The Interval Structure of Optimal Disclosure

Yingni Guo* Eran Shmaya†

Abstract

A fully committed sender persuades a receiver to accept a project through disclosing information regarding a payoff-relevant state. The receiver's payoff from acceptance increases in the 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. Hence, the most pessimistic types reject when the receiver benefits from acceptance the most. This mechanism is optimal even if the sender can condition the disclosure mechanism on the receiver's reported type.

JEL: D81, D82, D83

Keywords: Bayesian persuasion, privately informed receiver, nested intervals

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

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

1 Introduction

$\geqslant$

A sender (she) promotes a project to a receiver (he). The receiver's payoff from accepting this project depends on an unknown state. The sender can design a disclosure mechanism revealing information about the state. This is the Bayesian persuasion problem analyzed by Rayo and Segal (2010) and Kamenica and Gentzkow (2011).

Frequently, besides the sender's disclosure mechanism, the receiver has access to external sources of information regarding the underlying state. The realization of these external sources is the receiver's private information. We are interested in how the sender designs her disclosure mechanism when the receiver is privately informed about the state.

Many applications fit this description. For instance, a theorist tries to persuade the chair to hire a candidate. The chair himself reads the job market paper and acquires a private yet imperfect signal about the quality of the candidate. How shall the theorist make her recommendation to the chair? In a different example, a media outlet tries to promote a candidate. The audience acquires information about the candidate from various channels other than this media outlet. How shall the outlet structure its recommendation?

Environment. Formally, Sender promotes a project to Receiver who decides whether to accept or to reject. If Receiver accepts, his utility increases in the project's quality, i.e. the *state*. If he rejects, his utility is normalized to zero. Therefore, there is a threshold state such that Receiver benefits from the project if the state is above the threshold and loses otherwise. Moreover, the higher the state is, the more enthusiastic Receiver is about accepting the project. The lower the state is, the more detrimental Receiver finds it to accept the project. Sender's payoff, on the other hand, is one if Receiver accepts and zero otherwise. Hence, Sender simply wants the project to be accepted.

Receiver does not know the state but has access to an external information source. We refer to the information from this external source as Receiver's *type*. We assume that the higher the state is, the more likely it is that Receiver is of a higher rather than of a lower type. Therefore, a higher type for Receiver suggests a higher-quality project. We can rank the types in terms of how favorable they are for an acceptance decision: a higher type is more optimistic than a lower type about the state being favorable for an acceptance decision.

Sender designs a disclosure mechanism which consists of a signal space and a mapping from the state space to the set of distributions over the signal space. Sender commits to this disclosure mechanism. Receiver updates his belief about the state based on his private information and Sender's signal. Then, he takes an action, accept or reject. We refer to such a procedure as "public persuasion": different types must observe the same signal. We later address the case of "private persuasion", where Sender can ask Receiver to report his type and she then randomizes a signal from a distribution that depends on the state and the reported type.

Main results. Consider first the case of no private information for Receiver. Roughly speaking, in this case it is well known that the set of states on which Sender recommends acceptance is an interval which is bounded from below but unbounded from above. The lower bound is chosen such that Sender obtains all surplus and Receiver's expected payoff from acceptance is zero. In particular, Receiver approves for sure in those states under which he benefits from an acceptance decision.

Our first main result states that, in the case of private information, each type still approves on an interval of states, and the interval expands as the type increases. However, in this case the more pessimistic types' acceptance intervals are typically bounded from above, so that they receive a rejection recommendation under those states where Receiver benefits from an acceptance the most. This stands in contrast to the solution when Receiver has no private information.

To illustrate the structure and the intuition of the optimal mechanism, we consider the case in which Receiver's type is either high or low. Sender could design a pooling mechanism under which either both types are recommended to accept or both types are recommended to reject. Sender could also design a separating mechanism under which with positive probability only the high type is recommended to accept. When Receiver's external information is sufficiently informative, the high type is much easier to be persuaded than the low type is, so a separating mechanism delivers a higher payoff.

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

The states are divided into "positive states" and "negative states" depending on the sign of Receiver's utility from acceptance. Receiver would like to accept only on positive states, and Sender would like him to accept on both positive and negative states. The set of states on which Sender recommends each type to accept includes some positive and some negative states. The reason that this set is an interval is that, 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.

The reason that the additional states on which only the high type accepts should be the extreme states of his acceptance interval is as follows. When the state is above the threshold state, as the state increases, Receiver becomes more likely to be the high type. Therefore, the loss for Sender from excluding the low type decreases. Moreover, the high type puts more weight on a higher state over a lower state than the low type does. Therefore, when Sender recommends only the high type rather than both types to accept, the benefit for the high type's incentive constraint net of the cost on the low type's incentive constraint grows as the state increases. Therefore, for those states above the threshold state, Sender recommends only the high type to accept when the state is sufficiently high. When the state is below the threshold state, as the state decreases, Sender is more certain that Receiver is the low type. Excluding the low type becomes more costly. However, it also becomes more costly (in terms of the low type's incentive constraint) to include the low type since he puts more weight on a lower state over a higher state than the high type does. On the other hand, it becomes cheaper (in terms of the high type's incentive constraint) to recommend just the high type to accept, since he puts less weight on a lower state over a higher state than a low type does. Therefore, for those states below the threshold state, as the state decreases, Sender recommends only the high type to accept when the state is sufficiently low.

This is not yet a proof since we still need to show that both desirable properties can be achieved simultaneously in a publicly incentive-compatible mechanism. A key feature to prove this is that the two properties we need—that each acceptance set is an interval so the high type's interval contains the low type's interval and that the states on which only the high type accepts are the extreme states in his interval—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.

We generalize this observation to the setting with continuous types. The optimal public mechanism takes the form of nested intervals: (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. The lowest type is recommended to accept for an interval of states concentrated around the threshold state. His interval is bounded from below and typically bounded from above as well. The upper bound of the highest type's interval coincides with the highest possible state.

Our second main result states that this optimal public mechanism is also optimal among private mechanisms. Under private persuasion, Sender can design the disclosure mechanism conditional on the type reported by Receiver. We show that Sender does not benefit from inducing Receiver's private information before designing the disclosure mechanism. The key observation is that the two properties that pertain to public persuasion—that each type accepts on an interval of states and that the most efficient way to deter a higher type from mimicking a lower type is to additionally include extreme states—are desirable in the case of private persuasion as well.

The nested-interval structure gives us an algorithm to find the optimal disclosure. We only need to characterize the upper and lower bound of each type's interval. We can use the techniques from the standard screening problems to solve for the optimal mechanism. In Section 3, we illustrate how to solve for the optimal mechanism via finite-type and continuous-type examples. In Section 4, we show that our results can be generalized to broader payoff settings.

Related literature. Our paper is related to the literature on persuasion. Rayo and Segal (2010) and Kamenica and Gentzkow (2011) study optimal persuasion between a sender and a receiver.[1] The receiver has no external source of information. We study the information design problem in which the receiver has an external source of information. Kamenica and Gentzkow (2011) extends their baseline analysis to situations where the receiver has private information. They show that the geometric method can be readily generalized to the public persuasion setting. However, it is difficult to characterize the optimal mechanism using the geometric method. Our framework follows Kamenica and Gentzkow (2011), with the additional structure over the state space and the receiver's private information that enables us to derive the interval structure of the optimal mechanism, and to prove that focusing on public mechanism is without loss at the optimum.

Kolotilin (2016) also examines the optimal public persuasion when the receiver has private information. He uses a linear programming approach and provides conditions under which full and no revelation are optimal. Alonso and Câmara (2016*a*) examine the optimal persuasion when the sender's prior differs from the receiver's. In their benchmark model,

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

both priors are common knowledge so there is no private information. Hence, the analysis and main results are quite different from ours. They extend the analysis to the case where the sender is uncertain about the receiver's prior and characterize the conditions under which the sender benefits from persuasion.

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

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

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

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

2 Environment and main results

Let $\mathcal{S}$ be a set of *states*, $\mathcal{T}$ a set of *types* equipped with σ-finite measures μ, λ and consider a distribution over $\mathcal{S} \times \mathcal{T}$ with density f w.r.t. $\mu \times \lambda$.[2] In all our examples these spaces will either be discrete spaces or subsets of the real line equipped with Lebesgue's measure. Let $u\colon \mathcal{S} \to \mathbf{R}$ be a bounded Borel function representing *Receiver's utility* from accepting. Receiver's utility from rejecting is 0 at every state. Note that we do not assume that u is one to one or continuous. Thus, the state in our framework represents not just Receiver's payoff but also Sender's belief about Receiver's type. It would perhaps be better to use the term *Sender's type* instead of state, but we keep the term state in reverence to previous literature.

A *disclosure mechanism* is given by a triple $(\mathcal{X}, \kappa, r)$ where $\mathcal{X}$ is a set of *signals*, κ is a Markov kernel[3] from $\mathcal{S}$ to $\mathcal{X}$ and $r : \mathcal{X} \times \mathcal{T} \to \{0, 1\}$ is a *recommendation function*: When the state is s, the mechanism randomizes a signal x according to $\kappa(s)$ and recommends that type t accepts if and only if $r(x, t) = 1$.

Remark 1. Alternatively we could define a direct mechanism as a function from states to distributions over sets of types. However, doing so would require us to define a Borel structure over the powerset of the set of types.

If the mechanism is public then Receiver of any type t observes the signal x produced by the mechanism and chooses an action. A *strategy* for type t is given by $\sigma : \mathcal{X} \to \{0, 1\}$. Let $\sigma_t^* = r(\cdot, t)$ be the strategy that follows the mechanism's recommendation for type t. We say that the mechanism is *publicly incentive-compatible* if

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

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

For a publicly incentive-compatible mechanism the recommendation function r is almost determined by the signaling structure $(\mathcal{X}, \kappa)$: Receiver t is recommended to accept if his conditional expectation given the signal is positive and reject if it is negative. We say "almost determined" because of possible indifference and because the conditional expectation is defined up to a probability 0 event. We omit the formal statement of this assertion, in order not to get into the technical intricacies involving conditional distributions.

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

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

2.1 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 publicly incentive-compatible mechanisms. Note that we assume 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 also for the case that the density functions of Sender and Receiver are different but we need the common prior assumption for our results.

From now on we assume that $\mathcal{S}, \mathcal{T} \subset \mathbf{R}$ and that u is monotone increasing.

If Receiver has no private information (that is, if $\mathcal{T}$ is a singleton) then, up to some possible nuisance if the state space has atoms, the optimal mechanism is deterministic (that is, it does not randomize) and the set of states on which the mechanism recommends to accept is an interval of the form $[\pi, \infty)$. The threshold π is chosen such that Sender gets all surplus and Receiver's expected utility from accepting is 0. [4] In this section we provide a generalization of this observation for the case of private information. As we shall see, in this case each type t still accepts on an interval of states, but the intervals are typically bounded and expand as t increases. While the lowest type would get payoff 0, higher types receive some information rent.

We need the following additional assumption.

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

Let us say that the mechanism is a *cutoff mechanism* if $\mathcal{X} = \mathcal{T} \cup \{\infty\}$ and the recommendation function is given by $r(x,t) = 1 \leftrightarrow t \geqslant x$. That is, a cutoff mechanism announces a type x and recommends all higher types to accept. A *deterministic* cutoff mechanism is such that $\kappa(s)$ is Dirac's measure on $z(s)$ for some function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$. It is a standard argument that under Assumption 1 every publicly incentive-compatible mechanism is essentially a cutoff mechanism.[5] (We do not formalize and prove the last assertion because

[4] If the state space has an atom at π then the optimal mechanism randomizes the recommendation when the state is π.

[5] For any realized signal x, a low type is more pessimistic about the state distribution than a high type is. If a low type accepts after some x, any higher type prefers to accept as well. Each x effectively specifies

we do not need it. We say "essentially" because the mechanism can do weird stuff when Receiver is indifferent and on a probability 0 event.)

Our structural theorem says that in addition to being a cutoff mechanism, Sender's optimal mechanism has the property that the set of states on which each type accepts is an interval. For a deterministic cutoff mechanism this means that the cutoff function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$ is U-shaped. Recall that z is *U-shaped* if there exists some s_0 such that z is monotone decreasing for $s \leqslant s_0$ and monotone increasing for $s \geqslant s_0$. Equivalently, z is U-shaped if the set $\{s : z(s) \leqslant t\}$ is an interval for every t. For the case of a non-deterministic cutoff randomization, let us say that a Markov kernel κ from $\mathcal{S}$ to $\mathcal{T} \cup \{\infty\}$ is U-shaped if there exists some $s_0 \in \mathcal{S}$ such that for every $s' \leqslant s \leqslant s_0$ and every $s_0 \leqslant s \leqslant s'$, every $t \in \text{support}(\kappa(s))$ and $t' \in \text{support}(\kappa(s'))$ satisfy $t \leqslant t'$.

Theorem 2.1. *Under Assumption 1, the optimal mechanism is a cutoff mechanism with a U-shaped cutoff function. If μ is nonatomic then the optimal mechanism is also deterministic and the cutoff function is increasing on the set $\{s : u(s) \geqslant 0\}$ and decreasing on the set $\{s : u(s) \leqslant 0\}$.*

The main step in proving Theorem 2.1 is proving that the set of states in which each type accepts is an interval. The fact that the mechanism is a cutoff mechanism, which holds for every publicly incentive-compatible mechanism, implies that these intervals expand as t increases.

The i.m.l.r. part of Assumption 1 is essential for Theorem 2.1. The continuity part is not. Without it, we would have to modify the definition of a cutoff mechanism: in addition to announcing the cutoff type the mechanism would have to announce whether the cutoff type is supposed to accept or reject. Types above the cutoff accept and types below the cutoff reject.

2.2 Private mechanisms

If the mechanism is private, then Receiver does not observe the realized signal x, but can only report some type t' and observe $r(x, t')$. This restricts the set of possible deviation strategies. We say that the mechanism is *privately incentive-compatible* if (1) holds for every type t where the argmax ranges over all strategies σ of the form $\sigma = \bar{\sigma}(r(x, t'))$ for some type $t' \in \mathcal{T}$ and some $\bar{\sigma} : \{0, 1\} \to \{0, 1\}$. Clearly every mechanism that is publicly incentive-compatible is

a cutoff type such that types above this cutoff type accept while types below this cutoff type reject. If the type distribution admits atoms, the realized signal also specifies what the cutoff type should do.

also privately incentive-compatible, so the optimal privately incentive-compatible mechanism gives Sender a weakly higher payoff than the optimal publicly incentive-compatible one. Our next result shows that Sender does not do better under privately incentive-compatible mechanisms.

Theorem 2.2. *Under Assumption 1, no privately incentive-compatible mechanism can give a higher payoff to Sender than the optimal publicly incentive-compatible mechanism.*

For a mechanism $(\mathcal{X}, \kappa, r)$, if Receiver of type t follows the recommendation, his *individual acceptance probability at state s* is given by

$$\rho(s,t) = \int r(x,t)\ \kappa(s, \mathrm{d}x). \tag{3}$$

For every function $\rho : \mathcal{S} \times \mathcal{T} \to [0,1]$ there exists a mechanism whose individual acceptance probabilities are given by ρ. Indeed we can choose $\mathcal{X} = [0,1]$ with $\kappa(s) = \text{Uniform}(0,1)$ for every s and $r(x,t) = 1$ if and only if $x < \rho(s,t)$. When discussing properties of a mechanism that depends only on the individual acceptance probabilities given by the function ρ, we sometimes abuse terminology by referring to ρ as the mechanism. One such property is private incentive-compatibility. Indeed, in some papers (e.g., Kolotilin et al. (2016)) a private mechanism is defined by the function ρ. Our definitions have the advantages that the definition of a mechanism is the same for both the public and private cases and that they render the logical implication between private and public incentive-compatibility straightforward.

2.3 Downward incentive-compatibility

In the theory of mechanism design, it is a typical situation that the only binding incentive compatible conditions are the downward conditions, that type t will not mimic a lower type t'. Moreover, for the case of discrete type space, the adjacent downward conditions are sufficient. In this section we establish the corresponding results in our setup.

Fix a mechanism $(\mathcal{X}, \kappa, r)$ and let

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

be the utility for Receiver t from mimicking Receiver t'. Let us say that the mechanism is *downward incentive-compatible* if

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

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

The first condition says that Receiver t prefers to follow the mechanism recommendation to him over following the recommendation to Receiver t'. The second condition says that the lowest type $\underline{t}$ prefers to follow the mechanism's recommendation to him over always rejecting. Clearly, any publicly or privately incentive-compatible mechanism is downward incentive-compatible.

The following lemma, which follows from the definition of i.m.l.r., is familiar from the theory of mechanism design: If $t'' \leqslant t' \leqslant t$ are types such that type t' does not want to mimic type t'', then type t prefers to mimic t' over mimicking t''.

Lemma 2.3. *Under Assumption 1, every cutoff mechanism has the following property:*

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

Corollary 2.4. *If* $\mathcal{T} = \{\underline{t} = t_0 < t_1 < \cdots < t_n\}$ *is discrete, then in condition* (4) *in the definition of downward incentive-compatibility it is sufficient to consider the case where* $t = t_k$ *and* $t' = t_{k-1}$ *for every* $1 \leqslant k \leqslant n$.

The following lemma establishes the fact that when looking for the optimal mechanism it is sufficient to consider the downward incentive-compatible conditions.

Lemma 2.5. *For every cutoff mechanism* κ *that is downward incentive-compatible, the public mechanism induced by this mechanism has the property that type* t *accepts when* $t \geqslant x$.

The lemma asserts that when the cutoff declared by the mechanism is x, all types which are supposed to accept (i.e., types t such that $t \geqslant x$) will accept. Possibly lower types will also accept.

Proof of Lemma 2.5. Fix a type t. We need to show that in the induced mechanism Receiver 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 $B = \{\underline{t}\}$ as these sets generate the Borel sets. Indeed, for

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

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

where the inequality follows from downward incentive-compatibility when $t = t'$ and from Lemma 2.3 extends to $t \geqslant t'$. The case $B = \{\underline{\mathrm{t}}\}$ follows by a similar argument from (5) $\square$

From Lemma 2.5 above, it follows that when searching for the optimal publicly incentive-compatible mechanism we can do the optimization over the set of downward incentive-compatible cutoff mechanisms. This set is still not a very pleasant set to work with (and it is of infinite dimension when $\mathcal{S}$ is infinite), but from Theorem 2.1 we also know that we can restrict attention to U-shaped cutoff mechanisms.

3 Examples

3.1 The binary-type case

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

$$f(s,H) = 1/2 + \phi(s - 1/2), \text{ and } f(s,L) = 1/2 - \phi(s - 1/2) \text{ for } 0 \leqslant s \leqslant 1. \tag{6}$$

Type L prefers to accept absent further information if $\zeta \leqslant \frac{3-\phi}{6}$. We thus assume that the opposite is the case.

Our theorem 2.1 shows that the optimal mechanism takes the form of nested intervals. It can be described by four thresholds $0 \leqslant \underline{\pi}_H \leqslant \underline{\pi}_L \leqslant \zeta \leqslant \overline{\pi}_L \leqslant \overline{\pi}_H \leqslant 1$ such that type L accepts when $s \in [\underline{\pi}_L, \overline{\pi}_L]$ and type H accepts when $s \in [\underline{\pi}_H, \overline{\pi}_H]$. We write the optimization problem that maximizes Sender's payoff over the downward incentive-compatible U-shaped mechanisms. The variables are the endpoints of the acceptance intervals, and by Corollary

2.4 we have one incentive constraint for every type:

$$\begin{aligned}
\underset{\underline{\pi}_H,\underline{\pi}_L,\overline{\pi}_L,\overline{\pi}_H}{\text{Maximize}} \quad & \int_{\underline{\pi}_L}^{\overline{\pi}_L} f(s,L)\,\mathrm{d}s + \int_{\underline{\pi}_H}^{\overline{\pi}_H} f(s,H)\,\mathrm{d}s \\
\text{subject to} \quad & 0 \leqslant \underline{\pi}_H \leqslant \underline{\pi}_L \leqslant \zeta \leqslant \overline{\pi}_L \leqslant \overline{\pi}_H \leqslant 1, \\
& \int_{\underline{\pi}_L}^{\overline{\pi}_L} f(s,L)u(s)\,\mathrm{d}s \geqslant 0, \\
& \int_{\underline{\pi}_H}^{\underline{\pi}_L} f(s,H)u(s)\,\mathrm{d}s + \int_{\overline{\pi}_L}^{\overline{\pi}_H} f(s,H)u(s)\,\mathrm{d}s \geqslant 0,
\end{aligned} \tag{7}$$

where $f(s,L)$ and $f(s,H)$ are given by (6). Hence, our result reduces Sender's problem to a finite dimensional constrained optimization. In the case of our example the problem can be solved explicitly. We prefer to provide a qualitative description of the solution. It is clear that the optimal solution satisfies $\overline{\pi}_H = 1$ (otherwise increasing $\overline{\pi}_H$ would increase Sender's utility without violating the constraints) and similarly $\overline{\pi}_L < \overline{\pi}_H \leftrightarrow \underline{\pi}_H < \underline{\pi}_L$. Thus, there are two possibilities for the optimal solution: If $\overline{\pi}_H = \overline{\pi}_L = 1$, Sender pools the two types. She recommends both to accept if $s \in [\underline{\pi}_L, 1]$ and both to reject otherwise. If $\underline{\pi}_H < \underline{\pi}_L \leqslant \overline{\pi}_L < \overline{\pi}_H = 1$, Sender offers a separating mechanism. The following proposition shows 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 that the separating mechanism gives Sender a higher payoff.

Proposition 3.1. *There exists an increasing function $\Phi(\cdot)$ such that the optimal mechanism is separating if $\phi > \Phi(\zeta)$ and is pooling if $\phi < \Phi(\zeta)$.*

Figure 1 shows the structure of the optimal mechanism when separating is optimal. [6] In terms of implementation, Sender needs three different messages in her message space. We illustrate the implementation in the context of the communication between the theorist and the chair. If the state is in the green region, the theorist would claim that the candidate is a solid one and recommend that an offer is made regardless of the chair's private information. If the state is in the red region, the theorist would claim that the candidate is creative but also risky and recommend the chair to make the offer if her type is high. For the remaining states, the theorist recommends rejection. The chair is always willing to obey the theorist's recommendation. Moreover, the theorist does not extract all the surplus since the chair strictly prefers to accept when his type is high and the theorist reveals that the candidate

[6] The parameter values are $\phi = 3/5$ and $\zeta = 7/10$.

is solid (i.e., the state is in the green region). In other words, while the low type is always indifferent after an acceptance recommendation, the high type's interim payoff is strictly positive.

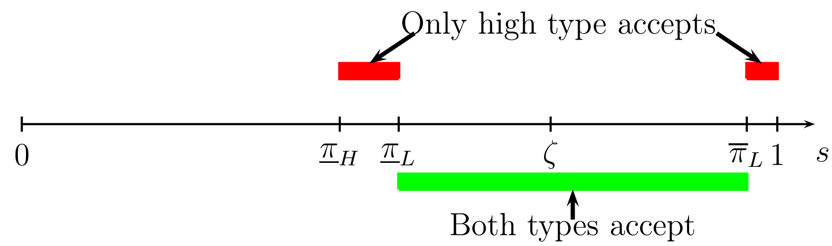

Figure 1: Optimal mechanism for binary types

3.2 Example revisited: a comparison with cav-V approach

We compare our approach and the general concavification (cav-V) approach, following Aumann and Maschler (1995) and Kamenica and Gentzkow (2011, Section VI A), by revisiting the example from Section 3.1. We first apply the general cav-V approach to this setting. For convenience we assume $\phi = 1$. 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. Receiver of type H accepts if $\int s(s-\zeta)\,\gamma(\mathrm{d}s) \geqslant 0$. Receiver of type L accepts if $\int(1-s)(s-\zeta)\,\gamma(\mathrm{d}s) \geqslant 0$. Therefore Sender's payoff from the belief γ induced by her signal is given by

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

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

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

where the distributions $\gamma_1, \gamma_2, \gamma_3$ belong to three regions that define V. Finding the optimal combination amounts to solving the following infinite dimensional LP problem with variables

$g_1, g_2, g_3 \in L^\infty(\gamma_0)$ that correspond to the densities of $\alpha_1\gamma_1, \alpha_2\gamma_2, \alpha_3\gamma_3$ w.r.t. γ_0. (It follows from (8) that γ_i are absolutiley continuous w.r.t. γ_0.)

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

In contrast, as shown in Section 3.1, our approach reduces Sender's problem to a finite dimensional constrained optimization, which is much simpler than the infinite dimensional LP problem above.

3.3 A continuous-type example

We next illustrate the main results by a continuous-type example.

The type space is $[0,1]$ and the state space $[-1/2, 1/2]$. The (s,t) distribution is

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

The distribution $f(s,t)$ is piecewise linear in s with a kink at $s = 0$. 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 is. Receiver's utility from accepting is $u(s) = -\eta < 0$ if $s < 0$ and is $u(s) = s$ if $s \geqslant 0$. Receiver prefers to accept if and only if the state is positive. Here, the parameter $\eta > 0$ measures how costly it is for Receiver to accept when the state is negative. The lowest type's payoff if he accepts under his prior belief is

$$\int_{-1/2}^{0} f(s,0)(-\eta)\mathrm{d}s + \int_0^{1/2} f(s,0) s \mathrm{d}s = \frac{3 - 12\eta - 2\phi}{24 - 6\phi}.$$

Sender's problem is trivial when the lowest type prefers to accept absent further information, i.e., $\eta \leqslant (3-2\phi)/12$. We thus assume that the opposite is the case:

Assumption 2. *The lowest type rejects in the absence of further information, i.e.* $\eta >$

$(3-2\phi)/12$.

Based on Theorem 2.1, we let $[\underline{\pi}(t), \overline{\pi}(t)]$ denote the interval of states under which type t accepts. The U-shaped structure implies that $\overline{\pi}(t)$ weakly increases in t whereas $\underline{\pi}(t)$ weakly decreases in t. Moreover, $\underline{\pi}(0) \leqslant 0 \leqslant \overline{\pi}(0)$. Type t's payoff if he reports t' and then follows Sender's recommendation is

$$\int_{\underline{\pi}(t')}^{\overline{\pi}(t')} f(s,t)u(s)\mathrm{d}s.$$

Thus, whenever differentiable, $\{\underline{\pi}(t), \overline{\pi}(t)\}$ must satisfy the (differential) incentive constraint:

$$u(\underline{\pi}(t))f(\underline{\pi}(t),t)\underline{\pi}'(t) = u(\overline{\pi}(t))f(\overline{\pi}(t),t)\overline{\pi}'(t).$$

Substituting the distribution $f(s,t)$ and Receiver's utility function into the incentive constraint, we obtain the following condition:

$$-\eta\underline{\pi}'(t) = \overline{\pi}(t)\left(2\phi\overline{\pi}(t)(2t-1)+1\right)\overline{\pi}'(t),$$

which allows us to solve for $\underline{\pi}(t)$ as a function of $\overline{\pi}(t)$ for all $t \in [0,1]$ and $\underline{\pi}(1)$:

$$\underline{\pi}(t) = \underline{\pi}(1) + \int_t^1 \frac{4\phi\overline{\pi}(z)^3}{-3\eta}\mathrm{d}z + \frac{\overline{\pi}(t)^2(4\phi\overline{\pi}(t)(2t-1)+3) - \overline{\pi}(1)^2(4\phi\overline{\pi}(1)+3)}{-6\eta}. \tag{9}$$

Recall that $U(0,0) \geqslant 0$ denote type 0's interim payoff which is given by:

$$U(0,0) = \int_{\underline{\pi}(0)}^0 f(s,0)(-\eta)\mathrm{d}s + \int_0^{\overline{\pi}(0)} f(s,0)s\mathrm{d}s = \frac{-4\eta\underline{\pi}(0)}{\phi-4} + \frac{2\overline{\pi}(0)^2(4\phi\overline{\pi}(0)-3)}{3(\phi-4)}.$$

Combining this equation and (9), we solve for value of $\underline{\pi}(1)$ as a function of $\overline{\pi}(t)$ for all $t \in [0,1]$ and $U(0,0)$:

$$\underline{\pi}(1) = \int_0^1 \frac{4\phi\overline{\pi}(z)^3}{3\eta}\mathrm{d}z + \frac{4\phi\overline{\pi}(1)^3 + 3\overline{\pi}(1)^2}{-6\eta} + \frac{\phi-4}{-4\eta}U(0,0).$$

This allows us to write $\underline{\pi}(t)$ as a function of $\overline{\pi}(t)$ for all $t \in [0,1]$ and $U(0,0)$.

Sender's payoff is given by

$$\int_0^1 \int_{\underline{\pi}(t)}^{\overline{\pi}(t)} f(s,t)\mathrm{d}s\mathrm{d}t.$$

Substituting $\underline{\pi}(t)$ into Sender's payoff and simplifying, we can write Sender's payoff as

$$\int_0^1 \left\{ \frac{8\left(\phi(1-2t)+2(-2\phi t+\phi-4)\tanh^{-1}\left(\frac{\phi(t-1)}{\phi t+4}\right)\right)}{-3\eta(\phi(2t-1)+4)}\overline{\pi}(t)^3 + \right.$$

$$\left.\left(4+\frac{2(1-8\eta)}{\eta(\phi(2t-1)+4)}\right)\overline{\pi}(t)^2+\frac{4}{2\phi t-\phi+4}\overline{\pi}(t)\right\}\mathrm{d}t+\frac{3(4-\phi)\tanh^{-1}\left(\frac{\phi}{4}\right)}{-3\eta\phi}U(0,0). \quad (10)$$

It is easily verified that Sender's payoff decreases in $U(0,0)$ so Sender optimally sets it to be zero. Sender maximizes his payoff by choosing $\overline{\pi}(t)$ to maximize the integrand pointwise. If we ignore the constraint that $\overline{\pi}(t)$ is below $1/2$ (the upper bound of the state space), then the integrand is maximized at:

$$\overline{\pi}^*(t) := \frac{2\eta}{2\eta\phi(1-2t)-1+\sqrt{(2\eta\phi-4\eta\phi t+1)^2+8\eta(2\phi t-\phi+4)\log\left(\frac{\phi+4}{2\phi t-\phi+4}\right)}}. \quad (11)$$

When Receiver's information is not informative enough, the integrand is maximized at the highest state $1/2$ for every t. In this case, the optimal mechanism is pooling.

Proposition 3.2. *There exists an increasing function $\Phi(\cdot)$ so that Sender pools all types if $\phi \leqslant \Phi(\eta)$: Sender sets $\overline{\pi}(t)$ to be $1/2$, and the allocation $\underline{\pi}(t)$ is a constant which is chosen such that type 0 is indifferent.*

Figure 2 shows how the optimal mechanism varies as (ϕ,η) vary. When the (ϕ,η) lies below the solid line, the lowest type is willing to accept without any further information from Sender. Hence, Sender's problem is nontrivial if and only if (ϕ,η) is 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 $1/2$ when Receiver has the lowest type 0. This term $\overline{\pi}^*(t)$ increases to $1/2$ at some type $\hat{t}$. We show that the integral (10) is maximized at $\overline{\pi}^*(t)$ when $t \leqslant \hat{t}$ and at $1/2$ when $t > \hat{t}$. Given this allocation of $\overline{\pi}(t)$, the allocation $\underline{\pi}(t)$ can be derived based on the lowest type's binding participation constraint and the (differential) incentive constraints. Note that we have ignored the constraint that $\underline{\pi}(t) \geqslant -1/2$. This constraint might bind when η is small enough, that is, the cost of accepting in a negative state is sufficiently small. We show in Lemma 5.2 in the appendix that both $\overline{\pi}^*(t)$ and the corresponding $\underline{\pi}(t)$ increases in η. Therefore, when η is large enough,

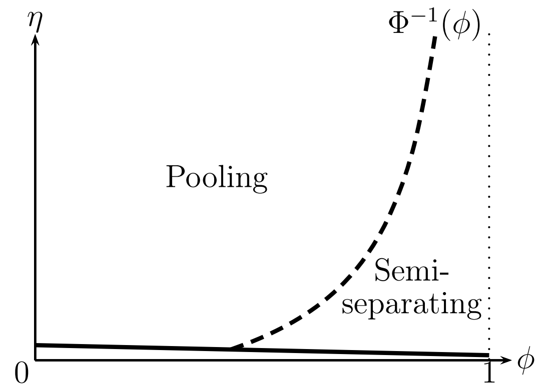

Figure 2: Pooling is optimal when ϕ is small

the constraint $\underline{\pi}(t) \geqslant -1/2$ does not bind.[7]

Proposition 3.3. *Suppose that $\phi > \Phi(\eta)$. There exists $\bar{\eta}(\phi)$ such that when $\eta > \bar{\eta}(\phi)$,*

(i) $\overline{\pi}(t) = \begin{cases} \overline{\pi}^*(t) & \text{for } t \leqslant \hat{t} \\ \frac{1}{2} & \text{for } t > \hat{t}, \end{cases}$ *where $\hat{t}$ is the critical type such that $\overline{\pi}^*\left(\hat{t}\right) = \frac{1}{2}$;*

(ii) $\underline{\pi}(t)$ is determined by the incentive constraints and the lowest type's binding participation constraint.

Figure 3 illustrate the optimal mechanism in which Sender separates the lower types.[8] 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 illustrate $\underline{\pi}(t)$ and $\overline{\pi}(t)$ 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 $\overline{\pi}(t)$ and $\underline{\pi}(t)$ move up as η increases.

The right-hand side of Figure 3 shows how to implement the optimal mechanism in a publicly incentive-compatible mechanism by illustrating the case with η being $1/5$. The solid curve corresponds to the cutoff function $z(s)$. For all the states below $\underline{\pi}(1)$, Sender recommends rejection. For any state s above $\underline{\pi}(1)$, Sender announces $z(s)$ and recommends types above $z(s)$ to accept. For a small segment of states surrounding s_0, $z(s)$ equals the lowest type 0, so all the 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.

[7] When η is large enough, the highest type prefers to reject without further information. In this case, $\underline{\pi}(t) \geqslant -1/2$ does not bind. Hence, there is a non-empty parameter region such that $\underline{\pi}(t) \geqslant -1/2$ does not bind.

[8] The parameter values are $\phi = 1$, $\eta = 1/5$ for the solid line, and $\eta = 1/2$ for the dashed line.

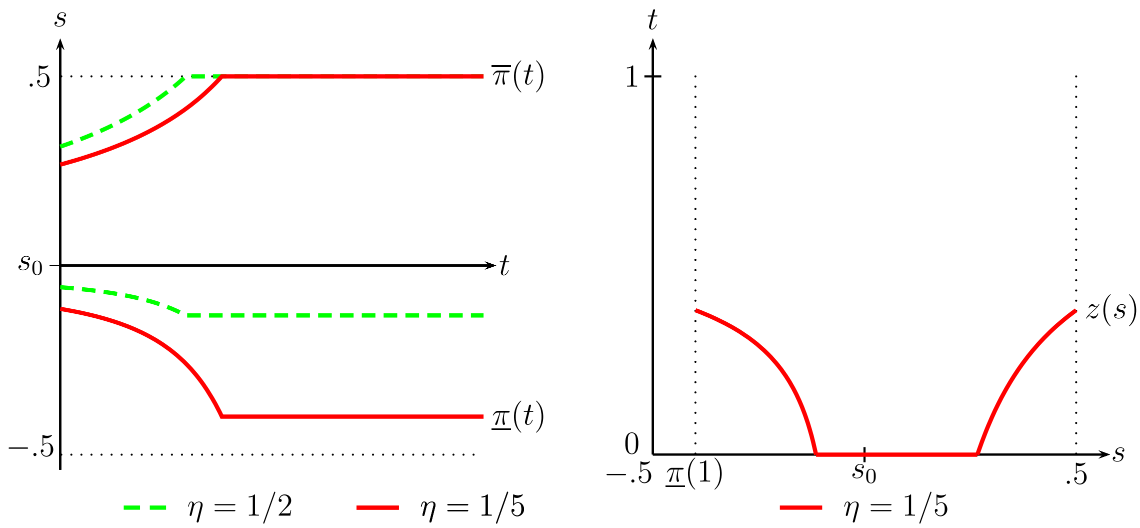

Figure 3: Optimal mechanism and its implementation

4 Extensions

In this section, we generalize our results to the setting where Sender's payoff depends on the state. For example, in the context of our theorist-chair example, this is the case if the theorist's utility from hiring the candidate increases in the candidate's quality.

Let $v(s) > 0$ denote Sender's payoff upon acceptance. Thus, Sender's problem is now

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

among all publicly incentive-compatible mechanisms.

Our theorems still hold when we replace the assumption that u is monotone with the assumption that u/v is monotone. Indeed, the new environment can be reduces to our previous case in which Sender's utility is constant: Let $\tilde{\mu}$ be the measure over $\mathcal{S}$ given by $\mathrm{d}\tilde{\mu} = v\mathrm{d}\mu$ and let $\tilde{u}(s) = u(s)/v(s)$. Then Sender's problem and incentive compatibility of the mechanisms are the same if we replace Sender's utility v by constant utility, the underlying measure μ by $\tilde{\mu}$, and Receiver's utility u by $\tilde{u}$.

5 Proofs

5.1 Proofs for Section 2

We can assume w.l.o.g. that $\mathcal{S} = \mathbf{R}$, that μ is Lebesgue's measure and that $u(0) = 0$.

For a given mechanism with individual acceptance probabilities $\rho(s,t)$ let

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

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

It will be convenient to add a semi-continuity assumption on the cutoff function of our mechanisms. In this Section, whenever we talk about U-shaped cutoff mechanism given by the cutoff function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$ we also assume that z is l.s.c. (lower semi-continuous). This ensures that the states on which each type stops is a closed interval. Moreover, there is a one-to-one correspondence between U-shaped l.s.c. functions $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$ and pairs of non-decreasing u.s.c. (upper semi-continuous) functions $p, n : \mathcal{T} \to \mathbf{R}_+$ such that

$$z(s) \leqslant t \leftrightarrow -n(t) \leqslant s \leqslant p(t). \tag{12}$$

Note that the semi-continuity assumption is vacuous when $\mathcal{T}$ is discrete. The proof of the theorems will use the following lemma and Lemma 2.5.

Lemma 5.1. *For every downward incentive-compatible mechanism there exists a l.s.c. U-shaped function $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$, such that the cutoff mechanism given by z is downward incentive-compatible and weakly dominates the original mechanism.*

Consider the space Z of all U-shaped l.s.c. functions $z : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$. 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 incentive-compatible mechanisms is a closed subset, and Sender's payoff function $z \mapsto \int_{t \geqslant z(s)} f(s,t) \;\; \mu(\mathrm{d}s)\,\lambda(\mathrm{d}t)$ is continuous. Therefore there exists a U-shaped l.s.c. functions $z^* : \mathcal{S} \to \mathcal{T} \cup \{\infty\}$ which is optimal to Sender. By Lemma 5.1 the cutoff mechanism induced by z^* is optimal for Sender among all downward incentive-compatible mechanisms. To prove Theorem 2.1 and Theorem 2.2 it remains to show that z^* is publicly incentive-compatible. By Lemma 2.5 in the publicly incentive-compatible mechanism induced by z^* Receiver t accepts (at least) on the event $\{s : z^*(s) \leqslant t\}$. Note that the set of all publicly incentive-compatible mechanisms is a subset of that of all privately incentive-compatible mechanisms which is a subset of the set of all downward incentive-compatible mechanisms. Because z^* is optimal among all downward incentive-compatible mechanisms, Receiver t cannot accept on a larger event than

$\{s : z^*(s) \leqslant t\}$.

Proof of Lemma 5.1. For every type t let $n(t)$ and $p(t)$ be such that $n(t), p(t) \geqslant 0$ and

$$\begin{aligned} \int_0^\infty f(s,t)u(s)\rho(s,t)\,\mathrm{d}s &= \int_0^{p(t)} f(s,t)u(s)\,\mathrm{d}s, \text{ and} \\ \int_{-\infty}^0 f(s,t)u(s)\rho(s,t)\,\mathrm{d}s &= \int_{-n(t)}^0 f(s,t)u(s)\,\mathrm{d}s. \end{aligned} \tag{13}$$

The existence of $p(t), n(t)$ for every type t follows from the assumption on μ. From (13) and i.m.l.r. we get that

$$\begin{aligned} \int_0^\infty f(s,t)u(s)\rho(s,t')\,\mathrm{d}s &\geq \int_0^{p(t')} f(s,t)u(s)\,\mathrm{d}s, \text{ and} \\ \int_{-\infty}^0 f(s,t)u(s)\rho(s,t')\,\mathrm{d}s &\geq \int_{-n(t')}^0 f(s,t)u(s)\,\mathrm{d}s \end{aligned} \tag{14}$$

for $t' < t$. From (13), (14) and (4) it follows that

$$\int f(s,t)u(s)\left(\mathbf{1}_{[-n(t),p(t)]} - \mathbf{1}_{[-n(t'),p(t')]}\right)\,\mathrm{d}s \geqslant 0 \tag{15}$$

for every $t' < t$.

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

$$\int_0^\infty f(s,t)\rho(s,t)\,\mathrm{d}s \leqslant \int_{-n(t)}^{p(t)} f(s,t)\,\mathrm{d}s \tag{16}$$

by the Neyman-Pearson Lemma.

Thus, the mechanism that recommends type t to accept when the state is in the interval $[-n(t), p(t)]$ is downward incentive-compatible (15) and is better for Sender (16). Note however that this is not yet a cutoff mechanism.

Let $\bar{p}(t) = \inf_{\varepsilon>0} \sup\{p(r) \colon r < t+\varepsilon\}$ and $\bar{n}(t) = \inf_{\varepsilon>0} \sup\{n(r) \colon r < t+\varepsilon\}$ be the smallest non-decreasing and u.s.c. functions that dominate p and n respectively and let z be given by (12):

$$z(s) \leqslant t \leftrightarrow -\bar{n}(t) \leqslant s \leqslant \bar{p}(t).$$

Clearly the new mechanism given by z is better for Sender. We claim that this mechanism is downward incentive-compatible.

Let $t' < t$. we first need to show that for the mechanism $(\bar{n}, \bar{p})$ type t will not mimic a lower type t', i.e., that

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

Let t_k, t'_k be such that

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

If $t_\infty \leqslant t'$ then $\bar{n}(t') = \bar{n}(t)$ and from (12) we get that $\{t' \leqslant z(s) < t\} \subseteq \{s \geqslant 0\}$ so (17) holds. Therefore we can assume that $t' < t_\infty$ and therefore $t'_k < t_k$ for every k.

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

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

By a similar argument, we get that $\limsup_{k\to\infty} n(t'_k) \leqslant \bar{n}(t')$ and $\limsup_{k\to\infty} n(t_k) = \bar{n}(t)$. Combined with the condition that $u(s) \leqslant 0$ for any $s \leqslant 0$, we get that

$$\limsup_{k\to\infty} \int_{-n(t_k)}^{-n(t'_k)} f(s, t_k)u(s) \, \mathrm{d}s \leqslant \int_{-\bar{n}(t_\infty)}^{-\bar{n}(t')} f(s, t_\infty)u(s) \, \mathrm{d}s.$$

Therefore

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

Given i.m.l.r., $\int f(s,t)u(s) \left(\mathbf{1}_{[-\bar{n}(t), \bar{p}(t)]} - \mathbf{1}_{[-\bar{n}(t'), \bar{p}(t')]} \right) \mathrm{d}s$ must be positive as well since $t \geqslant t_\infty$. This proves (17).

Finally, we need to show that the lowest type gets payoff at least 0 from obeying under the mechanism $(\bar{p}, \bar{n})$. In the case of discrete type space this follows from the corresponding property of the original mechanism since in this case $\bar{p}(\underline{t}) = p(\underline{t})$ and $\bar{n}(\underline{t}) = n(\underline{t})$. In the general case we need to appeal to a similar argument to the one we used to prove the downward incentive compatibility conditions using converging sequences of types. We omit this argument. □

5.2 Proofs for Section 3

Proof of Proposition 3.1. We first note that, under our assumption that type L rejects without additional information, both IC constraints in (7) bind. Indeed, for type H this holds trivially in the pooling case and 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 2.1 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 (7) are determined by a single variable: If $\overline{\pi}_L = y$ for some $\zeta \leqslant y \leqslant 1$ then $\underline{\pi}_L = \ell_L(y)$ 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 $\underline{\pi}_H = \ell_H(y)$ 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 derivatives are given by:

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

Clearly ℓ_L is monotone decreasing and i.m.l.r. implies that ℓ_H is monotone increasing.

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

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

So the 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 $\ell'_L(y) < 0 < \ell'_H(y)$, the left-hand side increases in y due to the i.m.l.r. property 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.

Substituting $f(s,H), f(s,L), u(s)$, we obtain 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 side is monotone increasing in ϕ, as desired. □

Proof of Proposition 3.2. We let $G(\overline{\pi}(t))$ denote the integrand in (10). (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

$$b_1(t) := \frac{8\left(\phi(1-2t) + 2(-2\phi t + \phi - 4)\tanh^{-1}\left(\frac{\phi(t-1)}{\phi t + 4}\right)\right)}{3\eta(\phi(1-2t) - 4)},$$

$$b_2(t) := \left(4 + \frac{2(1-8\eta)}{\eta(\phi(2t-1)+4)}\right),$$

$$b_3(t) := \frac{4}{2\phi t - \phi + 4}.$$

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 is 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 $t' \in (0,1)$ be the value which solves $b_1(t) = 0$. If $\eta\phi \leqslant 1/2$, $b_2(t)$ is always positive. Otherwise, $b_2(t)$ is positive if and only if $t \geqslant \frac{2\eta\phi - 1}{4\eta\phi}$.

When $t < t'$, 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

$$\frac{2\eta}{2\eta\phi(1-2t)-1+\sqrt{(2\eta\phi-4\eta\phi t+1)^2+8\eta(2\phi t-\phi+4)\log\left(\frac{\phi+4}{2\phi t-\phi+4}\right)}}. \tag{18}$$

The value $G(\overline{\pi}(t))$ is maximized at this positive root if it is below 1/2 and at 1/2 otherwise. When $t > t'$, we want to show that $b_2(t)$ is positive for any $t > t'$. This is the case when $\eta\phi \leqslant 1/2$. If $\eta\phi > 1/2$, it is readily verified that the value of $b_1(t)$ at $t = \frac{2\eta\phi-1}{4\eta\phi}$ is strictly negative. This implies that $\frac{2\eta\phi-1}{4\eta\phi} < t'$. Hence, $b_2(t) > 0$ for any $t > t'$. We contend that $G'(\overline{\pi})$ is strictly positive for any $t > t'$ so $G(\overline{\pi})$ is maximized at $\overline{\pi}(t) = 1/2$.

Next, we show that the positive root (18) given $t < t'$ increases in t. The denominator of (18) decreases in t if

$$2\eta\phi(1-2t)-1+\sqrt{(2\eta\phi-4\eta\phi t+1)^2+8\eta(2\phi t-\phi+4)\log\left(\frac{\phi+4}{2\phi t-\phi+4}\right)}$$
$$+4\tanh^{-1}\left(\frac{\phi(t-1)}{\phi t+4}\right)+4>0,$$

which is implied by the condition that the denominator (18) is positive and the observation that $4\tanh^{-1}\left(\frac{\phi(t-1)}{\phi t+4}\right)+4$ is positive for all $t \in [0,1]$. This completes the proof that (18) increases in t for $t < t'$. Moreover, since $b_1(t)$ drops to zero at t', the positive root (18) approaches infinity at t'. The implies that (18) reaches 1/2 before t'. We let $\hat{t}$ be the value at which (18) reaches 1/2.

If (18) is above 1/2 at $t = 0$, The value $G(\overline{\pi}(t))$ is maximized at $\overline{\pi}(t) = 1/2$ for every t. Thus, Sender pools all types. This is the case when

$$\frac{2\eta}{2\eta\phi+\sqrt{(1+2\eta\phi)^2+8\eta(\phi-4)\log\left(\frac{8}{\phi+4}-1\right)}+1} > \frac{1}{2},$$

which can be simplified to

$$\eta > \frac{1}{2}\left(\frac{(\phi-4)\log\left(\frac{8}{\phi+4}-1\right)}{1-\phi}-1\right).$$

The right-hand side increases in ϕ, so there exists an increasing function $\Phi(\cdot)$ such that pooling is optimal when $\phi < \Phi(\eta)$ and semi-separating is optimal otherwise. $\square$

Lemma 5.2. *Suppose that $\phi > \Phi(\eta)$. Both $\overline{\pi}(t)$ and $\underline{\pi}(t)$ increases pointwise in η. Thus, there exists $\bar{\eta}(\phi)$ such that the constraint $\underline{\pi}(t) \geqslant -1/2$ does not bind if and only if $\eta > \bar{\eta}(\phi)$.*

Proof of Lemma 5.2. When t equals $1/2$, the value of $b_1(t)$ defined in the proof of Proposition 3.2 is $\frac{8\log\left(\frac{4}{\phi+4}\right)}{3\eta}$, which is negative given that $\eta > 0$ and $\phi > 0$. When t equals $1/2$, the value of (18) equals:

$$\frac{2\eta}{\sqrt{1+32\eta\log\left(\frac{\phi+4}{4}\right)}-1},$$

which is greater than $1/2$. 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 $1/\overline{\pi}(t)$ with respect to η is negative if and only if

$$\begin{aligned} &2\eta\phi(1-2t) + 8\eta(\phi(1-2t)-4)\tanh^{-1}\left(\frac{\phi(t-1)}{\phi t+4}\right) + 1 \\ &\geqslant \sqrt{(2\eta\phi(1-2t)+1)^2 + 8\eta(\phi(1-2t)-4)\log\left(\frac{2\phi t-\phi+4}{\phi+4}\right)}. \end{aligned}$$

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:

$$\underline{\pi}(0) = \frac{-2\eta\left(-2\eta\phi + 3\sqrt{(1+2\eta\phi)^2 + 8\eta(\phi-4)\log\left(\frac{8}{\phi+4}-1\right)} - 3\right)}{3\left(2\eta\phi + \sqrt{(1+2\eta\phi)^2 + 8\eta(\phi-4)\log\left(\frac{8}{\phi+4}-1\right)} - 1\right)^3}.$$

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 explicitly:

$$\underline{\pi}(t,\eta) = \frac{\overline{\pi}(t,\eta)^2(4\phi(2t-1)\overline{\pi}(t,\eta)+3) - 8\phi\int_0^t \overline{\pi}(z,\eta)^3\,dz}{-6\eta}.$$

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

$$a_1 x(t,\eta)^3 + a_2 x(t,\eta)^2 + a_3 x(t,\eta) + a_4 > 0, \tag{19}$$

where

$$\begin{aligned}
a_1 &= 2\eta(\phi(6t-3)-8)-3,\\
a_2 &= (\eta\phi(4t-2)-1)(2\eta(7\phi(1-2t)-24)-1),\\
a_3 &= 2\eta(\phi(2t-1)(-2\eta(\phi(2t-1)(10\eta(\phi(1-2t)-8)+23)+144)+7)+72)+3,\\
a_4 &= (\eta\phi(4t-2)+1)(2\eta\phi(1-2t)+1)(4\eta(\phi(1-2t)(\eta(\phi(1-2t)-8)+2)-12)-1).
\end{aligned}$$

It is easy to show that $a_2>0$ and that $x(t,\eta)>1$. Moreover, the cubic inequality (19) is satisfied when $x(t,\eta)=1$. We rewrite the left-hand side of (19) 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 (19) is satisfied. Next, we prove that $(a_1 x(t,\eta)^2 + a_3)$ is positive.

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

$$\frac{4(2\eta\phi(1-2t)(\eta\phi(1-2t)+2)+1)}{2\eta(3\phi-6\phi t+8)+3} + \log\left(\frac{2\phi t-\phi+4}{\phi+4}\right) > 0. \tag{20}$$

It is easy to verify that the left-hand side of (20) is convex in η. The inequality (20) holds when η equals zero. We are interested in the parameter region when $\overline{\pi}(t)<1/2$. For fixed

ϕ and t, $\overline{\pi}(t)$ is smaller than 1/2 if

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

When we substitute $\eta = \underline{\eta}$ into (20), the inequality is satisfied for any $t \in (0, 1/2)$ and $\phi \in (0, 1)$.

The left-hand side of (20) increases in η if and only if

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

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

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

Given that $\eta > 0$, we obtain that the left-hand side of (20) decreases in η when $\eta < \tilde{\eta}$ and increases in η when $\eta \geqslant \tilde{\eta}$. We have to discuss two cases, depending on whether $\tilde{\eta}$ is above or below $\underline{\eta}$:

1. If $\tilde{\eta} \geqslant \underline{\eta}$, the left-hand side of (20) increases in η for any $\eta \in \left(0, \underline{\eta}\right)$. We have shown that the left-hand side of (20) is positive at $\eta = 0$ and $\eta = \underline{\eta}$. Hence, the left-hand side of (20) is positive for any $\eta \in \left(\underline{\eta}, 0\right)$.

2. If $\tilde{\eta} \in \left(0, \underline{\eta}\right)$, then the minimum of the left-hand side of (20) is achieved when $\eta = \tilde{\eta}$. Hence, we need to show that the left-hand side of (20) is positive when $\eta = \tilde{\eta}$. Substituting $\eta = \tilde{\eta}$ into (20), the inequality (20) is satisfied if and only if

$$\left(4\phi(2t-1)(\phi(6t-3)-16) - (\phi(3-6t)+8)^2\log\left(\frac{\phi+4}{2\phi t-\phi+4}\right)\right)\sqrt{128-9\phi^2(1-2t)^2}$$
$$> 4\phi(2t-1)\left(9\phi^2(1-2t)^2-128\right).$$

   The condition $\tilde{\eta} < \underline{\eta}$ holds only if $\phi > 3/5$ and $t < 1/4$. When we restrict attention to this parameter region, the inequality above holds.

Combining the two cases above, we have shown that (20) holds. □

Proof of Proposition 3.2. This follows based on Lemma 5.2 and the proof of Proposition 3.2. □

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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