The Interval Structure of Optimal Disclosure

Yingni Guo* Eran Shmaya†

September 09, 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 the monopoly screening problem. The mechanism is optimal even if the sender conditions the disclosure mechanism on the receiver's reported type.

JEL: D81, D82, D83

Keywords: Information design, 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

Research on the Bayesian persuasion problem (Rayo and Segal (2010) and Kamenica and Gentzkow (2011)) studies sender-receiver games with commitment on the sender's side. It is used to model situations when an informed and interested party (the sender) tries to influence the action of another party (the receiver). Frequently, besides the information disclosed by the sender, the receiver has access to various external sources of information. In this paper we consider the Bayesian persuasion problem in which this external information is private.

For instance, a lobbyist tries to sway a legislator's position on an issue. The legislator himself has gained knowledge from working on similar issues in the past. He may also have access to various current information sources. In a different example, a media outlet tries to promote a political 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 personal 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 sender is uncertain about the receiver's opinion or his optimism level about the issue/candidate/agenda.

Environment. We consider an environment with two players: Sender and Receiver. Sender promotes a project to Receiver, who decides whether to accept or reject it. Receiver's utility from accepting increases with the project's quality, i.e., the *state*. His utility 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. Correspondingly, a higher type is more optimistic than a lower type that the state favors a decision to accept. 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. He then takes an action: to accept or to reject.

Note that the state in our framework captures two aspects of the game: it determines not only Receiver's utility from accepting but also Sender's belief about Receiver's type. Both aspects have implications for designing the optimal mechanism. The first aspect appears in all information design papers. The second aspect appears only in setups, like ours, in which

Receiver also has some private information about the state.

Main results. To illustrate the structure and the intuition of the optimal mechanism, we consider the case in which Receiver has two possible types: high or low. Sender could pool the two types such that both types always take the same action after any signal. Sender could also design a separating mechanism under which, after some signal, only the high type (who is more optimistic) accepts. When Receiver's external information is sufficiently informative, the high type is much more optimistic and thus much easier to persuade than the low type, which suggests that a separating mechanism may provide a higher payoff to Sender.

Under any 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 on a nested-interval structure. This structure 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, the high type is the only type that accepts when the state, viewed from Receiver's perspective, is 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 in which 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 decide based on his own private information. 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 makes 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 shared by 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 endpoints of each type's interval. Moreover, the interval structure offers a natural connection between (i) the 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-utility states as a buyer's utility from consuming a good of some quality level. Receiver's disutility from accepting on an interval of negative-utility states is the buyer's disutility from paying some price. Receiver's level of optimism (a.k.a., his 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 correspondence enables us to use the techniques from the screening problem to solve for the optimal mechanism.

In the baseline model, we adopt relatively simple payoff functions in order to illustrate the essence of our solution. Later on, we extend the results to more general payoff setups by allowing both Sender's payoff and Receiver's utility to depend on both the state and Receiver's type. We also show that our solution stays optimal in the environment in which Receiver first reports his type and Sender can disclose different information to different reported types.

Intuition for the interval structure. We again illustrate the basic intuition with the binary-type case. We first observe that the additional states in which only the high type accepts should be the extreme states of his acceptance set, because this is the cheapest way to maintain the incentives for different types. For states above the threshold state, the high type puts more weight on a higher state over a lower state than the low type does. Therefore, when only the high type rather than both types accepts, 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*. For states below the threshold, the low type puts more weight on a lower state over a higher state than the high type does. Therefore, when the low type is excluded so that only the high type accepts, 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, only the high type accepts when the state is sufficiently low.

To sum up, the high type is the only type who accepts (i) in sufficiently good states, because the low type does not believe these good states to be likely, and (ii) in sufficiently bad states, because the high type does not believe these bad states to be likely.

We then 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 utility from accepting. The set of states in which each type accepts 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 state, because Sender's payoff is less sensitive to the states than Receiver's.

In order to complete the proof, it is essential that the two properties we need—(i) that the states in which only the high type accepts are the extreme states in his acceptance set, and (ii) that each type's acceptance set is an interval such that the high type's interval contains the low type's interval—can be achieved simultaneously. One key observation in proving this is the fact 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 information design. 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 private information about the state. This is motivated by the observation that the receiver typically has multiple sources of information. Kamenica and Gentzkow (2011) extend the geometric method to situations in which the receiver has private information. Nonetheless, it is generally difficult to solve for the optimal mechanism using the geometric method. We use a concentration argument to establish the interval structure of the optimal mechanism; once established, that structure enables us to explicitly solve for the optimal mechanism.

Kolotilin (2017) also examines optimal persuasion when the receiver has private information. He provides a linear programming approach and establishes conditions under which full or no revelation is optimal. Kolotilin et al. (2017) assume that the receiver privately learns about his threshold for accepting. The receiver's threshold is independent of the state, and his utility is additive in the state and his threshold. Our paper differs in that the receiver's type is informative about the state.

In terms of the assumption about the receiver's private information, our paper is closely related to Li and Shi (2017) and Bergemann, Bonatti and Smolin (2017). Like ours, both papers assume that the buyer has private information about the state. Li and Shi (2017)

[1]Rayo and Segal (2010) assume that the receiver's taste is his private information and examine optimal persuasion when all types observe the same signal by the sender.

consider a seller who sells a product to a buyer whose type is his incomplete information about the state. For each type, the seller designs an experiment along with a strike price and an advance payment. They show that disclosing different information to different types dominates full disclosure. Bergemann, Bonatti and Smolin (2017) consider a monopolist who sells experiments to a buyer. The monopolist designs a menu of experiments and a tariff function to maximize his profit. Both papers examine the optimal design when the seller can disclose different information to different types. In contrast, we focus on settings where transfers are not allowed and our sender attempts to sway the receiver's action. Moreover, we focus on settings in which the sender discloses the same information to all types.

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. Our paper is most closely related to the work in Arieli and Babichenko (2016), who study optimal persuasion when the receivers have different thresholds for acceptance. They assume that receivers' thresholds are common knowledge, so there is no private information. Hence, their analysis and main results are quite different from ours.

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)). The setup in Chen (2009) is the closest to our setup; 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. In terms of the receiver's private information, our paper is also related to signaling games with a privately informed receiver. The nonmonotonicity of our optimal mechanism is also related to the countersignalling equilibrium in signaling games with privately informed receivers (e.g., Feltovich, Harbaugh and To (2002) and Angeletos, Hellwig and Pavan (2006)).

Structure of the paper. In section 2 we illustrate the essence of our problem and solution with a simple example. Section 3 presents the model and main results; we explain there how we use our theorem to connect Sender's problem to the monopoly screening problem. Section 4 discusses the binary-type case. (The finite-type case is no more complicated than the binary-type case.) It 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 use the techniques from the screening problem to solve the continuous-type case. The solutions in sections 4 and 5 allow us to identify the necessary and sufficient conditions for pooling or separating to be optimal. Section 6 extends our results to broader payoff and information setups. We further show that our mechanism remains optimal even if Receiver first reports his type and Sender discloses information differently depending on the reported type. Section 7 contains the proofs.

2 An example

To fix our idea, we illustrate our problem and main results by considering the communication between a theorist and his department chair. The theorist promotes a candidate to the chair, and the chair makes the hiring decision. The chair's utility from hiring depends on the state s which is uniform on $[0, 1]$. The chair's utility from hiring is $s - 3/4$. The theorist's payoff is one if the candidate is hired. If not, both players get a zero payoff.

The theorist 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 $1/2$. The theorist extracts the entire surplus. The chair's expected utility 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 announces that the candidate is a solid one. When the state is in the green region, the theorist announces that the candidate is creative but also risky. These regions are such that, given their private information, both types wants to hire if they learn that the state is in the red region, and only the type H wants to hire if they learn that the state is in the green region. The theorist does not extract the entire surplus, since type H gets a positive utility from hiring when the state is in the red region.

A central concept in our solution is the notion of nested intervals. Both types hire on an interval of states: type L hires only in the red region and type H hires in both the red

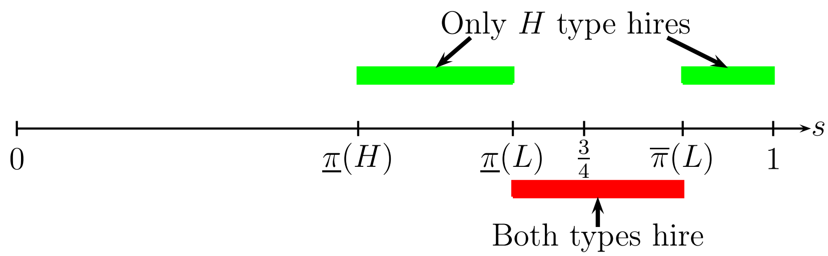

Figure 1: Optimal disclosure for binary types

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 quite 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 quite bad (i.e., in the green region), the theorist lets the chair decide based on his own private information.

To implement such an interval structure, the theorist announces at each state the lowest type who is still willing to hire. We call this type "the cutoff type" for 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.)

3 Environment and main results

Let $\mathcal{S}$, the set of *states*, be an interval on the real line equipped with Lebesgue's measure μ; let $\mathcal{T}$, the set of *(Receiver's) types*, be a subset of the real line equipped with a σ-finite measure λ of full support. In all our examples $\mathcal{T}$ is either a discrete space equipped with the counting measure or an interval equipped with Lebesgue's measure.[2] Consider a distribution over $\mathcal{S} \times \mathcal{T}$ with density f with respect to $\mu \times \lambda$. We make the following assumption about the type space $\mathcal{T}$ and the density f:

Assumption 1. *The set of types $\mathcal{T}$ is closed and bounded from below, with the lowest type denoted by $\underline{t}$. 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 $t' < t$.*

Receiver's type is his private information. Different types have different beliefs about

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

the state distribution. Given assumption 1, the higher t is, the more optimistic Receiver is about the state distribution.

Let $u\colon \mathcal{S} \to \mathbf{R}$ be a bounded and nondecreasing function representing *Receiver's utility* from accepting. Receiver's utility from rejecting is zero at every state. We assume that Sender's payoff is one if Receiver accepts and is zero otherwise. We adopt this payoff setup in order to simplify the exposition. In section 6.3, we extend our results to more general payoff setups by allowing both Sender's payoff and Receiver's utility to depend on both the state and Receiver's type.

Example 1. We extend the theorist-chair example from section 2. The set of states is $\mathcal{S} = [0, 1]$. The chair's utility from accepting is $u(s) = s - \zeta$, where ζ is a parameter. The type space is $\mathcal{T} = \{L, H\}$ with $L < H$. The joint density of states and types is given by

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

The parameter $\phi \in [0, 1]$ captures the accuracy of Receiver's private information. When $\phi = 0$, Receiver has no private information. As ϕ increases, Receiver's type becomes more informative about the state.

Example 1 shows that the state in our framework captures two aspects of the game: it determines not only Receiver's utility from accepting but also Sender's belief about Receiver's type. In our example, when the state is s, the chair's type is H with probability $1/2 + \phi(s - 1/2)$ and L with probability $1/2 - \phi(s - 1/2)$. Both aspects of the state have implications for the optimal disclosure mechanism. For this reason, 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.

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 from $\mathcal{S}$ to $\mathcal{X}$,[3] 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 $\kappa(s, \cdot)$ and recommends that type t accept if and only if $r(x, t) = 1$.

There are many ways in which Sender can disclose additional information about the state. To illustrate, we give a few examples of signaling structures $(\mathcal{X}, \kappa)$:

[3]That is, $\kappa : \mathcal{S} \times \mathcal{B}(\mathcal{X}) \to [0, 1]$ 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 $\mathcal{X}$.

– Sender can fully reveal the state. The signaling structure can be represented by letting $\mathcal{X} = \mathcal{S}$ and $\kappa(s, \cdot) = \delta_s$, where δ_s is Dirac's atomic measure on s.

– Sender can reveal whether the state is above or below a threshold π. This can be represented by letting $\mathcal{X} = \{above, below\}$. For any $s \geqslant \pi$, $\kappa(s, \cdot) = \delta_{above}$. For any $s < \pi$, $\kappa(s, \cdot) = \delta_{below}$.

– More generally, for every arbitrary set B of states, Sender can reveal whether the state is in B or not.

– Lastly, consider an example in which Sender randomizes. Let $\mathcal{X} = \{above, below, null\}$ and let π be a threshold. For any $s \geqslant \pi$, $\kappa(s, \cdot) = 1/2\delta_{above} + 1/2\delta_{null}$. This means that when $s \geqslant \pi$, then with probability 1/2 Sender says "above" and with probability 1/2 Sender says "null." For any $s < \pi$, $\kappa(s, \cdot) = 1/2\delta_{below} + 1/2\delta_{null}$. Under this mechanism, Sender sometimes reveals whether the state is above or below π and sometimes reveals nothing.

Once Receiver observes the signal x, the recommendation function $r(x, \cdot)$ contains no further information about the state. We could also define a disclosure mechanism to be a signaling structure $(\mathcal{X}, \kappa)$. However, we choose to include the recommendation function in our definition of a disclosure mechanism, because it allows us to succinctly define incentive compatibility and the Sender's problem.

3.2 Incentive compatibility

Each type observes the signal and decides whether to accept or not. Incentive compatibility means that, after observing the signal, each type will follow the mechanism's recommendation. We now proceed to formally define incentive compatibility.

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 *incentive-compatible (IC)* if, for every type t,

$$\sigma_t^* \in \arg\max \int f(s, t) u(s) \left(\int \sigma(x) \ \kappa(s, \mathrm{d}x) \right) \ \mu(\mathrm{d}s), \tag{1}$$

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

For an 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 utility conditional on the signal is positive, and to reject if it is negative.[4]

3.3 Sender's optimal mechanism

The *Sender's problem* is:

$$\text{Maximize} \iint f(s,t) \left(\int r(x,t)\ \kappa(s,\mathrm{d}x) \right)\ \mu(\mathrm{d}s)\lambda(\mathrm{d}t) \tag{2}$$

among all 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 (1) and (2). 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 theorem.

If Receiver has no private information (that is, if $\mathcal{T}$ is a singleton), then the optimal mechanism reveals whether or not the state is above some $\underline{\pi}$ and recommends accepting only when the state is above $\underline{\pi}$. The set of states in which Receiver accepts is an interval of the form $[\underline{\pi}, \infty)$ which is not bounded from above. The threshold $\underline{\pi}$ is chosen such that Receiver's expected utility from accepting is zero.

In this section we solve for the optimal disclosure when Receiver has private information about the state. As we shall see, each type t still accepts on an interval of states and the interval expands as t increases. However, the acceptance intervals are typically bounded from above. While the lowest type would get a zero utility, higher types receive some information rent. We formally define these properties of the optimal mechanism here, and provide the proof in section 7.

We say that the mechanism is a *cutoff mechanism* if

(i) the signal space is the type space plus infinity, i.e., $\mathcal{X} = \mathcal{T} \cup \{\infty\}$;

(ii) the mechanism recommends that type t accept if and only if $t \geqslant x$, i.e., $r(x,t) = 1$ if and only if $t \geqslant x$.

[4]We say "almost determined" because of the possible indifference and because the conditional expected utility 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.

Thus, a cutoff mechanism announces a cutoff type x and recommends that type x and all higher types accept. If the mechanism announces infinity to be the cutoff type, then it recommends that all types reject. A *deterministic cutoff mechanism* is such that $\kappa(s, \cdot)$ is Dirac's measure on $z(s)$ for some function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$. Hence, a deterministic cutoff mechanism performs no randomization and, when the state is s, announces a cutoff type $z(s)$.

Under a deterministic cutoff mechanism, the *acceptance set* of type t (that is, the set of states in which type t accepts) is $\{s : t \geqslant z(s)\}$. By the definition of a cutoff mechanism, whenever the mechanism recommends that type t accept, it also recommends that any higher type accept. Therefore, for a deterministic cutoff mechanism, type t's acceptance set is a subset of a higher type's acceptance set. Lastly, we say that the mechanism *recommends accepting on intervals* if the acceptance sets are intervals for every type.

Figure 2 illustrates two deterministic cutoff mechanisms for our theorist-chair example. The y-axis corresponds to the signal space $\mathcal{X} = \{L, H\} \cup \{\infty\}$. In both mechanisms, the solid curve illustrates the cutoff function $z(s)$. The dashed line illustrates type H's acceptance set, and the dotted line gives type L's acceptance set. The mechanism on the right-hand side recommends accepting on intervals, since both types' acceptance sets are intervals. On the left-hand side, neither type accepts on an interval.

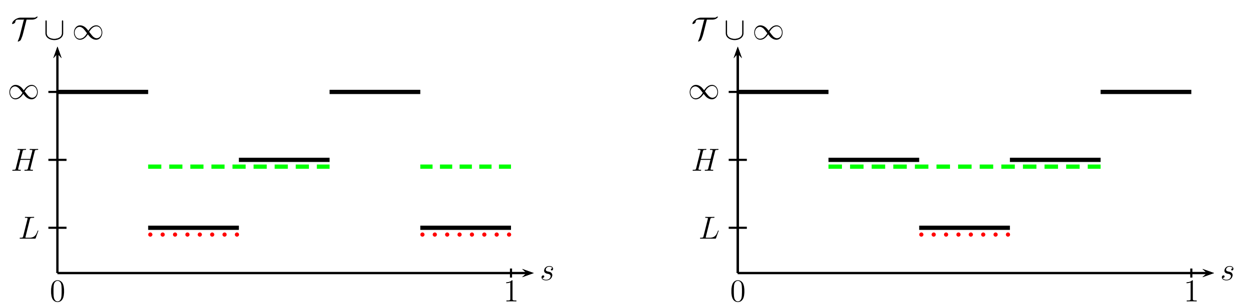

Figure 2: Examples of deterministic cutoff mechanisms

The following theorem is the main result of the paper. The theorem states that Sender's optimal mechanism is a deterministic cutoff mechanism that recommends accepting on intervals. As a result, we can formulate the Sender's problem in terms of the endpoints of each type's acceptance interval.

Theorem 3.1. *Under assumption 1, the optimal IC mechanism is a deterministic cutoff mechanism that recommends accepting on intervals. The acceptance intervals* $[\underline{\pi}(t), \overline{\pi}(t)]$ *are* *the solution for the following optimization problem:*

$$\begin{aligned} &\textit{Maximize} \quad \int W\left(\underline{\pi}(t), \overline{\pi}(t), t\right) \ \lambda(\mathrm{d}t) \\ &\textit{subject to} \quad U\left(\underline{\pi}(t), \overline{\pi}(t), t\right) \geqslant U\left(\underline{\pi}(t'), \overline{\pi}(t'), t\right) \ \textit{for all } t' \leqslant t, \\ &\qquad\qquad\quad U\left(\underline{\pi}(\underline{t}), \overline{\pi}(\underline{t}), \underline{t}\right) \geqslant 0, \end{aligned} \tag{3}$$

over all functions $\underline{\pi}, \overline{\pi} : \mathcal{T} \to \mathbf{R}$ *such that* $\underline{\pi}$ *is monotone decreasing,* $\overline{\pi}$ *is monotone increasing, and* $\underline{\pi}\left(\underline{t}\right) \leqslant \overline{\pi}\left(\underline{t}\right)$. *The functions* $W, U : \mathbf{R}^3 \to \mathbf{R}$ *are given by*

$$W\left(\underline{q}, \overline{q}, t\right) = \int_{\underline{q}}^{\overline{q}} f(s,t) \ \mu(\mathrm{d}s), \ \textit{and} \ U\left(\underline{q}, \overline{q}, t\right) = \int_{\underline{q}}^{\overline{q}} u(s) f(s,t) \ \mu(\mathrm{d}s).$$

It is a standard argument that under assumption 1 every IC mechanism is essentially a cutoff mechanism. We can replace each signal with the lowest type that is still willing to accept given that signal. Due to the i.m.l. ratio assumption, all higher types are willing to accept. (We do not formalize and prove this assertion because we do not need it. We say "essentially" because of the possible indifference and the intricacy on zero-probability events.) The contribution of theorem 3.1 is that the optimal mechanism has the additional property of recommending that each type accept on an interval. Moreover, when choosing the endpoints of each type's interval, Sender needs to ensure only that (i) each type's utility from accepting on his interval is weakly higher than that from accepting on a lower type's interval, and that (ii) the lowest type's utility from accepting is weakly positive.

The fact that the optimal mechanism does not randomize follows from the assumption of a nonatomic state space; we show in section 6.1 how to relax this assumption. The i.m.l. ratio part of assumption 1 is essential for theorem 3.1. (See the counterexample in section 6.4.) However, the assumption that $f(s,t)$ is continuous in t is not essential. Without it, we only need to modify the definition of a cutoff mechanism: in addition to announcing the cutoff type, the mechanism needs to announce whether the cutoff type is supposed to accept or reject.

3.4 Optimal disclosure and the monopoly screening problem

In this section we use theorem 3.1 to establish a connection between the Sender's problem and the (monopoly) screening problem, as in Mussa and Rosen (1978). For ease of exposition, we assume that the state space $\mathcal{S}$ is the real line and that $u(0) = 0$. We refer to states $s \geqslant 0$ as "positive states" and states $s \leqslant 0$ as "negative states," so that Receiver gains positive utility from accepting on positive states and negative utility from accepting on negative states.

From theorem 3.1 we know that under the optimal mechanism each type accepts on an interval of states. When Receiver of type t accepts on the interval $\left[\underline{q}, \overline{q}\right]$ with $\underline{q} < 0 < \overline{q}$, his (normalized) utility is given by:

$$U\left(\underline{q}, \overline{q}, t\right) = \overline{U}\left(\overline{q}, t\right) + \underline{U}\left(\underline{q}, t\right).$$

Here $\overline{U}\left(\overline{q}, t\right) = \int_0^{\overline{q}} f(s,t)u(s)\,\mathrm{d}s$ and $\underline{U}\left(\overline{q}, t\right) = \int_{\underline{q}}^0 f(s,t)u(s)\,\mathrm{d}s$ are, respectively, the utility from accepting on the positive states in $[0, \overline{q}]$ and the disutility from accepting on the negative states in $\left[\underline{q}, 0\right]$. Therefore, we can think of Receiver of type t as a buyer whose taste for quality is measured by his private type t. A type t buyer gets utility $\overline{U}(\overline{q}, t)$ from consuming a good of quality $\overline{q}$ and gets disutility $\underline{U}\left(\underline{q}, t\right)$ from paying price $-\underline{q}$. (In the screening problem, the disutility $\underline{U}\left(\underline{q}, t\right)$ from paying $-\underline{q}$ is usually assumed to be $-\underline{q}$.)

Consider now this screening problem: a seller offers the buyer of type t the quality $\overline{\pi}(t)$ and charges the price $-\underline{\pi}(t)$. The constraints of the Sender's problem given in (3) reflect the familiar "downward" IC constraints such that no type t would mimic a lower type t' in the screening problem. Thus, the Sender's problem in our setup has the same variables and the same constraints as the screening problem. However, the objective function in the Sender's problem is somewhat different from that in the monopoly screening problem: in our disclosure problem, Sender gets a payoff when Receiver accepts in positive states, while in the screening problem the seller does not benefit directly from the very fact that the buyer consumes a higher quality. Nonetheless, the mathematical essence of the two problems is still the same, since the objective functions have similar forms. As in the case of the screening problem, this is a constraint optimization problem for a discrete-type space and a control problem for a continuous-type space. Under some regularity assumptions, the monotonicity constraints do not bind, so we can continue without resorting to the control theory. In section 5, we show how to use standard techniques from the screening problem to solve our disclosure problem with a continuous-type space.

We emphasize that, while the connection between the Sender's problem and the monopoly screening problem is intuitive, it relies on two assertions in theorem 3.1 that are not immediate. First, the theorem asserts that the optimal mechanism recommends each type to accept on an interval, which allows us to interpret the endpoints of the intervals as quality and price. Second, the theorem asserts that the IC constraints of the monopoly screening problem are sufficient for the Sender's problem, even though in our environment Receiver has many more deviation strategies other than just following the mechanism's recommendation for some lower type. Given a cutoff mechanism, each type t can choose an arbitrary set of announced cutoff types after which he accepts.

4 The binary-type case

In this section, we fully characterize the optimal mechanism for the binary-type case. The analysis for any finite type space is similar. We show that theorem 3.1 reduces the Sender's problem to a finite-dimensional constrained-maximization problem and thus allows us to provide the sufficient and necessary condition for pooling to be optimal. We then explain why a solution to the Sender's problem, even in the simplest case of a binary-type space, is outside the scope of previous papers.

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

Theorem 3.1 shows that the optimal mechanism takes the form of nested intervals. We choose the endpoints $\underline{\pi}(L), \overline{\pi}(L), \underline{\pi}(H)$, and $\overline{\pi}(H)$ of the intervals such that type L accepts when $s \in [\underline{\pi}(L), \overline{\pi}(L)]$ and type H accepts when $s \in [\underline{\pi}(H), \overline{\pi}(H)]$. The endpoints are the solution to the following problem, in which we have one incentive constraint for each type:

$$
\begin{aligned}
\underset{\underline{\pi}(H),\underline{\pi}(L),\overline{\pi}(L),\overline{\pi}(H)}{\text{Maximize}} \quad & \int_{\underline{\pi}(L)}^{\overline{\pi}(L)} f(s, L)\,\mathrm{d}s + \int_{\underline{\pi}(H)}^{\overline{\pi}(H)} f(s, H)\,\mathrm{d}s \\
\text{subject to} \quad & 0 \leqslant \underline{\pi}(H) \leqslant \underline{\pi}(L) \leqslant \zeta \leqslant \overline{\pi}(L) \leqslant \overline{\pi}(H) \leqslant 1, \\
& \int_{\underline{\pi}(H)}^{\underline{\pi}(L)} f(s, H)u(s)\,\mathrm{d}s + \int_{\overline{\pi}(L)}^{\overline{\pi}(H)} f(s, H)u(s)\,\mathrm{d}s \geqslant 0, \\
& \int_{\underline{\pi}(L)}^{\overline{\pi}(L)} f(s, L)u(s)\,\mathrm{d}s \geqslant 0.
\end{aligned} \tag{4}
$$

Therefore, theorem 3.1 reduces the Sender's problem to a finite-dimensional constrained optimization. The structure of the problem allows us to further reduce the number of dimensions to one. First, the optimal solution satisfies $\overline{\pi}(H) = 1$; otherwise, increasing $\overline{\pi}(H)$ would increase Sender's payoff without violating the constraints. Second, type H's constraint must bind; otherwise, increasing $\overline{\pi}(L)$ would increase Sender's payoff without violating the constraints. Lastly, type L's constraint must bind; if not, decreasing $\underline{\pi}(L)$

would benefit Sender while still satisfying the constraints. We are left with one variable and can derive the condition under which Sender pools the two types and the condition under which he offers a separating mechanism.

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

$$\frac{f(1,H)}{f(1,L)} - \frac{f\left(\underline{\pi}^*(L),H\right)}{f\left(\underline{\pi}^*(L),L\right)} < 1 - \frac{u\left(\underline{\pi}^*(L)\right)}{u(1)}, \tag{5}$$

where $\underline{\pi}^*(L)$ is such that type L is indifferent between accepting on $[\underline{\pi}^*(L),1]$ and rejecting, i.e., $\int_{\underline{\pi}^*(L)}^{1} f(s,L)u(s)\,\mathrm{d}s = 0$. If (5) holds, then the mechanism recommends that both types accept on $[\pi^*(L),1]$.

Condition (5) states that Sender does not benefit if he marginally shrinks type L's interval in order to expand type H's interval. Starting from the pooling interval $[\underline{\pi}^*(L),1]$, Sender can replace type L's interval by

$$\left[\underline{\pi}^*(L) + \frac{f(1,L)}{f\left(\underline{\pi}^*(L),L\right)}\frac{u(1)}{-u\left(\underline{\pi}^*(L)\right)}\varepsilon, 1-\varepsilon\right] \quad \text{for small } \varepsilon > 0,$$

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 (4) to show that this condition is also sufficient.

Returning to example 1, corollary 4.1 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. *In example 1, if $\zeta \leqslant \frac{3-\phi}{6}$, then the optimal disclosure mechanism always recommends acceptance. If $\zeta > \frac{3-\phi}{6}$, then 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 $f(s,H), f(s,L)$, and $u(s)$ into condition (5) 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. ☐

4.1 Example revisited: A comparison with the cav-V approach

We now compare our approach with the concavification (cav-V) approach, following Aumann and Maschler (1995) and Kamenica and Gentzkow (2011, Section VI A), by revisiting our example 1 in section 3. We first apply the cav-V approach to this example. For convenience we assume that $\phi = 1$, so $f(s, H) = s$ 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 $\int s(s-\zeta)\,\gamma(\mathrm{d}s) \geqslant 0$, where s is type H's density function and $s - \zeta$ is Receiver's utility. Type L accepts if $\int (1-s)(s-\zeta)\,\gamma(\mathrm{d}s) \geqslant 0$, where $1 - s$ is type L's density function.

Therefore, Sender's payoff from the distribution γ induced by his signal is given by

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

The three regions that define V correspond respectively to the distributions under which both types accept, the distributions under which neither type accepts, and the distributions under which only type H accepts.

Note that, even though Receiver's optimal action in this example depends only on his expectation of the state conditional on 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 expectation $\int s\,\gamma(\mathrm{d}s)$ alone. Thus, the example is not in the class of the persuasion problems analyzed by Gentzkow and Kamenica (2016) and Dworczak and Martini (2017).

The cav-V approach states that the optimal payoff 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:

$$\gamma_0 = \alpha_1\gamma_1 + \alpha_2\gamma_2 + \alpha_3\gamma_3, \tag{6}$$

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 $g_1(\cdot)$, $g_2(\cdot)$, and $g_3(\cdot) \in L^\infty(\gamma_0)$ that correspond respectively to the densities of $\alpha_1\gamma_1$, $\alpha_2\gamma_2$, and $\alpha_3\gamma_3$ with respect to γ_0. (It follows from (6) that γ_i are absolutely continuous with respect to γ_0.) The Sender's problem is:

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

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

5 A continuous-type example

We next illustrate how to use theorem 3.1 and the techniques from the screening problem to solve for the optimal mechanism with a continuous-type example.

The type space is $\mathcal{T} = [0,1]$ and the state space is $\mathcal{S} = [-1,1]$. Receiver prefers to accept if and only if $s \geqslant 0$. We assume that $u(s)$ is constant over negative states, i.e., there exists $\eta > 0$ such that $u(s) = -\eta$ for every $s < 0$. Thus, the parameter η measures Receiver's loss when he accepts in a negative state. 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 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 utility 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 the following simplifying assumption:

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

Let $\overline{U}^*(\overline{q},t) = \int_0^{\overline{q}} f(s,t)u(s)\,\mathrm{d}s / \underline{f}(t)$ be the utility of type t from accepting on positive states $[0,\overline{q}]$, normalized by $\underline{f}(t)$. Type t's (normalized) utility from accepting on $[\underline{q},\overline{q}]$ is $\overline{U}^*(\overline{q},t) + \eta\underline{q}$. By the argument of section 3.4, we can view type t as a buyer with a quasilinear utility function who gets utility $\overline{U}^*(\overline{q}, t) + \eta \underline{q}$ from consuming quality $\overline{q}$ and paying price $-\eta \underline{q}$. 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 the IC constraints hold if and only if $\overline{\pi}$ is monotone-increasing and

$$-\eta \underline{\pi}(t) = \overline{U}^*(\overline{\pi}(t), t) - \int_0^t \overline{U}_2^*(\overline{\pi}(\tau), \tau)\, \mathrm{d}\tau, \tag{7}$$

where $\overline{U}_2^*$ is the derivative of $\overline{U}^*$ with respect to the type. Note that given $\overline{\pi}(t)$, the IC constraints determine the "price" $-\eta \underline{\pi}(t)$ 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 price.) Assumption 2 makes sure that this constraint does not bind.

For every $\overline{q} > 0$, we let $\overline{F}(\overline{q}, t) = \int_0^{\overline{q}} f(s, t)\, \mathrm{d}s$ be Sender's payoff if Receiver accepts in $[0, \overline{q}]$. Sender's payoff is given by

$$\begin{aligned}
\int_0^1 \int_{\underline{\pi}(t)}^{\overline{\pi}(t)} f(s,t) \mathrm{d}s \mathrm{d}t &= \int_0^1 \left(-\underline{f}(t) \underline{\pi}(t) + \int_0^{\overline{\pi}(t)} f(s,t) \mathrm{d}s \right) \mathrm{d}t \\
&= \int_0^1 \left(\frac{\overline{U}^*(\overline{\pi}(t), t) \underline{f}(t) - \overline{U}_2^*(\overline{\pi}(t), t) \int_t^1 \underline{f}(\tau) \mathrm{d}\tau}{\eta} + \overline{F}(\overline{\pi}(t), t) \right) \mathrm{d}t.
\end{aligned} \tag{8}$$

Thus, the Sender's problem in our setup has the same structure as the standard monopoly screening problem with quasilinear utility: we maximize (8) 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 assumptions, the monotonicity constraints do not bind, so we can continue without resorting to the control theory. Lastly, we fully characterize the optimal disclosure for a class of distributions, provide the necessary and sufficient condition for pooling to be optimal, and illustrate how the optimal mechanism is typically implemented with a continuous-type space.

We now assume that the joint density 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.

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

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

Substituting into the Sender's problem (8) and simplifying, we can write Sender's payoff as

$$\int_0^1 \left\{ \frac{2\left(\frac{\phi(2t-1)}{\phi(2t-1)+4} - \log\left(\frac{\phi+4}{\phi(2t-1)+4}\right)\right)}{3\eta}\overline{\pi}(t)^3 + \frac{\eta\phi(2t-1)+1}{\eta(\phi(2t-1)+4)}\overline{\pi}(t)^2 + \frac{2}{\phi(2t-1)+4}\overline{\pi}(t) \right\} \mathrm{d}t. \tag{9}$$

Sender maximizes his 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:

$$\overline{\pi}^*(t) := \frac{2\eta}{\eta\phi(1-2t) - 1 + \sqrt{(\eta\phi(1-2t)+1)^2 + 4\eta(\phi(2t-1)+4)\log\left(\frac{\phi+4}{\phi(2t-1)+4}\right)}}. \tag{10}$$

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 3 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 (9) is maximized at $\overline{\pi}^*(t)$ when $t \leqslant \hat{t}$ and at 1 when $t > \hat{t}$. Given this $\overline{\pi}(t)$, we can derive $\underline{\pi}(t)$ based on the incentive constraints (7). We show in lemma 7.4 (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,*

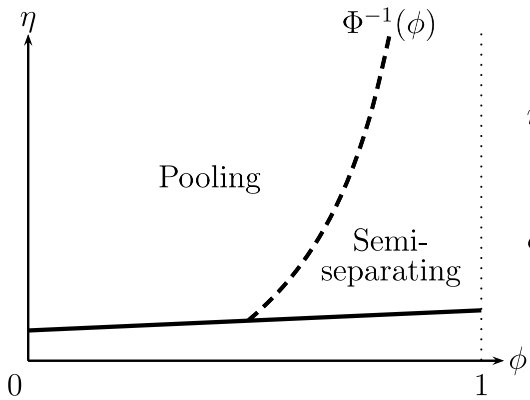

Figure 3: Pooling is optimal when ϕ is small.

(i) $\overline{\pi}(t) = \begin{cases} \overline{\pi}^*(t) & \textit{for } t \leqslant \hat{t} \\ 1 & \textit{for } t > \hat{t}, \end{cases}$ *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* (7)*.*

Figure 4 illustrates the optimal mechanism in which Sender separates the lower types.[5] 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 4 illustrates how to implement the optimal mechanism by examining the case in which η is 1. The solid curve corresponds to the cutoff function $z(s)$. For all the states below $\underline{\pi}(1)$, Sender recommends that all types reject. 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.

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

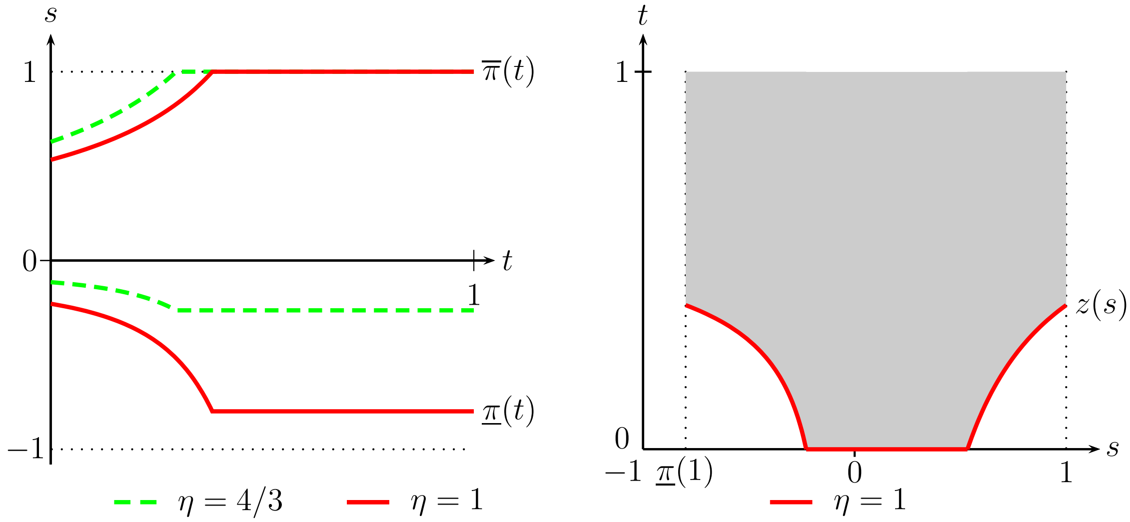

Figure 4: Optimal mechanism and its implementation

6 Discussion

6.1 State space with atoms

In this section we describe the changes that are needed to extend our theorem 3.1 to the case in which the state space $\mathcal{S}$ is a subset of the real line with a supporting measure μ that might include atoms. The main motivation is the atomic case in which $\mathcal{S}$ is a finite subset with μ being the counting measure. The definitions provided earlier of a disclosure mechanism, incentive compatibility, and the Sender's problem carry through to this general framework. The main difference from the nonatomic case is that the optimal mechanism is no longer deterministic, so some randomization is needed. Because of that, we need to modify the definition of what we mean by "accepting on intervals."

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

$$\rho(s,t) = \int r(x,t)\ \kappa(s, \mathrm{d}x). \tag{11}$$

We say that the 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 $s \notin [\underline{\pi}, \overline{\pi}]$. This definition imposes no constraints on the acceptance probabilities at $\underline{\pi}$ and $\overline{\pi}$. For a deterministic mechanism, this definition means that the set of states in which it recommends that type t accept is an interval. For a nondeterministic mechanism, the acceptance probabilities at $\underline{\pi}$ and $\overline{\pi}$ might be strictly between 0 and 1. Lastly, we say that the mechanism *recommends accepting on intervals* if it recommends that each type accept on an interval.

The following example features a finite-state space. We give a mechanism which recommends accepting on intervals.

Example 2. Let $\mathcal{S} = \{-2, -1, 1, 2\}$ with utility $u(s) = s$, and let $\mathcal{T} = \{L, H\}$. The prior over states and types is given by:

	-2	-1	1	2
H	$3/20$	$3/10$	$3/20$	$3/20$
L	$1/12$	$1/12$	$1/24$	$1/24$

Consider a cutoff mechanism which announces the cutoff type L with probability 1 if the state is 1, and with probability 1/2 if the state is -1. The mechanism announces the cutoff type H with probability 1 if the state is 2, and with probability 1/2 if the state is -1 or -2. This cutoff mechanism is IC. The acceptance probabilities $\rho(s, t)$ are given by:

	-2	-1	1	2
$\rho(s, H)$	$1/2$	1	1	1
$\rho(s, L)$	0	$1/2$	1	0

Both types accept on an interval. For each type, the acceptance probability at the lower endpoint of his interval is 1/2.

With these definitions we can modify theorem 3.1 as follows: Under assumption 1, the optimal IC mechanism is a cutoff mechanism that recommends accepting on intervals.

This holds because, by Skorokhod's representation theorem, every state space can naturally be transformed to an interval state space equipped with Lebesgue's measure. 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. The state is s_0 if $s \in [-1, 0]$, and s_1 if $s \in [0, 1]$. Conversely, we 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 $[-1, 1]$. This correspondence preserves the IC properties and Sender's payoff (although it may transform a deterministic mechanism into a nondeterministic mechanism).

6.2 Private incentive compatibility

Up to now we assumed that the information disclosed to all types is the same. We now consider a different environment in which Receiver first reports his type and Sender can disclose different information to different types. Our definition of a disclosure mechanism is amenable to this environment. Specifically, each type can report any type $t' \in \mathcal{T}$; instead of observing the signal x, he observes the recommendation $r(x, t')$ for the reported type.

This environment restricts the set of possible deviation strategies, and gives rise to a weaker notion of incentive compatibility, which we call "private incentive compatibility." Formally, a mechanism is *privately incentive-compatible (or privately IC)* if (1) 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 $\bar{\sigma} : \{0, 1\} \to \{0, 1\}$.

For our purposes, we use the same definition of a disclosure mechanism for (i) the environment in which Sender discloses the same information to all types and (ii) the environment in which Sender discloses different information to different types. We make the distinction between the two environments at the level of incentive compatibility. This approach makes straightforward the logical implication between incentive compatibility and private incentive compatibility.

It is easy to see that every mechanism that is IC is also privately IC. The following example shows that the converse is not true.

Example 3. Let $\mathcal{S} = \{-1000, 1, 10\}$ with utility $u(s) = s$, and let $\mathcal{T} = \{L, H\}$. The prior over states and types 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 that L accept if the state is 1 and reject otherwise. This mechanism is privately IC, but is not IC because if Sender announces the recommendation to both types, then Receiver will want to accept whenever it is recommended that some type should accept.

The mechanism in example 3 is, of course, not optimal. The following corollary shows that Sender does no better under privately IC mechanisms than under IC mechanisms.

Corollary 6.1. *No privately IC mechanism gives a higher payoff to Sender than the optimal IC mechanism.*

It is interesting to compare this result with a result of Kolotilin et al. (2017). In their setup, Receiver has no private information about the state, but privately learns about his threshold for accepting. Receiver's threshold is independent of the state, and 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 privately IC mechanism are implementable by an IC one. This is not true in our private information setup: not every privately IC mechanism in our setup can be duplicated by an IC one. Indeed, under the privately IC mechanism in example 3, type H accepts with the interim probability $1/11$ and type L accepts with the interim probability $20/41$, but there exists no IC mechanism with these interim acceptance probabilities.

6.3 More general setups

We make the following modifications to our model:

– We assume that Receiver's utility from accepting depends on his type. We use $u(s,t)$ to denote type t's utility when he accepts in 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 type t accepts in 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 the Sender's problem in this new environment but stating them should be straightforward. For theorem 3.1, 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 $u(s,t) \geqslant 0$ for $s \geqslant s_0$, and that $u(s,t) \leqslant 0$ for $s \leqslant s_0$. 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 increasing in s with the assumption that u/v is monotone increasing.

6.4 The i.m.l. ratio assumption

The following example shows that without the i.m.l. ratio assumption—ranking of types with respect to the monotone-likelihood-ratio order—theorem 3.1 may not hold. In this example, type H's distribution over states first-order stochastically dominates type L's distribution.

Example 4. Let $\mathcal{S} = \{-1000, -4, -3, 3\}$ with Receiver's utility from accepting given by $u(s) = s$, and let $\mathcal{T} = \{T, B\}$. The density over states and types is given by:

	-1000	-4	-3	3
T	$1/10$	$1/10$	$1/10$	$1/10$
B	$3/10$	$1/30$	$2/9$	$2/45$

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

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

Thus, the optimal IC mechanism does not recommend accepting on intervals.

6.5 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 the Sender's problem. The following example shows that theorem 3.1 may not hold without this assumption.

Example 5. Assume that $\mathcal{S} = \{-2, -1, 1\}$ with $u(s) = s$, and $\mathcal{T} = \{L, H\}$. 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 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. Thus, the optimal IC mechanism does not recommend that H accept on an interval.

7 Proofs

7.1 Proof of theorem 3.1 and corollary 6.1

We assume w.l.o.g. that $u(0) = 0$. Let $\chi(t', t)$ be the utility for type t if he follows the mechanism's recommendation for t':

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

The mechanism is *downward incentive-compatible (or downward IC)* iff (i) every t prefers to follow the recommendation for his type than to follow the recommendation for any lower type; and (ii) the lowest type $\underline{t}$ prefers to follow the recommendation for his type over always rejecting:

$$\chi(t, t) \geqslant \chi(t', t), \text{ for every types } t' \leqslant t \in \mathcal{T}, \text{ and} \tag{13}$$

$$\chi(\underline{t}, \underline{t}) \geqslant 0. \tag{14}$$

Recall that $\rho(s, t)$ is type t's acceptance probability at state s:

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

Let $\nu(t)$ be the normalized probability that type t accepts:

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

We say that a mechanism *weakly dominates* another mechanism if for every t the former has a weakly higher $\nu(t)$. We say that a mechanism *dominates* another mechanism if the former weakly dominates the latter and for some t the former has a strictly higher $\nu(t)$. Lastly, we say a cutoff function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$ is *U-shaped* if there exists some s_0 such that z is monotone-decreasing for $s \leqslant s_0$, and monotone-increasing for $s \geqslant s_0$. A deterministic cutoff mechanism has a U-shaped cutoff function if and only if it recommends accepting on intervals. We also call such a mechanism a U-shaped cutoff mechanism.

Theorem 3.1 is the immediate consequence of the following theorem:

Theorem 7.1. *The optimal downward IC mechanism is a deterministic cutoff mechanism that recommends accepting on intervals. The cutoff function $z(s)$ is increasing on the set*

$\{s : u(s) \geqslant 0\}$ *and decreasing on the set* $\{s : u(s) \leqslant 0\}$*. This optimal mechanism is IC.*

The proof of theorem 7.1 will use lemmata 7.2 and 7.3. Lemma 7.2 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 7.3 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.

Using lemmata 7.2 and 7.3, we prove theorem 7.1 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 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 in which $\mathcal{T}$ is a continuous-type space, 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 lemmata 7.2 and 7.3.

Proof of theorem 7.1. 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 $\iint_{z(s)\leqslant t} f(s,t)\ \mu(\mathrm{d}s)\,\lambda(\mathrm{d}t)$ 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 $\nu_z(t) = \int_{t\geqslant z(s)} f(s,t)\ \mu(\mathrm{d}s)$ is a right-continuous function of t, which follows from the continuity assumption on f. Because ν_z is right-continuous and λ has full support, the maximum z^* cannot be dominated by any $z \in \tilde{Z}$, i.e., if $\nu_z(t) \geqslant \nu_{z^*}(t)$ for every t, then $\nu_z = \nu_{z^*}$.

By lemma 7.2 the cutoff mechanism induced by z^* is optimal for Sender among all downward IC mechanisms. We must still show that z^* is IC. By lemma 7.3 the acceptance set of each type t under the IC mechanism induced by z^* is at least the event $\{s : z^*(s) \leqslant t\}$. Therefore, if the cutoff mechanism z^* were not IC, then it would be dominated by an IC mechanism. By lemma 7.2 again, this 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 $z \in \tilde{Z}$. $\square$

Lemma 7.2. *For every downward IC mechanism there exists a deterministic cutoff mechanism such that it recommends accepting on intervals, is downward IC, and weakly dominates the original mechanism. The cutoff function $z(s)$ is increasing on the set $\{s : u(s) \geqslant 0\}$ and decreasing on the set $\{s : u(s) \leqslant 0\}$.*

The proof of lemma 7.2 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

$$z(s) \leqslant t \longleftrightarrow \underline{\pi}(t) \leqslant s \leqslant \overline{\pi}(t), \tag{15}$$

by $z(s) = \min\{t \colon s \in [\underline{\pi}(t), \overline{\pi}(t)]\}$. The case of a continuous-type space has an additional complication, because in order to obtain a cutoff function z such that (15) holds, we need an additional continuity assumption on $\underline{\pi}(t), \overline{\pi}(t)$.[6] 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.

[6]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.

Proof of lemma 7.2. 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

$$\begin{aligned} \int_0^\infty f(s,t)u(s)\rho(s,t)\,\mu(\mathrm{d}s) &= \int_0^{\overline{p}(t)} f(s,t)u(s)\,\mu(\mathrm{d}s), \text{ and} \\ \int_{-\infty}^0 f(s,t)u(s)\rho(s,t)\,\mu(\mathrm{d}s) &= \int_{\underline{p}(t)}^0 f(s,t)u(s)\,\mu(\mathrm{d}s). \end{aligned} \tag{16}$$

From (16) and the i.m.l. ratio assumption, it follows that:

$$\begin{aligned} \int_0^\infty f(s,t)u(s)\rho(s,t')\,\mu(\mathrm{d}s) &\geqslant \int_0^{\overline{p}(t')} f(s,t)u(s)\,\mu(\mathrm{d}s), \text{ and} \\ \int_{-\infty}^0 f(s,t)u(s)\rho(s,t')\,\mu(\mathrm{d}s) &\geqslant \int_{\underline{p}(t')}^0 f(s,t)u(s)\,\mu(\mathrm{d}s) \end{aligned} \tag{17}$$

for $t' < t$. From (16), (17), and (13), it follows that:

$$\int f(s,t)u(s)\left(\mathbf{1}_{[\underline{p}(t),\overline{p}(t)]} - \mathbf{1}_{[\underline{p}(t'),\overline{p}(t')]}\right)\mu(\mathrm{d}s) \geqslant 0 \tag{18}$$

for every $t' < t$.

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

$$\int_{-\infty}^\infty f(s,t)\rho(s,t)\,\mu(\mathrm{d}s) \leqslant \int_{\underline{p}(t)}^{\overline{p}(t)} f(s,t)\,\mu(\mathrm{d}s) \tag{19}$$

by the Neyman-Pearson Lemma.

Thus, the mechanism with acceptance intervals $[\underline{p}(t), \overline{p}(t)]$ is downward IC (18) and weakly dominates the original mechanism (19). 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 $\underline{\pi}(t) = \inf_{\varepsilon>0} \inf\{\underline{p}(t') \colon t' < t + \varepsilon\}$ be the right-continuous and monotone functions that dominate $\underline{p}$ and $\overline{p}$, respectively, and let z be given by (15):

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

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 $[\underline{p}(t), \overline{p}(t)]$. 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 prefers to accept on his interval $[\underline{\pi}(t), \overline{\pi}(t)]$ over accepting on t''s interval $[\underline{\pi}(t'), \overline{\pi}(t')]$, i.e., that

$$\int f(s,t)u(s)\left(\mathbf{1}_{[\underline{\pi}(t),\overline{\pi}(t)]} - \mathbf{1}_{[\underline{\pi}(t'),\overline{\pi}(t')]}\right)\,\mu(\mathrm{d}s) \geqslant 0. \tag{20}$$

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 (15) it follows that $\{t' \leqslant z(s) < t\} \subseteq \{s \geqslant 0\}$, so (20) 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 (20).

Finally, we need to show that the lowest type gets a utility of at least zero from obeying under the mechanism $(\overline{\pi}, \underline{\pi})$. In the case of a discrete-type space, 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. □

Lemma 7.3. *For every cutoff mechanism κ that is downward IC, the IC mechanism induced by this mechanism has the property that type t accepts when $t \geqslant x$.*

Proof of lemma 7.3. We first argue that, for every cutoff mechanism, if $t'' \leqslant t' \leqslant t$ are types such that type t' prefers to follow the recommendation for his type than to follow the recommendation for t'', then type t prefers to follow the recommendation for t' than to follow the recommendation for t''.

$$\chi(t', t') \geqslant \chi(t'', t') \to \chi(t', t) \geqslant \chi(t'', t), \text{ for every } t'' \leqslant t' \leqslant t. \tag{21}$$

This follows from the definition of i.m.l. ratio.

Fix a type t. We need to show that in the induced 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

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

where the inequality follows from (i) downward incentive-compatibility when $t = t'$ and (ii) (21), which extends to $t \geqslant t'$. The case $B = \{\underline{t}\}$ follows by a similar argument from (14). □

Proof of corollary 6.1. Downward IC is a weaker notion than privately IC in the sense that every mechanism that is privately IC is also downward IC. This corollary follows directly from theorem 7.1. Because the optimal downward IC mechanism is IC, it is also the optimal privately IC mechanism. □

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 (4) 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 (4) 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 (9). (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 (10). 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 7.4. *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 7.4. 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 (10) 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

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

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

$$a_1 x(t,\eta)^3 + a_2 x(t,\eta)^2 + a_3 x(t,\eta) + a_4 > 0, \tag{22}$$

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 (22) is satisfied when $x(t,\eta) = 1$. We next rewrite the left-hand side of (22) 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 (22) 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

$$-\frac{2(\eta\phi(2t-1)(\eta\phi(2t-1)-4)+2)}{\eta(\phi(6t-3)-8)-3} + \log\left(\frac{2\phi t-\phi+4}{\phi+4}\right) > 0. \tag{23}$$

It is easy to verify that the left-hand side of (23) is convex in η. The inequality (23) 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 (23), the inequality is satisfied for any $t \in (0, 1/2)$ and $\phi \in (0, 1)$.

The left-hand side of (23) 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 (23) 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 (23) decreases in η for any $\eta \in (0, \overline{\eta})$. We have shown that the left-hand side of (23) is positive at $\eta = 0$ and $\eta = \overline{\eta}$. Hence, the left-hand side of (23) 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 (23) is achieved when $\eta = \tilde{\eta}$. We need to show that (23) 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 (23), the inequality (23) 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 (23) holds. □

References

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan. 2006. "Signaling in a Global Game: Coordination and Policy Traps." *Journal of Political Economy*, 114(3): 452–484.

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. 2017. "The Design and Price of Information." American Economic Review, forthcoming.

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.

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.

Feltovich, Nick, Richmond Harbaugh, and Ted To. 2002. "Too cool for school? Signalling and countersignalling." *RAND Journal of Economics*, 33(4): 630–649.

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

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

Kolotilin, Anton. 2017. “Optimal Information Disclosure: A Linear Programming Approach.” Theoretical Economics, forthcoming.

Kolotilin, Anton, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnyuk. 2017. “Persuasion of a Privately Informed Receiver.” Econometrica, forthcoming.

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.

Li, Hao, and Xianwen Shi. 2017. “Discriminatory Information Disclosure.” American Economic Review, forthcoming.

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.

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.