson, [32]), we can restrict attention to direct and incentive compatible mechanisms which implement an action as a function of sequential reports by the agent about type and state, and induce the agent to report truthfully. By a straightforward extension of Holmström's [14, 15] delegation principle, any such mechanism can be implemented by an *incentive compatible menu of delegation sets* (D_1, \dots, D_I) , where $D_i \subseteq \mathbb{R}$ specifies a set of contractible actions that the agent is permitted to take. Under such a menu, the agent selects a delegation set D_i from the menu before he observes the state, and then, after having observed the state, chooses an action x in D_i . A menu is incentive compatible if agent type i selects D_i from the menu.¹⁷

¹⁴Throughout the paper, we use the terms "increasing" and "decreasing" for weak monotonicity. When we mean strict monotonicity, we explicitly write "strictly increasing" and "strictly decreasing".

¹⁵Indeed, if $a'(x) = -x$, then the agent's favorite action is $x_A(\omega) = \omega$ while the principal's is $\omega - b(\omega)$. Note that it is not substantial that the bias formally enters V rather than U , but turns out to be more convenient analytically. Note also that the specification includes the familiar case with quadratic preferences: Define $a(x) = -1/2 \cdot x^2$, and add the action-irrelevant constant $-1/2 \cdot \omega^2$ to U .

¹⁶While we focus on the case with type-independent bias in the main part of the analysis, we will discuss the case with type-dependent bias $b_i(\omega)$ in Section 6.

¹⁷Focusing on incentive compatible delegation menus is restrictive only in that it rules out "stochastic" mechanisms which may implement lotteries over actions. For a detailed proof of the delegation principle in the current context, see our working paper Kováč and Krähmer [22]. As is common when applying the revelation principle, we also assume that type i , when indifferent between D_i and D_j , selects D_i for

Formally, denote the action the agent chooses from a delegation set D in state ω by¹⁸

$$x_D(\omega) \in \arg \max_{x \in D} U(\omega, x). \quad (4)$$

Conditional on type i , the agent's and the principal's expected utility from D are therefore

$$U_i(D) \equiv \int_{\underline{\omega}}^{\bar{\omega}} U(\omega, x_D(\omega)) dF_i(\omega), \quad V_i(D) \equiv \int_{\underline{\omega}}^{\bar{\omega}} V(\omega, x_D(\omega)) dF_i(\omega). \quad (5)$$

A delegation menu (D_1, \dots, D_I) is incentive compatible if for all i, j :

$$IC_{i,j} : \quad U_i(D_i) \geq U_i(D_j). \quad (6)$$

The principal's problem, referred to as \mathcal{P} , can then be written as

$$\mathcal{P} : \quad \max_{D_1, \dots, D_I} \sum_i \mu_i V_i(D_i) \quad s.t. \quad IC_{i,j} \quad \text{for all } i, j.$$

In general, the solution to \mathcal{P} will not be unique.¹⁹ This is so, as one can always add *redundant* actions to the solution which would not be chosen by the agent in any state. We therefore restrict attention to *minimal* delegation menus where any action in a delegation set is strictly chosen in some state and unchosen (redundant) actions are removed.^{20,21} We denote a solution to \mathcal{P} by (D_1^*, \dots, D_I^*) . We say that *static delegation is optimal* if there is a solution with $D_i^* = D_j^*$ for all i, j . Otherwise, *sequential delegation is optimal*.

Delegation sets that are intervals will be crucial for our analysis. Since $x_A(\omega)$ is strictly increasing in ω , a delegation set D is an interval if, and only if there are $\omega_L \leq \omega_H$ so that $D = [x_A(\omega_L), x_A(\omega_H)]$. We say D is *truncated from above* if $\omega_L = \underline{\omega}$ and $\omega_H < \bar{\omega}$.

Assumptions and static trade-offs

Throughout the paper, we impose the following assumptions on the bias function and the distributions (and we will not mention them explicitly in the statement of results).

A1. $f_i(\omega) + (b(\omega)f_i(\omega))' > 0$ for all ω and i .

A2. $b(\underline{\omega}) \geq 0$ and $b(\bar{\omega}) > 0$.

A3. $\int_{\underline{\omega}}^{\bar{\omega}} 1 - F_i(\omega) - b(\omega)f_i(\omega) d\omega > 0$ for all i .

sure. This prevents stochastic allocations to result from mixing by the agent. See also Remark 2 below.

¹⁸Without 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 Amador and Bagwell [1].

¹⁹Analogously as in Homlström [15], Theorem 1, it can be shown that a solution to \mathcal{P} exists.

²⁰Amador and Bagwell [1] proceed in the same fashion.

²¹Observe that removing these actions does not upset incentive compatibility. In fact, it relaxes incentive compatibility, as removing actions from D_j only reduces agent type i 's incentive to select it.

For the benchmark case in which the agent’s type i is publicly known, Assumption A1 will guarantee that it is suboptimal to insert a “gap” in a delegation set, and A3 will guarantee that it is suboptimal to implement the action that would be best for the principal in the absence of the agent. Jointly, A1–A3 will guarantee that for publicly known type, the optimal delegation set is a non-degenerate interval truncated from above.²²

To illuminate this, consider the principal’s basic trade-off. On the one hand, the principal benefits from a larger delegation set since this allows her to better utilize the agent’s information. On the other hand, due to conflicting interests, she is hurt from a larger delegation set since this allows the agent to take more biased actions. To capture these effects formally, let

$$u_D(\omega) = U(\omega, x_D(\omega)) = \max_{x \in D} U(\omega, x) \quad (7)$$

be the agent’s utility in state ω . A familiar envelope and integration by parts argument delivers that the principal’s expected utility from a delegation set is²³

$$V_i(D) = \underbrace{u_D(\underline{\omega})b(\underline{\omega})f_i(\underline{\omega})}_{J_i^I(D)} + \underbrace{\int_{\underline{\omega}}^{\bar{\omega}} u_D(\omega)[f_i(\omega) + (b(\omega)f_i(\omega))'] d\omega}_{J_i^{II}(D)} - \underbrace{u_D(\bar{\omega})b(\bar{\omega})f_i(\bar{\omega})}_{LC_i(D)} + C, \quad (8)$$

where C is a constant independent of u_D . Expression (8) says that the principal’s expected utility is a weighted average of the agent’s utility u_D across states: In the lowest state $\underline{\omega}$, the agent’s utility receives the weight $b(\underline{\omega})f_i(\underline{\omega})$, and in interior states $\omega \in (\underline{\omega}, \bar{\omega})$, it receives weight $f_i(\omega) + (b(\omega)f_i(\omega))'$. By Assumptions A1 and A2, these weights are *positive*, and therefore, the terms $J_i^I(D)$ and $J_i^{II}(D)$ jointly can be interpreted as a beneficial *information effect* that results from increasing the agent’s discretion. We will refer to them respectively as *first* and *second information effect*. Moreover, we will refer to $f_i(\omega) + (b(\omega)f_i(\omega))'$ as *virtual likelihood*.

In contrast, in the largest state $\bar{\omega}$, the agent’s utility receives the weight $-b(\bar{\omega})f_i(\bar{\omega})$, which is *negative* by A2. Hence, the principal’s utility is smaller the better off the agent in state $\bar{\omega}$, and therefore, $LC_i(D)$ can be interpreted as a costly *loss of control effect* that results from increasing the agent’s discretion.

²²Within the class of preferences (1) and (2), Amador and Bagwell [3] provide necessary and sufficient conditions for the optimality of interval delegation. Our conditions A1–A3 resemble the necessary conditions in [3] and are only slightly more restrictive. Thus, within the preference class (1) and (2), relaxing A1–A3 would essentially require to allow for non-intervals to be optimal with public types, which would significantly complicate the analysis. Moreover, as shown in [3], for a negative bias at $\underline{\omega}$ and $\bar{\omega}$, optimal intervals are truncated from below. Our positive bias assumption A2 primarily reduces case distinctions but is not substantial for the nature of our arguments (though specific conclusions might differ).

²³To our knowledge, the argument was first developed in Kováč and Mylovánov [21]. See the appendix for details.

Benchmarks

We now consider two important benchmark cases. In the first benchmark, the agent's type is publicly known, in which case we have:

Lemma 1. *Let type i be publicly known. Then the optimal delegation set is an interval truncated from above, i.e., $D_i^0 = [x_A(\underline{\omega}), x_A(\omega_i^0)]$, where $\omega_i^0 \in (\underline{\omega}, \bar{\omega})$ is the (unique) solution to the equation*

$$\int_{\omega_i^0}^{\bar{\omega}} 1 - F_i(\omega) - b(\omega)f_i(\omega) d\omega = 0. \quad (9)$$

In the second benchmark, the principal does not elicit the agent's type and is restricted to offer the same delegation set D to all types. We call a menu (D, \dots, D) a *static delegation menu*. The principal's expected utility from such a menu is $\int_{\underline{\omega}}^{\bar{\omega}} V(\omega, x_D(\omega)) dF^{st}(\omega)$ where $F^{st} = \sum_i \mu_i F_i$ is the average distribution of the state, averaged across types.

Lemma 2. *The optimal static delegation menu consists of intervals of the form $D^{st} = [x_A(\underline{\omega}), x_A(\omega^{st})]$, where $\omega^{st} < \bar{\omega}$ is given by (9) with F_i replaced by F^{st} .*

Since Assumptions A1 and A3 carry over to F^{st} , Lemma 2 follows immediately from 1. The intuition behind Lemma 1 can be seen from (8). In state $\bar{\omega}$, the agent chooses the highest action available, say \bar{x} . Since the bias is positive in $\bar{\omega}$, it is easy to see that $\bar{x} < x_A(\bar{\omega})$ at an optimal delegation set. "Filling all gaps below \bar{x} " is then beneficial, because allowing the agent to pick all smaller actions $x < \bar{x}$ does not affect $u_D(\bar{\omega})$ and so does not affect the loss of control effect. But, it (weakly) improves $u_D(\omega)$ in the other states $\omega \in [\underline{\omega}, \bar{\omega})$ and so (weakly) raises both information effects. Thus, interval delegation of the form $[x_A(\underline{\omega}), \bar{x}]$ is optimal. Finally, increasing \bar{x} raises both the second information effect and the loss of control effect. Equation (9) is the first order condition where these two countervailing forces are in balance, and Assumption A3 ensures that ω_i^0 is interior.²⁴

As a convention, we from now on (re-)label the agent's types so that²⁵

$$\omega_1^0 < \dots < \omega_I^0. \quad (10)$$

Thus, the optimal delegation menu with public types would violate incentive compatibility, because "low" types i would have incentives to mimic "high" types $j > i$. (Remark 3 below states a sufficient condition on the distributions F_i for the ordering (10).)

²⁴While Lemma 1 follows therefore directly from Amador and Bagwell [3], we give an alternative, more elementary, proof that extends the domain of objective functions for which the sufficient conditions in [3] deliver the optimality of interval delegation. The main benefit of this exercise will, however, accrue in our analysis of the sequential delegation problem where such objective functions naturally arise in the form of Lagrangians.

²⁵To reduce case distinction, we focus on the case that the ω_i^0 's are strictly ordered.

Next, we present an example that is used to illustrate our results throughout.

Example. Consider the problem of regulating a monopolist who chooses a price x facing a linear demand $q = A - x$ and constant marginal costs ω , where $A > 1$ and $\omega \in [\underline{\omega}, \bar{\omega}] = [0, 1]$. At the outset, the monopolist has private beliefs F_i about his future marginal costs. Once production starts, the monopolist learns the actual value of ω . The monopolist maximizes his profit, while the regulator's objective is a weighted average of consumer surplus and profit, with weights ξ and $1 - \xi$ respectively, where $\xi \in [0, 1]$ measures the extent to which the regulator is “pro-consumer”. Thus,²⁶

$$U(\omega, x) = \omega x + (A - x)x, \quad V(\omega, x) = \omega x + (A - x)x - \gamma(A - \omega)x, \quad (11)$$

with linear bias function $b(\omega) = \gamma(A - \omega)$, where $\gamma \geq 0$ and $\gamma = \xi/(2 - 3\xi)$. The agent's favorite action is then the monopoly price $x_A(\omega) = \frac{1}{2}(A + \omega)$. For $\gamma = 0$, the regulator is entirely “pro-business”, while for $\gamma = 1$, she maximizes total surplus.

Consider the case with two types $i = 1, 2$ and distributions

$$F_1(\omega) = \omega, \quad F_2(\omega) = \kappa_3\omega^3 + \kappa_2\omega^2 + (1 - \kappa_3 - \kappa_2)\omega, \quad (12)$$

where $\kappa_3, \kappa_2 \geq 0$ and $\kappa_3 + \kappa_2 < 1$. Since f_2/f_1 is increasing, type 1's marginal costs are smaller in the likelihood ratio order, and hence type 1 corresponds to the “more efficient” type. (This also ensures (10), see Remark 3 below.) Assumption A2 is clearly satisfied, and A1 is satisfied if and only if $\gamma < 1$, while A3 is satisfied if and only if $\gamma < 1/(2A - 1)$. Then, Lemma 1 and 2 apply, and optimal static delegation involves a price cap. \square

Remarks

Remark 1. If the second inequality in A2 is violated, i.e., if $b(\bar{\omega}) \leq 0$, then all weights in (8) are non-negative. Thus, the principal's expected utility is maximized when $u_D(\omega)$ is maximized point-wise for any $\omega \in [\underline{\omega}, \bar{\omega}]$. Hence, “full delegation” is optimal: $D_i^0 = [x_A(\underline{\omega}), x_A(\bar{\omega})]$ for all i . In this case, full delegation also solves the sequential problem \mathcal{P} .

Remark 2. While our restriction to delegation menus, and thus to deterministic mechanisms follows most of the delegation literature, stochastic mechanisms may especially help to relax our novel “dynamic” incentive constraints $IC_{i,j}$.²⁷ The restriction to deter-

²⁶A straightforward computation yields monopoly profit $(x - \omega)(A - x)$ and consumer surplus $\frac{1}{2}(A - x)^2$. We thus set $a(x) = (A - x)x$. To obtain the principal's utility of the form (2), we normalize the regulator's objective by the factor $2/(2 - 3\xi)$ and set $b(\omega) = \xi/(2 - 3\xi) \cdot (A - \omega)$. For the normalization factor to be positive, we require $\xi \in [0, \frac{2}{3}]$, and thus $\gamma \geq 0$. We also omit terms that do not depend on the price.

²⁷The technical reason is that the principal maximizes the objective (8), which is linear in $u_{D_i}(\omega)$, $\omega \in [0, 1]$, subject to the linear constraints (6). Samuelson [37] also studies a mechanism design problem with a linear objective and linear constraints and shows examples where stochastic mechanisms are optimal. This suggests that stochastic mechanisms could be optimal in specific cases in our setting, too.

ministic mechanisms thus stacks the deck against sequential delegation to be optimal. We focus on deterministic mechanisms to facilitate comparison with the literature and, not least, due to serious tractability issues. An applied justification for ruling out stochastic mechanisms is the often articulated concern that they may be hard to enforce in practice.

Remark 3. By (9), the upper threshold ω_i^0 depends both on the distribution and the bias function. Therefore, there is, in general, no order of distributions that, by itself, would ensure the ordering (10). However, for constant bias, (9) implies that ω_i^0 is the point where the bias is equal to the *mean residual life*

$$mrl_i(\omega) = \frac{\int_{\omega}^{\bar{\omega}} 1 - F_i(\tilde{\omega}) d\tilde{\omega}}{1 - F_i(\omega)}. \quad (13)$$

Hence, for constant bias, (10) holds if F_i has decreasing mean residual life for all i , and F_i dominates F_{i-1} in the mean residual life order: $mrl_i(\omega) \geq mrl_{i-1}(\omega)$ for all $\omega \in [\underline{\omega}, \bar{\omega}]$. A sufficient condition for this is that F_i dominates F_{i-1} in the likelihood ratio order, i.e., f_i/f_{i-1} is increasing (see Shaked and Shanthikumar, [41], Theorem 1.C.1, p. 43 and Theorem 2.A.1, p. 83).

3 Upward shifts and single-crossing

We now turn to the analysis of the sequential delegation problem \mathcal{P} . We first consider the question under what conditions it is feasible and profitable to elicit the agent's type i . In standard screening problems, this is closely linked to monotonicity and single-crossing conditions. In this section, we provide analogues of these conditions for our setting.

To motivate these conditions, we illustrate the main issues in the two types case. Consider the question whether it is feasible and profitable for the principal to deviate from the static menu (D^{st}, D^{st}) , where $D^{st} = [x_A(\underline{\omega}), x_A(\omega^{st})]$, and instead elicit the agent's type by offering type 2 a different delegation set $\tilde{D}_2 \neq D^{st}$ while maintaining $D_1 = D^{st}$. Given that D^{st} is an interval truncated from above, any different set $\tilde{D}_2 \neq D^{st}$ that maintains incentive compatibility must contain actions which are larger than $x_A(\omega^{st})$ (otherwise, type 2 would pick $D_1 = D^{st}$), but must not include all actions in $[x_A(\underline{\omega}), x_A(\omega^{st})]$ (otherwise, type 1 would pick D_2). In this sense, \tilde{D}_2 provides more discretion than D^{st} over large actions, but less discretion over small actions. Generalizing this notion for arbitrary delegation sets \tilde{D} and D , we say that \tilde{D} is an *upward (discretion) shift of D* if there is $\hat{\omega} \in [\underline{\omega}, \bar{\omega}]$ so that

$$\begin{aligned} u_{\tilde{D}}(\omega) &\leq u_D(\omega) && \text{when } \omega \in [\underline{\omega}, \hat{\omega}], \\ u_{\tilde{D}}(\omega) &\geq u_D(\omega) && \text{when } \omega \in [\hat{\omega}, \bar{\omega}]. \end{aligned} \quad (14)$$

This means that in small states the agent is better off under D while in large states, he is better off under \tilde{D} . Consequently, since the agent's ideal action is increasing in the state, \tilde{D} contains at least one action which is at least as large as the largest action in D , and D contains at least one action which is at most as large as the smallest action in \tilde{D} . In this sense, \tilde{D} shifts discretion from small to large actions relative to D . In particular, any set \tilde{D} is an upward shift of an interval truncated from above provided \tilde{D} contains one action at least as large as the upper end of this interval.²⁸

Returning to our illustration with two types, when is it feasible to offer type 2 a delegation set \tilde{D}_2 which is an upward shift of D^{st} ? Suppose that discretion has been reallocated between D^{st} and \tilde{D}_2 so that type 1 is just indifferent between selecting D^{st} and \tilde{D}_2 (IC_{12} is binding). This means that if type 1 selects \tilde{D}_2 , then his expected gain from the increased discretion over large actions is just compensated by his expected loss from the reduced discretion over small actions. Intuitively, this maintains type 2's incentive constraint if type 2's beliefs dominate type 1's beliefs in a sufficiently strong stochastic dominance order. An appropriate order turns out to be the likelihood ratio order, which requires that f_2/f_1 is increasing. Type 2 then assigns higher likelihood than type 1 to large states and lower likelihood to small states. Since the agent's ideal action is increasing in the state, type 2, when selecting \tilde{D}_2 , is therefore more likely than type 1 to utilize the larger discretion over large actions, and he is less likely to suffer from the reduced discretion over smaller actions. In expected terms, type 2 therefore benefits more from the upward discretion shift than type 1, and since type 1 was indifferent by construction, type 2 prefers the upward discretion shift.

We generalize these considerations in part (a) of the next lemma where type i corresponds to type 2 and type j to type 1 in the example. Part (b) covers the logically identical reverse case.

Lemma 3. *Let \tilde{D} be an upward shift of D and consider types i, j .*

(a) *If f_i/f_j is increasing, then $U_j(\tilde{D}) \geq U_j(D)$ implies that $U_i(\tilde{D}) \geq U_i(D)$.*

(b) *If f_i/f_j is decreasing, then $U_j(\tilde{D}) \leq U_j(D)$ implies that $U_i(\tilde{D}) \leq U_i(D)$.*

Part (a) implies that if type i 's delegation set D_i is replaced by an upward shift (D replaced by \tilde{D}) in a way so that type j 's incentive constraint $IC_{j,i}$ is maintained ($U_j(D) = U_j(\tilde{D})$), then type i 's incentive constraint $IC_{i,j}$ is automatically satisfied. Part (b) has a similar interpretation. The monotone likelihood ratio conditions in part (a) and (b) are thus the analogues to the ‘‘Spence–Mirrlees’’ single-crossing condition in a standard screening problem.

²⁸Let $D = [x_A(\underline{\omega}), x_A(\omega')]$ be an interval with largest action $x_A(\omega')$ for some ω' . Suppose \tilde{D} contains an action larger or equal to $x_A(\omega')$. Let \tilde{y} be the smallest such action. Let $\hat{\omega}$ be the state in which the agent is indifferent between $x_A(\omega')$ and \tilde{y} . Then $u_{\tilde{D}}(\omega) \leq u_D(\omega)$ for $\omega \leq \hat{\omega}$, while $u_{\tilde{D}}(\omega) \geq u_D(\omega)$ for $\omega \geq \hat{\omega}$. Thus, \tilde{D} is indeed an upward shift of D .

We next turn to the question how shifting discretion upwards in an incentive compatible way affects the principal's utility. There are three effects, corresponding to the two information and the loss of control effects in expression (8).

Lemma 4. *Let \tilde{D} be an upward shift of D . Then for all types i , $J_i^I(\tilde{D}) \leq J_i^I(D)$ and $LC_i(\tilde{D}) \geq LC_i(D)$.*

The lemma says that shifting discretion upwards (weakly) decreases the first information effect, J_i^I , and (weakly) increases the loss of control effect, LC_i . Both forces (weakly) lower the principal's utility. Intuitively, when there is less (more) discretion over small (large) actions, then in the lowest (largest) state, a less (more) favorable action is available for the agent, thus $u(\underline{\omega})$ decreases and so does J_i^I ($u(\bar{\omega})$ increases and so does LC_i).

Lemma 4 makes clear that the only channel by which the principal can benefit from an incentive compatible upward shift is the second information effect, J_i^{II} . As the next lemma shows, whether this effect goes up or down depends on the monotonicity of the *virtual likelihood ratio* which will play a crucial role in our analysis:

$$\rho_{i,j}(\omega) = \frac{f_i(\omega) + (b(\omega)f_i(\omega))'}{f_j(\omega)}. \quad (15)$$

Lemma 5. *Let \tilde{D} be an upward shift of D and consider types i, j .*

- (a) *If $\rho_{i,j}$ is increasing, then $U_j(\tilde{D}) \geq U_j(D)$ implies $J_i^{II}(\tilde{D}) \geq J_i^{II}(D)$.*
- (b) *If $\rho_{i,j}$ is decreasing, then $U_j(\tilde{D}) \leq U_j(D)$ implies $J_i^{II}(\tilde{D}) \leq J_i^{II}(D)$.*

The reasoning behind Lemma 5 is the same as behind Lemma 3, yet since in the expression for J_i^{II} , the agent's utility is weighed by the *virtual likelihood*, what matters now is not the likelihood ratio but the *virtual likelihood ratio*.

Lemma 5 suggests an intuitive interpretation of the steepness of the virtual likelihood ratio as a measure of the costs and benefits of screening the agent. In the context of our motivating question, offering type 2 an upward shift \tilde{D}_2 of D^{st} which maintains type 1's incentive constraint will, all else equal, be less profitable, the less steep is $\rho_{2,1}$. To see this, suppose that type 1's beliefs put more weight on large states in the sense that they change from f_1 to \hat{f}_1 so that \hat{f}_1/f_1 is increasing, implying that the virtual likelihood ratio becomes less steep which means that $\hat{\rho}_{2,1}/\rho_{2,1}$ is increasing. This increases type 1's incentive to select \tilde{D}_2 instead of D^{st} , as he now considers it more likely to benefit from the larger discretion over larger actions and less likely to be hurt from the smaller discretion over smaller actions. Therefore, type 1's incentive compatibility constraint becomes more stringent: under a less steep virtual likelihood ratio, screening costs are higher. Similarly, if $f_2 + (bf_2)'$ changes to $\hat{f}_2 + (b\hat{f}_2)'$ so that $[\hat{f}_2 + (b\hat{f}_2)']/[f_2 + (bf_2)']$

is decreasing, the virtual likelihood ratio becomes less steep. Thus, shifting discretion upwards from D^{st} to \tilde{D}_2 becomes less profitable for the principal.

Note that the virtual likelihood ratio depends jointly on the bias and the distribution. Thus, the effect of the bias on the virtual likelihood ratio, and thus on the costs and benefits of screening, is in general not clear-cut.

Example (ctd). In the price regulation example where $b(\omega) = \gamma(A - \omega)$,

$$\rho_{i,j}(\omega) = (1 - \gamma) \frac{f_i(\omega)}{f_j(\omega)} + \gamma(A - \omega) \frac{f'_i(\omega)}{f'_j(\omega)}. \quad (16)$$

If types are ranked by the monotone likelihood ratio (f_i/f_j increasing for $i > j$), it follows that $\rho_{i,j}$ is increasing if the regulator is very pro-business (γ close to 0). For large $\gamma > 0$, the shape of $\rho_{i,j}$ is, in general, not clear-cut, however. For the specification (12),

$$\frac{d\rho_{2,1}(\omega)}{d\omega} = 6\kappa_3(1 - 3\gamma)\omega + 6\kappa_3A\gamma + 2\kappa_2(1 - 2\gamma). \quad (17)$$

Hence, $\rho_{2,1}$ is increasing if and only if $3A\kappa_3 \geq 6\kappa_3 + \kappa_2$ or $\gamma \leq (3\kappa_3 + \kappa_2)/(9\kappa_3 + 2\kappa_2 - 3A\kappa_3)$, that is, if the market size A is sufficiently large, or the regulator is sufficiently pro-business. Further, $\rho_{2,1}$ is decreasing if and only if $3A\kappa_3 < \kappa_2$ and $\gamma \geq \kappa_2/(2\kappa_2 - 3A\kappa_3)$, that is, if the market size A is sufficiently small, and the regulator is sufficiently pro-consumer. \square

4 Optimality of static delegation

In this section, we derive sufficient conditions for static delegation to be optimal. We show that static delegation is optimal whenever virtual likelihood ratios are decreasing and whenever the bias is small.

Proposition 1. *If $\rho_{i,j}$ is decreasing for all $i, j \in \{1, \dots, I\}$ such that $i > j$, then the optimal static menu (D^{st}, \dots, D^{st}) solves the principal's problem \mathcal{P} .*

The intuition can best be seen for two types $i = 1, 2$. Consider the relaxed problem where we ignore type 2's incentive constraint. Type 1's incentive constraint is then binding at the optimum of the relaxed problem: $U_1(D_1) = U_1(D_2)$. This is intuitive since with public types, $D_1^0 \subset D_2^0$, and so type 1's incentive constraint is violated. Moreover, D_1 is an interval truncated from above: $D_1 = [x_A(\underline{\omega}), x_A(\omega_1)]$, $\omega_1 < \bar{\omega}$. Intuitively, D_1 would contain gaps otherwise. Adding some actions to the gap would then relax type 1's incentive constraint and increase the principal's utility, because it would increase both information effects J_1^I and J_1^{II} in (8), but not affect the loss of control effect LC_1 (because when adding an action in a gap, the action in state $\bar{\omega}$ and thus $u_D(\bar{\omega})$ remain the same).

To see that the relaxed problem has a static solution, suppose to the contrary that $D_2 \neq D_1$. We argue that the principal can improve by offering the static menu (D_1, D_1) . Indeed, since D_1 is an interval truncated from above, D_2 is an upward shift of D_1 .²⁹ As a result, Lemma 4 implies that if D_2 is replaced by D_1 , the first information effect increases and the loss of control effect decreases. Moreover, since $\rho_{2,1}$ is decreasing and since type 1's incentive constraint is binding, part (b) of Lemma 5 implies that also the second information effect goes up. Hence, all effects work in the direction of static delegation.

In other words, if the virtual likelihood is decreasing, the benefits from offering type 2 an upward shift D_2 of D_1 are smaller than the associated costs of providing incentives for type 1 not to select D_2 . Intuitively, conditional on facing type 2, the principal considers large states less (virtually) likely than type 1 and small states more (virtually) likely than type 1. Thus, in expected terms, the increased discretion over large actions, which boosts the agent's utility $u(\omega)$ in large states, benefits the principal less than type 1, and the reduced discretion over small actions, which depresses the agent's utility $u(\omega)$ in small states, hurts the principal more than type 1. Therefore, the principal's benefits increasing type 2's discretion over large actions are outweighed by the associated reduction of discretion over small actions which is necessary to achieve incentive compatibility.

To account for the case with more than two types, our actual proof of Proposition 1 in the appendix is based on the Kuhn-Tucker theorem. Our proof strategy is to (i) identify a relaxed problem so that (ii) the corresponding Lagrangian is maximized by the static delegation menu. Both steps are not straightforward and intertwined.

To illustrate, suppose that the only incentive constraints involving type i in the relaxed problem are $IC_{i,j}$ and $IC_{k,i}$, $j \neq i$, $k \neq i$, and denote the multipliers by $\lambda_{i,j}$ and $\lambda_{k,i}$. Due to the additive structure, the Lagrangian can then be written as

$$\mathcal{L} = \mu_i V_i(D_i) + \lambda_{i,j} [U_i(D_i) - U_i(D_j)] - \lambda_{k,i} [U_k(D_k) - U_k(D_i)] + C_1, \quad (18)$$

where C_1 does not depend on D_i (but still on D_m , $m \neq i$). Now recall that $U_i(D) = \int_{\underline{\omega}}^{\bar{\omega}} u_D(\omega) f_i(\omega) d\omega$ for each type i . Together with (8), this delivers that

$$\begin{aligned} \mathcal{L} = & \mu_i u_{D_i}(\underline{\omega}) b(\underline{\omega}) f_i(\underline{\omega}) - \mu_i u_{D_i}(\bar{\omega}) b(\bar{\omega}) f_i(\bar{\omega}) + \\ & + \int_{\underline{\omega}}^{\bar{\omega}} u_{D_i}(\omega) \{ \mu_i [f_i(\omega) + (b(\omega) f_i(\omega))'] + \lambda_{i,j} f_i(\omega) - \lambda_{k,i} f_k(\omega) \} d\omega + C_2, \end{aligned} \quad (19)$$

where C_2 does not depend on D_i . Observe that this expression essentially looks like (8), yet with a different, the *Lagrangian virtual likelihood* $\mu_i [f_i(\omega) + (b(\omega) f_i(\omega))'] + \lambda_{i,j} f_i(\omega) - \lambda_{k,i} f_k(\omega)$, which now depends on the multipliers. The issue stated in point (ii) above is when expression (19) is maximized by an interval delegation set, and in particular, by

²⁹See footnote 28 for a detailed argument.

D^{st} . We stress that the sufficient conditions for interval delegation to be optimal identified by Amador and Bagwell [3] (see footnote 22) are not directly applicable here, since the objective (19) is not of the form (8), as it contains additional terms with the multipliers. We therefore extend the sufficient conditions in [3] to a domain of objective functions that (potentially) includes objectives of the form (19) (see Lemma A.2 in the Appendix for details). That is, if the Lagrangian virtual likelihood satisfies these conditions, an interval maximizes (19), and, in addition, if the multipliers have the “right” magnitude, this interval is D^{st} .

Since the Lagrangian virtual likelihood depends on the multipliers, whether it satisfies our sufficient conditions depends on the set of constraints we select for the relaxed problem, which is the issue under point (i) above. Interestingly, considering the familiar relaxed problem that imposes the adjacent constraints $IC_{1,2}, \dots, IC_{I-1,I}$ does not work. Instead, we adopt an idea developed in Krähmer and Strausz [26] and consider the (non-adjacent) constraints that no type ℓ with $\omega_\ell^0 < \omega^{st}$ has an incentive to mimic any type h with $\omega_h^0 > \omega^{st}$.³⁰ Under the assumption that $\rho_{i,j}$ is decreasing, it can then be shown that the Lagrangian virtual likelihoods induced by these constraints satisfy our sufficient conditions (from Lemma A.2). In fact, as the proof reveals, it is sufficient to require that $\rho_{h,\ell}$ is decreasing only for such types ℓ and h .

Example (ctd). For the specification (12), as remarked above, $\rho_{2,1}$ is decreasing if and only if $3A\kappa_3 < \kappa_2$ and $\gamma \geq \kappa_2/(2\kappa_2 - 3A\kappa_3)$. In this case, static delegation is optimal by Proposition 1. As a numerical example, consider $A = \frac{7}{5}$, $\kappa_3 = 0$, and $\kappa_2 = \frac{19}{20}$. Then these conditions, together with Assumptions A1–A3 are satisfied, if and only if $\gamma \in [\frac{1}{2}, \frac{9}{14}]$. \square

An interesting question is what happens when the parties’ conflict of interest as measured by the bias is small. Proposition 1 may not be applicable in this case, because virtual likelihood ratios may not be decreasing. In fact, for small $b(\omega)$ the virtual likelihood ratio $\rho_{i,j}$ is approximately equal to the likelihood ratio f_i/f_j . Hence, increasing likelihood ratios—which is a sufficient condition for our types ordering (10) (see Remark 3)—are inconsistent with decreasing virtual likelihood ratios for small bias. As we now show, however, static delegation is always optimal for small bias, irrespective of the shape of the virtual likelihood ratio.

Proposition 2. *Consider a bias $b(\omega)$ so that Assumptions A1–A3 are satisfied. Then static delegation is optimal when the bias is equal to $\alpha b(\omega)$ and $\alpha > 0$ is sufficiently small.*

If the bias is small, then the optimal delegation set with public types is close to “full” delegation for *all* types ($D_i = [x_A(\underline{\omega}), x_A(\bar{\omega})]$). Intuitively, there is therefore little benefit

³⁰[26] use this idea to show a result with a similar flavour, that is, in a unit good sequential screening problem with money, a “static” contract is optimal when the agent has an ex post outside option.

in screening types. To illustrate the underlying issue, consider, in the two types case, a deviation from the optimal static menu and suppose the principal offers D^{st} to type 1 and an upward shift \tilde{D}_2 of D^{st} to type 2 so that type 1 is kept indifferent. By Lemma 4, the difference between the first information effects is then negative, $J_2^I(\tilde{D}_2) - J_2^I(D_2^{st}) \leq 0$, and the difference between the loss of control effects is positive, $LC_2(\tilde{D}_2) - LC_2(D_2^{st}) \geq 0$. While both effects thus work in favor of static delegation, it is evident from (8) that both differences converge to zero, as the bias gets small. Moreover, the difference between the second information effects, $J_2^{II}(\tilde{D}_2) - J_2^{II}(D_2^{st})$, also converges to zero. This is so, because as the bias gets small, D^{st} converges to “full” delegation, that is, to the set $[x_A(\underline{\omega}), x_A(\bar{\omega})]$. As a consequence, the benefits from permitting agent type 2 to choose actions which are larger than those offered by D^{st} become smaller and smaller since the likelihood of the set of (large) states in which the increased discretion could be utilized goes to zero. Our proof shows that, as the bias gets small, the difference between the second information effects converges faster than the differences between the first information and the loss of control effects. This implies that, as the bias gets small, static delegation is optimal.

We point out that Proposition 2 critically rests on the fact that the optimal delegation sets with public types converge to the same set for small bias, and is not universally true for all sequential delegation settings. In our working paper (Kováč and Krähmer, [22]), we explore a setting with two types where the support of F_1 is a proper subset of the support of F_2 and show that even for small bias, sequential delegation may be optimal.

5 Sequential delegation

We now ask whether sequential delegation can be optimal. We establish a necessary condition and a sufficient condition for the case with two types $i = 1, 2$. As we argue later, insofar as only the question whether sequential delegation can be optimal at all is concerned, these conditions can, *mutatis mutandis*, also be applied for more types. For two types, we will then also identify properties of the shape of an optimal sequential delegation menu. For reasons outlined below, this is much less tractable with more types.

We begin by deriving an intuitive sufficient condition for the optimality of sequential delegation. To describe the basic idea, consider the optimal static delegation menu with $D_1 = D_2 = [x_A(\underline{\omega}), x_A(\omega^{st})]$. If types were public, then since $\omega^{st} < \omega_2^0$, the principal would benefit from extending the upper end of D_2 in order to be closer to the optimal delegation set for type 2. Since types are private, however, if the principal wants to extend the upper end of D_2 , she must remove some other actions in D_2 so as to deter type 1 from picking D_2 . Removing actions makes it more costly for the agent to pick D_2 , and so if actions are removed whose removal is very costly for type 1 in expectation, there is more scope for extending the upper end of D_2 without violating incentive compatibility.

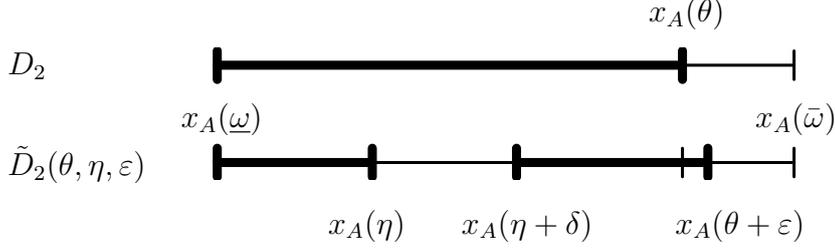


Figure 1: Marginal sequential delegation.

However, removing actions from D_2 is not only costly for type 1, but by (8) also for the principal herself since it reduces the information effect. Thus, when removing an action, the principal has to put her own and type 1's costs in proportion. Intuitively, the principal benefits from replacing a set of actions within D_2 by a set of actions at the upper end of D_2 if her relative gain from extending the interval, relative to agent type 1's gain, is larger than her relative costs from removing actions, relative to agent type 1's costs.

We now make these considerations precise. Take a static delegation menu with $D_1 = D_2 = [x_A(\underline{\omega}), x_A(\theta)]$ for some $\theta < \bar{\omega}$ (for example, think of $\theta = \omega^{st}$ from Lemma 2). We create a sequential delegation menu by leaving D_1 unchanged, and modifying D_2 by slightly extending its upper end and including a small gap starting at the action $x_A(\eta)$, where $\eta \in (\underline{\omega}, \theta]$, within D_2 so that type 1 remains indifferent. Formally, let

$$D_2(\theta, \eta, \varepsilon) = [x_A(\underline{\omega}), x_A(\eta)] \cup [x_A(\eta + \delta), x_A(\theta + \varepsilon)], \quad (20)$$

where $\varepsilon > 0$ and $\delta = \delta(\varepsilon) > 0$ is the unique value such that $U_1(D_1) = U_1(D_2(\theta, \eta, \varepsilon))$. We refer to this modification as *marginal sequential delegation* (see Figure 1). A calculation (shown in the appendix) yields that the effect of marginal sequential delegation on the principal's expected utility, conditional on type 2, is equal to

$$\frac{\partial V_2(D_2(\theta, \eta, \varepsilon))}{\partial \varepsilon} \Big|_{\varepsilon=0} = \left[-\rho_{2,1}(\eta) \int_{\theta}^{\bar{\omega}} 1 - F_1(\omega) d\omega + \int_{\theta}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega) f_2(\omega) d\omega \right] x'_A(\theta). \quad (21)$$

Therefore, marginal sequential delegation is profitable for the principal if

$$R_{2,1}(\omega) \equiv \frac{\int_{\theta}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega) f_2(\omega) d\omega}{\int_{\theta}^{\bar{\omega}} 1 - F_1(\omega) d\omega} > \rho_{2,1}(\eta). \quad (22)$$

This inequality captures precisely the principal's trade-off. The expected gain from extending the upper end of D_2 by ε can be computed to be $\varepsilon \cdot \int_{\theta}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega) f_2(\omega) d\omega$ for the principal and $\varepsilon \cdot \int_{\theta}^{\bar{\omega}} 1 - F_1(\omega) d\omega$ for agent type 1. Thus, the left hand side of (22) represents the principal's relative gain from extending the interval, relative to agent type 1's gain. Moreover, it turns out that the expected cost of including a δ -gap around

$x_A(\eta)$ is $\delta^3 \cdot [f_2(\eta) + (b(\eta)f_2(\eta))']$ for the principal and $\delta^3 \cdot f_1(\eta)$ for agent type 1. Thus, the right hand side of (22) represents the principal's relative costs from removing actions from D_2 , relative to agent type 1's costs.

By construction, marginal sequential delegation is incentive compatible for type 1. So far, we have (silently) ignored the question whether it is also incentive compatible for type 2. Observe that the set $D_2(\theta, \eta, \varepsilon)$ is an upward shift of the interval $D_1 = [x_A(\underline{\omega}), x_A(\theta)]$. We now impose that f_2/f_1 be increasing. By part (a) of Lemma 3, this ensures that type 2's incentive constraint is automatically satisfied under marginal sequential delegation.

Lemma 6. *If f_2/f_1 is increasing, marginal sequential delegation is incentive compatible.*

We can now state our characterization result for two types which essentially says that sequential delegation is optimal if and only if marginal sequential delegation is optimal. Recall the definition of $R_{2,1}$ in (22).

Proposition 3. (a) *If f_2/f_1 is increasing, then*

$$R_{2,1}(\omega^{st}) > \min_{\eta \leq \omega^{st}} \rho_{2,1}(\eta) \quad (23)$$

is sufficient for sequential delegation to be optimal.

(b) *If $R_{2,1}$ is decreasing, then (23) is necessary for sequential delegation to be optimal.*

Part (a) follows straightforwardly from our previous considerations. If marginal sequential delegation is feasible and profitable, then some sequential delegation menu is optimal for problem \mathcal{P} .

Part (b) is more interesting. Clearly, if (23) is violated, then *marginal* sequential delegation is not profitable. In other words, static delegation is *locally* optimal. In principle, this still leaves room for a *global* modification to be profitable. The condition that $R_{2,1}$ is decreasing ensures that this is not the case, and plays the role of a second order condition. Because the principal's expected utility is, in general, not concave, an additional condition is needed to ensure that a local is also a global optimum.

To prove part (b), we show that static delegation is optimal if condition (23) is violated. As with Proposition 1, we use the Kuhn-Tucker theorem to establish this. The result does not directly follow from Proposition 1, as the conditions that $R_{2,1}$ is decreasing and the reverse of (23) are weaker than the requirement in Proposition 1 that $\rho_{2,1}$ be decreasing. For the two types case, these weaker conditions are sufficient, however.³¹

Condition (23) is somewhat tedious to check since ω^{st} is only implicitly given. As a corollary of Proposition 3, we now provide conditions that do not depend on ω^{st} .

³¹Note that for the two types case, Proposition 1 follows from part (b) of Proposition 3 by Lemma A.4 in the appendix.

Lemma 7. *Let f_2/f_1 be increasing, $R_{2,1}$ be decreasing, and $\rho_{2,1}$ be increasing. If $R_{2,1}(\omega_1^0) > \rho_{2,1}(\underline{\omega})$, then there is a critical $\hat{\mu}_2 \in (0, 1]$ so that sequential delegation is optimal if and only if $\mu_2 \leq \hat{\mu}_2$.*

To see the result, note that under the conditions of the lemma, sequential delegation is optimal if and only if $R_{2,1}(\omega^{st}) > \rho_{2,1}(\underline{\omega})$ by Proposition 3. In the appendix, we show that ω^{st} as defined in (2) and understood as a function of μ_2 increases monotonically from ω_1^0 to ω_2^0 as μ_2 goes from 0 to 1. Hence, as $R_{2,1}$ is decreasing, $R_{2,1}(\omega_1^0) > \rho_{2,1}(\underline{\omega})$ implies that $R_{2,1}(\omega^{st}) > \rho_{2,1}(\underline{\omega})$ if and only if μ_2 is below a critical cut-off.

One may wonder whether there are environments in which all the conditions in Lemma 7 can be jointly satisfied. Our example shows that this is indeed the case.

Example (ctd). Consider the specification (12), with $A = \frac{7}{5}$, $\kappa_3 = \frac{19}{20}$, and $\kappa_2 = 0$. A straightforward computation reveals that Assumptions A1–A3 are satisfied if and only if $\gamma < \frac{5}{9}$. Moreover, f_2/f_1 is increasing, and $\rho_{2,1}$ is increasing (see (17)). Furthermore, $R_{2,1}$ is decreasing on $[0, 1]$ if and only if $\gamma \geq \frac{95}{251}$. Finally, it can be verified that $R_{2,1}(\omega_1^0) > \rho_{2,1}(\underline{\omega})$ for all $\gamma \in [\frac{95}{251}, \frac{5}{9})$. In this case, all assumptions of Lemma 7 are satisfied. \square

Proposition 3 provides conditions when sequential delegation is optimal with two types. We now study how a delegation menu looks like in this case. We begin with type 1's delegation set. As argued after Proposition 1, at the solution to the relaxed problem where we disregard type 2's incentive constraint, type 1's incentive constraint is binding (i.e., $U_1(D_1) = U_1(D_2)$) and D_1 is an interval, truncated from above. This implies that type 2's incentive constraint is automatically satisfied if f_2/f_1 is increasing. The reason is that since D_1 is an interval truncated from above, D_2 must be an upward shift of D_1 and therefore it follows from part (a) of Lemma 3 that $U_2(D_2) \geq U_2(D_1)$.

Lemma 8. *Suppose f_2/f_1 is increasing. There is a solution to \mathcal{P} so that type 1's incentive constraint is binding, and D_1^* is an interval:*

$$U_1(D_1^*) = U_1(D_2^*), \quad \text{and} \quad D_1^* = [x_A(\underline{\omega}), x_A(\omega_1^*)], \quad \text{where } \omega_1^* \in [\omega_1^0, \bar{\omega}]. \quad (24)$$

Next, we aim to narrow down the possible shapes of type 2's delegation set D_2^* at an optimum. Condition (22) suggests that D_2^* can, in general, look rather complicated. Intuitively, suppose that the upper endpoint of D_2^* is the agent's optimal action in state θ , $x_A(\theta)$. Recall that if (22) holds, it is beneficial to marginally extend the upper end of D_2 and insert a marginal gap around the action $x_A(\eta)$. Reversely, if it does not hold, it is beneficial to lower the upper end of D_2 and include action $x_A(\eta)$. Hence, at an optimum, D_2^* contains all actions $x_A(\eta)$ for which (22) is violated and no actions $x_A(\eta)$ for which (22) holds. Therefore, absent any structure on the shape of the right and left hand side in (22) it looks rather daunting to pin down D_2^* .

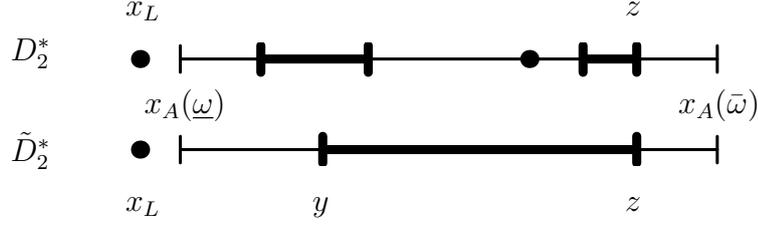


Figure 2: Example of construction of \tilde{D}_2^* .

Yet, we can narrow down the possible shapes of type 2's delegation set if the benefits ratio $R_{2,1}$ is decreasing, and the cost ratio $\rho_{2,1}$ is increasing. The former says that the marginal benefit of extending the upper end of D_2 becomes smaller the larger the upper end. The latter says that the marginal cost of inserting a gap around an action becomes smaller the smaller the action. Intuitively, the upper end of D_2 should thus be gradually extended and, starting at the bottom of D_2 , ever larger actions should be removed until the marginal benefits equal the marginal costs. The next result makes this precise.

Proposition 4. *Let f_2/f_1 be increasing, $R_{2,1}$ be decreasing, and $\rho_{2,1}$ be increasing. Then there is a solution to \mathcal{P} which satisfies (24) and such that there are actions $x_L \leq x_A(\underline{\omega})$, $y, z \in [x_A(\underline{\omega}), x_A(\bar{\omega})]$, $y \leq z$ so that*

- (a) $D_2^* = [y, z]$, or
- (b) $D_2^* = \{x_L\} \cup [y, z]$.

The basic idea to show the lemma is that for any D_2^* we can use Lemma 5(a) to construct a profitable modification of D_2^* of the form (a) or (b). For example, suppose D_2^* has a smallest action $x_L \leq x_A(\underline{\omega})$ and a largest action $z \leq x_A(\bar{\omega})$ but is not of the form (b) (see Figure 2). Then, because $U_1(D_1^*) = U_1(D_2^*)$ by Lemma 8, an intermediate value argument implies that there is a modification \tilde{D}_2^* of the first type in (b) so that type 1 stays indifferent between D_1^* and \tilde{D}_2^* . By construction, the first condition of Lemma 5(a) is met, and \tilde{D}_2^* is an upward shift of D_2^* (all “gaps” have been “merged” and moved to the bottom). Thus, since $\rho_{2,1}$ is increasing by assumption, Lemma 5(a) implies that the principal prefers \tilde{D}_2^* to D_2^* .

Thus, if sequential delegation is optimal, the optimal delegation set for type 2 is a (possibly degenerate) interval, truncated from below and above, as in (a), or it contains a single gap “at the bottom” as in (b). Moreover, since $IC_{1,2}$ is binding (Lemma 8), the largest action in D_2^* is larger than that in D_1^* , i.e., $x_A(\omega_1^*) < z$.

Under the condition that f_2/f_1 is increasing, $R_{2,1}$ is decreasing and that $\rho_{2,1}$ is increasing, Lemma 7 and Proposition 4 give a fairly complete qualitative picture of the solution to \mathcal{P} . In principle, one could now precisely pin down an optimal sequential delegation menu by optimizing over the possible forms (a) and (b) in Proposition 4.

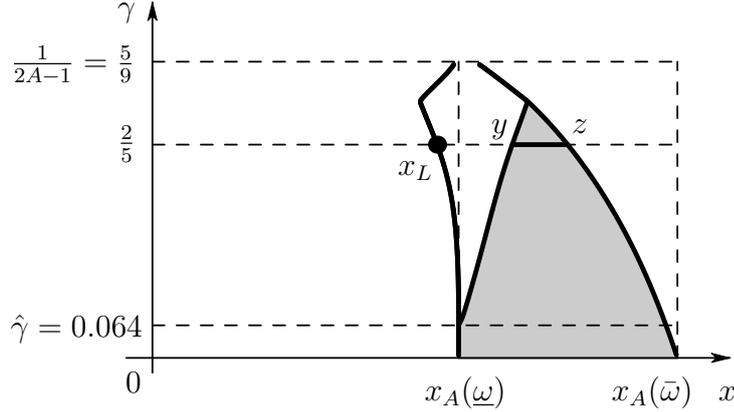


Figure 3: Type 2's optimal delegation set ($A = \frac{7}{5}$, $\kappa_3 = \frac{19}{20}$, $\kappa_2 = 0$, $\mu_2 = \frac{1}{20}$).

Example (ctd). Figure 3 illustrates type's 2 delegation set D_2^* in the optimal delegation menu (obtained numerically) for the regulation example with distributions given by (12). Given A , κ_3 , κ_2 , and μ_2 , consider some feasible γ on the vertical axis.³² Draw the corresponding horizontal line. D_2^* is then the intersection of that line with the plotted region (e.g., for $\gamma = \frac{2}{5}$, we have $D_2^* = \{x_L\} \cup [y, z] = \{0.650\} \cup [0.822, 0.949]$). The figure indicates that static delegation is optimal for sufficiently low values of γ ($\gamma \leq 0.064$), consistent with Proposition 2. Otherwise, sequential delegation is optimal. \square

We conclude this section with remarks on the robustness of our analysis.

Number of types

Our result on optimal sequential delegation for two types characterizes both, when sequential delegation is optimal (Proposition 3) and how an optimal delegation menu looks like (Proposition 4). To prove this, we have exploited the fact that when we ignore $IC_{2,1}$, then D_1 is an interval, truncated from above, and $IC_{1,2}$ is binding. This means that type 2's delegation set D_2 is necessarily an upward shift of D_1 . In other words, the solution (D_1, D_2) to the relaxed problem where $IC_{2,1}$ is ignored is monotone in the upward shift order. Under the single-crossing condition that f_2/f_1 is increasing, this implies that $IC_{2,1}$ is automatically satisfied by Lemma 3.

With more than two, say with three types, this logic fails. Even if we knew that at the relaxed problem where only $IC_{1,2}$ and $IC_{2,3}$ are imposed, all constraints are binding, then D_2 will typically not be an interval, and hence nothing guarantees that D_3 is an upward shift of D_2 . In other words, the difficulty with more than two types is to show that the optimal delegation sets for an appropriately relaxed problem are upward shifts of one another and would also solve the original problem.³³

³²Recall from the example above that Assumptions A1–A3 are satisfied if and only if $\gamma < 1/(2A-1) = \frac{5}{9} \approx 0.556$ (upper bound for γ in the figure) and that the assumptions of Lemma 7 as well as Proposition 4 hold for $\gamma \geq \frac{95}{251} \approx 0.378$.

³³We conjecture, however, that even in the case with more than two types, delegations sets will

On the other hand, it is comparatively easy to derive sufficient conditions for the optimality of sequential delegation along the lines of Proposition 3. For this, one only needs to identify a profitable modification of the optimal static menu. To sketch the idea, fix a type k and split the set of types in a group $i \leq k$ and a group $j > k$. Consider the marginal modification where all types $i \leq k$ receive the optimal static delegation set D^{st} and all types $j > k$ receive $D_2(\omega^{st}, \eta, \varepsilon)$ as defined in (20) where now $U_k(D^{st}) = U_k(D_2(\omega^{st}, \eta, \varepsilon))$.

The marginal modification is then profitable for the principal if the appropriately adjusted analogue to condition (23) holds.³⁴ The modification is incentive compatible if $U_i(D^{st}) \geq U_i(D_2(\omega^{st}, \eta, \varepsilon))$ for all $i \leq k$, and $U_j(D_2(\omega^{st}, \eta, \varepsilon)) \geq U_j(D^{st})$ for all $j > k$. Since $D_2(\omega^{st}, \varepsilon)$ is an upward shift of D^{st} , Lemma 3 implies that this is true if f_i/f_k is decreasing for all $i \leq k$, and f_j/f_k is increasing for all $j > k$.

Non-monotone virtual likelihood ratios

While our results in the two types case are most clear-cut when the virtual likelihood ratio $\rho_{2,1}$ is monotone, our analysis suggests insights also for the case when $\rho_{2,1}$ is not monotone. If, for example, $\rho_{2,1}$ is U-shaped with a minimum at $\hat{\eta} \in (\underline{\omega}, \bar{\omega})$, then the logic outlined before the statement of Proposition 4 can be adapted to show that D_2 has a gap around $x_A(\hat{\eta})$, that is, “in the middle” (rather than “at the bottom”). The reason is that if the principal has to remove some actions from D_2 under the constraint that type 1’s incentive constraint be binding, then it is most cost-effective to insert a gap around $x_A(\hat{\eta})$ where $\rho_{2,1}$ is smallest.

In our working paper (Kováč and Krähmer, [22]), we analyze the situation in which the agent receives an imperfect signal about the state with type 2 receiving a more precise signal than type 1, implying that F_2 is more spread out than F_1 . Virtual likelihood ratios are then not monotone. In fact, $\rho_{2,1}$ drops sharply at one point and increases sharply at some other larger point. Consistent with the considerations in the previous paragraph, we show in the working paper that if sequential delegation is optimal, type 2’s delegation set will be the disjoint union of an interval and a single action.³⁵

qualitatively look like in Proposition 4 as long as virtual likelihood ratios are increasing. This is based on the logic outlined before the statement of Proposition 4 which suggests that the optimal way to prevent low types from mimicking high types is to substitute small for large actions in the high types delegation set.

³⁴Because the modification affects the principal’s utility from all types $j > k$, instead of F_2 , one has to work with the conditional distribution of the state, conditional on $j > k$.

³⁵More precisely, with replacement noise, the induced distributions F_1 and F_2 , unlike in the current setup, do not have the same support, with the support of F_1 included in the support of F_2 . F_1 can, however, be approximated by a sequence of smooth distributions F_1^n that have the same support as F_2 . For n large, the virtual likelihood ratio drops sharply at the lower end of the support of F_1 and increases sharply at the upper end of the support of F_1 . Another difference to the current setup is that for small bias, Assumptions A1–A3 are violated for the average distribution F^{st} which implies that the optimal static delegation set may not be an interval. For this case, it turns out that sequential delegation is always optimal.

6 Type-dependent bias

In this section, we return to our comment in footnote 16 and consider the case that the bias function depends on the agent's type i . We first note that at no point in the analysis, we used that the bias function did not depend on i , and thus all our results remain true as long as Assumptions A1–A3 hold when $b(\omega)$ is replaced by $b_i(\omega)$.

An interesting special case arises when the agent has only private information about the bias, but the distribution of the state is common knowledge: $F_i = F$. The question is then whether it is optimal to elicit the bias from the agent. To explore this question, we consider the case with constant bias $b_i > 0$.³⁶ In independent work, Tanner [43] analyzes this case for the uniform distribution and for a subclass of the preferences (1) and (2), and finds that the principal does not elicit the bias. Our results, which are based on arguments very different from Tanner's, allow us to readily extend this finding to any distribution with log-concave density. Moreover, we argue that for the two types case, eliciting the bias is suboptimal if the density has decreasing mean residual life.

Lemma 9. *Consider the case with constant, type-dependent bias and type-independent distribution. Assume that (a) f is log-concave, or (b) there are two types ($i = 1, 2$) and F has decreasing mean residual life. Then it is optimal not to elicit the agent's bias.*

To understand (a), observe that the virtual likelihood ratio now becomes $\rho_{i,j}(\omega) = 1 + b_i \cdot f'(\omega)/f(\omega)$, which is decreasing if and only if f'/f is decreasing, which means that f is log-concave. Thus, part (a) is an immediate consequence of Proposition 1.

Part (b) of Lemma 9 is a consequence of Proposition 3(b). To see this, observe first that $f_2/f_1 = f/f = 1$ is now always (weakly) increasing. Furthermore,

$$R_{2,1}(\omega) = 1 - b_2 \frac{1 - F(\omega)}{\int_{\omega}^{\bar{\omega}} 1 - F(\tilde{\omega}) d\tilde{\omega}} \text{ is decreasing} \quad (25)$$

if and only if the mean residual life (see (13)) is decreasing. Finally, we verify in the appendix that the necessary condition (23) is violated with type-independent distribution and constant bias. Hence, Proposition 3(b) implies that eliciting the bias is not optimal.

Both, log-concave density and decreasing mean residual life are fairly mild requirements that are satisfied by a large class of distributions. Thus, together with Proposition 3, Lemma 9 makes clear that what matters for sequential delegation to be optimal is primarily the agent's private information about the distribution, not about the bias.

³⁶Assumptions A1 and A3 impose upper bounds on the feasible set of biases: $b_i < E\omega - \underline{\omega}$ and $f'(\omega)/f(\omega) > -1/b_i$ for all ω .

7 Conclusion

In this paper, we have provided insights into the nature of optimal delegation when the agent's private information arrives sequentially. Our results demonstrate the robustness, but also the limits, of interval delegation in this dynamic setting. While interval delegation remains optimal in a large class of environments, the agent's discretion may be restricted in richer ways than simply imposing minimum or maximum thresholds on his actions. In this sense, the principal's desire to elicit the agent's type provides a novel rationale for restricting the discretion of a privately informed agent, next to restraining agents from pursuing partisan interests or to motivate them to acquire costly information.

The Lagrangian techniques developed in this paper are likely to be useful in other optimal delegation settings whenever additional constraints on the agent's utility must be met. Examples are settings where the agent has to be supplied with a given level of expected utility. This may arise when there are participation constraints, but also in a dynamic setting where each period a new state is drawn, and eliciting the state in the current period requires to promise the agent a certain continuation value.

A Appendix

Preliminaries

Proof of formula (8). By standard arguments, $x_D(\omega)$ is increasing and satisfies the envelope condition³⁷

$$u_D(\omega_2) - u_D(\omega_1) = \int_{\omega_1}^{\omega_2} x_D(\omega) d\omega. \quad (26)$$

By (1) and (2), and using integration by parts, we obtain:

$$\begin{aligned} V_i(D) - U_i(D) &= \int_{\underline{\omega}}^{\bar{\omega}} [-b(\omega)x_D(\omega) + c(\omega)]f_i(\omega) d\omega \\ &= -b(\bar{\omega})f_i(\bar{\omega})u_D(\bar{\omega}) + b(\underline{\omega})f_i(\underline{\omega})u_D(\underline{\omega}) + \int_{\underline{\omega}}^{\bar{\omega}} (b(\omega)f_i(\omega))'u_D(\omega)d\omega + C, \end{aligned} \quad (27)$$

where $C = \int_{\underline{\omega}}^{\bar{\omega}} c(\omega)f_i(\omega) d\omega$ does not depend on D . This implies (8). \square

Lemma A.1. *Let*

$$\Gamma_i(\omega) \equiv \int_{\omega}^{\bar{\omega}} 1 - F_i(\tilde{\omega}) - b(\tilde{\omega})f_i(\tilde{\omega}) d\tilde{\omega}. \quad (28)$$

(a) *The equation $\Gamma_i(\omega) = 0$ has a unique solution in the interval $(\underline{\omega}, \bar{\omega})$, denoted ω_i^0 . Moreover, $\Gamma_i(\omega) > 0$ for all $\omega \in [\underline{\omega}, \omega_i^0)$, $\Gamma_i(\omega) < 0$ for all $\omega \in (\omega_i^0, \bar{\omega})$, and $\Gamma_i(\bar{\omega}) = 0$.*

³⁷The formal argument is identical to the argument in Kováč and Mylovanov [21] for quadratic preferences.

(b) Let D be a delegation set with maximal action $x_A(\hat{\theta}) = \max D$, $\hat{\theta} < \bar{\omega}$. In addition, for $\hat{\theta} \leq \theta$, let $D(\hat{\theta}, \theta) = D \cup [x_A(\hat{\theta}), x_A(\theta)]$. Then

$$\frac{\partial V_i(D(\hat{\theta}, \theta))}{\partial \theta} = \Gamma_i(\theta)x'_A(\theta). \quad (29)$$

Proof of Lemma A.1. (a) Γ_i is strictly convex, since $\Gamma_i''(\omega) = f_i(\omega) + (b(\omega)f_i(\omega))' > 0$ due to A1. Moreover, $\Gamma_i(\underline{\omega}) > 0$ by A3 and clearly $\Gamma_i(\bar{\omega}) = 0$, which together with strict convexity implies that Γ_i can attain the value 0 at most once in the interval $(\underline{\omega}, \bar{\omega})$. Because $\Gamma_i'(\bar{\omega}) = b(\bar{\omega})f_i(\bar{\omega}) > 0$, we have $\Gamma_i(\omega) < 0$ for ω sufficiently close to $\bar{\omega}$. The existence of a solution then follows from continuity. The remaining inequalities follow from the convexity of Γ_i and $\Gamma_i(\omega_i^0) = 0$.

(b) By (8),

$$\begin{aligned} V_i(D(\hat{\theta}, \theta)) &= c + \int_{\hat{\theta}}^{\theta} U(\omega, x_A(\omega))[f_i(\omega) + (b(\omega)f_i(\omega))'] d\omega \\ &\quad + \int_{\theta}^{\bar{\omega}} U(\omega, x_A(\theta))[f_i(\omega) + (b(\omega)f_i(\omega))'] d\omega - U(\bar{\omega}, x_A(\theta))b(\bar{\omega})f_i(\bar{\omega}), \end{aligned} \quad (30)$$

where c is a constant that does not depend on θ . By Leibniz' rule,

$$\begin{aligned} \frac{\partial V_i(D(\hat{\theta}, \theta))}{\partial \theta} &= U(\theta, x_A(\theta))[f_i(\theta) + (b(\theta)f_i(\theta))'] - U(\theta, x_A(\theta))[f_i(\theta) + (b(\theta)f_i(\theta))'] \\ &\quad + \int_{\theta}^{\bar{\omega}} \frac{\partial}{\partial \theta} U(\omega, x_A(\theta))[f_i(\omega) + (b(\omega)f_i(\omega))'] d\omega - b(\bar{\omega})f_i(\bar{\omega}) \frac{\partial}{\partial \theta} U(\bar{\omega}, x_A(\theta)) \\ &= \int_{\theta}^{\bar{\omega}} [\omega + a'(x_A(\theta))]x'_A(\theta)[f_i(\omega) + (b(\omega)f_i(\omega))'] d\omega \\ &\quad - b(\bar{\omega})f_i(\bar{\omega})[\bar{\omega} + a'(x_A(\theta))]x'_A(\theta) \\ &= \left[\int_{\theta}^{\bar{\omega}} (\omega - \theta)[f_i(\omega) + (b(\omega)f_i(\omega))'] d\omega - b(\bar{\omega})f_i(\bar{\omega})(\bar{\omega} - \theta) \right] x'_A(\theta), \end{aligned} \quad (31)$$

where the third equality follows from the agent's first-order condition, $\theta + a'(x_A(\theta)) = 0$. The desired equality (29) follows now from integration by parts. \square

Proofs for Section 2

The proof of Lemma 1 is based on the following lemma.

Lemma A.2. Let $g : [\underline{\omega}, \bar{\omega}] \rightarrow \mathbb{R}$ be a differentiable function, $\underline{g} \in \mathbb{R}$ be a constant, and let $\omega^0 \in [\underline{\omega}, \bar{\omega}]$. Assume that the following conditions hold:

- (c1) $g'(\omega) \leq 0$ for all $\omega \in [\underline{\omega}, \omega^0]$.
- (c2) $\int_{\omega}^{\bar{\omega}} g(\bar{\omega}) d\bar{\omega} \leq 0$ for all $\omega \in [\omega^0, \bar{\omega}]$, with equality at ω^0 .
- (c2') $\underline{g} \geq 0$.

Then the delegation set $D^0 = [x_A(\underline{\omega}), x_A(\omega^0)]$ solves the maximization problem

$$\max_D u_D(\underline{\omega})\underline{g} + u_D(\bar{\omega})g(\bar{\omega}) + \int_{\underline{\omega}}^{\bar{\omega}} u_D(\omega)(-g'(\omega)) d\omega. \quad (32)$$

Proof of Lemma A.2. The envelope condition (26) and integration by parts imply

$$u_D(\bar{\omega})g(\bar{\omega}) + \int_{\omega^0}^{\bar{\omega}} u_D(\omega)(-g'(\omega)) d\omega = u_D(\omega^0)g(\omega^0) + \int_{\omega^0}^{\bar{\omega}} x_D(\omega)g(\omega) d\omega. \quad (33)$$

The objective function (32) can therefore be written as

$$u_D(\underline{\omega})\underline{g} + u_D(\omega^0)g(\omega^0) + \int_{\underline{\omega}}^{\omega^0} u_D(\omega)(-g'(\omega)) d\omega + \int_{\omega^0}^{\bar{\omega}} x_D(\omega)g(\omega) d\omega. \quad (34)$$

The first three terms in (34) are linear in u_D . We now show that all coefficients at u_D are non-negative. The coefficient in the first term, \underline{g} , is non-negative by (c2'). Moreover, all coefficients in the third term, $-g'(\omega)$, are non-negative by (c1). Finally, as to the second term, observe that $g(\omega^0)$ is equal to $-\bar{G}'(\omega^0)$, where $\bar{G}(\omega) = \int_{\omega}^{\bar{\omega}} g(\tilde{\omega}) d\tilde{\omega}$. By (c2), $\bar{G}(\omega^0) = 0$, while $\bar{G}(\omega) \leq 0$ for all $\omega > \omega^0$. Thus, $\bar{G}'(\omega^0) \leq 0$.

As the next step, we show that the fourth term in (34) is bounded from above by 0. Using integration by parts for the *Riemann-Stieltjes integral* we obtain³⁸

$$\int_{\omega^0}^{\bar{\omega}} x_D(\omega)g(\omega) d\omega = -x_D(\bar{\omega})\bar{G}(\bar{\omega}) + x_D(\omega^0)\bar{G}(\omega^0) + \int_{\omega^0}^{\bar{\omega}} \bar{G}(\omega) dx_D(\omega). \quad (35)$$

Now the first and the second term on the right hand side are equal to zero by definition of \bar{G} and due to (c2), respectively. Moreover, the third term is non-positive, as x_D is increasing and due to (c2). Thus, indeed $\int_{\omega^0}^{\bar{\omega}} x_D(\omega)g(\omega) d\omega \leq 0$.

It follows from the previous arguments that (34) is maximized by the delegation set $D = [\underline{\omega}, \omega_i^0]$, as for this set: $x_D(\omega) = x_A(\omega)$ on $[\underline{\omega}, \omega_i^0]$ and $x_D(\omega)$ is constant on $[\omega^0, \bar{\omega}]$. The former implies that the first three terms in (34) are maximized point-wise, while the fourth term attains its upper bound, as $\int_{\omega^0}^{\bar{\omega}} g(\omega) d\omega = 0$ due to (c2). \square

Proof of Lemma 1. The existence and uniqueness of the solution ω_i^0 follow from part (a) of Lemma A.1. Now we verify that conditions (c1), (c2), (c2') from Lemma A.2 are satisfied when we set $g(\omega) = 1 - F_i(\omega) - b(\omega)f_i(\omega)$ and $\underline{g} = f_i(\omega)b(\underline{\omega})$. Clearly, (c1) and (c2') are satisfied by Assumptions A1 and A2. Finally, (c2) follows directly from part (a) of Lemma A.1. \square

³⁸Recall that x_D is increasing.

Proofs for Section 3

Proof of Lemma 3. (a) Let us denote $\Delta(\omega) = u_{\tilde{D}}(\omega) - u_D(\omega)$. Then

$$\begin{aligned}
U_i(\tilde{D}) - U_i(D) &= \int_{\underline{\omega}}^{\bar{\omega}} \Delta(\omega) f_i(\omega) d\omega \\
&= \int_{\underline{\omega}}^{\hat{\omega}} \Delta(\omega) \frac{f_i(\omega)}{f_j(\omega)} f_j(\omega) d\omega + \int_{\hat{\omega}}^{\bar{\omega}} \Delta(\omega) \frac{f_i(\omega)}{f_j(\omega)} f_j(\omega) d\omega \\
&\geq \frac{f_i(\hat{\omega})}{f_j(\hat{\omega})} \int_{\underline{\omega}}^{\hat{\omega}} \Delta(\omega) f_j(\omega) d\omega + \frac{f_i(\hat{\omega})}{f_j(\hat{\omega})} \int_{\hat{\omega}}^{\bar{\omega}} \Delta(\omega) f_j(\omega) d\omega \\
&= \frac{f_i(\hat{\omega})}{f_j(\hat{\omega})} \int_{\underline{\omega}}^{\bar{\omega}} \Delta(\omega) f_j(\omega) d\omega = \frac{f_i(\hat{\omega})}{f_j(\hat{\omega})} [U_j(\tilde{D}) - U_j(D)] \geq 0, \quad (36)
\end{aligned}$$

where $\hat{\omega}$ is given by (14), since \tilde{D} is an upward shift of D . The inequality in the third line follows from the inequalities: $f_i(\omega)/f_j(\omega) \leq f_i(\hat{\omega})/f_j(\hat{\omega})$ (since f_i/f_j is increasing) and $\Delta(\omega) \leq 0$ (due to (14)) for $\omega \leq \hat{\omega}$, while $f_i(\omega)/f_j(\omega) \geq f_i(\hat{\omega})/f_j(\hat{\omega})$ and $\Delta(\omega) \geq 0$ for $\omega \geq \hat{\omega}$. The final inequality follows by assumption. This establishes part (a).

Part (b) follows from an analogous chain of inequalities as in part (a) but, since f_i/f_j is now decreasing, the inequalities are now reversed. This completes the proof. \square

Proof of Lemma 4. By (14), $u_{\tilde{D}}(\underline{\omega}) \leq u_D(\underline{\omega})$ and $u_{\tilde{D}}(\bar{\omega}) \geq u_D(\bar{\omega})$, implying the claim. \square

Proof of Lemma 5. The proof is identical to the proof of Lemma 3 where we replace $f_i(\omega)$ by $f_i(\omega) + (b(\omega)f_i(\omega))'$. \square

Proofs for Section 4

The proof of Proposition 1 makes use of the following two lemmata. Lemma A.3 is a slightly modified version of Luenberger's [29] sufficiency condition (Theorem 1, p. 220) tailored to our maximization problem over menus of delegation sets.

Lemma A.3. *Let Ω be an arbitrary set ($\Omega \neq \emptyset$). Let $f : \Omega \rightarrow \mathbb{R}$ and $G : \Omega \rightarrow \mathbb{R}^k$ (where $k \in \mathbb{N}$) be two functions and let $y \in \mathbb{R}^k$. Suppose there exists $\lambda \in \mathbb{R}^k$, $\lambda \geq 0$ and an element $x_0 \in \Omega$ such that $G(x_0) = y$ and*

$$f(x_0) + \lambda^T G(x_0) \geq f(x) + \lambda^T G(x) \quad (37)$$

for all $x \in \Omega$. Then x_0 solves the problem $\max_{x \in \Omega} f(x)$ s.t. $G(x) \geq y$.

Lemma A.4. *For all i, j , define*

$$R_{i,j}(\theta) = \frac{\int_{\theta}^{\bar{\omega}} 1 - F_i(\omega) - b(\omega)f_i(\omega) d\omega}{\int_{\theta}^{\bar{\omega}} 1 - F_j(\omega) d\omega}. \quad (38)$$

If $\rho_{i,j}$ is decreasing, then $R_{i,j}$ is decreasing, and $R_{i,j}(\omega) \leq \rho_{i,j}(\omega)$ for all $\omega \in [\underline{\omega}, \bar{\omega}]$.

Proof of Lemma A.3. Let $x \in \Omega$ be such that $G(x) \geq y$. Then, since $\lambda \geq 0$, it follows that $\lambda^T G(x) \geq \lambda^T y = \lambda^T G(x_0)$. This, together with (37), yields

$$f(x_0) + \lambda^T G(x_0) \geq f(x) + \lambda^T G(x) \geq f(x) + \lambda^T G(x_0). \quad (39)$$

Hence, $f(x_0) \geq f(x)$, and thus x_0 indeed solves $\max_{x \in \Omega} f(x)$ s.t. $G(x) \geq y$. \square

Proof of Lemma A.4. We proceed in three steps.

Step 1. We show a general statement: If $g_1, g_2 : [\underline{\omega}, \bar{\omega}] \rightarrow \mathbb{R}$ are integrable functions such that $g_1 > 0$ and $g_2(\omega)/g_1(\omega)$ is decreasing in ω , then $\int_{\theta}^{\bar{\omega}} g_2(\omega) d\omega / \int_{\theta}^{\bar{\omega}} g_1(\omega) d\omega$ is decreasing in θ and it is bounded from above by $g_2(\theta)/g_1(\theta)$.

To see this, let $\theta_1, \theta_2 \in [\underline{\omega}, \bar{\omega}]$ with $\theta_1 < \theta_2$. Monotonicity and positivity of g_1 imply

$$\frac{\int_{\theta_2}^{\bar{\omega}} g_2(\omega) d\omega}{\int_{\theta_2}^{\bar{\omega}} g_1(\omega) d\omega} = \frac{\int_{\theta_2}^{\bar{\omega}} \frac{g_2(\omega)}{g_1(\omega)} \cdot g_1(\omega) d\omega}{\int_{\theta_2}^{\bar{\omega}} g_1(\omega) d\omega} \leq \frac{\int_{\theta_2}^{\bar{\omega}} \frac{g_2(\theta_2)}{g_1(\theta_2)} \cdot g_1(\omega) d\omega}{\int_{\theta_2}^{\bar{\omega}} g_1(\omega) d\omega} = \frac{g_2(\theta_2)}{g_1(\theta_2)}. \quad (40)$$

By an analogous argument,

$$\frac{g_2(\theta_2)}{g_1(\theta_2)} \leq \frac{\int_{\theta_1}^{\theta_2} g_2(\omega) d\omega}{\int_{\theta_1}^{\theta_2} g_1(\omega) d\omega}. \quad (41)$$

Now (40) and (41) imply that

$$\frac{\int_{\theta_2}^{\bar{\omega}} g_2(\omega) d\omega}{\int_{\theta_2}^{\bar{\omega}} g_1(\omega) d\omega} \leq \frac{\int_{\theta_1}^{\theta_2} g_2(\omega) d\omega + \int_{\theta_2}^{\bar{\omega}} g_2(\omega) d\omega}{\int_{\theta_1}^{\theta_2} g_1(\omega) d\omega + \int_{\theta_2}^{\bar{\omega}} g_1(\omega) d\omega} = \frac{\int_{\theta_1}^{\bar{\omega}} g_2(\omega) d\omega}{\int_{\theta_1}^{\bar{\omega}} g_1(\omega) d\omega}. \quad (42)$$

Hence, $\int_{\theta}^{\bar{\omega}} g_2(\omega) d\omega / \int_{\theta}^{\bar{\omega}} g_1(\omega) d\omega$ is decreasing. It is bounded by $g_2(\theta)/g_1(\theta)$ due to (40).

Step 2. We show that $[1 - F_i(\theta) - b(\theta)f_i(\theta)]/[1 - F_j(\theta)]$ is decreasing in θ . For this observe that

$$\frac{1 - F_i(\theta) - b(\theta)f_i(\theta)}{1 - F_j(\theta)} = \frac{\int_{\theta}^{\bar{\omega}} f_i(\omega) + (b(\omega)f_i(\omega))' d\omega}{\int_{\theta}^{\bar{\omega}} f_j(\omega) d\omega} - \frac{b(\bar{\omega})f_i(\bar{\omega})}{1 - F_j(\theta)}. \quad (43)$$

Now, the first term on the right hand side is decreasing due to Step 1 with $g_2 = f_i + (bf_i)'$ and $g_1 = f_j$, while the second term (including the minus sign) is decreasing, as $b(\bar{\omega})f_i(\bar{\omega}) \geq 0$, and $1 - F_j(\theta)$ is decreasing. Moreover, it follows from Step 1 and from $b(\bar{\omega})f_i(\bar{\omega}) \geq 0$ that $[1 - F_i(\theta) - b(\theta)f_i(\theta)]/[1 - F_j(\theta)] \leq [f_i(\theta) + (b(\theta)f_i(\theta))'] / f_j(\theta)$.

Step 3. Finally, let $g_2 = 1 - F_i - bf_i$ and $g_1 = 1 - F_j$. That $R_{i,j}(\theta) = \int_{\theta}^{\bar{\omega}} g_2(\omega) d\omega / \int_{\theta}^{\bar{\omega}} g_1(\omega) d\omega$ is decreasing now follows by the monotonicity established in Step 1 and 2. That $R_{i,j} \leq \rho_{i,j}$ follows from the bound established in Step 1 and the final inequality in Step 2. \square

*Proof of Proposition 1.*³⁹ Define the non-empty sets⁴⁰

$$L = \{\ell \in \{1, \dots, I\} \mid \omega_\ell^0 < \omega^{st}\}, \quad H = \{h \in \{1, \dots, I\} \mid \omega_h^0 > \omega^{st}\}. \quad (44)$$

We show that (D^{st}, \dots, D^{st}) solves the relaxed problem

$$\mathcal{R} : \quad \max_{(D_1, \dots, D_I)} \sum_i \mu_i V_i(D_i) \quad s.t. \quad IC_{\ell, h} \quad \text{for all } \ell \in L, h \in H. \quad (45)$$

Since (D^{st}, \dots, D^{st}) satisfies all ignored constraints, it then also solves problem \mathcal{P} .⁴¹

Let $\lambda_{\ell, h}$ be the multiplier associated to $IC_{\ell, h}$, and let λ be the $|L| \cdot |H| \times 1$ vector of $\lambda_{\ell, h}$'s. The Lagrangian for problem \mathcal{R} is

$$\mathcal{L}(D_1, \dots, D_I, \lambda) = \sum_{\ell \in L} \left(\mu_\ell V_\ell(D_\ell) + \sum_{h \in H} \lambda_{\ell, h} [U_\ell(D_\ell) - U_\ell(D_h)] \right) + \sum_{h \in H} \mu_h V_h(D_h). \quad (46)$$

Using (8) and the fact that $U_i(D_j) = \int_{\underline{\omega}}^{\bar{\omega}} u_{D_j}(\omega) f_i(\omega) d\omega$, it follows from inspection that $\mathcal{L}(D_1, \dots, D_I, \lambda) = \sum_i \mathcal{L}_i(D_i, \lambda)$, where for $\ell \in L$ and $h \in H$,

$$\begin{aligned} \mathcal{L}_\ell(D_\ell, \lambda) &= \mu_\ell u_{D_\ell}(\underline{\omega}) f_\ell(\underline{\omega}) b(\underline{\omega}) - \mu_\ell u_{D_\ell}(\bar{\omega}) f_\ell(\bar{\omega}) b(\bar{\omega}) \\ &\quad + \int_{\underline{\omega}}^{\bar{\omega}} u_{D_\ell}(\omega) \left\{ \mu_\ell [f_\ell(\omega) + (b(\omega) f_\ell(\omega)')] + \sum_{h \in H} \lambda_{\ell, h} f_\ell(\omega) \right\} d\omega, \end{aligned} \quad (47)$$

$$\begin{aligned} \mathcal{L}_h(D_h, \lambda) &= \mu_h u_{D_h}(\underline{\omega}) f_h(\underline{\omega}) b(\underline{\omega}) - \mu_h u_{D_h}(\bar{\omega}) f_h(\bar{\omega}) b(\bar{\omega}) \\ &\quad + \int_{\underline{\omega}}^{\bar{\omega}} u_{D_h}(\omega) \left\{ \mu_h [f_h(\omega) + (b(\omega) f_h(\omega)')] - \sum_{\ell \in L} \lambda_{\ell, h} f_\ell(\omega) \right\} d\omega. \end{aligned} \quad (48)$$

In what follows we show that there exists $\lambda \geq 0$ (component-wise) such that D^{st} maximizes $\mathcal{L}_\ell(D_\ell, \lambda)$ for each $\ell \in L$, as well as $\mathcal{L}_h(D_h, \lambda)$ for each $h \in H$. This will imply that the menu (D^{st}, \dots, D^{st}) maximizes $\mathcal{L}(D_1, \dots, D_I, \lambda)$. According to Lemma A.3, (D^{st}, \dots, D^{st}) solves problem \mathcal{R} .⁴²

We proceed in three steps. In Step 1, we propose the multipliers. We derive them from a restricted problem allowing only for intervals. In Steps 2 and 3, we show that for the proposed multipliers, D^{st} maximizes $\mathcal{L}_\ell(D_\ell, \lambda)$ and $\mathcal{L}_h(D_h, \lambda)$.

³⁹The selection of the constraints of the relaxed problem, and Step 1 of the proof below are adopted from the proof of Theorem 1 in Krähmer and Strausz [26].

⁴⁰To see non-emptiness, recall from Lemma 2 that $\sum_i \mu_i \Gamma_i(\omega^{st}) = 0$. Now, if $\omega \leq \omega_1^0$, then by Lemma A.1, $\Gamma_i(\omega) \geq 0$ (with strict inequality for all $i > 1$). Thus, $\sum_i \mu_i \Gamma_i(\omega) > 0$. Analogously $\sum_i \mu_i \Gamma_i(\omega) < 0$ for $\omega > \omega_I^0$. Therefore $\omega_1^0 < \omega^{st} < \omega_I^0$.

⁴¹If $\omega_i^0 = \omega^{st}$ for some type $i \in \{1, \dots, I\}$, we ignore all constraints for type i . Because ω^{st} maximizes the principal's expected utility conditional on such a type i , we omit it from further considerations. For the remainder of the proof, we for simplicity assume that there is no such type i .

⁴²Observe that for the delegation menu (D^{st}, \dots, D^{st}) all $IC_{\ell, h}$ constraints hold with equality, as required in Lemma A.3.

Step 1. We establish the multipliers $\lambda_{\ell,h}$. For this consider a restricted problem where all delegations sets D_i are intervals of the form $[x_A(\underline{\omega}), x_A(\hat{\omega}_i)]$, where $\hat{\omega}_i \in [\underline{\omega}, \bar{\omega}]$. As this problem differs from problem \mathcal{R} only in the domain, it has the same Lagrangian. Moreover, the maximization of each of the \mathcal{L}_i 's is just a single-variable maximization problem with respect to $\hat{\omega}_i$. By an analogous argument as in the proof of part (b) of Lemma A.1, the first-order conditions for the restricted problem are

$$\ell \in L : \quad \mu_\ell \Gamma_\ell(\omega^{st}) + \sum_{h \in H} \lambda_{\ell,h} \int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega = 0, \quad (49)$$

$$h \in H : \quad \mu_h \Gamma_h(\omega^{st}) - \sum_{\ell \in L} \lambda_{\ell,h} \int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega = 0, \quad (50)$$

with $\Gamma_i(\omega)$ defined by (28). This is a system of I linear equations in the $|L| \cdot |H|$ unknowns $\lambda_{\ell,h}$. We show that it has a non-negative solution λ , that is, $\lambda_{\ell,h} \geq 0$. To see this, it is useful to write the system in matrix notation $A\lambda = \beta$. Let β be the $I \times 1$ vector with component $\beta_i = -\mu_i \Gamma_i(\omega^{st})$. Consider the index m with $\lambda_m = \lambda_{\ell,h}$, and define α_m as the $I \times 1$ vector with the entry $\int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega$ in the ℓ -th and the entry $-\int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega$ in the h -th row and 0's elsewhere:

$$\alpha_m^T = \left(0, \dots, 0, \int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega, 0, \dots, 0, -\int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega, 0, \dots, 0 \right). \quad (51)$$

$\begin{array}{ccc} \uparrow & & \uparrow \\ \ell & & h \end{array}$

Finally, let $A = (\alpha_1, \dots, \alpha_{|L| \cdot |H|})$. It can now be seen by inspection that the system (49), (50) is equivalent to $A\lambda = \beta$.

To see that $A\lambda = \beta$ has a non-negative solution, recall from Farkas' lemma that exactly one of the following two statements is true:

$$\text{There exists a } \lambda \in \mathbb{R}^{|L| \cdot |H|} \text{ such that } A\lambda = \beta \text{ and } \lambda \geq 0. \quad (52)$$

$$\text{There exists a } y \in \mathbb{R}^I \text{ such that } y^T A \geq 0 \text{ and } y^T \beta < 0. \quad (53)$$

We show that the second statement is not true. Assume, to the contrary, it is true, i.e., such a y exists. Then $y^T \alpha_m \geq 0$ for all $m \in \{1, \dots, |L| \cdot |H|\}$, or equivalently $(y_\ell - y_h) \int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega \geq 0$. Hence, $y_\ell \geq y_h$ for all $\ell \in L, h \in H$, and thus,

$$\max_{h \in H} y_h \leq \min_{\ell \in L} y_\ell. \quad (54)$$

Recall that $\omega_\ell^0 \leq \omega^{st}$ and $\omega_h^0 \geq \omega^{st}$ by definition of L and H . Then by part (a) of

Lemma A.1, we have $\beta_\ell = -\mu_\ell \Gamma_\ell(\omega^{st}) \geq 0$ and $\beta_h = -\mu_h \Gamma_h(\omega^{st}) \leq 0$. Thus,

$$y^T \beta = \sum_{\ell \in L} \beta_\ell y_\ell + \sum_{h \in H} \beta_h y_h \geq \min_{\ell \in L} y_\ell \sum_{\ell \in L} \beta_\ell + \max_{h \in H} y_h \sum_{h \in H} \beta_h \geq \max_{h \in H} y_h \sum_{i \in L \cup H} \beta_i = 0, \quad (55)$$

where the second inequality follows from (54), and the final equality from the definition of ω^{st} and β .

Step 2. We show that for the multipliers established in Step 1, D^{st} maximizes $\mathcal{L}_\ell(D_\ell, \lambda)$ for all $\ell \in L$. For this we apply Lemma A.2 where we set

$$g(\omega) = \mu_\ell(1 - F_\ell(\omega) - b(\omega)f_\ell(\omega)) + \sum_{h \in H} \lambda_{\ell,h}(1 - F_\ell(\omega)) \quad \text{and} \quad \underline{g} = \mu_\ell f_\ell(\underline{\omega})b(\underline{\omega}). \quad (56)$$

Then $g(\bar{\omega}) = -\mu_\ell b(\bar{\omega})f_\ell(\bar{\omega})$ and $-g'(\omega) = \mu_\ell(f_\ell(\omega) + (b(\omega)f_\ell(\omega))') + \sum_{h \in H} \lambda_{\ell,h}f_\ell(\omega)$, and thus the objective function in (32) becomes equal to $\mathcal{L}_\ell(D_\ell, \lambda)$.

By Lemma A.2, it is thus sufficient to verify that conditions (c1), (c2), (c2') are satisfied for $\omega^0 = \omega^{st}$. Clearly, condition (c2') is satisfied by Assumption A2. It follows from Assumption A1 and from $\lambda \geq 0$ that $g'(\omega) < 0$ for all $\omega \in [\underline{\omega}, \bar{\omega}]$. This shows that (c1) is satisfied. Moreover, it also implies that $\bar{G}(\omega) = \int_{\underline{\omega}}^{\bar{\omega}} g(\tilde{\omega}) d\tilde{\omega}$ is convex. Now recall that the first-order condition (49) can be rewritten as $\bar{G}(\omega^{st}) = 0$. This together with the convexity of \bar{G} and $\bar{G}(\bar{\omega}) = 0$ implies that $\bar{G}(\omega) \leq 0$ for all $\omega \in [\omega^{st}, \bar{\omega}]$. Thus, indeed (c2) is satisfied.

Step 3. We show that for the multipliers established in Step 1, D^{st} maximizes $\mathcal{L}_h(D_h, \lambda)$ for all $h \in H$. We again apply Lemma A.2 where we set

$$g(\omega) = \mu_h(1 - F_h(\omega) - b(\omega)f_h(\omega)) - \sum_{\ell \in L} \lambda_{\ell,h}(1 - F_\ell(\omega)) \quad \text{and} \quad \underline{g} = \mu_h f_h(\underline{\omega})b(\underline{\omega}). \quad (57)$$

Then $g(\bar{\omega}) = -\mu_h b(\bar{\omega})f_h(\bar{\omega})$ and $-g'(\omega) = \mu_h(f_h(\omega) + (b(\omega)f_h(\omega))') - \sum_{\ell \in L} \lambda_{\ell,h}f_\ell(\omega)$, and thus the objective function in (32) becomes equal to $\mathcal{L}_h(D_h, \lambda)$.

By Lemma A.2, it is thus sufficient to verify that conditions (c1), (c2), (c2') are satisfied for $\omega^0 = \omega^{st}$. First, (c2') is satisfied by Assumption A2. Second, we verify (c1). Assume, to the contrary, that (c1) does not hold, i.e., that $g'(\omega) > 0$ for some $\omega = \tilde{\omega} < \omega^{st}$. Let us rewrite

$$\begin{aligned} -g'(\omega) &= \mu_h[f_h(\omega) + (b(\omega)f_h(\omega))'] - \sum_{\ell \in L} \lambda_{\ell,h}f_\ell(\omega) \\ &= \left[\mu_h - \sum_{\ell \in L} \lambda_{\ell,h} \frac{1}{\rho_{h,\ell}(\omega)} \right] [f_h(\omega) + (b(\omega)f_h(\omega))']. \end{aligned} \quad (58)$$

Then it follows from the assumption that $\rho_{h,\ell}$ is decreasing and from Assumption A1, that

the first factor (in square brackets) is decreasing.⁴³ Thus, $g'(\omega) > 0$ for all $\omega > \tilde{\omega}$. As $b(\bar{\omega}) > 0$, by Assumption A2, we have $g(\bar{\omega}) = -\mu_h b(\bar{\omega}) f_h(\bar{\omega}) \leq 0$, and we obtain $g(\omega) = g(\bar{\omega}) - \int_{\omega}^{\bar{\omega}} g'(\tilde{\omega}) d\tilde{\omega} < 0$ for all $\omega \in (\tilde{\omega}, \bar{\omega})$. Since $\tilde{\omega} < \omega^{st}$, this implies $\int_{\omega^{st}}^{\bar{\omega}} g(\omega) d\omega < 0$, which contradicts the first-order condition (50).

Finally, we verify (c2). Consider first $\omega \in [\omega_h^0, \bar{\omega}]$. Recall that then $\Gamma_h(\omega) < 0$ due to part (a) of Lemma A.1. As $\lambda_{\ell,h} \geq 0$, it follows immediately that $\int_{\omega}^{\bar{\omega}} g(\omega) d\omega \leq 0$ for all $\omega \in [\omega_h^0, \bar{\omega}]$. Thus, it remains to show the desired inequality in (c2) for $\omega \in [\omega^{st}, \omega_h^0]$. In this case, part (a) of Lemma A.1 implies that $\Gamma_h(\omega) > 0$. Let us now rewrite

$$\int_{\omega}^{\bar{\omega}} g(\tilde{\omega}) d\tilde{\omega} = \mu_h \Gamma_h(\omega) - \sum_{\ell \in L} \lambda_{\ell,h} \int_{\omega}^{\bar{\omega}} 1 - F_{\ell}(\tilde{\omega}) d\tilde{\omega} = \left[\mu_h - \sum_{\ell \in L} \lambda_{\ell,h} \frac{\int_{\omega}^{\bar{\omega}} 1 - F_{\ell}(\tilde{\omega}) d\tilde{\omega}}{\Gamma_h(\omega)} \right] \Gamma_h(\omega). \quad (59)$$

By assumption, $\rho_{h,\ell}$ is decreasing, and by Lemma A.4, $\Gamma_h(\omega) / \int_{\omega}^{\bar{\omega}} 1 - F_{\ell}(\tilde{\omega}) d\tilde{\omega}$ is decreasing. Thus, for $\omega \in [\omega^{st}, \omega_h^0]$ the term in the square brackets is decreasing (as $\Gamma_h(\omega) > 0$), while it is equal to 0 for $\omega = \omega^{st}$ (due to (50)). Hence, indeed $\int_{\omega}^{\bar{\omega}} g(\tilde{\omega}) d\tilde{\omega} \leq 0$ for $\omega \in [\omega^{st}, \omega_h^0]$. This verifies (c2) and completes the proof. \square

Proof of Proposition 2. Let $b(\omega)$ be such that Assumptions A1–A3 are satisfied. We show first that A1–A3 are also satisfied when the bias is $\alpha b(\omega)$, where $\alpha \in (0, 1]$. First, consider A1. It is clearly satisfied when $(b(\omega)f_i(\omega))' \geq 0$. It is also satisfied when $(b(\omega)f_i(\omega))' < 0$, because then $f_i(\omega) + (\alpha b(\omega)f_i(\omega))' = f_i(\omega) + \alpha(b(\omega)f_i(\omega))' \geq f_i(\omega) + (b(\omega)f_i(\omega))' > 0$. Second, A2 is clearly satisfied. Third, consider A3. It is clearly satisfied when $\int_{\omega}^{\bar{\omega}} b(\omega)f_i(\omega) d\omega < 0$. It is also satisfied when $\int_{\omega}^{\bar{\omega}} b(\omega)f_i(\omega) d\omega \geq 0$, because then $\int_{\omega}^{\bar{\omega}} 1 - F_i(\omega) - \alpha b(\omega)f_i(\omega) d\omega \geq \int_{\omega}^{\bar{\omega}} 1 - F_i(\omega) - b(\omega)f_i(\omega) d\omega > 0$.

Before proceeding with the proof, let us prove the following useful result:

$$\alpha \frac{\int_{\omega^{st}}^{\bar{\omega}} b(\omega)f_i(\omega) d\omega}{\int_{\omega^{st}}^{\bar{\omega}} 1 - F_i(\omega) d\omega} \rightarrow 1 \quad \text{for all } i \text{ as } \alpha \rightarrow 0. \quad (60)$$

To see this, recall from Lemma 2 that ω^{st} satisfies $\int_{\omega^{st}}^{\bar{\omega}} 1 - F^{st}(\omega) - \alpha b(\omega)f^{st}(\omega) d\omega = 0$. Thus,

$$\alpha = \frac{\int_{\omega^{st}}^{\bar{\omega}} 1 - F^{st}(\omega) d\omega}{\int_{\omega^{st}}^{\bar{\omega}} b(\omega)f^{st}(\omega) d\omega}, \quad (61)$$

and it follows from the *Implicit function theorem* that $\omega^{st} \rightarrow \bar{\omega}$ as $\alpha \rightarrow 0$.⁴⁴ By (61), the

⁴³As $\omega_i^0 \neq \omega^{st}$ for all $i \in L \cup H$, each of the conditions (49), (50), and thus also each of the Lagrangians (47), (48), contains at least one multiplier that is positive. In other words: (i) For any $\ell \in L$ there is $h \in H$ such that $\lambda_{\ell,h} > 0$. (ii) For any $h \in H$ there is $\ell \in L$ such that $\lambda_{\ell,h} > 0$.

⁴⁴In order to apply the *Implicit function theorem*, we need to verify that the derivative of the right hand side of (61) with respect to ω^{st} is not equal to zero. Indeed, a straightforward computation reveals that the derivative is equal to $-1/(2b(\bar{\omega})) \neq 0$.

ratio in (60) can be rewritten as

$$\frac{\int_{\omega^{st}}^{\bar{\omega}} 1 - F^{st}(\omega) d\omega}{\int_{\omega^{st}}^{\bar{\omega}} 1 - F_i(\omega) d\omega} \cdot \frac{\int_{\omega^{st}}^{\bar{\omega}} b(\omega) f_i(\omega) d\omega}{\int_{\omega^{st}}^{\bar{\omega}} b(\omega) f^{st}(\omega) d\omega} \rightarrow \frac{f^{st}(\bar{\omega})}{f_i(\bar{\omega})} \cdot \frac{b(\bar{\omega}) f_\ell(\bar{\omega})}{b(\bar{\omega}) f^{st}(\bar{\omega})} = 1, \quad (62)$$

where the limit for $\omega^{st} \rightarrow \bar{\omega}$ is obtained by using *L'Hospital's rule* for each fraction (twice for the first one and once for the second one). This proves (60).

Let us now denote

$$m_1 = \min_{\omega, i} f_i(\omega), \quad m_2 = \max_{\omega, i} f_i(\omega), \quad m_3 = \max_{\omega, i} |(b(\omega) f_2(\omega))'|, \quad \underline{\mu} = \min_i \mu_i. \quad (63)$$

Minima and maxima exist due to continuity. Moreover, $\underline{\mu} > 0$ and, due to positive densities, also $m_1 > 0$. Now, consider $\varepsilon > 0$ such that

$$\varepsilon < \frac{\underline{\mu}}{m_2 + \underline{\mu}} m_1. \quad (64)$$

By (60), there is $\bar{\alpha} > 0$ so that for all $\alpha \in (0, \bar{\alpha})$:

$$\alpha m_3 < \varepsilon, \quad \text{and} \quad \left| -1 + \alpha \frac{\int_{\omega^{st}}^{\bar{\omega}} b(\omega) f_i(\omega) d\omega}{\int_{\omega^{st}}^{\bar{\omega}} 1 - F_i(\omega) d\omega} \right| < \varepsilon \quad \text{for all } i. \quad (65)$$

We are now in the position to show that static delegation is optimal when $\alpha \in (0, \bar{\alpha})$. We proceed as in the proof of Proposition 1. Fix $\alpha \in (0, \bar{\alpha})$ (and thus the corresponding ω^{st}) and consider the relaxed problem \mathcal{R} with constraints $IC_{\ell, h}$ as defined in the proof of Proposition 1. We again show that D^{st} maximizes $\mathcal{L}_\ell(D_\ell, \lambda)$ as well as $\mathcal{L}_h(D_h, \lambda)$ for each $\ell \in L$, $h \in H$. The desired claim then follows from Lemma A.3

In Step 1, we establish the multipliers by the first-order conditions (49) and (50). In addition, (49) and (65) imply:

$$\sum_{\ell \in L} \sum_{h \in H} \lambda_{\ell, h} = \sum_{\ell \in L} \mu_\ell \left[-1 + \alpha \frac{\int_{\omega^{st}}^{\bar{\omega}} b(\omega) f_\ell(\omega) d\omega}{\int_{\omega^{st}}^{\bar{\omega}} 1 - F_\ell(\omega) d\omega} \right] < \sum_{\ell \in L} \mu_\ell \cdot \varepsilon < \varepsilon. \quad (66)$$

In Step 2, we show that D^{st} maximizes $\mathcal{L}_\ell(D_\ell, \lambda)$ with λ established in Step 1. Because $\alpha b(\omega)$ satisfies A1–A3, as remarked above, we can use the same proof as in Step 2 in the proof of Proposition 1.

In Step 3, we show that D^{st} maximizes $\mathcal{L}_h(D_h, \lambda)$. We apply Lemma A.2 in the same way as in Step 3 in the proof of Proposition 1 and we set

$$g(\omega) = \mu_h(1 - F_h(\omega) - \alpha b(\omega) f_h(\omega)) - \sum_{\ell \in L} \lambda_{\ell, h}(1 - F_\ell(\omega)), \quad \text{and} \quad \underline{g} = \mu_h \alpha f_h(\underline{\omega}) b(\underline{\omega}). \quad (67)$$

We now verify (c1), (c2), (c2'). Clearly, (c2') is satisfied. In order to verify (c1), we have

$$\begin{aligned} -g'(\omega) &= \mu_h[f_h(\omega) + \alpha(b(\omega)f_h(\omega))'] - \sum_{\ell \in L} \lambda_{\ell,h} f_\ell(\omega) \\ &\geq \mu_h(m_1 - \alpha m_3) - m_2 \sum_{\ell \in L} \lambda_{\ell,h} \geq \mu_h(m_1 - \varepsilon) - m_2 \varepsilon > 0, \end{aligned} \quad (68)$$

where the last inequality follows from (64), and the previous ones from (63), (65), and (66). Thus, we have $g'(\omega) < 0$ for all $\omega \in [\underline{\omega}, \bar{\omega}]$, and this verifies (c1). Verification of (c2) then follows from an identical argument as in Step 2 of Proposition 1. (Again $\alpha b(\omega)$ satisfies A1–A3, as remarked above.) This completes the proof. \square

Proofs for Section 5

Proof of formula (21). For small $\varepsilon > 0$ and arbitrary $\delta > 0$ consider the delegation set

$$\tilde{D}_2(\theta, \eta, \delta, \varepsilon) = [x_A(\underline{\omega}), x_A(\eta)] \cup [x_A(\eta + \delta), x_A(\theta + \varepsilon)], \quad (69)$$

and denote the principal's and agent type 1's expected utility from this delegation set respectively as $\tilde{V}_2(\theta, \eta, \delta, \varepsilon)$ and $\tilde{U}_1(\theta, \eta, \delta, \varepsilon)$.

Observe that for $\delta(\varepsilon)$ defined in the text in (20), $D_2(\theta, \eta, \varepsilon) = \tilde{D}_2(\theta, \eta, \delta(\varepsilon), \varepsilon)$. In particular, $V_2(D_2(\theta, \eta, \varepsilon)) = \tilde{V}_2(\theta, \eta, \delta(\varepsilon), \varepsilon)$, and since, by assumption, $\tilde{U}_1(\theta, \eta, \delta(\varepsilon), \varepsilon) = U_1(D_2(\theta, \eta, \varepsilon)) = U_1(D_1)$, we obtain

$$\frac{\partial V_2(D_2(\theta, \eta, \varepsilon))}{\partial \varepsilon} = -\frac{\frac{\partial}{\partial \delta} \tilde{V}_2(\theta, \eta, \delta(\varepsilon), \varepsilon)}{\frac{\partial}{\partial \delta} \tilde{U}_1(\theta, \eta, \delta(\varepsilon), \varepsilon)} \cdot \frac{\partial}{\partial \varepsilon} \tilde{U}_1(\theta, \eta, \delta(\varepsilon), \varepsilon) + \frac{\partial}{\partial \varepsilon} \tilde{V}_2(\theta, \eta, \delta(\varepsilon), \varepsilon). \quad (70)$$

We now compute the partial derivatives on the right hand side. By Lemma A.1,⁴⁵

$$\frac{\partial}{\partial \varepsilon} \tilde{V}_2(\theta, \eta, \delta, \varepsilon) = \int_{\theta+\varepsilon}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega)f_2(\omega) d\omega \cdot x'_A(\theta + \varepsilon), \quad (71)$$

$$\frac{\partial}{\partial \varepsilon} \tilde{U}_1(\theta, \eta, \delta, \varepsilon) = \int_{\theta+\varepsilon}^{\bar{\omega}} 1 - F_1(\omega) d\omega \cdot x'_A(\theta + \varepsilon). \quad (72)$$

To compute the partial derivatives with respect to δ , let ω' be the state in which the agent is indifferent between the actions $x_A(\eta)$ and $x_A(\eta + \delta)$. Hence, for states $\omega \in (\eta, \omega')$, the agent chooses $x_A(\eta)$ from D_2 , and for $\omega \in (\omega', \eta + \delta)$, he chooses $x_A(\eta + \delta)$. For $\omega \in (\eta + \delta, \theta + \varepsilon)$, he chooses his favorite action, and in all other states, he chooses an

⁴⁵The expression for $\frac{\partial}{\partial \varepsilon} \tilde{U}_1$ follows by setting $b = 0$ in part (b) of Lemma A.1.

action that does not depend on δ . Thus, type 1's expected utility from D_2 is

$$\begin{aligned}\tilde{U}_1(\theta, \eta, \delta, \varepsilon) &= \int_{\eta}^{\omega'} U(\omega, x_A(\eta)) f_1(\omega) d\omega + \int_{\omega'}^{\eta+\delta} U(\omega, x_A(\eta + \delta)) f_1(\omega) d\omega \\ &\quad + \int_{\eta+\delta}^{\theta+\varepsilon} U(\omega, x_A(\omega)) f_1(\omega) d\omega + C,\end{aligned}\tag{73}$$

where C is a constant that does not depend on δ . Therefore, by Leibniz' rule,

$$\begin{aligned}\frac{\partial}{\partial \delta} \tilde{U}_1(\theta, \eta, \delta, \varepsilon) &= U(\omega', x_A(\eta)) f_1(\omega') \frac{\partial \omega'}{\partial \delta} - U(\omega', x_A(\eta + \delta)) f_1(\omega') \frac{\partial \omega'}{\partial \delta} \\ &\quad + U(\eta + \delta, x_A(\eta + \delta)) f_1(\eta + \delta) \\ &\quad + \int_{\omega'}^{\eta+\delta} \frac{\partial}{\partial x} U(\omega, x_A(\eta + \delta)) f_1(\omega) d\omega \cdot x'_A(\eta + \delta) \\ &\quad - U(\eta + \delta, x_A(\eta + \delta)) f_1(\eta + \delta).\end{aligned}\tag{74}$$

Because of the agent's indifference between $x_A(\eta)$ and $x_A(\eta + \delta)$ in state ω' , the first two terms cancel. In addition, the third and the fifth term cancel as well. Moreover,

$$\frac{\partial}{\partial x} U(\omega, x_A(\eta + \delta)) = \omega + a'(x_A(\eta + \delta)) = \omega - (\eta + \delta)\tag{75}$$

due the first order condition for x_A . Thus,

$$\frac{\partial}{\partial \delta} \tilde{U}_1(\theta, \eta, \delta, \varepsilon) = \int_{\omega'}^{\eta+\delta} [\omega - (\eta + \delta)] f_1(\omega) d\omega \cdot x'_A(\eta + \delta).\tag{76}$$

With identical steps, we also obtain

$$\frac{\partial}{\partial \delta} \tilde{V}_2(\theta, \eta, \delta, \varepsilon) = \int_{\omega'}^{\eta+\delta} [\omega - (\eta + \delta)] [f_2(\omega) + (b(\omega) f_2(\omega))'] d\omega \cdot x'_A(\eta + \delta).\tag{77}$$

Now, if ε goes to zero, so does $\delta(\varepsilon)$. Using this in (70) and collecting terms delivers

$$\begin{aligned}\frac{\partial V_2(D_2(\theta, \eta, \varepsilon))}{\partial \varepsilon} \Big|_{\varepsilon=0} &= \left[- \lim_{\varepsilon \rightarrow 0} \frac{\frac{\partial}{\partial \delta} \tilde{V}_2(\theta, \eta, \delta(\varepsilon), \varepsilon)}{\frac{\partial}{\partial \delta} \tilde{U}_1(\theta, \eta, \delta(\varepsilon), \varepsilon)} \cdot \int_{\theta}^{\bar{\omega}} 1 - F_1(\omega) d\omega \right. \\ &\quad \left. + \int_{\theta}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega) f_2(\omega) d\omega \right] x'_A(\theta),\end{aligned}\tag{78}$$

provided the limit exists. We now show that it does exist and, in fact, equals $\rho_{2,1}(\eta)$ which will then imply formula (21). Indeed, when δ goes to zero, then ω' converges to η and, hence, the integrals in (76) and (77) converge to zero. Applying *L'Hospital's Rule*

twice, we obtain

$$\lim_{\varepsilon \rightarrow 0} \frac{\frac{\partial}{\partial \delta} \tilde{V}_2(\theta, \eta, \delta(\varepsilon), \varepsilon)}{\frac{\partial}{\partial \delta} \tilde{U}_1(\theta, \eta, \delta(\varepsilon), \varepsilon)} = \lim_{\delta \rightarrow 0} \frac{\int_{\omega'}^{\eta+\delta} [\omega - (\eta + \delta)] [f_2(\omega) + (b(\omega)f_2(\omega))'] d\omega}{\int_{\omega'}^{\eta+\delta} [\omega - (\eta + \delta)] f_1(\omega) d\omega} = \rho_{2,1}(\eta), \quad (79)$$

and this completes the proof. \square

Proof of Lemma 6. The proof follows from the main text. \square

Proof of Proposition 3. (a) The proof follows from the main text.

(b) We show that if (23) is violated, then static delegation is optimal. We proceed similarly as in the proof of Proposition 1, with $L = \{1\}$ and $H = \{2\}$. The relaxed problem \mathcal{R} now involves only one constraint, $IC_{1,2}$. We again show that D^{st} maximizes $\mathcal{L}_1(D_1, \lambda_{1,2})$ as well as $\mathcal{L}_2(D_2, \lambda_{1,2})$. The desired claim then follows from Lemma A.3.

In Step 1, we establish the value of the multiplier. From (50) we have:

$$\lambda_{1,2} = \frac{\mu_2 \Gamma_2(\omega^{st})}{\int_{\omega^{st}}^{\bar{\omega}} 1 - F_1(\omega) d\omega}. \quad (80)$$

In Step 2, we show that for $\lambda_{1,2}$ established in Step 1, the delegation set D^{st} maximizes the Lagrangian $\mathcal{L}_1(D_1, \lambda_{1,2})$. The proof is identical to Step 2 in Proposition 1.

In Step 3, we show that D^{st} maximizes $\mathcal{L}_2(D_2, \lambda_{1,2})$. We use Lemma A.2 with

$$g(\omega) = \mu_2(1 - F_2(\omega) - b(\omega)f_2(\omega)) - \lambda_{1,2}(1 - F_1(\omega)) \quad \text{and} \quad \underline{g} = \mu_2 f_2(\underline{\omega}) b(\underline{\omega}). \quad (81)$$

Then $g(\bar{\omega}) = -\mu_2 b(\bar{\omega}) f_2(\bar{\omega})$ and $-g'(\omega) = \mu_2(f_2(\omega) + (b(\omega)f_2(\omega))') - \lambda_{1,2} f_1(\omega)$, and, thus, the objective function in (32) becomes equal to $\mathcal{L}_2(D_2, \lambda_{1,2})$.

We verify that conditions (c1), (c2), (c2') are satisfied for $\omega = \omega^{st}$. Clearly, condition (c2') is satisfied. In order to verify (c1), we use (80) and obtain:

$$\begin{aligned} -g'(\omega) &= \mu_2[f_2(\omega) + (b(\omega)f_2(\omega))'] - \lambda_{1,2} f_1(\omega) \\ &= \mu_2 \left[\rho_{2,1}(\omega) - \frac{\Gamma_2(\omega^{st})}{\int_{\omega^{st}}^{\bar{\omega}} 1 - F_1(\omega) d\omega} \right] f_1(\omega). \end{aligned} \quad (82)$$

Our assumption that (23) is violated implies that the term in the square brackets is non-negative for all $\omega \leq \omega^{st}$. Thus, indeed (c1) holds. Finally, the proof of (c2) is almost identical to the proof of (c2) in Step 3 in Proposition 1.⁴⁶ This completes the proof. \square

Proof of Lemma 7. We first show that ω^{st} as a function of μ_2 increases monotonically

⁴⁶The only difference is that now we directly assume that $\Gamma_2(\omega)/\int_{\omega}^{\bar{\omega}} 1 - F_1(\bar{\omega}) d\bar{\omega}$ is decreasing, and thus we do not need to prove it.

from ω_1^0 to ω_2^0 as μ_2 goes from 0 to 1. For this, let us define the following function

$$\Gamma^{st}(\omega, \mu_2) = (1 - \mu_2)\Gamma_1(\omega) + \mu_2\Gamma_2(\omega) = \Gamma_1(\omega) + \mu_2[\Gamma_2(\omega) - \Gamma_1(\omega)]. \quad (83)$$

Recall from Lemmata 1 and 2 that, for a given μ_2 , ω^{st} is a unique solution to the equation $\Gamma^{st}(\omega^{st}, \mu_2) = 0$. Moreover, as argued in footnote 40, $\omega^{st} \in (\omega_1^0, \omega_2^0)$.

Let $\tilde{\mu}_2 > \mu_2$ and consider $\tilde{\omega}^{st}, \omega^{st} \in (\omega_1^0, \omega_2^0)$ such that $\Gamma^{st}(\tilde{\omega}^{st}, \tilde{\mu}_2) = 0$ and $\Gamma^{st}(\omega^{st}, \mu_2) = 0$. We show that $\tilde{\omega}^{st} \geq \omega^{st}$. Indeed, as $\omega_1^0 < \tilde{\omega}^{st} < \omega_2^0$, Lemma A.1 implies $\Gamma_1(\tilde{\omega}^{st}) < 0$ and $\Gamma_2(\tilde{\omega}^{st}) > 0$ so that $\Gamma_2(\tilde{\omega}^{st}) - \Gamma_1(\tilde{\omega}^{st}) > 0$. Thus, (83) implies that $\Gamma^{st}(\tilde{\omega}^{st}, \mu_2) < \Gamma^{st}(\tilde{\omega}^{st}, \tilde{\mu}_2) = 0$. Hence, Lemma A.1 applied to Γ^{st} yields $\tilde{\omega}^{st} > \omega^{st}$, as desired.

In order to show the convergence, consider now the inverse relationship, i.e., μ_2 as a function of ω^{st} , given by $\mu_2 = -\Gamma_1(\omega^{st})/[\Gamma_2(\omega^{st}) - \Gamma_1(\omega^{st})]$. If $\omega^{st} \rightarrow \omega_1^0$, then $\Gamma_1(\omega^{st}) \rightarrow 0$ and $\mu_2 \rightarrow 0$. Finally, if $\omega^{st} \rightarrow \omega_2^0$, then $\Gamma_2(\omega^{st}) \rightarrow 0$ and $\mu_2 \rightarrow 1$.

The remainder of the proof now follows from the argument in the main text. \square

Proof of Lemma 8. Consider the relaxed problem

$$\mathcal{R} : \quad \max_{D_1, D_2} \mu_1 V_1(D_1) + \mu_2 V_2(D_2) \quad s.t. \quad IC_{1,2}.$$

To prove the claim, it suffices to show that there is a solution (D_1^r, D_2^r) to \mathcal{R} so that:

- (a) $D_1^r = [x_A(\underline{\omega}), x_A(\hat{\omega})]$ for some $\hat{\omega} \in [\underline{\omega}, \bar{\omega}]$;
- (b) $IC_{1,2}$ is binding: $U_1(D_1^r) = U_1(D_2^r)$;
- (c) $U_2(D_2^r) \geq U_2(D_1^r)$.

Proof of (a). By contradiction, suppose that $D_1^r \neq [x_A(\underline{\omega}), x_A(\hat{\omega})]$ for all solutions to \mathcal{R} . If $x_{D_1^r}(\bar{\omega}) > x_A(\bar{\omega})$, then our assumptions on U imply that there is a unique action $\bar{x} < x_A(\bar{\omega})$ so that $U(\bar{\omega}, \bar{x}) = U(\bar{\omega}, x_{D_1^r}(\bar{\omega}))$. If $x_{D_1^r}(\bar{\omega}) \leq x_A(\bar{\omega})$, we define $\bar{x} = x_{D_1^r}(\bar{\omega})$. Because U has the single-crossing property $\partial^2 U / \partial \omega \partial x = 1 > 0$, it follows by definition of \bar{x} that $U(\omega, \bar{x}) \geq U(\omega, x_{D_1^r}(\bar{\omega}))$ for all ω .

Now, define

$$\tilde{D}_1^r = \begin{cases} [x_A(\underline{\omega}), \bar{x}] & \text{if } \bar{x} \in [x_A(\underline{\omega}), x_A(\bar{\omega})], \\ \{x_A(\underline{\omega})\} & \text{if } \bar{x} \notin [x_A(\underline{\omega}), x_A(\bar{\omega})]. \end{cases} \quad (84)$$

It is easy to see that the considerations above imply:

$$u_{\tilde{D}_1^r}(\omega) \geq u_{D_1}(\omega) \quad \text{for all } \omega \in [\underline{\omega}, \bar{\omega}]. \quad (85)$$

We now argue that offering \tilde{D}_1^r instead of D_1^r would relax $IC_{1,2}$ and improve the principal's expected utility from type 1. Hence, there is also a solution to \mathcal{R} in which D_1^r is an interval of the form $[x_A(\underline{\omega}), x_A(\hat{\omega})]$, a contradiction.

To see this, observe first that by (85), the modification evidently increases type 1's expected utility: $U_1(\tilde{D}_1^r) \geq U_1(D_1^r)$, and hence $IC_{1,2}$ is relaxed by the modification.

To see that the modification is (weakly) profitable, suppose first that $\bar{x} \in [x_A(\underline{\omega}), x_A(\bar{\omega})]$. In this case, we have $u_{D_1^r}(\bar{\omega}) = u_{\tilde{D}_1^r}(\bar{\omega})$. Hence, since $b(\bar{\omega}) > 0$ by A2, the third term in (8), LC_1 , is the same for \tilde{D}_1^r and D_1^r . Moreover, by (85) and due to A1 and A2, the two first terms in (8), J_1^I and J_2^I , are larger or equal for \tilde{D}_1^r than for D_1^r . This implies the claim.

Suppose next that $\bar{x} \notin [x_A(\underline{\omega}), x_A(\bar{\omega})]$ (and hence $\bar{x} < x_A(\underline{\omega})$). With the same argument as in the previous paragraph, we can infer that $V_1(\{\bar{x}\}) \geq V_1(D_1^r)$. It is therefore sufficient to show that $V_1(\tilde{D}_1^r) = V_1(\{x_A(\underline{\omega})\}) > V_1(\{\bar{x}\})$. Indeed, note that the principal's expected utility from a singleton $\{x_s\}$ is

$$V_1(\{x_s\}) = \int_{\underline{\omega}}^{\bar{\omega}} [\omega x_s + a(x_s) - b(\omega)x_s] f_1(\omega) d\omega + C, \quad (86)$$

where C does not depend on x_s . Now for $x_s < x_A(\underline{\omega})$, we obtain for the derivative:

$$\frac{\partial V_1(\{x_s\})}{\partial x_s} = \int_{\underline{\omega}}^{\bar{\omega}} [\omega - b(\omega)] f_1(\omega) d\omega + a'(x_s) > \underline{\omega} + a'(x_A(\underline{\omega})) = 0, \quad (87)$$

where the inequality follows from A3 and the strict concavity of a , and the final equality from the agent's first order condition $\omega + a'(x_A) = 0$. Hence, $V_1(\{x_s\})$ is increasing in x_s when $x_s \leq x_A(\underline{\omega})$. Since $\bar{x} < x_A(\underline{\omega})$, we get $V_1(\{x_A(\underline{\omega})\}) > V_1(\{\bar{x}\})$, as desired.

Proof of (b). Suppose to the contrary that $U_1(D_1^r) > U_1(D_2^r)$ for all solutions to \mathcal{R} that satisfy (a). Note first that then $\hat{\omega} \leq \omega_1^0$, as otherwise, if $\hat{\omega} > \omega_1^0$, the principal could slightly decrease the upper endpoint $x_A(\hat{\omega})$ of D_1^r without violating $IC_{1,2}$. Indeed, this would be profitable, since by part (b) of Lemma A.1: $\partial V_1([x_A(\underline{\omega}), x_A(\hat{\omega})]) / \partial \hat{\omega} = \Gamma_1(\hat{\omega}) x_A'(\hat{\omega}) < 0$, where the inequality follows from $\hat{\omega} > \omega_1^0$ and part (a) of Lemma A.1.

Second, we show that there is solution to \mathcal{R} with $D_2^r \subseteq [x_A(\underline{\omega}), x_A(\hat{\omega})]$. Otherwise, since $D_1^r = [x_A(\underline{\omega}), x_A(\hat{\omega})]$ and $U_1(D_1^r) > U_1(D_2^r)$, the set D_2^r displays a "gap", that is, there are $y, z \in D_2^r$ with $y < z$ so that $(y, z) \cap D_2^r = \emptyset$. Then, for $\varepsilon > 0$ sufficiently small, the principal could add the interval of actions $(y, y + \varepsilon)$ to D_2^r without violating $IC_{1,2}$. Moreover, by (8), such a change would increase the principal's expected utility conditional on facing type 2 (since it leaves LC_2 unaffected and (weakly) increases J_2^I and J_2^I). This shows that there is a solution with $D_2^r \subseteq [x_A(\underline{\omega}), x_A(\hat{\omega})]$.

Finally, we show that there is a solution with $D_2^r = [x_A(\underline{\omega}), x_A(\hat{\omega})] = D_1^r$ which then contradicts the initial assumption that $U_1(D_1^r) > U_1(D_2^r)$ for all solutions. Indeed, otherwise, the previous paragraph implies that a solution to \mathcal{R} exists with $D_2^r \subset [x_A(\underline{\omega}), x_A(\hat{\omega})]$. Let $x_A(\hat{\omega}) = \max D_2^r$, $\hat{\omega} \leq \hat{\omega}$. We argue that it would then be profitable to replace D_2^r by $\hat{D}_2^r = [x_A(\underline{\omega}), x_A(\hat{\omega})]$. Indeed, this operation (trivially) maintains $IC_{2,1}$. To see that it also improves the principal's expected utility conditional on type 2, we can use the same arguments as in the proof of part (a) to show that for the interval $\hat{D}_2^r = [x_A(\underline{\omega}), x_A(\hat{\omega})]$, we

have $V_2(\hat{D}_2^r) \geq V_2(D_2^r)$. Now, if $\hat{\omega} = \omega$, then $\hat{D}_2^r = D_2^r$, and the claim is shown. If $\hat{\omega} < \omega$, then by part (b) of Lemma A.1, we have that $\partial V_2([x_A(\underline{\omega}), x_A(\hat{\omega})])/\partial \hat{\omega} = \Gamma_2(\hat{\omega})x'_A(\hat{\omega}) > 0$, where the inequality follows since $\hat{\omega} < \omega \leq \omega_1^0 < \omega_2^0$ and part (a) of Lemma A.1. (Recall that $\hat{\omega} \leq \omega_1^0$ by the first paragraph of the proof of part (b).) This implies that $V_2(\hat{D}_2^r) \geq V_2(D_2^r)$ and is what we wanted to show.

Proof of (c). Consider a solution (D_1^r, D_2^r) to \mathcal{R} that satisfies (a) and (b). Then D_2^r is an upward shift of D_1^r . Therefore, since $U_1(D_1^r) = U_1(D_2^r)$ by part (b), and since f_2/f_1 is increasing by assumption, part (a) of Lemma 3 implies the claim. \square

Proof of Proposition 4. We proceed in two steps.

Step 1. We show that there is an optimal delegation menu that satisfies (24) and where D_2^* contains no action $x_H > x_A(\bar{\omega})$. Assume to the contrary that D_2^* contains some action $x_H > x_A(\bar{\omega})$ in every optimal menu. In this case, because redundant actions are removed by convention, the agent selects x_H from D_2^* in state $\bar{\omega}$, and we can write $D_2^* = D'_2 \cup \{x_H\}$ for $D'_2 = D_2^* \cap (-\infty, x_A(\bar{\omega}))$. Let $y = \max D'_2 < x_A(\bar{\omega})$. Our assumptions on U imply that there is a unique action $\bar{x} < x_A(\bar{\omega})$ so that $U(\bar{\omega}, \bar{x}) = U(\bar{\omega}, x_H)$, and since x_H is chosen over y in state $\bar{\omega}$, we have $y < \bar{x}$.

Let us define $\psi(z) = U_1(D'_2 \cup \{z\})$. As shown in the proof of part (a) of Lemma 8 By construction, agent type 1 prefers the delegation set $D'_2 \cup \{\bar{x}\}$ over D_2^* , and prefers D_2^* over D'_2 : $\psi(\bar{x}) = U_1(D'_2 \cup \{\bar{x}\}) \geq U_1(D_2^*) \geq U_1(D'_2) = \psi(y)$. Since ψ is continuous, it follows from the intermediate value theorem that there is an action $\tilde{y} \in [y, \bar{x}]$ so that

$$U_1(D'_2 \cup \{\tilde{y}\}) = \psi(\tilde{y}) = U_1(D_2^*). \quad (88)$$

We now argue that the modification is feasible and (weakly) profitable for the principal. Indeed, the modification is incentive compatible, as $IC_{1,2}$ remains binding by construction. Also, $IC_{2,1}$ is satisfied by part (a) of Lemma 3, because $D'_2 \cup \{\tilde{y}\}$ is an upward shift of D_1^* (since $IC_{1,2}$ is binding and D_1^* is an interval, truncated from above; see footnote 28) and f_2/f_1 is increasing.

We now show that replacing D_2^* by $D'_2 \cup \{\tilde{y}\}$ is (weakly) beneficial by showing that

$$\Delta \equiv V_2(D'_2 \cup \{\tilde{y}\}) - V_2(D_2^*) = [V_2(D'_2 \cup \{\tilde{y}\}) - V_2(D'_2)] - [V_2(D_2^*) - V_2(D'_2)] \geq 0. \quad (89)$$

To compute Δ , let $\tilde{\theta}$ be the unique state so that the agent is indifferent between y and \tilde{y} :

$$\tilde{\theta}y + a(y) = \tilde{\theta}\tilde{y} + a(\tilde{y}). \quad (90)$$

Note that for $\omega \leq \tilde{\theta}$, the agent chooses the same action from $D'_2 \cup \{\tilde{y}\}$ and D'_2 , and for

$\omega > \tilde{\theta}$, he chooses action \tilde{y} from $D'_2 \cup \{\tilde{y}\}$, while he chooses y from D'_2 . Hence, by (8),

$$\begin{aligned}
V_2(D'_2 \cup \{\tilde{y}\}) - V_2(D'_2) &= \int_{\tilde{\theta}}^{\bar{\omega}} [\omega\tilde{y} + a(\tilde{y}) - (\omega y + a(y))][f_2(\omega) + (b(\omega)f_2(\omega)')] d\omega \\
&\quad - [\bar{\omega}\tilde{y} + a(\tilde{y}) - (\bar{\omega}y + a(y))]b(\bar{\omega})f_2(\bar{\omega}) \\
&= (\tilde{y} - y) \int_{\tilde{\theta}}^{\bar{\omega}} [\omega - \tilde{\theta}][f_2(\omega) + (b(\omega)f_2(\omega)')] d\omega \\
&\quad - (\tilde{y} - y)[\bar{\omega} - \tilde{\theta}]b(\bar{\omega})f_2(\bar{\omega}) \\
&= (\tilde{y} - y) \int_{\tilde{\theta}}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega)f_2(\omega) d\omega, \tag{91}
\end{aligned}$$

where the second equality follows from (90) and the final equality from integration by parts. Likewise, since $D_2^* = D'_2 \cup \{x_H\}$,

$$V_2(D_2^*) - V_2(D'_2) = (x_H - y) \int_{\theta_H}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega)f_2(\omega) d\omega, \tag{92}$$

where θ_H is the state in which the agent is indifferent between y and x_H . Hence,

$$\Delta = (\tilde{y} - y) \int_{\tilde{\theta}}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega)f_2(\omega) d\omega - (x_H - y) \int_{\theta_H}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega)f_2(\omega) d\omega. \tag{93}$$

Agent type 1's utility differences $U_1(D'_2 \cup \{\tilde{y}\}) - U_1(D'_2)$ and $U_1(D_2^*) - U_1(D'_2)$ can be computed with analogous steps, and by (88) we obtain:

$$(\tilde{y} - y) \int_{\tilde{\theta}}^{\bar{\omega}} 1 - F_1(\omega) d\omega = (x_H - y) \int_{\theta_H}^{\bar{\omega}} 1 - F_1(\omega) d\omega. \tag{94}$$

Plugging this into (93) delivers

$$\begin{aligned}
\Delta &= (x_H - y) \int_{\theta_H}^{\bar{\omega}} 1 - F_1(\omega) d\omega \times \\
&\quad \times \left[\frac{\int_{\tilde{\theta}}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega)f_2(\omega) d\omega}{\int_{\tilde{\theta}}^{\bar{\omega}} 1 - F_1(\omega) d\omega} - \frac{\int_{\theta_H}^{\bar{\omega}} 1 - F_2(\omega) - b(\omega)f_2(\omega) d\omega}{\int_{\theta_H}^{\bar{\omega}} 1 - F_1(\omega) d\omega} \right]. \tag{95}
\end{aligned}$$

Since $y < x_H$ and $\tilde{\theta} < \theta_H$, this expression is non-negative because $R_{2,1}$ is decreasing by assumption, as desired.

Step 2. We show that there is an optimal delegation menu where D_2^* is of the form (a) or (b) in the lemma. Assume to the contrary that every solution (D_1^*, D_2^*) to \mathcal{P} which satisfies (24) violates conditions (a) and (b). We derive a contradiction by constructing a solution (D_1^*, \tilde{D}_2^*) to \mathcal{P} that satisfies (a) or (b).

Because $D_2^* \cap (x_A(\bar{\omega}), \infty) = \emptyset$ by Step 1, if D_2^* violates (a) and (b), then it holds:

(i) $D_2^* \cap [x_A(\underline{\omega}), x_A(\bar{\omega})] = \emptyset$; or

(ii) $\hat{z} = \max D_2^* < x_A(\bar{\omega})$, and there is at least one other action in $D_2^* \cap [x_A(\underline{\omega}), x_A(\bar{\omega})]$.

In case (i), $IC_{2,1}$ would clearly be violated since D_1^* satisfies (24). In case (ii), let $\hat{y} = \min(D_2^* \cap [x_A(\underline{\omega}), x_A(\bar{\omega})]) < \hat{z}$. Since D_2^* is not of the form (a) and (b), we have

$$D_2^* \cap [x_A(\underline{\omega}), x_A(\bar{\omega})] \subset [\hat{y}, \hat{z}]. \quad (96)$$

We now construct a solution (D_1^*, \tilde{D}_2^*) to \mathcal{P} so that \tilde{D}_2^* is of the form (b) in the lemma. If D_2^* contains an action smaller than $x_A(\underline{\omega})$, call this action x_L . Otherwise, let x_L be the action smaller than $x_A(\underline{\omega})$ so that type $\underline{\omega}$ is indifferent between x_L and \hat{y} . (Because of our assumptions on U such an action is uniquely defined.) Add this action to D_2^* . Because, by construction, this does not affect the agent's expected utility from the delegation set, we abuse notation and call the resulting set again D_2^* . Now define

$$\hat{D}_2 = \{x_L\} \cup [\hat{y}, \hat{z}]; \quad \hat{\hat{D}}_2 = \{x_L\} \cup \{\hat{z}\}. \quad (97)$$

Then, by construction, $U_1(\hat{D}_2) \geq U_1(D_2^*) \geq U_1(\hat{\hat{D}}_2)$. Therefore, since the agent's utility is continuous, by the intermediate value theorem there is a $y \in (\hat{y}, \hat{z})$ so that $U_1(D_2^*) = U_1(\{x_L\} \cup [y, \hat{z}])$. Now, let

$$\tilde{D}_2^* = \{x_L\} \cup [y, \hat{z}]. \quad (98)$$

We now argue that the principal (weakly) benefits from replacing D_2^* by \tilde{D}_2^* . Indeed, the same argument as after (88) work to show that the modification is incentive compatible.

To see that it is (weakly) profitable, observe that, by construction, \tilde{D}_2^* is an upward shift of D_2^* that satisfies $u_{\tilde{D}_2^*}(\underline{\omega}) = u_{D_2^*}(\underline{\omega})$ and $u_{\tilde{D}_2^*}(\bar{\omega}) = u_{D_2^*}(\bar{\omega})$. The latter implies that $J_2^I(\tilde{D}_2^*) = J_2^I(D_2^*)$ and $LC_2(\tilde{D}_2^*) = LC_2(D_2^*)$. Moreover, since $U_1(D_1^*) = U_1(D_2^*) = U_1(\tilde{D}_2^*)$ and since $\rho_{2,1}$ is increasing by assumption, the upward shift property implies that $J_2^{II}(\tilde{D}_2^*) \geq J_2^{II}(D_2^*)$ by Lemma 5(a). Therefore, by (8), the modification weakly improves the principal's utility. Hence, (D_1^*, \tilde{D}_2^*) is a solution to \mathcal{P} . As \tilde{D}_2^* is of the form (b) in the lemma, we obtain a contradiction. \square

Proofs for Section 6

Proof of Lemma 9. Part (a) follows from the main text. As to part (b), in light of the arguments after the statement of the lemma, it is sufficient to verify that condition (23) is violated, which now becomes

$$-\frac{1 - F(\omega^{st})}{\int_{\omega^{st}}^{\bar{\omega}} 1 - F(\tilde{\omega}) d\tilde{\omega}} > \min_{\eta \leq \omega^{st}} \frac{f'(\eta)}{f(\eta)}. \quad (99)$$

In fact, recall from (9) and Lemma 1 that, with constant bias and $F_i = F$,

$$\frac{\int_{\omega^{st}}^{\bar{\omega}} 1 - F(\tilde{\omega}) d\tilde{\omega}}{1 - F(\omega^{st})} = \mu_1 b_1 + \mu_2 b_2. \quad (100)$$

But, Assumption A1 implies that $f'(\omega)/f(\omega) > -1/b_i$ for $i = 1, 2$ and $\omega \in [\underline{\omega}, \bar{\omega}]$. Hence, $f'(\omega)/f(\omega) > -1/\max\{b_1, b_2\} > -1/(\mu_1 b_1 + \mu_2 b_2)$, which together with (100) yields a contradiction to (99). \square

References

- [1] Alonso, R. and N. Matouscheck, 2008. Optimal Delegation. *Rev. Econ. Stud.* 75 (1), 259–293.
- [2] Alonso, R., W. Dessein, and N. Matouschek, 2008. When Does Coordination Require Centralization? *Amer. Econ. Rev.* 98 (1), 145–179.
- [3] Amador, M. and K. Bagwell, 2013. The Theory of Optimal Delegation with An Application to Tariff Caps. *Econometrica* 81 (4), 1541-1599.
- [4] Baron, D. and D. Besanko, 1984. Regulation and Information in a Continuing Relationship. *Info. Econ. Pol.* 1, 267–302.
- [5] Battaglini, M., 2005. Long-Term Contracting with Markovian Consumers. *Amer. Econ. Rev.* 95, 637–658.
- [6] Bester, H., 2009. Externalities, Communication and the Allocation of Decision Rights. *Econ. Theory* 41, 269–296.
- [7] Bester, H. and D. Krähmer, 2014. Exit Options and the Allocation of Authority. Unpublished working paper.
- [8] Courty, P. and H. Li, 2000. Sequential Screening. *Rev. Econ. Stud.* 67, 697–717.
- [9] Dai, C., T. R. Lewis, and G. Lopomo, 2006. Delegating Management to Experts. *RAND J. Econ.* 37, 503–520.
- [10] Dessein, W., 2002. Authority and Communication in Organizations. *Rev. Econ. Stud.* 69, 811–838.
- [11] Esö, P. and B. Szentes, 2007a. The Price of Advise. *RAND J. Econ.* 38, 863–880.
- [12] Esö, P. and B. Szentes, 2007b. Optimal Information Disclosure in Auctions and the Handicap Auction. *Rev. Econ. Stud.* 74, 705–731.

- [13] Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani, 2009. Mediation, Arbitration, and Negotiation. *J. Econ. Theory* 144, 1397–1420.
- [14] Holmström, B., 1977. On Incentives and Control in Organizations. Ph.D. thesis, Stanford University.
- [15] Holmström, B., 1984. On the Theory of Delegation. In: *Bayesian Models in Economic Theory*. Ed. by M. Boyer, and R. Kihlstrom. North-Holland, New York.
- [16] Hoffmann, F. and R. Inderst, 2011. Presale Information. *J. Econ. Theory* 146 (6), 2333–2355.
- [17] Inderst, R. and M. Peitz, 2012. Informing Consumers about their own Preferences. *Int. J. Ind. Organ.* 30, 417–428.
- [18] Ivanov, M., 2010. Informational Control and Organizational Design. *J. Econ. Theory* 145 (2), 721–751.
- [19] Koessler, F., D. Martimort, 2012. Optimal Delegation with Multi-dimensional Decisions. *J. Econ. Theory* 147 (5), 1850–1881.
- [20] Kolotilin, A., H. Li, and W. Li, 2013. Optimal Limited Authority for Principal. *J. Econ. Theory* 148 (6), 2344–2382.
- [21] Kováč, E. and T. Mylovanov, 2009. Stochastic Mechanisms in Settings without Monetary Transfers: The Regular Case. *J. Econ. Theory* 144 (4), 1373–1395.
- [22] Kováč, E. and D. Krähmer, 2014. Optimal Sequential Delegation. Unpublished working paper, University of Bonn.
- [23] Krähmer, D., 2006. Message-Contingent Delegation. *J. Econ. Behav. Organ.* 60, 490–506.
- [24] 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.
- [25] Krähmer, D. and R. Strausz, 2011. Optimal Procurement Contracts with Pre-Project Planning. *Rev. Econ. Stud.* 78, 1015–1041.
- [26] Krähmer, D. and R. Strausz, 2015a. Optimal Sales Contracts with Withdrawal Rights. *Rev. Econ. Stud.* 82, 762–790.
- [27] Krähmer and Strausz, 2015b. Ex Post Information Rents in Sequential Screening. *Games Econ. Behav.* 90, 257–273.

- [28] Liang, P., 2013. Optimal Delegation via a Strategic Intermediary. *Games Econ. Behav.* 82, 15–30.
- [29] Luenberger, F., 1969. *Optimization by Vector Space Methods*. John Wiley & Sons: New York.
- [30] Martimort, D. and A. Semenov, 2006. Continuity in Mechanism Design without Transfers. *Econ. Letters* 93, 182–189.
- [31] Melumad, N. D., Shibano, T., 1991. Communication in Settings with no Transfers. *RAND J. Econ.* 22 (2), 173–198.
- [32] Myerson, R., 1986. Multistage Games with Communication. *Econometrica* 54, 323–358.
- [33] Mylovanov, T., 2008. Veto-based Delegation. *J. Econ. Theory* 138, 297–307.
- [34] Nocke, V., M. Peitz, and F. Rosar, 2011. Advance-purchase Discounts as a Price Discrimination Device. *J. Econ. Theory* 146, 141–162.
- [35] Pavan, A., I. Segal, and J. Toikka, 2014. Dynamic Mechanism Design: A Myersonian Approach. *Econometrica* 82 (2), 601–653.
- [36] Riordan, M. H. and D. E. Sappington, 1987. Information, Incentives and Organizational Mode. *Quart. J. Econ.* 102, 243–264.
- [37] Samuelson, W., 1984. Bargaining under Asymmetric Information. *Econometrica* 52, 995–1005.
- [38] Sappington, D., 2002. Price regulation. In: *The Handbook of Telecommunications Economics, Volume I: Structure, Regulation, and Competition*. Ed. by Cave M., S. Majumdar, and I. Vogelsang. Amsterdam: Elsevier Science Publishers, 225–293.
- [39] Semenov, A. P., 2008. Bargaining in the Appointment Process, Constrained Delegation and the Political Weight of the Senate. *Public Choice* 136, 165–180.
- [40] Semenov, A. P., 2014. Delegation to a Potentially Uninformed Agent. Unpublished working paper, Ottawa University.
- [41] Shaked, M. and J.G. Shanthikumar, 2007. *Stochastic Orders*. Springer: New York.
- [42] Szalay, D., 2005. The Economics of Clear Advice and Extreme Options. *Rev. Econ. Stud.* 72, 1173–1198.
- [43] Tanner, N., 2014. Optimal Delegation under Uncertain Bias. Unpublished working paper, Yale University.