Bayesian Persuasion[†]

By Emir Kamenica and Matthew Gentzkow*

When is it possible for one person to persuade another to change her action? We consider a symmetric information model where a sender chooses a signal to reveal to a receiver, who then takes a noncontractible action that affects the welfare of both players. We derive necessary and sufficient conditions for the existence of a signal that strictly benefits the sender. We characterize sender-optimal signals. We examine comparative statics with respect to the alignment of the sender's and the receiver's preferences. Finally, we apply our results to persuasion by litigators, lobbyists, and salespeople. (JEL D72, D82, D83, K40, M31)

Suppose one person, call him Sender, wishes to persuade another, call her Receiver, to change her action. If Receiver is a rational Bayesian, can Sender persuade her to take an action *he* would prefer over the action she was originally going to take? If Receiver understands that Sender chose what information to convey with the intent of manipulating her action for his own benefit, can Sender still gain from persuasion? If so, what is the optimal way to persuade?

These questions are of substantial economic importance. As Donald McCloskey and Arjo Klamer (1995) emphasize, attempts at persuasion command a sizable share of our resources. Persuasion, as we will define it below, plays an important role in advertising, courts, lobbying, financial disclosure, and political campaigns, among many other economic activities.

Consider the example of a prosecutor trying to convince a judge that a defendant is guilty. When the defendant is indeed guilty, revealing the facts of the case will tend to help the prosecutor's case. When the defendant is innocent, revealing facts will tend to hurt the prosecutor's case. Can the prosecutor structure his arguments, selection of evidence, etc. so as to increase the probability of conviction by a rational judge *on average*? Perhaps surprisingly, the answer to this question is yes. Bayes's Law restricts the expectation of posterior beliefs but puts no other constraints on

their distribution. Therefore, so long as the judge's action is not linear in her beliefs, the prosecutor may benefit from persuasion.

To make this concrete, suppose the judge (Receiver) must choose one of two actions: to *acquit* or *convict* a defendant. There are two states of the world: the defendant is either *guilty* or *innocent*. The judge gets utility 1 for choosing the just action (convict when guilty and acquit when innocent) and utility 0 for choosing the unjust action (convict when innocent and acquit when guilty). The prosecutor (Sender) gets utility 1 if the judge convicts and utility 0 if the judge acquits, regardless of the state. The prosecutor and the judge share a prior belief $\Pr(guilty) = 0.3$.

The prosecutor conducts an investigation and is required by law to report its full outcome. We can think of the choice of the investigation as consisting of the decisions on whom to subpoena, what forensic tests to conduct, what questions to ask an expert witness, etc. We formalize an investigation as distributions $\pi(\cdot \,|\, guilty)$ and $\pi(\cdot \,|\, innocent)$ on some set of signal realizations. The prosecutor chooses π and must honestly report the signal realization to the judge.

If there is no communication (or equivalently, if π is completely uninformative), the judge always acquits because guilt is less likely than innocence under her prior. If the prosecutor chooses a fully informative investigation, one that leaves no uncertainty about the state, the judge convicts 30 percent of the time. The prosecutor can do better, however. As we show below, his uniquely optimal investigation is a binary signal:

$$(1) \qquad \pi(i\,|\,innocent) \;=\; \tfrac{4}{7} \quad \pi(i\,|\,guilty) \;=\; 0$$

$$\pi(g\,|\,innocent) \;=\; \tfrac{3}{7} \quad \pi(g\,|\,guilty) \;=\; 1.$$

This leads the judge to convict with probability 60 percent. Note that the judge knows 70 percent of defendants are innocent, yet she convicts 60 percent of them! She does so even though she is fully aware that the investigation was designed to maximize the probability of conviction.

In this paper, we study the general problem of persuading a rational agent by controlling her informational environment. We consider a symmetric information setting with an arbitrary state space and action space, an arbitrary prior, and arbitrary state-dependent preferences for both Sender and Receiver. Sender chooses an informative signal about the state of the world, Receiver observes a realization from this signal, and then she takes an action. Throughout the analysis, we prohibit Sender from making transfers or affecting Receiver's payoffs in any way. We focus on two questions: (i) when does there exist a signal that strictly benefits Sender, and (ii) what is an optimal signal from Sender's perspective?

A key assumption of our model is that Sender cannot distort or conceal information once the signal realization is known. This allows us to abstract from the incentive compatibility issues that are the focus of much of the previous literature on strategic communication. We discuss the precise meaning of this assumption and the settings where it is likely to apply when we introduce the general model below.

We begin by establishing a result that simplifies our analysis. We show that we can reexpress the problem of choosing an optimal signal as a search over distributions of

posteriors subject to the constraint that the expected posterior is equal to the prior. This reformulation of Sender's problem provides a useful geometric approach to deriving the optimal signal.

When does there exist a signal that strictly benefits Sender? Consider why the prosecutor in the example benefits from the opportunity to provide information to the judge. Since the judge is rational, providing information must sometimes make her more convinced and sometimes less convinced that the defendant is guilty. The former will strictly improve the prosecutor's payoff if the information is strong enough to induce conviction. The latter, however, will not reduce the prosecutor's payoff, since the judge already acquits the defendant by default. The net effect is to increase the prosecutor's payoff in expectation. We show that in general Sender benefits from persuasion whenever (i) Receiver does not take Sender's preferred action by default (in a sense we make precise below) and (ii) Receiver's action is constant in some neighborhood of beliefs around the prior. When these conditions hold, Sender can benefit by sending a signal that induces a better action with positive probability and balances this with a worse belief that leaves Receiver's action unchanged. We also show that whether Sender benefits from persuasion depends in a natural way on the concavity or convexity of Sender's payoff as a function of Receiver's beliefs.

We next turn to studying optimal signals. We use tools from convex analysis to characterize the optimal signal for any given set of preferences and initial beliefs. We show that no disclosure of information is optimal when Sender's payoff is concave in Receiver's beliefs, and full disclosure is optimal when Sender's payoff is convex in Receiver's beliefs.

We then generalize two important properties of the optimal signal in the example above. Notice, first, that when the judge chooses the prosecutor's least-preferred action (*acquit*), she is certain of the state. That is, she never acquits guilty defendants. Otherwise, we would have $\pi(i\,|\,guilty) > 0$. But then the prosecutor could increase his payoff by decreasing $\pi(i\,|\,guilty)$ and increasing $\pi(g\,|\,guilty)$; this would strictly increase the probability of g and would only increase the willingness of the judge to convict when she sees g. We establish that, in general, whenever Receiver takes Sender's least-preferred action, she knows with certainty that the state is one where this action is uniquely optimal.

Second, notice that when the judge convicts, she is exactly indifferent between convicting and acquitting. If she strictly preferred to convict upon seeing g, the prosecutor could increase his payoff by slightly decreasing $\pi(i\,|\,innocent)$ and increasing $\pi(g\,|\,innocent)$; this would increase the probability of g and leave the judge's optimal action given the message unchanged, thus increasing the probability of conviction. We show that, in general, whenever Receiver has an interior posterior, she is effectively indifferent between two actions.

We also examine how the extent of information transmission and Sender's gain from persuasion depend on the alignment of Sender's and Receiver's preferences. In contrast to previous work on strategic communication, we find that making preferences more aligned can reduce the extent of communication in equilibrium.

We next apply our results to two examples. First, we examine the type of studies that a lobbying group might fund so as to influence a benevolent politician. Then, we analyze a seller's provision of free trials or other information to potential consumers.

In the final section of the paper, we discuss extensions of our results to settings where Receiver has private information and where there are multiple receivers.

The observation that Bayesian updating only restricts the expectation of posteriors has been made before and has been utilized in a variety of contexts. The work most closely related to our paper is Robert J. Aumann and Michael B. Maschler (1995) and Isabelle Brocas and Juan D. Carrillo (2007). Aumann and Maschler (1995) employ formal methods very similar to ours to analyze repeated games of incomplete information. They study the value to a player of knowing which game is being played when the other player lacks this knowledge, a fixed zero-sum game is repeated *ad infinitum*, players maximize their long-run nondiscounted average payoffs, and payoffs are not observed. The fact that the informed player's initial actions have no impact on his long-run average payoffs (and can thus be treated as nothing but a signal) combined with a focus on Nash equilibria (which implicitly allow for commitment) makes Aumann and Maschler's problem mathematically analogous to ours.

Brocas and Carrillo (2007) analyze the gain to Sender from controlling the flow of public information in a setting with a binary state space and information that consists of a sequence of symmetric binary signals. Our consideration of a more general environment allows us to provide new intuitions about when Sender benefits from persuasion and about the optimal informational environment from Sender's perspective. For example, we establish that whenever Receiver has finitely many actions and there is some information Sender would share, Sender benefits from persuasion. Also, we derive several novel properties that beliefs induced by an optimal signal must satisfy.

This paper also relates to a broader literature on optimal information structures. Luis Rayo and Ilya Segal (2010) characterize the optimal disclosure policy under specific assumptions about preferences and about Receiver's outside option. Michael Ostrovsky and Michael Schwarz (2010) examine the equilibrium design of grade transcripts and the resulting information about quality of students when schools compete to place their students in good jobs. Johannes Hörner and Andrzej Skrzypacz (2005) demonstrate how sequential revelation of partially informative signals can increase payments to a Sender who is trying to sell his information to Receiver. Finally, the results in our paper connect to several other strands of research, including work on optimal advertising, self-signalling, and contract theory.

I. A Model of Persuasion

A. *Setup*

Receiver has a continuous utility function $u(a, \omega)$ that depends on her action $a \in A$ and the state of the world $\omega \in \Omega$. Sender has a continuous utility function $v(a, \omega)$ that depends on Receiver's action and the state of the world. Sender and Receiver share a prior $\mu_0 \in \text{int}(\Delta(\Omega))$.[1] The action space A is compact and the state space Ω is finite. The latter assumption is mainly for ease of exposition. In the online Appendix, we show that our central characterization result extends to the case where Ω is any compact metric space.

---

[1] $\text{int}(X)$ denotes the interior of set X and $\Delta(X)$ the set of all probability distributions on X.

A signal π consists of a finite realization space S and a family of distributions $\{\pi(\cdot\,|\,\omega)\}_{\omega\in\Omega}$ over S. Sender chooses a signal. Receiver observes Sender's choice of the signal and a signal realization $s \in S$, and then takes her action. Our solution concept is Sender-preferred subgame perfect equilibrium: given Sender's choice of π and a signal realization s, Receiver forms the posterior μ_s using Bayes's rule and takes an action from the set $a^*(\mu_s) = \arg\max_{a\in A} E_{\mu_s}[u(a, \omega)]$. We assume that there are at least two actions in A and that for any action a there exists a μ such that $a \in a^*(\mu)$. If Receiver is indifferent between some actions at a given belief, we assume she takes an action that maximizes Sender's expected utility. This assumption simplifies notation and provides a useful benchmark in case of multiple equilibria. We let $\hat{a}(\mu)$ denote Receiver's equilibrium action at belief μ.[2] We refer to $\hat{a}(\mu_0)$ as the *default action*. Taking Receiver's behavior as given, Sender chooses a signal π which maximizes his expected utility.[3] In the remainder of the paper, we use the term "equilibrium" to mean a Sender-preferred subgame perfect equilibrium.

A common special case is where ω is a real-valued random variable, Receiver's action depends only on the expectation $E_\mu[\omega]$, rather than the entire distribution μ, and Sender's preferences over Receiver's actions do not depend on ω. This holds, for example, if $u(a, \omega) = -(a - \omega)^2$ and $v(a, \omega) = a$. When these conditions are satisfied, we will say that Sender's payoff depends only on the expected state.

We define the *value* of a signal to be the equilibrium expectation of $v(a, \omega)$ under that signal. The *gain* from a signal is the difference between its value and the equilibrium expectation of $v(a, \omega)$ when Receiver obtains no information. *Sender benefits from persuasion* if there is a signal with a strictly positive gain. A signal is *optimal* if no other signal has higher value. Clearly, in equilibrium Sender selects an optimal signal.

Given a signal, each signal realization s leads to a posterior belief $\mu_s \in \Delta(\Omega)$. Accordingly, each signal leads to a distribution over posterior beliefs. We denote a distribution of posteriors by $\tau \in \Delta(\Delta(\Omega))$. Algebraically, a signal π induces τ if $\text{Supp}(\tau) = \{\mu_s\}_{s\in S}$ and

$$\mu_s(\omega) = \frac{\pi(s\,|\,\omega)\mu_0(\omega)}{\sum_{\omega'\in\Omega}\pi(s\,|\,\omega')\mu_0(\omega')} \text{ for all } s \text{ and } \omega$$

$$\tau(\mu) = \sum_{s:\mu_s=\mu}\sum_{\omega'\in\Omega}\pi(s\,|\,\omega')\mu_0(\omega') \text{ for all } \mu.$$

We say a belief μ is induced by a signal if τ is induced by that signal and $\tau(\mu) > 0$. A distribution of posteriors is *Bayes plausible* if the expected posterior probability equals the prior:

$$\sum_{Supp(\tau)}\mu\tau(\mu) = \mu_0.$$

---

[2] If there is more than one action in $a^*(\mu)$ that maximizes Sender's expected utility, we let $\hat{a}(\mu)$ denote an arbitrary element of that set. This allows us to use convenient notation such as $v(\hat{a}(\mu), \omega)$.

[3] Note that Sender cannot benefit by using mixed strategies; for any randomization over signals, there is a single signal that leads to identical outcomes.

B. *Simplifying the Problem*

A distribution of posteriors induced by a signal determines a distribution of Receiver's actions. From Sender's perspective, any signal that induces the same distribution of actions conditional on the state must have the same value. To determine whether there exists a signal with some given value, therefore, it is sufficient to ask whether there exists a distribution of Receiver's posteriors that can be induced by a signal and that generates expected utility for Sender equal to that value. Hence, we need to know (i) which distributions of posteriors can be induced, and (ii) what is the expected utility generated by each distribution of posteriors?

Bayesian rationality requires that any equilibrium distribution of Receiver's beliefs be Bayes-plausible. Our first proposition below shows that this is the *only* restriction imposed by Bayesian rationality. That is, for any Bayes-plausible distribution of posteriors there is a signal that induces it.[4]

When both Sender and Receiver hold some belief μ, Sender's expected utility is equal to

$$\hat{v}(\mu) \equiv E_\mu v(\hat{a}(\mu), \omega).$$

Since Sender's and Receiver's beliefs coincide, Sender's utility from any signal which induces a distribution of posteriors τ is simply the expectation of $\hat{v}$ under τ, i.e., $E_\tau \hat{v}(\mu)$.

Combining the two observations above allows us to greatly simplify the analysis of our game. We simplify the analysis further by noting that, without loss of generality, we can restrict our attention to a particular class of signals. Say that a signal is *straightforward* if $S \subseteq A$ and Receiver's equilibrium action equals the signal realization. In other words, a straightforward signal produces a "recommended action" and Receiver always follows the recommendation.

PROPOSITION 1: *The following are equivalent:*

  (i) *There exists a signal with value v^*;*

 (ii) *There exists a straightforward signal with value v^*;*

(iii) *There exists a Bayes-plausible distribution of posteriors τ such that $E_\tau \hat{v}(\mu) = v^*$.*

Detailed proofs of all propositions are in Appendix A. The basic idea behind this proposition, however, is simple. The equivalence of (i) and (ii) is closely analogous to the revelation principle (e.g., Roger B. Myerson 1979), except that the revelation principle applies to problems where players' information is a given, while our problem is that of designing the informational environment. To see the equivalence between (i) and (iii), given a τ let S index Supp(τ) and consider a signal

---

[4] Strictly speaking, because S is finite, this is true only for distributions of posteriors with finite support. This, however, is not a substantive restriction, since Ω is also finite.

$\pi(s \,|\, \omega) = \mu_s(\omega)\,\tau(\mu_s)/\mu_0(\omega)$. Simple algebra reveals that π induces τ. As we mentioned earlier, this equivalence shows that Bayes plausibility is the only restriction on the equilibrium distribution of posteriors.[5] This fact is closely related to Eran Shmaya and Leeat Yariv's (2009) concurrent work that identifies which sequences of distributions of posteriors are consistent with Bayesian rationality.

The key implication of Proposition 1 is that to evaluate whether Sender benefits from persuasion and to determine the value of an optimal signal we need only ask how $E_\tau \hat{v}(\mu)$ varies over the space of Bayes-plausible distributions of posteriors.[6]

COROLLARY 1: *Sender benefits from persuasion if and only if there exists a Bayes-plausible distribution of posteriors τ such that*

$$E_\tau \hat{v}(\mu) \; > \; \hat{v}(\mu_0).$$

The value of an optimal signal is

$$\max_\tau E_\tau \hat{v}(\mu)$$

$$\text{s.t.} \sum_{Supp(\tau)} \mu \, d\tau(\mu) \; = \; \mu_0.$$

Note that Corollary 1 does not by itself tell us that an optimal signal exists. Our focus on Sender-preferred equilibria, however, implies that $\hat{v}$ is upper semicontinuous which in turn ensures the existence of an optimal signal.

We introduce a final definition that will be useful in the analysis that follows. Let V be the *concave closure* of $\hat{v}$:

$$V(\mu) \; \equiv \; \sup\{z \,|\, (\mu, z) \in co(\hat{v})\},$$

where $co(\hat{v})$ denotes the convex hull of the graph of $\hat{v}$. Note that V is concave by construction. In fact, it is the smallest concave function that is everywhere weakly greater than $\hat{v}$.[7] Figure 1 shows an example of the construction of V. In this figure, the state space is binary, and we identify a distribution μ with the probability of one of the states. Specifying this probability of course uniquely pins down the entire distribution μ.

To see why V is a useful construct, observe that if $(\mu', z) \in co(\hat{v})$, then there exists a distribution of posteriors τ such that $E_\tau \mu = \mu'$ and $E_\tau \hat{v}(\mu) = z$. Thus, by Proposition 1, $co(\hat{v})$ is the set of (μ, z) such that if the prior is μ, there exists a signal with value z. Hence, $V(\mu)$ is the largest payoff Sender can achieve with any signal when the prior is μ.
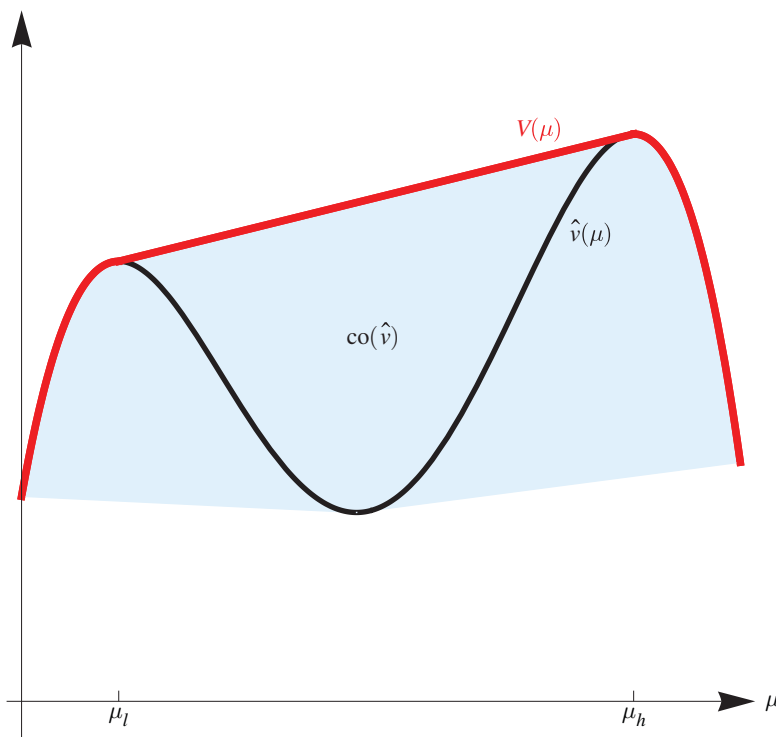
FIGURE 1. AN ILLUSTRATION OF CONCAVE CLOSURE

COROLLARY 2: *The value of an optimal signal is $V(\mu_0)$. Sender benefits from persuasion if and only if $V(\mu_0) > \hat{v}(\mu_0)$.*

Figure 2 shows the function $\hat{v}$, the optimal signal, and the concave closure V in the motivating example from the introduction. In the figure, μ denotes the probability that the state is *guilty*. As panel A shows, $\hat{v}$ is a step function: the prosecutor's expected payoff is 0 whenever μ is less than 0.5 (since the judge will choose *acquit*) and 1 whenever μ is greater than or equal to 0.5 (since the judge will choose *convict*). As panel B shows, the optimal signal induces two posterior beliefs. When the judge observes i, her posterior belief is $\mu = 0$ and $\hat{v}(0) = 0$. When the judge observes g, her posterior belief is $\mu = 0.5$ and $\hat{v}(0.5) = 1$. The distribution τ over these beliefs places probability 0.4 on $\mu = 0$ and probability 0.6 on $\mu = 0.5$. Hence, the prosecutor's expected utility is $E_\tau \hat{v}(\mu) = 0.6$. The distribution τ is Bayes plausible since $\mu_0 = 0.3 = 0.4 \times 0 + 0.6 \times 0.5$. As panel C shows, the concave closure V is equal to 2μ when $\mu \leq 0.5$ and constant at 1 when $\mu > 0.5$. It is clear that $V(\mu_0) > \hat{v}(\mu_0)$ and that the value of the optimal signal is $V(\mu_0)$. Finally, following the proof of Proposition 1, it is easy to compute the signal that induces the optimal τ: $\pi(s|\omega) = \mu_s(\omega) \tau(\mu_s)/\mu_0(\omega)$. This yields the signal in equation (1).

C. *The Commitment Assumption*

Most existing models of strategic communication consider settings where Sender is perfectly informed and has a limited ability to commit to communicate to Receiver
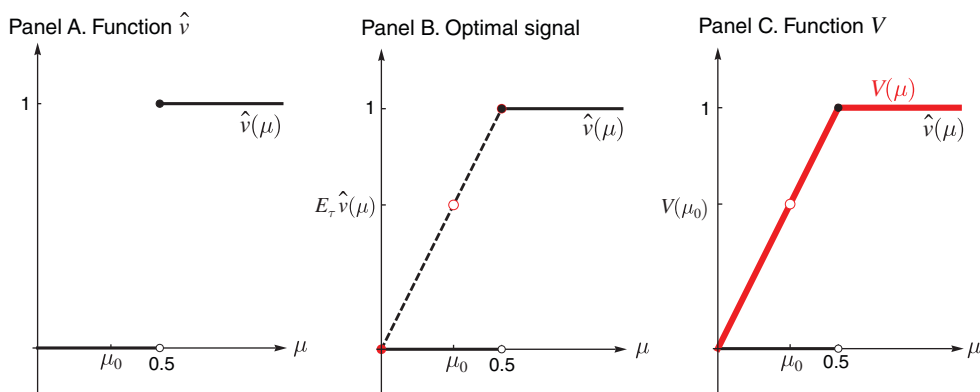
Figure 2. The Motivating Example

what he knows. For example, in Vincent Crawford and Joel Sobel (1982), he can costlessly send any message regardless of what he knows. In Sanford J. Grossman (1981) and Paul Milgrom (1981), he must tell the truth but not necessarily the whole truth. In Navin Kartik (2009), he suffers a lying cost for reporting messages that are far from the truth. In A. Michael Spence (1973), his cost of sending a high education message is decreasing in his true ability.

Our model gives Sender more commitment power than these papers because it assumes the realization of the signal is truthfully communicated to Receiver. This makes our environment effectively nonstrategic. Receiver solves a simple decision problem given the information she receives, and Sender needn't worry about Receiver's interpretation of his actions. In particular, Sender's choice of a signal establishes the meaning of messages and eliminates the possibility of babbling equilibria and related coordination failures that are common in other models of strategic communication.

Although this commitment assumption is strong, it is weaker than it may seem at first. For instance, it may seem restrictive to assume that Sender can generate signals that are arbitrarily informative. This assumption, however, is innocuous under an appropriate interpretation of the state ω. We can define ω to be the realization of the most informative signal Sender can generate. Then, his choice of π is equivalent to a choice of how much to garble information that is potentially available.

It may also seem restrictive to assume that Sender can commit to limit the information he gathers. In many situations, we might worry that when Sender chooses a partially informative signal, he may wish to deviate and generate additional information following certain signal realizations. As we show later, however, under the optimal signal, Sender would never want to generate any further information regardless of the initial signal realization (cf. Lemma 2 in the Appendix).

Finally, we can also weaken the assumption that Sender will truthfully communicate any signal realization, and allow Sender to conceal some or all of his information, so long as the messages he does send are verifiable. Specifically, consider an alternate game where Sender chooses signal π, privately observes its realization s,

and then sends a message $m \in \mathcal{P}(S)$ such that $s \in m$.[8] Receiver knows what signal π Sender chose and observes the message m.[9] Given any preferences and initial beliefs, the set of equilibrium outcomes of this alternate game and the set of equilibrium outcomes of the game we study in this paper coincide. We present a proof of this claim in the online Appendix.

There are a number of real-world settings in which our commitment assumption is a reasonable approximation. First, there are contexts where commitment is legally mandated. Our motivating example of a prosecutor persuading a judge is a good example of such a context. In *Brady v. Maryland*,[10] the Supreme Court of the United States ruled that a prosecutor violates the Due Process Clause of the Fourteenth Amendment when he fails to disclose material evidence favorable to the accused. Since a prosecutor maximizing convictions would always willingly choose to report any evidence unfavorable to the accused, our assumption that he discloses any evidence, i.e., any signal realization, seems justifiable.[11]

Second, there are settings where Sender cannot avoid learning all the available information but is able to commit to a stochastic disclosure rule $\pi : \Omega \rightarrow \Delta(S)$. Examples include a school choosing a coarse grading policy for students, or a rating agency choosing a scoring procedure for hospitals.

Third, procedures for information gathering in organizations often involve commitment, either formally through contracts, or informally through reputation. Firms commit to the information they will seek out for performance reviews. Academic departments follow fixed rules about the information they will solicit for midterm or tenure reviews.

Fourth, there are many settings where the decision to conduct tests or studies is publicly observable, and the results are verifiable ex post. Tobacco companies, for example, commission third party studies on the health effects of smoking. They can limit the precision of the information they gather and may find it costly to suppress or distort the results after the fact. Medical research provides another, especially clear example. In order to be eligible for consideration by any major peer-reviewed biomedical journal, a randomized clinical trial must be registered in a publicly accessible trials registry before any subjects are enrolled.[12] Similarly, the design of drug trials must be submitted ahead of time to the FDA. In both cases, the design must be registered before any information can be gathered, and deviating from the proposed design renders a study invalid.

Fifth, in many situations, both Sender's choice of the signal and the signal realization are directly observed by Receiver. Examples include a soft drink maker running a public taste test and a software maker providing a limited trial version of its software.

---

[8] $\mathcal{P}(X)$ denotes the set of all subsets of X.

[9] This game is clearly related to persuasion games studied in Milgrom and John Roberts (1986) with an important difference that Sender can ex ante choose to be imperfectly informed.

[10] 373 US 83, 87 (1963).

[11] That said, this Supreme Court ruling amplifies the importance of our assumption that the prosecutor fully controls what information is generated. For example, if an unexpected witness comes forward, *Brady v. Maryland* would require the prosecutor to report this even if doing so is suboptimal for him. The extension of the model to cases where Receiver might obtain additional information (Subsection A) allows us to address this issue.

[12] http://www.icmje.org/publishing_10register.html

Even in settings where Sender cannot commit, our model may still prove useful. In the online Appendix, we show that Sender's gain from persuasion in our model is weakly greater than his gain in any alternative communication game, including all those mentioned in the first paragraph of this subsection. Hence, the results in this paper provide an upper bound on gains from communication that are possible regardless of Sender's commitment power.

II. When Does Sender Benefit from Persuasion?

Corollary 2 provides a necessary and sufficient condition for Sender to benefit from persuasion in terms of the concave closure V. In any problem where we can graph the function $\hat{v}$, it is straightforward to construct V and determine the prior beliefs, if any, at which $V(\mu_0) > \hat{v}(\mu_0)$. In Figure 1, Sender benefits from persuasion for any $\mu_0 \in (\mu_l, \mu_h)$, and does not benefit from persuasion for any $\mu_0 \leq \mu_l$ or $\mu_0 \geq \mu_h$. In Figure 2, Sender benefits from persuasion for any $\mu_0 < 0.5$ —i.e., at any prior belief at which the judge does not convict by default. In this section, we show ways to determine whether $V(\mu_0) > \hat{v}(\mu_0)$ in cases when the state space is too large to graph $\hat{v}$. We first do this in terms of the properties of $\hat{v}$, and then in terms of the primitives of our model, namely Sender and Receiver's preferences and initial beliefs.

Corollaries 1 and 2 tell us that Sender benefits from persuasion if and only if there exists a τ such that $E_\tau(\hat{v}(\mu)) > \hat{v}(E_\tau(\mu))$. Whether this is the case is naturally tied to the concavity or convexity of $\hat{v}$.

REMARK 1: *If $\hat{v}$ is concave, Sender does not benefit from persuasion for any prior. If $\hat{v}$ is convex and not concave, Sender benefits from persuasion for every prior.*

Observe that in the simple case where Sender's payoff does not depend on the state, $\hat{v}(\mu) = v(\hat{a}(\mu))$. The concavity or convexity of $\hat{v}$ then depends on just two things: whether Receiver's action $\hat{a}(\mu)$ is concave or convex in μ, and whether Sender's payoff $v(a)$ is concave or convex in a. If both $\hat{a}$ and v are concave, Sender does not benefit from persuasion. If both $\hat{a}$ and v are convex and at least one of them is not concave, Sender benefits from persuasion.

Often, $\hat{v}$ will be neither convex nor concave. This is true, for example, in our motivating example as shown in Figure 2. As we discussed earlier, the fact that Sender benefits from persuasion in that example hinges on (i) the fact that Receiver does not take Sender's preferred action by default, and (ii) the fact that Receiver's action is constant in a neighborhood around the prior. We now show that these two conditions, suitably generalized, play a crucial role more broadly.

To generalize (i), say *there is information Sender would share* if $\exists\, \mu$ such that

$$(2) \qquad\qquad \hat{v}(\mu) \; > \; E_\mu v\big(\hat{a}(\mu_0), \omega\big).$$

In other words, there is a μ such that, if Sender had private information that led him to believe μ, he would prefer to share this information with Receiver rather than have Receiver act based on μ_0. Note that when v does not depend on ω, there is information Sender would share as long the default action is not dominant, i.e.,

as long as $v(\hat{a}(\mu_0)) < v(a)$ for some $a \in A$. This is the sense in which equation (2) generalizes condition (i).

To generalize (ii), we say Receiver's *preference is discrete* at belief μ if Receiver's expected utility from her preferred action $\hat{a}(\mu)$ is bounded away from her expected utility from any other action, i.e., if there is an $\varepsilon > 0$ s.t. $\forall a \neq \hat{a}(\mu)$, $E_\mu u(\hat{a}(\mu), \omega) > E_\mu u(a, \omega) + \varepsilon$. When A is finite, Receiver's preference is not discrete only if Receiver is exactly indifferent between two distinct actions. Our main result in this section is that the generalization of (i) is necessary, while generalizations of (i) and (ii) are jointly sufficient, for Sender to benefit from persuasion.

PROPOSITION 2: *If there is no information Sender would share, Sender does not benefit from persuasion. If there is information Sender would share and Receiver's preference is discrete at the prior, Sender benefits from persuasion. If A is finite, Receiver's preference is discrete at the prior generically.*

The first part of the proposition is easy to see. If there is no information Sender would share, any realization of any signal leads Receiver to take an action Sender weakly dislikes relative to the default action. Hence, a completely noninformative signal is optimal.

The intuition for the second part is as follows. Because there is information that Sender would share we can find a belief μ_h such that $\hat{v}(\mu_h) > \sum_\omega v(\hat{a}(\mu_0), \omega)\mu_h(\omega)$. Moreover, the discreteness of Receiver's preference implies that there is a belief near the prior, say μ_l, such that $\hat{a}(\mu_l)$ is equal to Receiver's default action and μ_0 is on the segment between μ_l and μ_h. That mixing point μ_l and μ_h produces a strictly positive gain is obvious in a case like the motivating example where Sender's payoff v does not depend on the state. The argument is more subtle when v does depend on the state. The key observation is that for any *given* action by Receiver, Sender's utility is linear in μ. In particular, $\sum_\omega v(\hat{a}(\mu_0), \omega)\,\mu(\omega)$ is linear in μ. This implies that mixing μ_l with μ_h yields a strictly positive gain.

Note that the second part of the proposition is of no use if A is connected. In that case, the continuity of u implies that Receiver's preference cannot be discrete at any belief. In contrast, as the last part of the proposition establishes, when A is finite, the set of beliefs at which Receiver's preference is not discrete is Lebesgue measure zero in $\Delta(\Omega)$. This holds because Receiver is indifferent between two actions only at finitely many beliefs, which means that she is generically not indifferent at the prior. The last part of Proposition 2 is not meant to suggest, however, that there is some form of discontinuity in Sender's benefit from persuasion as we move from large finite choice sets to infinite ones. As the action space becomes large, the gain from persuasion may become arbitrarily small.

Sender's Payoff Depends only on the Expected State.—We have shown that when $\hat{v}$ can be graphed, inspection of the graph can show directly whether Sender benefits from persuasion. The domain of $\hat{v}$, however, is $\Delta(\Omega)$. This means that it is possible to easily depict $\hat{v}$ only when there are two or three states. When there are more states, our propositions still apply, but one cannot approach the problem by simply studying the graph of $\hat{v}$. When Sender's payoff depends only on the expected state, however, a natural conjecture is that we could learn about Sender's gain from

persuasion by graphing Sender's payoff as a function of the expected state $E_\mu[\omega]$ rather than as a function of μ directly. If so, we would have a simple two-dimensional representation of this subclass of problems regardless of the size of the state space.

When Sender's payoff depends only on the expected state, there exists a $\tilde{v} : \mathbb{R} \to \mathbb{R}$ such that $\tilde{v}(E_\mu[\omega]) = \hat{v}(\mu)$. Let $\tilde{V}$ be the concave closure of $\tilde{v}$. The following Proposition shows that the conjecture above is correct: we can determine whether Sender benefits from persuasion simply by inspecting $\tilde{V}$ and $\tilde{v}$. In Section V below, we provide an example of how this result can greatly simplify the analysis of problems with a large state space.

PROPOSITION 3: *When Sender's payoff depends only on the expected state, Sender benefits from persuasion if and only if $\tilde{V}(E_{\mu_0}[\omega]) > \tilde{v}(E_{\mu_0}[\omega])$.*

To see that $\tilde{V}(E_{\mu_0}[\omega]) \leq \tilde{v}(E_{\mu_0}[\omega])$ implies Sender cannot benefit from persuasion, we need only note that for any Bayes-plausible τ, $E_\tau[\hat{v}(\mu)] \leq \tilde{V}(E_{\mu_0}[\omega]) \leq \tilde{v}(E_{\mu_0}[\omega]) = \hat{v}(\mu_0)$. The proof of the converse is more involved. If $\tilde{V}(E_{\mu_0}[\omega]) > \tilde{v}(E_{\mu_0}[\omega])$, we know there is a τ s.t. $E_\tau[E_\mu[\omega]] = E_{\mu_0}[\omega]$ and $E_\tau[\hat{v}(\mu)] > \hat{v}(\mu_0)$. If this τ were Bayes-plausible, we could construct a signal that induces it and we would be done. The trouble is that $E_\tau[E_\mu[\omega]] = E_{\mu_0}[\omega]$ does not guarantee that τ is Bayes plausible. To construct a signal with a strictly positive gain, we show that it is always possible to find a belief μ' such that $E_{\mu'}[\omega] = E_{\mu_0}[\omega]$ and a Bayes-plausible τ' that is a mixture of τ and μ'. Since $E_\tau[\hat{v}(\mu)] > \hat{v}(\mu_0)$ and $\hat{v}(\mu') = \hat{v}(\mu_0)$, we know that $E_{\tau'}[\hat{v}(\mu)] > \hat{v}(\mu_0)$.[13]

III. Optimal Signals

Corollary 2 shows that the value of an optimal signal is $V(\mu_0)$. When it is possible to graph $\hat{v}$ and V, we can read $V(\mu_0)$ and the gain $V(\mu_0) - \hat{v}(\mu_0)$ off the graph directly. The graphs of $\hat{v}$ and V also identify the beliefs induced by the optimal signal—these are the points on $\hat{v}$ whose convex combination yields value $V(\mu_0)$. In Figure 1, we see that the optimal signal induces beliefs μ_l and μ_h. As we explained at the end of Subsection IB, once these beliefs are identified, it easy to compute the optimal signal which generates them.

Determining the optimal signal is also easy when $\hat{v}$ is convex or concave. Say that no disclosure is optimal if $\mu_s = \mu_0$ for all s realized with positive probability under the optimal signal. If, at the other extreme, μ_s is degenerate[14] for all s realized with positive probability under the optimal signal, say that full disclosure is optimal. Then, if $\hat{v}$ is (strictly) concave, no disclosure is (uniquely) optimal, and if $\hat{v}$ is (strictly) convex, full disclosure is (uniquely) optimal. This observation follows directly from the definitions of convexity and concavity.

---

[13] As the detailed proof in Appendix A shows, we can also establish a result somewhat stronger than Proposition 3. Suppose there exists a linear $T : \Delta(\Omega) \to \mathbb{R}^k$ and a $\tilde{v} : \mathbb{R}^k \to \mathbb{R}$ s.t. $\hat{v}(\mu) = \tilde{v}(T\mu)$. Then, Sender benefits from persuasion if and only if $\tilde{v}$ is below its concave closure at $T\mu_0$. We focus on the case where $T\mu = E_\mu[\omega]$ only to simplify the exposition of the result.

[14] We say μ is degenerate if there is an ω s.t. $\mu(\omega) = 1$.

When $\hat{v}$ cannot be visualized and is neither convex nor concave, we can still develop insight about optimal signals by noting general properties of beliefs induced by optimal signals. Recall that in the motivating example: (i) whenever the judge chooses the prosecutor's least-preferred action (*acquit*), she is certain of the state; and (ii) whenever she chooses an action that is not the prosecutor's least-preferred (*convict*), she is indifferent between that action and a worse one.

The first of these properties holds in general. Define an action $\underline{a}$ to be a *worst action* if $v(\underline{a}, \omega) < v(a, \omega)$ for all ω and all $a \neq \underline{a}$. Say that Receiver is *certain of her action* at a belief μ if $\mu(\omega) = 0$ for all ω s.t. $\{\hat{a}(\mu)\} \neq \arg\max u(a, \omega)$. If $a = \hat{a}(\mu)$ say that μ *leads* to a. The proof of the following proposition follows the logic we described in the introduction with regard to the motivating example.

PROPOSITION 4: *If an optimal signal induces a belief μ that leads to a worst action, Receiver is certain of her action at μ.*

The second of the properties above, on the other hand, is not completely general when v depends on ω. Sender's preferences could then take a form such that two actions a_1 and a_2 are both induced by an optimal signal, but at the belief at which Receiver takes a_1 Sender is exactly indifferent between a_1 and a_2. Then the logic described in the introduction—where Sender could improve his payoff by shifting probability mass away from a belief that induced a strict preference for one action in order to increase the probability of an alternative action—need not hold, since Sender is exactly indifferent to this change. However, if this occurs under a particular utility function v, it will not occur for a slight perturbation of v. The following assumption rules out this pathological case.

ASSUMPTION 1: *There exists no action a s.t. (i) $\forall \mu$, $\hat{v}(\mu) \leq E_\mu v(a, \omega)$ and (ii) $\exists\ \mu$ s.t. $a \neq \hat{a}(\mu)$ and $\hat{v}(\mu) = E_\mu v(a, \omega)$.*

The intuition that we provided for property (ii) applies to any interior belief. The same intuition, however, also applies to any belief which leads to an action that Sender cannot improve upon. Say an action a is *best attainable* if $E_\mu v(a, \omega) > \hat{v}(\mu)$ for all μ s.t. $a \neq \hat{a}(\mu)$.

PROPOSITION 5: *Suppose Assumption 1 holds and Sender benefits from persuasion. If a belief μ induced by an optimal signal is either (i) interior or (ii) leads to a best-attainable action, then Receiver's preference is not discrete at μ.*

Recall that when the action space is finite, Receiver's preference is not discrete at μ means that she is indifferent between two actions at μ. Hence, in the motivating example, Proposition 5 implies that the judge is indifferent between convicting and acquitting when she convicts.

The motivating example suggests an additional property of the optimal signal. Because the prosecutor's payoff is (weakly) increasing in the judge's posterior belief that the state is *guilty*, it is meaningful to talk about beliefs that place more weight on *innocent* as being "worse" from the prosecutor's perspective. A different way to look at properties (i) and (ii) is that the prosecutor chooses an investigation

that induces the worst possible belief consistent with a given action by the judge—certainty of innocence when the action is *acquit*, and indifference when the action is *convict*. In the online Appendix, we show that when Sender's payoffs are monotonic in Receiver's beliefs, there is a general sense in which Sender always induces the worst belief consistent with a given action.

Sender's Payoff Depends only on the Expected State.—Recall that we established earlier that when Sender's payoff depends only on the expected state, there is a function $\tilde{v}$ s.t. $\tilde{v}(E_\mu[\omega]) = \hat{v}(\mu)$, and Sender benefits from persuasion if and only if $\tilde{V}(E_{\mu_0}[\omega]) > \tilde{v}(E_{\mu_0}[\omega])$. We might conjecture that the value of an optimal signal is $\tilde{V}(E_{\mu_0}[\omega])$. This conjecture turns out to be false. Recall from the discussion of Proposition 3 that even though we know there is always a τ s.t. $E_\tau[E_\mu[\omega]] = E_{\mu_0}[\omega]$ and $E_\tau[\hat{v}(\mu)] = \tilde{V}(E_{\mu_0}[\omega])$, such τ need not be Bayes plausible. In order to show that Sender could benefit from persuasion, we had to mix the beliefs in the support of τ with another belief μ' such that $E_{\mu'}[\omega] = E_{\mu_0}[\omega]$. This reduces the value of the signal strictly below $\tilde{V}(E_{\mu_0}[\omega])$.

For a concrete example, suppose that $A = [0, 1]$, $\Omega = \{-1, 0, 1\}$, $u(a, \omega) = -(a - \omega)^2$, and $v(a, \omega) = a^2$. In this case, $\hat{v}(\mu) = \tilde{v}(E_\mu[\omega]) = (E_\mu[\omega])^2$. Hence, $\tilde{V}$ is constant at 1 and in particular $\tilde{V}(E_{\mu_0}[\omega]) = 1$. Yet, whenever the prior puts any weight on $\omega = 0$, the value of any mechanism is strictly less than 1.

Hence, when Sender's payoff depends only on the expected state, we can use $\tilde{v}$ to determine whether Sender benefits from persuasion, but $\tilde{V}(E_{\mu_0}[\omega])$ is only an upper bound on the value of an optimal signal. To characterize the optimal signal more precisely, we need to study $\hat{v}$ directly or derive its properties from Propositions 4 and 5.

IV. Preference Alignment

We say that preferences (u', v') are *more aligned* than (u, v) if for any a and any μ

$$E_\mu v(\hat{a}(\mu), \omega) \geq E_\mu v(a, \omega) \Rightarrow$$

$$E_\mu v'(\hat{a}'(\mu), \omega) \geq E_\mu v'(a, \omega),$$

where $\hat{a}'(\mu)$ denotes Receiver's equilibrium action when her belief is μ and her utility is u'. Note that the functional form used to index alignment in the main example of Crawford and Sobel (1982) satisfies this definition.

It is easy to establish that Sender necessarily obtains higher utility if preferences are more aligned. Consider moving from (u, v) to a more aligned (u', v). We hold v constant so that Sender's utility is measured on a constant scale. Plugging in $a = \hat{a}(\mu)$ to the definition above, it is immediate that $E_\mu v(\hat{a}'(\mu), \omega) \geq E_\mu v(\hat{a}(\mu), \omega)$ for any μ. This means $\hat{v}'(\mu) \geq \hat{v}(\mu)$ for all μ, so the value of the optimal signal is weakly higher under (u', v) than under (u, v). Note, however, that when greater alignment changes the default action, it is possible for the gain from persuasion to be lower under more aligned preferences.

The impact of alignment on the amount of information communicated in equilibrium is also ambiguous. On the one hand, the more Receiver responds to information

in a way consistent with what Sender would do, the more Sender benefits from providing information. On the other hand, when preferences are more aligned Sender can provide less information and still sway Receiver's action in a desirable direction. Hence, making preferences more aligned can make the[15] optimal signal either more or less informative.

V. Examples

A. *Lobbying*

Consider a setting where a lobbying group commissions a study with the goal of influencing a benevolent politician. Such studies are common in practice. The tobacco lobby has spent large sums funding studies about the health effects of smoking (Joaquin Barnoya and Stanton A. Glantz 2006). Drug companies spend over $14 billion a year funding clinical trials of their products (Hamilton Moses III et al. 2005). Would it make sense for the lobbyist to commission such studies even if the politician is rational and knows the lobbyist is providing information with the goal of influencing her decision? Would the optimal study in this case be biased toward supporting the lobbyist's position or fully revealing of the true state?

Here, our commitment assumption requires that the information the lobbyist gathers is observable to the politician, and that the lobbyist can make verifiable claims about what the information shows. Recall from Section C that it is not necessary to assume the lobbyist truthfully reports everything he knows, so long as any claims he does make are verifiable.

We assume that the politician (Receiver) chooses a unidimensional policy $a \in [0, 1]$. The state $\omega \in [0, 1]$ is the socially optimal policy. The lobbyist (Sender) is employed by an interest group whose preferred action $a^* = \alpha\omega + (1 - \alpha)\omega^*$, with $\alpha \in [0, 1]$, depends on ω but is biased toward a specific policy $\omega^* > 1$. The politician's payoff is $u = -(a - \omega)^2$ and the lobbyist's payoff is $v = -(a - a^*)^2$. We are interested in the way the equilibrium depends on the extent of preference disagreement, as captured by α and ω^*. It is easy to see that an increase in α or a decrease in ω^* makes preferences more aligned according the definition in Section IV.

Since $u = -(a - \omega)^2$ we know that $\hat{a}(\mu) = E_\mu[\omega]$. Given this $\hat{a}$, simple algebra reveals that

$$\hat{v}(\mu) = -(1 - \alpha)^2 \omega^{*2} + 2(1 - \alpha)^2 \omega^* E_\mu[\omega] - \alpha^2 E_\mu[\omega^2]$$

$$+ (2\alpha - 1)(E_\mu[\omega])^2.$$

The expectation of $-(1 - \alpha)^2\omega^{*2} + 2(1 - \alpha)^2\omega^* E_\mu[\omega] - \alpha^2 E_\mu[\omega^2]$ is constant across all Bayes-plausible τ's, so we can treat $\hat{v}(\mu)$ as a constant plus $(2\alpha - 1)(E_\mu[\omega])^2$.

We can now easily solve for the optimal signal; $\hat{v}$ is linear in μ when $\alpha = \frac{1}{2}$, strictly convex when $\alpha > \frac{1}{2}$, and strictly concave when $\alpha < \frac{1}{2}$. From the discussion in Section III, this implies that when $\alpha > \frac{1}{2}$, i.e., the lobbyist's preferences are

---

[15] This remark assumes the optimal signal is unique. When there are multiple optimal signals, there is an additional complication of comparing the informativeness of sets of signals.

sufficiently aligned with those of the politician, full disclosure is uniquely optimal.[16] When $\alpha < \frac{1}{2}$, i.e., the lobbyist's and the politician's preferences diverge significantly, no disclosure is uniquely optimal. When $\alpha = \frac{1}{2}$ all signals yield the same value. There is thus a natural sense in which some alignment of preferences is necessary for information to be communicated in equilibrium even when Sender has the ability to commit.

Note, however, that the optimal signal is independent of ω^*. This is important because ω^* also captures a form of disagreement between the lobbyist and the politician. We might have expected communication to be difficult when ω^* is much greater than one. Unlike α, however, ω^* does not affect the way the lobbyist's payoff *varies* across realizations of a signal. The loss the lobbyist suffers from high values of ω^* is thus a sunk cost and does not affect the decision of how best to persuade. This is another illustration of the ambiguous effect of preference alignment in our setting.[17]

An interesting feature of this example is that the lobbyist either commissions a fully revealing study or no study at all. This contrasts with the observation that industry-funded studies often seem to produce results more favorable to the industry than independent studies. The model suggests that commissioning such biased studies when policymakers are rational may not be optimal from the industry's perspective. If such studies are optimal, it is likely that the intended constituency is not fully rational, the lobbyists cannot commit not to distort study results ex post, or the payoff functions differ from those in this simple example in a way that makes asymmetric signals optimal.

B. *Supplying Product Information*

Firms often provide information to consumers to help them learn about the match between their preferences and the characteristics of the firm's products. Car dealers allow consumers to test drive cars. Pharmaceutical advertising informs consumers of the conditions for which a particular drug is likely to be effective. Software producers allow consumers to download trial versions of their products. In each of these cases, firms decide what kind of signal consumers can obtain—the length and frequency of test drives, the details included in ads, the number of features that are available in the free version of software, and so on.

Our commitment assumption is natural in this setting because the consumer observes the signal and its realization directly. The assumption of common priors is

---

[16] Even if we remove the restriction that $\alpha \in [0, 1]$, it remains the case that $\hat{v}$ is convex whenever $\alpha > \frac{1}{2}$. Hence, even if the interest group's preferences differ greatly from socially optimal ones by being overly sensitive to ω, full disclosure remains uniquely optimal.

[17] This observation also highlights a difference between our setting and delegation models where Receiver has commitment power (e.g., Bengt R. Holmstrom 1984, Wouter Dessein 2002, Ricardo Alonso and Niko Matouschek 2008). In those models, Sender has no commitment power, but Receiver commits ex ante how she will respond to messages from Sender. Nahum D. Melumad and Toshiyuki Shibano (1991) analyze such a game with the same preferences as in this example and show that the amount of information communicated in equilibrium depends on ω^* as well as α. Hence, the way in which preference disagreement affects communications depends on the nature of commitment. Note that when there is a worst action for Sender and both parties have complete commitment, full disclosure is always an equilibrium: Receiver would commit to take that worst action at any nondegenerate belief and thus force Sender to send a fully informative signal.

also natural when the state is the match between a particular consumer's preferences and a product rather than the product's overall average quality.

Tracy R. Lewis and David E. M. Sappington (1994), Simon P. Anderson and Regis Renault (2006), and Justin P. Johnson and David P. Myatt (2006) study the question of how much information firms should provide to their potential customers. Here, we derive the optimal signal in a simple version of the Lewis and Sappington (1994) model.[18]

A firm (Sender) faces a single consumer (Receiver) who decides whether to buy one unit of the firm's product or not. If the consumer does not purchase, she gets an outside option which yields utility $\underline{u} \in [0, 1]$. The consumer gets utility $\omega \in [0, 1]$ from buying the product. The state ω indexes the match quality between consumer's taste and the product (holding constant the price and other characteristics which we treat as exogenous). The firm and the consumer share a prior μ_0 about ω. The consumer is risk neutral and hence will buy the product if and only if $E_\mu[\omega] \geq \underline{u}$ where μ is her posterior belief about ω. To make things interesting, we suppose that the default action is not to buy, i.e., $E_{\mu_0}[\omega] < \underline{u}$. The firm chooses the type of interaction that the consumer can have with the product before deciding whether to purchase it. Formally, it chooses a signal $\pi \colon [0, 1] \to \Delta(S)$ that the consumer can observe about ω.

We again begin by computing $\hat{v}$. Denoting the decision to buy with 1 and the alternative with 0, we have

$$\hat{a}(\mu) = \begin{cases} 0 & \text{if } E_\mu[\omega] < \underline{u} \\ 1 & E_\mu[\omega] \geq \underline{u} \end{cases}$$

$$\hat{v}(\mu) = \begin{cases} 0 & \text{if } E_\mu[\omega] < \underline{u} \\ 1 & E_\mu[\omega] \geq \underline{u} \end{cases}.$$

In this example, $\hat{v}$ is difficult to visualize. However, since Sender's payoff depends only on the expected state, we can depict $\tilde{v}$ and $\tilde{V}$ (Figure 3). Since $\tilde{V}(E_{\mu_0}[\omega]) > \tilde{v}(E_{\mu_0}[\omega])$, Proposition 3 tells us there exists an advertising campaign that increases the firm's revenue. Also, as we discussed at the end of Section III, while we cannot determine the optimal campaign by examining $\tilde{v}$, we know that $\tilde{V}(E_{\mu_0}[\omega])$ is an upper bound on the market share that can be achieved by any advertising campaign.

Although we cannot easily solve for the optimal signal in this example, the propositions in Section III allow us to characterize it. An optimal signal induces two possible beliefs, whose expectations are illustrated by $E_{\mu_l}[\omega]$ and $E_{\mu_h}[\omega]$ in Figure 3.[19] We know that any consumer who buys the product will be just indifferent between buying and not buying (Proposition 5), so that $E_{\mu_h}[\omega] = \underline{u}$. We also know

---

[18] The Lewis and Sappington (1994) model is more general than the one below because they endogenize prices and allow nonunit demand. However, they allow firms to select from a limited, parameterized set of signals, whereas we study the optimal choice from the set of all possible signals.

[19] Recall that an optimal signal need not have more realizations than $|A|$.

Panel A. Function $\tilde{v}$ Panel B. Function $\tilde{V}$
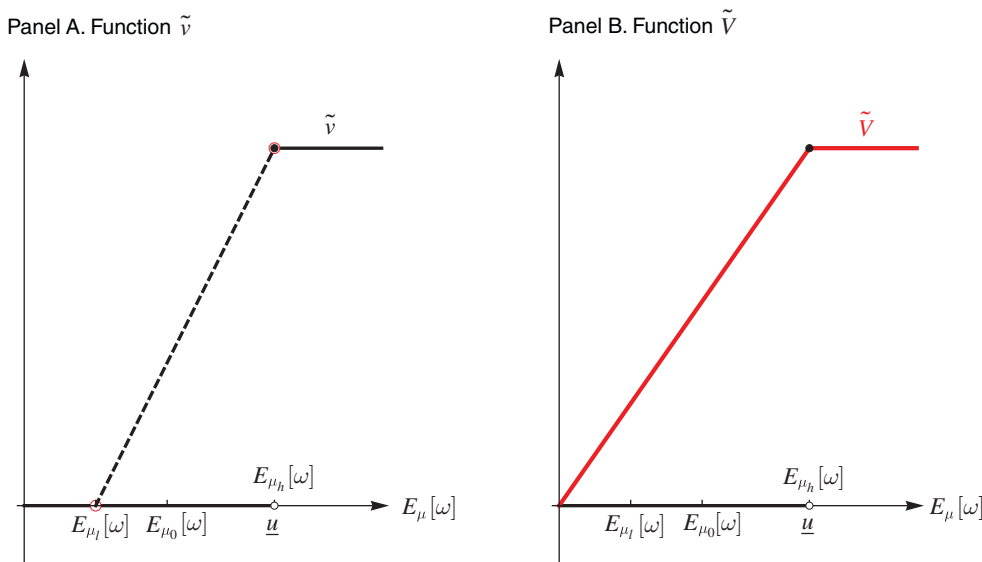

FIGURE 3. ADVERTISING TO INCREASE SALES

that any consumer who does not buy will have a posterior that puts zero weight on the possibility that $\omega \geq \underline{u}$ (Proposition 4), so that $\mathrm{Supp}(\mu_l) \subseteq [0, \underline{u})$. Note that, as discussed in the final part of Section III, we cannot simply read the value of $E_{\mu_l}[\omega]$ off the figure and assume $E_{\mu_l}[\omega] = 0$. For most priors, μ_l will have to assign positive weight to $\omega \in (0, \underline{u})$ so that $E_{\mu_l}[\omega] > 0$. By Bayes plausability, since $E_{\mu_h}[\omega] = \underline{u} > E_{\mu_0}[\omega]$, we must have $E_{\mu_l[\omega]} < E_{\mu_0}[\omega]$.

The firm's optimal strategy is thus to allow a trial which separates consumers into two groups: those who are sure the product is not for them, and those who are just positive enough that they will choose to buy. A car salesman, for example, might want to explain frankly that certain classes of buyers should not buy a given car; the credibility that he gains by doing so will then increase the willingness to pay of the remaining customers, ideally just to the point that they will be willing to purchase. Such optimal information provision allows the firm to extract all the consumer surplus. Note that this would not be the case if, in line with previous literature, we had restricted our attention to a parameterized subset of signals.

VI. Extensions

A. *Receiver's Private Information*

Extending our analysis to situations where Receiver has private information is straightforward. Suppose that Sender and Receiver share a prior μ_0 at the outset of the game, and then Receiver privately observes a realization $r \in R$ from some signal $\chi(\cdot|\omega)$. Sender then chooses a signal; Receiver observes its realization and takes her action. We assume that Receiver cannot communicate any of her private information to Sender before Sender selects his signal.

The only way in which Receiver's private information changes our analysis is that we can no longer construct a deterministic function $\hat{a}(\mu)$ which specifies Receiver's action at any belief she shares with Sender. Rather, for any Sender's belief μ, Receiver's optimal action $\hat{a}(\mu, r)$ depends on the realization of her private signal and so is stochastic from Sender's perspective. When his posterior is μ, Sender assigns probability $\chi(r|\omega)\,\mu(\omega)$ to the event that Receiver's signal is r and the state is ω. Hence, Sender's expected utility when his posterior is μ is:

$$(3) \qquad \hat{v}(\mu) \ = \ \sum_{\omega \in \Omega} \sum_{r \in R} v\big(\hat{a}(\mu, r), \omega\big)\ \chi(r|\omega)\ \mu(\omega).$$

A different situation that involves private information is when Receiver's preferences depend on some parameter $\theta \in \Theta$ which is unrelated to ω. Again, the impact of private information is that Receiver's optimal action $\hat{a}(\mu, \theta)$ is stochastic from Sender's perspective. We can therefore define $\hat{v}$ by integrating over this uncertainty, similarly to equation (3).

Finally, suppose that Receiver can gather some additional information following the realization of Sender's signal. For example, in our motivating example, when the judge observes signal realization g, gathering even a small amount of additional information would increase her payoff. In this case $\hat{v}(\mu)$ will be Sender's expected payoff integrating over the information that Receiver gathers following a signal realization that induces μ.

In all of these cases, our approach applies directly with respect to the reformulated $\hat{v}$. In particular, our key simplifying results—Proposition 1, Corollary 1, Corollary 2—still hold. Aside from the fact that constructing $\hat{v}$ is slightly more complicated, the analysis of the problem in terms of the properties of $\hat{v}$ and its concave closure V proceeds exactly as before. However, because private information will "smooth out" $\hat{v}$, some of our characterization results, such as that Receiver's preference is never discrete at any interior μ induced by an optimal signal, will no longer hold.

B. *Multiple Receivers*

In many settings of interest—politicians persuading voters, firms advertising to consumers, auctions—our assumption that there is a single Receiver is unrealistic. Suppose there are n receivers. For ease of exposition we maintain our common prior assumption, which in this setting means that Sender and all receivers share a prior μ_0 over Ω. Sender's utility is now a function of each receiver's action: $v(a_1,...,a_n,\omega)$.

There are two classes of multiple-receiver models where our results can be extended quite easily. The first is one where Sender sends separate (possibly correlated) messages to each receiver, Sender's utility is separable across receivers' actions, and each receiver cares only about her own action. In this case, we can simply apply our approach separately to Sender's problem vis-à-vis each receiver. The second class is where Sender can persuade only by revealing *public* information. In this case, we can compute the outcome following any (common) posterior μ and set $\hat{v}(\mu)$ to Sender's expected utility from that outcome.

There is an important third class of multiple-receiver models, however, where our results do not extend easily: those where the receivers care about each other's

actions and Sender can send private signals to individual receivers. The crucial problem with this case is that for a given set of beliefs that receivers hold after observing their messages, the receivers' actions may vary as a function of the *signal* that produced those beliefs. In an auction, for example, a bidder with a given belief may behave differently if she believes that other bidders are receiving highly informative signals than if she believes they are receiving uninformative signals. This means that the key simplifying step in our analysis—reducing the problem of finding an optimal signal to one of maximizing over distributions of posterior beliefs—does not apply.

VII. Conclusion

There are two ways to induce a person to do something. One is to provide incentives, by which we mean anything which changes marginal utility—explicit payments, coercion, or supply of complementary goods. The other is to persuade, by which we mean anything which changes beliefs.

In this paper, we study persuasion in a setting where both Sender and Receiver are rational Bayesians. Perhaps surprisingly, the scope for persuasion under rationality is large. Our results characterize the form persuasion should take in such a setting. If persuasion in the real world departs from this characterization—for example, if Receivers are systematically harmed by persuasive activity—this suggests that limited rationality may be at play.

Appendix: Proofs

A. *Proof of Proposition* 1

By definition, (ii) implies (i) and (iii). We first show that (i) implies (ii). Given a signal π with value v^*, let $S^a = \{s \mid \hat{a}(\mu_s) = a\}$ for each a. Consider a signal with $S' = A$ and $\pi'(a \mid \omega) = \sum_{s \in S^a} \pi(s \mid \omega)$. In other words, π' "replaces" each signal realization with a recommendation of the action that the signal realization induced. Since a was an optimal response to each $s \in S^a$, it must also be an optimal response to the realization a from π'. Hence, the distribution of Receiver's actions conditional on the state under π' is the same as under π. It remains to show that (iii) implies (i). Consider any v^* and any Bayes-plausible τ s.t. $E_\tau \hat{v}(\mu) = v^*$. Since Ω is finite, Caratheodory's Theorem implies there exists a Bayes-plausible τ^* with finite support such that $E_{\tau^*} \hat{v}(\mu) = v^*$. Now, define S so that $\text{Supp}(\tau^*) = \{\mu_s\}_{s \in S}$ and let $\pi(s \mid \omega) = \mu_s(\omega) \tau^*(\mu_s)/\mu_0(\omega)$. Then simple algebra shows that π is indeed a signal and induces τ^*.

B. *Proof of Proposition* 2

Suppose there is no information Sender would share: $\forall \mu,\ \hat{v}(\mu) \leq \sum_\omega v(\hat{a}(\mu_0), \omega)\mu(\omega)$. Given a signal π that induces some τ, its value is $\sum_{s \in S} \tau_s \hat{v}(\mu_s) \leq \sum_{s \in S} \tau_s \sum_\omega v(\hat{a}(\mu_0), \omega)\mu_s(\omega) = \hat{v}(\mu_0)$. Hence, Sender does not benefit from persuasion.

Now, suppose there is information Sender would share and Receiver's preference is discrete at the prior. Since u is continuous in ω, $\sum u(\hat{a}(\mu_0), \omega)\mu(\omega)$ is continuous

in μ. Therefore, since Receiver's preference is discrete at the prior, $\exists \delta > 0$ s.t. for all μ in a δ-ball around μ_0, $\hat{a}(\mu) = \hat{a}(\mu_0)$. Denote this ball by B_δ. Since there is information Sender would share, $\exists \mu_h$ s.t. $\hat{v}(\mu_h) > \sum_\omega v(\hat{a}(\mu_0), \omega) \mu_h(\omega)$. Consider a ray from μ_h through μ_0. Since μ_0 is not on the boundary of $\Delta(\Omega)$, there exists a belief on that ray, μ_l s.t. $\mu_l \in B_\delta$ and $\mu_0 = \gamma \mu_l + (1 - \gamma)\mu_h$ for some $\gamma \in (0, 1)$. Now, consider the Bayes-plausible distribution of posteriors $\tau(\mu_l) = \gamma, \tau(\mu_h) = 1 - \gamma$. Since $\hat{a}(\mu_0) = \hat{a}(\mu_l), \hat{v}(\mu_l) = \sum_\omega v(\hat{a}(\mu_0), \omega) \mu_l(\omega)$. Hence, $E_\tau[\hat{v}(\mu)] = \gamma \hat{v}(\mu_l) + (1 - \gamma)\hat{v}(\mu_h) > \gamma\sum_\omega v(\hat{a}(\mu_0), \omega)\mu_l(\omega) + (1 - \gamma)\sum_\omega v(\hat{a}(\mu_0), \omega)\mu_h(\omega) = \hat{v}(\mu_0)$. Therefore, Sender benefits from persuasion.

It remains to show that if A is finite, Receiver's preference is discrete at the prior generically. Suppose A is finite. We begin with the following Lemma:

LEMMA 1: *If Receiver's preference at a belief μ is not discrete, there must be an action $a \neq \hat{a}(\mu)$ such that $\sum u(\hat{a}(\mu), \omega)\mu(\omega) = \sum u(a, \omega)\mu(\omega)$.*

PROOF:
Suppose there is no such action. Then, we can define an $\varepsilon > 0$ by

$$\varepsilon = \frac{1}{2} \min_{a \neq \hat{a}(\mu)} \left\{ \sum u(\hat{a}(\mu), \omega) \; \mu(\omega) - \sum u(a, \omega) \; \mu(\omega) \right\}.$$

Since A is finite, the minimum is obtained. But then, $\sum u(\hat{a}(\mu), \omega) \; u(\omega) > \sum u(a, \omega) \; \mu(\omega) + \varepsilon \, \forall a \neq \hat{a}(\mu)$, which means that Receiver's preference is discrete at μ.

Given this Lemma, since there are only finitely many pairs of actions a, a', and since the union of a finite number of measure-zero sets has measure zero, it will suffice to show that given any distinct a and a', the set $\{\mu \, | \sum u(a, \omega) \; \mu(\omega) = \sum u(a', \omega) \; \mu(\omega)\}$ has measure zero. Given any distinct a and a', index states by i and let $\beta_i = u(a, \omega_i) - u(a', \omega_i)$. Let $\beta = [\beta_1, ..., \beta_n]$ and $\mu = [\mu(\omega_1), ..., \mu(\omega_n)]$. We need to show that the set $\{\mu \, | \, \beta'\mu = 0\}$ has measure zero. Recall that for any action a there exists a μ s.t. $a^*(\mu) = \{a\}$. That means that there is necessarily an ω s.t. $u(a, \omega \neq u(a', \omega)$. Hence, there is at least one $\beta_i \neq 0$. Therefore, β is a linear transformation of rank 1. Hence, the kernel of β is a vector space of dimension $n - 1$. Therefore, $\{\mu \, | \, \beta'\mu = 0\}$ is measure zero with respect to the Lebesgue measure on $\mathbb{R}^n$.

C. *Proof of Proposition* 3

As we mentioned in footnote 13, we will establish a somewhat stronger proposition which implies Proposition 3. Suppose that there exists a linear transformation $T: \Delta(\Omega) \to \mathbb{R}^k$ s.t. $\hat{v}(\mu) = \tilde{v}(T\mu)$. Let $\tilde{V}$ denote the concave closure of $\tilde{v}$. Then,

PROPOSITION 6: *Sender benefits from persuasion if and only if $\tilde{V}(T\mu_0) > \tilde{v}(T\mu_0)$.*

PROOF:
Suppose $\tilde{V}(T\mu_0) > \tilde{v}(T\mu_0)$. That implies there exists a z s.t. $z > \tilde{v}(T\mu_0)$ and $(T\mu_0, z) \in co(\tilde{v})$. Hence, there exists a set $(t_i)_{i=1}^{k+1}$ $w/t_i \in \text{Image}(T)$ and weights $\gamma \in \Delta^{k+1}$ s.t. $\sum_i \gamma_i t_i = T\mu_0$ and $\sum_i \gamma_i \tilde{v}(t_i) > \tilde{v}(T\mu_0)$. For each i, select any μ_i

from $T^{-1}t_i$. Let $\mu_a = \sum_{i=1}^{k+1} \gamma_i \mu_i$. Since T is linear, $T\mu_a = T\sum_i \gamma_i \mu_i = \sum_i \gamma_i T\mu_i$ $= \sum_i \gamma_i t_i = T\mu_0$. Since μ_0 is not on the boundary of Δ^n, there exists a belief μ_b and a $\lambda \in (0, 1)$ s.t. $\lambda \mu_a + (1 - \lambda)\mu_b = \mu_0$. Since T is linear, $T\mu_b = 1/(1 - \lambda)(T\mu_0 - \lambda T\mu_a)$. Therefore, since $T\mu_a = T\mu_0$, we have $T\mu_b = T\mu_0$. Hence, $\tilde{v}(T\mu_0) = \tilde{v}(T\mu_b)$. Now, consider a signal that induces the distribution of posteriors $\tau(\mu_i) = \lambda \gamma_i$ for $i = 1, \dots, k + 1$ and $\tau(\mu_b) = 1 - \lambda$. Since $\mu_a = \sum_{i=1}^{k+1} \gamma_i \mu_i$ and $\lambda \mu_a + (1 - \lambda)\mu_b = \mu_0$, this τ is Bayes-plausible. The value of a signal that induces this τ is

$$\sum_i \lambda \gamma_i \hat{v}(\mu_i) + (1 - \lambda)\hat{v}(\mu_b)$$

$$= \lambda \sum_i \gamma_i \tilde{v}(t_i) + (1 - \lambda)\tilde{v}(T\mu_0)$$

$$> \lambda \tilde{v}(T\mu_0) + (1 - \lambda)\tilde{v}(T\mu_0)$$

$$= \hat{v}(\mu_0).$$

Hence, Sender benefits from persuasion. Now suppose $\tilde{V}(T\mu_0) \leq \tilde{v}(T\mu_0)$. For any Bayes-plausible distribution of posteriors τ, $E_\tau[\mu] = \mu_0$ implies $E_\tau[T\mu] = T\mu_0$, so $E_\tau[\hat{v}(\mu)] = E_\tau[\tilde{v}(T\mu)] \leq \tilde{V}(T\mu_0) \leq \tilde{v}(T\mu_0) = \hat{v}(\mu_0)$. Hence, by Corollary 1 Sender does not benefit from persuasion.

D. *Proof of Proposition* 4

Suppose that $v(\underline{a}, \omega) < v(a, \omega)$ for all ω and all $a \neq \underline{a}$. Let $\underline{\Omega} = \{\omega | \hat{a}(\mu_\omega) = \underline{a}\}$, where $\mu_\omega(\omega) = 1$. Let $\overline{\Omega}$ be the complement of $\underline{\Omega}$. Suppose contrary to Proposition 4 that an optimal signal induces τ and there is a belief μ' s.t. $\tau(\mu') > 0$, $\hat{a}(\mu') = \underline{a}$ and $\exists \overline{\omega} \in \overline{\Omega}$ s.t. $\mu'(\overline{\omega}) > 0$. We can express μ' as a convex combination of $\overline{\mu}$ and $\underline{\mu}$

where $\overline{\mu}(\overline{\omega}) = 1$ and $\underline{\mu}(\omega) = \begin{cases} \mu'(\omega)/(1 - \mu'(\overline{\omega})) & \text{if } \omega \neq \overline{\omega} \\ 0 & \text{if } \omega = \overline{\omega} \end{cases}$. If we "replace"

μ' with a mixture of $\overline{\mu}$ and $\underline{\mu}$, this will yield a higher value since $\overline{\mu}$ induces an action Sender strictly prefers to $\underline{a}$ while $\hat{a}(\underline{\mu})$ cannot be any worse for Sender than $\hat{a}(\mu') = \underline{a}$. Formally, consider the following distribution of beliefs:

$$\tau^*(\overline{\mu}) = \mu'(\overline{\omega})\tau(\mu') + \tau(\overline{\mu})$$

$$\tau^*(\underline{\mu}) = (1 - \mu'(\overline{\omega}))\tau(\mu') + \tau(\underline{\mu})$$

$$\tau^*(\mu) = \tau(\mu) \text{ if } \mu \notin \{\mu', \underline{\mu}, \overline{\mu}\}.$$

Simple algebra reveals that τ^* is Bayes plausible and yields a higher value than τ does.

E. *Proof of Proposition* 5

LEMMA 2: *If μ' is induced by an optimal signal, $V(\mu') = \hat{v}(\mu')$.*

PROOF:

Suppose that an optimal signal induces τ and there is some μ' s.t. $\tau(\mu') > 0$ and $V(\mu') > \hat{v}(\mu')$. Since $(\mu', V(\mu')) \in co(\hat{v})$, there exists a distribution of posteriors τ' such that $E_{\tau'}\mu = \mu'$ and $E_{\tau'}\hat{v}(\mu) = V(\mu')$. But then we can then take all the weight from μ' and place it on τ' which would yield higher value while preserving Bayes plausibility. Formally, consider the distribution of posteriors

$$\tau^*(\mu) = \begin{cases} \tau(\mu')\,\tau'(\mu) & \text{if } \mu \in \text{Supp}(\tau')\backslash\text{Supp}(\tau) \\ \tau(\mu) + \tau(\mu')\,\tau'(\mu) & \text{if } \mu \in \text{Supp}(\tau') \cap \text{Supp}(\tau) \\ \tau(\mu) & \text{if } \mu \in \text{Supp}(\tau)\backslash(\text{Supp}(\tau') \cup \{\mu'\}). \end{cases}$$

By construction, τ^* is Bayes plausible and yields a higher value than τ does.

LEMMA 3: *Suppose μ_l and μ_r are induced by an optimal signal and $\mu_m = \gamma\mu_l + (1 - \gamma)\mu_r$ for some $\gamma \in [0, 1]$. Then, $\hat{v}(\mu_m) \leq \gamma\hat{v}(\mu_l) + (1 - \gamma)\hat{v}(\mu_r)$.*

PROOF:

Suppose to the contrary that τ is induced by an optimal signal, $\tau(\mu_l), \tau(\mu_r) > 0$, and $\hat{v}(\mu_m) > \gamma\hat{v}(\mu_l) + (1 - \gamma)\hat{v}(\mu_r)$. Then we can take some weight from μ_l and μ_r and place it on μ_m which would yield higher value while preserving Bayes plausibility. Formally, pick any $\varepsilon \in (0, 1)$. Let $\varepsilon' = \varepsilon\tau(\mu_l)/\tau(\mu_r)$. Consider an alternative τ^* defined by:

$$\tau^*(\mu_l) = (1 - \gamma\varepsilon)\tau(\mu_l)$$

$$\tau^*(\mu_r) = (1 - (1 - \gamma)\varepsilon')\tau(\mu_r)$$

$$\tau^*(\mu_m) = \tau(\mu_m) + \varepsilon\tau(\mu_l)$$

$$\tau^*(\mu) = \tau(\mu) \text{ if } \mu \notin \{\mu_l, \mu_m, \mu_r\}.$$

Simple algebra reveals that τ^* is Bayes plausible and yields a higher value than τ does.

Say that action a is induced dominant if $\forall \mu, \hat{v}(\mu) \leq \sum_\omega v(a, \omega)\mu(\omega)$. Say that a is strictly induced dominant if $\forall \mu$ s.t. $a \neq \hat{a}(\mu), \hat{v}(\mu) < \sum_\omega v(a, \omega)\mu(\omega)$. Say that a is weakly but not strictly dominant (wnsd) if it is induced dominant and $\exists \mu$ s.t. $a \neq \hat{a}(\mu)$, and $\hat{v}(\mu) = \sum_\omega v(a, \omega)\mu(\omega)$. Note that there is information Sender would share if and only if $\hat{a}(\mu_0)$ is not induced dominant, and that Assumption 1 states that there are no wnsd actions.

LEMMA 4: *Suppose that Assumption 1 holds. Let μ be an interior belief induced by an optimal signal. Then, either:* (*i*) *Receiver's preference at μ is not discrete, or* (*ii*) *$\hat{a}(\mu)$ is strictly induced dominant.*

PROOF:

Suppose that Assumption 1 holds and μ is an interior belief induced by an optimal signal. Now, suppose Receiver's preference at μ is discrete. By Proposition 2, we know that if μ were the prior either: (i) there would be no information Sender would want to share, i.e., $\hat{a}(\mu)$ is induced dominant; or (ii) Sender would benefit from persuasion. But, Sender could not benefit from persuasion if μ were the prior because by Lemma 2 we know $V(\mu) = \hat{v}(\mu)$. Thus, $\hat{a}(\mu)$ is induced dominant so by Assumption 1 it is strictly induced dominant.

LEMMA 5: *Suppose Sender benefits from persuasion, μ is an interior belief induced by an optimal signal, and $\hat{a}(\mu)$ is strictly induced dominant. Then, Receiver's preference at μ is not discrete.*

PROOF:

Suppose Sender benefits from persuasion, μ is an interior belief induced by an optimal signal, and $\hat{a}(\mu)$ is strictly induced dominant. Since the set of beliefs that induces any particular action is convex, when Sender benefits from persuasion, any optimal signal must induce at least two distinct actions. Therefore, there must be a μ' induced by the signal at which $\hat{a}(\mu) \neq \hat{a}(\mu')$. Now, suppose that Receiver's preference at μ is discrete. Then, there is an $\varepsilon > 0$ s.t. $\hat{a}(\varepsilon\mu' + (1 - \varepsilon)\mu) = \hat{a}(\mu)$. Let $\mu_m = \varepsilon\mu' + (1 - \varepsilon)\mu$. Since both μ and μ' are induced by an optimal signal, Lemma 3 tells us that $\hat{v}(\mu_m) \leq \varepsilon\hat{v}(\mu') + (1 - \varepsilon)\hat{v}(\mu)$. But, $\hat{v}(\mu_m) = \sum_\omega v(\hat{a}(\mu), \omega)\, \mu_m(\omega) = \varepsilon\sum_\omega v(\hat{a}(\mu), \omega)\, \mu'(\omega) + (1 - \varepsilon)\hat{v}(\mu)$. Hence, the last inequality is equivalent to $\sum_\omega v(\hat{a}(\mu), \omega)\, \mu'(\omega) \leq \hat{v}(\mu')$, which means $\hat{a}(\mu)$ is not strictly induced dominant.

Combining Lemma 4 and Lemma 5, we know that if Assumption 1 holds, Sender benefits from persuasion, and μ is an interior belief induced by an optimal signal, Receiver's preference at μ is not discrete. It remains to show is that if Assumption 1 holds, Sender benefits from persuasion, and a belief μ induced by an optimal signal leads to a best-attainable action, then Receiver's preference at μ is also not discrete. Suppose to the contrary Assumption 1 holds, Sender benefits from persuasion, an optimal signal π induces μ^* which leads to a best-attainable a^*, and Receiver's preference is discrete at μ^*. Let τ be the distribution of posteriors induced by π. Since Receiver's preference is discrete at μ^*, there is an $\varepsilon > 0$ s.t. $\mu' \equiv \varepsilon\mu_0 + (1 - \varepsilon)\mu^*$ also leads to a^*. Moreover, since τ puts positive measure on at most finitely many beliefs, we can select μ' s.t. τ does not put positive measure on μ'. Now, let τ' be defined as follows:

$$\tau'(\mu') = \frac{\tau(\mu^*)}{1 - \varepsilon + \varepsilon\tau(\mu^*)}$$

$$\tau'(\mu) = \frac{1 - \varepsilon}{1 - \varepsilon + \varepsilon\tau(\mu^*)}\tau(\mu) \text{ for } \mu \notin \{\mu^*, \mu'\}$$

$$\tau'(\mu^*) = 0.$$

Simple algebra shows τ' is Bayes plausible and has higher value than τ, so we've reached a contradiction.

REFERENCES

Alonso, Ricardo, and Niko Matouschek. 2008. "Optimal Delegation." *Review of Economic Studies*, 75(1): 259–93.

Anderson, Simon P., and Regis Renault. 2006. "Advertising Content." *American Economic Review*, 96(1): 93–113.

Aumann, Robert J., and Michael B. Maschler. 1995. *Repeated Games with Incomplete Information*. Cambridge, MA: MIT Press.

Barnoya, Joaquin, and Stanton A. Glantz. 2006. "The Tobacco Industry's Worldwide ETS Consultant's Project: European and Asian Components." *European Journal of Public Health*, 16(1): 69–77.

Brocas, Isabelle, and Juan D. Carrillo. 2007. "Influence through Ignorance." *RAND Journal of Economics*, 38(4): 931–47.

Caplin, Andrew, and John Leahy. 2004. "The Supply of Information by a Concerned Expert." *Economic Journal*, 114(497): 487–505.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica*, 50(6): 1431–51.

Dessein, Wouter. 2002. "Authority and Communication in Organizations." *Review of Economic Studies*, 69(4): 811–38.

Eliaz, Kfir, and Ran Spiegler. 2006. "Can Anticipatory Feelings Explain Anomalous Choices of Information Sources?" *Games and Economic Behavior*, 56(1): 87–104.

Ganuza, Juan-Jose, and Jose S. Penalva. 2010. "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions." *Econometrica*, 78(3): 1007–30.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics*, 24(3): 461–83.

Hiriart-Urruty, Jean-Baptiste, and Claude Lemaréchal. 2004. *Fundamentals of Convex Analysis*. New York: Springer.

Holmstrom, Bengt. 1984. "On the Theory of Delegation." In *Bayesian Models in Economic Theory*, ed. M. Boyer and R. E. Kihlstrom, 115–41. New York: Elsevier Science.

Horner, Johannes, and Andrzej Skrzypacz. 2005. "Selling Information." Yale University Cowles Foundation Discussion Paper 1743.

Johnson, Justin P., and David P. Myatt. 2006. "On the Simple Economics of Advertising, Marketing, and Product Design." *American Economic Review*, 96(3): 756–84.

Kartik, Navin. 2009. "Strategic Communication with Lying Costs." *Review of Economic Studies*, 76(4): 1359–95.

Kőszegi, Botond. 2006. "Emotional Agency." *Quarterly Journal of Economics*, 121(1): 121–55.

Lewis, Tracy R., and David E. M. Sappington. 1994. "Supplying Information to Facilitate Price Discrimination." *International Economic Review*, 35(2): 309–27.

McCloskey, Donald, and Arjo Klamer. 1995. "One Quarter of GDP Is Persuasion." *American Economic Review*, 85(2): 191–95.

Melumad, Nahum D., and Toshiyuki Shibano. 1991. "Communication in Settings with No Transfers." *RAND Journal of Economics*, 22(2): 173–98.

Milgrom, Paul. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics*, 12(2): 380–91.

Milgrom, Paul, and John Roberts. 1986. "Relying on the Information of Interested Parties." *RAND Journal of Economics*, 17(1): 18–32.

Moses, Hamilton III, E. Ray Dorsey, David H. M. Matheson, and Samuel O. Thier. 2005. "Financial Anatomy of Biomedical Research." *Journal of the American Medical Association*, 294(11): 1333–42.

Myerson, Roger B. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica*, 47(1): 61–73.

Ostrovsky, Michael, and Michael Schwarz. 2010. "Information Disclosure and Unraveling in Matching Markets." *American Economic Journal: Microeconomics*, 2(2): 34–63.

Rayo, Luis, and Ilya Segal. 2010. "Optimal Information Disclosure." *Journal of Political Economy*, 118(5): 949–87.

Shmaya, Eran, and Leeat Yariv. 2009. "Foundations for Bayesian Updating." Unpublished.

Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*, 87(3): 355–74.