PERSUASION MEETS DELEGATION

ANTON KOLOTILIN
School of Economics, UNSW Business School

ANDRIY ZAPECHELNYUK
School of Economics, University of Edinburgh

A principal can restrict an agent's information (the persuasion problem) or discretion (the delegation problem). We study these two problems under standard single-crossing assumptions on the agent's marginal utility. We show that these problems are equivalent on the set of monotone stochastic mechanisms, implying, in particular, the equivalence of deterministic delegation and monotone partitional persuasion. We also show that the monotonicity restriction is superfluous for linear persuasion and linear delegation, implying their equivalence on the set of all stochastic mechanisms. Finally, using tools from the persuasion literature, we characterize optimal delegation mechanisms, thereby generalizing and extending existing results in the delegation literature.

KEYWORDS: Persuasion, delegation, discriminatory disclosure.

1. INTRODUCTION

A principal has two ways to influence decisions of an agent: delegation and persuasion. The delegation literature, initiated by Holmström (1977, 1984), studies the design of decision rules, with applications to organizational decision processes (Dessein, 2002), monopoly regulation policies (Alonso and Matouschek, 2008), and international trade agreements (Amador and Bagwell, 2013). The persuasion literature, set in motion by Kamenica and Gentzkow (2011), studies the design of information disclosure rules, with applications to grade disclosure policies (Ostrovsky and Schwarz, 2010), internet advertising strategies (Rayo and Segal, 2010), and forensic tests (Kamenica and Gentzkow, 2011).

This paper shows that, under standard assumptions, the delegation and persuasion problems are equivalent, thereby bridging the two strands of literature. The implication is that the existing insights and results in one problem can be used to understand and solve the other problem. To connect delegation and persuasion, we introduce a third problem, called *discriminatory disclosure*. In general, a discriminatory disclosure problem is less constrained than a persuasion problem and more constrained than a delegation problem. Under standard assumptions, all three problems are equally constrained and thus equivalent.

Persuasion, delegation, and discriminatory disclosure problems describe interactions between a principal (she) and an agent (he). In persuasion, utilities depend on the agent's decision

Anton Kolotilin: akolotilin@gmail.com
Andriy Zapechelnyuk: azapech@gmail.com

We are grateful to Tymofiy Mylovanov, with whom we are working on related projects. We thank Ricardo Alonso, Kyle Bagwell, Benjamin Brooks, Deniz Dizdar, Piotr Dworczak, Alexander Frankel, Drew Fudenberg, Gabriele Gratton, Yingni Guo, Bengt Holmström, Emir Kamenica, Navin Kartik, Ming Li, Hongyi Li, Simon Loertscher, Carlos Oyarzun, Alessandro Pavan, Eric Rasmusen, Philip Reny, Andrzej Skrzypacz, Joel Sobel, Bruno Strulovici, and Thomas Tröger for helpful comments and suggestions. We also thank participants at various seminars and conferences. Part of this research was carried out while Anton Kolotilin was visiting MIT Sloan School of Management, whose hospitality is greatly appreciated. Anton Kolotilin gratefully acknowledges support from the Australian Research Council Discovery Early Career Research Award DE160100964 and from MIT Sloan's Program on Innovation in Markets and Organizations. Andriy Zapechelnyuk gratefully acknowledges support from the Economic and Social Research Council Grant ES/N01829X/1.

and a state of the world. The principal designs a test that generates informative messages about the state. The agent observes a message and chooses a decision. In delegation, utilities depend on the agent's decision and his private type. The principal designs a menu of decisions or, more generally, a menu of lotteries over decisions. The agent chooses a lottery from the menu. In discriminatory disclosure, utilities depend on the state, the agent's private type, and the agent's binary action, 0 or 1. The principal designs a menu of tests. The agent chooses a test from the menu, observes a message from the chosen test, and chooses action 0 or 1. We assume that the sets of decisions, states, and types are intervals of the real line, and that the agent's utility function satisfies standard single-crossing assumptions.

Appealing to the revelation principle, we restrict attention to direct mechanisms in the three problems. Our general equivalence result holds for monotone stochastic mechanisms, which have the following interpretation. In persuasion, a higher state generates a higher lottery over recommended decisions with respect to first-order stochastic dominance. In delegation, a higher reported type is assigned a higher lottery over decisions with respect to first-order stochastic dominance. In discriminatory disclosure, action 1 is recommended with a higher probability when the state is higher and the reported type is lower.

For each primitive of one problem, our equivalence result explicitly constructs an equivalent primitive of the other two problems. Up to normalization, this construction equates the marginal utilities in persuasion and delegation with the utilities in discriminatory disclosure. Moreover, the agent's type in delegation and discriminatory disclosure becomes the decision in persuasion, and the state in persuasion and discriminatory disclosure becomes the decision in delegation. Intuitively, decisions in delegation and states in persuasion play the same role because the principal controls discretion over decisions in delegation and information about states in persuasion.

To sketch the intuition for the equivalence, consider a monotone mechanism in discriminatory disclosure. On the one hand, this mechanism can be represented as a cutoff-state mechanism. For each reported type of the agent, a cutoff is drawn from a lottery, and then the agent is recommended action 1 whenever the state is above the cutoff. Describing these lotteries over cutoff states as lotteries over decisions, we obtain a delegation problem. On the other hand, this mechanism can be represented as a cutoff-type mechanism. For each state, a cutoff is drawn from a lottery, and then the agent is recommended action 1 whenever his reported type is below the cutoff. Describing these lotteries over cutoff types as lotteries over recommended decisions, we obtain a persuasion problem. In general, discriminatory disclosure is more constrained than delegation, because it has an additional obedience constraint that the agent prefers to take the action recommended by the chosen test. Moreover, discriminatory disclosure is less constrained than persuasion, because it allows the principal to design a menu of tests rather than a single test. The most challenging part of our equivalence result is to show that all three problems are in fact equally constrained when the agent's utility satisfies single-crossing assumptions.

Much of the literature studies *linear* problems. In linear persuasion, the marginal utilities are linear in the state (e.g., Gentzkow and Kamenica 2016, Kolotilin 2018, and Dworczak and Martini 2019). Similarly, in linear discriminatory disclosure, the utilities are linear in the state (e.g., Kolotilin et al. 2017, Bergemann and Morris 2019, and Candogan and Strack 2023). Finally, in linear delegation, the marginal utilities are linear in the decision (e.g., Alonso and Matouschek 2008, Kováč and Mylovanov 2009, and Amador and Bagwell 2013). By extending Strassen's theorem to include an additional monotone likelihood ratio property, we show that the monotonicity restriction on mechanisms is without loss of generality in linear problems, and thus our equivalence result holds for all stochastic mechanisms.[1] Using tools from the

[1]Recently, Kleiner et al. (2021) show a connection between delegation with quadratic utilities (a special case of linear delegation) and linear persuasion. See Section 5.4 for a detailed discussion.

persuasion literature, we provide necessary and sufficient conditions for the optimality of a candidate delegation mechanism. For familiar threshold delegation mechanisms, our conditions coincide with those in Alonso and Matouschek (2008) and Amador and Bagwell (2013), but we impose weaker differentiability assumptions.

The literature also studies *monotone deterministic* problems, where the principal designs a monotone partition of the state space in persuasion, a delegation set of decisions in delegation, and a menu of deterministic cutoff tests in discriminatory disclosure. The equivalence of these problems immediately follows from our main result. Using standard conditions for the optimality of full disclosure in persuasion, we provide novel conditions for the optimality of full discretion in delegation, subsuming existing conditions. Furthermore, by translating a tractable nonlinear setting from persuasion (Rayo 2013, Onuchic and Ray 2023) to delegation, we derive new necessary and sufficient conditions for the optimality of a candidate delegation set in this setting.

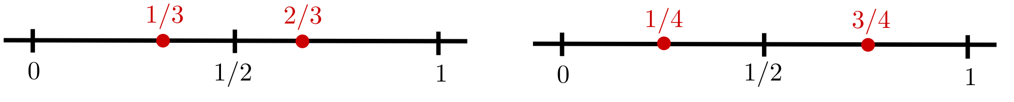
Recent literature considers a delegation problem where the agent has an outside option (Zapechelnyuk 2020, Kartik et al. 2021, Amador and Bagwell 2022, Saran 2024), whereas there is no such constraint in the standard delegation problem (Holmström 1977, 1984, Alonso and Matouschek 2008, and Amador and Bagwell 2013). Under natural Inada-type conditions on the utilities, our results apply to standard delegation and delegation with outside option, in both linear and monotone deterministic cases.

To illustrate our results, we solve a classical monopoly regulation problem in which a welfare-maximizing regulator (principal) restricts production choices of a monopolist (agent) who privately knows his cost. This problem is studied by Baron and Myerson (1982) as a mechanism design problem with transfers and by Alonso and Matouschek (2008) as a delegation problem without transfers. Amador and Bagwell (2022) further extend the analysis by including the monopolist's participation constraint.[2] We provide novel and simple conditions for the optimality of a price cap among stochastic or deterministic mechanisms.

2. EXAMPLE

Before presenting our formal setting and results, we illustrate the equivalence between persuasion and delegation in a simple example.

Consider first a delegation problem. In this problem, a principal commits to a set of decisions from which a privately informed agent chooses. The utilities depend on the decision $s \in \mathbb{R}$ and the agent's private type $t \in [0, 1]$ that is uniformly distributed on $[0, 1]$. The principal's utility is $V(s, t)$ and the agent's utility is $U(s, t) = -(s - t)^2$.



(a) Delegation problem (b) Persuasion problem

FIGURE 1.—The agent's choices in the delegation and persuasion problems.

Suppose the principal lets the agent choose one of two decisions, $1/3$ or $2/3$. The agent optimally chooses decision $1/3$ if his type is $t < 1/2$ and decision $2/3$ if his type is $t > 1/2$, as

[2]Applying the results of Halac and Yared (2022), the analysis can be further extended to allow for limited enforcement and money burning.

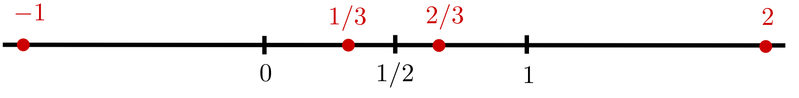
shown in Figure 1(a). The agent's optimal decision as a function of his type is thus

$$s_D^*(t) = \begin{cases} 1/3, & \text{if } t < 1/2, \\ 2/3, & \text{if } t > 1/2. \end{cases}$$

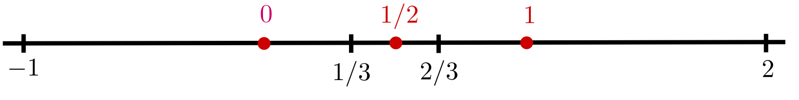
Consider now a persuasion problem. In this problem, the agent is free to choose any decision $s \in \mathbb{R}$ and is initially uninformed about the state $t \in [0,1]$ that is uniformly distributed on $[0,1]$. The principal designs the agent's information about the state. The utilities are the same as in the delegation problem.

Clearly, the principal cannot induce the agent to choose $s_D^*(t)$ for each state t in persuasion. Indeed, if the principal lets the agent know only whether t is below or above $1/2$, then the induced decisions are $1/4$ and $3/4$, as shown in Figure 1(b). Alternatively, the principal can induce the agent to choose only decisions $1/3$ and $2/3$. But each of the decisions $1/3$ and $2/3$ is necessarily induced with positive probabilities for both $t > 1/2$ and $t < 1/2$.

This example illustrates that the instruments of delegation and persuasion work differently in a given environment. Nevertheless, we show that the persuasion and delegation problems are mathematically equivalent. To relate these two problems, we swap the roles of the variables, so the type in delegation is identified with the decision in persuasion, and the state in persuasion is identified with the decision in delegation. We also appropriately associate the utilities in the two problems. For simplicity, in this section, we assume that the agent's utility is $U(s,t) = -(s-t)^2$ in both delegation and persuasion problems, in which case we only need to associate the principal's utilities V_D and V_P.



(a) Delegation problem



(b) Persuasion problem

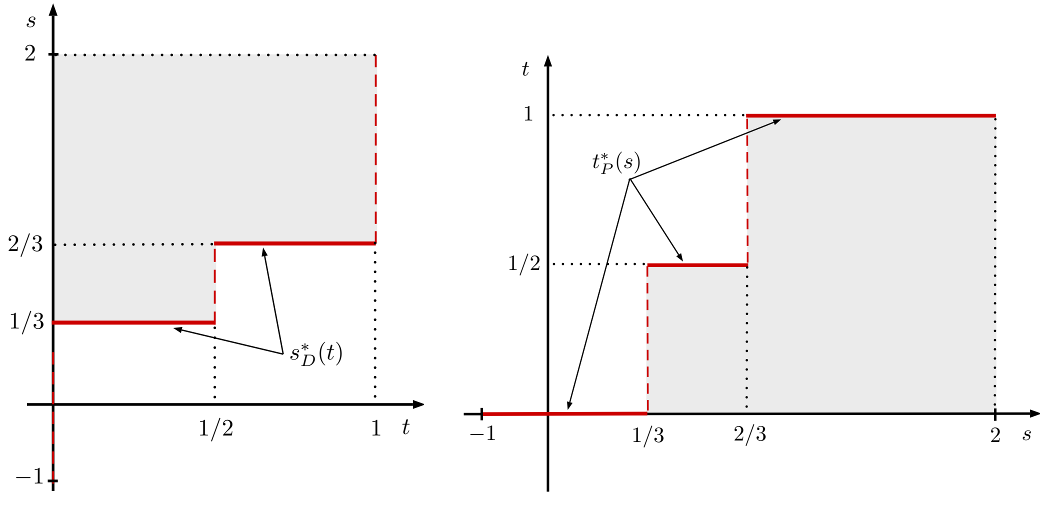
FIGURE 2.—The agent's choices in the equivalent delegation and persuasion problems.

For illustration, we now construct a persuasion problem that is equivalent to the delegation problem in our example. The first step is to restrict the agent's decisions in the delegation problem to an interval. Specifically, we reduce the decision set from $\mathbb{R}$ to the interval $[-1,2]$. The bounds of this interval are chosen arbitrarily, but far enough, so that the boundary decisions (and therefore any decisions outside the interval) are never chosen by the agent. Consider now the set of permitted decisions $S^* = \{-1, 1/3, 2/3, 2\}$, as shown in Figure 2(a). The agent's optimal decision for each type t is $s_D^*(t)$ as in the original problem, because $s = 1/3$ is preferred to $s = -1$ and $s = 2/3$ is preferred to $s = 2$ for each $t \in [0,1]$.

The second step is to swap the roles of s and t. So, in the persuasion problem, the agent chooses decision $t \in [0,1]$, whereas s is a state that is uniformly distributed on $[-1,2]$. Let the agent's information be a monotone partition described by the set of cutoff states $S^* = \{-1, 1/3, 2/3, 2\}$, so the agent knows whether the state is between -1 and $1/3$, between $1/3$ and $2/3$, or between $2/3$ and 2, as shown in Figure 2(b). That is, the permitted decisions in delegation (dots in Figure 2(a)) become the cutoff states in persuasion (vertical bars in Figure 2(b)). The agent's optimal decisions are shown as dots in Figure 2(b). If the agent knows that the state is between $1/3$ and $2/3$, then his optimal decision is the posterior expected state $1/2$. If the agent knows that the state is between -1 and $1/3$, then his optimal decision is 0, because this is the closest decision in $[0,1]$ to the posterior expected state $-1/3$. Similarly, if the agent knows that the state is between $2/3$ and 2, then his optimal decision is 1, because this is the closest decision in $[0,1]$ to the posterior expected state $4/3$. That is, the cutoff types in delegation (vertical bars in Figure 2(a)) become the induced decisions in persuasion (dots in Figure 2(b)). To summarize, the agent's optimal decision as a function of the state is

$$t_P^*(s) = \begin{cases} 0, & \text{if } s < 1/3, \\ 1/2, & \text{if } 1/3 < s < 2/3, \\ 1, & \text{if } s > 2/3. \end{cases}$$

The key observation is that $s_D^*(t)$ and $t_P^*(s)$ are inversely related, as shown in Figure 3.



(a) Delegation problem

(b) Persuasion problem

FIGURE 3.—The agent's decision functions in the equivalent delegation and persuasion problems.

The last step is to associate the utility functions in the two problems. Let

$$v_P(s,t) = \frac{\partial}{\partial t} V_P(s,t) \quad \text{and} \quad v_D(s,t) = -\frac{\partial}{\partial s} V_D(s,t),$$

so that $v_P(s,t)$ and $-v_D(s,t)$ are the principal's marginal utilities in the persuasion and delegation problems. Without loss of generality, we normalize the principal's utilities to zero from decision $t=0$ in persuasion and from decision $s=2$ in delegation, $V_P(0,s)=0$ and $V_D(2,t)=0$, so their expected utilities in the two problems are

$$\mathbb{E}[V_P] = \int_{-1}^{2} \underbrace{\int_{0}^{t_P^*(s)} v_P(s,t)\mathrm{d}t}_{V_P(t_P^*(s),s)} f_P(s)\mathrm{d}s \quad \text{and} \quad \mathbb{E}[V_D] = \int_{0}^{1} \underbrace{\int_{s_D^*(t)}^{2} v_D(s,t)\mathrm{d}s}_{V_D(s_D^*(t),t)} f_D(t)\mathrm{d}t,$$

where $f_D(t) = 1$ is the uniform density on $[0,1]$ and $f_P(s) = 1/3$ is the uniform density on $[-1,2]$. Notice that if the principal's marginal utility in the persuasion problem satisfies

$$v_P(s,t)f_P(s) = v_D(s,t)f_D(t), \tag{1}$$

then we obtain $\mathbb{E}[V_P] = \mathbb{E}[V_D]$, because the areas over which the marginal utilities are integrated (shaded areas in Figure 3) are exactly the same.

We have illustrated that, when the principal's marginal utilities satisfy (1), the delegation problem where the agent chooses from a set of *deterministic* decisions is equivalent to the persuasion problem where the agent's information is a *monotone partition* of the state space. Specifically, it is optimal to induce the agent's decision $s_D^*(t)$ for each type t in delegation if and only if it is optimal to induce the agent's decision $t_P^*(s)$ for each state s in persuasion, where $t_P^*(s)$ is the inverse of $s_D^*(t)$.

This equivalence holds more generally. It extends to monotone stochastic delegation and persuasion problems when the agent's marginal utility is single-crossing. In stochastic delegation, the delegation mechanism is described by a conditional probability function $\pi_D(s|t) = \mathbb{P}(\text{decision} < s|\text{type} = t)$. Similarly, in stochastic persuasion, the persuasion mechanism is described by a conditional probability function $\pi_P(t|s) = \mathbb{P}(\text{decision} > t|\text{state} = s)$. In what follows, we show that, when (1) holds, mechanism π_D is optimal in the delegation problem if and only if mechanism π_P is optimal in the persuasion problem, where $\pi_P(t|s) = \pi_D(s|t)$.[3] To connect delegation and persuasion, we introduce a third problem, called *discriminatory disclosure*, and show that all three problems are equivalent.

3. THREE PROBLEMS

This section introduces three principal-agent problems: a persuasion problem, a delegation problem, and a discriminatory disclosure problem, which are labeled by letters P, D, and I in the notation. To simplify the exposition, all functions in the paper are assumed to be bounded, left-continuous in variables labeled s and y, and right-continuous in variables labeled t and x.

3.1. *Persuasion Problem.*

The agent's utility $U_P(s,t)$ and principal's utility $V_P(s,t)$ depend on the agent's decision $t \in T = [0,1]$ and the state of the world $s \in S = [0,1]$, with the boundary conditions $U_P(s,0) = 0$ and $V_P(s,0) = 0$ for all $s \in S$. The state is uniformly distributed. The only substantive assumptions here are that the decision and state are one-dimensional.[4] We assume that

[3] In our example, the agent's decisions $s_D^*(t)$ and $t_P^*(s)$ are deterministic and can be expressed by $\pi_D(s|t)$ and $\pi_P(t|s)$ that take values 0 or 1. Specifically, $\pi_D(s|t) = 1$ in the shaded area in Figure 3(a) and $\pi_P(t|s) = 1$ in the shaded area in Figure 3(b). When $t_P^*(s)$ is the inverse of $s_D^*(t)$, the shaded areas coincide, so $\pi_P(t|s) = \pi_D(s|t)$.

[4] Suppose that the agent's utility $U(y,x)$ and principal's utility $V(y,x)$ depend on decision $x \in [\underline{x}, \overline{x}]$ and state $y \in [\underline{y}, \overline{y}]$, where y has a distribution $F(y) = \mathbb{P}(\text{state} < y)$. Let F^{-1} be the generalized inverse of F. By the Skorokhod

utilities $U_P(s,t)$ and $V_P(s,t)$ are absolutely continuous in decision t, so that

$$U_P(s,t)=\int_0^t u_P(s,\tilde{t})\mathrm{d}\tilde{t} \quad \text{and} \quad V_P(s,t)=\int_0^t v_P(s,\tilde{t})\mathrm{d}\tilde{t} \tag{2}$$

where $u_P(s,t)$ and $v_P(s,t)$ are marginal utilities. Given our normalizations, a pair (u_P,v_P) fully describes the problem.

The principal and agent are initially uninformed about the state. The principal designs a test that generates informative messages about the state. The agent observes a message, updates his beliefs about the state, and chooses a decision. By the revelation principle argument, we can assume that these messages are decision recommendations. That is, the principal chooses a *persuasion mechanism* $\pi_P(t|s)$ that provides a stochastic decision recommendation conditional on each state,

$$\pi_P(t|s)=\mathbb{P}\left(\text{decision}>t|\text{state}=s\right).$$

We frequently use Bayes' rule which states that, for all functions $w(s,t)$, we have

$$\int_{S\times T} w(s,t)(-\pi_P(\mathrm{d}t|s))\mathrm{d}s=\int_{T\times S} w(s,t)\pi_P(\mathrm{d}s|t)(-\pi_P(\mathrm{d}t)), \tag{3}$$

where, with abuse of notation, $\pi_P(t)=\mathbb{P}(\text{decision}>t)$ is the marginal probability and $\pi_P(s|t)=\mathbb{P}(\text{state}<s|\text{decision}=t)$ is (a version of) the conditional probability induced by the uniform distribution of s and the conditional probability $\pi_P(t|s)$. We write negative signs in (3) because $\pi_P(t)$ and $\pi_P(t|s)$ are decreasing in t.[5]

The key constraint on the persuasion mechanism is that the agent prefers to choose a recommended decision given his beliefs induced by this recommendation. This *incentive-compatibility constraint* is

$$\int_S U_P(s,t)\pi_P(\mathrm{d}s|t)\geq\int_S U_P(s,\hat{t})\pi_P(\mathrm{d}s|t), \quad \text{for all } \hat{t}\in T \text{ and } \pi_P\text{-almost all } t\in T. \tag{IC$_P$}$$

The agent may have profitable deviations for a π_P-negligible set of recommendations.

The principal chooses a persuasion mechanism π_P to

$$\text{maximize} \quad W_P(\pi_P)=\int_{S\times T} V_P(s,t)(-\pi_P(\mathrm{d}t|s))\mathrm{d}s \quad \text{subject to (IC}_P\text{)}.$$

3.2. *Delegation Problem.*

The agent's utility $U_D(s,t)$ and principal's utility $V_D(s,t)$ depend on the agent's decision $s\in S=[0,1]$ and the agent's private type $t\in T=[0,1]$, with the boundary conditions $U_D(1,t)=0$

---

representation, if s is uniformly distributed on $[0,1]$, then $y=F^{-1}(s)$ has distribution F. To obtain our setting, let $t=(x-\underline{x})/(\overline{x}-\underline{x})$, let s be uniformly distributed, let $U_P(s,t)=U(F^{-1}(s),\underline{x}+(\overline{x}-\underline{x})t)-U(F^{-1}(s),\underline{x})$, and let $V_P(s,t)=V(F^{-1}(s),\underline{x}+(\overline{x}-\underline{x})t)-V(F^{-1}(s),\underline{x})$. Kolotilin and Zapechelnyuk (2019, Section 6.3) illustrate this change of variables in the prosecutor-judge example of Kamenica and Gentzkow (2011) where F is binary.

[5]For illustration, we have $\mathbb{P}\left(\text{decision}\in(t_1,t_2]|\text{state}=s\right)=\pi_P(t_1|s)-\pi_P(t_2|s)=\int_{(t_1,t_2]}(-\pi_P(\mathrm{d}t|s))$ for all $s\in S$ and all $t_1,t_2\in T$ such that $t_1<t_2$.

and $V_D(1,t) = 0$ for all $t \in T$. The type is uniformly distributed.[6] We assume that utilities $U_D(s,t)$ and $V_D(s,t)$ are absolutely continuous in decision s, so that

$$U_D(s,t) = \int_s^1 u_D(\tilde{s},t)\mathrm{d}\tilde{s} \quad \text{and} \quad V_D(s,t) = \int_s^1 v_D(\tilde{s},t)\mathrm{d}\tilde{s} \tag{4}$$

where $-u_D(s,t)$ and $-v_D(s,t)$ are marginal utilities. Given our normalizations, a pair (u_D, v_D) fully describes the problem.

We consider a delegation problem where the agent can always choose extreme decisions $s = 0$ and $s = 1$. As suggested by the example in Section 2 and shown formally in Section 5.2 and Appendix A, this assumption is typically non-binding when the agent's and principal's utilities are defined on a sufficiently large interval of decisions, so that the extreme decisions are never chosen. Moreover, this assumption allows to incorporate additional constraints, such as the agent's participation constraint.

Formally, the principal designs a menu of lotteries over decisions which must contain the two degenerate lotteries that assign probability one to decisions $s = 0$ and $s = 1$. The agent privately observes his type and chooses a lottery from the menu. By the revelation principle argument, we can label each lottery in the menu by the type of the agent who is recommended to choose this lottery. That is, the principal chooses a *delegation mechanism* $\pi_D(s|t)$ that assigns to the agent's reported type a lottery over decisions,

$$\pi_D(s|t) = \mathbb{P}\left(\text{decision} < s | \text{type} = t\right).$$

The key constraint on the delegation mechanism is that the agent prefers to choose a lottery assigned to his type rather than the lottery assigned to any other type, or decisions $s = 0$ and $s = 1$. This *incentive-compatibility constraint* is

$$\int_S U_D(s,t)\pi_D(\mathrm{d}s|t) \geq \max\left\{\int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t}),\, U_D(0,t),\, U_D(1,t)\right\}, \tag{IC$_D$}$$
$$\text{for all } \hat{t} \in T \text{ and almost all } t \in T.$$

The principal chooses a delegation mechanism π_D to

$$\text{maximize} \quad W_D(\pi_D) = \int_{T\times S} V_D(s,t)\pi_D(\mathrm{d}s|t)\mathrm{d}t \quad \text{subject to (IC}_D\text{)}.$$

3.3. *Discriminatory Disclosure Problem.*

The agent chooses one of two actions, $a = 0$ or $a = 1$. The agent's utility $u_I(s,t)$ and principal's utility $v_I(s,t)$ from $a = 1$ depend on the state $s \in S = [0,1]$ and the agent's private type $t \in T = [0,1]$; the utilities from $a = 0$ are normalized to zero. The state and type are independently and uniformly distributed. Given our normalizations, a pair (u_I, v_I) fully describes the problem.

The principal and agent are initially uninformed about the state. The principal designs a menu of tests which generate informative messages about the state. The agent privately observes his type, chooses a test from the menu, observes a message from the chosen test, updates his beliefs

[6] As in the persuasion problem, the only substantive assumptions here are that the decision and type are one-dimensional (see Footnote 4).

about the state, and chooses $a=0$ or $a=1$. By the revelation principle argument, we can label each test in the menu by the type of the agent who is recommended to choose this test, and we can assume that test messages are action recommendations. That is, the principal chooses a *disclosure mechanism* π_I that asks the agent to report his type and then recommends him a stochastic action conditional on his report t and the state s:

$$\pi_I(s,t)=\mathbb{P}(\text{action}=1|\text{state}=s,\text{type}=t).$$

The key constraint on the disclosure mechanism is that the agent prefers to report his true type and to choose a recommended action. This *incentive-compatibility constraint* is

$$\int_S u_I(s,t)\pi_I(s,t)\mathrm{d}s \geq \int_S u_I(s,t)\left(\hat{a}_0(1-\pi_I(s,\hat{t}))+\hat{a}_1\pi_I(s,\hat{t})\right)\mathrm{d}s, \quad \text{for all } \hat{a}_0,\hat{a}_1\in\{0,1\}, \text{ all } \hat{t}\in T, \text{ and almost all } t\in T. \tag{IC$_I$}$$

The principal chooses a disclosure mechanism π_I to

$$\text{maximize} \quad W_I(\pi_I)=\int_{T\times S} v_I(s,t)\pi_I(s,t)\mathrm{d}s\mathrm{d}t \quad \text{subject to (IC}_I\text{)}.$$

4. EQUIVALENCE

4.1. *Main Result*

This section shows that the three problems are equivalent. Our notion of equivalence identifies the mechanisms in the three problems. For each triple of such mechanisms, this notion requires that (a) the principal gets the same expected utility in all three problems, and (b) incentive compatibility either holds or fails simultaneously in all three problems.[7]

Although persuasion mechanism π_P, delegation mechanism π_D, and disclosure mechanism π_I have different meanings in the three problems, we can identify them as follows:

$$\pi_P(t|s)=\pi_D(s|t)=\pi_I(s,t), \quad \text{for all } s\in S \text{ and all } t\in T. \tag{E$_\pi$}$$

Since, by the definition of π_P and π_D, we have $\pi_P(1|s)=0$ and $\pi_D(0|t)=0$, for (E$_\pi$) to hold, we impose the following normalizations:

$$\pi_P(t|0)=\pi_D(s|1)=\pi_I(0,t)=\pi_I(s,1)=0, \quad \text{for all } s\in S \text{ and all } t\in T. \tag{5}$$

In persuasion, $\pi_P(t|0)=0$ is w.l.o.g. because state $s=0$ occurs with zero probability. In delegation, $\pi_D(s|1)=0$ is w.l.o.g. because type $t=1$ occurs with zero probability, and decision $s=1$ is always available to the agent. In discriminatory disclosure, $\pi_I(0,t)=0$ and $\pi_I(s,1)=0$ are w.l.o.g. because $s=0$ and $t=1$ occur with zero probability.

Let Π_P, Π_D, and Π_I be the sets of all persuasion mechanisms π_P, delegation mechanisms π_D, and disclosure mechanisms π_I that satisfy (5). Since, by definition, $\pi_P(t|s)$ is decreasing in t, the set $\Pi_P\subset\Pi_I$ consists of all functions in Π_I that are decreasing in t. Similarly, since, by definition, $\pi_D(s|t)$ is increasing in s, the set $\Pi_D\subset\Pi_I$ consists of all functions in Π_I that are increasing in s.

[7]Our main result holds (with the same proof) under a stronger notion of equivalence (as in Manelli and Vincent 2010, Gershkov et al. 2013, and Kolotilin et al. 2017), which preserves not only the expected utility of the principal, but also the *interim* expected utilities of both the principal and the agent.

TABLE I

THREE EQUIVALENT PROBLEMS

	Persuasion	Delegation	Discriminatory Disclosure
Variable s	State $\sim U[0,1]$	Decision $\in [0,1]$	State $\sim U[0,1]$
Variable t	Decision $\in [0,1]$	Type $\sim U[0,1]$	Type $\sim U[0,1]$
Variable a	—	—	Action $\in \{0,1\}$
Agent's utility	$\int_0^t u(s,\tilde{t})\mathrm{d}\tilde{t}$	$\int_s^1 u(\tilde{s},t)\mathrm{d}\tilde{s}$	$au(s,t)$
Principal's utility	$\int_0^t v(s,\tilde{t})\mathrm{d}\tilde{t}$	$\int_s^1 v(\tilde{s},t)\mathrm{d}\tilde{s}$	$av(s,t)$
Mechanism π	$\mathbb{P}\,(\text{decision} > t\|s)$	$\mathbb{P}\,(\text{decision} < s\|t)$	$\mathbb{P}(\text{action} = 1\|s,t)$

DEFINITION 1: Two problems (u_K, v_K) and (u_N, v_N), with $K, N \in \{P, D, I\}$, are *equivalent* if, for all $\pi_K, \pi_N \in \Pi_K \cap \Pi_N$ satisfying (E_π), we have:
(a) $W_K(\pi_K) = W_N(\pi_N)$;
(b) π_K satisfies (IC_K) $\iff$ π_N satisfies (IC_N).
All three problems are equivalent if each pair of them is equivalent.

Note that the equivalence between problems (u_K, v_K) and (u_N, v_N) is defined on the restricted set of mechanisms $\Pi_K \cap \Pi_N$, because if mechanisms $\pi_K \in \Pi_K$ and $\pi_N \in \Pi_N$ satisfy (E_π), then $\pi_K, \pi_N \in \Pi_K \cap \Pi_N$. Thus, the equivalence of persuasion and discriminatory disclosure is defined on the set Π_P of mechanisms that are decreasing in t, the equivalence of delegation and discriminatory disclosure is defined on the set Π_D of mechanisms that are increasing in s, and the equivalence of persuasion and delegation, as well as the equivalence of all three problems, is defined on the set $\Pi_M = \Pi_P \cap \Pi_D$ of mechanisms that are increasing in s and decreasing in t.

We refer to mechanisms in Π_M as *monotone*. Monotone mechanisms have a natural interpretation. Under a monotone persuasion mechanism, a higher state generates a higher lottery over recommended decisions with respect to first-order stochastic dominance. Under a monotone delegation mechanism, a higher reported type is assigned a higher lottery over decisions with respect to first-order stochastic dominance. Under a monotone disclosure mechanism, action $a = 1$ is recommended with a higher probability when the state is higher and the reported type is lower.

Our main result shows that the three problems are equivalent when the agent's utility satisfies the standard single-crossing assumptions (Milgrom and Shannon, 1994, Quah and Strulovici, 2012, Anderson and Smith, 2024). A function $u(s,t)$ is:
(i) *upcrossing in s* if, for each t,

$$u(s_1,t) \geq (>)\, 0 \implies u(s_2,t) \geq (>)\, 0 \quad \text{whenever } s_2 > s_1;$$

(ii) *aggregate downcrossing in t* if, for each probability distribution $\lambda \in \Delta(S)$,

$$\int_S u(s,t_1)\lambda(\mathrm{d}s) \leq (<)\, 0 \implies \int_S u(s,t_2)\lambda(\mathrm{d}s) \leq (<)\, 0 \quad \text{whenever } t_2 > t_1.$$

In particular, (i) and (ii) hold if $u(s,t)$ is increasing in s and decreasing in t. In persuasion, upcrossing of u_P in s means that the agent's optimal decision is increasing in the state, and aggregate downcrossing of u_P in t means that the agent's utility is single peaked in the decision for any beliefs about the state. In delegation, upcrossing of u_D in s means that the agent's utility is single peaked in the decision for any type, and aggregate downcrossing of u_D in t means that a higher type of the agent prefers a higher lottery over decisions.

THEOREM 1: *A persuasion problem (u_P, v_P), a delegation problem (u_D, v_D), and a discriminatory disclosure problem (u_I, v_I) are equivalent if*

$$u_P,\ u_D,\ \text{and } u_I \text{ are upcrossing in } s \text{ and aggregate downcrossing in } t, \tag{SC}$$

$$(u_P, v_P) = (u_D, v_D) = (u_I, v_I). \tag{E}$$

Theorem 1 establishes the equivalence between the three problems on the set of monotone mechanisms. Table I describes three equivalent problems for a given (u, v). Thus, if a monotone mechanism solves one problem, it also solves the other two equivalent problems.

Theorem 1 follows from Lemmas 1 and 2 below. Lemma 1 establishes the equivalence between delegation and discriminatory disclosure problems when the agent's utilities satisfy single-crossing in s.

LEMMA 1: *Problems (u_D, v_D) and (u_I, v_I) are equivalent if (u_D, v_D) and (u_I, v_I) satisfy (E) and u_D and u_I are upcrossing in s.*

Lemma 2 establishes the equivalence between persuasion and discriminatory disclosure problems when the agent's utilities satisfy aggregate single-crossing in t.

LEMMA 2: *Problems (u_P, v_P) and (u_I, v_I) are equivalent if (u_P, v_P) and (u_I, v_I) satisfy (E) and u_P and u_I are aggregate downcrossing in t.*

Jointly, the conditions of single-crossing in s and t imposed separately in Lemmas 1 and 2 are precisely the conditions imposed in Theorem 1, in which case all three problems are equivalent. We prove Lemmas 1 and 2 in Sections 4.2 and 4.3 below.

REMARK 1: In the literature, distributions of random variables are usually not normalized to be uniform on $[0, 1]$. Let F be a distribution of the state $y \in Y = [\underline{y}, \overline{y}]$ in persuasion and discriminatory disclosure, and let G be a distribution of the agent's type $x \in X = [\underline{x}, \overline{x}]$ in delegation and discriminatory disclosure. Theorem 1 then applies after the change of variables as in Footnote 4.[8] In particular, if F and G admit strictly positive densities f and g, then persuasion, delegation, and discriminatory disclosure problems are equivalent if u_P, u_D, and u_I are upcrossing in y and aggregate downcrossing in x, and

$$\begin{aligned} u_P(y,x)f(y) &= u_D(y,x)g(x) = u_I(y,x)f(y)g(x), \\ v_P(y,x)f(y) &= v_D(y,x)g(x) = v_I(y,x)f(y)g(x). \end{aligned} \tag{6}$$

where (u_P, v_P) and $(-u_D, -v_D)$ are the marginal utilities in persuasion and delegation, and (u_I, v_I) are the utilities from action $a = 1$ in discriminatory disclosure.

4.2. *Delegation and Discriminatory Disclosure*

To connect delegation and discriminatory disclosure, we represent disclosure mechanisms as cutoff-state mechanisms. A disclosure mechanism $\pi_I \in \Pi_I$ is a *deterministic cutoff-state mechanism* if for each reported type t there exists a cutoff state s_t such that action 1 is recommended

[8] In expressions (2), (4), (5), and (IC_D), $s = 0$ ($s = 1$) and $t = 0$ ($t = 1$) should be replaced with $y = \underline{y}$ ($y = \overline{y}$) and $x = \underline{x}$ ($x = \overline{x}$).

if and only if $s > s_t$. Under this mechanism, when the agent is truthful and obedient, the agent's and principal's utilities conditional on type t are

$$U_S(s_t,t) = \int_{s_t}^{1} u_I(\tilde{s},t)\mathrm{d}\tilde{s} \quad \text{and} \quad V_S(s_t,t) = \int_{s_t}^{1} v_I(\tilde{s},t)\mathrm{d}\tilde{s}, \quad \text{for all } t \in T. \tag{7}$$

A disclosure mechanism $\pi_I \in \Pi_I$ is a *(stochastic) cutoff-state mechanism* if for each reported type there exists a probability distribution of cutoffs such that action 1 is recommended if and only if the state is above the cutoff.

The key observation is that each disclosure mechanism $\pi_I(s,t)$ that is increasing in s can be represented as a cutoff-state mechanism where the distribution of cutoffs conditional on reported type t is $\mathbb{P}(\text{cutoff} < s|\text{type} = t) = \pi_I(s,t)$. Indeed, under such a distribution of cutoffs, the probability that action 1 is recommended conditional on state s and type t equals the probability that the cutoff is less than s, which is precisely $\pi_I(s,t)$. When the agent is truthful and obedient, the agent's and principal's utilities conditional on type t are

$$\int_S U_S(s,t)\pi_I(\mathrm{d}s,t) \quad \text{and} \quad \int_S V_S(s,t)\pi_I(\mathrm{d}s,t), \quad \text{for all } t \in T.$$

To prove part (a) of Lemma 1, we show that the utilities conditional on type t are the same in the delegation problem $(u_D, v_D) = (u_I, v_I)$ with $\pi_D(s|t) = \pi_I(s,t)$. To prove part (b) of Lemma 1, we show that, in general, the delegation problem is a relaxation of the discriminatory disclosure problem, but these problems become equivalent if the assumptions of upcrossing and monotonicity are imposed. Intuitively, when translated to discriminatory disclosure, (IC_D) prohibits two types of deviations: *obedient misreporting*, where the agent misreports his type and chooses the recommended action; and *disregarding*, where the agent disregards the recommendation and chooses the best of the two actions. In addition, (IC_I) prohibits the third type of deviations: *disobedient misreporting*, where the agent misreports his type and chooses the action opposite to the recommendation. Thus, (IC_D) is weaker than (IC_I). For the converse, suppose that the agent's utility u_I is upcrossing in s and the disclosure mechanism π_I is increasing in s. Then disobedient misreporting can never be better for the agent than disregarding. Thus, (IC_D) is equivalent to (IC_I), and part (b) of Lemma 1 follows.[9]

PROOF OF LEMMA 1: Consider a discriminatory disclosure problem (u_I, v_I) with $\pi_I \in \Pi_I$ and a delegation problem (u_D, v_D) with $\pi_D \in \Pi_D$ such that

$$u_I(s,t) = u_D(s,t), \quad v_I(s,t) = v_D(s,t), \quad \text{and} \quad \pi_I(s,t) = \pi_D(s|t), \tag{E_D}$$

where u_D and u_I are upcrossing in s. Then, for all $t \in T$, we have

$$\begin{aligned} \int_S v_I(s,t)\pi_I(s,t)\mathrm{d}s &= \int_S V_S(s,t)\pi_I(\mathrm{d}s,t) = \int_S \left(\int_s^1 v_I(\tilde{s},t)\mathrm{d}\tilde{s} \right) \pi_I(\mathrm{d}s,t) \\ &= \int_S \left(\int_s^1 v_D(\tilde{s},t)\mathrm{d}\tilde{s} \right) \pi_D(\mathrm{d}s|t) = \int_S V_D(s,t)\pi_D(\mathrm{d}s|t), \end{aligned} \tag{8}$$

[9] Lemma 1 continues to hold under a weaker notion of upcrossing defined by Karlin and Rubin (1956): for each t, $u(s_1,t) > 0 \implies u(s_2,t) \geq 0$ whenever $s_2 > s_1$. Appendix C.1 provides an example where a mechanism π_D in delegation is incentive compatible, but the mechanism π_I given by (E_D) in discriminatory disclosure is not, when u_D is not upcrossing in s.

where the first equality is by the representation of the disclosure mechanism as a cutoff-state mechanism, the second equality is by (7), the third equality is by (E_D), and the fourth equality is by (4). Thus, part (a) follows from

$$W_I(\pi_I) = \int_{T\times S} v_I(s,t)\pi_I(s,t)\mathrm{d}s\mathrm{d}t = \int_{T\times S} V_D(s,t)\pi_D(\mathrm{d}s|t)\mathrm{d}t = W_D(\pi_D).$$

We now prove part (b). By the same logic as in (8), for all $\hat{a}_0, \hat{a}_1 \in \{0,1\}$ and all $t, \hat{t} \in T$,

$$\begin{aligned}&\int_S u_I(s,t)\left(\hat{a}_0(1-\pi_I(s,\hat{t})) + \hat{a}_1\pi_I(s,\hat{t})\right)\mathrm{d}s\\ &\quad = \hat{a}_0\left(U_D(0,t) - \int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t})\right) + \hat{a}_1\int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t}).\end{aligned} \tag{9}$$

Thus, π_I satisfies (IC_I) iff

$$\int_S U_D(s,t)\pi_D(\mathrm{d}s|t) \geq \hat{a}_0\left(U_D(0,t) - \int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t})\right) + \hat{a}_1\int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t}),$$
$$\text{for all } \hat{a}_0, \hat{a}_1 \in \{0,1\}, \text{ all } \hat{t} \in T, \text{ and almost all } t \in T.$$

Note that (IC_I) written for all $(\hat{a}_0, \hat{a}_1, \hat{t})$ such that $(\hat{a}_0, \hat{a}_1) \neq (1,0)$ is equivalent to (IC_D). To prove that (IC_I) is equivalent to (IC_D), suppose by contradiction that π_D satisfies (IC_D), but π_I violates (IC_I) for $(\hat{a}_0, \hat{a}_1) = (1,0)$. That is, there exist $t, \hat{t} \in T$ such that

$$\begin{aligned}U_D(0,t) - \int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t}) &> \int_S U_D(s,t)\pi_D(\mathrm{d}s|t)\\ &\geq \max\left\{\int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t}),\, U_D(0,t),\, 0\right\},\end{aligned} \tag{10}$$

where, by (9), the first inequality states that (IC_I) fails for $(\hat{a}_0, \hat{a}_1) = (1,0)$, and the second inequality states that (IC_D) holds, given that $U_D(1,t) = 0$ by (4). Next, we have

$$\begin{aligned}\int_S u_D(s,t)(1-\pi_D(s|\hat{t}))\mathrm{d}s &= \int_S u_I(s,t)(1-\pi_I(s,\hat{t}))\mathrm{d}s\\ &= U_D(0,t) - \int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t}) > 0,\end{aligned} \tag{11}$$

where the first equality is by (E_D), the second equality is by (9) evaluated at $(\hat{a}_0, \hat{a}_1) = (1,0)$, and the inequality is by (10). Since u_D is upcrossing in s, there exists $s_t \in S$ such that $u_D(s,t) \leq 0$ for $s < s_t$ and $u_D(s,t) \geq 0$ for $s > s_t$. Observe that $\pi_D(s_t|\hat{t}) < 1$, as otherwise we would have had

$$\int_S u_D(s,t)(1-\pi_D(s|\hat{t}))\mathrm{d}s = \int_0^{s_t} u_D(s,t)(1-\pi_D(s|\hat{t}))\mathrm{d}s \leq 0,$$

where the inequality holds because $u_D(s,t) \le 0$ for $s < s_t$. Thus,

$$\begin{aligned}\int_S u_D(s,t)\pi_D(s|\hat{t})\mathrm{d}s &= \frac{1}{1-\pi_D(s_t|\hat{t})}\left(\int_S u_D(s,t)(\pi_D(s|\hat{t}) - \pi_D(s_t|\hat{t}))\mathrm{d}s\right. \\ &\qquad \left. + \pi_D(s_t|\hat{t})\int_S u_D(s,t)(1-\pi_D(s|\hat{t}))\mathrm{d}s\right) \\ &\ge \frac{\pi_D(s_t|\hat{t})}{1-\pi_D(s_t|\hat{t})}\int_S u_D(s,t)(1-\pi_D(s|\hat{t}))\mathrm{d}s \ge 0,\end{aligned} \tag{12}$$

where the equality holds by rearrangement, the first inequality holds because $\pi_D(s|\hat{t})$ is increasing in s and $u_D(s,t) \le (\ge)\, 0$ for $s < (>)\, s_t$, and the second inequality holds by (11). Finally, we have

$$U_D(0,t) - \int_S U_D(s,t)\pi_D(\mathrm{d}s|\hat{t})\mathrm{d}s = U_D(0,t) - \int_S u_D(s,t)\pi_D(s|\hat{t})\mathrm{d}s \le U_D(0,t),$$

where the equality is by (E_D) and (9) with $(\hat{a}_0, \hat{a}_1) = (0,1)$, and the inequality is by (12). This inequality contradicts (10). Thus, π_I satisfies (IC_I) iff π_D satisfies (IC_D). *Q.E.D.*

4.3. *Persuasion and Discriminatory Disclosure*

To connect persuasion and discriminatory disclosure, we represent disclosure mechanisms as cutoff-type mechanisms. A disclosure mechanism $\pi_I \in \Pi_I$ is a *deterministic cutoff-type mechanism* if for each state s there exists a cutoff type t_s such that action 1 is recommended if and only if $t < t_s$. Under this mechanism, when the agent is truthful and obedient, the agent's and principal's utilities conditional on state s are

$$U_T(t_s, s) = \int_0^{t_s} u_I(s,\tilde{t})\mathrm{d}\tilde{t} \quad \text{and} \quad V_T(t_s, s) = \int_0^{t_s} v_I(s,\tilde{t})\mathrm{d}\tilde{t}, \quad \text{for all } s \in S. \tag{13}$$

A disclosure mechanism is a *(stochastic) cutoff-type mechanism* if for each state there exists a probability distribution of cutoffs such that action 1 is recommended if and only if the type is below the cutoff.

The key observation is that each disclosure mechanism $\pi_I(s,t)$ that is decreasing in t can be represented as a cutoff-type mechanism where the distribution of cutoffs conditional on state s is $\mathbb{P}(\text{cutoff} > t|\text{state} = s) = \pi_I(s,t)$. Indeed, under such a distribution of cutoffs, the probability that action 1 is recommended conditional on state s and type t equals the probability that the cutoff is greater than t, which is precisely $\pi_I(s,t)$. When the agent is truthful and obedient, the agent's and principal's utilities conditional on state s are

$$\int_T U_T(s,t)(-\pi_I(s,\mathrm{d}t)) \quad \text{and} \quad \int_T V_T(s,t)(-\pi_I(s,\mathrm{d}t)), \quad \text{for all } s \in S.$$

To prove part (a) of Lemma 2, we show that the utilities conditional on state s are the same in the persuasion problem $(u_P, v_P) = (u_I, v_I)$ with $\pi_P(t|s) = \pi_I(s,t)$. To prove part (b) of Lemma 2, we show that, in general, the discriminatory disclosure problem is a relaxation of the persuasion problem, but these problems become equivalent if the assumptions of aggregate downcrossing and monotonicity are imposed. Intuitively, when translated to discriminatory disclosure, (IC_P) prohibits agent's deviations when he observes the realized cutoff. In contrast, (IC$_I$) prohibits agent's deviations when he observes only that his type is above or below the cutoff. Thus, (IC$_I$) is weaker than (IC$_P$). For the converse, suppose that, in the discriminatory disclosure problem, the agent's utility u_I is aggregate downcrossing in t and the disclosure mechanism π_I is decreasing in t. Then the agent optimally chooses action 1 if and only if his type is below the realized cutoff, regardless of whether he observes the cutoff or only that his type is below or above the cutoff. Thus, (IC$_I$) is equivalent to (IC$_P$), and part (b) of Lemma 2 follows.[10]

PROOF OF LEMMA 2: Consider a discriminatory disclosure problem with (u_I, v_I) and $\pi_I \in \Pi_I$, and a persuasion problem with (u_P, v_P) and $\pi_P \in \Pi_P$ such that

$$u_I(s,t) = u_P(s,t), \quad v_I(s,t) = v_P(s,t), \quad \text{and} \quad \pi_I(s,t) = \pi_P(t|s), \tag{E$_P$}$$

where u_P and u_I are aggregate downcrossing in t. Then, for all $s \in S$, we have

$$\begin{aligned}\int_T v_I(s,t)\pi_I(s,t)\mathrm{d}t &= \int_T V_T(s,t)(-\pi_I(s,\mathrm{d}t)) = \int_T \left(\int_0^t v_I(s,\tilde t)\mathrm{d}\tilde t\right)(-\pi_I(s,\mathrm{d}t)) \\ &= \int_T \left(\int_0^t v_P(s,\tilde t)\mathrm{d}\tilde t\right)(-\pi_P(\mathrm{d}t|s)) = \int_T V_P(s,t)(-\pi_P(\mathrm{d}t|s)),\end{aligned}$$

where the first equality is by the representation of the disclosure mechanism as a cutoff-type mechanism, the second equality is by (13), the third equality is by (E$_P$), and the fourth equality is by (2). Thus, part (a) follows from

$$W_I(\pi_I) = \int_{S\times T} v_I(s,t)\pi_I(s,t)\mathrm{d}t\mathrm{d}s = \int_{S\times T} V_P(s,t)(-\pi_P(\mathrm{d}t|s))\mathrm{d}s = W_P(\pi_P).$$

We now prove part (b). First, suppose that π_P satisfies (IC$_P$). Then, for all $\hat t \in T$ and almost all $t \in T$, we have

$$0 \le \int_S (U_P(s,t) - U_P(s,\hat t))\pi_P(\mathrm{d}s|t) = \int_{\hat t}^t \left(\int_S u_P(s,\tilde t)\pi_P(\mathrm{d}s|t)\right)\mathrm{d}\tilde t.$$

Thus, as u_P is aggregate downcrossing in t, for almost all $t \in T$, we have

$$\int_S u_P(s,\tilde t)\pi_P(\mathrm{d}s|t) \ge (\le)\, 0 \quad \text{for } \tilde t < (>)\, t. \tag{14}$$

Next, let r denote the realized cutoff type, and define

$$A_{\hat a_0,\hat a_1,\hat t}(r) = \hat a_0 \mathbf{1}\{r \le \hat t\} + \hat a_1 \mathbf{1}\{r > \hat t\}. \tag{15}$$

Since the conditional distribution of r given s is $1 - \pi_I(s,r)$, we have

$$\int_T A_{\hat a_0,\hat a_1,\hat t}(r)(-\pi_I(s,\mathrm{d}r)) = \hat a_0(1 - \pi_I(s,\hat t)) + \hat a_1 \pi_I(s,\hat t). \tag{16}$$

[10] Appendix C.1 provides an example where a mechanism π_I in discriminatory disclosure is incentive compatible, but the mechanism π_P given by (E$_P$) in persuasion is not, when u_I is not aggregate downcrossing in t.

For all $\hat{a}_0, \hat{a}_1 \in \{0,1\}$, all $\hat{t} \in T$, almost all $t \in T$, we have

$$
\begin{aligned}
&-\int_S u_I(s,t)\left(\hat{a}_0(1-\pi_I(s,\hat{t})) + \hat{a}_1\pi_I(s,\hat{t})\right)\mathrm{d}s = \int_{S\times T} u_I(s,t)A_{\hat{a}_0,\hat{a}_1,\hat{t}}(r)\pi_I(s,\mathrm{d}r)\mathrm{d}s \\
&= \int_{S\times T} u_P(s,t)A_{\hat{a}_0,\hat{a}_1,\hat{t}}(r)\pi_P(\mathrm{d}r|s)\mathrm{d}s = \int_T A_{\hat{a}_0,\hat{a}_1,\hat{t}}(r)\left(\int_S u_P(s,t)\pi_P(\mathrm{d}s|r)\right)\pi_P(\mathrm{d}r) \\
&\geq \int_T \mathbf{1}\{r>t\}\left(\int_S u_P(s,t)\pi_P(\mathrm{d}s|r)\right)\pi_P(\mathrm{d}r) = \int_{S\times T} u_P(s,t)\mathbf{1}\{r>t\}\pi_P(\mathrm{d}r|s)\mathrm{d}s \\
&= \int_{S\times T} u_I(s,t)\mathbf{1}\{r>t\}\pi_I(\mathrm{d}r|s)\mathrm{d}s = \int_{S\times T} u_I(s,t)A_{0,1,t}(r)\pi_I(s,\mathrm{d}r)\mathrm{d}s \\
&= -\int_S u_I(s,t)\pi_I(s,t)\mathrm{d}s,
\end{aligned}
$$

where the first and last equalities are by (16), the second and fifth equalities are by (E_P), the third and forth equalities are by Bayes' rule (3), the inequality is by (14), and the sixths equality is by (15) with $(\hat{a}_0, \hat{a}_1, \hat{t}) = (0,1,t)$. Consequently, if π_P satisfies (IC_P), then π_I satisfies (IC_I).

Second, suppose that π_I satisfies (IC_I). Then for all $\hat{t} \in T$ and almost all $t > \hat{t}$ we have

$$
\begin{aligned}
0 &\geq \int_S u_I(s,t)(\pi_I(s,\hat{t}) - \pi_I(s,t))\mathrm{d}s = \int_S u_P(s,t)(\pi_P(\hat{t}|s) - \pi_P(t|s))\mathrm{d}s \\
&= \int_S \int_{(\hat{t},t]} u_P(s,t)(-\pi_P(\mathrm{d}\tilde{t}|s))\mathrm{d}s = \int_{(\hat{t},t]}\int_S u_P(s,t)\pi_P(\mathrm{d}s|\tilde{t})(-\pi_P(\mathrm{d}\tilde{t})),
\end{aligned}
\tag{17}
$$

where the inequality is by (IC_I) with $(\hat{a}_0, \hat{a}_1) = (0,1)$, the first equality is by (E_P), the second inequality is by the definition of π_P, and the third equality is by Bayes' rule (3). Similarly, for almost all $t \in T$ we have

$$
\begin{aligned}
0 &\geq \int_S u_I(s,t)\mathrm{d}s - \int_S u_I(s,t)\pi_I(s,t)\mathrm{d}s = \int_S u_P(s,t)(1-\pi_P(t|s))\mathrm{d}s \\
&= \int_S \int_{[0,t]} u_P(s,t)(-\pi_P(\mathrm{d}\tilde{t}|s))\mathrm{d}s = \int_{[0,t]}\int_S u_P(s,t)\pi_P(\mathrm{d}s|\tilde{t})(-\pi_P(\mathrm{d}\tilde{t})),
\end{aligned}
\tag{18}
$$

where the inequality is by (IC_I) with $(\hat{a}_0, \hat{a}_1) = (1,1)$, the first equality is by (E_P), the second inequality is by the definition of π_P, and the third equality is by Bayes' rule (3).
Thus, by (17) and (18), for π_P-almost all $\tilde{t}$ and all $\varepsilon > 0$, there exists $t \in [\tilde{t}, \tilde{t}+\varepsilon]$ such that

$$
\int_S u_P(s,t)\pi_P(\mathrm{d}s|\tilde{t}) \leq 0.
$$

By aggregate downcrossing of u_P in t, we obtain

$$
\int_S u_P(s,t)\pi_P(\mathrm{d}s|\tilde{t}) \leq 0 \quad \text{for } \pi_P\text{-almost all } \tilde{t} \text{ and all } t > \tilde{t}. \tag{19}
$$

By symmetric arguments, we obtain

$$
\int_S u_P(s,t)\pi_P(\mathrm{d}s|\tilde{t}) \geq 0 \quad \text{for } \pi_P\text{-almost all } \tilde{t} \text{ and all } t < \tilde{t}. \tag{20}
$$

Consequently, for π_P-almost all $\tilde{t} \in T$ and all $\hat{t} \in T$, we obtain

$$0 \leq \int_{\hat{t}}^{\tilde{t}} \left(\int_S u_P(s,t)\pi_P(\mathrm{d}s|\tilde{t}) \right) \mathrm{d}t = \int_S (U_P(s,\tilde{t}) - U_P(s,\hat{t}))\pi_P(\mathrm{d}s|\tilde{t}),$$

where the inequality is by (19) and (20), and the equality is by (2). Consequently, if π_I satisfies (IC_I), then π_P satisfies (IC_P). *Q.E.D.*

5. LINEAR CASE

In this section, we consider a popular subclass of the persuasion, delegation, and discriminatory disclosure problems, referred to as *linear* problems. For $K \in \{P,D,I\}$, a problem (u_K, v_K) is *linear* if[11]

$$u_K(s,t) = c(s) - b(t) \quad \text{and} \quad v_K(s,t) = \alpha c(s) - d(b(t)), \tag{L}$$

where $\alpha \in \mathbb{R}$, b and c are continuous and strictly increasing, d is continuous, and

$$c(0) \leq b(0) < b(1) \leq c(1).$$

Clearly, if u_K satisfies (L), then it satisfies (SC), so all our results apply.

In linear persuasion and linear discriminatory disclosure, the (marginal) utilities are linear in an increasing transformation c of the state. Similarly, in linear delegation, the marginal utilities are linear in an increasing transformation c of the decision. Intuitively, the linear problems are tractable, because the analysis depends only on the mean value of $c(s)$. For illustration, in linear persuasion, if the state is known to be in an interval $(s_1, s_2]$, then the agent-optimal and principal-optimal decisions depend only on $\int_{s_1}^{s_2} c(s)\mathrm{d}s/(s_2 - s_1)$. Similarly, in linear delegation, if the only permitted decisions are s_1 and s_2, then the agent-optimal and principal-optimal assignments of types to these decisions depend only on $\int_{s_1}^{s_2} c(s)\mathrm{d}s/(s_2 - s_1)$.

In the persuasion literature, the state is typically defined as $y = c(s)$, which has a distribution $F(y) = \mathbb{P}(c(s) < y) = c^{-1}(y)$ on the interval $Y = [\underline{y}, \overline{y}] = [c(0), c(1)]$. Analogously, in the delegation literature, the type is typically defined as $x = b(t)$, which has a distribution $G(x) = \mathbb{P}(b(t) \leq x) = b^{-1}(x)$ on the interval $X = [\underline{x}, \overline{x}] = [b(0), b(1)]$. Finally, in the discriminatory disclosure literature, both the state and the type are defined as $y = c(s)$ and $x = b(t)$, which have distributions F and G. Since c and b are strictly increasing and continuous, F and G are continuous and have full support on Y. In this section, we use the transformed state y and type x in line with the literature.

[11]Many of our results continue to hold under the weaker assumption that b and c are non-decreasing and are not necessarily continuous. Moreover, condition $c(0) \leq b(0) < b(1) \leq c(1)$ can be relaxed. Indeed, in persuasion, decisions t such that $b(t) < c(0)$ and $b(t) > c(1)$ can never be chosen. In delegation, types t such that $b(t) < c(0)$ and $b(t) > c(1)$ always choose decisions 0 and 1, respectively. In discriminatory disclosure, types t such that $b(t) < c(0)$ and $b(t) > c(1)$ always choose actions 1 and 0, respectively. Consequently, w.l.o.g., we can remove all t such that $b(t) \notin [c(0), c(1)]$ from consideration.

5.1. *Monotonization*

This section shows that the restriction to monotone mechanisms is w.l.o.g. in the three linear problems.[12] Specifically, for each incentive-compatible mechanism, there exists an incentive-compatible monotone mechanism such that the expected utilities are the same.[13]

THEOREM 2: *For each $K \in \{P, D, I\}$, each (u_K, v_K) that satisfies (L), and each $\pi_K \in \Pi_K$ that satisfies (IC_K), there exists a monotone mechanism $\hat{\pi}_K \in \Pi_M \subset \Pi_K$ that satisfies (IC_K) such that $W_K(\hat{\pi}_K) = W_K(\pi_K)$.*

It suffices to prove Theorem 2 for discriminatory disclosure, $K = I$, as then the result for persuasion, $K = P$, and delegation, $K = D$, follows from Lemmas 2 and 1.[14]

Before proving Theorem 2 for $K = I$, we introduce notation and present two key lemmas. We use the transformed state and type, $y = c(s)$ and $x = b(t)$, which have distributions F and G. The agent's and principal's utilities from $a = 1$ given state y and type x are

$$u(y, x) = y - x \quad \text{and} \quad v(y, x) = \alpha y - d(x), \quad \text{for all } y \in Y \text{ and all } x \in X.$$

By (L), $X \subseteq Y$, so Y is a common interval domain for y and x. Consider a mechanism $\pi(y, x)$ in variables y and x on $Y \times Y$ that satisfies the incentive-compatibility constraint

$$\int_Y (y - x)\pi(y, x)F(\mathrm{d}y) \geq \int_Y (y - x)\left(\hat{a}_0(1 - \pi(y, \hat{x})) + \hat{a}_1 \pi(y, \hat{x})\right) F(\mathrm{d}y), \tag{IC}$$
$$\text{for all } \hat{a}_0, \hat{a}_1 \in \{0, 1\} \text{ and all } x, \hat{x} \in Y.$$

When the agent's type is x, he chooses action 0 with interim probability $H_\pi(x)$ and obtains interim utility $U_\pi(x)$ given by

$$H_\pi(x) = \int_Y (1 - \pi(y, x))F(\mathrm{d}y), \quad U_\pi(x) = \int_Y (y - x)\pi(y, x)F(\mathrm{d}y), \quad \text{for all } x \in Y, \tag{21}$$

and the principal obtains interim utility $V_\pi(x)$ given by

$$\begin{aligned} V_\pi(x) &= \int_Y (\alpha y - d(x))\pi(y, x)F(\mathrm{d}y) \\ &= \alpha U_\pi(x) + (\alpha x - d(x))(1 - H_\pi(x)), \quad \text{for all } x \in X. \end{aligned} \tag{22}$$

[12] Appendix C.2 shows that the restriction to monotone mechanisms is not w.l.o.g. in nonlinear problems. In fact, in nonlinear persuasion, optimal mechanisms are often nonmonotone (Rayo and Segal 2010, Goldstein and Leitner 2018, Guo and Shmaya 2019, Kolotilin et al. 2024).

[13] This result (with the same proof) extends to the stronger notion of equivalence stated in Footnote 7. That is, for each incentive-compatible mechanism, there exists an incentive-compatible monotone mechanism that preserves the interim expected utilities of both the principal and the agent.

[14] Indeed, consider a persuasion problem (u_P, v_P) with mechanism π_P that satisfies (IC_P). By Lemma 2, in the discriminatory disclosure problem $(u_I, v_I) = (u_P, v_P)$, the disclosure mechanism $\pi_I(t, s) = \pi_P(t|s)$ satisfies (IC_I) and $W_I(\pi_I) = W_P(\pi_P)$. By Theorem 2 for $K = I$, there exists a monotone disclosure mechanism $\hat{\pi}_I$ that satisfies (IC_I) and $W_I(\hat{\pi}_I) = W_I(\pi_I)$. Again by Lemma 2, the monotone persuasion mechanism $\hat{\pi}_P(t|s) = \hat{\pi}_I(s, t)$ satisfies (IC_P) and $W_P(\hat{\pi}_P) = W_I(\hat{\pi}_I)$. The argument for $K = D$ is analogous.

Since the agent's interim utility is maximized under full disclosure of y, we have

$$
\begin{aligned}
U_\pi(x) &= \int_Y (y-x)\pi(y,x)F(\mathrm{d}y) \le \int_Y (y-x)\mathbf{1}\{y>x\}F(\mathrm{d}y) \\
&= \int_x^{\overline{y}} (y-x)F(\mathrm{d}y) = \int_x^{\overline{y}} (1-F(y))\mathrm{d}y, \quad \text{for all } x \in Y,
\end{aligned}
\tag{23}
$$

where the first and second equalities are by definition, the inequality is by pointwise maximization, and the last equality is by integration by parts.

Now, we present two key lemmas. Lemma 3 is the envelope characterization of incentive compatibility (Kolotilin et al., 2017, Lemma 1).

LEMMA 3—(Kolotilin et al., 2017): *A mechanism π satisfies* (IC) *if and only if*

$$
\begin{aligned}
&H_\pi \textit{ is increasing}, \quad U_\pi(\underline{y}) = \int_Y (1-F(y))\mathrm{d}y, \\
&U_\pi(x) = \int_x^{\overline{y}} (1-H_\pi(\tilde{x}))\mathrm{d}\tilde{x}, \quad \textit{for all } x \in Y.
\end{aligned}
\tag{24}
$$

Lemma 4 is the extension of Strassen's theorem with the additional monotone likelihood ratio property.[15] Strassen (1965) shows that if a distribution F is a mean-preserving spread of a distribution H, then there exists a joint distribution P of (x,y) such that (i) the marginal distributions of x and y are H and F and (ii) the expected value of y given x is x, meaning that (x,y) is a martingale. Müller and Rüschendorf (2001, Theorem 4.1) provide a constructive proof of Strassen's theorem. Importantly, they show that their constructed conditional distribution $P(y|x)$ increases in x with respect to first-order stochastic dominance. We show that their $P(y|x)$ increases in x with respect to the likelihood ratio order.[16]

LEMMA 4: *Let F and H be two distributions on $Y = [\underline{y}, \overline{y}] \subset \mathbb{R}$ that satisfy*

$$
\int_x^{\overline{y}} (1-H(\tilde{x}))\mathrm{d}\tilde{x} \le \int_x^{\overline{y}} (1-F(\tilde{x}))\mathrm{d}\tilde{x}, \quad \textit{for all } x \in Y, \textit{ with equality at } x = \underline{y}. \tag{MPS}
$$

There exists a conditional distribution $P(y|x)$, with $x, y \in Y$, that satisfies

$$
\int_{[\underline{y},\overline{y}]} P(y|x)H(\mathrm{d}x) = F(y), \quad \textit{for all } y \in Y, \tag{25}
$$

$$
\int_{[\underline{y},\overline{y}]} yP(\mathrm{d}y|x) = x, \quad \textit{for all } x \in Y, \tag{26}
$$

$$
\begin{gathered}
\int_{[y_1,y_2]} P(\mathrm{d}y|x_1) \int_{[y_3,y_4]} P(\mathrm{d}y|x_2) \ge \int_{[y_1,y_2]} P(\mathrm{d}y|x_2) \int_{[y_3,y_4]} P(\mathrm{d}y|x_1), \\
\textit{for all } \underline{y} \le x_1 < x_2 \le \overline{y} \textit{ and all } \underline{y} \le y_1 < y_2 < y_3 < y_4 \le \overline{y}.
\end{gathered}
\tag{27}
$$

[15] The proof of Lemma 4 and other omitted proofs are in Appendix B.

[16] One difficulty is that the likelihood ratio order is not an integral stochastic order, so we cannot use standard results on monotonicity of Markov processes (e.g., Müller and Stoyan, 2002, Section 5.2), as Müller and Rüschendorf do.

The last two ingredients for the proof of Theorem 2 are two simple claims. Claim 1 is a straightforward implication of Lemmas 3 and 4.

CLAIM 1: *For each distribution H that satisfies* (MPS)*, there exists a mechanism $\hat{\pi}(y,x)$ that is increasing in y and decreasing in x, satisfies* (IC)*, and $H_{\hat{\pi}}(x) = H(x)$ for all $x \in Y$.*

Indeed, if H satisfies (MPS), by Lemma 4, there exists a conditional distribution $P(y|x)$ that satisfies (25)–(27). By (27), $P(y|x)$ is increasing in x with respect to the likelihood ratio order. Since the likelihood ratio order is invariant under a permutation of variables x and y (e.g., Müller and Stoyan, 2002, Theorem 3.10.14), it follows that a conditional distribution $P(x|y)$, derived by Bayes' rule from $P(y|x)$ and $H(x)$, is increasing in y with respect to the likelihood ratio order. Therefore, it is also increasing in y with respect to first-order stochastic dominance (e.g., Müller and Stoyan, 2002, Theorem 3.10.16), meaning that $P(x|y)$ is decreasing in y.

Since a distribution $P(x|y)$ is increasing in x, it follows that the disclosure mechanism $\hat{\pi}$ given by

$$\hat{\pi}(y,x) = 1 - P(x|y), \quad \text{for all } y, x \in Y, \tag{28}$$

is increasing in y and decreasing in x. Moreover, it is easy to check that $H_{\hat{\pi}}(x) = H(x)$ and $U_{\hat{\pi}}(x) = \int_x^{\overline{y}} (1 - H(\tilde{x}))\mathrm{d}\tilde{x}$ for all $x \in Y$, so $\hat{\pi}$ satisfies (IC) by Lemma 3.

Claim 2 shows that, for each disclosure mechanism π_I that satisfies (IC_I), there exists a mechanism $\pi(y,x)$ that represents $\pi_I(s,t)$ in variables y and x on $Y \times Y$ and satisfies (IC). To prove Claim 2, we extend the set of types from X to Y, and let each type $x \in Y$ of the agent choose a report $\hat{x} \in X$ and actions $\hat{a}_0, \hat{a}_1 \in \{0,1\}$ to maximize his interim utility.

CLAIM 2: *For each mechanism π_I that satisfies* (IC_I) *there exists a mechanism π on $Y \times Y$ that satisfies* (IC) *and $\pi(y,x) = \pi_I(F(y), G(x))$ for all $y \in Y$ and all $x \in [\underline{x}, \overline{x})$.*

We finally prove Theorem 2. Consider any π_I that satisfies (IC_I). By Claim 2, there exists π on $Y \times Y$ that represents π_I in variables y and x, and satisfies (IC). As $x = \overline{x}$ occurs with zero probability, the principal obtains the same expected utility under π and π_I. By (23) and Lemma 3, $H = H_\pi$ satisfies (MPS). By Claim 1, there exists a monotone $\hat{\pi}$ that satisfies (IC), and $H_{\hat{\pi}} = H_\pi$. Then, by Lemma 3 and (22), $U_{\hat{\pi}} = U_\pi$ and $V_{\hat{\pi}} = V_\pi$. Consequently, the principal obtains the same expected utility under $\hat{\pi}$ and π:

$$W(\hat{\pi}) = \int_X V_{\hat{\pi}}(x) G(\mathrm{d}x) = \int_X V_\pi(x) G(\mathrm{d}x) = W(\pi).$$

Finally, $\hat{\pi}_I$ given by $\hat{\pi}_I(s,t) = \hat{\pi}(c(s), b(t))$ for all $s \in S$ and all $t \in T$ is monotone, satisfies (IC_I), and the principal obtains the same expected utility under $\hat{\pi}_I$ and $\hat{\pi}$.[17]

5.2. *Compactification*

In our delegation problem presented in Section 3.2, the principal has the constraint that the two extreme decisions, $s = 0$ and $s = 1$, are always available to the agent. However, the principal has no such constraint in a standard delegation setting (Holmström, 1977, 1984, Alonso

[17] By construction, $\hat{\pi}_I(s,t)$ is right-continuous in t and satisfies $\hat{\pi}_I(s,1) = 0$. By monotonicity of $\hat{\pi}_I$, $\tilde{\pi}_I$ given by $\tilde{\pi}_I(s,t) = \lim_{\tilde{s} \uparrow s} \hat{\pi}_I(\tilde{s},t)$ for all $s \in (0,1]$ and $\tilde{\pi}(0,t) = 0$ is left-continuous in s and coincides with $\hat{\pi}_I$ almost everywhere. So, by redefining $\hat{\pi}_I$ in this way if necessary, w.l.o.g., we can assume that $\hat{\pi}_I(s,t)$ is left-continuous in s, right-continuous in t, and satisfies (5).

and Matouschek, 2008, Amador and Bagwell, 2013), and only one such decision is available in a delegation setting where the agent has an outside option (Kartik et al., 2021, Amador and Bagwell, 2022, Saran, 2024). This section shows that, under natural Inada-type assumptions, which are satisfied in the delegation literature, both standard delegation and delegation with outside option can be represented as our delegation problem.

In all delegation settings we consider, the agent's decision s belongs to a closed interval of the real line, $S \subseteq \mathbb{R}$. We use the transformed type, $x = b(t)$, which has distribution G on $X = [\underline{x}, \overline{x}]$. Other primitives are the same as in Section 3.2. Letting c be defined on S, with $[0,1] \subset S$, the agent's and principal's marginal utilities satisfy (L), so, by (4), up to type-dependent constants, their utilities are given by

$$U(s,x) = xs - C(s) \quad \text{and} \quad V(s,x) = d(x)s - \alpha C(s), \\ \text{where } C(s) = \int_0^s c(\tilde{s})\mathrm{d}\tilde{s}. \tag{29}$$

For convenience, we change the variable $y = c(s)$. The (transformed) decisions y are in the set $Y_0 = c(S)$, with $X \subset Y_0$ by (L). In the transformed variables, the agent's and principal's utilities are given by

$$U(y,x) = xc^{-1}(y) - C(c^{-1}(y)) \quad \text{and} \quad V(y,x) = d(x)c^{-1}(y) - \alpha C(c^{-1}(y)), \\ \text{for all } y \in Y_0 \text{ and all } x \in X. \tag{30}$$

In *standard delegation*, the set of decisions is the real line $S = \mathbb{R}$, so that $Y_0 = (\underline{y}_0, \overline{y}_0) = c(\mathbb{R})$, and the agent has no outside option. So a delegation mechanism π must satisfy only the incentive-compatibility constraint

$$\int_{Y_0} U(y,x)\pi(\mathrm{d}y|x) \geq \int_{Y_0} U(y,x)\pi(\mathrm{d}y|\hat{x}), \quad \text{for all } x, \hat{x} \in X. \tag{IC$_0$}$$

In *delegation with outside option*, the set of decisions is a ray $S = [\underline{s}, \infty)$, so that $Y_0 = [\underline{y}, \overline{y}_0) = c([\underline{s}, \infty))$, and the agent can always choose the *outside option* $\underline{y} = c(\underline{s})$. So a delegation mechanism π must satisfy the incentive-compatibility constraint (IC$_0$) and the participation constraint

$$\int_{Y_0} U(y,x)\pi(\mathrm{d}y|x) \geq U(\underline{y}, x), \quad \text{for all } x \in X. \tag{IC$_1$}$$

In our delegation problem of Section 3.2, the set of decisions is a compact interval $S = [\underline{s}, \overline{s}]$, so that $Y_0 = [\underline{y}, \overline{y}] = [c(\underline{s}), c(\overline{s})]$, and the agent can always choose the extreme decisions $\underline{y}$ and $\overline{y}$.[18] So a delegation mechanism π must satisfy the incentive-compatibility constraint (IC$_0$), the participation constraint (IC$_1$), and the additional constraint

$$\int_{Y_0} U(y,x)\pi(\mathrm{d}y|x) \geq U(\overline{y}, x), \quad \text{for all } x \in X. \tag{IC$_2$}$$

To show that both standard delegation and delegation with outside option can be represented as our delegation, for each incentive-compatible mechanism π we find another incentive-compatible mechanism $\tilde{\pi}$ with support on a compact interval $Y = [\underline{y}, \overline{y}]$ such that the agent's

[18] In Section 3.2, w.l.o.g., we normalized $[\underline{s}, \overline{s}] = [0,1]$, but here it is convenient not to use this normalization.

and principal's interim utilities are the same:

$$\begin{aligned}\int_Y U(y,x)\tilde{\pi}(\mathrm{d}y|x) &= \int_{Y_0} U(y,x)\pi(\mathrm{d}y|x), \quad \text{for all } x\in X,\\ \int_Y V(y,x)\tilde{\pi}(\mathrm{d}y|x) &= \int_{Y_0} V(y,x)\pi(\mathrm{d}y|x), \quad \text{for all } x\in X.\end{aligned} \tag{31}$$

We first show that standard delegation can be represented as our delegation if $U(y,x)\to -\infty$ and $V(y,x)\to -\infty$ for all $x\in X$ as $y\to \overline{y}_0$ and as $y\to \underline{y}_0$. Say that a mechanism π is *undominated by* $V_0 : X\to \mathbb{R}$ if

$$\int_{Y_0} V(y,x)\pi(\mathrm{d}y|x)\geq V_0(x), \quad \text{for some } x\in X.$$

From the principal's optimization perspective, it is w.l.o.g. to consider undominated mechanisms. To this end, we can set V_0 to be the principal's interim or expected utility if a decision $y^*\in Y_0$ is implemented for all reports of the agent,

$$V_0(x) = V(y^*,x) \quad \text{or} \quad V_0(x) = \int_X V(y^*,\tilde{x})G(\mathrm{d}\tilde{x}), \quad \text{for some } y^*\in Y_0. \tag{32}$$

PROPOSITION 1: *Suppose that $Y_0 = (\underline{y}_0,\overline{y}_0)\subseteq \mathbb{R}$, $\alpha>0$, and*

$$\underline{y}_0 < x < \overline{y}_0 \quad \textit{and} \quad \alpha\underline{y}_0 < d(x) < \alpha\overline{y}_0, \quad \textit{for all } x\in X. \tag{33}$$

For each continuous V_0, there exist $\underline{y},\overline{y}\in Y_0$ (with $\underline{y}<\underline{x}<\overline{x}<\overline{y}$) such that the following holds. For each mechanism π that satisfies (IC$_0$) and is undominated by V_0, there exists another mechanism $\tilde{\pi}$ with support in $Y=[\underline{y},\overline{y}]$ that satisfies (IC$_0$)–(IC$_2$) (with strict inequalities in (IC$_1$) and (IC$_2$)) and (31).

We now show that delegation with outside option can be represented as our delegation if $U(y,x)\to -\infty$ for all $x\in X$ as $y\to \overline{y}_0$.

PROPOSITION 2: *Suppose that $Y_0 = [\underline{y},\overline{y}_0)\subset \mathbb{R}$ and*

$$\underline{y}\leq x < \overline{y}_0, \quad \textit{for all } x\in X. \tag{34}$$

Then there exists $\overline{y}\in Y_0$ (with $\overline{y}>\overline{x}$) such that the following holds. For each mechanism π that satisfies (IC$_0$) and (IC$_1$), there exists another mechanism $\tilde{\pi}$ with support in $Y=[\underline{y},\overline{y}]$ that satisfies (IC$_0$)–(IC$_2$) (with strict inequality in (IC$_2$)) and (31).

5.3. *Optimization*

Using the tools from the literature on linear persuasion, this section fully characterizes optimal mechanisms in standard delegation and delegation with outside option. Let the utilities be given by (30), and suppose that the type x has distribution G that admits a càdlàg density g, meaning that g is right-continuous and has left limits on $X=[\underline{x},\overline{x}]$.

We now impose assumptions of Section 5.2 to represent standard delegation and delegation with outside option as our delegation problem with an appropriately defined decision set $Y=[\underline{y},\overline{y}]\subset Y_0$. Our analysis then applies simultaneously to both variants of the delegation problem, with the understanding that Y differs in the two variants. In fact, since delegation with outside option features an additional constraint (namely, (IC$_1$)), we can always define Y to be larger in standard delegation than in delegation with outside option.

First, consider standard delegation. Suppose that the assumptions of Proposition 1 hold with V_0 given by (32). Let Y be as in Proposition 1. W.l.o.g., we restrict attention to delegation mechanisms $\pi(y|x)$ in variables y and x on $Y \times X$ that satisfy (IC$_0$)–(IC$_2$), with strict inequalities in (IC$_1$) and (IC$_2$). Let Π_0 be the set of all such delegation mechanisms.

Second, consider delegation with outside option $\underline{y}$. Suppose that the assumptions of Proposition 2 hold. Let Y be as in Proposition 2. W.l.o.g., we restrict attention to delegation mechanisms $\pi(y|x)$ in variables y and x on $Y \times X$ that satisfy (IC$_0$)–(IC$_2$), with strict inequality in (IC$_2$). Let Π_1 be the set of all such delegation mechanisms.

By (30), the agent's and principal's utilities, up to a strictly increasing affine transformation, are given by

$$U(y,x) = \int_y^{\overline{y}} (\tilde{y} - x) F(\mathrm{d}\tilde{y}) \quad \text{and} \quad V(y,x) = \int_y^{\overline{y}} (\alpha\tilde{y} - d(x)) F(\mathrm{d}\tilde{y}),$$
$$\text{where} \;\; F(y) = \frac{c^{-1}(y) - c^{-1}(\underline{y})}{c^{-1}(\overline{y}) - c^{-1}(\underline{y})}. \tag{35}$$

Next, define the *backward bias* $\nu : Y \to \mathbb{R}$ as

$$\nu(z) = \begin{cases} 0, & \text{if } z \in [\underline{y}, \underline{x}], \\ \int_{\underline{x}}^{z} (\alpha z - d(\tilde{x})) g(\tilde{x}) \mathrm{d}\tilde{x}, & \text{if } z \in (\underline{x}, \overline{x}], \\ \alpha z - \int_{\underline{x}}^{\overline{x}} d(\tilde{x}) g(\tilde{x}) \mathrm{d}\tilde{x}, & \text{if } z \in (\overline{x}, \overline{y}]. \end{cases} \tag{36}$$

The backward bias is introduced by Alonso and Matouschek (2008) and is frequently used in the literature on linear delegation.[19] In the equivalent persuasion problem, the same function ν represents the principal's *indirect utility*, which is widely used in the literature on linear persuasion. Indeed, in the equivalent persuasion problem, the agent's and principal's utilities are given by

$$U_P(y,x) = \int_{\underline{x}}^{x} (y - \tilde{x}) g(\tilde{x}) \mathrm{d}\tilde{x} \quad \text{and} \quad V_P(y,x) = \int_{\underline{x}}^{x} (\alpha y - d(\tilde{x})) g(\tilde{x}) \mathrm{d}\tilde{x}, \tag{37}$$

and the state y has distribution F. For each distribution $\lambda \in \Delta(Y)$, we have $\mathbb{E}_\lambda[U_P(y,x)] = U_P(\mathbb{E}_\lambda[y], x)$ and $\mathbb{E}_\lambda[V_P(y,x)] = V_P(\mathbb{E}_\lambda[y], x)$, so the utilities depend on posterior beliefs only through the posterior mean. Each posterior mean $z \in Y$ induces the agent to choose decision $x^*(z) = \underline{x}$ if $z \leq \underline{x}$, decision $x^*(z) = z$ if $z \in (\underline{x}, \overline{x}]$, and decision $x^*(z) = \overline{x}$ if $z > \overline{x}$. Thus, the principal's indirect utility $V_P(z, x^*(z))$ is precisely $\nu(z)$ given by (36).

Consider a delegation mechanism $\pi \in \Pi_0$ or $\pi \in \Pi_1$. By Lemma 1 and Claim 2, π can be extended to $Y \times Y$ such that (IC$_0$)–(IC$_2$) hold for all $x, \hat{x} \in Y$ (rather than just in X). Next, let a distribution H_π on Y be given by[20]

$$H_\pi(x) = \int_Y (1 - \pi(y|x)) F(\mathrm{d}y), \quad \text{for all } x \in Y, \tag{38}$$

[19]The utilities in Alonso and Matouschek (2008) are given by (29) with $c(s) = s$ and $\alpha = 1$. In this case, the agent's and principal's preferred decisions are x and $d(x)$. Alonso and Matouschek (2008) define the backward bias as $\nu(z) = G(z)(z - \mathbb{E}_G[d(x)|x \leq z])$ for all $z \in Y$, which coincides with (36).

[20]It follows from Lemmas 1 and 3, and Claim 2 that H_π is a distribution on Y.

and let $\mathrm{supp}(H_\pi)$ denote the support of H_π.

We now present two simple claims which show that the delegation problem can be solved by standard methods from the persuasion literature. Claim 3 shows that the principal's expected utility can be expressed as $\int_Y \nu(x)H_\pi(\mathrm{d}x)$.

CLAIM 3: *For $j = 0, 1$,*

$$W(\pi) = \int_Y \nu(x)H_\pi(\mathrm{d}x), \quad \textit{for all } \pi \in \Pi_j. \tag{39}$$

Claim 4 shows that the delegation problem can be expressed as a maximization over distributions H such that F is a mean-preserving spread of H. Although Claim 4 takes the same form for standard delegation and delegation with outside option, the optimal mechanisms are not the same, because the set Y and thus condition (MPS) differ in the two variants.

CLAIM 4: *For $j = 0, 1$, mechanism π maximizes W on Π_j if and only if H_π maximizes $\int_Y \nu(x)H(\mathrm{d}x)$ over distributions H that satisfy* (MPS).

We now characterize optimal delegation mechanisms by adapting Theorem 6 in Dworczak and Kolotilin (2024) from persuasion to delegation, and generalizing it to allow for nondifferentiable functions ν and nonmonotone mechanisms π. Let $\boldsymbol{\nu}'$ be the set of all generalized derivatives of ν:

$$\boldsymbol{\nu}'(x) = \left[\liminf_{\varepsilon\to 0} \frac{\nu(x+\varepsilon)-\nu(x)}{\varepsilon}, \limsup_{\varepsilon\to 0} \frac{\nu(x+\varepsilon)-\nu(x)}{\varepsilon}\right], \quad \text{for all } x \in Y.$$

As d is continuous and g is càdlàg, ν given by (36) has left and right derivatives, so $\boldsymbol{\nu}'(x)$ is simply the interval between $\nu'(x_-)$ and $\nu'(x_+)$.

THEOREM 3: *For $j = 0, 1$, mechanism π maximizes W on Π_j if and only if there exists a selection ν' from $\boldsymbol{\nu}'$ such that $p : Y \to \mathbb{R}$ given by*

$$p(y) = \sup_{x\in \mathrm{supp}(H_\pi)} (\nu(x) + \nu'(x)(y-x)), \quad \textit{for all } y \in Y, \tag{40}$$

satisfies

$$p(y) \geq \nu(y), \quad \textit{for all } y \in Y, \tag{41}$$

$$\int_Y p(y)F(\mathrm{d}y) = \int_Y \nu(x)H_\pi(\mathrm{d}x), \tag{42}$$

Condition (42) in Theorem 3 can be simplified when $\pi \in \Pi_j$ is monotone (i.e., $\pi(y|x)$ is decreasing in x for all $y \in Y$) or deterministic (i.e., $\pi(y|x) \in \{0, 1\}$ for all $y \in Y$ and all $x \in X$). For a monotone $\pi \in \Pi_j$, define a joint distribution $J_\pi \in \Delta(Y \times Y)$ as

$$J_\pi(y, x) = \int_{\underline{y}}^{y} (1 - \pi(\tilde{y}|x))F(\mathrm{d}\tilde{y}), \quad \text{for all } (y, x) \in Y \times Y. \tag{43}$$

For a deterministic $\pi \in \Pi_j$, there exists a corresponding compact delegation set $B \subset Y$ and agent's best-response function $y^*_B(x) \in \arg\max_{y\in B} U(y, x)$ such that $\pi(y|x) = \mathbf{1}\{y^*_B(x) < y\}$ for all $y \in Y$ and all $x \in X$. For each $y \in Y$, let

$$\overline{z}_B(y) = \inf\{\tilde{y} \in B \cup \{\overline{y}\} : \tilde{y} \geq y\} \quad \text{and} \quad \underline{z}_B(y) = \sup\{\tilde{y} \in B \cup \{\underline{y}\} : \tilde{y} < y\},$$

with the convention $\underline{z}_B(\underline{y}) = \underline{y}$, and let

$$x_B^*(y) = \mathbb{E}_F[y|y \in [\underline{z}_B(y), \overline{z}_B(y)]] = \begin{cases} y, & \text{if } \underline{z}_B(y) = \overline{z}_B(y), \\ \frac{\int_{\underline{z}_B(y)}^{\overline{z}_B(y)} yF(\mathrm{d}y)}{F(\overline{z}_B(y)) - F(\underline{z}_B(y))}, & \text{if } \underline{z}_B(y) < \overline{z}_B(y). \end{cases} \tag{44}$$

REMARK 2: For $j = 0, 1$, if $\pi \in \Pi_j$ is monotone, then condition (42) simplifies to

$$p(y) = \nu(x) + \nu'(x)(y - x), \quad \text{for } J_\pi\text{-almost all } (y, x) \in Y \times Y. \tag{45}$$

For $j = 0, 1$, if $\pi \in \Pi_j$ is deterministic with a corresponding compact delegation set B, then condition (42) simplifies to

$$p(y) = \nu(x_B^*(y)) + \nu'(x_B^*(y))(y - x_B^*(y)), \quad \text{for all } y \in Y. \tag{46}$$

Theorem 3 with Remark 2 easily yield the optimality conditions for *threshold delegation*, which permits all decisions on one side of a given threshold. Consider *floor delegation* which permits all decisions above y^*. Recall that the set of decisions is $Y_0 = (\underline{y}_0, \overline{y}_0)$ in standard delegation and $Y_0 = [\underline{y}, \overline{y}_0)$ in delegation with outside option $\underline{y}$. Let g be continuous on X, so that $\boldsymbol{\nu'}(x)$ is a singleton for all $x \notin \{\underline{x}, \overline{x}\}$. First, Theorem 3 with Remark 2 immediately imply that full discretion (i.e., $y^* \leq \underline{x}$) is optimal if and only if ν is convex on Y_0. Second, Corollary 1 shows when nontrivial floor delegation (i.e., $y^* > \underline{x}$) is optimal.

COROLLARY 1: *Let $y^* \in (\underline{x}, \overline{y}_0)$. Delegation set $\{\underline{y}\} \cup [y^*, \overline{y}_0)$ in delegation with outside option $\underline{y}$ (delegation set $[y^*, \overline{y}_0)$ in standard delegation) is optimal if and only if*

(a) ν is convex on $[y^*, \overline{y}_0)$,

(b) there exists $\nu'(z^*) \in \boldsymbol{\nu'}(z^*)$ such that $\nu(y) \leq \nu(z^*) + \nu'(z^*)(y - z^*)$ for all $y \leq y^*$ with equality at $y = y^*$,

where $z^* = \int_{\underline{y}}^{y^*} yF(\mathrm{d}y)/(F(y^*) - F(\underline{y}))$ in delegation with outside option ($z^* \in (\underline{y}_0, \underline{x})$ in standard delegation).

5.4. *Related Literature on Linear Delegation*

Recently, Kleiner et al. (2021) show that a standard delegation problem with quadratic utilities (a special case of linear delegation with $F(y) = y$) simplifies to maximizing $\int \nu(x)H(\mathrm{d}x)$ over H satisfying (MPS), thereby indirectly establishing the connection to linear persuasion where the same maximization problem arises (e.g., Kolotilin, 2018, Proposition 2). Instead, our equivalence applies to the general class of linear persuasion, linear delegation (covering both standard delegation and delegation with outside option), and linear discriminatory disclosure problems. Moreover, our equivalence identifies a direct mapping between the primitives of the equivalent problems (see Table I). Finally, thanks to Theorem 2, our equivalence holds on the space of primitive mechanisms π, and thus on a narrower space of induced distributions H, as in Kleiner et al. The latter space is narrower, as every π induces a unique H, but not every H is induced by a unique π. In particular, there exist H, distinct $\pi_P \in \Pi_P$ and $\pi_D \in \Pi_D$ inducing H, and ν such that H uniquely maximizes $\int \nu(x)H(\mathrm{d}x)$ subject to (MPS), but $\pi_P \notin \Pi_D$ and $\pi_D \notin \Pi_P$. Thus, the equivalence is stronger on the space of π than on the space of H. Our proof of the equivalence crucially relies on our Theorems 1 and 2, which have no counterparts in Kleiner et al.

The delegation literature has largely focused on the optimality of threshold delegation[21] in standard delegation, which corresponds to censorship in linear persuasion.[22] The most general result is due to Amador and Bagwell (2013, Propositions 1 and 2).[23] Our approach easily yields essentially the same result but under weaker differentiability assumptions (Corollary 1). We allow for stochastic mechanisms, whereas Amador and Bagwell (2013) restrict to deterministic mechanisms but permit money burning.[24] However, allowing for money burning does not change our conditions for the optimality of floor delegation if $\alpha \leq 1$, in which case money burning is a costlier incentive tool than stochastic mechanisms.

Special cases of linear delegation with outside option are studied in Zapechelnyuk (2020), Kartik et al. (2021), Amador and Bagwell (2022), and Saran (2024). As we consider the general case of linear delegation, our conditions for the optimality of threshold delegation (Corollary 1) subsume the corresponding conditions in these papers.

6. NONLINEAR DETERMINISTIC CASE

In this section, we consider another popular subclass of the persuasion, delegation, and discriminatory disclosure problems, referred to as *deterministic* problems. For $K \in \{P, D, I\}$, a monotone mechanism $\pi_K \in \Pi_M$ is *deterministic* if it takes values 0 or 1, and a problem K is deterministic if the principal is constrained to deterministic mechanisms.

As in Remark 1, we use variables $y \in Y = [\underline{y}, \overline{y}]$ and $x \in X = [\underline{x}, \overline{x}]$ that have strictly positive densities f and g, and assume that (6) holds. In discriminatory disclosure, the agent's and principal's utilities from action $a = 1$ are given by $u(y, x)$ and $v(y, x)$. In the other two problems, the utilities are determined by (6). In persuasion, the agent's and principal's utilities are given by

$$U_P(y,x) = \int_{\underline{x}}^{x} u(y,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x} \quad \text{and} \quad V_P(y,x) = \int_{\underline{x}}^{x} v(y,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x}. \tag{47}$$

In delegation, the agent's and principal's utilities are given by

$$U_D(y,x) = \int_{y}^{\overline{y}} u(\tilde{y},x)f(\tilde{y})\mathrm{d}\tilde{y} \quad \text{and} \quad V_D(y,x) = \int_{y}^{\overline{y}} v(\tilde{y},x)f(\tilde{y})\mathrm{d}\tilde{y}. \tag{48}$$

We impose strict single-crossing assumptions: u is *strictly upcrossing in* y if, for each $x \in X$,

$$u(y_1, x) \geq 0 \implies u(y_2, x) > 0 \quad \text{whenever } y_2 > y_1,$$

and u is *strictly aggregate downcrossing in* x if, for each probability distribution $\lambda \in \Delta(Y)$,

$$\int_Y u(y, x_1)\lambda(\mathrm{d}y) \leq 0 \implies \int_Y u(y, x_2)\lambda(\mathrm{d}y) < 0 \quad \text{whenever } x_2 > x_1.$$

[21]The delegation literature also shows when interval delegation (which permits all decisions in an interval) and gap delegation (which prohibits all decisions in an interval) are optimal. These results easily follow from our Theorem 3 with Remark 2 (see Kolotilin and Zapechelnyuk, 2019, Proposition 2).

[22]Censorship in linear persuasion is studied, among others, by Kamenica and Gentzkow (2011), Kolotilin (2015), Gentzkow and Kamenica (2016), Kolotilin et al. (2017), Kolotilin (2018), Dworczak and Martini (2019), Kleiner et al. (2021), Kolotilin et al. (2022), and Arieli et al. (2023).

[23]See also Amador et al. (2018). Alonso and Matouschek (2008) study the case with quadratic utilities ($F(y) = y$). More specialized cases are studied, among others, in Holmström (1977, 1984), Melumad and Shibano (1991), and Martimort and Semenov (2006).

[24]Goltsman et al. (2009), Kováč and Mylovanov (2009), and Kleiner et al. (2021) also allow for stochastic mechanisms but restrict attention to quadratic utilities ($F(y) = y$).

Under strict single-crossing, an incentive-compatible deterministic mechanism can be described by a subset B of S, representing a monotone partition in persuasion, a delegation set in delegation, and a menu of cutoff tests in discriminatory disclosure.[25] We next show that the set of all incentive-compatible deterministic mechanisms in the three problems is identified with

$$\mathcal{B} = \{B \subset Y : B \text{ is closed and } \{\underline{y}, \overline{y}\} \subset B\}.$$

In persuasion, an incentive-compatible deterministic mechanism is described by a monotone partition that divides the set of states Y into convex sets — singletons and intervals. The agent observes which partition element contains the state and chooses an optimal decision. A monotone partition is represented by a set $B \in \mathcal{B}$ of boundary points of all partition elements. Specifically, let

$$\overline{z}_B(y) = \inf\{\tilde{y} \in B : \tilde{y} \geq y\} \quad \text{and} \quad \underline{z}_B(y) = \sup\{\tilde{y} \in B : \tilde{y} < y\},$$

with the convention $\underline{z}_B(\underline{y}) = \underline{y}$. The partition element of B that contains state $y \in Y$ is the singleton $\{y\}$ (so the state is revealed) when $\underline{z}_B(y) = \overline{z}_B(y)$, and it is the interval $(\underline{z}_B(y), \overline{z}_B(y)]$ (so the state is pooled with other states in that interval) when $\underline{z}_B(y) < \overline{z}_B(y)$. For example, if $Y = [0,1]$ and $B = [0,1/2] \cup \{1\}$, then all states in $[0,1/2]$ are revealed, and all states in $(1/2,1]$ are pooled.

In delegation, an incentive-compatible deterministic mechanism is described by a delegation set $B \subset Y$. The agent privately observes his type and chooses an optimal decision from the delegation set. As the agent can always choose extreme decisions $y = \underline{y}$ and $y = \overline{y}$, these decisions are included in B, so $B \in \mathcal{B}$.

In discriminatory disclosure, an incentive-compatible deterministic mechanism is described by a menu $B \subset Y$ of cutoff tests. The agent privately observes his type x, chooses his preferred cutoff test b from the menu B, observes whether the state y is below or above b, and chooses an optimal action $a \in \{0,1\}$. As the agent can always ignore the test, the uninformative tests $b = \underline{y}$ and $b = \overline{y}$ are included in B, so $B \in \mathcal{B}$.

6.1. *Equivalence*

This section shows that in the case of deterministic mechanisms, our equivalence result takes a simple form: for each set $B \in \mathcal{B}$, the principal's expected utility is the same in the three equivalent problems. To state this result, we define the agent's best-response function and the principal's expected utility in the three problems for each set $B \in \mathcal{B}$. In persuasion with monotone partition B, when the state is y, the agent optimally chooses decision given by

$$x^*_B(y) \in \arg\max_{x \in X} \begin{cases} U_P(y,x), & \text{if } \underline{z}_B(y) = \overline{z}_B(y), \\ \int_{\underline{z}_B(y)}^{\overline{z}_B(y)} U_P(\tilde{y},x) f(\tilde{y}) \mathrm{d}\tilde{y}, & \text{if } \underline{z}_B(y) < \overline{z}_B(y), \end{cases} \tag{49}$$

so the principal's expected utility is

$$W_P(B) = \int_Y V_P(y, x^*_B(y)) f(y) \mathrm{d}y, \quad \text{for all } B \in \mathcal{B}.$$

[25] Our analysis can be extended from strict single-crossing to single-crossing. Strict single crossing ensures that the agent's best-response correspondence is single valued almost everywhere in the three problems, and, thus, for all B, the principal's expected utility is also single valued. Without strict single-crossing, the agent's best response and, thus, the principal's expected utility $W(B)$ are generally set valued. In this case, by Aumann's (1965) integration of correspondences, Corollary 2 in Section 6.1 would equate the *sets* of the principal's expected utilities in the three problems for all B, as shown in Kolotilin and Zapechelnyuk (2019, Theorem 1′).

In delegation with delegation set B, the agent with type x optimally chooses decision

$$y_B^*(x) \in \arg\max_{y\in B} U_D(y,x), \tag{50}$$

so the principal's expected utility is

$$W_D(B) = \int_X V_D(y_B^*(x),x)g(x)\mathrm{d}x, \quad \text{for all } B \in \mathcal{B}.$$

In discriminatory disclosure with menu of cutoff tests B, when the state is y, the agent with type x optimally chooses action

$$a_B^*(y,x) = \begin{cases} a_0^*(x), & \text{if } y \le b^*(x), \\ a_1^*(x), & \text{if } y > b^*(x), \end{cases} \tag{51}$$

where $(a_0^*(x), a_1^*(x), b^*(x)) \in \arg\max_{(a_0,a_1,b)\in\{0,1\}^2\times B} a_0 \int_{\underline{y}}^b u(y,x)f(y)\mathrm{d}y + a_1 \int_b^{\overline{y}} u(y,x)f(y)\mathrm{d}y,$

so the principal's expected utility is

$$W_I(B) = \int_{X\times Y} v(y,x)a_B^*(y,x)f(y)g(x)\mathrm{d}y\mathrm{d}x, \quad \text{for all } B\in\mathcal{B}.$$

In the deterministic case, Theorem 1 with Remark 1 simplifies as follows.

COROLLARY 2: *A deterministic persuasion problem (u_P, v_P, f), a deterministic delegation problem (u_D, v_D, g), and a deterministic discriminatory disclosure problem (u_I, v_I, f, g) satisfy $W_P(B) = W_D(B) = W_I(B)$ for all $B \in \mathcal{B}$ and all y_B^*, x_B^* and a_B^* given by (49)–(51) if u_P, u_D, and u_I are strictly upcrossing in y and strictly aggregate downcrossing in x, and satisfy (6).*

6.2. *Optimization*

Using the tools from the literature on nonlinear persuasion, this section derives two results on optimal deterministic delegation.[26] First, for general utility functions, Proposition 3 provides a sufficient condition for the optimality of *full discretion*, which permits all decisions. Second, for a special class of nonlinear utility functions, Proposition 4 provides necessary and sufficient conditions for the optimality of any candidate delegation set.

Our analysis applies to our delegation problem, as well as to standard delegation and delegation with outside option, after applying compactification analogously to Section 5.2 (see Appendix A). We henceforth restrict attention to our delegation problem.

For each decision $y \in Y$, define $x^*(y) = \arg\max_{x\in X} \int_{\underline{x}}^x u(y,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x}$. In the interior case, $x^*(y)$ is the type x whose preferred decision is y, so that $u(y,x^*(y)) = 0$.

[26] A promising avenue for future research is to establish new results in persuasion using tools from delegation. In particular, the Lagrangian method developed by Amador and Bagwell (2013) is widely used to study delegation problems with a nonlinear utility of the principal. Using our equivalence, this method can be applied to nonlinear monotone persuasion problems, which are currently not well understood.

PROPOSITION 3: *Let g, v, and $u_x = \partial u/\partial x$ be continuous functions. Delegation set $B = Y$ maximizes W on $\mathcal{B}$ if*

$$\frac{\int_x^{x^*(y_2)} v(y_2,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x}}{u(y_2,x)} \geq \frac{\int_{x^*(y_1)}^{x} v(y_1,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x}}{-u(y_1,x)}, \tag{52}$$

for all $y_1, y_2 \in Y$ and all $x \in X$ such that $u(y_1,x) < 0 < u(y_2,x)$.

Proposition 3 adapts the results from the persuasion literature. Kolotilin (2018, Proposition 1(ii)) and Kolotilin et al. (2024, Theorem 5) show that full disclosure is optimal among all (and thus among all monotone) persuasion mechanisms if, for all $\rho \in (0,1)$ and all $y_1, y_2 \in Y$, the principal prefers to split posterior λ that assigns probabilities ρ and $1-\rho$ to states y_1 and y_2 into two degenerate posteriors that assign probability 1 to y_1 and y_2.[27] This condition can be expressed as (52), because in persuasion the utilities are given by (47) and posterior λ induces decision x if $\rho u(y_1,x) + (1-\rho)u(y_2,x) = 0$. Thus, if (52) holds, then full discretion is optimal among all monotone (and thus among all deterministic) delegation mechanisms by Corollary 2.[28]

We now show that condition (52) is weaker than common conditions for full discretion in the delegation literature, and it applies to a broader class of utilities.

REMARK 3: Let $X \subset Y$, $u(y,x) = y - x$, and $v_y = \partial v/\partial y$ be a continuous function.[29] Then (52) holds if

$$\begin{gathered}\left(\min_{y,x \in Y \times X} v_y(y,x)\right) G(x) + v(x,x)g(x) \quad \text{is increasing in } x \text{ on } X, \\ v(\underline{x},\underline{x})g(\underline{x}) \geq 0 \ \text{ if } \underline{x} > \underline{y} \quad \text{and} \quad v(\overline{x},\overline{x})g(\overline{x}) \leq 0 \ \text{ if } \overline{x} < \overline{y}.\end{gathered} \tag{53}$$

Amador and Bagwell (2013, Proposition 1) and Kartik et al. (2021, Proposition 1) show that, in standard delegation and delegation with outside option, full discretion is optimal if (53) holds.[30] Their Lagrangian method uses the envelope characterization of incentive compatibility, which relies on $u(y,x) = y - x$, and the concavity of the Lagrangian, which relies on (53). Our approach is valid for a general u and does not rely on the concavity of the Lagrangian.

Consider now a popular subclass of nonlinear problems referred to as *linear* problems*. For comparability with the linear case of Section 5, we use variables s and t that are uniformly distributed on $S = T = [0,1]$, and replace $y = s$ and $x = t$ in the notation.

For $K \in \{P, D, I\}$, a problem (u_K, v_K) is *linear** if

$$u_K(s,t) = c(s) - b(t) \quad \text{and} \quad v_K(s,t) = e(c(s)) - \beta b(t), \tag{L*}$$

where $\beta \in \mathbb{R}$, b and c are continuous and strictly increasing, and e is continuous.

In contrast to (L), where u and v are linear in $c(s)$, here u and v are linear in $b(t)$. W.l.o.g., we normalize $c(0) = 0$ and $c(1) = 1$.

[27] Kolotilin (2018) and Kolotilin et al. (2024) assume that $x^*(y)$ is interior for all y, but the result easily extends to the general case with possible boundary solutions, as follows from the proof of Proposition 3.

[28] This result does not depend on f which determines the utilities in (48).

[29] In standard delegation and in delegation with outside option, $X \subset Y$ follows from Propositions $1'$ and $2'$.

[30] See also Amador and Bagwell (2022) for related results in delegation with outside option.

We now characterize optimal delegation sets in linear* delegation when $b(t)=t$. Define

$$\eta(s)=\int_0^s\left(e(c(\tilde{s}))-\frac{\beta c(\tilde{s})}{2}\right)\mathrm{d}\tilde{s},$$

and let $\operatorname{conv}\eta$ be the largest convex function on S such that $\operatorname{conv}\eta\le\eta$. Next, for each delegation set $B\in\mathcal{B}$, define

$$\eta_B(s)=\begin{cases}\eta(s), & \text{if } \underline{z}_B(s)=\overline{z}_B(s),\\ \frac{\overline{z}_B(s)-s}{\overline{z}_B(s)-\underline{z}_B(s)}\eta(\underline{z}_B(s))+\frac{s-\underline{z}_B(s)}{\overline{z}_B(s)-\underline{z}_B(s)}\eta(\overline{z}_B(s)), & \text{if } \underline{z}_B(s)<\overline{z}_B(s).\end{cases}$$

PROPOSITION 4: *Suppose that* (L*) *holds with* $b(t)=t$*. Delegation set* B *maximizes* W *on* $\mathcal{B}$ *if and only if* $\eta_B=\operatorname{conv}\eta$*.*

To prove Proposition 4, we apply Corollary 2 to recast this delegation problem as a persuasion problem, and then solve it using the standard ironing technique of Myerson (1981). A similar characterization of optimal monotone partitions in linear* persuasion, albeit in a slightly less general case, appears in Rayo (2013, Theorem 1) and Onuchic and Ray (2023, Theorem 1). Proposition 4 also complements the characterization of optimal delegation sets in Kartik et al. (2021). They consider linear* delegation with outside option in the special case where $c(s)=s$, and the principal's utility is type-independent ($\beta=0$) and concave in y ($e(y)$ is increasing), but, unlike in Proposition 4, they allow nonlinear $b(t)$.

7. APPLICATION TO MONOPOLY REGULATION

This section illustrates how the results in Sections 5.3 and 6.2 can be used to characterize the optimality of a quantity floor (which is equivalent to a price cap) in the classical problem of monopoly regulation.

There are a unit mass of consumers with unit demand, a monopolist (agent), and a regulator (principal). The monopolist chooses a quantity $s\in S=[0,1]$ at cost $(1-x)s$, where $x\in[0,1]$ is the monopolist's private type that has a distribution G with a continuously differentiable and strictly positive density g. The monopolist's profit is given by

$$U(s,x)=p(s)s-(1-x)s, \tag{54}$$

where $p(s)$ is the inverse demand. Assume that $p(s)$ is twice continuously differentiable and strictly decreasing on S, and the monopolist's marginal revenue $p(s)+p'(s)s$ is nonnegative and strictly decreasing on S. For convenience, let $p(0)=1$ and $p(1)+p'(1)=0$, so that $p(s)+p'(s)s\in[0,1]$.[31]

The regulator chooses a delegation set to maximize the weighted sum of the consumer and producer surpluses:

$$V(s,x)=\int_0^s(p(\tilde{s})-p(s))\mathrm{d}\tilde{s}+\beta U(s,x), \tag{55}$$

where $\beta\in[0,1]$ is a parameter. Quantities $s=0$ and $s=1$ are always included in the delegation set, because $s=0$ is the monopolist's outside option, whereas $s=1$ is either strictly worse than $s=0$ for the monopolist or it is the first-best choice for the principal, as Claim 5 shows.

[31]The substantive assumption here is that the marginal revenue is nonnegative. The interval of relevant types is equal to the interval of values of the marginal revenue. The types outside this interval are irrelevant as they make trivial choices. We normalize this interval to be $[0,1]$.

CLAIM 5: *For each $x \in [0,1]$, either $U(1,x) < U(0,x)$ or $V(s,x)$ is increasing in s.*

Let $c(s) = 1 - p(s) - p'(s)s$ and $e(c(s)) = p'(s)s + \beta c(s)$.[32] We have

$$u(s,x) = -\frac{\partial U(s,x)}{\partial s} = c(s) - x \quad \text{and} \quad v(s,x) = -\frac{\partial V(s,x)}{\partial s} = e(c(s)) - \beta x.$$

This problem is linear* (i.e., (L*) holds). This problem is also linear (i.e., (L) holds) if the elasticity of the slope of the inverse demand, $p''(s)s/p'(s)$, is constant. Then the inverse demand takes the form of $p(s) = 1 - s^k/(k+1)$ for $k > 0$, in which case, letting $y = s^k$, we obtain

$$u(y,x) = y - x \quad \text{and} \quad v(y,x) = \left(\beta - \frac{k}{k+1}\right) y - \beta x.$$

We now show that optimal delegation takes the form of a quantity floor (that is, any quantity above a certain threshold is permitted) if (1) the elasticity of the slope of the inverse demand is constant, the density of types is log-concave, and the weight on the producer surplus is not too large[33] or (2) this elasticity is decreasing and the density of types is uniform.[34]

COROLLARY 3: *There exists $s^* \in [0,1]$ such that delegation set $\{0\} \cup [s^*, 1]$ is optimal if one of the following two conditions hold:*
(1) $p''(s)s/p'(s) = k - 1$ *for some* $k > 0$, $g'(x)/g(x)$ *is decreasing, and* $\beta \leq 2k/(k+1)$*;*
(2) $p''(s)s/p'(s)$ *is decreasing and* $g(x) = 1$.

REFERENCES

ALONSO, RICARDO AND NIKO MATOUSCHEK (2008): "Optimal Delegation," *Review of Economic Studies*, 75, 259–293. [1, 2, 3, 20, 23, 26]

AMADOR, MANUEL AND KYLE BAGWELL (2013): "The Theory of Optimal Delegation with an Application to Tariff Caps," *Econometrica*, 81 (4), 1541–1599. [1, 2, 3, 21, 26, 28, 29]

——— (2022): "Regulating a Monopolist with Uncertain Costs without Transfers," *Theoretical Economics*, 17, 1719–1760. [3, 21, 26, 29, 31]

AMADOR, MANUEL, KYLE BAGWELL, AND ALEX FRANKEL (2018): "A Note on Interval Delegation," *Economic Theory Bulletin*, 6, 239–249. [26]

ANDERSON, A AND L SMITH (2024): "The Comparative Statics of Sorting," *American Economic Review*, 114, 709–751. [10]

ARIELI, ITAI, YAKOV BABICHENKO, RANN SMORODINSKY, AND TAKURO YAMASHITA (2023): "Optimal Persuasion via Bi-pooling," *Theoretical Economics*, 18, 15–36. [26]

AUMANN, ROBERT J (1965): "Integrals of Set-Valued Functions," *Journal of Mathematical Analysis and Applications*, 12, 1–12. [27]

BARON, DAVID P AND ROGER B MYERSON (1982): "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50, 911–930. [3]

BERGEMANN, DIRK AND STEPHEN MORRIS (2019): "Information Design: A Unified Perspective," *Journal of Economic Literature*, 57, 44–95. [2]

CANDOGAN, OZAN AND PHILIPP STRACK (2023): "Optimal Disclosure of Information to Privately Informed Agents," *Theoretical Economics*, 18 (3), 1225–1269. [2]

[32] Since the marginal revenue is strictly decreasing, c^{-1} is well defined. For each $z = c(s)$, function e is given by $e(z) = p'(c^{-1}(z))c^{-1}(z) + \beta z$.

[33] As this problem is linear, our result actually shows that a quantity floor is optimal among all stochastic delegation mechanisms (see Section 5.1). Amador and Bagwell (2022) also characterize the conditions for the optimality of a quantity floor, but they restrict attention to deterministic delegation mechanisms.

[34] Proposition 3 and Remark 3 easily yield conditions for the optimality of full discretion in the general case.

DESSEIN, WOUTER (2002): "Authority and Communication in Organizations," *Review of Economic Studies*, 69, 811–838. [1]
DWORCZAK, PIOTR AND ANTON KOLOTILIN (2024): "The Persuasion Duality," *Theoretical Economics*, forthcoming. [24]
DWORCZAK, PIOTR AND GIORGIO MARTINI (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127, 1993–2048. [2, 26, 42]
GENTZKOW, MATTHEW AND EMIR KAMENICA (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceedings*, 106, 597–601. [2, 26]
GERSHKOV, ALEX, JACOB K. GOEREE, ALEXEY KUSHNIR, BENNY MOLDOVANU, AND XIANWEN SHI (2013): "On the Equivalence of Bayesian and Dominant Strategy Implementation," *Econometrica*, 81, 197–220. [9]
GOLDSTEIN, ITAY AND YARON LEITNER (2018): "Stress Tests and Information Disclosure," *Journal of Economic Theory*, 177, 34–69. [18]
GOLTSMAN, MARIA, JOHANNES HÖRNER, GREGORY PAVLOV, AND FRANCESCO SQUINTANI (2009): "Mediation, Arbitration and Negotiation," *Journal of Economic Theory*, 144, 1397–1420. [26]
GUO, YINGNI AND ERAN SHMAYA (2019): "The Interval Structure of Optimal Disclosure," *Econometrica*, 87, 653–675. [18]
HALAC, MARINA AND PIERRE YARED (2022): "Fiscal Rules and Discretion Under Limited Enforcement," *Econometrica*, 90 (5), 2093–2127. [3]
HOLMSTRÖM, BENGT (1977): "On Incentives and Control in Organizations," (Ph.D. Thesis, Stanford University). [1, 3, 20, 26]
——— (1984): "On the Theory of Delegation," in *Bayesian Models in Economic Theory*, ed. by M. Boyer and R. Kihlstrom, New York: North-Holland, 115–141. [1, 3, 20, 26]
KAMENICA, EMIR AND MATTHEW GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615. [1, 7, 26]
KARLIN, SAMUEL AND HERMAN RUBIN (1956): "The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio," *Annals of Mathematical Statistics*, 27, 272–299. [12, 49]
KARTIK, NAVIN, ANDREAS KLEINER, AND RICHARD VAN WEELDEN (2021): "Delegation in Veto Bargaining," *American Economic Review*, 111, 4046–4087. [3, 21, 26, 29, 30]
KLEINER, ANDREAS, BENNY MOLDOVANU, AND PHILIPP STRACK (2021): "Extreme Points and Majorization: Economic Applications," *Econometrica*, 89, 1557–1593. [2, 25, 26]
KOLOTILIN, ANTON (2015): "Experimental Design to Persuade," *Games and Economic Behavior*, 90, 215–226. [26]
——— (2018): "Optimal Information Disclosure: A Linear Programming Approach," *Theoretical Economics*, 13, 607–636. [2, 25, 26, 29]
KOLOTILIN, ANTON, ROBERTO CORRAO, AND ALEXANDER WOLITZKY (2024): "Persuasion and Matching: Optimal Productive Transport," *Journal of Political Economy*, forthcoming. [18, 29]
KOLOTILIN, ANTON, TIMOFIY MYLOVANOV, AND ANDRIY ZAPECHELNYUK (2022): "Censorship as Optimal Persuasion," *Theoretical Economics*, 17, 561–585. [26]
KOLOTILIN, ANTON, TYMOFIY MYLOVANOV, ANDRIY ZAPECHELNYUK, AND MING LI (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964. [2, 9, 19, 26]
KOLOTILIN, ANTON AND ANDRIY ZAPECHELNYUK (2019): "Persuasion Meets Delegation," ArXiv:1902.02628. [7, 26, 27]
KOVÁČ, EUGEN AND TYMOFIY MYLOVANOV (2009): "Stochastic Mechanisms in Settings without Monetary Transfers: The Regular Case," *Journal of Economic Theory*, 144, 1373–1395. [2, 26]
MANELLI, ALEJANDRO M. AND DANIEL R. VINCENT (2010): "Bayesian and Dominant-Strategy Implementation in the Independent Private-Values Model," *Econometrica*, 78, 1905–1938. [9]
MARTIMORT, DAVID AND AGGEY SEMENOV (2006): "Continuity in Mechanism Design without Transfers," *Economics Letters*, 93 (2), 182–189. [26]
MELUMAD, NAHUM D AND TOSHIYUKI SHIBANO (1991): "Communication in Settings with No Transfers," *RAND Journal of Economics*, 22, 173–198. [26]
MILGROM, PAUL AND CHRIS SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180. [10]
MÜLLER, ALFRED AND LUDGER RÜSCHENDORF (2001): "On the Optimal Stopping Values Induced by General Dependence Structures," *Journal of Applied Probability*, 38, 672–684. [19, 35]
MÜLLER, ALFRED AND DIETRICH STOYAN (2002): *Comparison Methods for Stochastic Models and Risks*, John Wiley & Sons. [19, 20, 35]
MYERSON, ROGER B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73. [30]
ONUCHIC, PAULA AND DEBRAJ RAY (2023): "Conveying Value via Categories," *Theoretical Economics*, 18, 1407–1439. [3, 30]

OSTROVSKY, MICHAEL AND MICHAEL SCHWARZ (2010): "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, 2, 34–63. [1]

QUAH, JOHN AND BRUNO STRULOVICI (2012): "Aggregating the Single Crossing Property," *Econometrica*, 80, 2333–2348. [10]

RAYO, LUIS (2013): "Monopolistic Signal Provision," *BE Journal of Theoretical Economics*, 13, 27–58. [3, 30]

RAYO, LUIS AND ILYA SEGAL (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118, 949–987. [1, 18]

SARAN, RENE (2024): "A Dynamic Optimization Approach to Delegation with an Application to Volunteer Contracts," SSRN: 4703145. [3, 21, 26]

STRASSEN, VOLKER (1965): "The Existence of Probability Measures with Given Marginals," *Annals of Mathematical Statistics*, 36, 423–439. [19]

ZAPECHELNYUK, ANDRIY (2020): "Optimal Quality Certification," *American Economic Review: Insights*, 2, 161–176. [3, 26]

APPENDIX

APPENDIX A: COMPACTIFICATION IN DETERMINISTIC DELEGATION

Analogously to Section 5.2, this section shows that, under Inada-type assumptions, standard deterministic delegation and deterministic delegation with outside option can be represented as our deterministic delegation problem.

In standard delegation, the set of decisions is $Y_0 = (\underline{y}_0, \overline{y}_0)$. In delegation with outside option, the set of decisions is $Y_0 = [\underline{y}, \overline{y}_0)$, and the agent can always choose the outside option $\underline{y}$.

Say that a delegation set $B \subset Y_0$ is *undominated by* $V_0 : X \to \mathbb{R}$ if

$$V(y^*_B(x), x) \geq V_0(x), \quad \text{for some } x \in X.$$

Propositions 1′ and 2′ are the deterministic counterparts of Propositions 1 and 2.

PROPOSITION 1′: *Suppose that* $Y_0 = (\underline{y}_0, \overline{y}_0) \subseteq \mathbb{R}$, *and*

$$\max_{x \in X} U(y, x) \to -\infty \quad \textit{and} \quad \max_{x \in X} V(y, x) \to -\infty \quad \textit{as } y \to \underline{y}_0 \textit{ and } y \to \overline{y}_0. \tag{56}$$

For each continuous V_0, *there exist* $\underline{y}, \overline{y} \in Y_0$ *(with* $u(\underline{y}, \underline{x}) < 0 < u(\overline{y}, \overline{x})$*) such that the following holds. For each delegation set* B *that is undominated by* V_0, *there exists another delegation set* $\tilde{B} \subset Y = [\underline{y}, \overline{y}]$ *with* $\{\underline{y}, \overline{y}\} \subset \tilde{B}$ *such that* $y^*_{\tilde{B}}(x) = y^*_B(x)$ *for almost all* $x \in X$.

PROOF: Define

$$Z = \left\{ y \in Y_0 : \max_{x \in X} V(y, x) \geq \min_{x \in X} V_0(x) \right\}.$$

If Z is empty, then every $B \subset Y_0$ is dominated by V_0, and the proposition holds trivially. Assume henceforth that Z is nonempty. Let $\underline{z} = \inf Z$ and $\overline{z} = \sup Z$. By (56), compactness of X, and continuity of V and V_0, we have $\underline{y}_0 < \underline{z} \leq \overline{z} < \overline{y}_0$.

Next, let $\varepsilon > 0$ and define

$$\tilde{Y} = \left\{ y \in Y_0 : \max_{x \in X} U(y, x) \geq \min_{x \in X,\, z \in [\underline{z}, \overline{z}]} U(z, x) - \varepsilon \right\}.$$

As $[\underline{z}, \overline{z}] \subset \tilde{Y}$, $\tilde{Y}$ is nonempty. Let $\underline{y} = \inf \tilde{Y}$ and $\overline{y} = \sup \tilde{Y}$. By (56), compactness of X and $[\underline{z}, \overline{z}]$, and continuity of U, we have $\underline{z} > \underline{y} > \underline{y}_0$ and $\overline{z} < \overline{y} < \overline{y}_0$. Let $Y = [\underline{y}, \overline{y}]$. Thus, each type of the agent strictly prefers every decision in $[\underline{z}, \overline{z}]$ to every decision outside of Y. For each $B \subset Y_0$ undominated by V_0, let $\tilde{B} = (B \cap Y) \cup \{\underline{y}, \overline{y}\}$. Then, by (50) and strict aggregate downcrossing of u in x, we have $y^*_{\tilde{B}}(x) = y^*_B(x)$ for almost all $x \in X$. *Q.E.D.*

PROPOSITION 2′: *Suppose that* $Y_0 = [\underline{y}, \overline{y}_0) \subset \mathbb{R}$ *and*

$$\max_{x \in X} U(y, x) \to -\infty \quad \textit{as } y \to \overline{y}_0. \tag{57}$$

Then there exists $\overline{y} \in Y_0$ *(with* $u(\overline{y}, \overline{x}) < 0$*) such that the following holds. For each delegation set* B *with* $\underline{y} \in B$, *there exists another delegation set* $\tilde{B} \subset Y = [\underline{y}, \overline{y}]$ *with* $\{\underline{y}, \overline{y}\} \subset \tilde{B}$ *such that* $y^*_{\tilde{B}}(x) = y^*_B(x)$ *for almost all* $x \in X$.

PROOF: Let $\varepsilon > 0$ and define

$$\overline{y} = \inf\left\{y \in Y_0 : \max_{x\in X} U(y,x) \geq \min_{x\in X} U(\underline{y},x) - \varepsilon\right\}.$$

By (57), compactness of X and continuity of U, we have $\underline{y} < \overline{y} < \overline{y}_0$. Let $Y = [\underline{y}, \overline{y}]$. Thus, each type of the agent strictly prefers decision $\underline{y}$ to every decision outside of Y. For each $B \subset Y_0$ with $\underline{y} \in B$, let $\tilde{B} = (B \cap Y) \cup \{\overline{y}\}$. Then, by (50) and strict aggregate downcrossing of u in x, we have $y^*_{\tilde{B}}(x) = y^*_B(x)$ for almost all $x \in X$. *Q.E.D.*

APPENDIX B: PROOFS

B.1. *Proof of Lemma 4*

Consider a sequence of intervals $(\underline{x}_n, \overline{x}_n) \subset Y$ with $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, let

$$Q_n(x_{n+1}|x_n) = \begin{cases} \delta_{x_n}, & \text{if } x_n \notin (\underline{x}_n, \overline{x}_n), \\ \frac{\overline{x}_n - x_n}{\overline{x}_n - \underline{x}_n}\delta_{\underline{x}_n} + \frac{x_n - \underline{x}_n}{\overline{x}_n - \underline{x}_n}\delta_{\overline{x}_n}, & \text{if } x_n \in (\underline{x}_n, \overline{x}_n), \end{cases}$$

where δ_x, with $x \in Y$, denotes the degenerate distribution at x. Let $H_1 = H$, and for each $n \in \mathbb{N}$, let $H_{n+1}(x_{n+1}) = \int Q_n(x_{n+1}|x_n) H_n(\mathrm{d}x_n)$. This construction gives a finitely supported conditional distribution $P_n(x_{n+1}|x)$ such that $\int P_n(x_{n+1}|x)H(\mathrm{d}x) = H_{n+1}(x_{n+1})$ for all $x_{n+1} \in Y$, and $\int x_{n+1} P_n(\mathrm{d}x_{n+1}|x) = x$ for all $x \in Y$. In the proof of their Theorem 4.1, Müller and Rüschendorf (2001) show that a sequence of intervals $(\underline{x}_n, \overline{x}_n)$ can be chosen in such a way that $P_n(\cdot|x)$ converges weakly to $P(\cdot|x)$ such that $\int P(y|x)H(\mathrm{d}x) = F(y)$ for all $y \in Y$, and $\int y P(\mathrm{d}y|x) = x$ for all $x \in Y$.

Since the likelihood ratio is closed with respect to weak convergence (e.g., Müller and Stoyan, 2002, Theorem 1.4.9), it remains to show that $P_n(x_{n+1}|x)$ increases in x with respect to the likelihood ratio order. Since $P_n(x_{n+1}|x)$ has a finite support, by induction, it suffices to show that, for all intervals $(\underline{z}, \overline{z}) \subset Y$, all finite sets $Z \subset Y$, and all discrete probability densities $h_1(\cdot)$ and $h_2(\cdot)$ supported on Z that are ordered with respect to the likelihood ratio order,

$$h_1(y_1)h_2(y_2) \geq h_2(y_1)h_1(y_2), \quad \text{for all } y_2 > y_1, \tag{58}$$

we have that discrete probability densities $\tilde{h}_l(\cdot)$, with $l = 1, 2$, supported on $\tilde{Z} = Z \cup \{\underline{z}, \overline{z}\} \setminus (\underline{z}, \overline{z})$ and defined by

$$\tilde{h}_l(y) = \begin{cases} \sum_{\tilde{z}\in Z\cap[\underline{z},\overline{z}]} h_l(\tilde{z})\frac{\overline{z}-\tilde{z}}{\overline{z}-\underline{z}}, & \text{if } y = \underline{z}, \\ h_l(y), & \text{if } y \notin [\underline{z}, \overline{z}], \\ \sum_{\tilde{z}\in Z\cap[\underline{z},\overline{z}]} h_l(\tilde{z})\frac{\tilde{z}-\underline{z}}{\overline{z}-\underline{z}}, & \text{if } y = \overline{z}, \end{cases} \tag{59}$$

are also ordered with respect to the likelihood ratio order,

$$\tilde{h}_1(y_1)\tilde{h}_2(y_2) \geq \tilde{h}_2(y_1)\tilde{h}_1(y_2), \quad \text{for all } y_2 > y_1,$$

This follows from direct calculations for all possible cases with $y_2 > y_1$. The only non-trivial case is where $y_1 = \underline{z}$ and $y_2 = \overline{z}$. In this case, we have

$$\tilde{h}_1(\underline{z})\tilde{h}_2(\overline{z}) - \tilde{h}_2(\underline{z})\tilde{h}_1(\overline{z}) = \left(\sum_{\tilde{z}\in Z\cap[\underline{z},\overline{z}]} h_1(\tilde{z})\frac{\overline{z}-\tilde{z}}{\overline{z}-\underline{z}}\right)\left(\sum_{\tilde{z}\in Z\cap[\underline{z},\overline{z}]} h_2(\tilde{z})\frac{\tilde{z}-\underline{z}}{\overline{z}-\underline{z}}\right)$$

$$
\begin{aligned}
&-\left(\sum_{\tilde z\in Z\cap[\underline{z},\overline{z}]} h_2(\tilde z)\frac{\overline{z}-\tilde z}{\overline{z}-\underline{z}}\right)\left(\sum_{\tilde z\in Z\cap[\underline{z},\overline{z}]} h_1(\tilde z)\frac{\tilde z-\underline{z}}{\overline{z}-\underline{z}}\right)\\
&=\sum_{y_1,y_2\in Z\cap[\underline{z},\overline{z}]:\ y_1<y_2}\left(h_1(y_1)h_2(y_2)-h_2(y_1)h_1(y_2)\right)\frac{y_2-y_1}{\overline{z}-\underline{z}}\geq 0,
\end{aligned}
$$

where the first equality is by (59), the second equality is by rearrangement, and the inequality is by (58). *Q.E.D.*

B.2. *Proofs of Claims 1–4*

PROOF OF CLAIM 1: It suffices to show that $H_{\hat\pi}(x)=H(x)$ and $U_{\hat\pi}(x)=\int_x^{\overline{y}}(1-H(\tilde x))\mathrm{d}\tilde x$ for all $x\in Y$. Indeed, for all $x\in Y$, we have

$$
H_{\hat\pi}(x)=\int_Y(1-\hat\pi(y,x))F(\mathrm{d}y)=\int_Y P(x|y)F(\mathrm{d}y)=H(x),
$$

where the first equality is by definition, the second equality is by (28), and the third equality is by (25) and Bayes' rule. Moreover, for all $x\in Y$, we have

$$
\begin{aligned}
U_{\hat\pi}(x)&=\int_Y(y-x)\hat\pi(y,x)F(\mathrm{d}y)=\int_Y(y-x)(1-P(x|y))F(\mathrm{d}y)\\
&=\int_Y(y-x)\left(\int_{(x,\overline{y}]}P(\mathrm{d}\tilde x|y)\right)F(\mathrm{d}y)=\int_{(x,\overline{y}]}\int_Y(y-x)P(\mathrm{d}y|\tilde x)H(\mathrm{d}\tilde x)\\
&=\int_{(x,\overline{y}]}(\tilde x-x)H(\mathrm{d}\tilde x)=\int_{(x,\overline{y}]}(1-H(\tilde x))\mathrm{d}\tilde x,
\end{aligned}
$$

where the first equality is by the definition, the second equality is by (28), the third equality is by the Leibniz rule, the fourth equality is by Bayes' rule, the fifth equality is by (26), and the sixth equality is by integration by parts. *Q.E.D.*

PROOF OF CLAIM 2: Let π_I satisfy (IC$_I$). Since π_I is right-continuous in t, it follows that $\tilde\pi$ given by $\tilde\pi(y,x)=\pi_I(F(y),G(x))$ for all $(y,x)\in Y\times Y$ satisfies (IC) for all $x,\hat x\in[\underline{x},\overline{x})$. By the standard argument, (IC) with $(x,\hat x)\in\{(x_1,x_2),(x_2,x_1)\}$ and $(\hat a_0,\hat a_1)=(0,1)$ yields

$$
-(1-H_{\tilde\pi}(x_1))(x_2-x_1)\leq U_{\tilde\pi}(x_2)-U_{\tilde\pi}(x_1)\leq-(1-H_{\tilde\pi}(x_2))(x_2-x_1),
$$
$$
\text{for all } \underline{x}\leq x_1<x_2<\overline{x}.
$$

Hence, $H_{\tilde\pi}$ is increasing on $[\underline{x},\overline{x})$, and, by the envelope theorem,

$$
U_{\tilde\pi}(x_2)-U_{\tilde\pi}(x_1)=-\int_{x_1}^{x_2}(1-H_{\tilde\pi}(\tilde x))\mathrm{d}\tilde x,\quad \text{for all } \underline{x}\leq x_1<x_2<\overline{x}. \tag{60}
$$

Since $H_{\tilde\pi}$ and $U_{\tilde\pi}$ are monotone on $[\underline{x},\overline{x})$, the left limits $H_{\tilde\pi}(\overline{x}_-)=\lim_{x\uparrow\overline{x}}H_{\tilde\pi}(x)$ and $U_{\tilde\pi}(\overline{x}_-)=\lim_{x\uparrow\overline{x}}U_{\tilde\pi}(x)$ exist. Moreover, there exists a left-continuous function $\phi:Y\to[0,1]$ such that $H_{\tilde\pi}(\overline{x}_-)=\int_Y(1-\phi(y))F(\mathrm{d}y)$ and $U_{\tilde\pi}(\overline{x}_-)=\int_Y(y-\overline{x})\phi(y)F(\mathrm{d}y)$.

Let

$$\underline{x}^* = \min\left\{x \in [\underline{y}, \underline{x}] : U_{\tilde{\pi}}(\underline{x}) + (1 - H_{\tilde{\pi}}(\underline{x}))(\underline{x} - x) \geq \int_Y (y - x)F(\mathrm{d}y)\right\},$$
$$\overline{x}^* = \max\left\{x \in [\overline{x}, \overline{y}] : U_{\tilde{\pi}}(\overline{x}_-) + (1 - H_{\tilde{\pi}}(\overline{x}_-))(\overline{x} - x) \geq 0\right\},$$

which are well-defined because $U_{\tilde{\pi}}(\underline{x}) \geq \int_Y (y-x)F(\mathrm{d}y)$ by (IC) with $(\hat{a}_0, \hat{a}_1) = (1,1)$ and $U_{\tilde{\pi}}(\overline{x}_-) \geq 0$ by (IC) with $(\hat{a}_0, \hat{a}_1) = (0,0)$. Consider π given for each $(y,x) \in Y \times Y$ by

$$\pi(y,x) = \begin{cases} 1, & \text{if } x \in [\underline{y}, \underline{x}^*) \\ \tilde{\pi}(y, \underline{x}), & \text{if } x \in [\underline{x}^*, \underline{x}), \\ \tilde{\pi}(y,x), & \text{if } x \in [\underline{x}, \overline{x}), \\ \phi(y), & \text{if } x \in [\overline{x}, \overline{x}^*), \\ 0, & \text{if } x \in [\overline{x}^*, \overline{y}], \end{cases} \tag{61}$$

Since $\pi(y,x) = \tilde{\pi}(y,x)$ for all $(y,x) \in Y \times [\underline{x}, \overline{x})$, we have $H_\pi(x) = H_{\tilde{\pi}}(x)$ and $U_\pi(x) = U_{\tilde{\pi}}(x)$ for all $x \in [\underline{x}, \overline{x})$. By (61) and the monotonicity of $H_{\tilde{\pi}}$ on $[\underline{x}, \overline{x})$, H_π is increasing on Y.

Next, we have:

$$U_\pi(x) = 0 = \int_x^{\overline{y}} (1 - H_\pi(\tilde{x}))\mathrm{d}\tilde{x}, \quad \text{for all } x \in [\overline{x}^*, \overline{y}], \tag{62}$$

where the equalities are by the definition of π, H_π, and U_π;

$$U_\pi(x) = (\overline{x}^* - x)(1 - H_\pi(\overline{x})) = \int_x^{\overline{y}} (1 - H_\pi(\tilde{x}))\mathrm{d}\tilde{x}, \quad \text{for all } x \in [\overline{x}, \overline{x}^*), \tag{63}$$

where the first equality is by the definition of π, H_π, U_π, and $\overline{x}^*$, and the second is by (62);

$$U_\pi(x) = U_\pi(\overline{x}) + \int_x^{\overline{x}} (1 - H_{\tilde{\pi}}(\tilde{x}))\mathrm{d}\tilde{x} = \int_x^{\overline{y}} (1 - H_{\tilde{\pi}}(\tilde{x}))\mathrm{d}\tilde{x}, \quad \text{for all } x \in [\underline{x}, \overline{x}), \tag{64}$$

where the first equality is by (60) and the definition of $\pi(y, \overline{x})$, and the second is by (63);

$$U_\pi(x) = U_\pi(\underline{x}) + (\underline{x} - x)(1 - H(\overline{x})) = \int_x^{\overline{y}} (1 - H_{\tilde{\pi}}(\tilde{x}))\mathrm{d}\tilde{x}, \quad \text{for all } x \in [\underline{x}^*, \underline{x}), \tag{65}$$

where the first equality is by the definition of π, H_π, and U_π, and the second is by (64);

$$U_\pi(x) = U_\pi(\underline{x}^*) + (\underline{x}^* - x)(1 - H(\underline{y})) = \int_x^{\overline{y}} (1 - H_{\tilde{\pi}}(\tilde{x}))\mathrm{d}\tilde{x}, \quad \text{for all } x \in [\underline{y}, \underline{x}^*), \tag{66}$$

where the first equality is by the definition of π, H_π, U_π, and $\underline{x}^*$, and the second is by (65);

$$U_\pi(\underline{y}) = \int_Y (y - \underline{y})F(\mathrm{d}y) = \int_Y (1 - F(y))\mathrm{d}y, \tag{67}$$

where the first equality is by the definition of π and U_π, and the second is by integration by parts. Since π satisfies (24), it satisfies (IC) by Lemma 3. *Q.E.D.*

PROOF OF CLAIM 3: Let $j = 0, 1$ and let $\pi \in \Pi_j$. The principal's interim utility is

$$\begin{aligned} V_\pi(x) &= \alpha \int_x^{\overline{y}} (1 - H_\pi(\tilde{x}))\mathrm{d}\tilde{x} + (\alpha x - d(x))(1 - H_\pi(x)) \\ &= -\alpha x \int_{(x,\overline{y}]} H_\pi(\mathrm{d}\tilde{x}) + \alpha \int_{(x,\overline{y}]} \tilde{x} H_\pi(\mathrm{d}\tilde{x}) + (\alpha x - d(x)) \int_{(x,\overline{y}]} H_\pi(\mathrm{d}\tilde{x}) \\ &= \int_{(x,\overline{y}]} (\alpha\tilde{x} - d(x)) H_\pi(\mathrm{d}\tilde{x}), \quad \text{for all } x \in X, \end{aligned} \tag{68}$$

where the first equality is by (22), Lemmas 1 and 3, and Claim 2, the second equality is by integration by parts, and the last equality is by rearrangement. So, the principal's expected utility is

$$\begin{aligned} W(\pi) &= \int_X \left(\int_{(x,\overline{y}]} (\alpha\tilde{x} - d(x)) H_\pi(\mathrm{d}\tilde{x}) \right) g(x)\mathrm{d}x \\ &= \int_X \left(\int_{\underline{x}}^{\tilde{x}} (\alpha\tilde{x} - d(x)) g(x)\mathrm{d}x \right) H_\pi(\mathrm{d}\tilde{x}) + \int_{(\overline{x},\overline{y}]} \left(\int_X (\alpha\tilde{x} - d(x)) g(x)\mathrm{d}x \right) H_\pi(\mathrm{d}\tilde{x}) \\ &= \int_Y \nu(\tilde{x}) H_\pi(\mathrm{d}\tilde{x}), \end{aligned}$$

where the first equality is by (68) and the definition of $W(\pi)$, the second equality is by Fubini's theorem, and the third equality is by the definition of ν. *Q.E.D.*

PROOF OF CLAIM 4: Let $j = 0, 1$ and let $\pi \in \Pi_j$.

If. Suppose that H_π maximizes $\int_Y \nu(x)H(\mathrm{d}x)$ over distributions H that satisfy (MPS). We show that the principal's expected utility is higher under π than under any $\tilde{\pi} \in \Pi_j$. As $\tilde{\pi} \in \Pi_j$, it satisfies (IC_0)–(IC_2). Observe that H_π and $H_{\tilde{\pi}}$ satisfy (MPS) by Lemmas 1 and 3, Claim 2, and condition (23). The claim follows from

$$W(\tilde{\pi}) = \int_Y \nu(x) H_{\tilde{\pi}}(\mathrm{d}x) \leq \int_Y \nu(x) H_\pi(\mathrm{d}x) = W(\pi),$$

where the equalities hold by Claim 3, and the inequality holds by the optimality of H_π. Hence π is optimal on Π_j.

Only if. Suppose that π is optimal on Π_j. Consider any distribution H on Y that satisfies (MPS). By Lemma 1 and Claim 1, there exists a (monotone) delegation mechanism $\hat{\pi}$ on $Y \times Y$ with $H_{\hat{\pi}} = H$ that satisfies (IC_0)–(IC_2) for all $x, \hat{x} \in Y$. We have

$$\int_Y \nu(x) H_\pi(\mathrm{d}x) = W(\pi) \geq W(\hat{\pi}) = \int_Y \nu(x) H_{\hat{\pi}}(\mathrm{d}x) = \int_Y \nu(x) H(\mathrm{d}x),$$

where the first and second equalities are by Claim 3, the inequality is by the optimality of π, and the third equality is by $H_{\hat{\pi}} = H$. Thus, H_π maximizes $\int_Y \nu(x)H(\mathrm{d}x)$ over distributions H that satisfy (MPS). *Q.E.D.*

B.3. *Proofs of Propositions 1 and 2*

To simplify notation, we prove Propositions 1 and 2 using the original decision variable s with $S = \mathbb{R}$ in standard delegation and $S = [\underline{s}, \infty)$ in delegation with outside option. When we refer to constraints (IC_0)–(IC_2) and conditions (33)–(34), it is understood that they are expressed in variable s rather than y.

We first introduce some notations and prove a lemma. Let $k : X \to \mathbb{R}$ and $\ell : X \to \mathbb{R}$ be continuous functions. Define

$$
\begin{gathered}
Z = \left\{ s \in S : \max_{x \in X} \big(k(x)s - \ell(x) - C(s) \big) \geq 0 \right\}, \\
\underline{z} = \inf Z, \quad \overline{z} = \sup Z.
\end{gathered}
\tag{69}
$$

Note that $\underline{z}$ and $\overline{z}$ can be infinite if Z is unbounded or empty. Also, define

$$
\Lambda = \left\{ \lambda \in \Delta(S) : \max_{x \in X} \mathbb{E}_\lambda \big[k(x)s - \ell(x) - C(s) \big] \geq 0 \right\}. \tag{70}
$$

CLAIM 6: *If $\underline{z}$ and $\overline{z}$ are finite, and*

$$
\max_{x \in X} \big(k(x)\underline{z} - \ell(x) - C(\underline{z}) \big) = 0 \quad \textit{and} \quad \max_{x \in X} \big(k(x)\overline{z} - \ell(x) - C(\overline{z}) \big) = 0, \tag{71}
$$

then, for each $\lambda \in \Lambda$, there exists $\hat{\lambda} \in \Lambda$ such that $\operatorname{supp}(\hat{\lambda}) \subset [\underline{z}, \overline{z}]$, $\mathbb{E}_{\hat{\lambda}}[s] = \mathbb{E}_\lambda[s]$, *and* $\mathbb{E}_{\hat{\lambda}}[C(s)] = \mathbb{E}_\lambda[C(s)]$.

PROOF: Let $L(s) = \max_{x \in X} \big(k(x)s - \ell(x) \big)$. For all $s \in S$ and all $\lambda \in \Delta(S)$, we have

$$
\begin{aligned}
&\max_{x \in X} \big(k(x)s - \ell(x) - C(s) \big) = L(s) - C(s), \\
&\max_{x \in X} \mathbb{E}_\lambda \big[k(x)s - \ell(x) - C(s) \big] = L(\mathbb{E}_\lambda[s]) - \mathbb{E}_\lambda[C(s)].
\end{aligned}
\tag{72}
$$

Fix $\lambda \in \Lambda$. We have

$$
0 \leq L(\mathbb{E}_\lambda[s])] - \mathbb{E}_\lambda[C(s)] \leq L(\mathbb{E}_\lambda[s])] - C(\mathbb{E}_\lambda[s]),
$$

where the first inequality is by (72) and $\lambda \in \Lambda$, and the second inequality is by the convexity of C. Then, by (72) and the definition of $\underline{z}$ and $\overline{z}$, $\mathbb{E}_\lambda[s] \in [\underline{z}, \overline{z}]$. Moreover, if $\mathbb{E}_\lambda[s] = \underline{z}$ or $\mathbb{E}_\lambda[s] = \overline{z}$, then $\lambda = \delta_z$, so $\hat{\lambda} = \lambda$ is as required. Assume henceforth that $\underline{z} < \mathbb{E}_\lambda[s] < \overline{z}$. Let

$$
\theta = \frac{\mathbb{E}_\lambda[s] - \underline{z}}{\overline{z} - \underline{z}} \quad \text{and} \quad \tau = \frac{(1-\theta)C(\underline{z}) + \theta C(\overline{z}) - \mathbb{E}_\lambda[C(s)]}{(1-\theta)C(\underline{z}) + \theta C(\overline{z}) - C(\mathbb{E}_\lambda[s])}.
$$

We claim that $\tau \in [0, 1]$. Indeed, the numerator is smaller than the denominator because C is convex. Moreover, the denominator is strictly positive because $\underline{z} < \mathbb{E}_\lambda[s] < \overline{z}$ and C is strictly convex. Finally, the numerator is positive, because

$$
\begin{aligned}
(1-\theta)C(\underline{z}) + \theta C(\overline{z}) - \mathbb{E}_\lambda[C(s)] &= (1-\theta)L(\underline{z}) + \theta L(\overline{z}) - \mathbb{E}_\lambda[C(s)] \\
&\geq (1-\theta)L(\underline{z}) + \theta L(\overline{z}) - L(\mathbb{E}_\lambda[s]) \geq 0,
\end{aligned}
$$

where the equality is by (71), the first inequality is by the definition of Λ, and the second inequality is by the convexity of $L(s)$ and the definition of θ, which implies that $(1-\theta)\underline{z} + \theta\overline{z} = \mathbb{E}_\lambda[s]$.

Let $\hat{\lambda} \in \Lambda$ be given by

$$\hat{\lambda} = \tau \delta_{\mathbb{E}_\lambda[s]} + (1-\tau)(1-\theta)\delta_{\underline{z}} + (1-\tau)\theta\delta_{\overline{z}}.$$

By construction, $\text{supp}(\hat{\lambda}) \subset [\underline{z}, \overline{z}]$. Moreover, by the definition of θ and τ, we have $\mathbb{E}_{\hat{\lambda}}[s] = \mathbb{E}_\lambda[s]$ and $\mathbb{E}_{\hat{\lambda}}[C(s)] = \mathbb{E}_\lambda[C(s)]$. *Q.E.D.*

PROOF OF PROPOSITION 1: Let $k(x) = d(x)/\alpha$ and $\ell(x) = V_0(x)/\alpha$. By (29), we have

$$V(s,x) - V_0(x) = \alpha\big(k(x)s - \ell(x) - C(s)\big).$$

Observe that $Z \subset \mathbb{R}$ and $\Lambda \subset \Delta(\mathbb{R})$, given by (69) and (70), are the sets of deterministic and stochastic decisions that are undominated by V_0.

We now show that $\underline{z}$ and $\overline{z}$, given by (69), are finite. If Z is empty, that is, $V(s,x) - V_0(x) < 0$ for all $s \in \mathbb{R}$ and all $x \in X$, then there is no mechanism that is undominated by V_0, and the proposition holds trivially. Assume henceforth that Z is nonempty, and thus $\underline{z} \le \overline{z}$. By (33) and continuity of d and V_0, we have $\underline{z} > -\infty$ and $\overline{z} < \infty$.

Next, define

$$\underline{s}^* = \inf\left\{ s \in \mathbb{R} : \max_{x \in X,\, z \in [\underline{z}, \overline{z}]} (U(s,x) - U(z,x)) \ge 0 \right\},$$

$$\overline{s}^* = \sup\left\{ s \in \mathbb{R} : \max_{x \in X,\, z \in [\underline{z}, \overline{z}]} (U(s,x) - U(z,x)) \ge 0 \right\}.$$

By (33) and compactness of $[\underline{z}, \overline{z}]$, we have $\underline{z} \ge \underline{s}^* > -\infty$ and $\overline{z} \le \overline{s}^* < \infty$. Let $S^* = [\underline{s}^*, \overline{s}^*]$. By the definition of S^*, each type of the agent strictly prefers every decision in $[\underline{z}, \overline{z}]$ — and, thus, every lottery over $[\underline{z}, \overline{z}]$ — to every decision $s \notin S^*$.

Let π be a mechanism that satisfies (IC_0) and is undominated by V_0. Because $\underline{z}$ and $\overline{z}$ are finite and $S = \mathbb{R}$, condition (71) of Claim 6 is satisfied. Thus, by Claim 6, there exists a mechanism $\tilde{\pi}$ such that $\text{supp}(\tilde{\pi}) \subset [\underline{z}, \overline{z}] \subset S^*$, and

$$\mathbb{E}_{\tilde{\pi}(\cdot|x)}[s] = \mathbb{E}_{\pi(\cdot|x)}[s] \quad \text{and} \quad \mathbb{E}_{\tilde{\pi}(\cdot|x)}[C(s)] = \mathbb{E}_{\pi(\cdot|x)}[C(s)], \qquad \text{for all } x \in X.$$

By (29), equalities (31) hold for all $x \in X$. Finally, let $\varepsilon_1, \varepsilon_2 > 0$ be such that $c(\underline{s}^* - \varepsilon_1) < \underline{x}$ and $c(\overline{s}^* + \varepsilon_2) > \overline{x}$, and let $Y = [\underline{y}, \overline{y}] = [c(\underline{s}^* - \varepsilon_1), c(\overline{s}^* + \varepsilon_2)]$. By the definition of $\overline{s}^*$ and $\underline{s}^*$ and (33), $\tilde{\pi}$ satisfies (IC_1) and (IC_2) with strict inequalities. *Q.E.D.*

PROOF OF PROPOSITION 2: Let $k(x) = x$ and $\ell(x) = x\underline{s} - C(\underline{s})$. Thus, by (29), we have

$$U(s,x) - U(\underline{s},x) = k(x)s - \ell(x) - C(s). \tag{73}$$

Observe that $Z \subset [\underline{s}, \infty]$ and $\Lambda \subset \Delta([\underline{s}, \infty))$, given by (69) and (70), are the sets of deterministic and stochastic decisions that are preferred to the outside option $\underline{s}$ by at least one type of the agent.

Let π be a mechanism that satisfies (IC_0)–(IC_1). Observe that $\underline{z} = \underline{s}$, and, by (34), we have $\overline{z} < \infty$. Then, by $S = [\underline{s}, \infty)$, (69), and (73), the condition (71) of Claim 6 is satisfied. Thus, by Claim 6, there exists a mechanism $\tilde{\pi}$ such that $\text{supp}(\tilde{\pi}) \subset [\underline{z}, \overline{z}]$, and

$$\mathbb{E}_{\tilde{\pi}(\cdot|x)}[s] = \mathbb{E}_{\pi(\cdot|x)}[s] \quad \text{and} \quad \mathbb{E}_{\tilde{\pi}(\cdot|x)}[C(s)] = \mathbb{E}_{\pi_D(\cdot|x)}[C(s)], \qquad \text{for all } x \in X.$$

By (29), equalities (31) holds for all $x \in X$. Finally, let $\varepsilon > 0$ be such that $c(\overline{z} + \varepsilon) > \overline{x}$, and let $Y = [\underline{y}, \overline{y}] = [c(\underline{z}), c(\overline{z} + \varepsilon)]$. By the definition of $\overline{z}$ and (34), $\tilde{\pi}$ satisfies (IC_2) with strict inequality. *Q.E.D.*

B.4. *Proof of Theorem 3*

Let $j = 0,1$ and let $\pi \in \Pi_j$. We first prove three simple claims.

CLAIM 7: *Let $\nu' \in \boldsymbol{\nu'}$. If p satisfies (40), then it is continuous and convex.*

PROOF: Let p be given by (40). Then $|\nu'(x)| \leq L$ for all $x \in Y$ with $L \in \mathbb{R}$ given by

$$L = \sup_{x \in Y} |(\alpha x - d(x))g(x) + \alpha G(x)|,$$

where $L \in \mathbb{R}$, because d is continuous, g is càdlàg, and Y is compact. Hence, by (40), for each $y \in Y$, there exists a converging sequence x_n in $\operatorname{supp}(H_\pi)$, with converging $\nu'(x_n)$, such that

$$p(z) \geq \lim_{n\to\infty} (\nu(x_n) + \nu'(x_n)(z - x_n)), \quad \text{for all } z \in Y \text{, with equality at } z = y. \tag{74}$$

Then, p is continuous, because, by (74), for all $z \in Y$,

$$\begin{aligned} p(y) - p(z) &\leq \lim_{n\to\infty} (\nu(x_n) + \nu'(x_n)(y - x_n) - \nu(x_n) - \nu'(x_n)(z - x_n)) \\ &= \lim_{n\to\infty} \nu'(x_n)(y - z) \leq L|y - z|. \end{aligned}$$

Also, p is convex, since, by (74), for all $z, z' \in Y$ and all $\rho \in [0,1]$ with $\rho z + (1-\rho)z' = y$,

$$p(y) - \rho p(z) - (1-\rho)p(z') \leq \lim_{n\to\infty} \nu'(x_n)(y - \rho z - (1-\rho)z') = 0.$$

Q.E.D.

CLAIM 8: *If p is convex and satisfies (41) and (42), then, for all distributions H on Y that satisfy (MPS), we have*

$$\int_Y \nu(x)H(\mathrm{d}x) \leq \int_Y p(x)H(\mathrm{d}x) \leq \int_Y p(y)F(\mathrm{d}y) = \int_Y \nu(x)H_\pi(\mathrm{d}x). \tag{75}$$

PROOF: Let p be convex and satisfy (41) and (42), and let H be a distribution that satisfies (MPS). The first inequality holds by (41), the second inequality holds because p is convex and H satisfies (MPS), and the equality holds by (42). *Q.E.D.*

CLAIM 9: *If p is continuous and convex, and satisfies (41) and (42), then there exists a selection ν' from $\boldsymbol{\nu'}$ such that p_π given by (40) satisfies $p(y) \geq p_\pi(y)$ for all $y \in Y$.*

PROOF: Let p be continuous and convex, and satisfy (41) and (42). Observe that for $H = H_\pi$, all inequalities in (75) hold with equality. Hence, by the continuity of ν and p, we have

$$p(x) = \nu(x), \quad \text{for all } x \in \operatorname{supp}(H_\pi). \tag{76}$$

Fix $x \in \operatorname{supp}(H_\pi)$ such that $x < \overline{y}$. For all $y \in (x, \overline{y}]$ and all $\varepsilon \in (0,1]$, we have

$$\frac{p(y) - p(x)}{y - x} \geq \frac{p(x + \varepsilon(y - x)) - p(x)}{\varepsilon(y - x)} \geq \frac{\nu(x + \varepsilon(y - x)) - \nu(x)}{\varepsilon(y - x)},$$

where the first inequality is by the convexity of p and the second inequality is by (41) and (76). Taking the limit $\varepsilon \downarrow 0$ implies that

$$\frac{p(y) - p(x)}{y - x} \geq \nu'(x_+), \quad \text{for all } y \in (x, \overline{y}].$$

Since p is convex, taking the limit $y \downarrow x$ implies that $p'(x_+)$ is well defined and satisfies $p'(x_+) \geq \nu'(x_+)$. By a symmetric argument, for all $x \in \mathrm{supp}(H_\pi)$ such that $x > \underline{y}$, we have that $p'(x_-)$ is well defined and satisfies $p'(x_-) \leq \nu'(x_-)$.

If $x = \underline{y} \in \mathrm{supp}(H_\pi)$, then

$$p(y) \geq p(\underline{y}) + p'(\underline{y}_+)(y - \underline{y}) \geq \nu(\underline{y}) + \nu'(\underline{y})(y - \underline{y}), \quad \text{for all } y \in Y,$$

where the first inequality is by the convexity of p and the second inequality is by (41) and $p'(\underline{y}_+) \geq \nu'(\underline{y}_+) = \nu'(\underline{y})$. Similarly, if $x = \overline{y} \in \mathrm{supp}(H_\pi)$, then

$$p(y) \geq p(\overline{y}) + p'(\overline{y}_-)(y - \overline{y}) \geq \nu(\overline{y}) + \nu'(\overline{y})(y - \overline{y}), \quad \text{for all } y \in Y.$$

Finally, consider $x \in \mathrm{supp}(H_\pi)$ such that $x \in (\underline{y}, \overline{y})$. By the convexity of p, we have $p'(x_-) \leq p'(x_+)$. In both cases $\nu'(x_-) < \nu'(x_+)$ and $\nu'(x_-) \geq \nu'(x_+)$, the inequalities $p'(x_-) \leq \nu'(x_-)$, $p'(x_+) \geq \nu'(x_+)$, and $p'(x_-) \leq p'(x_+)$ imply that there exists $\nu'(x) \in \boldsymbol{\nu}'(x)$ such that $p'(x_-) \leq \nu'(x) \leq p'(x_+)$. Then, by the convexity of p and (41), we have

$$\begin{aligned} p(y) &\geq p(x) - p'(x_-)(x - y) \geq \nu(x) - \nu'(x)(x - y), \quad \text{for all } y < x, \\ p(y) &\geq p(x) + p'(x_+)(y - x) \geq \nu(x) + \nu'(x)(y - x), \quad \text{for all } y > x. \end{aligned}$$

In sum, there exists a selection ν' from $\boldsymbol{\nu}'$ such that

$$p(y) \geq \nu(x) + \nu'(x)(y - x), \quad \text{for all } x \in \mathrm{supp}(H_\pi) \text{ and all } y \in Y.$$

Thus, p_π given by (40) satisfies $p \geq p_\pi$. *Q.E.D.*

We now prove Theorem 3.

If. Suppose that there exists a selection ν' from $\boldsymbol{\nu}'$ such that p given by (40) satisfies (41) and (42). Then π is optimal by Claims 4, 7, and 8.

Only if. Suppose that π is optimal. By Claim 4, H_π maximizes $\int_Y \nu(x)H(\mathrm{d}x)$ over distributions H that satisfy (MPS). Thus, since ν given by (36) is Lipschitz continuous, Theorem 2 in Dworczak and Martini (2019) implies that there exists a continuous and convex function p on Y that satisfies (41) and (42). Next, by Claim 9, there exists a selection ν' from $\boldsymbol{\nu}'$ such that p_π given by (40) satisfies $p(y) \geq p_\pi(y)$ for all $y \in Y$. Then,

$$\int_Y p(y)F(\mathrm{d}y) = \int_Y \nu(x)H_\pi(\mathrm{d}x) \leq \int_Y p_\pi(x)H_\pi(\mathrm{d}x) \leq \int_Y p_\pi(y)F(\mathrm{d}y) \leq \int_Y p(y)F(\mathrm{d}y),$$

where the equality holds by (42), the first inequality holds because $p_\pi(x) \geq \nu(x)$ for all $x \in \mathrm{supp}(H_\pi)$ by (40), the second inequality holds because p_π is convex by Claim 7 and H_π satisfies (MPS), and the last inequality holds by $p \geq p_\pi$. So all inequalities hold with equality. Thus, $p = p_\pi$, by the continuity of p and p_π. *Q.E.D.*

B.5. *Proof of Remark 2*

Let $j = 0, 1$ and let $\pi \in \Pi_j$ be monotone. The marginal distributions of J_π are F and H_π, because

$$\begin{aligned} J_\pi(y, \overline{y}) &= \int_{\underline{y}}^{y} (1 - \pi(\tilde{y}|\overline{y}))F(\mathrm{d}\tilde{y}) = \int_{\underline{y}}^{y} F(\mathrm{d}\tilde{y}) = F(y), \quad \text{for all } y \in Y, \\ J_\pi(\overline{y}, x) &= \int_{\underline{y}}^{\overline{y}} (1 - \pi(y|x))F(\mathrm{d}y) = H_\pi(x), \quad \text{for all } x \in Y, \end{aligned} \tag{77}$$

where the first equalities in both lines hold by (43), the second equality in the first line holds by (IC_1), and the second equality in the second line holds by (38).

Consider π_P given by $\pi_P(x|y) = \pi(y|x)$ for all $y \in Y$ and all $x \in Y$. By assumption, $\pi(y|x)$ is a monotone delegation mechanism, so it is increasing and left-continuous in y and decreasing and right-continuous in x, and satisfies $\pi(\tilde{y}|\overline{y}) = 0$ by (IC_1). Thus, π_P is a monotone persuasion mechanism. By the definition of persuasion mechanisms and (43),

$$J_\pi(y,x) = \int_{\underline{y}}^{y} (1 - \pi_P(x|\tilde{y}))F(\mathrm{d}\tilde{y}) = \mathbb{P}(\text{state} < y, \text{decision} \leq x), \quad \text{for all } (y,x) \in Y \times Y.$$

By Lemmas 1 and 2 and Claim 2, for all $\hat{x} \in Y$ and H_π-almost all $x \in Y$, we have

$$\int_Y \left(\int_{\underline{y}}^{x} (y - \tilde{x})\mathrm{d}\tilde{x} \right) \pi_P(\mathrm{d}y|x) \geq \int_Y \left(\int_{\underline{y}}^{\hat{x}} (y - \tilde{x})\mathrm{d}\tilde{x} \right) \pi_P(\mathrm{d}y|x). \tag{78}$$

Condition (78) implies the first-order condition

$$\int_Y (y - x)\pi_P(\mathrm{d}y|x) = 0, \quad \text{for } H_\pi\text{-almost all } x \in Y.$$

Then, for all functions $\phi : Y \to \mathbb{R}$, we have

$$\int_{Y \times Y} \phi(x)(y - x)J_\pi(\mathrm{d}y, \mathrm{d}x) = \int_X \phi(x) \int_Y (y - x)\,(-\pi_P(\mathrm{d}y|x))H_\pi(\mathrm{d}x) = 0. \tag{79}$$

Fix $p \in P_\pi$. Then

$$\begin{aligned}\int_Y \nu(x)H_\pi(\mathrm{d}x) &= \int_{Y \times Y} \nu(x)J_\pi(\mathrm{d}y, \mathrm{d}x) = \int_{Y \times Y} (\nu(x) + \nu'(x)(y - x))J_\pi(\mathrm{d}y, \mathrm{d}x) \\ &\leq \int_{Y \times Y} p(y)J_\pi(\mathrm{d}y, \mathrm{d}x) = \int_Y p(y)F(\mathrm{d}y),\end{aligned} \tag{80}$$

where the first and last equalities are by (77), the second equality is by (79), and the inequality is by (40). Thus, (42) holds if and only if the inequality holds with equality, which is equivalent to (45).

Let $j = 0, 1$, let $\pi \in \Pi_j$ be deterministic, and let $B \subset Y$ be a corresponding compact delegation set. Then an extension of π from $Y \times X$ to $Y \times Y$ that satisfies (IC_0)–(IC_2) for all $x, \hat{x} \in Y$ is given by

$$\pi(y|x) = \mathbf{1}\{x_B^*(y) > x\}, \quad \text{for all } (y,x) \in Y \times Y. \tag{81}$$

Then, by (38), (43) and (81),

$$H_\pi(x) = \int_Y \mathbf{1}\{x_B^*(y) \leq x\}F(\mathrm{d}y) \quad \text{and} \quad J_\pi(y,x) = \int_{\underline{y}}^{y} \mathbf{1}\{x_B^*(\tilde{y}) \leq x\}F(\mathrm{d}\tilde{y}),$$
$$\text{for all } y \in Y \text{ and all } x \in Y.$$

Since $f(y) > 0$ for all $y \in Y$ and H_π is a distribution of $x_B^*(y)$ where y has distribution F, we have

$$x_B^*(y) \in \mathrm{supp}(H_\pi), \quad \text{for all } y \in Y. \tag{82}$$

We thus obtain

$$\begin{aligned}\int_Y \nu(x)H_\pi(\mathrm{d}x) &= \int_{Y\times Y} \nu(x)J_\pi(\mathrm{d}y,\mathrm{d}x) = \int_{Y\times Y} (\nu(x)+\nu'(x)(y-x))J_\pi(\mathrm{d}y,\mathrm{d}x) \\ &= \int_Y (\nu(x_B^*(y))+\nu'(x_B^*(y))(y-x_B^*(y)))F(\mathrm{d}y) \le \int_Y p(y)F(\mathrm{d}y),\end{aligned} \tag{83}$$

where the first equality is by (77), the second equality is by (79), the third equality is because J_π is a joint distribution of $(y, x_B^*(y))$ where y has distribution F, and the inequality is by (40) and (82). Thus, (42) holds if and only if inequality (83) holds with equality, which is equivalent to (46) by (40) and the continuity of p on Y. *Q.E.D.*

B.6. *Proof of Corollary 1*

We first prove Corollary 1 for delegation with outside option, and then explain how the proof changes in standard delegation.

Only if. Suppose that delegation set $\{\underline{y}\}\cup[y^*,\overline{y}_0)$ is optimal. As follows from Proposition 2, there exists $Y=[\underline{y},\overline{y}]\subset[\underline{y},\overline{y}_0)$, such that the agent's best response for all $x\in X$ is the same under delegation sets $\{\underline{y}\}\cup[y^*,\overline{y}_0)$ and $B=\{\underline{y}\}\cup[y^*,\overline{y}]$. By (44), we have

$$x_B^*(y)=\begin{cases} y, & \text{if } y\ge y^*,\\ z^*, & \text{if } y<y^*,\end{cases} \quad \text{where} \quad z^*=\frac{1}{F(y^*)-F(\underline{y})}\int_{\underline{y}}^{y^*} F(\mathrm{d}y).$$

Hence, by (46), we have

$$p(y)=\begin{cases} \nu(y), & \text{for all } y\in(y^*,\overline{y}],\\ \nu(z^*)+\nu'(z^*)(y-z^*), & \text{for all } y\in[\underline{y},y^*].\end{cases} \tag{84}$$

We thus obtain (a) by (40) and Claim 7, and (b) by (41).

If. Suppose that conditions (a) and (b) hold. Then, p given by (84) satisfies (40), (41), and (46), so delegation set $\{\underline{y}\}\cup[y^*,\overline{y}_0)$ is optimal.

In standard delegation, for the *only if* part, we instead apply Proposition 1 with V_0 given by (32) and observe that $z^*\in(\underline{y},\underline{x})$, because the agent with any type $x\in X$ strictly prefers y^* to $\underline{y}$. The rest of the proof is the same. *Q.E.D.*

B.7. *Proof of Corollary 2*

Let $B\in\mathcal{B}$ be a monotone partition in persuasion. Since u_P is strictly aggregate down-crossing in x, $x_B^*(y)$ given by (49) is uniquely defined for all $y\in Y$. Redefine $x_B^*(\underline{y})=\underline{x}$, which is w.l.o.g. because state $y=\underline{y}$ occurs with zero probability and is always revealed, as $\underline{z}_B(\underline{y})=\overline{z}_B(\underline{y})=\underline{y}$ by definition. The corresponding persuasion mechanism is given by

$$\pi_P(x|y)=\mathbf{1}\{x_B^*(y)>x\}, \quad \text{for all } y\in Y \text{ and all } x\in X.$$

Observe that (i) π_P is left-continuous in y, because partition intervals $(\underline{z}_B(y),\overline{z}_B(y)]$ are closed on the right; (ii) π_P is right-continuous in x because π_P is defined using a strict inequality; (iii) π_P satisfies the normalization $\pi_P(x|\underline{y})=0$ for all $x\in X$ because $x_B^*(\underline{y})=\underline{x}$; (iv) π_P is monotone, because u_P is upcrossing in y and thus $x_B^*(y)$ is increasing in y; (v) π_P is incentive-compatible, because the agent's decision is optimal for each partition element, except possibly for $\{\underline{y}\}$. In sum, mechanism π_P is monotone, deterministic, and incentive-compatible.

Consider deterministic mechanisms π_D and π_I given by $\pi_D(y|x) = \pi_I(y,x) = \pi_P(x|y)$ for all $y \in Y$ and all $x \in X$. Note that

$$\pi_I(y,x) = \pi_D(y|x) = \mathbf{1}\{y^*(x) < y\}, \quad \text{for all } y \in Y \text{ and all } x \in X,$$

$$\text{where} \quad y^*(x) = \begin{cases} \inf\{y \in Y : x_B^*(y) > x\}, & \text{if } x < x_B^*(\overline{y}), \\ \overline{y}, & \text{if } x \geq x_B^*(\overline{y}). \end{cases}$$

By Theorem 1, π_D and π_I are incentive-compatible and satisfy $W_P(\pi_P) = W_D(\pi_D) = W_I(\pi_I)$. Since π_D is incentive-compatible, we have, for almost all $x \in X$,

$$U_D(y^*(x),x) \geq \max\{\sup_{\hat{x}\in X} U_D(y^*(\hat{x}),x), U_D(\underline{y},x), U_D(\overline{y},x)\} = \max_{y\in B} U_D(y,x),$$

and thus $y^*(x)$ satisfies (50). Next, by strict aggregate downcrossing of u_D in x, y_B^* that satisfies (50) is uniquely defined for almost all x, so $y^* = y_B^*$ almost everywhere. Similarly, since π_I is incentive-compatible, we have, for almost all $x \in X$,

$$\int_{y^*(x)}^{\overline{y}} u_I(y,x)f(y)\mathrm{d}y \geq \max_{(a_0,a_1,b)\in\{0,1\}^2\times B} a_0 \int_{\underline{y}}^{b} u_I(y,x)f(y)\mathrm{d}y + a_1 \int_{b}^{\overline{y}} u_I(y,x)f(y)\mathrm{d}y,$$

and thus $a^*(y,x) = \mathbf{1}\{y > y^*(x)\}$ satisfies (51). Next, by strict aggregate downcrossing of u_I in x, for a_B^* that satisfies (51), we have $a^* = a_B^*$ almost everywhere. *Q.E.D.*

B.8. *Proof of Proposition 3 and Remark 3*

By Corollary 2, it suffices to show that if (52) holds, then full disclosure maximizes the principal's expected utility in the equivalent persuasion problem where the utilities are given by (47). Under full disclosure, each state $y \in Y$ is revealed, so the principal chooses decision $x^*(y)$. As decisions $x \notin [x^*(\underline{y}), x^*(\overline{y})]$ can never be chosen, w.l.o.g., assume that $X = [\underline{x}, \overline{x}] = [x^*(\underline{y}), x^*(\overline{y})]$. For each $x \in X$, let $y^*(x) \in Y$ be such that $u(y^*(x),x) = 0$.

CLAIM 10: *Define a function $q : X \to \mathbb{R}$ by*

$$q(x) = \begin{cases} 0, & \text{if } x = \underline{x} \text{ and } u(\underline{y},\underline{x}) < 0 \text{ or } x = \overline{x} \text{ and } u(\overline{y},\overline{x}) > 0, \\ -\frac{v(y^*(x),x)g(x)}{u_x(y^*(x),x)}, & \text{otherwise}. \end{cases} \tag{85}$$

Condition (52) holds iff

$$E(y,x) = V_P(y,x^*(y)) - V_P(y,x) - q(x)u(y,x) \geq 0, \text{ for all } y \in Y \text{ and all } x \in X. \tag{86}$$

PROOF: Suppose that (86) holds. By rearrangement, for all $y_1, y_2 \in Y$ and all $x \in X$ such that $u(y_1,x) < 0 < u(y_2,x)$, we have

$$\frac{V_P(y_2,x^*(y_2)) - V_P(y_2,x)}{u(y_2,x)} \geq q(x) \geq \frac{V_P(y_1,x) - V_P(y_1,x^*(y_1))}{-u(y_1,x)}, \tag{87}$$

yielding (52) by (47). Conversely, suppose that (52) holds. There are four cases to consider depending on whether $u(\underline{y},\underline{x})$ and $u(\overline{y},\overline{x})$ are equal to 0. By symmetry, it suffices to consider the case $u(\underline{y},\underline{x}) < 0 = u(\overline{y},\overline{x})$. Fix any $x \in [\underline{x},\overline{x})$. By (47) and (52), we have

$$\inf_{y_2 > y^*(x)} \frac{V_P(y_2, x^*(y_2)) - V_P(y_2, x)}{u(y_2, x)} \geq \sup_{y_1 < y^*(x)} \frac{V_P(y_1, x) - V_P(y_1, x^*(y_1))}{-u(y_1, x)}. \tag{88}$$

Hence there exists $q(x) \in \mathbb{R}$ bounded above by the left-hand side of (88) and below by the right-hand side of (88), so (87) holds for all $y_1, y_2 \in Y$ such that $u(y_1, x) < 0 < u(y_2, x)$. If $x = \underline{x}$, the right-hand side of (87) is 0, so (87) holds with $q(\underline{x}) = 0$. If $x \in (\underline{x},\overline{x})$, L'Hôpital's rule for $y_2 \downarrow y^*(x)$ and $y_1 \uparrow y^*(x)$ implies that $q(x) = -v(y^*(x), x)g(x)/u_x(y^*(x), x)$. Rearranging (87) yields $E(y,x) \geq 0$ for all $y \neq y^*(x)$, and thus for all $y \in Y$ by continuity in y. Moreover, by continuity in x, we have $E(y,\overline{x}) \geq 0$ for all $y \in Y$. *Q.E.D.*

For each incentive-compatible persuasion mechanism π_P, define

$$J_{\pi_P}(y,x) = \mathbb{P}(\text{state} < y, \text{decision} \leq x) = \int_{\underline{y}}^{y} (1 - \pi_P(x|\tilde{y}))F(\mathrm{d}\tilde{y}), \quad \text{for all } (y,x) \in Y \times X.$$

If (52) holds, then so does (86), by Claim 10. Then, the principal gets a higher expected utility under full disclosure than under π_P, because

$$\begin{aligned}
\int V_P(y,x)(-\pi_P(\mathrm{d}x|y)F(\mathrm{d}y) &= \int V_P(y,x) J_{\pi_P}(\mathrm{d}y, \mathrm{d}x) \\
&= \int (V_P(y,x) + q(x)u(y,x)) J_{\pi_P}(\mathrm{d}y, \mathrm{d}x) \\
&\leq \int V_P(y, x^*(y)) J_{\pi_P}(\mathrm{d}y, \mathrm{d}x) = \int V_P(y, x^*(y)) F(\mathrm{d}y),
\end{aligned}$$

where the first and last inequalities are by the definition of J_{π_P}, the second equality is by incentive compatibility of π_P and the definition of $q(\underline{x})$ and $q(\overline{x})$ in (85), and the inequality is by (86).

We now prove Remark 3. Suppose that $u(y,x) = y - x$ and (53) holds. There are four cases to consider depending on whether $\underline{y}$ and $\overline{y}$ are equal to $\underline{x}$ and $\overline{x}$. By symmetry, it suffices to consider the case $\underline{y} < \underline{x}$ and $\overline{y} = \overline{x}$. By (85), $q(x) = v(x,x)g(x)$ for $x \in (\underline{x},\overline{x}]$ and $q(\underline{x}) = 0$. Denote $\kappa = \min_{y,x \in Y \times X} v_y(y,x)$. Note that $E(x,x) = 0$. By Claim 10, (52) holds if $E(y,x) \geq 0$ for all (y,x). If $x = \underline{x}$, then $E(y,\underline{x}) = 0$ for $y \leq \underline{x}$ and $E(y,\underline{x}) \geq 0$ for $y > \underline{x}$, because

$$\begin{aligned}
E_y(y,\underline{x}) &= v(y,y)g(y) + \int_{\underline{x}}^{y} v_y(y,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x} \geq v(y,y)g(y) + \kappa G(y) - \kappa G(\underline{x}) \\
&\geq v(y,y)g(y) + \kappa G(y) - \kappa G(\underline{x}) - v(\underline{x},\underline{x})g(\underline{x}) \geq 0,
\end{aligned}$$

where the first inequality is by the definition of κ, and the second and third inequalities are by (53). If $x \in (\underline{x},\overline{x}]$ and $y \in (\underline{x},\overline{x}]$, then $E(y,x) \geq 0$ because, for $y \geq (\leq) x$, we have

$$\begin{aligned}
E_y(y,x) &= v(y,y)g(y) + \int_{x}^{y} v_y(y,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x} - v(x,x)g(x) \\
&\geq (\leq) v(y,y)g(y) + \kappa G(y) - \kappa G(x) - v(x,x)g(x) \geq (\leq) 0,
\end{aligned}$$

where the first inequality is by the definition of κ, and the second inequality is by (53). Finally, if $x \in (\underline{x}, \overline{x}]$ and $y \in [\underline{y}, \underline{x}]$, then $E(y,x) \geq 0$ because

$$\begin{aligned} E_y(y,x) &= -\int_{\underline{x}}^{x} v_y(y,\tilde{x})g(\tilde{x})\mathrm{d}\tilde{x} - v(x,x)g(x) \leq \kappa G(\underline{x}) - \kappa G(x) - v(x,x)g(x) \\ &\leq v(\underline{x},\underline{x})g(\underline{x}) + \kappa G(\underline{x}) - \kappa G(x) - v(x,x)g(x) \leq 0, \end{aligned}$$

where the first inequality is by the definition of κ, and the second and third inequalities are by (53). *Q.E.D.*

B.9. *Proof of Proposition 4*

By Corollary 2, it suffices to characterize optimal monotone partitions in the equivalent persuasion problem where

$$U_P(s,t) = c(s)t - \frac{t^2}{2} \quad \text{and} \quad V_P(s,t) = e(c(s))t - \frac{\beta t^2}{2}. \tag{89}$$

A monotone partition $B \in \mathcal{B}$ is represented by a countable set of pooling intervals $(\underline{b}_i, \overline{b}_i]$. The remaining states $\tilde{B} = S \setminus \left(\bigcup_i (\underline{b}_i, \overline{b}_i]\right)$ are revealed.

Let $m_i = \int_{\underline{b}_i}^{\overline{b}_i} c(s)\mathrm{d}s / (\overline{b}_i - \underline{b}_i)$. Note that the left derivative of η_B is

$$\eta_B'(s) = \begin{cases} e(c(s)) - \frac{\beta c(s)}{2}, & \text{if } s \in \tilde{B}, \\ \frac{1}{\overline{b}_i - \underline{b}_i} \int_{\underline{b}_i}^{\overline{b}_i} \left(e(c(\tilde{s})) - \frac{\beta c(\tilde{s})}{2}\right) \mathrm{d}\tilde{s}, & \text{if } s \in (\underline{b}_i, \overline{b}_i]. \end{cases} \tag{90}$$

We have

$$\begin{aligned} W_P(B) &= \int_{\tilde{B}} V_P(s,s)\mathrm{d}s + \sum_i \int_{\underline{b}_i}^{\overline{b}_i} V_P(s,m_i)\,\mathrm{d}s \\ &= \int_{\tilde{B}} \left(e(c(s))c(s) - \frac{\beta (c(s))^2}{2}\right)\mathrm{d}s + \sum_i \int_{\underline{b}_i}^{\overline{b}_i} \left(e(c(s))m_i - \frac{\beta m_i^2}{2}\right)\mathrm{d}s \\ &= \int_{\tilde{B}} c(s)\eta_B(\mathrm{d}s) + \sum_i \int_{\underline{b}_i}^{\overline{b}_i} c(s)\eta_B(\mathrm{d}s) = \int_S c(s)\eta_B(\mathrm{d}s) \\ &= \eta_B(1) - \int_S \eta_B(s)c(\mathrm{d}s) \leq \eta(1) - \int_S \operatorname{conv}\eta(s)c(\mathrm{d}s), \end{aligned} \tag{91}$$

where the first equality is by the definition of W_P, the second equality is by (89), the third equality is by (90) and $\int_{\underline{b}_i}^{\overline{b}_i} m_i\mathrm{d}s = \int_{\underline{b}_i}^{\overline{b}_i} c(s)\mathrm{d}s$, the fourth equality is by $\tilde{B} \cup \left(\bigcup_i (\underline{b}_i, \overline{b}_i]\right) = S$, the fifth equality is by integration by parts and normalizations $c(0) = 0$ and $c(1) = 1$, and the inequality is by the definitions of η_B and $\operatorname{conv}\eta$, and by $\eta_B(1) = \eta(1)$. Finally, let B^* be such that $\operatorname{conv}\eta = \eta_{B^*}$. That is, $\operatorname{conv}\eta$ is linear on $(\underline{b}_i^*, \overline{b}_i^*]$ for each i and $\operatorname{conv}\eta = \eta$ on $\tilde{B}^*$. By (91) and the strict monotonicity of c, we have $W_P(B) \leq W_P(B^*)$ for $B \in \mathcal{B}$, with equality if and only if $\eta_B = \operatorname{conv}\eta$. *Q.E.D.*

B.10. *Proofs of Claim 5 and Corollary 3*

PROOF OF CLAIM 5: Suppose that $U(1,x) \geq U(0,x)$, that is, $1-x \leq p(1)$. Then

$$\frac{\partial V(s,x)}{\partial s} = \beta p(s) - (1-\beta)p'(s)s - \beta(1-x) \geq \beta p(s) - (1-\beta)p'(s)s - \beta p(1) \geq 0,$$

where the first inequality is by $1-x \leq p(1)$, and the second inequality is by $p(s) \geq p(1)$, $p'(s) \leq 0$, and $\beta \in [0,1]$. *Q.E.D.*

PROOF OF COROLLARY 3: *Part (1).* We have

$$\nu'(y) = \left(\beta - \frac{k}{k+1}\right) G(y) - \frac{k}{k+1} g(y) y,$$

$$\nu''(y) = g(y) y \left(\left(\beta - \frac{2k}{k+1}\right)\frac{1}{y} - \left(\frac{k}{k+1}\right)\frac{g'(y)}{g(y)}\right).$$

Given (1), the expression in the parentheses is increasing. Corollary 1 implies that there exists $s^* \in [0,1]$ such that delegation set $\{0\} \cup [s^*,1]$ is optimal.

Part (2). We have

$$\eta'(s) = e(c(s)) - \frac{\beta}{2}c(s) = \frac{\beta}{2}(1-p(s)) + \left(1-\frac{\beta}{2}\right)p'(s)s,$$

$$\eta''(s) = (1-\beta)p'(s) + \left(1-\frac{\beta}{2}\right)p''(s)s = p'(s)\left(1-\beta+\left(1-\frac{\beta}{2}\right)\frac{p''(s)s}{p'(s)}\right).$$

Given (2), the expression in the parentheses is decreasing, and $p'(s) < 0$. Proposition 4 implies that there exists $s^* \in [0,1]$ such that delegation set $\{0\} \cup [s^*,1]$ is optimal. *Q.E.D.*

APPENDIX C: COUNTEREXAMPLES

C.1. *Failure of Equivalence Without Single-Crossing Utilities*

First, we show that Lemma 1 does not hold if $u_D(s,t)$ is not upcrossing in s. Consider a delegation problem with (u_D, v_D) given by

$$u_D(s,t) = \begin{cases} 1, & (s,t) \in [0,\frac{1}{3}] \times [\frac{1}{2},1], \\ -1, & (s,t) \in (\frac{1}{3},1] \times [\frac{1}{2},1], \\ 0, & \text{otherwise,} \end{cases} \quad v_D(s,t) = \begin{cases} 3, & (s,t) \in (\frac{1}{3},1] \times [0,\frac{1}{2}), \\ -1, & \text{otherwise.} \end{cases}$$

Note that $u_D(s,t)$ is not upcrossing in s. Let

$$\pi_D(s,t) = \begin{cases} 1, & (s,t) \in (\frac{1}{3},1] \times [0,\frac{1}{2}), \\ 0, & \text{otherwise.} \end{cases}$$

Mechanism π_D satisfies (IC_D). However, in the discriminatory disclosure problem with $(u_I, v_I) = (u_D, v_D)$, mechanism $\pi_I = \pi_D$ violates (IC_I). Indeed, the agent with a type $t \in [1/2, 1]$ strictly prefers to misreport his type and choose the action opposite to the recommendation. Moreover, the principal's expected utility of 1 attained by incentive-compatible delegation mechanism π_D is not attained by any incentive-compatible disclosure mechanism.

Second, we show that Lemma 2 does not hold if $u_I(s,t)$ is not aggregate downcrossing in t. Consider a discriminatory disclosure problem with (u_I, v_I) given by

$$u_I(s,t) = \begin{cases} -1, & (s,t) \in [0, \frac{1}{3}] \times [\frac{1}{2}, 1], \\ 1, & (s,t) \in (\frac{1}{3}, 1] \times [\frac{1}{2}, 1], \\ 0, & \text{otherwise}, \end{cases} \qquad v_I(s,t) = \begin{cases} 3, & (s,t) \in [0, \frac{1}{3}] \times [0, \frac{1}{2}), \\ -1, & \text{otherwise}. \end{cases}$$

Note that $u_I(s,t)$ is not aggregate downcrossing in t (although it satisfies a weaker notion of aggregate downcrossing defined by Karlin and Rubin, 1956). Let

$$\pi_I(s,t) = \begin{cases} 1, & (s,t) \in (0, \frac{1}{3}] \times [0, \frac{1}{2}), \\ \frac{1}{2}, & (s,t) \in (\frac{1}{3}, 1] \times [0, 1), \\ 0, & \text{otherwise}. \end{cases}$$

Mechanism π_I satisfies (IC_I). However, in the persuasion problem with $(u_P, v_P) = (u_I, v_I)$, mechanism $\pi_P = \pi_I$ violates (IC_P). Indeed, when $s \in (1/3, 1]$, the agent is recommended decision $t = 0$ with probability $1/2$. But, conditional on this recommendation, the agent infers that $s \in (1/3, 1]$, in which case he strictly prefers $t = 1$. Moreover, the principal's expected utility of $1/6$ attained by incentive-compatible disclosure mechanism π_I is not attained by any incentive-compatible persuasion mechanism.

C.2. *Suboptimality of Monotone Mechanisms*

First, we show that the principal's expected utility can be strictly higher under non-monotone disclosure and persuasion mechanisms than under any delegation mechanism. Consider a discriminatory disclosure problem with (u_I, v_I) given by

$$u_I(s,t) = 0, \;\; (s,t) \in [0,1] \times [0,1], \;\; \text{and} \;\; v_I(s,t) = \begin{cases} 1, & (s,t) \in [0, \frac{1}{2}] \times [0, 1], \\ -1, & (s,t) \in (\frac{1}{2}, 1] \times [0, 1]. \end{cases}$$

The principal's maximum utility of $1/2$ is attained by the first-best disclosure mechanism

$$\pi_I(s,t) = \begin{cases} 1, & (s,t) \in (0, \frac{1}{2}] \times [0, 1), \\ 0, & \text{otherwise}. \end{cases}$$

In the persuasion problem with $(u_P, v_P) = (u_I, v_I)$, mechanism $\pi_P = \pi_I$ also maximizes the principal's expected utility. Now consider the delegation problem with $(u_D, v_D) = (u_I, v_I)$. Note that $\pi_D(s|t) = \pi_I(s,t)$ is not a well-defined delegation mechanism, because π_I is not increasing in s. Moreover, since $V_D(s,t) = \max\{-s, s-1\} \leq 0$ for all (s,t), the principal's expected utility in delegation is at most 0.

Second, we show that the principal's expected utility can be strictly higher under non-monotone disclosure and delegation mechanisms than under any persuasion mechanism. Consider a discriminatory disclosure problem with (u_I, v_I) given by

$$u_I(s,t) = 0, \;\; (s,t) \in [0,1] \times [0,1], \;\; \text{and} \;\; v_I(s,t) = \begin{cases} -1, & (s,t) \in [0,1] \times [0, \frac{1}{2}), \\ 1, & (s,t) \in [0,1] \times [\frac{1}{2}, 1]. \end{cases}$$

The principal's maximum utility of $1/2$ is attained by the first-best disclosure mechanism

$$\pi_I(s,t) = \begin{cases} 1, & (s,t) \in (0,1] \times [\frac{1}{2}, 1), \\ 0, & \text{otherwise.} \end{cases}$$

In the delegation problem with $(u_D, v_D) = (u_I, v_I)$, mechanism $\pi_D = \pi_I$ also maximizes the principal's expected utility. Now consider the persuasion problem with $(u_P, v_P) = (u_I, v_I)$. Note that $\pi_P(t|s) = \pi_I(s,t)$ is not a well-defined persuasion mechanism, because π_I is not decreasing in t. Moreover, since $V_P(s,t) = \max\{-t, t-1\} \leq 0$ for all (s,t), the principal's expected utility in persuasion is at most 0.