The Price of Cheap Talk

Elliot Lipnowski Doron Ravid*

February, 2017

Preliminary and Incomplete

Abstract

We study a model of cheap talk with one substantive assumption: the sender's preferences are state-independent. Our key observation is that this setting is amenable to the belief-based approach familiar from models of persuasion with commitment. This approach allows us to assess the value of commitment, address several classic questions about cheap talk for the state-independent case, and explicitly solve for sender-optimal equilibria in a large class of examples. A key product is a geometric characterization of the value of cheap talk, described by the *quasiconcave* envelope of the sender's indirect utility.

1 Introduction

Much of our economy's resources are spent on persuasion.[1] To persuade an audience, however, requires credibility. As such, it is no surprise that a large and growing literature studies optimal persuasion under guaranteed credibility,

*Elliot Lipnowski: Becker Friedman Institute, University of Chicago, lipnowski@uchicago.edu, http://elliotlipnowski.com/. Doron Ravid: Department of Economics, University of Chicago, dravid@uchicago.edu, https://sites.google.com/site/ravidd/. We would like to thank Emir Kamenica for his invaluable comments and suggestions. We would also like to thank Benjamin Brooks, Laurent Mathevet and Stephen Morris for insightful discussions.

[1] See McCloskey and Klamer (1995), for example.

i.e. when the persuader has full commitment power (e.g. Rayo and Segal (2010), Kamenica and Gentzkow (2011), Alonso and Câmara (2016), and Ely (2017)). However, many who wish to persuade have no power to commit themselves and must communicate via cheap talk alone (Crawford and Sobel (1982), Green and Stokey (2007)). The effect of cheap talk communication on optimal persuasion is the subject of the present paper.

We consider optimal persuasion in a general cheap talk model with one substantive assumption: The sender's preferences do not depend on her information. This assumption is justified in many applications. Salespeople want to sell products with higher commissions, politicians want to get elected, job candidates want to get hired, etc. Thus, we start with a receiver facing a decision problem with incomplete information. The relevant information is available to an informed sender with a fixed agenda. In other words, the sender cares only about the receiver's action. Wanting to influence this action, the sender communicates with the receiver via cheap talk.

To illustrate, consider an investor who consults his broker to learn about an asset. The broker knows the amount of the asset that the investor should hold, and she earns a commission proportional to the volume of the investor's trades. Thus, the broker wants the investor to trade as high a volume as possible, regardless of the state of the market. We answer questions such as: Can the broker benefit from cheap talk with the investor? If so, what is the broker's benefit from cheap talk? What is the investor's benefit? What will broker say, and how will the investor respond?

We begin by noting that our model is amenable to the belief-based approach adopted by the persuasion literature (e.g. Kamenica and Gentzkow (2011), Alonso and Câmara (2016), Ely (2017)). In particular, equilibrium communication can be described via an *information policy*, i.e. a distribution over the receiver's posterior belief whose mean is equal to the prior. The observation that every information policy is induced by some communication protocol, and vice versa, is at the heart of the belief-based approach. We then use this approach to succinctly describe which ex-ante payoffs and information policies are consistent with an equilibrium (Lemma 1). Moreover, the sender's value from her most preferred equilibrium can be characterized geometrically in a similar vein to the full commitment case of Kamenica and Gentzkow (2011) and Aumann et al. (1995) (Theorem 1).

Theorem 2 shows that cheap talk is only as persuasive as the least persuasive message the sender can deliver. More precisely, say that s is securable if the sender's ex-post payoffs from some information policy are always higher than s. Then Theorem 2 says that a payoff better than no communication can be obtained via cheap talk if and only if the payoff is securable. Note that the theorem relies on there being *some* information policy that secures the sender a payoff that is bounded from below by s. However, this information policy need not be an equilibrium. Still, the theorem says that one can find an equilibrium respecting the same lower bound provided that such an information policy exists. Intuitively, said information policy leads to posteriors that provide too much sender-beneficial information to the receiver. By reducing said information posterior-by-posterior, one can construct an equilibrium information policy that respects the original policy's lower bound on the sender's payoffs.

To understand our results, consider the following simplified version of the broker example. Suppose the investor owns $\frac{1}{3}$ of the available asset and is considering whether to sell all that he owns, increase his holdings to 1, or keep his current position. Getting a commission proportional to the investor's trades, the broker wants to maximize the absolute difference between the investor's final asset holdings, $a \in \left\{0, \frac{1}{3}, 1\right\}$, and his initial position, $\frac{1}{3}$. Assume that the investor's optimal asset holdings, θ, are known to the broker. For simplicity, suppose that θ is either 0 or 1, meaning that the investor should either own the entire asset or none of it. Finally, suppose that, upon consulting with his broker, the investor sets his holdings to be the value closest to his expectation of θ among the three options under consideration.

In this example, Theorem 2 implies is that the broker benefits from cheap talk if and only if fully revealing the state secures a payoff strictly higher than babbling. To see why, let μ_0 be the probability that the investor should buy the entire asset, and therefore the expected value of θ under the investor's prior. Note that revealing the state to the investor is equivalent to the broker telling the investor to *sell* for sure when the state is 0, and to *buy* for sure when the state is 1. A *buy* message leads the investor buys the entire asset, while a *sell* message leads the investor to sell everything. Hence, the broker's ex-post payoffs are $\frac{2}{3}$ when sending a *buy* message and $\frac{1}{3}$ following a *sell* message. Hence, revealing the state secures a payoff of $\frac{1}{3}$. Such a payoff is strictly higher than the broker's payoff under babbling if and only if $\frac{1}{6} < \mu_0 < \frac{2}{3}$. We now show that one can use this restriction to construct a cheap talk equilibrium that outperforms babbling. To do so, reduce the informativeness of the *buy* signal by having the broker tell the investor to *buy* according to:

$$\Pr\{buy|\theta = 0\} = \frac{1}{2}\left(\frac{\mu_0}{1-\mu_0}\right)$$
$$\Pr\{buy|\theta = 1\} = 1$$

while telling the investor to *sell* with the complementary probabilities. As with perfect state revelation, the investor's unique best response following a *sell* message is to sell all his holdings. However, the investor's posterior given a *buy* message assigns a probability of $\frac{2}{3}$ that she should buy everything, making him indifferent between buying and keeping her current position. Being indifferent upon hearing a *buy* message, the investor can mix: he keeps his position with probability $\frac{2}{3}$, and buys the entire asset with probability $\frac{1}{3}$. Note that this mixing means that the broker has no incentive to lie, since her payoff is $\frac{1}{3}$ regardless of her message. We have therefore constructed a cheap talk equilibrium which is strictly better for the broker than babbling.

We now argue that, if the broker benefits from cheap talk, then revealing the state must secure her a payoff strictly higher than under μ_0. In other words, it must be that $\frac{1}{6} < \mu_0 < \frac{2}{3}$. Indeed, note that the investor is willing to buy the entire asset under his prior whenever $\mu_0 \geq \frac{2}{3}$. Since this is the best possible outcome for the broker, the broker can benefit from cheap talk only if $\mu_0 < \frac{2}{3}$. Suppose then that $\mu_0 < \frac{2}{3}$. Note any equilibrium has some message that leads the investor to assign a probability weakly below μ_0 to the state being 1. Under such posteriors, the investor is never willing to buy the entire asset. As such, the highest ex-post payoff the broker hope to secure is $\frac{1}{3}$. By Theorem 2, the broker's equilibrium payoff cannot be above $\frac{1}{3}$. But $\frac{1}{3}$ is the broker's payoff under babbling whenever $\mu_0 \leq \frac{1}{6}$. Hence, the broker can benefit from cheap talk only when her prior belief that she should buy the entire asset is strictly between $\frac{1}{6}$ and $\frac{2}{3}$.

Theorem 2 also tells us that commitment power would benefit the broker whenever $0 < \mu_0 < \frac{2}{3}$. Consider for example the message protocol used in the above constructed equilibrium. With commitment power, the broker could use the exact same protocol, but replace the investor's mixing with the broker's preferred behavior when he is indifferent, buying the entire asset for sure after seeing a *buy* message. This gives the broker an ex-ante payoff of $\frac{1}{3} + \frac{\mu_0}{2}$. However, an argument identical to the one of the previous paragraph implies that for $\mu_0 < \frac{2}{3}$ the highest payoff that the broker can secure is $\frac{1}{3}$. As such, we have found a payoff that is feasible ex-ante, but cannot be secured ex-post. The ability to commit would therefore strictly benefit the broker.

The above example highlights that cheap may hurt the sender by forcing the receiver to mix. In Section 3, we show that such mixing generally hurts the sender, but that mixing is often required in equilibrium. Intuitively, having the receiver mix generates commitment power to the sender by preventing her from steering the receiver towards her preferred actions. But this commitment power comes at a cost, since the sender's preferred actions are then taken with a lower probability. Increasing this probability is therefore one way in which the sender could benefit from commitment.

Section 4 generalizes the above broker example. In particular, we allow the investor to hold any share of the asset, and let the ideal asset holdings have an arbitrary distribution over the unit interval. As in the simple example above, the broker-preferred equilibrium has the broker sending the receiver either a *buy* or a *sell* message, and the receiver acting accordingly. In fact, we show that the two-message structure is sufficient for describing all possible equilibrium outcomes, even in the richer many-state model. Specifically, every equilibrium broker-investor payoff pair can be obtained via such a two-message equilibrium. Using this observation, we provide a complete characterization of the player's possible equilibrium payoffs in our general example.

We then turn to discussing some classical questions from the cheap talk literature, showing that the belief-based approach allows these questions to be fully settled in our environment. We begin by providing necessary and sufficient conditions for full information revelation, and continue by examining when effective communication is possible.

There are two ways in which cheap talk communication can be effective. First, the sender's messages convey information to the receiver. In other words, cheap talk equilibria can be *informative*. Second, different messages lead the receiver to take different actions. That is, cheap talk equilibria can be *influential.* In a specialization of our model Chakraborty and Harbaugh (2010) apply an ingenious fixed point argument to name sufficient conditions for an influential equilibrium to exist. Using the belief-based approach, we extend their logic to obtain sufficient conditions for there to be an informative equilibrium in our general model. By dealing with the remaining special cases, Proposition 5 provides necessary and sufficient conditions for the existence of an informative equilibrium. Virtually the same conditions characterize the existence of an influential equilibrium in Chakraborty and Harbaugh (2010)'s special setting. We present those conditions in Proposition 6.

Section 5.3 considers how cheap talk influences persuasive information. We begin by noting that, for any information policy that secures the sender a payoff better than no communication, our proof of Theorem 2 constructs a weakly less informative cheap talk equilibrium. There are, however, cases in which cheap talk does not influence the persuasive information policy. Consider the messaging protocol we constructed in the simple broker example above. When the receiver mixes, this protocol gives the broker best cheap talk equilibrium. When the receiver uses a pure strategy, one can use Kamenica and Gentzkow (2011)'s results to show that the *same information policy* protocol gives the broker's best value under commitment. Hence, while losing commitment reduces the broker's value, it need not change the nature of persuasive information. Proposition 7 identifies a property of persuasive information that allows it to be optimal under both commitment and cheap talk while leading to a value loss under cheap talk.

2 Persuasion without Commitment

Our model is an abstract cheap talk model with the one substantive restriction that the sender has state-independent preferences. Thus, we have two players: a sender (S) and a receiver (R). The game begins with the realization of an unknown state, $\theta \in \Theta$, which S observes. After observing the state, S sends R a message, $m \in M$. R then observes m (but not θ) and decides which action, $a \in A$, to take. While R's payoffs from a depend on the state, S's payoffs do not.

We impose a few technical restrictions on our model. Both Θ and A are compact metrizable spaces that contain at least two elements. The state, θ, follows some full-support distribution $\mu_0 \in \Delta\Theta$,[2] which is known to both players. Both players' utility functions are continuous, where we take $\pi : A \to \mathbb{R}$ to be S's utility and $u : A \times \Theta \to \mathbb{R}$ to be R's. Finally, we assume that the message space, M, is some Polish space which contains both $\Delta\Theta$ and ΔA.

A strategy for S maps each state of the world to a distribution over messages. A strategy for R specifies a mixed action for R conditional on every message R may observe. We are interested in studying the game's equilibria, by which we mean perfect Bayesian equilibria. Thus, an **equilibrium** consists of three measurable maps: a strategy $\sigma : \Theta \to \Delta M$ for S, a strategy $\rho : M \to \Delta A$ for R, and a belief system $\beta : M \to \Delta\Theta$ for R, such that:

1. β is obtained from μ_0, given message policy σ, using Bayes' rule whenever possible.[3]

2. $\rho(m)$ is supported on $\arg\max_{a \in A} u(a, \beta(m))$ for all $m \in M$.

3. $\sigma(\theta)$ is supported on $\arg\max_{m \in M} \pi(\rho(m))$ for all $\theta \in \Theta$.

[2]For a Polish space Y we let ΔY denote the set of all Borel probability measures over Y, endowed with the weak* topology. For $\gamma \in \Delta Y$, we let $\text{supp}(\gamma)$ denote the support of γ. For $f : Y \to \mathbb{R}$ bounded and measurable, we let $f(\gamma) := \int_Y f \, \mathrm{d}\gamma$.

[3]i.e. $\int_{\hat{\Theta}} \sigma(\hat{M}|\cdot) \, \mathrm{d}\mu_0 = \int_\Theta \int_{\hat{M}} \beta(\hat{\Theta}|\cdot) \, \mathrm{d}\sigma(\cdot|\theta) \, \mathrm{d}\mu_0(\theta)$ for every Borel $\hat{\Theta} \subseteq \Theta$ and $\hat{M} \subseteq M$.

Any triple $\mathcal{E} = (\sigma, \rho, \beta)$ induces a joint distribution, $\mathbb{P}_{\mathcal{E}}$, over realized states, messages, and actions,[4] which in turn induces (through β and ρ, respectively) distributions over R's equilibrium beliefs and chosen mixed action. As in Crawford and Sobel (1982), a **babbling equilibrium** is any equilibrium in which S does not communicate any information, i.e. σ is constant.

We begin our analysis by noting that our model is amenable to the belief-based approach used in the information design literature. This approach uses the ex-ante distribution over R's posterior beliefs, $p \in \Delta\Delta\Theta$, as a substitute for both S's strategy and the equilibrium belief system. Clearly, every belief system and strategy for S generate some such distribution over R's posterior belief. By Bayes' rule, this posterior distribution averages to the prior, μ_0. That is, $p \in \Delta\Delta\Theta$ satisfies $\int_{\Delta\Theta} \mu \, \mathrm{d}p(\mu) = \mu_0$. We refer to any p that averages back to the prior as an **information policy**. Thus, only information policies can originate from some σ and β. The fundamental result underlying the belief-based approach is that every information policy can be generated by some σ and β.[5] Let $\mathcal{I}(\mu_0)$ denote the set of all information policies.

The belief-based approach allows us to focus on the game's equilibrium outcomes. Formally, an **outcome** is a triplet, $(p, s, r) \in \Delta\Delta\Theta \times \mathbb{R} \times \mathbb{R}$, representing R's posterior distribution (p), S's ex-ante payoff (s), and R's ex-ante payoff (r). An outcome is an **equilibrium outcome** if it corresponds to an equilibrium.[6] In contrast, an outcome is a **commitment outcome** if it is corresponds to some triple (σ, ρ, β) satisfying the first two of three conditions for equilibrium above. In other words, commitment outcomes do not require S's behavior to be incentive-compatible.

Lemma 1 below characterizes the game's equilibrium outcomes. To state

---

[4] Specifically, $\mathcal{E} = (\sigma, \rho, \beta)$ induces measure $\mathbb{P}_{\mathcal{E}} \in \Delta(\Theta \times M \times A)$, which assigns probability $\mathbb{P}_{\mathcal{E}}(\hat{\Theta} \times \hat{M} \times \hat{A}) = \int_{\hat{\Theta}} \int_{\hat{M}} \rho(\hat{A}|\cdot) \, \mathrm{d}\sigma(\cdot|\theta) \, \mathrm{d}\mu_0(\theta)$ for every Borel $\hat{\Theta} \subseteq \Theta, \hat{M} \subseteq M, \hat{A} \subseteq A$.

[5] For example, see Benoit and Dubra (2011) or Kamenica and Gentzkow (2011).

[6] i.e. If there exists an equilibrium $\mathcal{E} = (\sigma, \rho, \beta)$ such that $p = (\mathrm{marg}_M \mathbb{P}_{\mathcal{E}}) \circ \beta^{-1}$, $s = \pi(\mathrm{marg}_A \mathbb{P}_{\mathcal{E}})$, and $r = u\left(\mathrm{marg}_{A\times\Theta} \mathbb{P}_{\mathcal{E}}\right)$.

the theorem, let $U(\mu)$ be R's maximal payoff when her belief is μ,

$$\begin{aligned} U &: \Delta\Theta \to \mathbb{R} \\ &\mu \mapsto \max_{\alpha \in \Delta A} u(\alpha \otimes \mu) = \max_{a \in A} u(a, \mu); \end{aligned}$$

and let $V(\mu)$ be the set of continuation values that S can attain following a messages that leads R to have a posterior μ,

$$\begin{aligned} V &: \Delta\Theta \rightrightarrows \mathbb{R} \\ &\mu \mapsto \pi\left(\arg\max_{\alpha \in \Delta A} u(\alpha \otimes \mu)\right) = \text{co } \pi\left(\arg\max_{a \in A} u(a, \mu)\right). \end{aligned}$$

Notice that (appealing to Berge's theorem) U is continuous, and V is a Kakutani correspondence[7]

Lemma 1. *The outcome (p, s, r) is an equilibrium outcome if and only if:*

1. $p \in \mathcal{I}(\mu_0)$, *i.e.* $\int_{\Delta\Theta} \mu \, \mathrm{d}p(\mu) = \mu_0$;
2. $s \in \bigcap_{\mu \in \text{supp}(p)} V(\mu)$;
3. $r = \int_{\Delta\Theta} U \; dp$.

The three conditions correspond exactly to the three conditions defining perfect Bayesian equilibrium. Notice that $V(\mu) \times \{U(\mu)\}$ is the set of payoff pairs attainable in a babbling equilibrium given prior μ. Thus, unlike the general cheap talk model, understanding the babbling equilibrium outcomes (at different hypothetical priors) is enough to fully describe the equilibrium outcomes of the very cheap talk setting.

Our main focus is cheap talk's influence on persuasion. Formally, an equilibrium is **strictly persuasive** if it is strictly better for S than every babbling equilibrium. A **most persuasive equilibrium** is an ex-ante S-preferred equilibrium. Similarly, a **most persuasive commitment outcome** is an ex-ante S-preferred commitment outcome.

[7]i.e. nonempty-, compact-, and convex- valued upper hemicontinuous correspondence.

Theorem 1 characterizes S's value from a most persuasive equilibrium. Let $v := \max V : \Delta\Theta \to \mathbb{R}$. Note that v is upper semicontinuous, since u is continuous and therefore V is upper hemicontinuous by Berge's theorem. We take $\bar{v} : \Delta\Theta \to \mathbb{R}$ and $\hat{v} : \Delta\Theta \to \mathbb{R}$ to denote the **quasiconcave envelope** and **concave envelope** of v, respectively. That is, $\bar{v}$ (resp. $\hat{v}$) is the pointwise lowest quasiconcave (concave) and upper semicontinuous function which majorizes v.[8] Since quasiconcavity follows from concavity, the quasiconcave envelope lies (weakly) below the concave envelope. By Kamenica and Gentzkow (2011), $\hat{v}$ gives S's payoff from a most persuasive commitment policy. Theorem 1 below shows that, under cheap talk, S's optimal value reduces to $\bar{v}$.

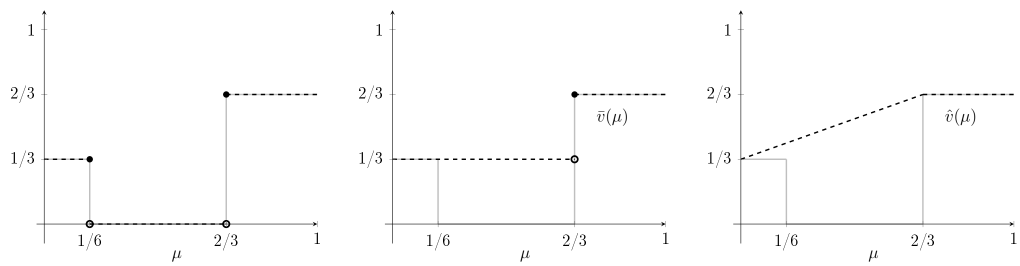
Theorem 1. *A most persuasive equilibrium exists, and it provides value $\bar{v}(\mu_0)$ to the sender.*

Theorem 1 enables us to graphically compare the value of cheap talk persuasion with persuasion with commitment. Figure 1 below graphs v, $\bar{v}$ and $\hat{v}$ as a function of μ for the simple broker example from the introduction. Since the state is binary, we identify each posterior μ with the probability it assigns to θ being 1. All figures have the broker's value correspondence, V, in light gray. The dashed black represents v in the left figure, v's quasiconcave envelope in the middle figure, and v's concave envelope in the right figure. These correspond to the highest value S can obtain from babbling, cheap talk, and commitment, respectively.

In cases with many states, a graphical comparison is often difficult to accomplish. In such cases, qualitative comparison of full commitment and cheap talk persuasion becomes more difficult. We therefore turn to a different characterization of S's value from cheap talk persuasion, based on the following definition.

Definition 1. Given an information policy p and sender payoff s: say p **achieves** s if $\int_{\Delta\Theta} v \; \mathrm{d}p \geq s$, and p **secures** s if $v|_{\mathrm{supp}(p)} \geq s$. Say a sender

[8] In the appendix, we show that $\bar{v}$ is well-defined in Lemma 4. In the case that Θ is finite, the qualifier "upper semicontinuous" can be omitted from the definition without change.



(a) Babbling: v (b) Cheap talk: $\bar{v}$ (c) Commitment: $\hat{v}$

Figure 1: The simple broker example

payoff s can be achieved (resp. secured) if there exists some Bayes-plausible information policy which achieves (secures) s.

That S can achieve a payoff s in equilibrium only if s can be secured is immediate. To see why, note that state-independent preferences require indifference between any on-path message. As such, S can achieve a high value s in equilibrium only if s is her payoff from any on-path message. Thus, the equilibrium information policy must secure s. However, not every information policy securing s is an equilibrium. In particular, one could secure s with an information policy that gives a payoff higher than s after certain messages, as was the case for full information in our simple broker example. Still, this information policy can be used to construct another, less informative, policy which yields s in an equilibrium.

Theorem 2. *Suppose $s \geq v(\mu_0)$. Then there exists an equilibrium inducing sender payoff s if and only if s can be secured. In particular,*

$$\bar{v}\left(\mu_0\right) = \max_{p \in \mathcal{I}(\mu_0)} \inf_{\mu \in \operatorname{supp} p} v\left(\mu\right).$$

The theorem allows us to understand S's equilibrium payoffs without finding any equilibrium outcome. To learn whether some high payoff is feasible in equilibrium, it is sufficient to look at which values can be secured by some (equilibrium or non-equilibrium) information policy. One can use the theorem to compare the value of cheap talk persuasion to two natural benchmarks:

babbling and full commitment.[9] An equilibrium is **extremely persuasive** if it gives S her payoff under full commitment, i.e. if $s = \hat{v}(\mu_0)$. Say **cheap talk is costless** if such an equilibrium exists, i.e. $\bar{v}(\mu_0) = \hat{v}(\mu_0)$, and that **cheap talk is valuable** if $\bar{v}(\mu_0) > v(\mu_0)$. In other words, cheap talk is costless if a most persuasive equilibrium is extremely persuasive, and cheap talk is valuable if a most persuasive equilibrium is strictly persuasive. The following corollary provides necessary and sufficient conditions both for valuable and for costless cheap talk.

Corollary 1. *Cheap talk is valuable if and only if some payoff $s > v(\mu_0)$ can be secured. Cheap talk is costless if and only if every payoff which can be achieved can be secured.*

The proof is rather straightforward. For the first part, note that cheap talk is valuable if and only if some sender payoff $s > v(\mu_0)$ is attainable in equilibrium. By Theorem 2, this is equivalent to s being securable. For the second part, note that, in light of the theorem, cheap talk is costless if and only if the most persuasive commitment payoff can be secured. But in this case any achievable payoff can be secured, thereby completing the proof.

3 Mixed Messages

What kind of messages does S send to R in equilibrium? Kamenica and Gentzkow (2011) show that, with full commitment power, all equilibrium payoffs can be generated by an equilibrium in which S recommends to R which pure action to take. The following proposition shows that a similar result holds with cheap talk, with one important caveat: S must be allowed to recommend that R takes a *mixed action.*

Proposition 1. *If (p, s, r) is an equilibrium outcome, then there exists an equilibrium (σ, ρ, β) which induces ex-ante payoffs (s, r) and such that there exists closed $X \subseteq \Delta A$ with $\sigma(X|\cdot) = 1$ and $\rho(\cdot|\alpha) = \alpha$ for all $\alpha \in X$.*

[9] If, rather than choosing what to say to the receiver, S were choosing an experiment, then the appropriate limited commitment model would not be one of cheap talk. Rather, it would be an informed principal problem, as studied by Perez-Richet (2014).

The proof follows a similar line to similar results in mechanism design (Myerson (1979)), persuasion (Kamenica and Gentzkow (2011)), and other information design settings (Bergemann and Morris (2016)). One can take any equilibrium and pool all the messages leading to the same mixed action without changing R's or S's incentives. However, S's incentives do change if one tries to use the induced distribution over pure action recommendations to replace a mixed action. In particular, if the mixed action includes two actions that give S two different payoffs, recommending the low payoff action will never be incentive-compatible. As such, some equilibrium payoffs may be infeasible to implement using pure action recommendations. This issue would be sidestepped if the players had access to a disinterested mediator, as in Lugo (2016), in which case the revelation principle reasoning of Myerson (1986) would apply.

More than a technical issue separating these models, the possible requirement that R mix can have payoff consequences for S. To see why, consider the argument in the above paragraph, and suppose that S could commit. Then one could increase S's payoffs by replacing any mixed action recommendation with S's most preferred action in the mixture's support. If doing this is impossible, then S's incentive constraint must be binding, which means S must be losing value from limited commitment. We formalize this result in Proposition 2 below.

Proposition 2. *Suppose that (σ, ρ, β) is a most persuasive equilibrium, and that $(\sigma, \tilde{\rho}, \beta)$ is not an equilibrium for any receiver strategy $\tilde{\rho}$ without mixing.*[10] *Then cheap talk is costly.*

At this stage, the reader may wonder whether there are families of cases in which optimal cheap talk persuasion requires R to mix. To find such a class, consider the simple broker example presented in the introduction. In this example, the investor can take one of three actions, each of which leads to a different payoff for the broker. In other words, the broker's payoffs are a one-to-one function of the investor's actions. This property turns out to be

[10] Say $\tilde{\rho}$ entails no mixing if, for all $m \in M$, there is some $a \in A$ with $\tilde{\rho}(a|m) = 1$.

sufficient for cheap talk persuasion to require mixing.

Lemma 2. *If π is one-to-one and cheap talk is valuable, then every most persuasive equilibrium entails mixing by the receiver.*

To understand the intuition behind the lemma, suppose cheap talk is valuable. By Proposition 1, there is some most persuasive equilibrium using mixed action recommendations. Since cheap talk is valuable, S must be recommending to R least two distinct mixed actions in equilibrium (as the equilibrium would otherwise be payoff-equivalent to babbling) . By incentive compatibility, S's expected payoffs from each mixed action must be the same. But since π is one-to-one there are is at most one pure action giving each payoff, meaning that R must be mixing.

Combining Lemma 2 with Proposition 2, one obtains Corollary 2 below. This Corollary shows that commitment is valuable for S whenever π is one-to-one.

Corollary 2. *Suppose π is one-to-one. Then one of the following must hold:*

1. *Persuasion is impossible:* $\hat{v}(\mu_0) = \bar{v}(\mu_0) = v(\mu_0)$.
2. *Cheap talk is costly:* $\hat{v}(\mu_0) > \bar{v}(\mu_0)$.

The above corollary has strong implications for the world of finitely many actions, in which S's payoffs are generically one-to-one. In the finite action setting, the above corollary suggests that cheap talk is generically costly whenever it is valuable.

4 The Broker Example

We now consider a richer model of the broker-investor relationship. Our investor has pre-existing holdings $\bar{a} \in (0, \frac{1}{2})$ of an asset of unknown quality $\theta \in \Theta = [0,1]$ (of prior distribution $\mu_0 \in \Delta\Theta$) and must decide a final amount $a \in A = [0,1]$ to hold. A broker, knowing the "ideal" position that the investor should take, is tasked with advising the investor. The broker has state-independent preferences: her payoff accrues from brokerage fees, and so is proportional to the net volume of trade, $\pi(a) = |a - \bar{a}|$. The investor wants to match the ideal holdings level, and may internalize the broker's fees: $u(a, \theta) = -\frac{1}{2}(a - \theta)^2 - \kappa\pi(a)$ for some $\kappa \geq 0$. In this section, we fully characterize the set of equilibrium payoff pairs.

First, it is straightforward to characterize optimal behavior by the investor.

Claim 1. For any posterior belief $\mu \in \Delta\Theta$, the investor's best response satisfies $\arg\max_{a \in A} u(a, \mu) = \{a^*(\mu)\}$, where

$$a^*(\mu) = \begin{cases} \int_\Theta \theta \, \mathrm{d}\mu(\theta) + \kappa & : \int_\Theta \theta \, \mathrm{d}\mu(\theta) - \bar{a} \leq -\kappa \\ \bar{a} & : \int_\Theta \theta \, \mathrm{d}\mu(\theta) - \bar{a} \in [-\kappa, \kappa] \\ \int_\Theta \theta \, \mathrm{d}\mu(\theta) - \kappa & : \int_\Theta \theta \, \mathrm{d}\mu(\theta) - \bar{a} \geq \kappa. \end{cases}$$

It immediately follows that V is single-valued with $v(\mu) = \left[\left| \int_\Theta \theta \, \mathrm{d}\mu(\theta) - \bar{a} \right| - \kappa \right]_+$.

Our next observation about the structure of this game is that the players will exhibit the same preferences over any two equilibria.

Claim 2. The tuple (p, s, r) is an equilibrium outcome if and only if $r = u(\bar{a}, \mu_0) + \frac{1}{2}s^2$ and $v|_{\mathrm{supp}(p)} = s$.

In light of the above claim, all that remains is to characterize the set of equilibrium values for the broker. We know that $v(\mu_0)$ is an equilibrium payoff for the broker (with babbling), and that any other equilibrium payoff is at least as good (since v is quasiconvex). Moreover, Theorem 2 tells us that the equilibrium payoff set is convex (see Corollary 6), and Theorem 1 fully characterizes the highest equilibrium payoff for the broker as $\bar{v}(\mu_0)$. In principle, we now understand fully the equilibrium payoff set of the broker game.

Claim 3. There exists an equilibrium generating payoff s to the broker and r to the investor if and only if $v(\mu_0) \leq s \leq \bar{v}(\mu_0)$ and $r = u(\bar{a}, \mu_0) + \frac{1}{2}s^2$. In particular, the unique Pareto optimal equilibrium payoff pair is $\left(\bar{v}(\mu_0),\ u(\bar{a}, \mu_0) + \frac{1}{2}\bar{v}(\mu_0)^2\right)$.

While exact, the above characterization leaves something to be desired. In general, computing the quasiconcave envelope of a function (especially one defined on an infinite-dimensional space, as ours is since Θ is infinite) may be demanding or intractable. Relatedly, it would be desirable to understand more about what a most persuasive equilibrium actually looks like. That is, one might like to know, "What sort of information will the broker provide?"

To address this, we observe that the broker example has two noteworthy features. First, the state is one-dimensional, just as in the seminal CS model. Second, the investor's optimal behavior depends only on her expectation of the state. This feature of receiver preferences (which is true also of the benchmark example of CS) is immediate from the statement of Claim 1. These two features, together with state-independence, characterize the setting of Gentzkow and Kamenica (2016).[11] Just as Gentzkow and Kamenica (2016) show in the case with commitment, we will see that they enable a more complete analysis using the belief-based approach.

A central concept to our simple characterization of a most persuasive equilibrium is the notion of a cutoff equilibrium.

Definition 2. Suppose $\Theta \subseteq \mathbb{R}$. Given $q \in [0,1]$, the q-**cutoff policy** is the (necessarily unique) information policy $p \in \mathcal{I}(\mu_0)$ of the form $p = q\delta_{\mu_L} + (1-q)\delta_{\mu_R}$, for $\mu_L, \mu_R \in \Delta\Theta$ with $\max \text{supp}(\mu_L) \leq \min \text{supp}(\mu_R)$. Say $p \in \mathcal{I}(\mu_0)$ is a **cutoff policy** if it is the q-cutoff policy for some $q \in [0,1]$.

The q-cutoff policy reports whether the state is in the bottom q quantiles or the top $1-q$ quantiles, as measured according to the prior. More concretely, the sender simply reports whether the state is above or below some well-calibrated cutoff.[12] Such policies are an intuitive form of information, but they are also tractable. As the following proposition shows, pairing the setting of Gentzkow and Kamenica (2016) with the assumption that the sender cannot commit *almost* restricts the search for most persuasive equilibria to that of a cutoff

[11] The investor's best response is also unique at every belief in our setting. The setting of Gentzkow and Kamenica (2016) has a different special feature, namely that the action set is finite. As is clear from the statements of Proposition 3 and Corollary 3 below, neither additional feature is important for our results.

[12] This is exactly correct in the case that μ_0 is atomless; if the cutoff is itself a state with positive prior probability, then S's message may need to be random conditional on the cutoff state itself occuring.

policy.

Proposition 3. *Suppose that* $\Theta \subseteq \mathbb{R}$ *and that there exists* $a^* : co(\Theta) \rightrightarrows A$ *such that* $\arg\max_{a\in A} u(a,\mu) = a^*\left(\int_\Theta \theta \, \mathrm{d}\mu(\theta)\right)$ *for every* $\mu \in \Delta\Theta$. *Then, for every sender value* $s \geq v(\mu_0)$ *which can be secured, there is some equilibrium outcome* (p^*, s, r) *such that* $|\mathrm{supp}(p^*)| \leq 2$ *and* p^* *is a garbling of a cutoff policy.*

This simplification is not available in the world of persuasion with commitment. The scope to send more than two messages will expand (whenever there are more than two states) the distribution of sender values that can be induced. It is special to the cheap talk case that the sender has no way of profiting from this in equilibrium.[13]

Finally, the broker example has one additional feature which simplifies the form of a most persuasive equilibrium: all credible information that the broker conveys is to her benefit. This yields further structure for equilibrium.

Corollary 3. *Suppose the premise of Proposition 3 holds, and further suppose that* v *is weakly quasiconvex. Then there is a most persuasive equilibrium outcome* (p^*, s, r) *such that* p^* *is a cutoff policy.*

Applying the above corollary to our broker example tells us that some most persuasive equilibrium entails a cutoff strategy. Our exact functional form—in particular, that $\theta \mapsto (|\theta - \bar{a}| - \kappa)_+$ is strictly decreasing (resp. increasing) to the left (right) of its set of minimizers—provides more detail.

Claim 4. If $\int_\Theta \theta \, \mathrm{d}\mu_0(\theta) \geq 2\bar{a}$, then a babbling equilibrium is a Pareto dominant (and therefore most persuasive) equilibrium. If $\int_\Theta \theta \, \mathrm{d}\mu_0(\theta) < 2\bar{a}$, then there is a unique $q \in (0,1)$ such that the q-cutoff policy $p = q\delta_{\mu_L} + (1-q)\delta_{\mu_R}$ satisfies $\bar{a} - \int_\Theta \theta \, \mathrm{d}\mu_L(\theta) = \int_\Theta \theta \, \mathrm{d}\mu_H(\theta) - \bar{a}$, and the unique equilibrium outcome with belief distribution p is a Pareto dominant equilibrium.

[13] For example, take $\Theta = A = [0,1]$, $u(a,\theta) = -\frac{1}{2}(a-\theta)^2$, μ_0 to be the uniform measure, and $\pi(a) = a^2$. While no cheap talk equilibrium can outperform babbling, a committed sender would strictly prefer to reveal the state, as v is strictly convex. Thus, a restriction to finitely many messages would entail a payoff loss to the sender.

In summary, the belief-based approach (with Lemma 1 at its core) enables a complete analysis of this rich example of persuasion without commitment. Using this approach, we show that a Pareto optimal equilibrium communication protocol takes the form of a simple cutoff rule: the investor is advised to sell some of his position at low states and bolster his pre-existing position at high states.

5 Discussion

5.1 Fully Revealing Equilibrium

When can S communicate to R all that she knows? This question has been asked in a variety of settings. For example, Baliga and Morris (2002) and Hagenbach et al. (2014) study when can players communicate what they know before a strategic interaction. Mathis (2008) provides conditions under which a sender can reveal her information when information is certifiable. Renault et al. (2013), Golosov et al. (2014), and Margaria and Smolin (2016) study, among other things, the possibility of full information revelation when S and R's interaction is dynamic.

An equilibrium (σ, ρ, β) is **fully revealing** if, for every $m \in M$, there is some $\theta \in \Theta$ such that $\beta(\theta|m) = 1$. The following proposition characterizes when our model admits a fully revealing equilibrium.

Proposition 4. *There exists a fully revealing equilibrium if and only if*

$$\sup_{\theta \in \Theta} \min V(\delta_\theta) \leq \inf_{\theta \in \Theta} \max V(\delta_\theta).$$

The proof is straightforward. Note first that, if the above inequality does not hold, there must be two states such that S necessarily strictly prefers R to be completely convinced of one rather than the other. As such, revealing the less beneficial state to R cannot be incentive-compatible. Suppose now that the above inequality holds. Since V is a Kakutani correspondence, there is a payoff S can obtain after revealing the state to R, regardless which state it is.

The existence of a fully revealing equilibrium then follows from Lemma 1.

5.2 When is Communication Effective?

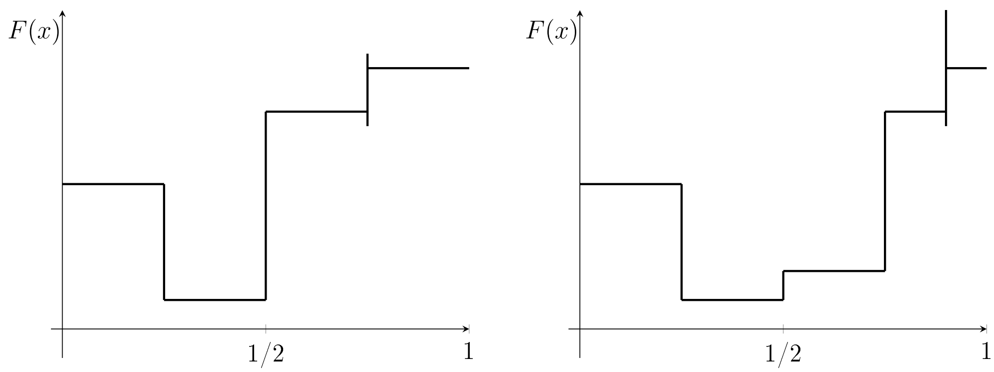
As mentioned in the introduction, communication can be effective either by providing R with information or by influencing R's actions. An equilibrium $\mathcal{E}$ is **informative** if R's beliefs change on path, i.e. if the induced receiver belief distribution is not equal to δ_{μ_0}. An equilibrium $\mathcal{E}$ is **influential** if S's messages influence R's actions on-path, meaning that there exists no sender behavior $\alpha \in \Delta A$ such that $\mathbb{P}_{\mathcal{E}}\{\rho = \alpha\} = 1$. The possibility of influential communication is the subject of the present section.

Our analysis in this section is based on Lemma 3 below, which owes an intellectual debt to Chakraborty and Harbaugh (2010). By applying the famous Borsuk-Ulam Theorem, Chakraborty and Harbaugh (2010) show that a special case of our model always admits an influential equilibrium. There are two features of Chakraborty and Harbaugh (2010)'s specialization that are essential for using Borsuk-Ulam: (1) that the state belongs to a multidimensional Euclidean space, and (2) that S's value is a (single-valued) continuous function of R's beliefs. Our belief-based approach shows that the structure of the first feature is mirrored by *any* setting with more than two states. For the second feature, we show that a weaker condition which applies in the general case—S's possible values form a Kakutani correspondence of R's beliefs—lets one apply similar fixed point reasoning. We apply this reasoning to prove Lemma 3 below.

Lemma 3. *If $\mu_0 = \sum_{i=1}^{3} p_i \mu_i$ for some $\mu_1, \mu_2, \mu_3 \in \Delta\Theta$ and $p_1, p_2, p_3 > 0$, then there exist distinct $q, q' \in \Delta\{1,2,3\}$ such that $p = \frac{1}{2}(q + q')$, and there is an equilibrium whose generated information policy is supported on $\left\{\sum_{i=1}^{3} q_i \mu_i,\ \sum_{i=1}^{3} q'_i \mu_i\right\}$.*

We now use Lemma 3 to find necessary and sufficient conditions for communication to be effective. Our conditions completely characterize when informative equilibria exist. We then show that virtually the same conditions complete Chakraborty and Harbaugh (2010)'s characterization for there to be an influential equilibrium in their setting. The characterization is based on the following monotonicity condition.

Definition 3. Say a correspondence $F : I \rightrightarrows \mathbb{R}$ on a real interval is **monotone through** $x \in I$ if $\forall y_0, y_1 \in I$ with $y_0 < x < y_1$, we have $\max F(y_0) < \min F(y_1)$. Given a convex subset X of some real linear space, say $F : X \rightrightarrows \mathbb{R}$ is monotone through $x \in X$, if there is some affine injection $\psi : X \to \mathbb{R}$ such that $F \circ \psi^{-1}$ is monotone through $\psi(x)$.



(a) Monotone through $\frac{1}{2}$ (b) Not monotone through $\frac{1}{2}$

Figure 2: Illustration of the monotonicity condition of Definition 3

Figure 2 illustrates Definition 3 for correspondences defined on the unit interval. It plots two correspondences, one that is not monotone through $\frac{1}{2}$ and one that is. When I is an interval in $\mathbb{R}$, a correspondence is monotone through x if all values on the left of x are strictly higher than all the values on the left of x, or vice versa. Hence, a correspondence is monotone through all x in an open interval if and only if the correspondence is strictly monotone.

Proposition 5. *There exists no informative equilibrium if and only if* $|\Theta| = 2$ *and* V *is monotone through* μ_0.

Proposition 5 shows that state-independent cheap talk can almost always be informative. With three or more states, one can partition the state space into three. Lemma 3 then implies that there is an equilibrium in which S gives R information about said partition. It then remains to identify when is informative communication is impossible under a binary state. This turns out to be the case in which S unambiguously prefers to increase the probability that R assigns to one of the states (relative to the prior) rather than the other. Any message that decreases the belief of said state cannot then credibly be sent by S, meaning that no information can be communicated.

A similar proof gives Proposition 6 below. Proposition 6 provides necessary and sufficient conditions under which there exists an influential equilibrium in Chakraborty and Harbaugh (2010)'s setting. Chakraborty and Harbaugh (2010)'s setting has both the state and the action spaces being the same convex subset of $\mathbb{R}^N$, has a prior which admits a density, and has R choosing an action equal to his expectation of the state. Chakraborty and Harbaugh (2010) show that there is an influential equilibrium in this setting whenever $N \geq 2$. We show that the case of $N = 1$ behaves similarly to the binary state case of the previous proposition.

Proposition 6. *Suppose $A = \Theta \subseteq \mathbb{R}^N$ ($N \in \mathbb{N}$) is convex with nonempty interior, μ_0 admits a density, and $\forall \mu \in \Delta\Theta$:*

$$\arg\max_{a \in A} \int_\Theta u(a, \cdot)\ d\mu = \left\{ \int_\Theta \theta\ d\mu(\theta) \right\}$$

Then, there exists no influential equilibrium if and only if $N = 1$ and $\pi|_{\Theta^\circ}$ is monotone through $a_0 = \int_\Theta \theta\ d\mu_0(\theta)$.

Chakraborty and Harbaugh (2010) also show that, if one assumes that S's payoff is strictly quasiconvex in R's action, their setting is even better behaved. In particular, persuasive communication is easy for the sender to achieve, and its limits are straightforward for the analyst to delineate. The reason is that, in Chakraborty and Harbaugh (2010)'s setting, every influential equilibrium generates a mean preserving spread of R's actions. Strict quasiconvexity of S's payoff in R's action then immediately implies that every influential equilibrium is strictly persuasive. As such, the only case in which cheap talk is not valuable is when $\pi|_{\Theta^\circ}$ is monotone through $a_0 = \int_\Theta \theta\ \mathrm{d}\mu_0(\theta)$. Combining our monotonicity condition with the strong structure implied by strict quasiconvexity gives a concise condition for cheap talk to be valuable. We summarize this in Corollary 4 below.

Corollary 4. *Suppose the premise of Proposition 6 holds, and further suppose that* π *is strictly quasiconvex. Then cheap talk is valuable if and only if either* $N = 1$ *with* $\pi(\min \Theta)$, $\pi(\max \Theta) > \pi(a_0)$*, or* $N > 1$ *.*

The Corollary follows from Proposition 6, together with two observations. First, under strict quasiconvexity with a one-dimensional state, π being monotone through a_0 reduces to the simple condition: $\pi(a_0) < \pi(\min \Theta)$, $\pi(\max \Theta)$[14]. Second, as Chakraborty and Harbaugh (2010) note, quasiconvexity of the sender's objective immediately implies that every influential equilibrium benefits the sender (relative to babbling).

5.3 The informative content of persuasive cheap talk

We conclude our discussion with examining the informative content of persuasive cheap talk. To set the stage, we note that cheap talk generally constrains information. In particular, the proof of Theorem 2 tells us that, for every information policy that secures a high sender value, there is a less informative equilibrium policy which acheives it. There is therefore a sense in which acheiving a value through equilibrium involves communicating less information than securing it.

Corollary 5. *Sender value* $s \geq v(\mu_0)$ *is secured by* $p \in \mathcal{I}(\mu_0)$ *if and only if it can be secured by some equilibrium policy* p^* *which is weakly less Blackwell-informative than* p*.*

There are, however, cases in which cheap talk does not influence the most persuasive information policy. Consider the policy identified in the simple broker example from the introduction. This information policy is both a most persuasive cheap talk policy and the best policy for S under commitment. It turns out that the information policy's dual role comes from it inducing two

[14] This is because strict quasiconvexity in a one-dimensional space implies that each of $\pi|_{[\min \Theta, a_0]}$ and $\pi|_{[a_0, \max \Theta]}$ is maximized at one of the two end points of its domain.

consecutive payoffs for S. More precisely, say that $s > s^{'}$ are **consecutive** if $\pi(A) \cap \left[s^{'}, s\right] = \left\{s, s^{'}\right\}$. In addition, say p *is* **supported on two consecutive payoffs** if there are two consecutive s and $s^{'}$ such that $p\left\{\mu : v(\mu) \in \left\{s, s^{'}\right\}\right\} = 1$. Proposition 7 below shows that, whenever an information policy is both a most persuasive commitment policy and supported on two consecutive payoffs, then it is also a most persuasive cheap talk policy.

Proposition 7. *Let $\bar{p}$ be a most persuasive commitment information policy. If $\bar{p}$ is supported on two consecutive payoffs, then $\bar{p}$ is a most persuasive equilibrium information policy.*

To understand Proposition 7, consider a most persuasive commitment policy, $\bar{p}$, which is supported on two consecutive payoffs, $s > s^{'}$. The proof involves two steps. The first step is to show that $\bar{p}$ is an equilibrium outcome. Since $\bar{p}$ is optimal, whenever $\bar{p}$ induces R to take an s action, R must also be willing to take another action that gives S a lower payoff. Otherwise, one could increase the weight $\bar{p}$ places on s, by replacing s posteriors in which R is not indifferent with other posteriors that are closer to the prior. But R's indifference then means that R can replace the s action with a mixture that yields S a payoff of $s^{'}$, thereby making $\bar{p}$ incentive-compatible for S.

The second step is to show that $\bar{p}$ must be a most persuasive equilibrium policy. Suppose, for a contradiction, that some equilibrium yields sender payoff $s^{''}$ strictly larger than $s^{'}$. Since $\bar{p}$ is an optimal commitment policy that yields a payoffs in $\left(s^{'}, s\right)$, it must be that $s^{''}$ lies strictly below s, and is therefore strictly between $s^{'}$ and s. Since s and $s^{'}$ are consecutive, $s^{''}$ can only be obtained via R mixing. So, in the equilibrium in question, S mixes after *every* on-path message. But then, at every on-path belief, the sender has incentive compatible pure action which provides the sender a payoff strictly greater than s'', and therefore (since s'' lies between the two consecutive payoffs s and s'') at least as high as s. But this implies that p secures S a payoff of s, strictly larger than S's commitment payoff from $\bar{p}$, a contradiction.

6 Conclusion

In this paper, we study cheap talk under the assumption that the sender has a fixed agenda. Adopting a belief-based approach, we derive a complete characterization of feasible equilibrium outcomes. Our approach allows us to describe the sender-preferred equilibrium value geometrically, and to contrast this value with that of persuasion with commitment. The tractability of our model enables us to fully solve a rich class of examples, and to speak to the broader literature on cheap talk communication.

References

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

Aumann, Robert J, Michael Maschler, and Richard E Stearns, *Repeated games with incomplete information*, MIT press, 1995.

Baliga, Sandeep and Stephen Morris, "Co-ordination, spillovers, and cheap talk," *Journal of Economic Theory*, 2002, *105* (2), 450–468.

Benoit, J.P. and Juan Dubra, "Apparent Overconfidence," *Econometrica*, 2011.

Bergemann, Dirk and Stephen Morris, "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economics*, 2016, *11* (2), 487–522.

Chakraborty, Archishman and Rick Harbaugh, "Persuasion by cheap talk," *American Economic Review*, 2010, *100* (5), 2361–2382.

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

Ely, Jeffrey C, "Beeps," *American Economic Review*, 2017, *107* (1), 31–53.

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

Golosov, Mikhail, Vasiliki Skreta, Aleh Tsyvinski, and Andrea Wilson, "Dynamic strategic information transmission," *Journal of Economic Theory*, 2014, *151*, 304–341.

Green, Jerry R and Nancy L Stokey, "A two-person game of information transmission," *Journal of Economic Theory*, 2007, *135* (1), 90–104.

Hagenbach, Jeanne, Frédéric Koessler, and Eduardo Perez-Richet, "Certifiable Pre-Play Communication: Full Disclosure," *Econometrica*, 2014, *82* (3), 1093–1131.

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

Lugo, Andrés Salamanca, "The Value of Mediated Communication," 2016.

Margaria, Chiara and Alex Smolin, "Dynamic Communication with Biased Senders," 2016.

Mathis, Jérôme, "Full revelation of information in Sender–Receiver games of persuasion," *Journal of Economic Theory*, 2008, *143* (1), 571–584.

McCloskey, Donald and Arjo Klamer, "One Quarter of GDP is Persuasion," *The American Economic Review*, May 1995, *85* (2), 191–195.

Myerson, Roger B, "Incentive compatibility and the bargaining problem," *Econometrica: Journal of the Econometric Society*, 1979, pp. 61–73.

_ , "Multistage games with communication," *Econometrica: Journal of the Econometric Society*, 1986, pp. 323–358.

Perez-Richet, Eduardo, "Interim bayesian persuasion: First steps," *The American Economic Review*, 2014, *104* (5), 469–474.

Phelps, Robert R, *Lectures on Choquet's theorem*, Springer Science & Business Media, 2001.

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

Renault, Jérôme, Eilon Solan, and Nicolas Vieille, "Dynamic sender–receiver games," *Journal of Economic Theory*, 2013, *148* (2), 502–534.

A Appendix

A.1 Proofs for Section 2

A.1.1 Proof of Lemma 1

Proof. First take any equilibrium (σ, ρ, β), and let (p, s, y) be the induced outcome. That $p \in \mathcal{I}(\mu_0)$ follows directly from the Bayesian property. Define the interim payoff,

$$\begin{aligned} \hat{s} : M &\to \mathbb{R}^2 \\ m &\mapsto \hat{s}_m = \pi\left(\rho(m)\right). \end{aligned}$$

Sender incentive-compatibility tells us that there exists some $M^* \subseteq M$ such that $\int_\Theta \beta(M^*|\cdot)\ \mathrm{d}\mu_0 = 1$ and for every $m \in M^*$ and $m' \in M$, we have $\hat{s}_m \geq \hat{s}_{m'}$. This implies that $\hat{s}_m = \hat{s}_{m'}$ for every $m, m' \in M^*$; that is, there is some $\hat{s}^*$ such that $\hat{s}|_{M^*} = \hat{s}^*$. But then,

$$s = \int_\Theta \int_M \pi\left(\rho(m)\right)\ \mathrm{d}\sigma(m|\theta)\ \mathrm{d}\mu_0(\theta) = \int_\Theta \int_{M^*} \pi\left(\rho(m)\right)\ \mathrm{d}\sigma(m|\theta)\ \mathrm{d}\mu_0(\theta) = \int_\Theta \int_{M^*} \hat{s}^*\ \mathrm{d}\mu_0(\theta) = \hat{s}^*,$$

so that, by receiver incentive-compatibility, $s \in V(\beta(\cdot|m))$ for every $m \in M^*$. By definition of p, then, $s \in V(\mu)$ for p-almost every $\mu \in \Delta\Theta$. As V is upper hemicontinuous, it follows that $s \in \bigcap_{\mu \in \mathrm{supp}(p)} V(\mu)$.

Now, turning to the receiver,

$$\begin{aligned}
r &= \int_\Theta \int_M u\left(\rho(m), \theta\right)\ \mathrm{d}\sigma(m|\theta)\ \mathrm{d}\mu_0(\theta) \\
&= \int_\Theta \int_M u\left(\rho(m) \otimes \beta(m)\right)\ \mathrm{d}\sigma(m|\theta)\ \mathrm{d}\mu_0(\theta)\ \text{(by the Bayesian property)} \\
&= \int_\Theta \int_M U(\beta(m))\ \mathrm{d}\mu_0(\theta)\ \text{(by receiver incentive-compatibility)} \\
&= \int_{\Delta\Theta} U\ \mathrm{d}p,
\end{aligned}$$

as required.

Now, suppose (p, s, r) satisfies the three conditions. Define the compact set $B := \operatorname{supp}(p)$. It is well-known (see Benoit and Dubra (2011) or Kamenica and Gentzkow (2011)) that $p \in \mathcal{I}(\mu_0)$ implies there is some sender strategy σ and Bayes-consistent β that induce distribution p over posterior beliefs.[15] Without disrupting the Bayesian property, we may without loss assume that $\beta(m) \in B$ for all $m \in M$. Now, define the correspondence

$$\begin{aligned}
&\alpha^* : B \rightrightarrows \Delta A \\
&\qquad \mu \mapsto \{\alpha \in \Delta A :\ \pi(\alpha) = s\} \cap \arg\max_{\alpha \in \Delta A} u(\alpha \otimes \mu).
\end{aligned}$$

By hypothesis, α^* is nonempty-valued by definition of s. Continuity of u and Berge's theorem then tell us α^* is upper hemicontinuous. By the Kuratowski & Ryll-Nardzewski Theorem, α^* admits a measurable selection $\alpha : B \to \Delta A$. We can then define the receiver strategy $\sigma := \alpha \circ \beta$, which satisfies receiver incentive-compatibility by definition of α^*. By definition of α^*, $\int_{\Delta\Theta} U\ \mathrm{d}p = s$. Also, by construction, $\int_A \pi\ \mathrm{d}\rho(\cdot|m) = s$ for every $m \in M$, so that every sender strategy is incentive-compatible. Therefore, (σ, ρ, β) is an equilibrium generating outcome (p, s, r). □

[15] In particular, there is one with $\sigma(\Delta\Theta|\theta) = 1$ for all $\theta \in \Theta$ and $\beta(\cdot|\mu) = \mu$ for all $\mu \in \Delta\Theta$.

A.1.2 The quasiconcave envelope is well-defined

Lemma 4. *The function $\bar{v}$ is well-defined. That is, there exists a pointwise lowest quasiconcave upper semicontinuous function which majorizes v.*

Proof. Let $\mathcal{F} := \{f : \Delta\Theta :\ f \geq v,\ f \text{ is quasiconcave and upper semicontinuous}\}$, and let $\bar{v}$ be the pointwise infimum of $\mathcal{F}$. By construction, $\bar{v}$ lies above v and lies below any $f \in \mathcal{F}$. Moreover, any infimum of upper semicontinuous functions is itself upper semicontinuous, as a union of open sets is open. Thus, all that remains is to show that $\bar{v}$ as defined is quasiconcave. Given a function $f : \Delta\Theta \to \mathbb{R}$,

$$\begin{aligned} f \text{ is quasiconcave} &\iff \{\mu \in \Delta\Theta : f(\mu) \geq f(\tilde{\mu})\} \text{ is convex } \forall \tilde{\mu} \in \Delta\Theta \\ &\iff \{\mu \in \Delta\Theta : f(\mu) \geq s\} \text{ is convex } \forall s \in f(\Delta\Theta) \\ &\iff \{\mu \in \Delta\Theta : f(\mu) \geq s\} \text{ is convex } \forall s \in \mathbb{R}, \end{aligned}$$

where is the last equivalence holds because a union of a chain of convex sets is convex. Therefore,

$$\begin{aligned} \bar{v} \text{ is quasiconcave} &\iff \{\mu \in \Delta\Theta : \bar{v}(\mu) \geq s\} \text{ is convex } \forall s \in \mathbb{R} \\ &\iff \bigcap_{f \in \mathcal{F}} \{\mu \in \Delta\Theta : f(\mu) \geq s\} \text{ is convex } \forall s \in \mathbb{R} \end{aligned}$$

The result follows because an intersection of convex sets is convex. □

It is worth noting that, in the finite state case (just as in the case with commitment), the qualifier "upper semicontinuous" can be omitted.

Fact 1. *If $|\Theta| < \infty$, then $\bar{v}$ is the pointwise lowest quasiconcave function which majorizes v.*

Proof. Notice that:

$$\begin{aligned}
\bar{v}(\mu) &= \max\left\{ s: \ p \in \mathcal{I}(\mu), \ s \in \bigcap_{\nu \in \text{supp}(p)} V(\nu) \right\} \text{ (by Theorem 1)} \\
&= \max\left\{ s: \ \mu \in \overline{\text{co}} V^{-1}(s) \right\} \text{ (by Phelps (2001) Proposition 1.2)} \\
&= \max\left\{ s: \ \mu \in \text{co}\overline{V^{-1}(s)} \right\} \quad (\text{as } |\Theta| < \infty) \\
&= \max\left\{ s: \ \mu \in \text{co} V^{-1}(s) \right\} \quad (\text{as } V \text{ is upper hemicontinuous}).
\end{aligned}$$

Suppose $\tilde{v} : \Delta\Theta \to \mathbb{R}$ is a quasiconcave function which majorizes v, and take any $\mu \in \Delta\Theta$. By the above, there exists some finite-support $p \in \mathcal{I}(\mu)$ such that $\bar{v}(\mu) \in V(\nu)$ for every $\nu \in \text{supp}(p)$. Moreover, there exists no $\epsilon > 0$ such that $\bar{v}(\mu) + \epsilon \in V(\nu)$ for every $\nu \in \text{supp}(p)$. As V is a Kakutani correspondence, it must be that $\max V(\nu^*) = \bar{v}(\mu)$ for some $\nu^* \in \text{supp}(p)$. In particular, $v(\nu^*) = \bar{v}(\mu) \leq v|_{\text{supp}(p)}$. Therefore:

$$\tilde{v}(\mu) = \tilde{v}\left(\sum_{\nu \in \text{supp}(p)} p(\nu)\nu \right) \geq \min_{\nu \in \text{supp}(p)} \tilde{v}(\nu) \geq \min_{\nu \in \text{supp}(p)} v(\nu) = \bar{v}(\mu).$$

So, in the finite state world, $\bar{v}$ is indeed the pointwise lowest quasiconcave function above v. □

A.1.3 Theorem 1

Proof. Let $V^* : \Delta\Theta \rightrightarrows \mathbb{R}$ map any given prior to the associated set of equilibrium values for the sender, and let $v^* := \sup V^*$. Our goal is to show that $v^* = \bar{v}$.

First notice that $v \in V^*$, witnessed by a babbling equilibrium. Therefore, V^* is a nonempty-valued correspondence, and $v^* \geq v$.

Now, consider any sequence $(\mu_n, s_n)_n$ from the graph of V^* which converges to some $(\mu, s) \in (\Delta\Theta) \times \mathbb{R}$. For each n, Lemma 1 delivers some $p_n \in \mathcal{I}(\mu_n)$ such that $s_n \in V(\nu)$ for p_n-almost every $\nu \in \Delta\Theta$. By compactness, some sub-sequence of $(p_n)_n$ converges to some $p \in \Delta\Delta\Theta$. The correspondence $\mathcal{I}$

is upper hemicontinuous, so that $p \in \mathcal{I}(\mu)$. Continuity of u then tells us that $s \in V(\nu)$ for p-almost every $\nu \in \Delta\Theta$. Appealing to the above lemma again then tells us that $s \in V^*(\mu)$. Therefore, V^* has closed graph: it is an upper hemicontinuous, compact-valued correspondence. As two immediate by-products, we learn that an optimal information policy exists for every prior (as V^* is compact-valued), and that v^* is an upper semicontinuous function.

Next, consider any $\mu, \mu' \in \Delta\Theta$ with $s := v^*(\mu') \leq v^*(\mu)$, and any $\lambda \in [0,1]$. We now show that $v^*((1-\lambda)\mu' + \lambda\mu) \geq s$ too. If $v((1-\lambda)\mu' + \lambda\mu) \geq s$, then there is nothing to show because $v^* \geq v$; so assume to the contrary that $v((1-\lambda)\mu' + \lambda\mu) < s$.

Let us now show that $s \in V((1-\hat{\lambda})\mu' + \hat{\lambda}\mu)$ for some $\hat{\lambda} \in [\lambda, 1]$. Assume not, for a contradiction. Then, V being convex valued, it follows that $V((1-\hat{\lambda})\mu' + \hat{\lambda}\mu) \subseteq (-\infty, s)$ or $V((1-\hat{\lambda})\mu' + \hat{\lambda}\mu) \subseteq (s, \infty)$ for each $\hat{\lambda} \in [\lambda, 1]$. So, defining the continuous function $g := 0|_{(-\infty,s)} \cup 1|_{(s,\infty)} : \mathbb{R} \setminus \{s\} \to \mathbb{R}$, the function

$$\begin{aligned} [\lambda, 1] &\to \mathbb{R} \\ \hat{\lambda} &\mapsto g \circ V((1-\hat{\lambda})\mu' + \hat{\lambda}\mu) \end{aligned}$$

is a continuous function (as V is upper hemicontinuous) which maps the connected set $[\lambda, 1]$ onto the disconnected set $\{0, 1\}$, a contradiction.

So there is some $\hat{\lambda} \in [\lambda, 1]$ such that $s \in V((1-\hat{\lambda})\mu' + \hat{\lambda}\mu)$. By Lemma 1, there are then $p' \in \mathcal{I}(\mu')$ and $p \in \mathcal{I}((1-\hat{\lambda})\mu' + \hat{\lambda}\mu)$ such that $s \in V(\nu)$ for p'-almost every and p-almost every $\nu \in \Delta\Theta$. Applying Lemma 1 again, now to $\frac{\hat{\lambda}-\lambda}{\hat{\lambda}} p' + \frac{\lambda}{\hat{\lambda}} p$, tells us that $s \in V^*\left((1-\lambda)\mu' + \lambda\mu\right)$. Therefore, $v^*((1-\lambda)\mu' + \lambda\mu) \geq s$, and v^* is quasiconcave.

Finally, take any upper semicontinuous quasiconcave $f : \Delta\Theta \to \mathbb{R}$ which majorizes v and any prior $\mu \in \Delta\Theta$. By the lemma, there is some $p \in \mathcal{I}(\mu)$ such that $v^*(\mu) \in V(\nu)$ for p-almost every $\nu \in \Delta\Theta$. Therefore, for such ν, we have: $v^*(\mu) \leq v(\nu) \leq f(\nu)$. Letting $B \subseteq \Delta\Theta$ denote the support of p, upper semicontinuity of f implies $v^*(\mu) \leq f|_B$. But then, as f is quasiconcave and upper semicontinuous, $v^*(\mu) \leq f|_{\overline{\text{co}}B}$. As p witnesses $\mu \in \overline{\text{co}}B$, we learn that

$v^*(\mu) \leq f(\mu)$. But μ and f were arbitrary; it follows that $v^* \leq \bar{v}$.

And so we know that v^* is an upper semicontinuous, quasiconcave function with $v \leq v^* \leq \bar{v}$. It follows from the definition of $\bar{v}$ that $v^* = \bar{v}$, completing the proof. □

A.1.4 Theorem 2

Proof. The "if" direction follows directly from Lemma1: for any equilibrium outcome (p, s, r), information policy p secures payoff s. Now, we prove the "only if" direction. This is trivial (taking a sender-preferred babbling equilibrium) for $s = v(\mu_0)$, so focus on the case of $s > v(\mu_0)$.

Let $p \in \mathcal{I}(\mu_0)$ secure s and $B := \text{supp}\,(p)$, and notice that $v(\mu) \geq s$ for every $\mu \in B$ since v is upper semicontinuous. Define the measurable function,

$$\begin{aligned} \lambda : B &\to [0,1] \\ \mu &\mapsto \inf\left\{\hat{\lambda} \in [0,1] :\ v\left((1-\hat{\lambda})\mu_0 + \hat{\lambda}\mu\right) \geq s\right\}. \end{aligned}$$

As V is upper hemicontinuous, it must be that $s \in V([1-\lambda(\mu)]\mu_0 + \lambda(\mu)\mu)$ for every $\mu \in S$.

Notice that there is some number $\epsilon > 0$ such that $\lambda \geq \epsilon$ uniformly. If there were no such ϵ, then there would be a sequence $\{\mu_n\}_n \subseteq B$ such that $\lambda(\mu_n)$ converges to zero. But then, the sequence $\{([1-\lambda(\mu_n)]\mu_0 + \lambda(\mu_n)\mu_n,\ s)\}_n$ from the graph of V would converge to (μ_0, s). As V is upper hemicontinuous, this would contradict $s > v(\mu_0)$. Therefore, such an ϵ exists, and $\frac{1}{\lambda}$ is a bounded function.

Now, define $p^* \in \Delta\Delta\Theta$ by letting $p^*(\hat{B}) := (\int_{\Delta\Theta} \frac{1}{\lambda}\, \mathrm{d}p)^{-1} \cdot \int_{\Delta\Theta} \frac{1}{\lambda(\mu)} \mathbf{1}_{[1-\lambda(\mu)]\mu_0 + \lambda(\mu)\mu \in \hat{B}}\, \mathrm{d}p(\mu)$ for every Borel $\hat{B} \subseteq \Delta\Theta$. Direct computation shows that $p^* \in \mathcal{I}(\mu_0)$. By Lemma 1, there is an equilibrium generating sender value s and information policy p^*.

The last point comes from applying the above to $s = \bar{v}(\mu_0)$. □

For convenience, we record an easy consequence of the above theorem.

Corollary 6. *The set of sender payoffs s induced by some equilibrium is a compact interval.*

Proof. Let S^* be the set of equilibrium sender payoffs, $S_+ := \{s \in S^* : \ s \geq \max V(\mu_0)\}$, $S_- := \{s \in S^* : \ s \leq \min V(\mu_0)\}$, and $S_0 := \{s \in S^* : \ \min V(\mu_0) \leq s \leq \max V(\mu_0)\}$.

As V is convex-valued, $S_0 = S^* \cap V(\mu_0)$. By considering babbling equilibrium, we see that $S_0 = V(\mu_0) = [\min V(\mu_0), \max V(\mu_0)]$.

It follows immediately from Theorem 2 that S_+ is convex. From Theorem 1, S^* has a maximum value s_+. By an identical argument, S_- is convex and has a minimum value s_-.[16]

Therefore, $S^* = [s_-, \min V(\mu_0)] \cup [\min V(\mu_0), \max V(\mu_0)] \cup [\max V(\mu_0), s_+] = [s_-, s_+]$. □

A.2 Proofs for Section 3

A.2.1 Proposition 1

Proof. Let (p, s, r) be an equilibrium outcome, witnessed by equilibrium $\hat{\mathcal{E}} = (\hat{\sigma}, \hat{\rho}, \hat{\beta})$. Let $\mathbb{M} := \text{marg}_M \mathbb{P}_{\hat{\mathcal{E}}}$ $X := \text{supp}\,[\mathbb{M} \circ \hat{\rho}^{-1}] \subseteq \Delta A$, and fix arbitrary $(\hat{\alpha}, \hat{\mu}) \in \text{supp}\left[\mathbb{M} \circ (\hat{\rho}, \hat{\beta})^{-1}\right]$; in particular, $\hat{\alpha} \in X$. By continuity of $u$and receiver incentive-compatibility, $\hat{\alpha} \in \arg\max_{\alpha \in \Delta A} u(\alpha \otimes \hat{\mu})$. Defining $\rho' : M \to \Delta A$ [resp. $\beta' : M \to \Delta\Theta$] to agree with $\hat{\rho}$ [$\hat{\beta}$] on-path and take value $\hat{\alpha}$ [$\hat{\mu}$] off-path, there is then an equilibrium $\mathcal{E}' = (\hat{\sigma}, \rho', \beta')$ such that $\mathbb{P}_{\mathcal{E}'} = \mathbb{P}_{\hat{\mathcal{E}}}$ and $\rho'(\cdot|m) \in X$ for every $m \in M$.

[16] Notice that the only properties of V used in the proofs—that it is a Kakutani correspondence — is also true of $-V$.

Now, define:

$$
\begin{aligned}
\sigma : \Theta &\to \Delta X \subseteq \Delta M \\
\theta &\mapsto \hat{\sigma}(\cdot|\theta) \circ \rho'^{-1} \\
\rho : M &\to X \subseteq \Delta A \\
m &\mapsto \begin{cases} m & : \ m \in X \\ \hat{\alpha} & : \ m \notin X \end{cases} \\
\beta : M &\to \Delta\Theta \\
m &\mapsto \begin{cases} \mathbb{E}_{m \sim \mathbb{M}} \left[\beta(m) \middle| \rho(m) \right] & : \ m \in X \\ \hat{\mu} & : \ m \notin X. \end{cases}
\end{aligned}
$$

By construction, (σ, ρ, β) is an equilibrium generating outcome $(\tilde{p}, s, r)$ for some garbling $\tilde{p}$ of p. □

A.2.2 Proposition 2

Proof. Suppose cheap talk is costless. Letting (p, s, r) be the outcome induced by the equilibrium (σ, ρ, β), Proposition 1guarantees that $\int_{\Delta\Theta} v \; \mathrm{d}p = s$. This implies that $v(\mu) = s$ for p-almost every μ. Then $\pi^{-1}(s) \cap \arg\max_{a \in A} u(a, \mu)$ is nonempty for p-almost every μ. As u, π are continuous, the correspondence

$$
\begin{aligned}
\mathrm{supp}(p) &\rightrightarrows A \\
\mu &\mapsto \pi^{-1}(s) \cap \arg\max_{a \in A} u(a, \mu)
\end{aligned}
$$

is then nonempty-valued and upper hemicontinuous. By the Kuratowski & Ryll-Nardzewski Theorem, it therefore admits a measurable selection $\alpha : \mathrm{supp}(p) \to A$. Fix some $\hat{\mu} \in \mathrm{supp}(p)$. Now, define a new belief map

$$
\begin{aligned}
\tilde{\beta} : M &\to \Delta\Theta \\
m &\mapsto \begin{cases} \beta(m) & : \ \beta(m) \in \mathrm{supp}(p), \\ \hat{\mu} & : \ \text{otherwise}, \end{cases}
\end{aligned}
$$

and a new receiver strategy

$$\begin{aligned} \tilde{\rho} : M &\to \Delta\Theta \\ m &\mapsto \delta_{\alpha(\beta(m))}. \end{aligned}$$

By construction, $(\sigma, \tilde{\rho}, \tilde{\beta})$ is an equilibrium with no receiver mixing. □

A.2.3 Lemma 2

Proof. Suppose π is one-to-one, and that there is a most persuasive equilibrium (σ, ρ, β) with no mixing, which induces an outcome (p, s, r). Let $A^* := \{a \in A : \ \delta_a \in \operatorname{supp}[\operatorname{marg}_{\Delta A}\mathbb{P}]\}$. By sender incentive-compatibility, $\pi|_{A^*}$ is constant. But then, $A^* = \{a\}$ for some action a, as π is injective. By receiver incentive-compatibility, $a \in \arg\max_{\hat{a} \in A} u(\hat{a}, \mu)$ for every p-almost every μ. Therefore, $a \in \arg\max_{\hat{a} \in A} u(\hat{a}, \mu_0)$ too. This implies that $s \in V(\mu_0)$, so that a babbling equilibrium is most persuasive, i.e. cheap talk is not valuable. □

A.2.4 Corollary 2

Proof. Suppose cheap talk is valuable. Lemma 2 then tells us that every most persuasive equilibrium entails mixing by the receiver. But then, Proposition 2tells us that cheap talk is costly.

So if cheap talk is valuable, it is costly. Said differently: if cheap talk is costless, it is not valuable. That is, if $\hat{v}(\mu_0) = \bar{v}(\mu_0)$, then $\bar{v}(\mu_0) = v(\mu_0)$. The result follows directly. □

A.3 Proofs for Section 4

For this subsection, define $E : \Delta\Theta \to \mathrm{co}\Theta$ as the barycenter map, with $E\mu := \int_\Theta \theta \ \mathrm{d}\mu(\theta)$.

A.3.1 Claim 1

Proof. For any $a \in A \setminus \{\bar{a}\}$,

$$\frac{\mathrm{d}}{\mathrm{d}a}u(a,\mu) = -\int_{\Delta\Theta}(a-\theta)\ \mathrm{d}\mu(\theta) - \kappa \cdot \mathrm{sign}(a-\bar{a}) = [E\mu - \kappa \cdot \mathrm{sign}(a-\bar{a})] - a.$$

If $E\mu > \bar{a}+\kappa$, then $u(\cdot,\mu)$ is increasing on $[0,\bar{a}]$, so that $\arg\max_{a\in A} u(a,\mu) \subseteq [\bar{a},1]$. But $u(\cdot,\mu)$ is concave on $[\bar{a},1]$, so that the first-order approach is valid. Therefore, $\arg\max_{a\in A} u(a,\mu) = \{E\mu-\kappa\}$. Identical reasoning tells us that $\arg\max_{a\in A} u(a,\mu) = \{E\mu+\kappa\}$ when $E\mu < \bar{a}-\kappa$.

If $|E\mu| \leq \kappa$, then $u(\cdot,\mu)$ is increasing on $[0,\bar{a}]$ and decreasing on $[\bar{a},1]$, so that $\arg\max_{a\in A} u(a,\mu) = \{\bar{a}\}$. □

A.3.2 Claim 2

Proof. If $\mu \in \Delta\Theta$ and $a = a^*(\mu)$, then:

$$\begin{aligned}
U(\mu) - u(\bar{a},\mu_0) &= \tfrac{1}{2}\int_\Theta (\bar{a}^2 - 2\bar{a}\theta + \theta^2)\ \mathrm{d}\mu_0 - \tfrac{1}{2}\int_\Theta (a^2 - 2a\theta + \theta^2)\ \mathrm{d}\mu - \kappa|a-\bar{a}| \\
&= \int_\Theta \left(\tfrac{1}{2}\theta^2 - \bar{a}\theta\right)[\mathrm{d}\mu_0(\theta) - \mathrm{d}\mu(\theta)] \\
&\quad + \tfrac{1}{2}(\bar{a}^2 - a^2) + (a-\bar{a})\int_\Theta \theta\ \mathrm{d}\mu(\theta) - \kappa|a-\bar{a}| \\
&= \int_\Theta \left(\tfrac{1}{2}\theta^2 - \bar{a}\theta\right)[\mathrm{d}\mu_0(\theta) - \mathrm{d}\mu(\theta)] \\
&\quad + (a-\bar{a})\left[-\tfrac{1}{2}(\bar{a}+a) + \int_\Theta \theta\ \mathrm{d}\mu(\theta) - \kappa\cdot\mathrm{sign}(a-\bar{a})\right] \\
&= \int_\Theta \left(\tfrac{1}{2}\theta^2 - \bar{a}\theta\right)[\mathrm{d}\mu_0(\theta) - \mathrm{d}\mu(\theta)] + (a-\bar{a})\left[-\tfrac{1}{2}(\bar{a}+a) + a\right] \\
&= \int_\Theta \left(\tfrac{1}{2}\theta^2 - \bar{a}\theta\right)[\mathrm{d}\mu_0(\theta) - \mathrm{d}\mu(\theta)] + \tfrac{1}{2}(a-\bar{a})^2 \\
&= \int_\Theta \left(\tfrac{1}{2}\theta^2 - \bar{a}\theta\right)[\mathrm{d}\mu_0(\theta) - \mathrm{d}\mu(\theta)] + \tfrac{1}{2}[\pi(a)]^2.
\end{aligned}$$

For any $p \in \mathcal{I}(\mu_0)$ with $v|_{\text{supp}(p)} = \pi \circ a^*|_{\text{supp}(p)} = s$:

$$\begin{aligned}\int_{\Delta\Theta} U \text{ d}p &= u(\bar{a}, \mu_0) + \tfrac{1}{2}s^2 + \int_{\Delta\Theta}\int_{\Theta} \left(\tfrac{1}{2}\theta^2 - \bar{a}\theta\right)[\text{d}\mu_0(\theta) - \text{d}\mu(\theta)] \text{ d}p(\mu) \\ &= u(\bar{a}, \mu_0) + \tfrac{1}{2}s^2,\end{aligned}$$

where the last equality holds because $p \in \mathcal{I}(\mu_0)$ and integration against $\theta \mapsto \frac{1}{2}\theta^2 - \bar{a}\theta$ is linear and continuous. The claim then follows directly from Lemma 1. □

A.3.3 Proposition 3

Proof. If $s = v(\mu_0)$, then $p^* = \delta_{\mu_0}$ proves the claim, so suppose without loss that $s > v(\mu_0)$. Define:

$$\begin{aligned}\theta_0 &:= a^*(\mu_0), \\ \theta_L &:= \max\left\{\theta \in \text{co}\Theta: \ \theta < \theta_0 \text{ and } s \in \pi\left(a^*\left(E\mu\right)\right)\right\}, \\ \theta_R &:= \min\left\{\theta \in \text{co}\Theta: \ \theta > \theta_0 \text{ and } s \in \pi\left(a^*\left(E\mu\right)\right)\right\}.\end{aligned}$$

As p secures $s > v(\mu_0)$ and v is upper semicontinuous, we know that both θ_L, θ_R are well-defined and distinct from $E(\mu_0)$, and that $p \circ E^{-1}\left[\Theta \setminus (\theta_L, \theta_R)\right] = 0$. Bayes plausibility then tells us that $q := p \circ E^{-1}(-\infty, \theta_L) \in (0, 1)$. Let $p' = q\delta_{\mu_L} + (1-q)\delta_{\mu_R}$ denote the q-cutoff policy.

Now, notice that, $\mu_L \in \arg\min_{\mu \in \Delta\Theta,\ q\mu \leq \mu_0} E(\mu)$, which implies that

$$qE(\mu_L) \leq \int_{E^{-1}(-\infty, \theta_L)} E \text{ d}p \leq \int_{E^{-1}(-\infty, \theta_L)} \theta_L \text{ d}p = q\theta_L.$$

That is, $E(\mu_L) \leq \theta_L$. An identical argument shows that $E(\mu_R) \geq \theta_R$.

It follows that the two-posterior information policy p' has some (necessarily two-posterior) garbling $p^* \in \mathcal{I}(\mu_0)$ such that $p^* \circ E^{-1}$ is supported on.$\{\theta_L, \theta_R\}$. As a^* is upper hemicontinuous and $\pi$is continuous, it follows that value s∈ $V(\mu)$ for both $\mu \in \text{supp}(p)$. Lemma 1 then tells us that $(p^*, s, \int_{\Delta\Theta} U \ dp^*)$ is an equilibrium outcome as required. □

A.3.4 Corollary 3

In the event that v is quasiconvex, any information that can be credibly conveyed is of benefit to the sender. Quasiconvexity is therefore sufficient, in light of the above proposition, for a simple cutoff policy to arise in a most persuasive equilibrium.

Proof. As a babbling equilibrium is a cutoff equilibrium (with cutoff 0 or 1, equivalently), there is nothing to show if $\bar{v}(\mu_0) = v(\mu_0)$. So assume without loss that $s := \bar{v}(\mu_0) > v(\mu_0)$. By hypothesis, there is some $v^* : \text{co}\Theta \to \mathbb{R}$ such that $v = v^* \circ E$. Moreover, v^* is weakly quasiconvex and upper semicontinuous because v is. By the above proposition, the securable sender value $\bar{v}(\mu_0)$ can be secured by a garbling p of a cutoff policy, such that $p \circ E^{-1}$ is supported on $\underline{\theta}, \bar{\theta}$ for some $\min\Theta \leq \underline{\theta} \leq \bar{\theta} \leq \max\Theta$. As $s > v^* \circ E(\mu_0)$ and v^* is quasiconvex, it follows that v^* is weakly decreasing on $[\min\Theta, \underline{\theta}]$ and weakly increasing on $[\bar{\theta}, \max\Theta]$. Therefore, value s can be secured by the cutoff policy itself, rather than the garbling p.

We now show that there exists an equilibrium yielding sender payoff s and a cutoff policy as its belief distribution. For any $q \in (0,1)$, let $p^q = q\delta_{\mu_L^q} + (1-q)\delta_{\mu_R^q}$ be the q-cutoff policy. For $i \in \{L, R\}$, define

$$\begin{aligned} f_i : (0,1) &\to \mathbb{R} \\ q &\mapsto E(\mu_i^q). \end{aligned}$$

Notice that f_L, f_R are both strictly increasing and continuous, with $f_L < f_R$ globally. As v^* is quasiconvex, there is some $\hat{\theta} \in \text{co}\Theta$ such that v^* is weakly decreasing on $[\min\Theta, \hat{\theta}]$ and weakly increasing on $[\hat{\theta}, \max\Theta]$. The set $Q := \left\{q \in (0,1) : \ f_L(q) < \hat{\theta} < f_H(q)\right\}$ is therefore open, convex, and nonempty. Also notice that, as some cutoff policy secures $s > \min v^*[0,1]$, it must be that $\min\Theta < \hat{\theta} < \max\Theta$. Therefore, $Q = (q_0, q_1)$ for some $0 < q_0 < q_1 < 1$. Finally, we know that any cutoff policy which secures s must have a cutoff belonging to Q.

Now, $g := v^* \circ f_R - v^* \circ f_L : (0,1) \to \mathbb{R}$ is weakly increasing on Q, so that there exists $q^* \in [q_0, q_1]$ such that $g|_{[q_0,q^*]} \le 0 \le g|_{[q^*,q_1]}$. Therefore, $\min\{v^* \circ f_L,\ v^* \circ f_R\}$ is weakly increasing on $[q_0, q^*]$ and weakly decreasing on $[q^*, q_1]$. We conclude (because some cutoff policy does) that the q^*-cutoff policy secures s. Moreover, the Kakutani property of V ensures that $s \in V(\mu_L^{q^*}) \cap V(\mu_R^{q^*})$. The result now follows from Lemma 1. □

A.3.5 Claim 4

Proof. For any $q \in (0,1)$, let $p^q = q\delta_{\mu_L^q} + (1-q)\delta_{\mu_R^q}$ be the q-cutoff policy.

First, suppose $E\mu_0 \ge 2\bar{a}$. Then, for any $q \in (0,1)$,

$$[E(\mu_R^q) - \bar{a}] - [\bar{a} - E(\mu_L^q)] = E(\mu_L^q) + E(\mu_R^q) - 2\bar{a} > 0 + E\mu_0 - 2\bar{a} \ge 0.$$

Therefore, there can be no equilibrium generating a q-cutoff policy unless it satisfies $|E(\mu_R^q) - \bar{a}|, |E(\mu_L^q) - \bar{a}| \le \kappa$. Therefore, appealing to Corollary 3, no equilibrium is more persuasive than a babbling equilibrium.

Now, suppose $E\mu_0 < 2\bar{a}$. Observe that $E(\mu_L^q)$ and $E(\mu_R^q)$ are both strictly increasing and continuous in q, and that $E(\mu_L^q) + E(\mu_R^q)$ is $< 2\bar{a}$ for $q \approx 0$ and $> 2\bar{a}$ for $q \approx 1$ (by hypothesis, and since $\bar{a} \le \frac{1}{2}$). Intermediate value theorem then delivers a unique $q \in (0,1)$ such that $E(\mu_R^q) - \bar{a} = \bar{a} - E(\mu_L^q)$. From the functional form of v and by Lemma 1, there is an equilibrium outcome of the form (s, r, p^q). Any equilibrium inducing a cutoff policy $p^{\hat{q}}$ for some $\hat{q} \ne q$ would have to be (by the functional form of v again) yield S a payoff of zero. Therefore, appealing to Corollary 3, (s, r, p^q) is a most persuasive equilibrium outcome. □

A.4 Proofs for Section 5

A.4.1 Proof of Proposition 4

Proof. First, let us show that there exists a fully revealing equilibrium if and only if there exists an equilibrium outcome (p, s, r) with $p\{\delta_\theta :\ \theta \in \Theta\} = 1$. The "only if" direction is immediate. For the "if direction", suppose $\hat{\mathcal{E}} = (\hat{\sigma}, \hat{\rho}, \hat{\beta})$ is an equilibrium generating belief distribution p with $p\{\delta_\theta :\ \theta \in \Theta\} =$

1. Fix some $(\hat{\alpha}, \hat{\mu}) \in \text{supp}\left[(\text{marg}_{\text{M}}\mathbb{P}_{\hat{\mathcal{E}}}) \circ (\hat{\rho}, \hat{\beta})^{-1}\right]$. Defining $\rho : M \to \Delta A$ [resp. $\beta : M \to \Delta\Theta$] to agree with $\hat{\rho}$ [$\hat{\beta}$] on-path and take value $\hat{\alpha}$ [$\hat{\mu}$] off-path, there is then an equilibrium $(\hat{\sigma}, \rho, \beta)$ such that $\mathbb{P}_{\hat{\sigma},\rho,\beta} = \mathbb{P}_{\hat{\mathcal{E}}}$ and $\beta(\cdot|m) \in \text{supp}\left[(\text{marg}_{\text{M}}\mathbb{P}_{\hat{\mathcal{E}}}) \circ \hat{\beta}^{-1}\right] = \{\delta_\theta\}_{\theta\in\Theta}$ for every $m \in M$.

As equilibrium belief distributions must belong to $\mathcal{I}(\mu_0)$ by Lemma 1, an equilibrium is fully revealing if and only if it induces a belief distribution of p, where $p(B) = \mu_0 \{\theta \in \Theta : \ \delta_\theta \in B\}$ for every Borel $B \subseteq \Delta\Theta$. But then, by Lemma 1, such an equilibrium exists if and only if there exists some sender payoff s such that $s \in V(\mu)$ for every $\mu \in \text{supp}(p)$. As μ_0 is of full support, $\text{supp}(p) = \{\delta_\theta\}_{\theta\in\Theta}$, so that existence of fully revealing equilibrium is equivalent to

$$\emptyset \neq \bigcap_{\theta\in\Theta} V(\delta_\theta) = \bigcap_{\theta\in\Theta} [\min V(\delta_\theta),\ \max V(\delta_\theta)] = \left[\max_{\theta\in\Theta} \min V(\delta_\theta),\ \min_{\theta\in\Theta} \max V(\delta_\theta)\right].$$

The result follows directly. □

A.4.2 Proof of Lemma 3

Proof. Let $\epsilon := \min\left\{p_1,\ p_2,\ \frac{1}{\sqrt{2}}p_3\right\}$. We can embed the circle $\mathbb{S} = \{(x, y) \in \mathbb{R}^2 :\ x^2 + y^2 = 1\}$ in the space of beliefs via the map

$$\begin{aligned} \phi &: \mathbb{S} \to \Delta\Theta \\ (x, y) &\mapsto (p_1 + \epsilon x)\mu_1 + (p_2 + \epsilon y)\mu_2 + [p_3 - \epsilon(x + y)]\mu_3. \end{aligned}$$

Next, define the function

$$\begin{aligned} f &: \mathbb{S} \to \mathbb{R} \\ z &\mapsto \max V(\phi(z)) - \min V(\phi(-z)). \end{aligned}$$

Two properties of f are immediate. First, f is upper semicontinuous because V is is upper hemicontinuous. Second, any $z \in \mathbb{S}$ satisfies $f(z) + f(-z) \geq 0$ because $\max V \geq \min V$.

We use the above two properties to show that there is some $z \in \mathbb{S}$ for which

$f(z)$ and $f(-z)$ are both non-negative. Assume otherwise, for a contradiction. Then every $z \in \mathbb{S}$ has the property that one of $f(z), f(-z)$ is strictly negative and (as $f(z) + f(-z) \geq 0$) the other strictly positive; in particular, f is globally nonzero. Since $\mathbb{S}$ is connected, there is then some $z \in \mathbb{S}$ which is a limit point of both $f^{-1}((0, \infty))$ and $f^{-1}((-\infty, 0))$. But then $-z$ shares this same property. Finally, upper semicontinuity tells us that both $f(z)$ and $f(-z)$ are non-negative, a contradiction.

Now, we have $z \in \mathbb{S}$ with $f(z), f(-z) \geq 0$. That is, $\max V(\phi(z)) \geq \min V(\phi(-z))$ and $\max V(\phi(-z)) \geq \min V(\phi(z))$. Said differently (recall, V is convex-valued), $V(\phi(z)) \cap V(\phi(-z)) \neq \emptyset$. Lemma 1 then guarantees existence of an equilibrium generating information policy $\frac{1}{2}\delta_{\phi(z)} + \frac{1}{2}\delta_{\phi(-z)}$. □

A.4.3 Proof of Proposition 5

Proof. First, suppose $|\text{supp}(\mu_0)| > 2$. Then, letting $E, F \subseteq \Theta$ be sufficiently small disjoint open neighborhoods around two distinct points in the support of μ_0, we may assume $\mu_0(E),\ \mu_0(F),\ \mu(\Theta \setminus [E \cup F]) > 0$. The lemma then yields an informative equilibrium.

For the rest of the proof, assume $|\text{supp}(\mu_0)| = 2$. To conserve notation, assume without loss that $\text{supp}(\mu_0) = \{0, 1\}$, and identify $\Delta\{0, 1\}$ with $[0, 1]$ in the obvious way. Assume without loss that $v(0) \leq v(1)$.

Suppose $V|_{\Delta[\text{supp}(\mu_0)]}$ does not change level at μ_0. Then there exist some $0 \leq \underline{y} < x < \bar{y} \leq 1$ such that $\max V(\underline{y}) > \min V(\bar{y})$. Define $\phi : [0, 1] \to [0, 1]^2$ via $\phi(\lambda) := (1 - \lambda)(\underline{y}, \bar{y}) + \lambda(0, 1)$, and define $G : [0, 1] \rightrightarrows \mathbb{R}$ via $G(\lambda) := V(\phi_2(\lambda)) - V(\phi_1(\lambda))$. Notice, G is a Kakutani correspondence because V is. By assumption, $G(0)$ contains a negative number and $G(1)$ contains a non-negative number. Therefore, $G(\lambda) \ni 0$ for some $\lambda \in [0, 1]$. Then there exists an equilibrium with support $\{\phi_1(\lambda), \phi_2(\lambda)\}$, which is therefore informative.

Suppose $V|_{\Delta[\text{supp}(\mu_0)]}$ is monotone through μ_0. Any informative equilibrium would have some $\underline{y} \in [0, \mu_0)$ and some $\bar{y} \in (\mu_0, 1]$ in its support. By hypothesis, $V(\underline{y}) \cap V(\bar{y}) = \emptyset$, so that Lemma 1 tells us it cannot be an equilibrium. This completes the proof. □

A.4.4 Proof of Proposition 6

Proof. First suppose $N > 1$. Then there exists a partition $\Theta = \sqcup\{E_1, E_2, E_3\}$ into three positive measure events such that a_1, a_2, a_3 are non-colinear, where a_i is the barycentre of μ_0 conditioned on E_i.[17] Lemma 3 then yields distinct $q, q' \in \Delta\{1,2,3\}$ and an equilibrium whose associated information policy has support $\left\{\sum_{i=1}^3 q_i\mu_i,\ \sum_{i=1}^3 q_i'\mu_i\right\}$. As a_1, a_2, a_3 are not co-linear, the barycentres $\sum_{i=1}^3 q_i a_i,\ \sum_{i=1}^3 q_i' a_i$ are distinct. Given the hypothesis that the receiver optimally chooses his expectation, such an equilibrium is influential.

For the rest of the proof, assume $N = 1$. To conserve notation, assume without loss that $\Theta = [0,1]$ and $\pi(0) \leq \pi(1)$.

Suppose $\pi|_{(0,1)}$ does not change level at a_0. We consider three exhaustive cases. (i) There exist $a_L, a_H \in (0,1)$ such that $a_L < a_0 < a_R$ and $\pi(a_L) = \pi(a_0) = \pi(a_H)$. (ii) There is no $a \in (a_0, 1)$ with $\pi(a) = \pi(a_0)$. (iii) There is no $a \in (0, a_0)$ with $\pi(a) = \pi(a_0)$.

Case (i): Without loss of generality, assume that $\frac{1}{\mu_0([0,a_L]\cup[a_H,1])}\int_{[0,a_L]\cup[a_H,1]} \theta \,\mathrm{d}\mu_0(\theta) \geq a_0$. By intermediate value theorem, there is some $\hat{a}_H \in [a_H, 1]$ such that

$$\tfrac{1}{\mu_0([0,a_L]\cup[\hat{a}_H,1])}\int_{[0,a_L]\cup[\hat{a}_H,1]} \theta \,\mathrm{d}\mu_0(\theta) = a_0.$$

So let

$$\begin{aligned}
\mu_L &:= \tfrac{1}{\mu_0([0,a_L])}\mu_0([0,a_L]\cap\cdot),\\
\mu_M &:= \tfrac{1}{\mu_0((a_L,\hat{a}_H))}\mu_0((a_L,\hat{a}_H)\cap\cdot),\\
\mu_H &:= \tfrac{1}{\mu_0([\hat{a}_H,1])}\mu_0([\hat{a}_H,1]\cap\cdot).
\end{aligned}$$

Let $p^* \in \mathcal{I}(\mu_0)$ denote the information policy induced by fully reveal-

[17] Fix an exposed point $a_1^0 \in \Theta$, and let $\theta_0 := \int_\Theta \theta \,\mathrm{d}\mu_0(\theta) \in \Theta^\circ$. There exist affine functions $\phi, \psi : \mathbb{R}^N \to \mathbb{R}$ such that $\Theta \cap \phi^{-1}(0) = \{a_1^0\}$, $\phi(\theta_0) = 1$, and $\psi(a_1^0) = \psi(\theta_0) = 0$. For any $\epsilon \in [0,1)$, let $E_1^\epsilon := \Theta \cap \phi^{-1}([0,\epsilon))$, $E_2^\epsilon := \Theta \cap \psi^{-1}([0,\infty))$, and $E_3^\epsilon := \Theta \setminus (E_1^\epsilon \cup E_2^\epsilon)$. For any $\epsilon \in [0,1)$ and $i \in \{1,2,3\}$ with $(\epsilon, i) \neq (0,1)$: as $\mu_0(E_i^\epsilon) > 0$, we can define $\mu_i^\epsilon := \frac{1}{\mu_0(E_i^\epsilon)}\mu_0(E_i^\epsilon \cap \cdot) \in \Delta\Theta$ and $a_i^\epsilon := \int_\Theta \theta \,\mathrm{d}\mu_i^\epsilon(\theta) \in E_i^\epsilon$. Notice that $\lim_{\epsilon\searrow 0}(a_1^\epsilon, a_2^\epsilon, a_3^\epsilon) = (a_1^0, a_2^0, a_3^0)$ is a colinear triple. Therefore, we can take $(E_1, E_2, E_3) = (E_1^\epsilon, E_2^\epsilon, E_3^\epsilon)$ for sufficiently small $\epsilon \in (0,1)$.

ing which member of the partition $\{[0, a_L], (a_L, \hat{a}_H), [\hat{a}_H, 1]\}$ contains the true state; it has support $\{\mu_L, \mu_M, \mu_H\}$. Let $a_i^* := \int_\Theta \theta \, \mathrm{d}\mu_i(\theta)$ for $i = L, M, H$. By construction, we know that $a_L^* \leq a_L$, $a_M^* = a_0$, and $a_H^* \geq a_H$. There is therefore some garbling p of p^* (one which fully reveals whether p^* would yield posterior μ_M, but only partially reveals whether p^* would yield posterior μ_L or μ_H) which induces actions a_L, a_0, a_H with positive probability. By Lemma 1, p is then an influential equilibrium.

Case (ii): Since π is continuous, $\pi((a_0, 1)) \subseteq (\pi(a_0), \infty)$. As π does not change level at a_0, we may assume that there are $a < a_0 < a'$ such that $\pi(a), \pi\left(a'\right) > \pi(a)$. By the intermediate value theorem, for sufficiently large $k \in \mathbb{N}$, we have $\pi(a_0) + \frac{1}{k} \in \pi([0, a_0)) \cap \pi((a_0, 1])$. So let $a_k^L := \max\left\{a \in [0, a_0] : \ \pi(a) \leq \pi(a_0) + \frac{1}{k}\right\}$ and $a_k^H := \min\left\{a \in [a_0, 1] : \ \pi(a) \leq \pi(a_0) + \frac{1}{k}\right\}$, which exist because π is continuous; also by continuity, $\pi(a_k^L) = \pi(a_k^L) = \pi(a_0) + \frac{1}{k}$. Also, notice that $(a_k^H)_k$ is strictly decreasing and (as $\pi((a_0, 1]) \not\ni \pi(a_0)$) converges to a_0; and $(a_k^L)_k$ is strictly increasing and therefore converges to some $a_\infty^L \in (0, a_0]$. Let $p^* \in \mathcal{I}(\mu_0)$ denote the information policy induced by fully revealing which member of the partition $\left\{[0, a_\infty^L], (a_\infty^L, 1]\right\}$ contains the true state, and let $a_1, a_2 \in [0, 1]$ with $a_1 < a_2$ be the induced actions from receiver best responding to p^*-supported posteriors. Then $a_1 < a_\infty^L$ and $a_2 > a_0$. For sufficiently large $k \in \mathbb{N}$, we have $a_1 < a_k^L < a_0 < a_k^H < a_2$. There is therefore some garbling p of p^* which induces actions a_k^L, a_k^H with positive probability. By Lemma 1, p is then an influential equilibrium.

Case (iii): By an argument analogous to that for (ii), there exists an influential equilibrium.

Finally, if $\pi|_{(0,1)}$ is monotone through a_0, then any influential equilibrium would have some in its support some belief with barycentre $\underline{a} \in (0, a_0)$ and some belief with barycentre $\bar{a} \in (a_0, 1)$. But then $\pi(\underline{a}) \neq \pi(\bar{a})$ by hypothesis, so that Lemma 1 tells us it cannot be an equilibrium. This completes the proof. □

A.4.5 Proof of Corollary 5

Proof. This statement verbatim is proven constructively in the proof of Theorem 2. □

A.4.6 Proof of Proposition 7

Proof. Suppose $\bar{p}$ is a most persuasive commitment policy that is supported on two consecutive actions. Let $B = \text{supp}\,(\bar{p})$. If $v\,(B) = \{s\}$, then Lemma 1 directly applies to tell us $(\bar{p}, s, \int_{\Delta\Theta} U \,\mathrm{d}\bar{p})$ is an equilibrium outcome. Now, suppose that $v\,(B) = \left\{s, s^{'}\right\}$ for $s > s^{'}$. Let:

$$B_0 = \{\mu \in B : \; V\,(\mu) = \{s\}\}$$
$$B_1 = \left\{\mu \in B : \; s^{'} \in V\,(\mu)\right\}.$$

As s and $s^{'}$ are consecutive, $s > s'$, and V convex-valued, it must be that $B_0 \cup B_1 = B$.

Suppose now, for a contradiction, that $B_0 \neq \varnothing$. For every $\epsilon \in (0, 1)$, define

$$B_\epsilon := \{\mu \in B_0 : V\,((1-\epsilon)\,\mu + \epsilon\mu_0) \subseteq (s', \infty)\} = \{\mu \in B_0 : V\,((1-\epsilon)\,\mu + \epsilon\mu_0) \subseteq [s, \infty)\}\,.$$

Upper hemicontinuity of V implies that $\bar{p}\,(B_\epsilon) > 0$ for sufficiently small $\epsilon \in (0, 1)$. Let $\lambda := 1 - \epsilon\,[1 - \bar{p}\,(B_\epsilon)] \in (0, 1)$, and define $p := \frac{1}{\lambda}\,[(1-\epsilon)\,\bar{p}\,(\cdot \backslash B_\epsilon) + \bar{p}\,(\cdot \cap B_\epsilon)] \in \Delta\Delta\Theta$. Next, let

$$f : \Delta\Theta \to \Delta\Theta$$
$$\mu \mapsto \mu + \epsilon \mathbf{1}_{\mu \in B_\epsilon}\,(\mu_0 - \mu)\,.$$

For $g \in \{v, \; \text{id}_{\Delta\Theta}\}$, we have

$$
\begin{aligned}
\int_{\Delta\Theta} g \,\mathrm{d}\left[p \circ f^{-1}\right] &= \int_{(\Delta\Theta)\setminus B_\epsilon} g \,\mathrm{d}p + \int_{B_\epsilon} g\left(\left(1-\epsilon\right)\mu + \epsilon\mu_0\right)\,\mathrm{d}p\left(\mu\right) \\
&= \tfrac{1-\epsilon}{\lambda}\int_{(\Delta\Theta)\setminus B_\epsilon} g \,\mathrm{d}\bar{p} + \tfrac{1}{\lambda}\int_{B_\epsilon} g\left(\left(1-\epsilon\right)\mu + \epsilon\mu_0\right)\,\mathrm{d}\bar{p}\left(\mu\right) \\
\implies \tfrac{\lambda}{\bar{p}(B_\epsilon)}\left(\int_{\Delta\Theta} g\,\mathrm{d}\left[p\circ f^{-1}\right] - \int_{\Delta\Theta} g\,\mathrm{d}\bar{p}\right) &= \tfrac{1}{\bar{p}(B_\epsilon)}\int_{B_\epsilon}\left[g\left(\left(1-\epsilon\right)\mu + \epsilon\mu_0\right) - (1-\epsilon)g(\mu)\right]\,\mathrm{d}\bar{p}\left(\mu\right) \\
&\quad -\epsilon\int_{\Delta\Theta} g\,\mathrm{d}\bar{p}.
\end{aligned}
$$

This yields two by-products. First, it tells us that $p \circ f^{-1} \in \mathcal{I}\left(\mu_0\right)$, as

$$
\begin{aligned}
\tfrac{\lambda}{\bar{p}(B_\epsilon)}\left(\int_{\Delta\Theta} \mu\,\mathrm{d}\left[p\circ f^{-1}\right]\left(\mu\right) - \mu_0\right) &= \tfrac{1}{\bar{p}(B_\epsilon)}\int_{B_\epsilon}\left[\left(\left(1-\epsilon\right)\mu + \epsilon\mu_0\right) - (1-\epsilon)\mu\right]\,\mathrm{d}\bar{p}\left(\mu\right) \\
&\quad -\epsilon\int_{\Delta\Theta}\mu\,\mathrm{d}\bar{p}(\mu) \\
&= \epsilon\mu_0 - \epsilon\mu_0 = 0.
\end{aligned}
$$

Next, it tells us that $\int_{\Delta\Theta} v\,\mathrm{d}\left[p\circ f^{-1}\right] > \int_{\Delta\Theta} v\,\mathrm{d}\bar{p}$, as

$$
\begin{aligned}
\tfrac{\lambda}{\bar{p}(B_\epsilon)}\left(\int_{\Delta\Theta} v\,\mathrm{d}\left[p\circ f^{-1}\right] - \int_{\Delta\Theta} v\,\mathrm{d}\bar{p}\right) &= \tfrac{1}{\bar{p}(B_\epsilon)}\int_{B_\epsilon}\left[v\left(\left(1-\epsilon\right)\mu + \epsilon\mu_0\right) - (1-\epsilon)v(\mu)\right]\,\mathrm{d}\bar{p}\left(\mu\right) \\
&\quad -\epsilon\int_{\Delta\Theta} v\,\mathrm{d}\bar{p} \\
&\geq \tfrac{1}{\bar{p}(B_\epsilon)}\int_{B_\epsilon}\left[s - (1-\epsilon)v(\mu)\right]\,\mathrm{d}\bar{p}\left(\mu\right) - \epsilon\int_{\Delta\Theta} v\,\mathrm{d}\bar{p} \\
&> \tfrac{1}{\bar{p}(B_\epsilon)}\int_{B_\epsilon}\left[s - (1-\epsilon)s\right]\,\mathrm{d}\bar{p}\left(\mu\right) - \epsilon s \\
&= \epsilon s - \epsilon s = 0.
\end{aligned}
$$

This contradicts $\bar{p}$ being optimal under commitment.

Therefore $p(B_1) = 1$, implying that $(\bar{p}, s', \int_{\Delta\Theta} U \,\mathrm{d}p)$ is an equilibrium outcome, by Lemma 1. Now, let (p^*, s^*, r^*) be a most persuasive equilibrium outcome. It cannot be that $s^* \in (s', s)$. Indeed, as $\{s, s'\}$ are consecutive and V is convex-valued with $v(\Delta\Theta) \subseteq \pi(A)$, having $(s', s) \cap \bigcap_{\mu\in\mathrm{supp}(p)} V(\mu) \neq \emptyset$

would imply $\bigcap_{\mu \in \mathrm{supp}(p)} V(\mu) \ni s$, contradicting (by Lemma 1) the equilibrium outcome (p^*, s^*, r^*) being most persuasive. It also cannot be that $s^* \geq s$, since that would mean $\int_{\Delta\Theta} v \ \mathrm{d}\bar{p} < \int_{\Delta\Theta} v \ \mathrm{d}p^*$, contradicting the definition of $\bar{p}$. Therefore, $s^* \leq s'$, proving that $(\bar{p}, s', \int_{\Delta\Theta} U \ \mathrm{d}p)$ is a most persuasive equilibrium. □