The Interval Structure of Optimal Disclosure

Yingni Guo[*] Eran Shmaya[†]

May 14, 2017

Abstract

A sender persuades a receiver to accept a project by disclosing information regarding a payoff-relevant state. The receiver has private information about the state, referred to as his type. We show that the sender-optimal mechanism takes the form of nested intervals: each type accepts on an interval of states and a more optimistic type's interval contains a less optimistic type's interval. This nested-interval structure offers a simple algorithm to solve for the optimal disclosure and connects our problem to monopoly screening problems. The mechanism is optimal even if the sender conditions the disclosure mechanism on the receiver's reported type.

JEL: D81, D82, D83

Keywords: Bayesian persuasion, privately informed receiver, nested intervals

[*]Department of Economics, Northwestern University, yingni.guo@northwestern.edu.

[†]MEDS, Kellogg, Northwestern University, e-shmaya@kellogg.northwestern.edu.

1 Introduction

A theorist recommends a candidate to her department chair, who decides whether or not to hire that candidate. The chair's payoff from hiring the candidate depends on the quality of the candidate. The theorist knows the true quality and can design a disclosure mechanism revealing more information about the candidate's quality to the chair. This is the Bayesian persuasion problem analyzed by Rayo and Segal (2010) and Kamenica and Gentzkow (2011).

Frequently, besides the information disclosed by the theorist, the chair has access to various external sources of information regarding the underlying quality. He can read the job market paper by himself or call and consult his friends from other institutions. What he learns from these external sources will remain the chair's private information. A chair who has received positive feedback is more optimistic about the quality and more willing to hire than a chair who has received some worrisome feedback. How will the theorist structure her recommendation to best persuade a chair with an unknown level of optimism? Under what conditions will she actively campaign for the candidate to be hired, and under what conditions will she step back and let the chair resort to his own private information and decide by himself?

This phenomenon of trying to persuade a privately informed receiver is pervasive in economic and political life. For instance, a lobbyist tries to sway a legislator's position on an issue. The legislator himself has access to various information sources. He may also have gained knowledge from working on similar issues in the past. In a different example, a media outlet tries to promote a candidate or an agenda. The audience acquires information about the candidate/agenda from various channels other than this media outlet, or they may also have first-hand experience which has a significant influence on their position, unbeknownst to the media outlet. In this paper we study how to design a disclosure mechanism when the receiver has diverse opinions about the issue/candidate/agenda.

Environment. Following Rayo and Segal (2010) and Kamenica and Gentzkow (2011), we consider an environment with two players: Sender (she) and Receiver (he). Sender promotes a project to Receiver, who decides whether to accept or reject it. Receiver's payoff from accepting increases with the project's quality, i.e., the *state*. His payoff is normalized to zero if he rejects. Therefore, there is a threshold state such that Receiver benefits from the project if the state is above the threshold, and loses otherwise. Sender, on the other hand, simply wants the project to be accepted.

Receiver does not know the state but has access to an external information source. The information from this source is Receiver's *type*. We assume that the higher the state, the more likely it is that Receiver has a higher rather than a lower type. Therefore, a higher type is more optimistic than a lower type that the state favors an acceptance decision. Sender designs a disclosure mechanism to reveal more information about the state and can commit to this mechanism. Receiver updates his belief based on his private information and Sender's signal. Then, he takes an action: to accept or to reject. Following Kolotilin et al. (2016), we call such an environment "public persuasion": different types observe the same signal. We later address the environment of "private persuasion," in which Sender asks Receiver to report his type and then randomizes a recommendation from a distribution that depends on the state and the type reported.

Main results. To illustrate the structure and the intuition of the optimal mechanism, we consider first the case in which Receiver has two possible types: high or low. Sender could design a pooling mechanism under which both types always receive the same recommendation. Sender could also design a separating mechanism under which, with positive probability, only the high type (who is more optimistic) accepts. When Receiver's external information is sufficiently informative, it is much easier to persuade the high type than the low type, so a separating mechanism delivers a higher payoff.

Under a separating mechanism, the high type will accept whenever the low type accepts. In addition, the high type will accept in some other states, of which some are good for Receiver and some are bad. Our main result states that the optimal separating mechanism takes a nested-interval structure. This entails two key properties. First, the set of states in which each type accepts is an interval. Second, the high type's interval contains the low type's; as a result, Sender recommends that only the high type accept when the state is, from Receiver's perspective, either quite good or quite bad.

The interval structure translates to an intuitive rule of optimal disclosure. Sender pushes for the project to be accepted in the intermediate states where the two types' beliefs do not differ much. When the state is quite good or quite bad, the two types' beliefs differ significantly. Sender pools these states, and lets Receiver resort to his own private information and decide by himself. One important implication is that, when a lobbyist faces a legislator or when the media addresses an audience with diverse viewpoints, both the lobbyist and the media will "spend their capital" on intermediate states and push for their preferred action. They pool relatively extreme states under which the legislator or the audience make use of their own information.

More generally, in an environment with more than two types we show that the optimal disclosure mechanism has the following structure: (i) each type is endowed with an interval of states under which this type accepts; (ii) a lower type's acceptance interval is a subset of a higher type's acceptance interval. This result relies on two features of our environment: a common prior belief between Sender and Receiver and the fact that higher types are more optimistic about the state than lower types.

The interval structure gives us a simple algorithm to find the optimal disclosure: we only need to characterize the upper and lower bound of each type's interval. Moreover, the interval structure offers a natural connection between (i) Sender's persuasion problem and (ii) the monopoly screening problem analyzed in Mussa and Rosen (1978). We can interpret Receiver's utility from accepting on an interval of positive-payoff states as a buyer's utility from some quality level. Receiver's disutility from accepting on an interval of negative-payoff states is the buyer's disutility from some amount of payment. Receiver's level of optimism (a.k.a., their type) corresponds to how much the buyer values quality. Once we impose the interval structure, Sender's problem in our setup has the same structure as the monopoly screening problem. In addition to the conceptual advantage of unifying these two seemingly different problems, this reduction enables us to use the techniques from the screening problems to solve for the optimal mechanism.

Our second main result states that this optimal mechanism under public persuasion is also optimal under private persuasion. Under private persuasion, Sender draws the recommendation conditional on the state and the reported type. We show that Sender does not benefit from eliciting Receiver's private information. The key observation is that, while the incentive-compatible constraints pertaining to private persuasion are in general weaker than those of public persuasion, for interval-structure mechanisms they are essentially the same. For this reason, unlike the environment in Kolotilin et al. (2016), in our environment the equivalence between private and public persuasion holds only for Sender's optimal mechanism, with its interval structure.

Intuition for the proof. The proof is sufficiently intuitive that we can write the core argument here. We first argue that Sender's payoff is maximized when each type accepts in states that are close to the threshold state. We divide the states into "positive states" and "negative states" depending on the sign of Receiver's payoff from accepting. The set of states on which Sender recommends that each type accept includes some positive and some negative states. For any given utility to Receiver from positive states (and similarly for any given disutility from negative states), Sender's payoff is maximized when these states are the ones which are closer to the threshold, because Sender's payoff is less sensitive to the states than Receiver's.

We next argue that the additional states on which only the high type accepts should be the extreme states of his acceptance interval, because this is the cheapest way to maintain the incentives for different types. For states above the threshold state, the higher the state, the more likely it is that Receiver's type is high. The high type puts more weight on a higher state than on a lower state than the low type does. Therefore, when Sender recommends that only the high type rather than both types accept, the benefit for the high type's incentive constraint net of the cost on the low type's incentive constraint grows as the state increases. Hence, for states above the threshold state, Sender recommends that only the high type accept when the state is sufficiently high. For states below the threshold, the lower the state, the more likely it is that Receiver's type is low. The low type puts more weight on a lower state over a higher state than the high type does. Therefore, when Sender excludes the low type and recommends that only the high type accept, the benefit for the low type's incentive constraint net of the cost on the high type's incentive constraint grows as the state decreases. Hence, for states below the threshold state, Sender recommends that only the high type accept when the state is sufficiently low.

In order to complete the proof, it is essential that the two properties we need—(i) that each type's acceptance set is an interval so that the high type's interval contains the low type's interval and (ii) that the states on which only the high type accepts are the extreme states in his interval—can be achieved simultaneously. One key observation in proving this is that these two desirable properties are compatible with each other. In fact, we invite the readers to convince themselves that if an interval strictly contains another interval, then the set of states in the former that are not included in the latter are precisely the extreme states on both sides.

Related literature. Our paper is related to the literature on persuasion. Rayo and Segal (2010) and Kamenica and Gentzkow (2011) study optimal persuasion between a sender and a receiver.[1] We study the information design problem in which the receiver has an external source of information. Kamenica and Gentzkow (2011) extend their baseline analysis to situations in which the receiver has private information. They show that the geometric method is readily generalized to the public persuasion setting. In general, it is difficult

[1]Rayo and Segal (2010) assume that the receiver has private information about his threshold or taste. They examine optimal public persuasion.

to characterize the optimal mechanism using the geometric method. For this reason, our framework follows Kamenica and Gentzkow (2011) overall, but we assume an additional structure over the state space and the receiver's private information; this enables us to derive the interval structure of the optimal mechanism, and to prove that focusing on public persuasion is without loss at the optimum.

Kolotilin (2016) also examines optimal public persuasion when the receiver has private information. He provides a linear programming approach and establishes conditions under which full or no revelation is optimal. Alonso and Câmara (2016*a*) examine optimal persuasion when the sender's prior differs from the receiver's. In their benchmark model, both priors are common knowledge so there is no private information. Hence, the analysis and main results are quite different from ours. They extend their analysis to the case where the sender is uncertain about the receiver's prior and characterize the conditions under which the sender benefits from persuasion.

Our result about the optimality of some public mechanism among all private mechanisms is also related to Kolotilin et al. (2016). In their setup, the receiver privately learns about his threshold for accepting. The receiver's threshold is independent of the state, and the receiver's utility is additive in the state and his threshold. They show that, given the independence and the additive payoff structure, any payoffs that are implementable by a private mechanism are implementable by a public one. Our paper differs in that the receiver's type is informative about the state. The strong equivalence result in Kolotilin et al. (2016) does not hold in our environment, but we show that when Receiver's type satisfies the monotone-likelihood-ratio property, the optimal private mechanism permits a public implementation.

Our interval structure is different from the interval mechanism in both Kolotilin et al. (2016) and Kolotilin (2016). In those two papers, an interval mechanism is characterized by two bounds: states below the lower bound are pooled into one signal, states above the upper bound are pooled into another, and states between the two bounds are revealed perfectly.

Our discussion of private persuasion is also related to Bergemann, Bonatti and Smolin (2015), who consider a monopolist who sells informative experiments to a buyer who has private information about the state. The monopolist designs a menu of experiments and a tariff function to maximize her profit. Our paper differs since (i) we do not allow for transfers and (ii) Sender attempts to sway Receiver's action.

Our model admits both the interpretation of a single receiver and that of a continuum of receivers. For this reason, our paper is also related to information design with multiple receivers. Lehrer, Rosenberg and Shmaya (2010), Lehrer, Rosenberg and Shmaya

(2013), Bergemann and Morris (2016*a*), Bergemann and Morris (2016*b*), Mathevet, Perego and Taneva (2016), and Taneva (2016) examine the design problem in a general environment. Schnakenberg (2015), Alonso and Câmara (2016*b*), Chan et al. (2016), and Guo and Bardhi (2016) focus on the voting context. Our paper is mostly closely related to Arieli and Babichenko (2016), who study optimal persuasion when the receivers have different thresholds for acceptance. They assume that receivers' thresholds are common knowledge. Since there is no private information, their definition of private persuasion does not entail reporting or self-selection by the receivers.

Our paper is also related to cheap talk communication (Crawford and Sobel (1982)) with a privately informed receiver (e.g., Seidmann (1990), Watson (1996), Olszewski (2004), Chen (2009), Lai (2014)). Chen (2009) is the closest to our setting; she assumes that the receiver has a signal about the state, and shows that nonmonotone equilibria may arise. However, our model differs in that we abstract away from incentive issues in disclosure and instead focus on the optimal disclosure mechanism.

Structure of the paper. In Section 2 we formally state our results for the theorist-chair example. Section 3 includes the general framework and our main results. Section 4 contains a detailed discussion of the binary-type case (the finite-type case is no more complicated than the binary-type case). This section also shows the computational advantage of our approach in comparison with the Cav-V approach of Aumann and Maschler (1995) and Kamenica and Gentzkow (2011). In Section 5 we explain how we use our theorem to reduce Sender's problem to an optimal screening problem, and use the techniques from the screening problems to solve for the optimal mechanism in a continuous-type example. The solutions to Sender's problem in Sections 4 and 5 allow us to identify the situations in which it is optimal to pool types, and those in which it is optimal to separate them. In Section 6, we extend our results to broader payoff and information settings. Section 7 contains the proofs.

2 An example

In this section, we illustrate our question and main results by using the theorist-chair example from the introduction. We assume that the state s is uniform on $[0,1]$. The chair's payoff from hiring is $s-3/4$. The theorist's payoff is 1 if the candidate is hired. If not, both players get a zero payoff.

The theorist knows the state s. She designs a mechanism for disclosing more information about the state and can commit to this mechanism. If the chair has no private information, the theorist will reveal whether s is above or below $1/2$. She extracts the entire surplus. The chair's expected payoff from hiring is zero.

The chair has some private information about the state, captured by his type: given s, the chair's type is H with probability s and L with probability $1 - s$. Naturally, the higher the state, the more likely that the chair's type is H. Thus, type H is more optimistic about s and easier to persuade than type L is. How does the theorist disclose information when different types have different levels of optimism?

We show that the optimal disclosure takes the form shown in Figure 1. The x-axis is the state. When the state is in the red region (concentrated around $3/4$), the theorist claims that the candidate is a solid one and recommends that the chair hire regardless of his type. When the state is in the green region, the theorist claims that the candidate is creative but also risky. She recommends that the chair hire only if his type is high. For the remaining states, the theorist recommends not hiring.

The chair is always willing to obey the theorist's recommendation. Conditional on s being in the red region, type L is indifferent between hiring and not. Conditional on s being in the green region, type H is indifferent. The theorist does not extract the entire surplus, since type H gets a positive payoff from hiring when the state is in the red region.

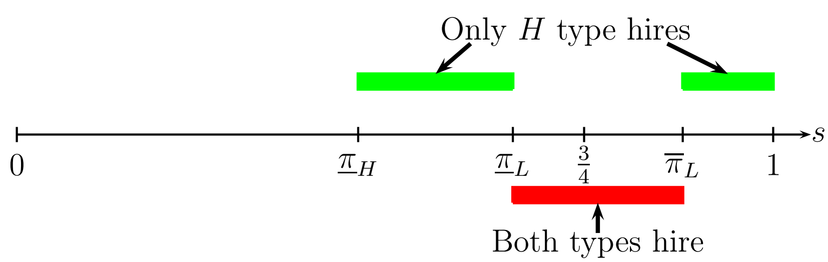

Figure 1: Optimal disclosure for binary types

A central concept in our solution is the notion of nested intervals. Both types hire on an interval of states: type L hires in the red region and type H hires in both the red and green regions. Type L's hiring interval is a subset of type H's. Compared to type H's interval, type L does not hire when the state is either quite good or bad.

Thus, facing a privately informed chair, the theorist pushes for the candidate to be hired only when the state is intermediate (i.e., in the red region) by recommending hiring regardless of the chair's type. When the state is quite good or bad (i.e., in the green region), the theorist lets the chair resort to his private information and decide by himself.

To implement such an interval structure, the theorist announces at each state the smallest type who is still willing to hire. We call this type "the cutoff type" at each state, since all types including and above this cutoff type will hire. (For instance, the cutoff type is L in the red region and H in the green region.) The interval structure translates to a U-shaped cutoff mechanism. These two notions—nested intervals and a U-shaped cutoff mechanism—are equivalent to each other. We show that for a large class of games, the optimal disclosure mechanism is given by a U-shaped cutoff mechanism.

3 Environment and main results

Let $\mathcal{S}$ be a set of *states*, $\mathcal{T}$ a set of *types* equipped with σ-finite measures μ, λ and consider a distribution over $\mathcal{S} \times \mathcal{T}$ with density f w.r.t. $\mu \times \lambda$.[2] In all our examples these spaces will be either discrete spaces or intervals in the real line equipped with Lebesgue's measure. Let $u\colon \mathcal{S} \to \mathbf{R}$ be a bounded Borel function representing *Receiver's payoff* from accepting. Receiver's payoff from rejecting is zero at every state. Note that the state in our framework represents not just Receiver's payoff but also Sender's belief about Receiver's type. It would perhaps be better to use the term "Sender's type" instead of state, but we keep the term *state* to conform to the prior literature.

In this section we first define the concept of disclosure mechanisms, and then define the incentive-compatible restrictions that correspond to two environments: (i) when the mechanism discloses the same information to all types, and (ii) when the mechanism can elicit Receiver's private information and disclose different information to different types.

3.1 Disclosure mechanisms

A *disclosure mechanism* is given by a triple $(\mathcal{X}, \kappa, r)$ where $\mathcal{X}$ is a set of *signals*, κ is a Markov kernel[3] from $\mathcal{S}$ to $\mathcal{X}$, and $r : \mathcal{X} \times \mathcal{T} \to \{0, 1\}$ is a *recommendation function*: when the state is s, the mechanism randomizes a signal x according to $\kappa(s)$ and recommends that type t accepts if and only if $r(x, t) = 1$. We say that a mechanism $(\mathcal{X}, \kappa, r)$ is *deterministic* if $\kappa(s)$ is Dirac's measure on $z(s)$ for some function $z : \mathcal{S} \to \mathcal{X}$. This means that the mechanism performs no randomization, so that the signal is a function of the state.

[2] All sets and all functions in the paper are, by assumption or by construction, Borel.

[3] Specifically, $\kappa : \mathcal{S} \times \mathcal{B}(\mathcal{X}) \to [0, 1]$ such that $\kappa(s, \cdot)$ is a probability measure over $\mathcal{X}$ for every $s \in S$, where $\mathcal{B}(\mathcal{X})$ is the sigma-algebra of Borel subsets of $\mathcal{X}$. We sometimes write $\kappa(s)$ for $\kappa(s, \cdot)$.

For a mechanism $(\mathcal{X}, \kappa, r)$, if type t follows the recommendation, then *type t's acceptance probability at state s* is given by

$$\rho(s,t) = \int r(x,t)\ \kappa(s, \mathrm{d}x). \tag{1}$$

For every function $\rho : \mathcal{S} \times \mathcal{T} \to [0,1]$, there exists a mechanism whose acceptance probabilities are given by ρ. Indeed, we can choose $\mathcal{X} = [0,1]$ with $\kappa(s) = \text{Uniform}(0,1)$ for every s and with $r(x,t) = 1$ if and only if $x < \rho(s,t)$. When talking about a mechanism's properties that depend only on the acceptance probabilities, we sometimes abuse terminology by referring to ρ as the mechanism. One such property, which we define in Section 3.4, is private incentive-compatibility.

3.2 Incentive compatibility

Incentive compatibility means that Receiver will follow the mechanism's recommendation. We now proceed to formally define incentive compatibility, under the assumption that Sender makes the signal x public and each type t chooses an action after observing x.

A *strategy* for type t is given by $\sigma : \mathcal{X} \to \{0,1\}$. Let $\sigma_t^* = r(\cdot, t)$ be the strategy that follows the mechanism's recommendation for type t. We say that the mechanism is *publicly incentive-compatible (or publicly IC)* if

$$\sigma_t^* \in \arg\max \int f(s,t) u(s) \left(\int \sigma(x)\ \kappa(s, \mathrm{d}x) \right) \mu(\mathrm{d}s) \tag{2}$$

for every type t, where the argmax ranges over all strategies σ. The expression inside the argmax is, up to normalization, the expected payoff of type t who follows σ.

For a publicly IC mechanism, the recommendation function r is almost determined by the signaling structure $(\mathcal{X}, \kappa)$: type t is recommended to accept if his expected payoff conditional on the signal is positive, and to reject if it is negative. We say "almost determined" because of the possible indifference and because the conditional expected payoff is defined up to an event with zero probability. We omit the formal statement of this assertion, in order not to get into the technical intricacies involving conditional distributions.

3.3 Sender's optimal mechanism

The *Sender's problem* is:

$$\text{Maximize} \iint f(s,t)\left(\int r(x,t)\ \kappa(s,\mathrm{d}x)\right)\ \mu(\mathrm{d}s)\lambda(\mathrm{d}t) \tag{3}$$

among all publicly IC mechanisms. Note that we assume a common prior between Sender and Receiver, which is reflected by the fact that the same density function f appears in (2) and (3). The problem would be well-defined for the case in which the density functions of Sender and Receiver are different, but we need the common prior assumption for our results. (See the counterexample in Section 6.3.)

From now on, we assume that $\mathcal{S}, \mathcal{T} \subset \mathbf{R}$, that μ and λ have full support, and that u is monotone-increasing.

If Receiver has no private information (that is, if $\mathcal{T}$ is a singleton), then—up to some possible nuisance if the state space has atoms—the optimal mechanism is deterministic (that is, it does not randomize) and the set of states on which the mechanism recommends accepting is an interval of the form $[\underline{\pi}, \infty)$. The threshold $\underline{\pi}$ is chosen such that Receiver's expected payoff from accepting is zero. In this section we provide a generalization of this observation for the case of private information. As we shall see, each type t still accepts on an interval of states and the interval expands as t increases. However, the most pessimistic types' intervals are typically bounded from above. While the lowest type would get a zero payoff, higher types receive some information rent.

A mechanism *recommends that type t accept on an interval* if there exists some $\underline{\pi} \leqslant \overline{\pi} \in \mathcal{S}$ such that $\rho(s,t) = 1$ whenever $s \in (\underline{\pi}, \overline{\pi})$, and $\rho(s,t) = 0$ whenever $s \notin [\underline{\pi}, \overline{\pi}]$. If the mechanism is deterministic, then this condition means that the set of states on which it recommends that type t accept is an interval. If the mechanism is not deterministic, then the mechanism may randomize at the endpoints of the interval.

We need the following additional assumption:

Assumption 1. *The set of types $\mathcal{T}$ is closed and bounded from below. The density function $f(s,t)$ is continuous in t and satisfies the* increasing monotone likelihood ratio (i.m.l. ratio)*, i.e., $f(s,t)/f(s,t')$ is (weakly) increasing in s for every $t' < t$.*

We say that the mechanism is a *cutoff mechanism* if $\mathcal{X} = \mathcal{T} \cup \{\infty\}$ and the recommendation function is given by $r(x,t) = 1 \longleftrightarrow t \geqslant x$. That is, a cutoff mechanism announces a type x and recommends that all higher types accept. A deterministic cutoff mechanism is such that $\kappa(s)$ is Dirac's measure on $z(s)$ for some function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$.

Our structural theorem states that Sender's optimal mechanism is a cutoff mechanism that recommends that each type accept on an interval. For a deterministic cutoff mechanism, this means that the cutoff function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$ is U-shaped. (The right-hand side of Figure 3 depicts a U-shaped cutoff function.) Recall that z is *U-shaped* if there exists some s_0 such that z is monotone-decreasing for $s \leqslant s_0$, and monotone-increasing for $s \geqslant s_0$. Equivalently, z is U-shaped if the set $\{s : z(s) \leqslant t\}$ is an interval for every t. For the case of a nondeterministic cutoff mechanism, we say that a Markov kernel κ from $\mathcal{S}$ to $\mathcal{T} \cup \{\infty\}$ is U-shaped if there exists some $s_0 \in \mathcal{S}$ such that for every $s' < s \leqslant s_0$ and for every $s_0 \leqslant s < s'$, every $t \in \text{support}(\kappa(s))$ and $t' \in \text{support}(\kappa(s'))$ satisfy $t \leqslant t'$.

Theorem 3.1. *The optimal publicly IC mechanism is a U-shaped cutoff mechanism. If μ is nonatomic, then this optimal mechanism is also deterministic and the cutoff function is increasing on the set $\{s : u(s) \geqslant 0\}$ and decreasing on the set $\{s : u(s) \leqslant 0\}$.*

It is a standard argument that under Assumption 1 every publicly IC mechanism is essentially a cutoff mechanism. (We do not formalize and prove the last assertion because we do not need it. We say "essentially" because of the possible indifference and the intricacy on zero-probability events.) Theorem 3.1 states that the optimal mechanism has the additional property that it recommends that each type accept on an interval (or on a "fractional interval," with some probability of acceptance at the endpoints of the interval in the case of an atomic state space).

The i.m.l. ratio part of Assumption 1 is essential for Theorem 3.1. (See the counterexample in Section 6.2.) The continuity part is not. Without it, we would have to modify the definition of a cutoff mechanism: in addition to announcing the cutoff type, the mechanism would have to announce whether the cutoff type is supposed to accept or reject.

3.4 Private incentive-compatibility

We now consider a different environment in which Sender discloses different information to different types. Thus, Receiver does not observe the realized signal x, but can only report some type t' and observe $r(x, t')$. This restricts the set of possible deviation strategies, and gives rise to a weaker notion of incentive compatibility which we call private incentive-compatibility.

We say that the mechanism is *privately incentive-compatible (or privately IC)* if (2) holds for every type t where the argmax ranges over all strategies σ of the form $\sigma(x) = \bar{\sigma}(r(x, t'))$

for some type $t' \in \mathcal{T}$ and some $\bar{\sigma} : \{0,1\} \to \{0,1\}$. We note that, in contrast with public incentive-compatibility, private incentive-compatibility depends only on the acceptance probabilities ρ given by (1).

It is easy to see that every mechanism that is publicly IC is also privately IC. (The example in Section 6.4 shows that not every privately IC mechanism can be duplicated by a publicly IC one.) Hence, the optimal privately IC mechanism gives Sender a weakly higher payoff than the optimal publicly IC one does. Our next result shows that Sender does not do better under privately IC mechanisms.

Theorem 3.2. *No privately IC mechanism gives a higher payoff to Sender than the optimal publicly IC mechanism.*

Our definitions of mechanisms and incentive compatibility are different from those of other papers (such as Kolotilin et al. (2016)) which define the notion of mechanisms differently in the public environment than in the private one. For our purposes, we use the same definition of mechanisms for both environments and make the distinction at the level of incentive compatibility. This approach renders the logical implication between private and public incentive-compatibility straightforward, and it simplifies the proofs of our theorems.

3.5 Downward incentive-compatibility

In this section, we introduce the notion of downward incentive-compatibility, that is, type t prefers to follow the mechanism's recommendation for his type over following its recommendation for a lower type t'. We prove Theorem 3.1 and Theorem 3.2 with the help of this notion and show that the only binding IC constraints are the downward ones. This simplifies Sender's problem further since other IC constraints can be ignored. (Moreover, for the case of discrete type spaces, the adjacent downward constraints are sufficient.) This also connects our problem to the standard screening problem, where it is typical to have only downward constraints bind.

We say that the mechanism $(\mathcal{X}, \kappa, r)$ is *downward incentive-compatible (or downward IC)* if (i) for every type t, (2) holds where the argmax ranges over all strategies σ of the form $\sigma(x) = r(x, t')$ for some type $t' \leq t$; and (ii) the lowest type $\underline{t}$ prefers to follow the recommendation for his type over always rejecting. It is easy to see that every privately (and, in particular, publicly) IC mechanism is downward IC.

Theorem 3.1 and Theorem 3.2 are immediate consequences of the following theorem:

Theorem 3.3. *The optimal downward IC mechanism is a U-shaped cutoff mechanism. This optimal mechanism is publicly IC. If μ is nonatomic, then this optimal mechanism is also deterministic and the cutoff function is increasing on the set $\{s : u(s) \geqslant 0\}$ and decreasing on the set $\{s : u(s) \leqslant 0\}$.*

It follows from Theorem 3.3 that when we search for the optimal publicly IC mechanism, we can optimize over the set of U-shaped cutoff mechanisms that are downward IC. This set is especially easy to work with in the case of finite type spaces, because in this case the adjacent downward-IC constraints are sufficient. We now proceed to formalize this fact.

Let

$$U(t', t) = \int f(s, t)u(s) \int r(x, t')\ \kappa(s, \mathrm{d}x)\, \mu(\mathrm{d}s)$$

be the utility for type t from mimicking t'. Then the mechanism is downward IC iff

$$U(t, t) \geqslant U(t', t),\ \text{for every types } t' \leqslant t \in \mathcal{T},\ \text{and} \tag{4}$$

$$U(\underline{t}, \underline{t}) \geqslant 0. \tag{5}$$

The following lemma, which follows from the definition of i.m.l. ratio, is familiar from the standard screening problem: if $t'' \leqslant t' \leqslant t$ are types such that type t' does not want to mimic type t'', then type t prefers to mimic t' over mimicking t''.

Lemma 3.4. *Every cutoff mechanism has the following property:*

$$U(t', t') \geqslant U(t'', t') \to U(t', t) \geqslant U(t'', t),\ \text{for every } t'' \leqslant t' \leqslant t.$$

Corollary 3.5. *If $\mathcal{T} = \{\underline{t} = t_0 < t_1 < \cdots < t_n\}$ is finite, then in condition (4) in the definition of downward incentive-compatibility, it is sufficient to consider the case where $t = t_k$ and $t' = t_{k-1}$ for every $1 \leqslant k \leqslant n$.*

In Sections 4 and 5, we show how to use our theorems to find the optimal publicly IC mechanism by analyzing the binary-type case and the continuous-type case, respectively.

4 The binary-type case

Consider the case in which $\mathcal{T} = \{H, L\}$ (high and low types) and $\mathcal{S} = [0, 1]$. Receiver's payoff u is strictly monotone-increasing, and $u(\zeta) = 0$ for some $\zeta \in [0, 1]$. Assumption 1 means that $\frac{f(s,H)}{f(s,L)}$ is monotone-increasing in s. Sender's problem is trivial if type L accepts without further information, so we assume otherwise.

Our Theorem 3.1 shows that the optimal mechanism takes the form of nested intervals. We choose the endpoints of the intervals $0 \leqslant \underline{\pi}_H \leqslant \underline{\pi}_L \leqslant \zeta \leqslant \overline{\pi}_L \leqslant \overline{\pi}_H \leqslant 1$ such that type L accepts when $s \in [\underline{\pi}_L, \overline{\pi}_L]$ and type H accepts when $s \in [\underline{\pi}_H, \overline{\pi}_H]$. We maximize Sender's payoff subject to downward incentive constraints. By Corollary 3.5 we have one incentive constraint for each type:

$$\begin{aligned}
\underset{\underline{\pi}_H, \underline{\pi}_L, \overline{\pi}_L, \overline{\pi}_H}{\text{Maximize}} \quad & \int_{\underline{\pi}_L}^{\overline{\pi}_L} f(s,L)\,\mathrm{d}s + \int_{\underline{\pi}_H}^{\overline{\pi}_H} f(s,H)\,\mathrm{d}s \\
\text{subject to} \quad & 0 \leqslant \underline{\pi}_H \leqslant \underline{\pi}_L \leqslant \zeta \leqslant \overline{\pi}_L \leqslant \overline{\pi}_H \leqslant 1, \\
& \int_{\underline{\pi}_L}^{\overline{\pi}_L} f(s,L)u(s)\,\mathrm{d}s \geqslant 0, \\
& \int_{\underline{\pi}_H}^{\underline{\pi}_L} f(s,H)u(s)\,\mathrm{d}s + \int_{\overline{\pi}_L}^{\overline{\pi}_H} f(s,H)u(s)\,\mathrm{d}s \geqslant 0.
\end{aligned} \tag{6}$$

Hence, our result reduces Sender's problem to a finite-dimensional constrained optimization. In the example in Section 4.1, the problem can be solved explicitly. Here, we first provide a qualitative description of the solution. It is clear that the optimal solution satisfies $\overline{\pi}_H = 1$; otherwise increasing $\overline{\pi}_H$ would increase Sender's payoff without violating the constraints. Similarly, $\overline{\pi}_L < \overline{\pi}_H \longleftrightarrow \underline{\pi}_H < \underline{\pi}_L$. Thus, there are two possibilities for the optimal solution: If $\overline{\pi}_H = \overline{\pi}_L = 1$, Sender pools the two types. She recommends that both types accept if $s \in [\underline{\pi}_L, 1]$ and that both reject otherwise. If $\underline{\pi}_H < \underline{\pi}_L \leqslant \overline{\pi}_L < \overline{\pi}_H = 1$, Sender offers a separating mechanism.

Proposition 4.1. *Pooling is optimal if and only if*

$$\frac{f(1,H)}{f(1,L)} - \frac{f(\underline{\pi}_L^*, H)}{f(\underline{\pi}_L^*, L)} < 1 - \frac{u(\underline{\pi}_L^*)}{u(1)}, \tag{7}$$

where $\underline{\pi}_L^*$ *is such that* $\int_{\underline{\pi}_L^*}^1 f(s,L)u(s)\,\mathrm{d}s = 0$. *In this case the mechanism recommends that both types accept on* $[\pi_L^*, 1]$.

The condition (7) states that Sender does not benefit from marginally shrinking type L's acceptance interval in order to expand type H's acceptance interval. Starting from the pooling interval $[\underline{\pi}_L^*, 1]$, Sender can replace type L's interval by $\left[\underline{\pi}_L^* + \frac{f(1,L)}{f(\underline{\pi}_L^*,L)} \frac{u(1)}{-u(\underline{\pi}_L^*)}\varepsilon, 1-\varepsilon\right]$ for small $\varepsilon > 0$, without violating type L's incentive constraint. Due to the i.m.l. ratio assumption, this change allows Sender to lower $\underline{\pi}_H$, so that type H accepts more often. Pooling is optimal only if Sender does not benefit from such an operation. We use (6) to show that this condition is also sufficient.

4.1 Example

We further assume that s is drawn uniformly from $[0, 1]$ and that $u(s) = s - \zeta$ where ζ is a parameter. Receiver's private information about the state is given by a coin toss with probability $1/2 + \phi(s - 1/2)$ for success. Receiver is of type H or type L when the outcome of the coin toss is success or failure, respectively. The parameter $\phi \in [0, 1]$ measures how informative Receiver's private information is. When $\phi = 0$, Receiver has no private information. As ϕ increases, type H is increasingly more optimistic than type L is. With this formulation, the joint density is given by

$$f(s, H) = 1/2 + \phi(s - 1/2), \text{ and } f(s, L) = 1/2 - \phi(s - 1/2) \text{ for } 0 \leqslant s \leqslant 1.$$

If $\zeta \leqslant \frac{3-\phi}{6}$, then type L prefers to accept absent further information. In this case the optimal mechanism always recommends that both types accept. We thus assume that $\zeta > \frac{3-\phi}{6}$.

From Proposition 4.1 we get Corollary 4.1, which states that separating is strictly optimal when ϕ is sufficiently large. Intuitively, the more type H's belief differs from type L's, the more likely it is that Sender obtains a higher payoff from separating by sometimes persuading only type H to accept.

Corollary 4.1. *There exists an increasing function $\Phi(\cdot)$ such that the optimal mechanism is separating if $\phi > \Phi(\zeta)$, and pooling if $\phi < \Phi(\zeta)$.*

Proof. Substituting $f(s, H), f(s, L), u(s)$ into the condition in Proposition 4.1, we find that this condition holds if and only if

$$\zeta \geqslant \frac{3\phi^3 + 13\phi^2 - (\phi - 1)^2\sqrt{9\phi^2 - 6\phi + 33} + 21\phi - 5}{8\phi(3\phi + 1)},$$

and the right-hand side is monotone-increasing in ϕ, as desired. □

4.2 Example revisited: a comparison with the cav-V approach

We now compare our approach with the general concavification (cav-V) approach, following Aumann and Maschler (1995) and Kamenica and Gentzkow (2011, Section VI A), by revisiting the example from Section 4.1. We first apply the general cav-V approach to this setting. For convenience we assume $\phi = 1$, and thus, $f(s,H) = s$ and $f(s,L) = 1 - s$. If Sender's signal induces distribution γ over states, then Receiver is of type H or L with probabilities $\int s\,\gamma(\mathrm{d}s)$ and $1 - \int s\,\gamma(\mathrm{d}s)$, respectively. Type H accepts if $\int s(s-\zeta)\,\gamma(\mathrm{d}s) \geqslant 0$. Type L accepts if $\int (1-s)(s-\zeta)\,\gamma(\mathrm{d}s) \geqslant 0$. Therefore, Sender's payoff from the belief γ induced by her signal is given by

$$V(\gamma) = \begin{cases} 1, & \text{if } \int (1-s)(s-\zeta)\,\gamma(\mathrm{d}s) \geqslant 0, \\ 0, & \text{if } \int s(s-\zeta)\,\gamma(\mathrm{d}s) < 0, \\ \int s\,\gamma(\mathrm{d}s), & \text{otherwise.} \end{cases}$$

Note that, even though Receiver's optimal action in this example depends only on his conditional mean of the state, given his private information and Sender's signal, Sender's payoff $V(\gamma)$ from a posterior γ over states is not a function of the posterior mean $\int s\,\gamma(\mathrm{d}s)$ alone. Thus, the example is not in the class of persuasion problems analyzed by Gentzkow and Kamenica (2016) and Dworczak and Martini (2017).

The cav-V approach states that the optimal utility for Sender is given by $\mathrm{cav}V(\gamma_0)$, where $\mathrm{cav}V$ is the concave envelope of V and γ_0 is the prior over states, which is uniform in our example. Because V is piecewise linear, the concavification is achieved at some convex combination:

$$\gamma_0 = \alpha_1\gamma_1 + \alpha_2\gamma_2 + \alpha_3\gamma_3, \tag{8}$$

where the distributions $\gamma_1, \gamma_2, \gamma_3$ belong to the three regions that define V. Finding the optimal combination amounts to solving the following infinite-dimensional LP problem with variables $g_1, g_2, g_3 \in L^\infty(\gamma_0)$ that correspond to the densities of $\alpha_1\gamma_1$, $\alpha_2\gamma_2$, and $\alpha_3\gamma_3$ w.r.t. γ_0. (It follows from (8) that γ_i are absolutely continuous w.r.t. γ_0.) Sender's problem is:

$$\begin{aligned} \underset{g_1,g_2,g_3}{\text{Maximize}} \quad & \int_0^1 g_1(s) + s g_2(s)\,\mathrm{d}s \\ \text{subject to} \quad & g_1(s) + g_2(s) + g_3(s) = 1, \ \forall s, \\ & \int_0^1 (1-s)(s-\zeta) g_1(s)\,\mathrm{d}s \geqslant 0, \\ & \int_0^1 s(s-\zeta) g_2(s)\,\mathrm{d}s \geqslant 0. \end{aligned}$$

In contrast, as shown in Section 4, our approach reduces Sender's problem to a finite-dimensional constrained optimization, which is much simpler than the infinite-dimensional LP problem above.

5 A continuous-type example

We next illustrate the main results by a continuous-type example. An additional goal is to show the relationship between Sender's problem in our setup and the monopoly screening problem, as in Mussa and Rosen (1978).

In the examples in this section the type space is $\mathcal{T} = [0, 1]$ and the state space is $\mathcal{S} = [-1, 1]$. Receiver's $u(s)$ is constant over negative states, i.e., there exists $\eta > 0$ such that $u(s) = -\eta$ for every $s < 0$. We assume that $u(s) \geqslant 0$ for $s \geqslant 0$, so that Receiver prefers to accept if and only if $s \geqslant 0$. Here, the parameter $\eta > 0$ measures how costly it is for Receiver to accept when the state is negative. The density function of every type over states is also constant over negative states, i.e., there exists some $\underline{f} : \mathcal{T} \to \mathbf{R}_+$ such that $f(s,t) = \underline{f}(t)$ for every $s < 0$. Finally, we also assume that the marginal distribution over types is uniform so that $\int_{-1}^{1} f(s,t)\,\mathrm{d}s = 1$ for every t. The last assumption is for convenience only.

The highest type's payoff if he accepts under his prior belief is

$$\int_{-1}^{1} f(s,1)u(s)\,\mathrm{d}s = -\eta\underline{f}(t) + \int_{0}^{1} f(s,1)u(s)\mathrm{d}s.$$

We make a simplifying assumption that the highest type rejects in the absence of further information.

Assumption 2. *The highest type rejects in the absence of further information.*

From Theorem 3.1 we know that each type accepts on an interval of states that contains state 0. We let $[\underline{\pi}(t), \overline{\pi}(t)]$ denote the interval for type t. Consider an arbitrary such interval $\left[\underline{q}, \overline{q}\right]$ with $-1 \leqslant \underline{q} \leqslant 0 \leqslant \overline{q} \leqslant 1$. The payoff of type t from accepting on this interval is given by

$$\int_{\underline{q}}^{\overline{q}} f(s,t)u(s)\mathrm{d}s = \underline{f}(t) \cdot \left(\overline{U}(\overline{q}, t) + \eta\underline{q}\right),$$

where $\overline{U}(\overline{q}, t) = \int_0^{\overline{q}} f(s,t)u(s)\,\mathrm{d}s/\underline{f}(t)$ is the utility of type t from accepting on $[0, \overline{q}]$, normalized by $\underline{f}(t)$. Thus, when we restrict attention to acceptance sets that are intervals, we can view type t as a buyer with quasilinear utility who gets utility $\overline{U}(\overline{q}, t) + \eta\underline{q}$ from receiving quantity $\overline{q}$ of a divisible good and paying $-\eta\underline{q}$ money. The private IC constraints in our framework translate to the standard IC constraints for selling the good when the buyer's type is private.

The Spence-Mirrlees sorting condition $\frac{\partial}{\partial \overline{q}, \partial t} \overline{U}(\overline{q}, t) \geqslant 0$ holds in our setup given the i.m.l. ratio assumption. From the theory of optimal screening we therefore know that IC constraints hold if and only if $\overline{\pi}$ is monotone-increasing and

$$- \eta \underline{\pi}(t) = \overline{U}(\overline{\pi}(t), t) - \int_0^t \overline{U}_2(\overline{\pi}(\tau), \tau) \, \mathrm{d}\tau, \tag{9}$$

where $\overline{U}_2$ is the derivative of $\overline{U}$ w.r.t. the type. Note that given $\overline{\pi}(t)$, the IC constraints determine the "payment" $-\eta \underline{\pi}(t)$ up to a constant, so we could add some $C \leqslant 0$ to the right-hand side; however as in the monopoly screening problem, it is optimal to choose $C = 0$. Also, our setup has an additional requirement that $\underline{\pi}(t) \geqslant -1$. (In terms of the monopoly screening problem, this would correspond to an upper bound on the payment.) Assumption 2 makes sure that this constraint does not bind.

For every $\overline{q} > 0$, we let $\overline{F}(\overline{q}, t) = \int_{s=0}^{\overline{q}} f(s, t) \, \mathrm{d}s$ be Sender's utility from accepting in $[0, \overline{q}]$. Sender's payoff is given by

$$\begin{aligned} \int_0^1 \int_{\underline{\pi}(t)}^{\overline{\pi}(t)} f(s,t) \mathrm{d}s \mathrm{d}t &= \int_0^1 \left(-\underline{f}(t) \underline{\pi}(t) + \int_0^{\overline{\pi}(t)} f(s,t) \mathrm{d}s \right) \mathrm{d}t \\ &= \int_0^1 \left(\frac{\overline{U}(\overline{\pi}(t), t) \underline{f}(t) - \overline{U}_2(\overline{\pi}(t), t) \int_t^1 \underline{f}(z) \mathrm{d}z}{\eta} + \overline{F}(\overline{\pi}(t), t) \right) \mathrm{d}t. \end{aligned} \tag{10}$$

Thus, Sender's problem in our setup has the same structure as the monopoly screening problem: we need to maximize (10) over all monotone-increasing functions $\overline{\pi} : [0, 1] \to [0, 1]$. As in the case of the monopoly screening problem, this is in general a control problem. Under some regularity assumption, the monotonicity constraints do not bind, so we can continue without resorting to the control theory.

We now assume that the joint density of (s, t) is

$$f(s,t) = \begin{cases} \frac{2}{\phi(2t-1)+4}, & \text{if } s \in [-1, 0), \\ \frac{2}{\phi(2t-1)+4} (\phi s (2t-1) + 1), & \text{if } s \in [0, 1]. \end{cases}$$

The parameter $\phi \in [0, 1]$ measures how informative Receiver's type is: when $\phi = 0$, Receiver has no private information. As ϕ increases, a higher type becomes progressively more optimistic than a lower type does. Receiver's utility from accepting is $u(s) = -\eta < 0$ if $s < 0$, and $u(s) = s$ if $s \geqslant 0$.

Sender's and type t's utilities if type t accepts on $[0, \overline{q}]$ are given by

$$\overline{F}(\overline{q}, t) = \frac{\overline{q}(\phi\overline{q}(2t-1)+2)}{\phi(2t-1)+4}, \text{ and } \overline{U}(\overline{q}, t) = \frac{1}{6}\overline{q}^2(2\phi\overline{q}(2t-1)+3).$$

Substituting into Sender's problem (10) and simplifying, we can write Sender's payoff as

$$\int_0^1 \left\{ \frac{2\left(\frac{\phi(2t-1)}{\phi(2t-1)+4} - \log\left(\frac{\phi+4}{\phi(2t-1)+4}\right)\right)}{3\eta}\overline{\pi}(t)^3 + \frac{\eta\phi(2t-1)+1}{\eta(\phi(2t-1)+4)}\overline{\pi}(t)^2 + \frac{2}{\phi(2t-1)+4}\overline{\pi}(t) \right\} \mathrm{d}t. \tag{11}$$

Sender maximizes her payoff by choosing $\overline{\pi}(t)$ to maximize the integrand pointwise. If we ignore the constraint $\overline{\pi}(t) \leqslant 1$, the integrand is maximized at:

$$\overline{\pi}^*(t) := \frac{2\eta}{\eta\phi(1-2t) - 1 + \sqrt{(\eta\phi(1-2t)+1)^2 + 4\eta(\phi(2t-1)+4)\log\left(\frac{\phi+4}{\phi(2t-1)+4}\right)}}. \tag{12}$$

When Receiver's type is not informative enough, $\overline{\pi}^*(t)$ is greater than 1. The integrand is maximized at the highest state 1 for every t. In this case, the optimal mechanism is pooling.

Proposition 5.1. *There exists an increasing function $\Phi(\cdot)$ such that Sender pools all types if and only if $\phi \leqslant \Phi(\eta)$: Sender sets $\overline{\pi}(t)$ to be 1, and $\underline{\pi}(t)$ is a constant which is chosen such that the lowest type is indifferent.*

Figure 2 shows how the optimal mechanism varies as (ϕ, η) vary. Assumption 2 states that the highest type rejects without further information, i.e., $6\eta \geqslant 2\phi + 3$. This corresponds to the parameter region above the solid line. The dashed line corresponds to the function $\Phi^{-1}(\cdot)$. To the right of the dashed line, semi-separating is optimal; to the left, pooling is optimal.

When $\phi > \Phi(\eta)$, the maximizer $\overline{\pi}^*(t)$ is below 1 when Receiver has the lowest type 0. We show that $\overline{\pi}^*(t)$ increases to 1 at some type $\hat{t}$. Thus, the integrand in (11) is maximized at $\overline{\pi}^*(t)$ when $t \leqslant \hat{t}$ and at 1 when $t > \hat{t}$. Given this $\overline{\pi}(t)$, $\underline{\pi}(t)$ can be derived based on the incentive constraints (9). We show in Lemma 7.3 (in Section 7.3) that both $\overline{\pi}^*(t)$ and the corresponding $\underline{\pi}(t)$ increase in η for any fixed t.

Proposition 5.2. *Suppose that $\phi > \Phi(\eta)$. Then,*

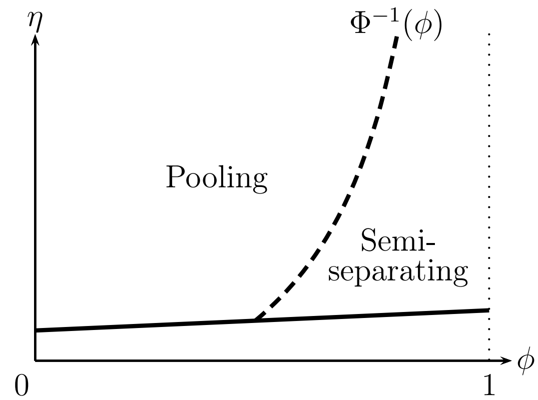

Figure 2: Pooling is optimal when ϕ is small

(i) $\overline{\pi}(t) = \begin{cases} \overline{\pi}^*(t) & \text{for } t \leqslant \hat{t} \\ 1 & \text{for } t > \hat{t}, \end{cases}$ *where* $\hat{t}$ *is the critical type such that* $\overline{\pi}^*\left(\hat{t}\right) = 1$*;*

(ii) $\underline{\pi}(t)$ *is determined by the incentive constraints* (9).

Figure 3 illustrates the optimal mechanism in which Sender separates the lower types.[4] The left-hand side illustrates each type t's acceptance interval $[\underline{\pi}(t), \overline{\pi}(t)]$, which expands as t increases. The solid curves correspond to $\overline{\pi}(t)$ and $\underline{\pi}(t)$ for a lower η, and the dashed ones for a higher η. As η increases, Sender is less capable of persuading Receiver to accept in unfavorable states. Hence, both allocations $\overline{\pi}(t), \underline{\pi}(t)$ increase.

The right-hand side of Figure 3 illustrates how to implement the optimal mechanism in a publicly IC mechanism by examining the case where η is 1. The solid curve corresponds to the U-shaped cutoff function $z(s)$. For all the states below $\underline{\pi}(1)$, Sender recommends rejection. For any state s above $\underline{\pi}(1)$, Sender announces $z(s)$ and recommends that types above $z(s)$ accept. For a small segment of states surrounding state 0, $z(s)$ equals the lowest type, so all types accept. For any higher state s, Sender always mixes it with a lower state so that type $z(s)$ is indifferent when he is pronounced to be the cutoff type.

6 Discussion

In this section, we extend our results to more general payoff and information settings. We then provide some examples to show (i) that Theorem 3.1 and Theorem 3.2 may not hold without the common prior or the i.m.l. ratio assumption and (ii) that not every privately

[4] The parameter values are $\phi = 1$, $\eta = 1$ for the solid line, and $\eta = 4/3$ for the dashed line.

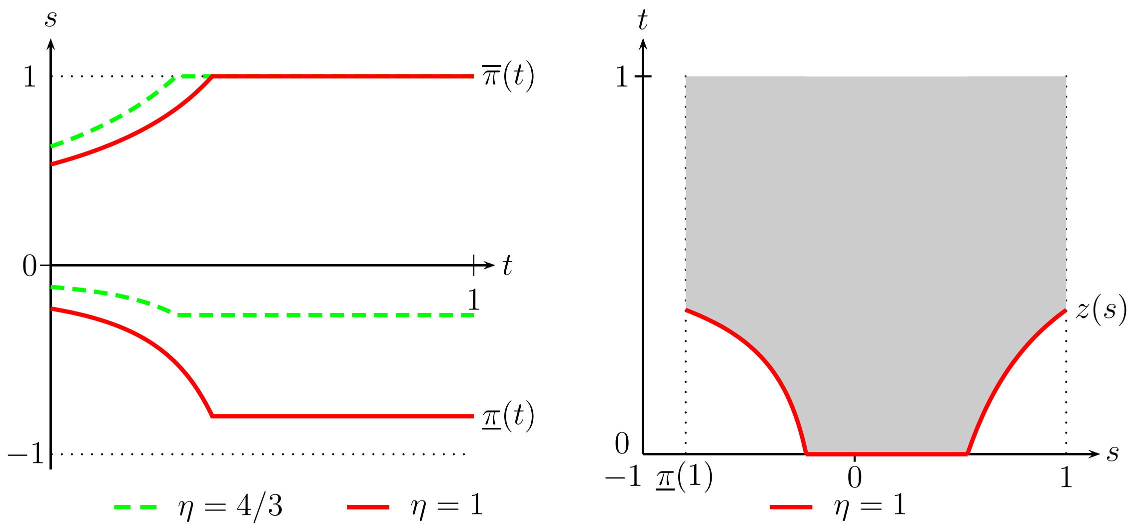

Figure 3: Optimal mechanism and its implementation

IC mechanism can be duplicated by a publicly IC one. All the assertions in the examples in this section were verified by solving the corresponding LP problems.

6.1 More general settings

We make the following modifications to the model:

- We assume that Receiver's payoff from accepting depends on his type. We use $u(s,t)$ to denote type t's payoff when accepting at state s.
- We assume that Sender's payoff from accepting depends on the state and on Receiver's type. We use $v(s,t)$ to denote Sender's payoff when Receiver t accepts at state s. We assume that $v(s,t) > 0$ for every s,t.
- We allow heterogeneous beliefs. We continue to denote Receiver's belief by $f(s,t)$, but we now denote Sender's belief by $g(s,t)$. We assume that $g(s,t) > 0$ for every s,t.

We omit the definitions of incentive compatibility and Sender's problem in this new environment but they should be straightforward. For Theorem 3.1, 3.2, and 3.3 and the proof to hold, we need two assumptions: First, for every type t, we need $\frac{f(s,t)u(s,t)}{g(s,t)v(s,t)}$ to be monotone in s. Second, for every $t' < t$, we need $\frac{f(s,t)u(s,t)}{f(s,t')u(s,t')}$ to be monotone in s. Note that the second assumption implies in particular that there exists some $s_0 \in \mathcal{S}$ such that $u(s,t) \geq 0$ for $s \geq s_0$ and $u(s,t) \leq 0$ for $s \leq s_0$. And, as before, in order for the cutoff type to always accept, we also need $u(s,t), v(s,t), f(s,t)$ and $g(s,t)$ to be continuous in t.

One interesting example occurs when, with all else fixed, Sender's payoff depends on the state. In the context of the theorist-chair example, this is the case if the theorist's payoff from hiring also increases in the state. Our theorems still hold when we replace the assumption that u is monotone in s with the assumption that u/v is monotone.

6.2 The i.m.l. ratio assumption

The following example shows that without the i.m.r. ratio assumption—ranking of types w.r.t. the monotone-likelihood-ratio order—Theorems 3.1 and 3.2 may not hold.

Example 1. Let $\mathcal{S} = \{-4, -3, 3\}$ with Receiver's payoff from accepting given by $u(s) = s$, and let $\mathcal{T} = \{T, B\}$. The density over states and types is given by

	-4	-3	3
T	$1/6,$	$1/6,$	$1/6$
B	$3/54,$	$20/54,$	$4/54$

The unique optimal publicly IC mechanism pools the two types and, if the state is s, recommends that both accept with probability $\rho(s)$ where

$$\rho(-4) = 12/17, \quad \rho(-3) = 1/17, \quad \rho(3) = 1. \tag{13}$$

The unique optimal privately IC mechanism recommends that T accept (with probability 1) if the state is 3 or -3, and reject if the state is -4. It recommends that B accept with probability $\rho(s)$ if the state is s where ρ is given by (13). Thus, the optimal publicly IC mechanism and the optimal privately IC mechanism give different payoffs to Sender, and neither recommends that type B accept on an interval.

6.3 The common-prior assumption

Our formulation implicitly assumed a common prior between Sender and Receiver, since the same density function f over states and types was used in the definitions of incentive compatibility and Sender's problem. The following example shows that Theorem 3.1 and Theorem 3.2 may not hold without this assumption.

Example 2. Assume that $\mathcal{S} = \{-2, -1, 1\}$ with $u(s) = s$, and $\mathcal{T} = \{H, L\}$. Let Receiver's and Sender's beliefs be given by the following density functions:

Receiver's belief

	-2	-1	1
H	1/10,	2/10,	2/10
L	4/12,	1/12,	1/12

Sender's belief

	-2	-1	1
H	8/20,	4/20,	4/20
L	2/20,	1/20,	1/20

Receiver's belief satisfies the i.m.l. ratio assumption and Sender believes that the type and the state are independent.

The unique optimal publicly IC mechanism gives up on L and recommends that H accept if the state is either -2 or 1, and reject if the state is -1. The unique optimal privately IC mechanism recommends that L accept if the state is either -1 or 1, and reject if the state is -2; it also recommends that H accept if the state is either -2 or 1, and reject if the state is -1.

Thus, the optimal publicly IC mechanism and the optimal privately IC mechanism give different payoffs to Sender, and neither recommends type H accept on an interval.

6.4 Equivalence between public and private incentive-compatibility

Theorem 3.2 implies that, for the optimal privately IC mechanism, there exists a publicly IC mechanism which is equivalent in the sense that each type has the same probability of accepting. The following example shows that such an equivalent mechanism does not necessarily exist for a nonoptimal privately IC mechanism. The example shows that the strong equivalence result in Kolotilin et al. (2016) does not hold in our setup.

Example 3. Let $\mathcal{S} = \{-1000, 1, 10\}$ with utility $u(s) = s$, and let $\mathcal{T} = \{H, L\}$. The prior over types and states is given by

	-1000	1	10
H	5/22,	5/22,	1/22
L	20/82,	20/82,	1/82

Consider a mechanism which recommends that H accept if the state is 10 and reject otherwise, and which recommends that L accept if the state is 1 and reject otherwise. This mechanism is privately IC (and of course not optimal). Under this mechanism, type H accepts with the interim probability 1/11 and type L accepts with the interim probability 20/41. However, no publicly IC mechanism exists with these interim acceptance probabilities.

7 Proofs

7.1 Proof of Theorem 3.3

We assume w.l.o.g. that $u(0) = 0$. We also assume w.l.o.g. that $\mathcal{S}$ is an interval and that μ is Lebesgue's measure. This is because the IC properties of a mechanism and Sender's payoff depend only on the joint distribution of the state and the mechanism's recommendations to each type; moreover, by Skorokhod's representation theorem, the random variable that represents the state can be implemented as a monotone function of an interval state space. For example, if the state space is binary with two equally probable states $s_0 < s_1$, then one can think of the state as a function of some $s \in [-1, 1]$ drawn from a uniform distribution. Thus, the state is s_0 if $s \in [-1, 0]$ and s_1 if $s \in [0, 1]$, and Sender can create s by randomizing from a uniform distribution on $[-1, 0]$ or $[0, 1]$ when the state is s_0 or s_1, respectively. This leads to a correspondence between mechanisms defined on the state space $\{s_0, s_1\}$ and mechanisms defined on the state space $[-1, 1]$, a correspondence that preserves the IC properties and Sender's payoff (but which may transform a deterministic mechanism into a nondeterministic mechanism).

For a given mechanism with acceptance probabilities ρ, let

$$v(t) = \int f(s,t)\rho(s,t)\ \mu(\mathrm{d}s)$$

be the normalized probability that type t accepts. We say that a mechanism ρ *weakly dominates* another mechanism ρ' if $v(t) \geqslant v'(t)$ for every type t. We say that ρ *dominates* ρ' if it weakly dominates ρ', and $v(t) > v'(t)$ for some t.

The proof of Theorem 3.3 will use Lemma 7.1 and 7.2. Lemma 7.1 establishes the fact that when one searches for the optimal downward IC mechanism, it is sufficient to search among U-shaped cutoff mechanisms that are downward IC. Lemma 7.2 asserts that, for any cutoff mechanism that is downward IC, if we publicly declare the cutoff type x, then all types which are supposed to accept (i.e., types t such that $t \geqslant x$) will still accept. It is possible that lower types will also accept.

Lemma 7.1. *For every downward IC mechanism there exists a U-shaped function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$, such that the cutoff mechanism given by z is downward IC and weakly dominates the original mechanism.*

Lemma 7.2. *For every cutoff mechanism κ that is downward IC, the publicly IC mechanism induced by this mechanism has the property that type t accepts when $t \geqslant x$.*

Using Lemmas 7.1 and 7.2, we prove Theorem 3.3 in three steps. First, among all U-shaped cutoff mechanisms that are downward IC, there exists one z^* that is optimal to Sender. Second, this mechanism z^* is not dominated by any other U-shaped cutoff mechanism that is also downward IC. Third, z^* is publicly IC.

For the case in which the type space $\mathcal{T}$ is finite, the first and second steps of the proof are immediate since the space of U-shaped cutoff mechanisms is finite-dimensional; this is because every U-shaped cutoff mechanism is given by the endpoints of the intervals. For the case where $\mathcal{T}$ is a continuum, we need to be more careful in establishing the existence of Sender's optimal mechanism and in showing that it is not dominated, because of the possible problem with sets of types of measure zero. Aside from these nuisances, the core of the proof is in the third step. This step uses Lemma 7.1 and Lemma 7.2.

Proof of Theorem 3.3. Consider the space Z of all U-shaped functions $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$ with minimum at 0. This space, viewed as a subspace of $L^\infty(\mathcal{S} \to \mathcal{T} \cup \{\infty\})$ equipped with the weak star topology, is compact. The set of such functions $z \in Z$ which give rise to downward IC mechanisms is a closed subset of Z. We denote this subset by $\tilde{Z}$. Sender's payoff $\iint_{z(s)\leq t} f(s,t) \ \mu(\mathrm{d}s)\,\lambda(\mathrm{d}t)$ is a continuous function of z. Therefore, there exists a $z^* \in \tilde{Z}$ which maximizes Sender's payoff.

For each $z \in \tilde{Z}$, the normalized probability that type t accepts $v_z(t) = \int_{t\geqslant z(s)} f(s,t) \ \mu(\mathrm{d}s)$ is a right-continuous function of t, which follows from the continuity assumption on f. Because v_z is right-continuous and λ has full support, the maximum z^* cannot be dominated by any $z \in \tilde{Z}$, i.e., if $v_z(t) \geq v_{z^*}(t)$ for every t, then $v_z = v_{z^*}$.

By Lemma 7.1 the cutoff mechanism induced by z^* is optimal for Sender among all downward IC mechanisms. We must still show that z^* is publicly IC. By Lemma 7.2 the acceptance set of each type t under the publicly IC mechanism induced by z^* is at least the event $\{s : z^*(s) \leqslant t\}$. Therefore, if the cutoff mechanism z^* were not publicly IC, then it would be dominated by a publicly IC mechanism. By Lemma 7.1 again, this publicly IC mechanism is itself weakly dominated by a U-shaped cutoff mechanism given by some $z \in \tilde{Z}$, which is a contradiction to the fact that z^* is not dominated by any $z \in \tilde{Z}$. □

The proof of Lemma 7.1 has two steps. We begin with an arbitrary downward IC mechanism. In the first step we concentrate the acceptance probabilities of each type t to an interval $\left[\underline{p}(t), \overline{p}(t)\right]$ around 0, in such a way that the utility (disutility) that each type gets from positive states (negative states) is the same as in the original mechanism. This step preserves the downward IC conditions and weakly increases the acceptance probability of each type.

In the second step we make the mechanism a cutoff mechanism by essentially letting each type accept at the union of his interval and all intervals of lower types. This creates the new acceptance intervals $[\underline{\pi}(t), \overline{\pi}(t)]$, which are nested in the sense that $\underline{\pi}$ is monotone-decreasing and $\overline{\pi}$ is monotone-increasing. If $\mathcal{T}$ is finite, then we can now define a cutoff function z that induces these acceptance intervals, i.e., such that

$$z(s) \leqslant t \longleftrightarrow \underline{\pi}(t) \leqslant s \leqslant \overline{\pi}(t), \tag{14}$$

by $z(s) = \min\{t \colon s \in [\underline{\pi}(t), \overline{\pi}(t)]\}$. The case of continuum types has an additional complication, because in order to obtain a cutoff function z such that (14) holds, we need an additional continuity assumption on $\underline{\pi}(t), \overline{\pi}(t)$.[5] This additional complication is due to our insistence that the cutoff type will accept. A more general definition, which allows the mechanism to recommend that the cutoff type either accept or reject, would have spared us some technical difficulties in the proof, but we chose to make the definitions easier for the reader.

Proof of Lemma 7.1. For every type t let $\underline{p}(t)$ and $\overline{p}(t)$ be such that $\underline{p}(t) \leqslant 0 \leqslant \overline{p}(t)$ and

$$\begin{aligned} \int_0^\infty f(s,t)u(s)\rho(s,t)\,\mu(\mathrm{d}s) &= \int_0^{\overline{p}(t)} f(s,t)u(s)\,\mu(\mathrm{d}s), \text{ and} \\ \int_{-\infty}^0 f(s,t)u(s)\rho(s,t)\,\mu(\mathrm{d}s) &= \int_{\underline{p}(t)}^0 f(s,t)u(s)\,\mu(\mathrm{d}s). \end{aligned} \tag{15}$$

From (15) and the i.m.l. ratio assumption, it follows that:

$$\begin{aligned} \int_0^\infty f(s,t)u(s)\rho(s,t')\,\mu(\mathrm{d}s) &\geqslant \int_0^{\overline{p}(t')} f(s,t)u(s)\,\mu(\mathrm{d}s), \text{ and} \\ \int_{-\infty}^0 f(s,t)u(s)\rho(s,t')\,\mu(\mathrm{d}s) &\geqslant \int_{\underline{p}(t')}^0 f(s,t)u(s)\,\mu(\mathrm{d}s) \end{aligned} \tag{16}$$

for $t' < t$. From (15), (16), and (4), it follows that:

$$\int f(s,t)u(s)\left(\mathbf{1}_{[\underline{p}(t),\overline{p}(t)]} - \mathbf{1}_{[\underline{p}(t'),\overline{p}(t')]}\right)\mu(\mathrm{d}s) \geqslant 0 \tag{17}$$

for every $t' < t$.

[5] For example, if $[\underline{\pi}(t), \overline{\pi}(t)] = \begin{cases} [-2,2], & \text{if } t > 1 \\ [-1,1] & \text{if } t \leq 1 \end{cases}$, then no such z exists.

In addition, the monotonicity of u and the fact that $0 \leqslant \rho(s,t) \leqslant 1$ imply that

$$\int_0^\infty f(s,t)\rho(s,t)\,\mu(\mathrm{d}s) \leqslant \int_{\underline{p}(t)}^{\overline{p}(t)} f(s,t)\,\mu(\mathrm{d}s) \tag{18}$$

by the Neyman-Pearson Lemma.

Thus, the mechanism with acceptance intervals $\big[\underline{p}(t), \overline{p}(t)\big]$ is downward IC (17) and weakly dominates the original mechanism (18). Note, however, that this is not yet a cutoff mechanism.

We now introduce nested acceptance intervals $[\underline{\pi}(t), \overline{\pi}(t)]$ which are downward IC and which will give rise to a cutoff mechanism. Let $\overline{\pi}(t) = \inf_{\varepsilon>0}\sup\{\overline{p}(t')\colon t' < t+\varepsilon\}$ and $\underline{\pi}(t) = \inf_{\varepsilon>0}\inf\big\{\underline{p}(t')\colon t' < t+\varepsilon\big\}$ be the right-continuous and monotone functions that dominate $\underline{p}$ and $\overline{p}$, respectively, and let z be given by (14):

$$z(s) \leqslant t \longleftrightarrow \underline{\pi}(t) \leqslant s \leqslant \overline{\pi}(t).$$

It is easy to see that the cutoff mechanism given by z with acceptance intervals $[\underline{\pi}(t), \overline{\pi}(t)]$ weakly dominates the mechanism with acceptance intervals $\big[\underline{p}(t), \overline{p}(t)\big]$. We claim that the former mechanism is downward IC.

Let $t' < t$. We first need to show that for the mechanism $(\underline{\pi}, \overline{\pi})$, type t will not mimic a lower type t', i.e., that

$$\int f(s,t)u(s)\left(\mathbf{1}_{[\underline{\pi}(t),\overline{\pi}(t)]} - \mathbf{1}_{[\underline{\pi}(t'),\overline{\pi}(t')]}\right)\,\mu(\mathrm{d}s) \geqslant 0. \tag{19}$$

Let t_k, t'_k be such that

$$\lim_{k\to\infty} t_k = t_\infty \leqslant t \text{ and } \underline{p}(t_k) \uparrow \underline{\pi}(t), \text{ and}$$
$$\lim_{k\to\infty} t'_k = t'_\infty \leqslant t' \text{ and } \overline{p}(t'_k) \uparrow \overline{\pi}(t').$$

If $t_\infty \leqslant t'$, then $\underline{\pi}(t') = \underline{\pi}(t)$; and from (14) it follows that $\{t' \leqslant z(s) < t\} \subseteq \{s \geqslant 0\}$, so (19) holds. Therefore, we can assume that $t' < t_\infty$ and therefore, $t'_k < t_k$ for every k.

From the definition of $\overline{\pi}$, the continuity assumption on f, and the fact that $\lim_{k\to\infty} t_k = t_\infty$, it follows that $\limsup_{k\to\infty}\overline{p}(t_k) \leqslant \overline{\pi}(t_\infty)$ and $\lim_{k\to\infty} f(s,t_k) = f(s,t_\infty)$. In addition, we know that $\limsup_{k\to\infty}\overline{p}(t'_k) = \overline{\pi}(t')$. From these properties and Fatou's Lemma, it follows that

$$\limsup_{k\to\infty} \int_{\overline{p}(t'_k)}^{\overline{p}(t_k)} f(s,t_k)u(s)\,\mu(\mathrm{d}s) \leqslant \int_{\overline{\pi}(t')}^{\overline{\pi}(t_\infty)} f(s,t_\infty)u(s)\,\mu(\mathrm{d}s).$$

By a similar argument, we have that $\limsup_{k\to\infty} \underline{p}(t'_k) \geqslant \underline{\pi}(t')$ and $\limsup_{k\to\infty} \underline{p}(t_k) = \underline{\pi}(t)$. When this result is combined with the condition that $u(s) \leqslant 0$ for any $s \leqslant 0$, we have that

$$\limsup_{k\to\infty} \int_{\underline{p}(t_k)}^{\underline{p}(t'_k)} f(s,t_k)u(s)\,\mu(\mathrm{d}s) \leqslant \int_{\underline{\pi}(t_\infty)}^{\underline{\pi}(t')} f(s,t_\infty)u(s)\,\mu(\mathrm{d}s).$$

Therefore, it follows that:

$$\begin{aligned}
&\int f(s,t_\infty)u(s)\left(\mathbf{1}_{[\underline{\pi}(t),\overline{\pi}(t)]} - \mathbf{1}_{[\underline{\pi}(t'),\overline{\pi}(t')]}\right)\,\mu(\mathrm{d}s)\\
\geqslant &\int_{\underline{\pi}(t_\infty)}^{\underline{\pi}(t')} f(s,t_\infty)u(s)\,\mu(\mathrm{d}s) + \int_{\overline{\pi}(t')}^{\overline{\pi}(t_\infty)} f(s,t_\infty)u(s)\,\mu(\mathrm{d}s)\\
\geqslant &\limsup_{k\to\infty}\left(\int_{\overline{p}(t'_k)}^{\overline{p}(t_k)} f(s,t_k)u(s)\,\mu(\mathrm{d}s) + \int_{\underline{p}(t_k)}^{\underline{p}(t'_k)} f(s,t_k)u(s)\,\mu(\mathrm{d}s)\right) \geqslant 0.
\end{aligned}$$

Given the i.m.l. ratio assumption, $\int f(s,t)u(s)\left(\mathbf{1}_{[\underline{\pi}(t),\overline{\pi}(t)]} - \mathbf{1}_{\underline{\pi}(t'),\overline{\pi}(t')]}\right)\,\mu(\mathrm{d}s)$ must be positive as well since $t \geqslant t_\infty$. This proves (19).

Finally, we need to show that the lowest type gets a payoff of at least zero from obeying under the mechanism $(\overline{\pi},\underline{\pi})$. In the case of discrete type spaces, this follows from the corresponding property of the original mechanism since in this case $\overline{\pi}(\underline{t}) = \overline{p}(\underline{t})$ and $\underline{\pi}(\underline{t}) = \underline{p}(\underline{t})$. In the general case we need to appeal to an argument similar to the one we used to prove the downward IC conditions using converging sequences of types. We omit this argument here. ☐

Proof of Lemma 7.2. Fix a type t. We need to show that in the induced publicly IC mechanism, type t accepts on the event $\{x \in B\}$ for every Borel subset $B \subseteq [\underline{t}, t]$, where x is the public signal produced by the mechanism. That is, we need to show that

$$\int f(s,t)u(s)\kappa(s,B)\,\mu(\mathrm{d}s) \geqslant 0.$$

It is sufficient to prove the assertion for sets B of the form $B = \{x : t'' < x \leqslant t'\}$ for some $\underline{t} \leqslant t'' < t' \leqslant t$ and for the set $B = \{\underline{t}\}$ since these sets generate the Borel sets. Indeed, for

$B = \{x : t'' < x \leqslant t'\}$, it holds that

$$\int f(s,t)u(s)\kappa(s,B)\,\mu(\mathrm{d}s) = U(t',t) - U(t'',t) \geqslant 0,$$

where the inequality follows from (i) downward incentive-compatibility when $t = t'$ and (ii) Lemma 3.4, which extends to $t \geqslant t'$. The case $B = \{\underline{t}\}$ follows by a similar argument from (5). □

7.2 Proof of Proposition 4.1

We first note that, under our assumption that type L rejects without additional information, both IC constraints in (6) bind. Indeed, for type H this holds trivially in the pooling case; in the separating case, if the constraint is not binding, then slightly increasing $\overline{\pi}_L$ would increase Sender's payoff without violating either type's IC constraint. For type L, we know from Theorem 3.1 that $u(\underline{\pi}_L) \leqslant 0$. If the IC constraint for type L is not binding, then making $\underline{\pi}_L$ slightly smaller will increase Sender's payoff without violating either type's IC constraint.

Since both IC constraints are binding and since $\underline{\pi}_H = 1$, the variables in Sender's problem (6) are determined by a single variable. If $\overline{\pi}_L = y$ for some $\zeta \leqslant y \leqslant 1$, then (i) $\underline{\pi}_L = \ell_L(y)$ where $\ell_L : [\zeta, 1] \to [0, \zeta]$ is given by

$$\int_{\ell_L(y)}^{y} f(s,L)u(s)\,\mathrm{d}s = 0;$$

and (ii) $\underline{\pi}_H = \ell_H(y)$ where $\ell_L : [\zeta, 1] \to [0, \zeta]$ is given by

$$\int_{\ell_H(y)}^{\ell_L(y)} + \int_{y}^{1} f(s,H)u(s)\,\mathrm{d}s = 0.$$

By the implicit function theorem, ℓ_L and ℓ_H are differentiable and their respective derivatives are given by:

$$\ell'_L(y) = \frac{u(y)}{u(\ell_L(y))}\frac{f(y,L)}{f(\ell_L(y),L)}, \quad \ell'_H(y) = \frac{u(y)}{-u(\ell_H(y))}\frac{f(y,H)f(\ell_L(y),L) - f(y,L)f(\ell_L(y),H)}{f(\ell_H(y),H)f(\ell_L(y),L)}.$$

It is easy to see that ℓ_L is monotone-decreasing and that the i.m.l. ratio assumption implies that ℓ_H is monotone-increasing.

In terms of the variable y, Sender's payoff is given by

$$R(y)=\int_{\ell_L(y)}^{y} f(s,L)\,\mathrm{d}s+\int_{\ell_H(y)}^{1} f(s,H)\,\mathrm{d}s.$$

Thus, Sender's payoff is differentiable. Substituting $\ell'_L(y)$ and $\ell'_H(y)$ into $R'(y)$, we obtain that $R'(y)$ is positive if and only if:

$$\frac{f(y,H)}{f(y,L)}-\frac{f(\ell_L(y),H)}{f(\ell_L(y),L)}<-u(\ell_H(y))\left(\frac{1}{u(y)}-\frac{1}{u(\ell_L(y))}\right).$$

Since $\ell'_L(y)\leqslant 0\leqslant \ell'_H(y)$, the left-hand side increases in y due to the i.m.l. ratio assumption and the right-hand side decreases in y since $u(s)$ increases in s. Therefore, $R'(y)$ is positive if and only if y is small enough. Therefore, pooling is optimal if and only if Sender's payoff achieves maximum at $y=1$, which is equivalent to $R'(1)\geqslant 0$. That is, the inequality above holds when y equals 1.

7.3 Proof of Propositions 5.1 and 5.2

Proof. We let $G(\overline{\pi}(t))$ denote the integrand in (11). (To simplify exposition, the dependence of the integrand on t is omitted.) $G(\overline{\pi}(t))$ is a cubic function in $\overline{\pi}(t)$:

$$G(\overline{\pi}(t)):=b_1(t)\overline{\pi}(t)^3+b_2(t)\overline{\pi}(t)^2+b_3(t)\overline{\pi}(t),$$

where

$$b_1(t):=\frac{2\left(\frac{2\phi t-\phi}{2\phi t-\phi+4}-\log\left(\frac{\phi+4}{2\phi t-\phi+4}\right)\right)}{3\eta},\quad b_2(t):=\frac{2\eta\phi t-\eta\phi+1}{\eta(2\phi t-\phi+4)},\quad b_3(t):=\frac{2}{\phi(2t-1)+4}.$$

The first-order condition is $G'(\overline{\pi}(t))=3b_1(t)\overline{\pi}(t)^2+2b_2(t)\overline{\pi}(t)+b_3(t)$. It can be readily verified that $b_3(t)>0$ for all $t\in[0,1]$. Moreover, $b_1(t)$ strictly increases in t; $b_1(0)<0$; and $b_1(1)>0$. We let $\tilde{t}\in(0,1)$ be the value which solves $b_1(t)=0$.

When $t\geqslant\tilde{t}$, $G'(\overline{\pi}(t))$ is strictly positive for any $\overline{\pi}(t)\geqslant 0$, so $G(\overline{\pi}(t))$ is maximized at $\overline{\pi}(t)=1$. When $t<\tilde{t}$, the equation $G'(\overline{\pi}(t))=0$ has a unique positive root since $b_1(t)<0$ and $b_3(t)>0$. This root is given by $\overline{\pi}^*(t)$ in (12). It can be easily verified that $\overline{\pi}^*(t)$ is monotone-increasing for $t\in\left[0,\tilde{t}\right]$ and goes to infinity as t approaches $\tilde{t}$.

If $\overline{\pi}^*(t)$ is above 1 at $t=0$, then $G(\overline{\pi}(t))$ is maximized at $\overline{\pi}(t)=1$ for every t. Thus, Sender pools all types when the following inequality holds:

$$\overline{\pi}^*(0) > 1 \longleftrightarrow \eta > \frac{(\phi-4)\log\left(\frac{\phi+4}{4-\phi}\right)}{\phi-1} - 1.$$

The right-hand side is monotone-increasing in ϕ, as desired.

If $\overline{\pi}^*(t)$ is below 1 at $t = 0$, there is a critical $\hat{t}$ which solves $\overline{\pi}(t) = 1$. Thus, $G(\overline{\pi}(t))$ is maximized at $\overline{\pi}^*(t)$ for $t \leqslant \hat{t}$ and at 1 for $t > \hat{t}$. □

Lemma 7.3. *Suppose that $\phi > \Phi(\eta)$. Both $\overline{\pi}(t)$ and $\underline{\pi}(t)$ increase pointwise in η. Thus, there exists $\bar{\eta}(\phi)$ such that the constraint $\underline{\pi}(t) \geqslant -1$ does not bind if and only if $\eta > \bar{\eta}(\phi)$.*

Proof of Lemma 7.3. When t equals 1/2, the value of $b_1(t)$, as defined in the proof of Proposition 5.1, is $\frac{2\log\left(\frac{4}{\phi+4}\right)}{3\eta}$, which is negative given that $\eta > 0$ and $\phi > 0$. When t equals 1/2, the value of (12) equals:

$$\frac{2\eta}{\sqrt{16\eta\log\left(\frac{\phi+4}{4}\right)+1}-1},$$

which is greater than 1. This implies that $\hat{t} < 1/2$. For the rest of the proof, we focus on the domain that $t \in [0, 1/2)$.

We first show that $\overline{\pi}(t)$ increases in η by showing that $1/\overline{\pi}(t)$ decreases in η. The derivative of $1/\overline{\pi}(t)$ w.r.t. η is negative if and only if

$$\eta(\phi - 2\phi t) + 2\eta(\phi(2t-1)+4)\log\left(\frac{\phi+4}{2\phi t - \phi + 4}\right) + 1$$
$$\geqslant \sqrt{(-2\eta\phi t + \eta\phi + 1)^2 + 4\eta(2\phi t - \phi + 4)\log\left(\frac{\phi+4}{2\phi t - \phi + 4}\right)}.$$

The left-hand side is concave in t, and it decreases in t at $t = 0$. Hence, the left-hand side decreases in t. Moreover, it is positive when $t = 1/2$, so the left-hand side is positive. Taking the power of both sides, we show that the inequality above holds.

We next show that $\underline{\pi}(0)$ increases in η. We solve for $\underline{\pi}(0)$ based on the condition that type 0's expected utility is zero:

$$\underline{\pi}(0) = \frac{2\eta\left(\eta\phi - 3\sqrt{(\eta\phi+1)^2 - 4\eta(\phi-4)\log\left(\frac{\phi+4}{4-\phi}\right)} + 3\right)}{3\left(\eta\phi + \sqrt{(\eta\phi+1)^2 - 4\eta(\phi-4)\log\left(\frac{\phi+4}{4-\phi}\right)} - 1\right)^3}.$$

It is easy to show that this term increases in η.

Lastly, we want to show that $\underline{\pi}(t)$ increases in η. To do so, we write $\underline{\pi}(t)$ (as well as $\overline{\pi}(t)$) as a function of t and η:

$$\underline{\pi}(t,\eta) = \frac{6\int_0^t \frac{2}{3}\phi\overline{\pi}(\tau,\eta)^3\, d\tau + \overline{\pi}(t,\eta)^2(\phi(2-4t)\overline{\pi}(t,\eta)-3)}{6\eta}.$$

We have shown above that $\underline{\pi}^{(0,1)}(0,\eta)$ is positive. Next, we show that $\underline{\pi}^{(1,1)}(t,\eta)$ is positive, that is, $\underline{\pi}^{(0,1)}(t,\eta)$ increases in t. This completes the proof that $\underline{\pi}^{(0,1)}(t,\eta)$ is positive, so $\underline{\pi}(t,\eta)$ increases in η.

We let $x(t,\eta)$ denote the square root term in $\overline{\pi}(t)$:

$$x(t,\eta) := \sqrt{(2\eta\phi(1-2t)+1)^2 + 8\eta(\phi(1-2t)-4)\log\left(\frac{2\phi t-\phi+4}{\phi+4}\right)}.$$

It can be easily verified that $x(t,\eta)-\eta\phi(2t-1)-1>0$. Given this condition and the fact that $t \in (0,1/2)$, $\phi \in (0,1)$, and $\eta > 0$, it follows that the derivative $\underline{\pi}^{(1,1)}(t,\eta)$ is positive if

$$a_1 x(t,\eta)^3 + a_2 x(t,\eta)^2 + a_3 x(t,\eta) + a_4 > 0, \tag{20}$$

where

$$\begin{aligned}
a_1 &= \eta(\phi(6t-3)-8)-3,\\
a_2 &= (\eta(\phi-2\phi t)+1)(\eta(7\phi(2t-1)+24)+1),\\
a_3 &= -(5\eta\phi(1-2t)+3)(\eta(\phi(2t-1)(\eta(\phi(2t-1)+8)-4)-24)-1),\\
a_4 &= -(\eta\phi(2t-1)-1)(\eta\phi(2t-1)+1)(\eta(\phi(2t-1)(\eta(\phi(2t-1)+8)-4)-24)-1).
\end{aligned}$$

It is easy to show that $a_2 > 0$ and that $x(t,\eta) > 1$. Moreover, the cubic inequality (20) is satisfied when $x(t,\eta) = 1$. We next rewrite the left-hand side of (20) as a quadratic function of $x(t,\eta)$:

$$a_2 x(t,\eta)^2 + (a_1 x(t,\eta)^2 + a_3)x(t,\eta) + a_4.$$

If we can show that $(a_1 x(t,\eta)^2 + a_3)$ is positive, then the quadratic function increases in $x(t,\eta)$ for any $x(t,\eta) \geqslant 1$. Given that the quadratic function is positive when $x(t,\eta) = 1$, the inequality (20) is satisfied. Next, we prove that $(a_1 x(t,\eta)^2 + a_3)$ is positive.

Substituting $x(t,\eta)$ into $(a_1 x(t,\eta)^2 + a_3)$, we find that this term is positive if

$$-\frac{2(\eta\phi(2t-1)(\eta\phi(2t-1)-4)+2)}{\eta(\phi(6t-3)-8)-3} + \log\left(\frac{2\phi t-\phi+4}{\phi+4}\right) > 0. \tag{21}$$

It is easy to verify that the left-hand side of (21) is convex in η. The inequality (21) holds when η equals zero. We are interested in the parameter region when $\overline{\pi}(t) < 1$. For fixed ϕ and t, $\overline{\pi}(t)$ is smaller than 1 if

$$\eta < \overline{\eta} := \frac{(\phi(2t-1)+4)\log\left(\frac{\phi+4}{2\phi t-\phi+4}\right)}{\phi(2t-1)+1} - 1.$$

When we substitute $\eta = \overline{\eta}$ into (21), the inequality is satisfied for any $t \in (0, 1/2)$ and $\phi \in (0, 1)$.

The left-hand side of (21) increases in η if and only if

$$\phi^2(1-2t)^2(\phi(6t-3)-8)\eta^2 - 6\phi^2(1-2t)^2\eta + 2(\phi(6t-3)+8) < 0.$$

The quadratic function on the left-hand side is concave and admits one positive root and one negative root. The positive root is given by

$$\tilde{\eta} := \frac{\sqrt{128-9\phi^2(1-2t)^2} + \phi(6t-3)}{\phi(2t-1)(\phi(6t-3)-8)}.$$

Given that $\eta > 0$, we obtain that the left-hand side of (21) decreases in η when $\eta < \tilde{\eta}$, and increases in η when $\eta \geqslant \tilde{\eta}$. We now have to discuss two cases, depending on whether $\tilde{\eta}$ is above or below $\overline{\eta}$:

1. If $\tilde{\eta} \geqslant \overline{\eta}$, the left-hand side of (21) decreases in η for any $\eta \in (0, \overline{\eta})$. We have shown that the left-hand side of (21) is positive at $\eta = 0$ and $\eta = \overline{\eta}$. Hence, the left-hand side of (21) is positive for any $\eta \in (0, \overline{\eta})$.

2. If $\tilde{\eta} \in (0, \overline{\eta})$, then the minimum of the left-hand side of (21) is achieved when $\eta = \tilde{\eta}$. We need to show that (21) holds when $\eta = \tilde{\eta}$. The condition $\tilde{\eta} < \overline{\eta}$ holds only if $\phi > 11/20$ and $t < 1/4$. When we substitute $\eta = \tilde{\eta}$ into (21), the inequality (21) is satisfied when we restrict attention to the parameter region $\phi > 11/20$ and $t < 1/4$.

By combining the two cases above, we have shown that (21) holds. □

References

Alonso, Ricardo, and Odilon Câmara. 2016*a*. "Bayesian Persuasion with Heterogeneous Priors." Working paper.

Alonso, Ricardo, and Odilon Câmara. 2016*b*. "Persuading Voters." *American Economic Review*, 106(11): 3590–3605.

Arieli, Itai, and Yakov Babichenko. 2016. "Private Bayesian Persuasion." Working paper.

Aumann, Robert J., and Michael Maschler. 1995. *Repeated Games with Incomplete Information.* The MIT Press.

Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin. 2015. "Designing and Pricing Information." Working paper.

Bergemann, Dirk, and Stephen Morris. 2016*a*. "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games." *Theoretical Economics*, 11(2): 487–522.

Bergemann, Dirk, and Stephen Morris. 2016*b*. "Information Design, Bayesian Persuasion, and Bayes Correlated Equilibrium." *American Economic Review: Papers and Proceedings*, 106(5): 586–591.

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang. 2016. "Pivotal Persuasion." Working paper.

Chen, Ying. 2009. "Communication with Two-sided Asymmetric Information." Working paper.

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

Dworczak, Piotr, and Giorgio Martini. 2017. "The Simple Economics of Optimal Persuasion." Working paper.

Gentzkow, Matthew, and Emir Kamenica. 2016. "A Rothschild-Stiglitz Approach to Bayesian Persuasion." *American Economic Review*, 5(106): 597–601.

Guo, Yingni, and Arjada Bardhi. 2016. "Modes of Persuasion Toward Unanimous Consent." Working paper.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Kolotilin, Anton. 2016. "Optimal Information Disclosure: A Linear Programming Approach." Working paper.

Kolotilin, Anton, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnyuk. 2016. "Persuasion of a Privately Informed Receiver." Working paper.

Lai, Ernest K. 2014. "Expert Advice for Amateurs." *Journal of Economic Behavior & Organization*, 103: 1–16.

Lehrer, Ehud, Dinah Rosenberg, and Eran Shmaya. 2010. "Signaling and Mediation in Games with Common Interests." *Games and Economic Behavior*, 68(2): 670–682.

Lehrer, Ehud, Dinah Rosenberg, and Eran Shmaya. 2013. "Garbling of Signals and Outcome Equivalence." *Games and Economic Behavior*, 81: 179–191.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. 2016. "Information Design: The Belief Hierarchy Approach." Working paper.

Mussa, Michael, and Sherwin Rosen. 1978. "Monopoly and Product Quality." *Journal of Economic Theory*, 18(2): 301–317.

Olszewski, Wojciech. 2004. "Informal Communication." *Journal of Economic Theory*, 117(2): 180–200.

Rayo, Luis, and Ilya Segal. 2010. "Optimal Information Disclosure." *Journal of Political Economy*, 118(5): 949–987.

Schnakenberg, Keith E. 2015. "Expert Advice to a Voting Body." *Journal of Economic Theory*, 160: 102–113.

Seidmann, Daniel. 1990. "Effective Cheap Talk with Conflicting Interests." *Journal of Economic Theory*, 50(2): 445–458.

Taneva, Ina. 2016. "Information Design." Working paper.

Watson, Joel. 1996. "Information Transmission When the Informed Party Is Confused." *Games and Economic Behavior*, 12(1): 143–161.