Econometrica, Vol. 88, No. 4 (July, 2020), 1631–1660

CHEAP TALK WITH TRANSPARENT MOTIVES

ELLIOT LIPNOWSKI
Department of Economics, Columbia University

DORON RAVID
Department of Economics, University of Chicago

We study a model of cheap talk with one substantive assumption: The sender's preferences are state independent. Our main observation is that such a sender gains credibility by degrading self-serving information. Using this observation, we examine the sender's benefits from communication, assess the value of commitment, and explicitly solve for sender-optimal equilibria in three examples. A key result is a geometric characterization of the value of cheap talk, described by the *quasi*concave envelope of the sender's value function.

KEYWORDS: Cheap talk, belief-based approach, securability, quasiconcave envelope, persuasion, information transmission, information design.

1. INTRODUCTION

HOW MUCH CAN AN EXPERT BENEFIT from strategic communication with an uninformed agent? A large literature, starting with Crawford and Sobel (1982) and Green and Stokey (2007), has studied this question, focusing on the case in which the expert's preferences depend on the state. However, many experts have state-independent preferences: Salespeople want to sell products with higher commissions; politicians want to get elected; lawyers want favorable rulings; and so on. This paper analyzes the extent to which such experts benefit from cheap talk.

We consider a general cheap-talk model with one substantive assumption: The sender has state-independent preferences. Thus, we start with a receiver facing a decision problem with incomplete information. The relevant information is available to an informed sender who cares only about the receiver's action. Wanting to influence this action, the sender communicates with the receiver using costless messages.

Other papers have studied cheap-talk communication between a sender and a receiver when the former has state-independent preferences.[1] The most relevant is Chakraborty and Harbaugh (2010). Looking at a multidimensional specialization of our model, they

Elliot Lipnowski: e.lipnowski@columbia.edu
Doron Ravid: dravid@uchicago.edu

This research was supported by National Science Foundation Grant SES-1730168. We would like to thank Emir Kamenica for his invaluable comments and suggestions. We would also like to thank Ben Brooks, Kevin Bryan, Odilon Câmara, Archishman Chakraborty, David Dillenberger, Françoise Forges, Pedro Gardete, Faruk Gul, Johannes Hörner, Navin Kartik, R. Vijay Krishna, Laurent Mathevet, Stephen Morris, Wolfgang Pesendorfer, Andrew Postlewaite, Ariel Rubinstein, Maher Said, Denis Shishkin, and Vasiliki Skreta for insightful discussions.

[1] Schnakenberg (2015) characterized when an expert can convince voters to implement a proposal, and when said communication harms the voting population. Margaria and Smolin (2018) proved a folk theorem for a repeated interaction in which both a sender and a receiver are long-lived. With a long-lived sender but short-lived receivers, Best and Quigley (2020) showed that only partitional information can be credibly revealed, and that well-chosen mediation protocols can restore the commitment solution for a patient sender. Chung and Harbaugh (2019) tested experimentally the predictions of a recommendation game similar to our leading example.

 https://doi.org/10.3982/ECTA15674

showed the sender can always communicate some information credibly and influence the receiver's actions by trading off dimensions.[2] Chakraborty and Harbaugh (2010) also observed that a need for sender indifference creates a role for quasiconvexity and quasiconcavity. In particular, their second theorem says that, in their environment, the sender likes (dislikes) influencing the receiver in equilibrium whenever the sender's utility is a quasiconvex (quasiconcave) function of the receiver's action.

Our main insight is that a sender with state-independent preferences gains credibility by degrading self-serving information, that is, by making messages that serve as profitable deviations less informative. To derive this insight, we take a belief-based approach, as is common in the literature on communication.[3] Thus, we summarize communication via its induced *information policy*, a distribution over receiver posterior beliefs that averages to the prior. Say that a payoff s is *sender beneficial* if it is larger than the sender's no-information payoff, and *securable* if the sender's lowest ex post payoff from some information policy is at least s. Theorem 1 shows a sender-beneficial payoff s can be obtained in equilibrium if and only if s is securable. Thus, although the information policy securing s need not itself arise in equilibrium, its existence is sufficient for the sender to obtain a payoff of s in some equilibrium. Intuitively, the securing policy leads to posteriors that provide too much sender-beneficial information to the receiver. By degrading said information posterior by posterior, one can construct an equilibrium information policy attaining the secured value.

To illustrate our main result, consider a political think tank that advises a lawmaker. The lawmaker is contemplating whether to pass one of two possible reforms, denoted by 1 and 2, or to maintain the status quo, denoted by 0. Evaluating each proposal's merits requires expertise, which the think tank possesses. Given the think tank's political leanings, it is known to prefer certain proposals to others. In particular, suppose the status quo is the think tank's least preferred option and the second reform is the think tank's favorite option. Hence, let $a \in \{0, 1, 2\}$ represent both the lawmaker's choice and the think tank's payoff from that choice. To choose to implement a reform, the lawmaker must be sufficiently confident that the reform is good. Suppose one reform is good and one is bad, where the state, $\theta \in \{\theta_1, \theta_2\}$, indicates the identity of the good reform. The lawmaker implements reform a whenever he assigns θ_a a probability strictly above $\frac{3}{4}$. At $\frac{3}{4}$, the lawmaker is indifferent between said reform and the status quo, which the lawmaker chooses when neither reform is sufficiently likely to be good. Both reforms are equally likely to be good under the prior.

Suppose the think tank could reveal the state to the lawmaker; that is, the think tank recommends that the lawmaker *implement 1* when the state is θ_1 and *implement 2* when the state is θ_2. Because following these recommendations is incentive-compatible for the lawmaker, the think tank's ex post payoff would be 1 when sending *implement 1* and 2 when sending *implement 2*. By contrast, under no information, the think tank's payoff is 0. Thus, revealing the state secures the think tank a payoff of 1, which is higher than its payoff under the prior. Notice that 1 is then the highest payoff that the think tank can secure, because no information policy always increases the probability that the lawmaker assigns to θ_2. One can therefore apply Theorem 1 to learn two things: (i) 1 is an upper bound on the think tank's equilibrium payoffs, and (ii) we can achieve this bound via a

[2]See Battaglini (2002) and Chakraborty and Harbaugh (2007) for applications of this idea in the case of state-dependent sender preferences.

[3]For example, see Aumann and Maschler (1995), Aumann and Hart (2003), Kamenica and Gentzkow (2011), Alonso and Câmara (2016), and Ely (2017).

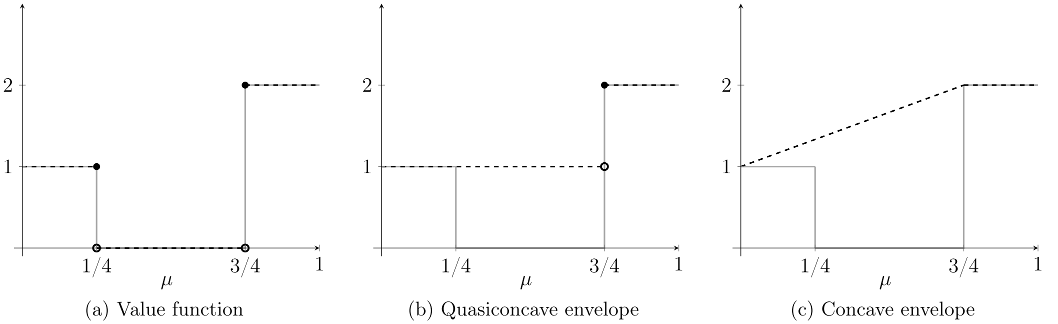
message-by-message garbling of said protocol. For (ii), consider what happens when the think tank sends the *implement 2* message according to

$$\mathbb{P}\{\text{implement } 2|\theta = \theta_1\} = \frac{1}{3},$$
$$\mathbb{P}\{\text{implement } 2|\theta = \theta_2\} = 1,$$

and sends *implement 1* with the complementary probabilities. As with perfect state revelation, choosing proposal 1 is the lawmaker's unique best response to *implement 1*. However, given *implement 2*, the lawmaker assigns a probability of $\frac{3}{4}$ to θ_2. Being indifferent, the lawmaker mixes between keeping the status quo and implementing 2 with equal probabilities. Such mixing results in indifference by the think tank, yielding an equilibrium.

In the general model, Theorem 1 allows us to geometrically characterize the sender's maximal benefit from cheap talk and compare this benefit with her benefit under commitment.[4] Kamenica and Gentzkow (2011) characterized the sender's benefit under commitment in terms of her *value function*, that is, the highest value the sender can obtain from the receiver's optimal behavior given his posterior beliefs. Specifically, they showed the sender's maximal commitment value is equal to the concave envelope of her value function. As we show in Theorem 2, replacing the concave envelope with the *quasi*concave envelope gives the sender's maximal value under cheap talk. Thus, the value of commitment is the difference between the concave and quasiconcave envelopes of the sender's value function.

Figure 1 visualizes the geometric comparison between cheap talk and commitment in the aforementioned think-tank example. Because the state is binary, the lawmaker's belief can be summarized by the probability it assigns to the second reform being good ($\theta = \theta_2$). Putting this probability on the horizontal axis, the figure plots the highest value the think tank can obtain from uninformative communication, cheap talk, and commitment. That is, the figure plots the think tank's value function (left), along with its quasiconcave (center) and concave (right) envelopes. The two envelopes describe how communication benefits the think tank by allowing it to connect points on the value function's graph. In contrast to communication with commitment, which enables the think tank to connect points



(a) Value function (b) Quasiconcave envelope (c) Concave envelope

FIGURE 1.—The simple think-tank example. The dashed lines represent the highest value the think tank can obtain from no information (left), cheap talk (center), and commitment (right).

[4]Our assumption of sender state-independent preferences is common in the literature on communication with hard evidence (e.g., Glazer and Rubinstein (2004, 2006), Hart, Kremer, and Perry (2017), Rappoport (2020)). Many such studies explore sufficient conditions for receiver- (rather than sender-) optimal equilibria to replicate receiver (rather than sender) commitment.

using any affine segment, only flat segments are allowed with cheap talk. The restriction to flat segments comes from the think tank's incentive constraints: Because the think tank's preferences are state independent, all equilibrium messages must yield the same payoff. As such, the think tank can only connect points with the same payoff coordinate; that is, only flat segments are feasible.

The geometric difference between cheap talk and commitment allows us to show that, in finite settings, almost all priors fall into one of two categories: Either the sender can get her first-best outcome with cheap talk, or she would strictly benefit from commitment. One can see this categorization holds in the simple think-tank example for almost all beliefs by using Figure 1. The figure clearly shows that unless the second reform is never good, the concave envelope lies above the quasiconcave envelope whenever the probability of the second reform being good is below $\frac{3}{4}$. Whenever the second reform is good with probability $\frac{3}{4}$ or above, the lawmaker is willing to implement the think tank's favorite reform under the prior, and so the two envelopes must coincide with the value function.

In Section 5, we use our results in three specific economic settings. In a richer version of the above think-tank example, we show a think tank's best equilibrium involves giving the lawmaker noisy recommendations, where the noise is calibrated to make the lawmaker indifferent between the recommended reform and the status quo. We also study a broker-investor relationship, in which an investor consults his broker about an asset, and the broker earns a fee proportional to the investor's trades. We identify a Pareto-dominant equilibrium in which the broker tells the investor whether his holdings should be above or below a fee-independent cutoff amount. Thus, the lower the broker's fee, the better off the investor, who pays less money for the same information. Lower fees have an ambiguous effect on the broker because they reduce her income per trade but increase equilibrium trade volume. We also conduct comparative statics in market volatility. Although higher volatility cannot hurt the broker, she strictly benefits from higher volatility only if she can effectively communicate about it to the investor. The investor's attitude toward higher volatility is ambiguous because it changes both the investor's prior uncertainty and the usefulness of the broker's information. Our third example is a symmetric version of the multiple-goods seller example of Chakraborty and Harbaugh (2010). Specifically, we consider a seller who wants to maximize the probability of selling one of her many products to a buyer. In this setting, we show the best the seller can do with cheap talk is tell the buyer the identity of her best product. Moreover, we show being able to benefit ex ante from providing the buyer with additional information about the best product is a necessary and sufficient condition for the seller to benefit from commitment.

In Section 6.1, we revisit Chakraborty and Harbaugh (2010). We point out that, absent their specific parametric structure, Chakraborty and Harbaugh's (2010) reasoning shows the sender can influence the receiver's estimate of any multidimensional statistic of the state. Whenever this estimate coincides with the receiver's best response, the sender can also influence the receiver's actions. Otherwise, Chakraborty and Harbaugh's (2010) reasoning delivers informative communication, which might not influence the receiver's actions, as long as three or more states exist.

To summarize, we contribute to the literature on cheap talk with state-independent sender preferences in three ways. First, we identify the ability to reduce the informativeness of profitable messages as a key channel through which the sender gains credibility. Using this channel, we obtain a complete characterization of the sender's payoff set. Second, we show quasiconcavity fully summarizes the sender's ability to benefit from communication. Third, we apply our results to generate new insights in economic applications.

2. CHEAP TALK WITH STATE-INDEPENDENT PREFERENCES

Our model is an abstract cheap-talk model with the substantive restriction that the sender has state-independent preferences. Thus, we have two players: a sender (S, she) and a receiver (R, he). The game begins with the realization of a random 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. Whereas R's payoffs depend on θ, S's payoffs do not.

We impose some technical restrictions on our model.[5] Each of Θ, A, and M is a compact metrizable space containing at least two elements, and M is sufficiently rich.[6] The state, θ, follows some full-support distribution $\mu_0 \in \Delta\Theta$, which is known to both players. Both players' utility functions are continuous, where we take $u_S : A \to \mathbb{R}$ to be S's utility and $u_R : A \times \Theta \to \mathbb{R}$ to be R's.

We are interested in studying the game's equilibria, by which we mean perfect Bayesian equilibria. 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 σ, using Bayes's rule;[7]
2. $\rho(m)$ is supported on $\arg\max_{a\in A} \int_\Theta u_R(a, \cdot)\,\mathrm{d}\beta(\cdot|m)$ for all $m \in M$; and
3. $\sigma(\theta)$ is supported on $\arg\max_{m\in M} \int_A u_S(\cdot)\,\mathrm{d}\rho(\cdot|m)$ for all $\theta \in \Theta$.

Any triple $\mathcal{E} = (\sigma, \rho, \beta)$ induces a joint distribution, $\mathbb{P}_\mathcal{E}$, over realized states, messages, and actions,[8] which, in turn, induces (through β and ρ, respectively) distributions over R's equilibrium beliefs and chosen mixed action.

The following are a few concrete examples of our setting.

EXAMPLE 1: Consider the following richer version of the think-tank example from the Introduction. Thus, S is a think tank that is advising a lawmaker (R) on whether to pass one of $n \in \mathbb{N}$ reforms or to pass none; that is, the lawmaker chooses from $A = \{0, 1, \dots, n\}$. A given reform $i \in \{1, \dots, n\}$ provides uncertain benefit $\theta_i \in [0, 1]$ to the lawmaker. From the lawmaker's perspective, reforms are ex ante identical: Their benefits are distributed according to an exchangeable prior μ_0 over $[0, 1]^n$, and each entails an implementation cost of c. Maintaining the status quo is costless but generates no benefits, $u_R(0, \theta) = 0$. The think tank prefers higher-indexed reforms to lower-indexed ones, and prefers some reform to no reform; that is, the think tank's payoffs are given by a strictly increasing function, $u_S : A \to \mathbb{R}$, where we normalize $u_S(0) = 0$.[9] We analyze this example in Section 5.1.

[5]Let us describe some notational conventions we adopt throughout the paper. For a compact metrizable space Y, we let ΔY denote the set of all Borel probability measures over Y, endowed with the weak* topology. Given $y \in Y$, we let $\delta_y \in \Delta Y$ denote a unit atom on y, $\delta_y\{y\} = 1$. For $\gamma \in \Delta Y$, we let $\operatorname{supp}\gamma$ denote the support of γ. For a set X, a transition $g : X \to \Delta Y$, a point $\bar{x} \in X$, and a Borel subset $\hat{Y} \subseteq Y$, we let $g(\hat{Y}|\bar{x}) := g(\bar{x})(\hat{Y})$. For a set Z, a function $h : X \to Z$, and a subset $\hat{X} \subseteq X$, we let $h(\hat{X}) := \{h(x) : x \in \hat{X}\}$. Finally, "co" refers to the convex hull, and "$\overline{\text{co}}$" refers to the closed convex hull.

[6]To simplify the statements of our results, we assume $M \supseteq A \cup \Delta A \cup \Delta\Theta$. S's attainable payoffs would be the same if we instead imposed either that $|M| \geq |A|$ or that Θ is finite and $|M| \geq |\Theta|$, by Proposition 2, Corollary 1, and Carathéodory's theorem.

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

[8]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 $\hat{\Theta} \subseteq \Theta$, $\hat{M} \subseteq M$, $\hat{A} \subseteq A$.

[9]This example is related to, but formally distinct from, the respective models of Che, Dessein, and Kartik (2013) and Chung and Harbaugh (2019). The former studies a project-selection model with state-dependent preferences for both players, and the latter tests experimentally a binary-state project-selection model with a stochastic receiver outside option.

EXAMPLE 2: R is an investor consulting a broker (S) about an asset. The broker knows the investor's ideal position in the asset, $\theta \in \Theta = [0, 1]$, which is distributed according to the atomless prior, μ_0. The investor's pre-existing position is $a_0 \in [0, 1]$. After consulting his broker, the investor chooses a new position in the asset, $a \in A = [0, 1]$. The broker's payoff accrues from brokerage fees proportional to the net volume of trade; that is, $u_S(a) = \phi|a - a_0|$ for some $\phi > 0$. The investor wants to match the ideal holdings level, but must pay the broker's fees: $u_R(a, \theta) = -\frac{1}{2}(a - \theta)^2 - u_S(a)$. In Section 5.2, we find a Pareto-dominant equilibrium and conduct comparative statics under the assumption that the investor's existing position is correct; that is, $a_0 = \int_\Theta \theta \, \mathrm{d}\mu_0(\theta)$.

EXAMPLE 3: A buyer (R) can take an outside option or buy one of N goods from a seller (S). The seller knows the vector $\theta = (\theta_1, \dots, \theta_n)$, where θ_i denotes the buyer's net value from product i. Product values are i.i.d. atomlessly distributed over $[0, 1]$. The seller wants to maximize the probability of a sale, but does not care which product is sold. Hence, the seller receives a value of 1 if the buyer chooses to purchase product $i \in \{1, \dots, n\}$, and 0 if the buyer chooses the outside option, which we denote by 0. Only the buyer knows her value from the outside option, ϵ, which is distributed independently from θ according to G, a continuous, full-support CDF over $[0, 1]$. Chakraborty and Harbaugh (2010) studied this example and showed the seller can always benefit from communication. In Section 5.3, we use our tools to expand on their analysis.

We analyze our model via the belief-based approach, commonly used in the communication 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's rule, this posterior distribution averages to the prior, μ_0. That is, $p \in \Delta\Delta\Theta$ satisfies $\int \mu \, \mathrm{d}p(\mu) = \mu_0$. 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.[10] Let $\mathcal{I}(\mu_0)$ denote the set of all information policies.

The belief-based approach allows us to focus on the game's outcomes. Formally, an *outcome* is a pair, $(p, s) \in \Delta\Delta\Theta \times \mathbb{R}$, representing R's posterior distribution, p, and S's ex ante payoff, s. An outcome is an *equilibrium outcome* if it corresponds to an equilibrium.[11] An equilibrium outcome is *informative* if R's posterior distribution is non-degenerate, $p \neq \delta_{\mu_0}$. In contrast to equilibrium, a triple (σ, ρ, β) is a *commitment protocol* if it satisfies the first two of the three equilibrium conditions above; and (p, s) is a *commitment outcome* if it corresponds to some commitment protocol. In other words, commitment outcomes do not require S's behavior to be incentive-compatible.

Using the belief-based approach, Aumann and Hart (2003) analyzed, among other things, the outcomes of the cheap-talk model with general S preferences over states and actions. When S's preferences are state independent, their characterization essentially specializes to Lemma 1 below,[12] which describes the game's equilibrium outcomes. To

[10]For example, see Aumann and Maschler (1995), Benoît and Dubra (2011), or Kamenica and Gentzkow (2011).

[11]That is, an equilibrium $\mathcal{E} = (\sigma, \rho, \beta)$ exists such that $p(\hat{B}) = \mathrm{marg}_M \mathbb{P}_\mathcal{E}[\beta^{-1}(\hat{B})]$ for every Borel $\hat{B} \subseteq \Delta\Theta$, and $s = \int_A u_S \, \mathrm{d}\mathrm{marg}_A \mathbb{P}_\mathcal{E}$.

[12]Because Aumann and Hart's (2003) setting is finite, we provide a direct independent proof of said lemma for the sake of completeness.

state the lemma, let $V(\mu)$ be S's possible continuation values from R having μ as his posterior,

$$V : \Delta\Theta \rightrightarrows \mathbb{R}$$

$$\mu \mapsto \operatorname{co} u_S\left(\arg\max_{a\in A} \int u_R(a,\cdot)\,\mathrm{d}\mu\right).$$

By Berge's theorem, V is a Kakutani correspondence, and the *value function*, $v(\cdot) := \max V(\cdot)$, is upper semicontinuous.[13]

LEMMA 1: *The outcome (p, s) is an equilibrium outcome if and only if*:
1. $p \in \mathcal{I}(\mu_0)$, *that is*, $\int \mu\,\mathrm{d}p(\mu) = \mu_0$, *and*
2. $s \in \bigcap_{\mu\in\operatorname{supp}(p)} V(\mu)$.

The lemma's conditions reflect the requirements of perfect Bayesian equilibrium. The first condition comes from the equivalence between Bayesian updating and p being an information policy. The second condition combines both players' incentive-compatibility constraints. For S, incentive compatibility requires her continuation value to be the same from all posteriors in p's support, meaning her ex ante value must be equal to her continuation value upon sending a message. For R, incentive compatibility requires that $V(\mu)$ contain S's continuation value from any message that leaves R at posterior belief μ. Therefore, S's ex ante value must be in $V(\mu)$ for all posteriors μ in p's support.

Our setting nests the model of Chakraborty and Harbaugh (2010). In their model, $\Theta = A \subseteq \mathbb{R}^N$ is a compact convex set with a nonempty interior, where $N > 1$, the prior admits a full-support density, and $\arg\max_{a\in A} \int u_R(a,\cdot)\,\mathrm{d}\mu = \{\int \theta\,\mathrm{d}\mu(\theta)\}$ for every $\mu \in \Delta\Theta$. Chakraborty and Harbaugh's (2010) main result is that this setting always admits an equilibrium in which S's messages influence R's actions. Using Lemma 1, one can generalize Chakraborty and Harbaugh's (2010) logic to show S can typically communicate information to R; that is, most versions of our model admit an informative equilibrium. Because our analysis does not rely on the existence of an informative equilibrium, we defer discussion of this result to Section 6.1.

Another insight of Chakraborty and Harbaugh (2010) is that the reliance of equilibrium communication on S indifference creates a role for quasiconcavity and quasiconvexity. In particular, they observed that a finite-support distribution can give a quasiconcave (quasiconvex) function a constant value only if said value is lower (higher) than the function's value at the distribution's mean. This observation has many useful implications. One implication is that in Chakraborty and Harbaugh's (2010) setting, S always benefits from influencing R's action in equilibrium when u_S is strictly quasiconvex. Another implication is that babbling is S's best (worst) equilibrium whenever v is quasiconcave (quasiconvex) and R's best response is unique for all beliefs.

In what follows, we show quasiconcavity completely summarizes S's ability to benefit from communication. More precisely, we prove S's maximal equilibrium payoff is given by the quasiconcave envelope of v (Theorem 2). This result is based on our main result (Theorem 1), presented in the next section.

[13]That is, V is a nonempty-, compact-, and convex-valued, upper hemicontinuous correspondence, and $v(\mu) := \max_{s\in V(\mu)} s$ is upper semicontinuous in $\mu \in \Delta\Theta$. Notice v is well defined (i.e., $\max V(\mu)$ exists) because $V(\mu)$ is nonempty and compact.

3. SECURABILITY

This section presents our main result, Theorem 1, which characterizes S's equilibrium payoffs. The characterization shows that as far as S's payoffs are concerned, one can ignore S's incentive constraints by focusing on S's least favorite message in any given information policy. Thus, using the theorem, one can use non-equilibrium information policies to reason about S's possible equilibrium payoffs.

Let p be an information policy, and take s to be some possible S payoff. Say that policy p *secures* s if $p\{v \geq s\} = 1$,[14] and that s is *securable* if an information policy exists that secures s, that is, if $\mu_0 \in \overline{\text{co}}\{v \geq s\}$. Our main result shows securability characterizes S's equilibrium values.

THEOREM 1—Securability: *Suppose* $s \geq v(\mu_0)$.[15] *Then, an equilibrium inducing sender payoff s exists if and only if s is securable.*

The key observation behind Theorem 1 is that one can transform any policy p that secures s into an equilibrium policy by degrading information. Specifically, we replace every supported posterior μ with a different posterior μ' that lies on the line segment between μ_0 and μ. Because μ' is between μ_0 and μ, replacing μ with μ' results in a weakly less informative signal. To ensure the resulting signal is an equilibrium, we take μ' to be the closest posterior to μ_0 among the posteriors between μ_0 and μ that make providing s incentive-compatible for R. Thus, this transformation replaces a potentially incentive-*incompatible* posterior μ with the incentive-*compatible* μ'. That μ' exists follows from two facts. First, s is between S's no-information value and her highest μ payoff, $v(\mu)$. Second, V is a Kakutani correspondence, admitting an intermediate value theorem.[16]

The above logic also identifies a class of equilibrium information policies that span all of S's equilibrium payoffs above $v(\mu_0)$. Say that *p barely secures s* if $\{v \geq s\} \cap \text{co}\{\mu, \mu_0\} = \{\mu\}$ holds for p-a.e. μ. In words, barely securing policies are policies that secure a payoff higher than what S can attain at any belief between any supported posterior and the prior. The construction behind Theorem 1 transforms every securing policy into a barely securing policy that is also an equilibrium. Because all equilibrium values are securable, we thus have that any high equilibrium value can be attained in an equilibrium with a barely securing policy. Moreover, because barely securing policies are left untouched by Theorem 1's transformation, every barely securing policy must then be an equilibrium.

Theorem 1 highlights the way incentives constrain S's ability to extract value from her information. Although S can always degrade self-serving information to guarantee incentives, the same cannot be done to information that is self-harming.[17] As such, S's highest value is determined by the best worst message she must send if she could commit. It follows S can do no better than no information if and only if she cannot avoid sending R

[14]Here, we use the standard notation: $\{v \geq s\} = \{\mu : v(\mu) \geq s\}$.

[15]Given our focus on S's benefits from cheap talk, we state the theorem for high S values. For $s \leq \min V(\mu_0)$, one replaces the requirement that s is securable with the existence of some $p \in \mathcal{I}(\mu_0)$ such that $p\{\min V \leq s\} = 1$.

[16]Theorem 1's proof is related (in that both use the intermediate value theorem to construct an equilibrium) to the proof of Chakraborty and Harbaugh's (2010) Theorem 4. Their theorem says that in Chakraborty and Harbaugh's (2010) specialization of our model, if u_S is strictly quasiconvex, a sequence of equilibria $\{\mathcal{E}_k\}_{k=1}^{\infty}$ exist such that $\mathcal{E}_k$ entails 2^k on-path messages, and S's value from $\mathcal{E}_k$ strictly increases in k. We thank an anonymous referee for making us aware of the relationship between these two results.

[17]The statement is true for S-beneficial payoffs. For S-harmful payoffs, the sender would degrade excessively self-harming information to guarantee incentives.

messages that are worse than providing no information. That is, the set of beliefs at which S attains a value strictly higher than no information does not contain the prior in its closed convex hull.[18]

Theorem 1 also yields a convenient formula for S's maximal equilibrium value, which we present in Corollary 1 below.

COROLLARY 1: *An S-preferred equilibrium exists, giving the sender a payoff of* $v^*(\mu_0)$, *where*

$$v^*(\cdot) := \max_{p \in \mathcal{I}(\cdot)} \inf v(\operatorname{supp} p).$$

Notice that $\inf v(\operatorname{supp} p)$ is the highest value that p secures. Thus, Corollary 1 says that maximizing S's equilibrium value is equivalent to maximizing the highest value S can secure across all information policies. In the next section, we provide a geometric characterization of v^*.

4. COMMITMENT'S VALUE IN COMMUNICATION

The current section uses Theorem 1 to examine the value of commitment in strategic communication. The main result of this section is Theorem 2, which geometrically characterizes S's maximal equilibrium value. 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 that majorizes v.[19] Because concavity implies quasiconcavity, the quasiconcave envelope lies (weakly) below the concave envelope. Figure 2 illustrates the definitions of the concave and quasiconcave envelopes for an abstract function.

As described in Aumann and Maschler (1995)[20] and Kamenica and Gentzkow (2011), $\hat{v}$ gives S's payoff from her favorite commitment outcome. Theorem 2 below shows $\bar{v}$ gives S's maximal value under cheap talk.

THEOREM 2—Quasiconcavification: *S's maximal equilibrium value is given by v's quasiconcave envelope*; *that is*,

$$v^* = \bar{v}.$$

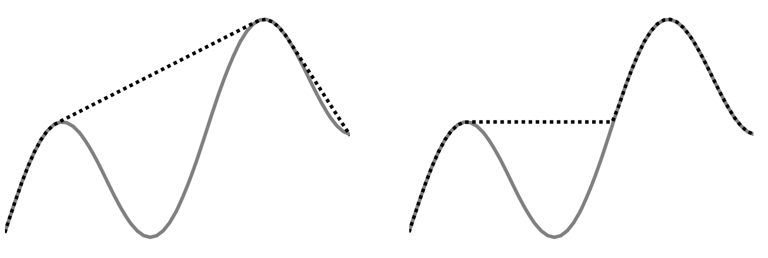


FIGURE 2.—A function with its concave (left) and quasiconcave (right) envelopes.

[18]Formally, $s \leq v(\mu_0)$ for all equilibrium outcomes if and only if for every $\epsilon > 0$, $\mu_0 \notin \overline{\text{co}}\{v \geq v(\mu_0) + \epsilon\}$. Schnakenberg (2015) showed a similar condition characterizes an expert's ability to sway voters to support her favorite of two policies.

[19]The Appendix contains a proof that, for finite Θ, the quasiconcave envelope is below *every* quasiconcave function that majorizes v, even those that are not upper semicontinuous.

[20]Also see Aumann and Maschler (1966).

PROOF OF THEOREM 2: We begin by showing v^* is a quasiconcave, upper semicontinuous function that majorizes v. That v^* majorizes v follows from existence of an uninformative equilibrium. For upper semicontinuity, we refer the reader to Lemma 5, which we prove in the Appendix.

We now argue v^* is quasiconcave. For this purpose, fix μ', μ'', and $\lambda \in (0,1)$, and consider the following observations. First, if $p' \in \mathcal{I}(\mu')$, and $p'' \in \mathcal{I}(\mu'')$, then $\lambda p' + (1-\lambda)p'' \in \mathcal{I}(\lambda\mu' + (1-\lambda)\mu'')$. Second, the support of the convex combination of two distributions is the union of their supports. Taken together, these observations imply the following inequality chain:

$$\begin{aligned} v^*\big(\lambda\mu' + (1-\lambda)\mu''\big) &= \max_{p\in\mathcal{I}(\lambda\mu+(1-\lambda)\mu')} \inf v(\operatorname{supp} p) \\ &\geq \max_{p'\in\mathcal{I}(\mu'),\, p''\in\mathcal{I}(\mu'')} \inf v\big(\operatorname{supp} p' \cup \operatorname{supp} p''\big) \\ &= \max_{p'\in\mathcal{I}(\mu'),\, p''\in\mathcal{I}(\mu'')} \min\big\{\inf v\big(\operatorname{supp} p'\big), \inf v\big(\operatorname{supp} p''\big)\big\} \\ &= \min\big\{v^*\big(\mu'\big), v^*\big(\mu''\big)\big\}, \end{aligned}$$

where the last equality follows from reasoning separately for p' and p''.

To show $v^* = \bar{v}$, it remains to show that v^* lies below any upper semicontinuous and quasiconcave $f : \Delta\Theta \to \mathbb{R}$ that majorizes v. Fixing some prior $\mu \in \Delta\Theta$, take $p \in \mathcal{I}(\mu)$ to be an information policy securing S's favorite equilibrium value, $v^*(\mu)$. By choice of p, we have that, for $D := \operatorname{supp} p$, both $\inf v(D) = v^*(\mu)$ and $\mu \in \overline{\operatorname{co}}D$. Combined with f being upper semicontinuous, quasiconcave, and above v, we have

$$f(\mu) \geq \inf f(\overline{\operatorname{co}}D) = \inf f(\operatorname{co} D) = \inf f(D) \geq \inf v(D) = v^*(\mu).$$

Because μ and f were arbitrary, our proof is complete. *Q.E.D.*

Theorem 2 provides a geometric comparison between communication's value under cheap talk and under commitment. With commitment, communication is only restricted by R's incentives and Bayes's rule. The value function's concave envelope describes the maximal payoff S can attain in this manner. Replacing the value function's concave envelope with its quasiconcave envelope expresses the value S loses in cheap talk due to her incentive constraints. Graphically, both envelopes allow S to extract value from connecting points on the graph of S's value correspondence. However, although with commitment S can connect points via any affine segment, cheap talk restricts her to flat ones. One can see the associated value loss for the Introduction's example in Figure 1: For priors $\mu \in (0, \frac{3}{4})$, S's highest cheap-talk value is 1, whereas with commitment, her highest payoff is given by $1 + \frac{4}{3}\mu$.

Corollary 2 below uses the geometric difference between cheap talk and commitment to show that in a finite setting, commitment is valuable for most priors. In particular, with finite actions and states, the following is true for all priors lying outside a measure-zero set: Either S attains her first-best feasible payoff, or S strictly benefits from commitment.

COROLLARY 2: *Suppose A and Θ are finite. Then, for Lebesgue-almost all $\mu_0 \in \Delta\Theta$, either $\bar{v}(\mu_0) = \max v(\Delta\Theta)$ or $\bar{v}(\mu_0) < \hat{v}(\mu_0)$.*

The intuition for the corollary is geometric: Except at S's first-best feasible payoff, the concave envelope, $\hat{v}$, must lie above the interior of any of the quasiconcave envelope's flat surfaces. To see why, notice any prior μ_0 in the interior of such a surface can be expressed as a convex combination of another belief on the same surface and a belief yielding S's first-best feasible value. Said formally, some $\lambda \in (0,1)$, μ, and μ' exist such that $\bar{v}(\mu) = \bar{v}(\mu_0)$, $\bar{v}(\mu') = \max v(\Delta\Theta)$, and $\mu_0 = \lambda\mu + (1-\lambda)\mu'$. Because $\bar{v}$ lies below $\hat{v}$, and because $\hat{v}$ is concave, we obtain

$$\bar{v}(\mu_0) < \lambda\bar{v}(\mu) + (1-\lambda)\bar{v}(\mu') \leq \lambda\hat{v}(\mu) + (1-\lambda)\hat{v}(\mu') \leq \hat{v}(\mu_0),$$

as required.

5. APPLICATIONS

5.1. *The Think Tank*

This section uses our results to analyze Example 1. We characterize the think tank's maximal equilibrium value and find an equilibrium in a barely securing policy that attains it. To ease notation, we assume in the main text that the probability that two reforms yield the same benefit to the lawmaker is zero.

In the single-reform case, neither player can do better than no information: In this case, think-tank indifference occurs only if the lawmaker's mixed action is constant on path. With multiple reforms, one can analyze the example using the claim below, made possible by Theorem 1.

CLAIM 1: *The following are equivalent, given* $k \in \{1, \dots, n\}$:
1. *The think tank can attain the value* $u_S(k)$ *in equilibrium.*
2. $\mathbb{E}_{\theta\sim\mu_0}[\max_{i\in\{k,\dots,n\}} \theta_i] \geq c$.
3. *The policy,* $p_k \in \mathcal{I}(\mu_0)$, *that reveals the random variable*

$$\mathbf{i}_k := \arg\max_{i\in\{k,\dots,n\}} \theta_i$$

to the lawmaker secures $u_S(k)$.

The claim says the think tank attaining a value of $u_S(k)$ in equilibrium is equivalent to two other conditions. First, always choosing the status quo is ex ante worse for the lawmaker than always choosing the best reform from $\{k, \dots, n\}$ (Part 2). Second, telling the lawmaker nothing but the identity of the best reform from $\{k, \dots, n\}$ secures $u_S(k)$ (Part 3).

Claim 1's Part 2 provides a simple necessary and sufficient condition for $u_S(k)$ to be an equilibrium value. Using this condition, we can find S's maximal value across all equilibria: it is given by $u_S(k^*)$, where

$$k^* := \max\left\{k \in \{1, \dots, n\} : \mathbb{E}_{\theta\sim\mu_0}\left[\max_{i\in\{k,\dots,n\}} \theta_i\right] \geq c\right\}.$$

That is, k^* is the highest k for which Part 2 holds. With k^* in hand, we can identify a best equilibrium for the think tank using the claim's Part 3. This part tells us the think tank's favorite equilibrium value, $u_S(k^*)$, is securable by the information policy, p_{k^*}, that reveals to the lawmaker the identity of the best reform from the set $\{k^*, \dots, n\}$. Thus, to find an equilibrium, we can take p_{k^*} and garble information message by message to obtain a new policy that barely secures $u_S(k^*)$. Doing so results in a policy that has the think tank randomizing between accurately recommending the lawmaker's best reform from $\{k^*, \dots, n\}$ with probability $1-\epsilon$, and recommending a uniformly drawn reform from $\{k^*, \dots, n\}$ with probability ϵ. By choosing ϵ appropriately, one can degrade information so as to make the lawmaker indifferent between the suggested recommendation and the status quo. The result is an equilibrium in which the lawmaker implements the suggested reform i with probability $\frac{u_S(k^*)}{u_S(i)}$ and maintains the status quo with complementary probability. Thus, all that remains is to calculate k^* and ϵ, which depend on the prior. For example, if $\theta_1, \dots, \theta_n$ are i.i.d. uniformly distributed on $[0, 1]$ and $c > \frac{1}{2}$,[21] then

$$k^* = \left\lfloor n - \frac{2c-1}{1-c} \right\rfloor \quad \text{and} \quad \epsilon = 2\left(1 - c - \frac{2c-1}{n-k^*}\right).$$

The policy p_{k^*} also yields an easy lower bound on commitment's value. Specifically, the value of commitment is at least the difference between k^* and the think tank's value function's expectation under p_{k^*},

$$\int v(\cdot)\,\mathrm{d}p_{k^*} - u_S(k^*) = \frac{1}{n-k^*+1}\sum_{i=k^*}^{n} u_S(i) - u_S(k^*),$$

which simplifies to $\frac{1}{2}(n-k^*)$ in the special case of $u_S(a) = a$.

5.2. *The Broker*

We now revisit the setting of Example 2 under the assumption that the investor's initial holdings are correct given her information, that is, that $a_0 = \int \theta\,\mathrm{d}\mu_0(\theta)$. Even without this assumption, characterizing optimal behavior by the investor is straightforward. For any posterior belief $\mu \in \Delta\Theta$, simple calculus yields that the investor's best response is unique and given by

$$a^*(\mu) = \begin{cases} \int \theta\,\mathrm{d}\mu(\theta) + \phi & : \int \theta\,\mathrm{d}\mu(\theta) - a_0 \le -\phi, \\ a_0 & : \int \theta\,\mathrm{d}\mu(\theta) - a_0 \in [-\phi, \phi], \\ \int \theta\,\mathrm{d}\mu(\theta) - \phi & : \int \theta\,\mathrm{d}\mu(\theta) - a_0 \ge \phi. \end{cases}$$

As such, V is a single-valued correspondence, with $v(\mu) = \phi[|\int \theta\,\mathrm{d}\mu(\theta) - a_0| - \phi]_+$.[22]

The above expression demonstrates that this example is a specific instance of a class of models in which $\Theta \subseteq \mathbb{R}$ and S's value function is a quasiconvex function of R's expectation of the state. The special one-dimensional structure of this class allows us to focus on *cutoff policies*. Formally, p is a *θ^*-cutoff policy* if it reports whether the state is above or below $\theta^* \in \Theta$.[23] The following proposition shows garblings of cutoff policies are sufficient to attain any S equilibrium value in one-dimensional settings.

[21] When $c \le \frac{1}{2}$, the think tank can obtain its first-best outcome under no information; that is, $v(\mu_0) = u_S(n)$.
[22] We let $[\cdot]_+ := \max\{\cdot, 0\}$.
[23] In Section B.2 in the Appendix, we provide a definition of cutoff policies (and prove a version of Claim 2) that applies for general priors. The two definitions coincide when the prior is atomless.

CLAIM 2: *Suppose* $\Theta \subseteq \mathbb{R}$, μ_0 *is atomless, and that* $v(\mu) = v_M(\int \theta \, \mathrm{d}\mu(\theta))$, *where* $v_M : \operatorname{co}\Theta \to \mathbb{R}$ *is weakly quasiconvex. Then, the following are equivalent for all* $s \geq v(\mu_0)$:

1. *S can attain payoff s in equilibrium.*
2. *The payoff s is securable by a cutoff policy.*

Moreover, an S-preferred equilibrium outcome (p, s) *exists such that* p *is a cutoff policy.*

We now apply the claim to our specific broker example. Notice the broker's value function is given by $v(\mu) = v_M(\int \theta \, \mathrm{d}\mu(\theta))$, where $v_M(\theta) = \phi[|\theta - a_0| - \phi]_+$. Because v_M is a convex function, Claim 2 implies an S-preferred equilibrium exists in which S uses a cutoff policy. Consider the median-cutoff policy, where the broker tells the investor whether the state is above or below the median. Let $\theta_<$ and $\theta_>$ denote the investor's expectation of the state conditional on it being below or above the median, respectively. Because $a_0 = \int \theta \, \mathrm{d}\mu_0(\theta) = \frac{1}{2}\theta_< + \frac{1}{2}\theta_>$, one has $|\theta_> - a_0| = |\theta_< - a_0|$, meaning $v_M(\theta_<) = v_M(\theta_>)$. Thus, the median cutoff policy is an equilibrium policy. Moreover, v_M decreases on $[\theta_<, a_0]$ and increases on $[a_0, \theta_>]$, and so no alternative cutoff policy can secure a higher value. Hence, Claim 2 tells us the median cutoff policy yields a broker-preferred equilibrium. We can therefore calculate the broker's maximal equilibrium payoff,

$$\bar{v}(\mu_0) = \phi\left[\frac{1}{2}(\theta_> - \theta_<) - \phi\right]_+. \tag{1}$$

In the median-cutoff equilibrium, the transmitted information does not depend on ϕ. This observation simplifies the task of conducting comparative statics in ϕ: The broker's maximal equilibrium payoff is single-peaked in ϕ, with the optimal ϕ being $\frac{1}{4}(\theta_> - \theta_<)$. Intuitively, increasing ϕ reduces trade but increases the broker's income per trade, with the latter effect dominating for low ϕ and the former dominating for high ϕ.

It is easy to see the broker's maximal equilibrium payoff increases with mean-preserving spreads of μ_0; that is, the more volatile the market is, the better off the broker. However, not all volatility is equal: Mean-preserving spreads strictly increase the broker's payoff if and only if they increase $\theta_> - \theta_<$. Thus, for the broker to strictly benefit from market volatility, she must be able to communicate about it to the investor.

How does the investor fare in the broker's preferred equilibrium? Simple algebra reveals the investor's payoff is $\frac{1}{2\phi^2}s^2 - \operatorname{Var}_{\theta\sim\mu_0}(\theta)$ in any equilibrium yielding the broker a payoff of s.[24] Two consequences are immediate. First, the investor's equilibrium payoffs increase with the broker's, meaning the broker's favorite equilibrium is Pareto-dominant. Second, the investor's payoffs in the Pareto-dominant equilibrium are given by

$$\frac{1}{2}\left\{\left[\frac{1}{2}(\theta_> - \theta_<) - \phi\right]_+\right\}^2 - \operatorname{Var}_{\theta\sim\mu_0}(\theta).$$

Notice the investor is always better off with lower brokerage fees: Because the broker's information does not change with ϕ, a lower ϕ means the investor pays less for the same information. By contrast, the investor's attitude toward higher prior volatility (in the sense of mean-preserving spreads) is ambiguous. Intuitively, increased market volatility both increases the investor's risk and increases the usefulness of the broker's recommendations. As such, higher volatility that does not change the broker's recommendations unambiguously hurts the investor.

[24] The reader can find said algebra in Appendix B.2.

5.3. *The Salesperson*

In this section, we return to Example 3. This example was first analyzed by Chakraborty and Harbaugh (2010),[25] who showed it always admits an influential equilibrium, that is, an equilibrium in which different messages lead to different action distributions by the buyer. Chakraborty and Harbaugh (2010) also noticed that every influential equilibrium in this setting benefits the seller due to quasiconvexity. In this section, we find a seller-preferred equilibrium and obtain a full characterization of when the seller benefits from commitment.

Because the buyer has private information, this example does not formally fall within our model. Our analysis, however, still applies.[26] Given a belief $\mu \in \Delta\Theta$, the buyer purchases the good with probability $\mathbb{P}\{\epsilon \leq \max_i \int \theta_i \,\mathrm{d}\mu(\theta)\} = G(\max_i \int \theta_i \,\mathrm{d}\mu(\theta))$. Hence, the seller's continuation value from sending a message that gives the buyer a posterior of μ is $v(\mu) := G(\max_i \int \theta_i \,\mathrm{d}\mu(\theta))$. Using the continuous function v as the seller's value function, we can directly apply our results to this example.

Applying Theorem 2 yields an upper bound on the seller's equilibrium values. To obtain this bound, define the continuous function $\bar{v}^*(\mu) := G(\int \max_{j\in\{1,\dots,n\}} \theta_j \,\mathrm{d}\mu(\theta))$. Being an increasing transform of an affine function, $\bar{v}^*$ is quasiconcave.[27] Moreover, because G is increasing, Jensen's inequality tells us

$$\bar{v}^*(\mu) = G\left(\int \max_{j\in\{1,\dots,n\}} \theta_j \,\mathrm{d}\mu(\theta)\right) \geq G\left(\max_{j\in\{1,\dots,n\}} \int \theta_j \,\mathrm{d}\mu(\theta)\right) = v(\mu).$$

In other words, $\bar{v}^*$ is a continuous quasiconcave function that majorizes the seller's value function, and so lies above the value function's quasiconcave envelope. Theorem 2 then implies $\bar{v}^*(\mu_0)$ is above any equilibrium seller value.

We now describe an equilibrium that attains the upper bound $\bar{v}^*(\mu_0)$. Let p^* be the information policy in which the seller tells the buyer the identity of the most valuable product.[28] Assuming the buyer believes the seller, the seller's expected value from recommending product i is

$$G\left(\mathbb{E}_{\theta\sim\mu_0}\left[\theta_i \,\middle|\, i \in \arg\max_{j\in\{1,\dots,n\}} \theta_j\right]\right) = G\left(\int \max_{j\in\{1,\dots,n\}} \theta_j \,\mathrm{d}\mu_0(\theta)\right) = \bar{v}^*(\mu_0),$$

where the first equality follows from product values being i.i.d. Notice all recommendations yield the seller the same value, meaning p^* is an equilibrium. Moreover, p^* attains the upper bound $\bar{v}^*(\mu_0)$ on the seller's equilibrium values. In other words, $(p^*, \bar{v}^*(\mu_0))$ is a seller-preferred equilibrium outcome.

The identified equilibrium is, in fact, Pareto dominant. To see why, notice that if the seller's equilibrium payoff is s, the buyer's expected utility from the best product is $G^{-1}(s)$ for any on-path message. Therefore, the buyer's utility in equilibrium is

[25] Chakraborty and Harbaugh (2014) studied a similar example in which the buyer has product-specific taste shocks.

[26] More generally, our results apply without change to the following model. R has a metric space Z of payoff parameters such that the distribution of $(\theta, z) \in \Theta \times Z$ is $\mu_0 \otimes \zeta_0$ for some $\zeta_0 \in \Delta Z$. R's payoffs are given by $u_R : A \times \Theta \times Z \to \mathbb{R}$ that is measurable over Z and continuous over $A \times \Theta$. In this extended model, $V : \Delta\Theta \rightrightarrows \mathbb{R}$ takes the form $V(\mu) = \int_Z \operatorname{co} u_S(\arg\max_{a\in A} \int_\Theta u_R(a, \theta, z)\,\mathrm{d}\mu(\theta))\,\mathrm{d}\zeta_0(z)$, a Kakutani correspondence.

[27] The logic yielding quasiconcavity of $\bar{v}^*$ is similar to Chakraborty and Harbaugh's (2010) observation that S's utility function in this example is quasiconvex as an increasing transformation of a convex function. We thank an anonymous referee for pointing out this connection.

[28] That is, the seller reveals the identity of $\arg\max_{i\in\{1,\dots,n\}} \theta_i$.

$\mathbb{E}[\max\{\epsilon, G^{-1}(s)\}]$. Hence, all equilibria are Pareto-ranked, and so any seller-best equilibrium is buyer-best as well.

When does the seller benefit from commitment? The answer depends on the relationship between G and its concave envelope, $\hat{G}$, evaluated at $t_0^* := \int \max_{j\in\{1,\ldots,n\}} \theta_j \, \mathrm{d}\mu_0(\theta)$.

CLAIM 3: *The seller benefits from commitment if and only if* $\hat{G}(t_0^*) > G(t_0^*)$.

To see that commitment can benefit the seller only if $\hat{G}(t_0^*) > G(t_0^*)$, observe that $\hat{v}^*(\mu) := \hat{G}(\int \max_{j\in\{1,\ldots,n\}} \theta_j \, \mathrm{d}\mu(\theta))$ is a continuous and concave function that lies everywhere above the seller's value function. Hence, the concave envelope of the seller's value function, $\hat{v}$, lies below $\hat{v}^*$. Thus, if the seller benefits from commitment,

$$\hat{G}(t_0^*) \geq \hat{v}(\mu_0) > \bar{v}(\mu_0) = G(t_0^*).$$

Conversely, suppose $\hat{G}(t_0^*) > G(t_0^*)$. Then, by reasoning analogous to Kamenica and Gentzkow's (2011) Proposition 3,[29] a seller with commitment power can strictly outperform p^* by providing additional information about the value of the best good. Thus, commitment always benefits the seller when $\hat{G}(t_0^*) > G(t_0^*)$.

Claim 3 reduces the question of whether commitment benefits the seller to comparing a one-dimensional function with its concave envelope. Such a comparison is simple when G is well behaved. In particular, if G admits a decreasing, increasing, or single-peaked density, G itself is concave, convex, or convex-concave, respectively, and so characterizing its concave envelope is straightforward.

CLAIM 4: *Suppose G admits a continuous density g.*

1. *If g is weakly decreasing, the seller does not benefit from commitment.*
2. *If g is nonconstant and weakly increasing, the seller benefits from commitment.*
3. *If g is strictly quasiconcave, the seller benefits from commitment if and only if* $g(t_0^*) > \frac{1}{t_0^*}\int_0^{t_0^*} g(t)\,\mathrm{d}t$.

The claim's first part says the seller does not benefit from commitment when g is decreasing, that is, when G is concave. The second part says that when G is convex and nonaffine, the seller always benefits from commitment. The third part discusses the seller's benefits from commitment when G is S-shaped. Specifically, it shows commitment is valuable in this case if and only if G's density at t_0^* is strictly larger than the average density up to t_0^*.

6. DISCUSSION

6.1. *Effective Communication*

In a seminal paper, Chakraborty and Harbaugh (2010) showed that a large special case of our model always admits an *influential equilibrium*, namely, an equilibrium in which

[29]Proposition 3 of Kamenica and Gentzkow (2011) assumes the state space is finite, and so does not directly apply here. However, the extension to this example is straightforward given that G is continuous. See Appendix B.3.

R's action is non-constant across S's on-path messages. In this section, we note their insight applies beyond their parametric setting, and implies informative communication is possible whenever three or more states exist.

We begin with a few definitions. A *statistic* is a continuous function T from Θ into some locally convex space $\mathcal{X}$. Say T is *multivariate* if its range is noncollinear, that is, the affine span of $T(\Theta)$ has dimension strictly greater than 1. Finally, given a belief $\mu \in \Delta\Theta$, its associated *estimate* of a statistic T is the barycenter $\int T \, \mathrm{d}\mu$.[30]

The above-defined objects arise naturally in Chakraborty and Harbaugh's (2010) setting. There, Θ and A are the same convex, multidimensional Euclidean set, and the prior admits a density. Moreover, R's unique optimal action given belief μ is his expectation of the state; that is, R chooses $a = \int T \, \mathrm{d}\mu$, where $T = \mathrm{id}_\Theta$. Chakraborty and Harbaugh (2010) showed an equilibrium exists in which the estimate of T, and therefore R's action, changes on path. Adapting Chakraborty and Harbaugh's (2010) logic, Proposition 1 highlights the key feature behind their result: T is multivariate.

PROPOSITION 1: *For any multivariate statistic T, an equilibrium outcome (p, s) exists such that the estimate of T is not p-almost surely constant.*

Observe Proposition 1 readily delivers an informative equilibrium whenever three or more states exist. The reason is that, in this case, the mapping $T(\theta) := \delta_\theta$ taking each state to a degenerate belief is a multivariate statistic (taking values in the span of $\Delta\Theta$). The proposition also yields an influential equilibrium whenever R's best response equals his estimate of a multivariate statistic, as is the case in Chakraborty and Harbaugh's (2010) model.

Proposition 1 delivers a generalization of another of Chakraborty and Harbaugh's (2010) insights: S always benefits from communication via cheap talk when v is a strictly quasiconvex function of R's estimate of a multivariate statistic. This conclusion roughly follows from the fact that a strictly quasiconvex function can be constant across a nondegenerate distribution only if it is strictly lower at the distribution's mean.[31]

6.2. *The Equilibrium Payoff Set*

Despite our focus on S's favorite equilibrium, our approach is useful for analyzing the entire equilibrium payoff set. To find S's payoff set, notice that because S's incentives are characterized by indifference, the game's equilibrium set of S strategies is the same regardless of whether S's objective is u_S or $-u_S$. Just as applying Theorem 1 to the original game characterizes S's high payoffs, one can apply the theorem to the game with S objective $-u_S$ to find S's low equilibrium payoffs. Under this objective, S's value function is given by $-w$, where $w(\cdot) := \min V(\cdot)$. Theorem 1 then implies $s \leq w(\mu_0)$ is an equilibrium payoff in the *original* game if and only if some $p \in \mathcal{I}(\mu_0)$ exists such that $p\{w \leq s\} = 1$. Applying Theorem 2 then tells us S's lowest equilibrium payoff is given by the quasiconvex of envelope of w, which we denote by $\underline{w}$.[32] The above reasoning gives S's entire equilibrium payoff set: s is an S equilibrium payoff if and only if $s \in [\underline{w}(\mu_0), \bar{v}(\mu_0)]$.

[30] Recall Phelps (2001, p. 1), the barycenter $\int T \, \mathrm{d}\mu$ is the unique $\tau \in \overline{\mathrm{co}} T(\Theta)$ such that $\varphi(\tau) = \int \varphi \circ T(\theta) \, \mathrm{d}\mu(\theta)$ for every continuous linear $\varphi : \mathcal{X} \to \mathbb{R}$.

[31] When R may have multiple best responses to a given belief, an additional step is needed. See Appendix C.1 of the Supplemental Material (Lipnowski and Ravid (2020)) for details.

[32] More precisely, $\underline{w}$ is the highest quasiconvex and lower semicontinuous function that is everywhere below w.

With S's equilibrium payoffs in hand, we can find R's possible equilibrium payoffs using two observations. First, one can implement any particular payoff profile in an equilibrium in which S recommends a pair of actions to R, and R responds by mixing only over the recommended actions. Second, if S's equilibrium payoff is s, S's recommended action pair must consist of one action yielding S a payoff above s, and one action yielding S a payoff below s. Taking s as given, we can thus reduce the number of action pairs that S may recommend in equilibrium. We discuss these observations more formally in Appendix C.2 of the Supplemental Material (Lipnowski and Revid (2020)).

6.3. *Long and Transparent Cheap Talk*

It is by now well-known that allowing multiple rounds of bilateral communication—that is, long cheap talk—expands the set of feasible equilibrium outcomes (e.g., see Forges (1990), Aumann and Hart (2003), and Krishna and Morgan (2004)). Forges (1990) characterized the long-cheap-talk payoff set in a striking example in which certain outcomes require infinitely many rounds of communication. Her characterization, which uses repeated-games techniques (e.g., see Hart (1985)), was generalized by Aumann and Hart (2003). Broadly, one can describe the long-cheap-talk outcome set in terms of separation by diconvex functions (Aumann and Hart (1986, 2003)). When S's preferences are state independent, one can obtain such a separating function for S's payoffs using Theorem 1. One can then show that every S payoff attainable in a Nash equilibrium with long cheap talk is also attainable in PBE of the one-shot cheap-talk game.[33] The same, however, is not true for R, who can benefit from long cheap talk. We refer the reader to Appendix C.3 for the formal details.

6.4. *Optimality of Full Revelation*

In this section, we ask when honesty is the best policy. More precisely, we provide a sufficient condition for full revelation to be an S-favorite equilibrium. To understand our conditions, starting with the commitment case is useful. When S can commit, full revelation is optimal whenever v is *nowhere concave*, that is, when every non-extreme prior, $\mu_0 \in \Delta\Theta \setminus \{\delta_\theta\}_{\theta\in\Theta}$, admits two beliefs, μ', μ'', and a $\lambda \in (0, 1)$, such that $\mu_0 = \lambda\mu' + (1-\lambda)\mu''$ and $v(\mu_0) < \lambda v(\mu') + (1-\lambda)v(\mu'')$. Intuitively, whenever v is nowhere concave, one can strictly improve on any non-full revelation policy by appropriately[34] splitting non-extreme beliefs in the policy's support. Hence, a non-full revelation policy cannot be optimal. Because an optimal policy exists, it must be full information.

Without commitment, one can use securability to obtain that full revelation is an S-favorite equilibrium whenever v is *nowhere quasiconcave*, that is, when, for every non-extreme prior, $\mu_0 \in \Delta\Theta \setminus \{\delta_\theta\}_{\theta\in\Theta}$, two beliefs, μ' and μ'', and a $\lambda \in (0, 1)$ exist, such that $\mu_0 = \lambda\mu' + (1-\lambda)\mu''$ and $v(\mu) < \min\{v(\mu'), v(\mu'')\}$. In fact, we show v being nowhere quasiconcave implies full revelation barely secures S's maximal equilibrium value. That

[33]Although the formal results therein are limited to finite settings, the Aumann and Hart (2003) setting is conceptually richer than ours, featuring a sender who may also make payoff-relevant decisions after communication concludes. For such settings, one can still show that one-shot *bilateral* communication is without loss for sender payoffs given state-independent preferences over action profiles. The driving observation is that jointly controlled lotteries deliver a Kakutani correspondence, to which one can then apply Theorem 1.

[34]In particular, one can benefit by splitting non-extreme beliefs in a measurable way. See Appendix C.4 for details.

full information secures S's maximal equilibrium value, $\bar{v}(\mu_0)$, under nowhere quasiconcavity is straightforward: By correctly splitting non-extreme beliefs, one can weakly increase the value secured by any non-full revelation policy. Showing full revelation *barely* secures $\bar{v}(\mu_0)$ requires a more subtle argument. We refer the reader to Appendix C.4 for the precise details.

We should remark that, whereas strict convexity is sufficient for nowhere concavity, strict *quasi*convexity of v is insufficient for v to be nowhere *quasi*concave. Indeed, full revelation can fail to be an equilibrium at any non-degenerate prior—even if v is strictly quasiconvex. The reason is that a strictly quasiconvex function can exhibit quasiconcavities on one-dimensional extreme subsets of its domain. We show such quasiconcavity is the only possible issue, however: A strictly quasiconvex v is nowhere quasiconcave if and only if it is nowhere quasiconcave on $\text{co}\{\delta_\theta, \delta_{\theta'}\}$ for all θ, θ'.[35]

Notice a nowhere quasiconcave v must also be nowhere concave. Therefore, whenever v is nowhere quasiconcave, full revelation is both an S-favorite equilibrium and S's unique optimal commitment policy. Note S could still benefit from commitment. The reason is that under cheap talk, R might need to break ties against S's interests due to S's incentive constraints. Appendix C.4 of the Supplemental Material contains such an example. The example also demonstrates that nowhere quasiconcavity is insufficient for full information to be S's unique favorite equilibrium. However, both issues disappear when R's best response to each belief is unique. Said differently, when R's best responses are unique, nowhere quasiconcavity of v is sufficient for full revelation to be the unique equilibrium attaining S's maximal commitment payoff. In this case, v's quasiconcave and concave envelopes coincide, that is, $\bar{v} = \hat{v}$.

APPENDIX A: OMITTED PROOFS: MAIN RESULTS

A.1. *Preliminaries and Additional Notation*

We begin by noting an abuse of notation that we use throughout the appendix. For a compact metrizable space Y, a Borel measure over it $\gamma \in \Delta Y$, and a γ-integrable function $f : Y \to \mathbb{R}$, we let $f(\gamma) = \int_Y f \, \mathrm{d}\gamma$.

We now document the (standard) notion of information ranking used throughout the paper. This definition is motivated by the Hardy–Littlewood–Polya–Blackwell–Stein–Sherman–Cartier theorem (see Phelps (2001)).

DEFINITION 1: Given $p, p' \in \Delta\Delta\Theta$, say *$p$ is more (Blackwell) informative than p'* if p is a mean-preserving spread of p', that is, if a measurable selector r of $\mathcal{I} : \Delta\Theta \rightrightarrows \Delta\Delta\Theta$ exists such that $p(D) = \int_{\Delta\Theta} r(D|\cdot) \, \mathrm{d}p'$ for all Borel $D \subseteq \Delta\Theta$.

Now, we record a useful measurable selection result.

LEMMA 2: *If $D \subseteq \Delta\Theta$ is Borel and $f : D \to \mathbb{R}$ is any measurable selector of $V|_D$, a measurable function $\alpha_f : D \to \Delta A$ exists such that, for all $\mu \in D$, the measure $\hat{\alpha} = \alpha_f(\cdot|\mu)$ satisfies*:

1. $u_S(\hat{\alpha}) = f(\mu)$;
2. $\hat{\alpha} \in \arg\max_{\alpha \in \Delta A} u_R(\alpha, \mu)$;
3. $|\text{supp}(\hat{\alpha})| \leq 2$.

[35] To prove this result, we note the generalization of Chakraborty and Harbaugh's (2010) ideas as in Proposition 1 implies a strictly quasiconvex v is not quasiconcave at any non-binary belief.

PROOF: The result follows readily from the measurable maximum theorem (Theorem 18.19 from Aliprantis and Border (2006)). Define

$$A^* : \Delta\Theta \rightrightarrows A,$$
$$\mu \mapsto \arg\max_{a\in A} u_R(a, \mu).$$

Notice A^* is nonempty-compact-valued and weakly measurable by the measurable maximum theorem. Applying the same theorem to $\mu \mapsto \arg\max_{a\in A^*(\mu)} u_S(a)$ and $\mu \mapsto \arg\min_{a\in A^*(\mu)} u_S(a)$, and noting $V = \mathrm{co}(u_S \circ A^*)$, delivers measurable selectors a_+ and a_- of A^* such that $u_S \circ a_+ = \max V$ and $u_S \circ a_- = \min V$.

But the same theorem delivers measurable selectors a_+ and a_- of A^* such that $u_S \circ a_+ = \max V$ and $u_S \circ a_- = \min V$. Now, define the measurable map:

$$\alpha_f : D \to \Delta A,$$
$$\mu \mapsto \begin{cases} \dfrac{v(\mu) - f(\mu)}{v(\mu) - \min V(\mu)}\delta_{a_-(\mu)} + \dfrac{f(\mu) - \min V(\mu)}{v(\mu) - \min V(\mu)}\delta_{a_+(\mu)} & : \min V(\mu) \neq f(\mu), \\ \delta_{a_-(\mu)} & : \min V(\mu) = f(\mu). \end{cases}$$

By construction, α_f is as desired. *Q.E.D.*

Next, we prove a variant of the intermediate value theorem, which is useful for our setting. This result is essentially proven in Lemma 2 of de Clippel (2008). Because the statement of that lemma is slightly weaker than we need, however, we provide a proof here for the sake of completeness.

LEMMA 3: *If* $F : [0, 1] \rightrightarrows \mathbb{R}$ *is a Kakutani correspondence with* $\min F(0) \leq 0 \leq \max F(1)$, *and* $\bar{x} = \inf\{x \in [0, 1] : \max F(x) \geq 0\}$, *then* $0 \in F(\bar{x})$.

PROOF: By definition of $\bar{x}$, some weakly decreasing $\{x_n^+\}_{n=1}^\infty \subseteq [\bar{x}, 1]$ exists that converges to $\bar{x}$ such that $\max F(x_n^+) \geq 0$ for every $n \in \mathbb{N}$. Define the sequence $\{x_n^-\}_{n=1}^\infty \subseteq [0, \bar{x}]$ to be the constant 0 sequence if $\bar{x} = 0$ and to be any strictly increasing sequence that converges to $\bar{x}$ otherwise. By definition of $\bar{x}$ (and, in the case of $\bar{x} = 0$, because $\min F(0) \leq 0$), it must be that $\min F(x_n^-) \leq 0 \leq \max F(x_n^+)$.

Passing to a subsequence if necessary, we may assume (as a Kakutani correspondence has compact range) $\{\max F(x_n^+)\}_{n=1}^\infty$ converges to some $y \in \mathbb{R}$, which would necessarily be nonnegative. Upper hemicontinuity of F then implies $\max F(\bar{x}) \geq 0$. An analogous argument shows $\min F(\bar{x}) \leq 0$. Because F is convex-valued, it follows that $0 \in F(\bar{x})$. *Q.E.D.*

A.2. *Proof for Section 2*

Below is the proof of Lemma 1, which initializes our belief-based approach. For finite states, the result can be easily proven from results in Aumann and Hart (2003). Although their ideas easily generalize to infinite state spaces such as ours, we include a direct proof here for completeness.

PROOF OF LEMMA 1: First, take any equilibrium (σ, ρ, β) and let (p, s) be the induced outcome. That $p \in \mathcal{I}(\mu_0)$ follows directly from the Bayesian property.

Define the interim payoff, $\hat{s} : M \to \mathbb{R}$ via $\hat{s}(m) := u_S(\rho(m))$. S incentive compatibility tells us some $M^* \subseteq M$ exists such that $\int_\Theta \beta(M^*|\cdot)\,\mathrm{d}\mu_0 = 1$, and for every $m \in M^*$ and

$m' \in M$, we have $\hat{s}(m) \geq \hat{s}(m')$. In particular, $\hat{s}(m) = \hat{s}(m')$ for every $m, m' \in M^*$; that is, some $\hat{s}^* \in \mathbb{R}$ exists such that $\hat{s}|_{M^*} = \hat{s}^*$. But

$$s = \int_\Theta \int_{M^*} u_S\big(\rho(m)\big)\,\mathrm{d}\sigma(m|\theta)\,\mathrm{d}\mu_0(\theta) = \int_\Theta \int_{M^*} \hat{s}^*\,\mathrm{d}\mu_0(\theta) = \hat{s}^*,$$

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 $\mu \in \Delta\Theta$. Because V is upper hemicontinuous, it follows that $s \in \bigcap_{\mu \in \mathrm{supp}(p)} V(\mu)$.

Now suppose (p, s) satisfies the three conditions. Define the compact set $D := \mathrm{supp}(p)$. It is well known (see Benoît and Dubra (2011) or Kamenica and Gentzkow (2011)) that every $p \in \mathcal{I}(\mu_0)$ exhibits some S strategy σ and Bayes-consistent belief map $\beta : M \to \Delta\Theta$ that induce distribution p over posterior beliefs.[36] Without disrupting the Bayesian property, we may without loss assume $\beta(m) \in D$ for all $m \in M$. Now let $\alpha = \alpha_s : D \to \Delta A$ be as given by Lemma 2. We can then define the receiver strategy $\sigma := \alpha \circ \beta$, which is incentive-compatible for R by definition of α. Finally, by construction, $\int_A u_S\,\mathrm{d}\rho(\cdot|m) = s$ for every $m \in M$, so that every S strategy is incentive-compatible. Therefore, (σ, ρ, β) is an equilibrium that generates outcome (p, s). *Q.E.D.*

A.3. *Proofs for Section 3*

A.3.1. *Proof of Theorem 1*

Below, we prove a lemma that is at the heart of Theorem 1. It constructs an equilibrium (a barely securing policy, which we then show to be compatible with equilibrium) of S value s from an arbitrary information policy securing s. The constructed equilibrium policy is less informative than the original policy and requires fewer messages to implement.

LEMMA 4: *Let $p \in \mathcal{I}(\mu_0)$ and $s \in \mathbb{R}$.*

1. *If p secures s and $s \geq v(\mu_0)$, some $p^* \in \mathcal{I}(\mu_0)$ exists such that p^* barely secures s, p^* is weakly less Blackwell-informative than p, and $|\mathrm{supp}(p^*)| \leq |\mathrm{supp}(p)|$.*
2. *If p barely secures s, (p, s) is an equilibrium outcome.*

PROOF: If $s = v(\mu_0)$, both results are trivial: In this case, the uninformative policy is the unique one that barely secures s. From this point, we focus on the case of $s > v(\mu_0)$.

Toward the first point, let $p \in \mathcal{I}(\mu_0)$ secure s, and $D := \mathrm{supp}(p)$. Notice $v(\mu) \geq s$ for every $\mu \in D$ because v is upper semicontinuous. Define the semicontinuous (and so measurable) function,

$$\begin{aligned} \lambda = \lambda_{p,s} : D &\to [0, 1], \\ \mu &\mapsto \inf\big\{\hat{\lambda} \in [0, 1] : v\big((1 - \hat{\lambda})\mu_0 + \hat{\lambda}\mu\big) \geq s\big\}. \end{aligned}$$

By Lemma 3, it must be that $s \in V([1 - \lambda(\mu)]\mu_0 + \lambda(\mu)\mu)$ for every $\mu \in D$.

Notice some number $\epsilon > 0$ exists such that $\lambda \geq \epsilon$ uniformly. If no such ϵ existed, a sequence $\{\mu_n\}_n \subseteq D$ would exist such that $\lambda(\mu_n)$ converges to zero. But the sequence

[36] In particular, such (σ, β) exist with $\sigma(\Delta\Theta|\theta) = 1$ for all $\theta \in \Theta$ and $\beta(\cdot|\mu) = \mu$ for all $\mu \in D$.

$\{([1-\lambda(\mu_n)]\mu_0+\lambda(\mu_n)\mu_n, s)\}_n$ from the graph of V would then converge to $(\mu_{0,} s)$. Because V is upper hemicontinuous, such convergence would contradict $s > v(\mu_0)$. Therefore, such an $\epsilon > 0$ exists, and so $\frac{1}{\lambda}$ is a bounded function.

Now, define $p^* = p^*_s \in \Delta\Delta\Theta$ by letting

$$p^*(\hat{D}) := \left(\int_{\Delta\Theta} \frac{1}{\lambda}\,\mathrm{d}p\right)^{-1} \cdot \int_{\Delta\Theta} \frac{1}{\lambda(\mu)} \mathbf{1}_{[1-\lambda(\mu)]\mu_0+\lambda(\mu)\mu\in\hat{D}}\,\mathrm{d}p(\mu)$$

for every Borel $\hat{D} \subseteq \Delta\Theta$. Direct computation shows $p^* \in \mathcal{I}(\mu_0)$, and p^* barely secures s by construction.

Last, we note p^* has the other required properties. The map $\mu \mapsto [1-\lambda(\mu)]\mu_0+\lambda(\mu)\mu$ is a surjection from $\mathrm{supp}(p^*)$ to $\mathrm{supp}(p)$, so that $|\mathrm{supp}(p^*)| \leq |\mathrm{supp}(p)|$. Also by construction, p^* is weakly less informative than $(1-\int_{\Delta\Theta}\lambda\,\mathrm{d}p)\delta_{\mu_0}+(\int_{\Delta\Theta}\lambda\,\mathrm{d}p)p$, which in turn is less informative than p. This proves (1).

Toward (2), suppose p barely secures s. That is, p-a.e. μ has $\{v \geq s\} \cap \mathrm{co}\{\mu, \mu_0\} = \{\mu\}$. For such μ, some subsequence of $\{v((1-2^{-n})\mu + 2^{-n}\mu_0)\}_{n=1}^{\infty} \subseteq [\min u_S(A), s]$ converges, leading to (as V is upper hemicontinuous) some element of $V(\mu)$ that is weakly less than s. Because $v(\mu) \geq s$ by hypothesis, and V is convex-valued, it follows that $s \in V(\mu)$. But upper hemicontinuity of V then implies $s \in V(\mu')$ for each $\mu' \in \mathrm{supp}(p)$, and Lemma 1 delivers an equilibrium that generates S value s and information policy p. *Q.E.D.*

We now prove the securability theorem (Theorem 1).

PROOF OF THEOREM 1: The "only if" direction follows directly from Lemma 1: For any equilibrium outcome (p, s), information policy p secures payoff s. The "if" direction is a direct consequence of (both parts of) Lemma 4. *Q.E.D.*

A.3.2. *Convexity of the Equilibrium Payoff Set, and Corollary 1*

Given Theorem 1, all that remains for proving Corollary 1 is that an S-best equilibrium exists, which follows from Corollary 3 below.

COROLLARY 3: *The set of sender equilibrium payoffs is a compact interval.*

PROOF: Let Π^* be the set of equilibrium S payoffs, $\Pi_+ := \{s \in \Pi^* : s \geq \max V(\mu_0)\}$, $\Pi_- := \{s \in \Pi^* : s \leq \min V(\mu_0)\}$, and $\Pi_0 := \{s \in \Pi^* : \min V(\mu_0) \leq s \leq \max V(\mu_0)\}$.

Because V is convex-valued, $\Pi_0 = \Pi^* \cap V(\mu_0)$. By considering uninformative equilibria, we see that $\Pi_0 = V(\mu_0) = [\min V(\mu_0), \max V(\mu_0)]$.

It follows immediately from Theorem 1 that Π_+ is convex. Letting $s_+ := \sup(\Pi_+) \geq v(\mu_0)$, a sequence $\{s_n\}_{n=1}^{\infty} \subseteq [v(\mu_0), s_+]$ exists that converges to s_+. Dropping to a subsequence, if necessary, we may assume some $\{p_n\}_{n=1}^{\infty} \subseteq \mathcal{I}(\mu_0)$ exists such that p_n secures s_n for each n, and $\{p_n\}_n$ converges to some $p_+ \in \mathcal{I}(\mu_0)$. But p_+ secures s_+ because v is upper semicontinuous, so that (by Theorem 1) $s_+ \in \Pi_+$. It follows that $\Pi_+ = [v(\mu_0), s_+]$, a compact interval. By an identical argument, Π_- is a compact interval, say, $[s_-, \min V(\mu_0)]$ as well.[37]

Therefore, $\Pi^* = [s_-, \min V(\mu_0)] \cup [\min V(\mu_0), \max V(\mu_0)] \cup [\max V(\mu_0), s_+] = [s_-, s_+]$. *Q.E.D.*

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

A.4. *Proofs for Section 4*

A.4.1. *Upper Semicontinuity of v^**

We prove here that v^* is upper semicontinuous, a fact that the main-text proof of Theorem 2 takes as given.

LEMMA 5: *v^* is upper semicontinuous.*

PROOF: Let $\tilde{v}^* : \Delta\Delta\Theta \to \mathbb{R}$ be given by $\tilde{v}^*(p) := \inf v(\operatorname{supp} p)$, so that $v^*(\mu) := \max_{p\in\mathcal{I}(\mu)} \tilde{v}^*(p)$ for every $\mu \in \Delta\Theta$. The correspondence $\operatorname{supp} : \Delta\Delta\Theta \rightrightarrows \Delta\Theta$ is lower hemicontinuous (Aliprantis and Border (2006, Theorem 17.14)). Because v is upper semicontinuous, it follows (Aliprantis and Border (2006, Lemma 17.29)) that $\tilde{v}^*$ is upper semicontinuous. Next, the correspondence $\mathcal{I} : \Delta\Theta \rightrightarrows \Delta\Delta\Theta$ is upper hemicontinuous because the barycenter map $(p \mapsto \int_{\Delta\Theta} \mu \,\mathrm{d}p(\mu))$ is continuous (Phelps (2001, Proposition 1.1)). Upper semicontinuity of v^* follows (Aliprantis and Border (2006, Lemma 17.30)). *Q.E.D.*

A.4.2. *Quasiconcave Envelope With Finite States*

The purpose of this section is to prove Corollary 4 below. The corollary says that, with finite states, $\bar{v}$ is the lowest quasiconcave function majorizing v. In other words, the "upper semicontinuous" qualifier in the definition of the quasiconcave envelope is necessary only when the state is infinite.

COROLLARY 4: *Suppose Θ is finite. Then, $\bar{v}$ lies below every quasiconcave function majorizing v.*

PROOF: Take any quasiconcave $f : \Delta\Theta \to \mathbb{R}$ majorizing v. We show $f \geq v^*$. The result then follows from $v^* = \bar{v}$ (Theorem 2). Fix some prior $\mu \in \Delta\Theta$ and let $p \in \mathcal{I}(\mu)$ be an information policy securing S's favorite equilibrium value, $v^*(\mu)$. Because Θ is finite, Carathéodory's theorem delivers a finite subset $D \subseteq \operatorname{supp} p$ whose convex hull includes the prior. Combined with f being a quasiconcave function majorizing v, we have that

$$v^*(\mu) = \inf v(\operatorname{supp} p) \leq \min v(D) \leq \min f(D) \leq f(\mu),$$

as required. *Q.E.D.*

A.4.3. *Corollary 2: Commitment Is Usually Valuable*

We now prove Corollary 2, for which it suffices to show that Lebesgue-almost every prior μ_0 has either $\bar{v}(\mu_0) = s^{\mathrm{FB}} := \max v(\Delta\Theta)$ or $\hat{v}(\mu_0) > \bar{v}(\mu_0)$.

PROOF: First, observe

$$\bar{v}(\Delta\Theta) = v^*(\Delta\Theta) \subseteq \mathrm{cl}\big[v(\Delta\Theta)\big] \subseteq \mathrm{cl}\big[u_S(A)\big] = u_S(A),$$

which is finite. Next, that $\bar{v}$ is quasiconcave implies $\{\bar{v} \geq s\}$ is convex for every $s \in u_S(A)$. Let

$$D := (\Delta\Theta)^\circ \setminus \bigcup_{s\in u_S(A)} \partial\{\bar{v} \geq s\}$$

be the set of full-support beliefs that are not on the boundary of any $\bar{v}$-upper contour set. Being the boundary of a bounded convex set in a $(|\Theta| - 1)$-dimensional space, the set

$\partial\{\bar{v} \geq s\}$ is a manifold of dimension strictly lower than $|\Theta| - 1$ for each $s \in u_S(A)$, and so has zero Lebesgue measure. Because $(\Delta\Theta)^\circ$ has full Lebesgue measure, the finite union $\Delta\Theta \setminus D$ is Lebesgue-null as well.[38]

Suppose $\mu_0 \in D$, and fix some belief $\mu \in \Delta\Theta$ such that some action in $u_S^{-1}(s^{\mathrm{FB}})$ is a best response for R to belief μ. By definition of D, sufficiently small $\epsilon \in (0, 1]$ will have $\epsilon\mu \leq \mu_0$ and $\bar{v}(\frac{\mu_0 - \epsilon\mu}{1-\epsilon}) \geq \bar{v}(\mu_0)$. But $\hat{v}$ being concave and lying above $\bar{v}$,

$$\hat{v}(\mu_0) \geq (1-\epsilon)\bar{v}\left(\frac{\mu_0 - \epsilon\mu}{1-\epsilon}\right) + \epsilon\bar{v}(\mu) \geq (1-\epsilon)\bar{v}(\mu_0) + \epsilon s^{\mathrm{FB}}.$$

Thus, the proof is complete: Either $\bar{v}(\mu_0) < \hat{v}(\mu_0)$ or $\bar{v}(\mu_0) = s^{\mathrm{FB}}$. *Q.E.D.*

APPENDIX B: OMITTED PROOFS: APPLICATIONS

B.1. *Proofs for Section 5.1: The Think Tank*

In this example, $A = \{0, \dots, n\}$, $\Theta = [0, 1]^n$, μ_0 is exchangeable, u_S is increasing with $u_S(0) = 0$, and

$$u_R(a, \theta) = \begin{cases} \theta_i - c & : a = i \in \{1, \dots, n\}, \\ 0 & : a = 0. \end{cases}$$

We now invest in some notation. For $\theta \in \Theta$ and $k \in \{1, \dots, n\}$, let $\theta_{k,n}^{(1)} := \max_{i \in \{k,\dots,n\}} \theta_i$ be the first-order statistic among reforms better (for S) than k. For finite $\hat{M} \subseteq M$, let $\mathcal{U}(\hat{M}) \in \Delta(\hat{M}) \subseteq \Delta M$ be the uniform measure over $\hat{M}$. Given $k \in \{1, \dots, n\}$, let

$$\sigma_k : \Theta \to \Delta\{k, \dots, n\} \subseteq \Delta M,$$
$$\theta \mapsto \mathcal{U}\Big(\arg \max_{i \in \{k,\dots,n\}} \theta_i\Big),$$

be the S strategy that reports the best reform from among those the think tank prefers to k; let $\beta_k : M \to \Delta\Theta$ be some belief map such that σ_k and β_k are together Bayes consistent; and let $p_k \in \mathcal{I}(\mu_0)$ be the associated information policy. For any measurable $f : \Theta \to [0, 1]$, let $\mathbb{E}_0 f(\theta) := \int f \, \mathrm{d}\mu_0$; and for $k \in \{1, \dots, n\}$ and $i \in \{k, \dots, n\}$, let $\mathbb{E}_i^k f(\theta) := \int f \, \mathrm{d}\beta_k(\cdot|i)$. Finally, for any $k \in \{1, \dots, n\}$, let $\hat{\theta}^k := \mathbb{E}_0 \theta_{k,n}^{(1)}$.

B.1.1. *Claim 1: Ranking the Best Reforms*

Toward the proof of Claim 1, we first show the following.

CLAIM: *Fix $k \in \{1, \dots, n\}$ and $i \in \{k, \dots, n\}$. Then, $i \in \arg\max_{a \in A} u_R(a, \beta_k(i))$ if and only if $\hat{\theta}^k \geq c$.*

PROOF: For a given $i \in \{k, \dots, n\}$, exchangeability of μ_0 implies the following four facts:

(1) $\mathbb{E}_0\theta_i = \mathbb{E}_0\theta_j = \mathbb{E}_i^k\theta_j$ for $j \in \{1, \dots, k-1\}$.
(2) $\mathbb{E}_0\theta_i \in \mathrm{co}\{\mathbb{E}_i^k\theta_i, \mathbb{E}_i^k\theta_j\}$ for $j \in \{k, \dots, n\} \setminus \{i\}$.

[38] By the same argument, $\Delta\Theta \setminus D$ is also nowhere dense.

(3) $\mathbb{E}_i^k \theta_i \geq \mathbb{E}_0 \theta_i$.
(4) $\mathbb{E}_i^k \theta_i = \hat{\theta}^k$.
The first three facts collectively tell us $\mathbb{E}_i^k \theta_i \geq \mathbb{E}_i^k \theta_j$ for $j \in \{1, \dots, n\} \setminus \{i\}$. As an implication, $i \in \arg\max_{a \in A} u_R(a, \beta_k(i))$ if and only if $\mathbb{E}_i^k \theta_i \geq c$. The fourth fact completes the proof of the claim. *Q.E.D.*

PROOF OF CLAIM 1: Now, we prove the three-way equivalence of Claim 1. First, that Part 2 implies Part 3 follows from the above claim. Next, that Part 3 implies Part 1 follows directly from Theorem 1. Now, to show Part 1 implies Part 2, consider any equilibrium yielding S value $u_S(k)$. In this equilibrium, every on-path message yields value $u_S(k)$ to S, implying some reform from $\{k, \dots, n\}$ is incentive-compatible for R. That is, R has an optimal strategy in which his gross benefit is one of $\{\theta_i\}_{i=k}^n$ almost surely. But R's ex ante payoff is no greater than the prior expectation of $\max_{i \in \{k, \dots, n\}} \theta_i - c$. This expectation is then nonnegative by R's incentives: He does not want to deviate to the status quo ex ante. Thus, Part 1 implies Part 2, completing the proof of Claim 1. *Q.E.D.*

B.1.2. *Construction of an S-Best Equilibrium*

Finally, Corollary 1 tells us the sender's best equilibrium value lies in $\{0, \dots, n\}$, so that the S-optimal equilibrium payoff is $u_S(k^*)$, where

$$k^* = \begin{cases} \max\{k \in \{1, \dots, n\} : \hat{\theta}^k \geq c\} & : \hat{\theta}^1 \geq c, \\ 0 & : \hat{\theta}^1 < c. \end{cases}$$

As described in Section 5.1, we can use the constructive proof of Theorem 1 to explicitly derive the modification of p_{k^*} that supports payoff $u_S(k^*)$ as an equilibrium payoff when $k^* > 0$. Let $\epsilon := \frac{\hat{\theta}^{k^*} - c}{\hat{\theta}^{k^*} - \hat{\theta}^n}$, and consider the truth-or-noise signal $\sigma^* := (1 - \epsilon)\sigma_{k^*} + \epsilon \mathcal{U}\{k^*, \dots, n\}$. That is, among the proposals that the think tank weakly prefers to k^*, it either reports the best (with probability $1 - \epsilon$, independent of the state) or a random one. Following a recommendation $i \in \{k, \dots, n\}$, the lawmaker is indifferent between reform i and no reform at all. He responds with $\rho(i|i) = \frac{u_S(k^*)}{u_S(i)}$ and $\rho(0|i) = 1 - \rho(i|i)$. The proof of Lemma 4 shows such play is in fact equilibrium play.

B.2. *Proofs for Section 5.2: The Broker*

B.2.1. *The One-Dimensional Model*

In this section, we look at a one-dimensional version of our model, which generalizes Example 2, analyzed in Section 5.2. Our task is to prove a generalization of Claim 2 that applies for all priors (including those exhibiting atoms).

Suppose $\Theta \subseteq \mathbb{R}$ and that some $v_M : \operatorname{co}\Theta \to \mathbb{R}$ exists such that $v = v_M \circ E$, where $E : \Delta\Theta \to \operatorname{co}\Theta$ maps each belief to its associated expectation of the state. This setting, which we call the *one-dimensional model*, was studied in Gentzkow and Kamenica (2016) and Dworczak and Martini (2019) under sender commitment power. We assume without loss that $\operatorname{co}\Theta = [0, 1]$, and denote the prior mean by $\theta_0 = E\mu_0$.

An important concept to simplify analysis of the one-dimensional model is the notion of a cutoff policy. Given $q \in [0, 1]$, the *q-quantile-cutoff policy* is the (necessarily unique) information policy $p^q \in \mathcal{I}(\mu_0)$ of the form $p^q = q\delta_{\mu_-^q} + (1 - q)\delta_{\mu_+^q}$, for $\mu_-^q, \mu_+^q \in \Delta\Theta$ with $\max\operatorname{supp}(\mu_-^q) \leq \min\operatorname{supp}(\mu_+^q)$; and let $\theta_-^q := E\mu_-^q$ and $\theta_+^q := E\mu_+^q$. Say $p \in \mathcal{I}(\mu_0)$ is a *cutoff policy* if it is the q-quantile-cutoff policy for some $q \in [0, 1]$. The following alternative characterization of cutoff policies, which is immediate, is useful for analyzing the one-dimensional model.

FACT 1: *For $q \in [0, 1]$, the belief μ_-^q (μ_+^q) is the unique solution to the program* $\min_{\mu\in\Delta\Theta:q\mu\leq\mu_0} E\mu$ ($\max_{\mu\in\Delta\Theta:(1-q)\mu\leq\mu_0} E\mu$).

The q-quantile-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, S simply reports whether the state is above or below some well-calibrated cutoff.[39] As the following claim shows, securability enables us to use cutoff policies to analyze many one-dimensional applications of interest, including the broker example.

CLAIM 5: *Suppose $\Theta \subseteq \mathbb{R}$ and $v_M : \text{co}\,\Theta \to \mathbb{R}$ is a weakly quasiconvex function such that $v = v_M \circ E$. Then, the following are equivalent for all $s \geq v(\mu_0)$:*
(i) *S can attain a payoff s in equilibrium.*
(ii) *The payoff s is securable by a cutoff policy.*
Moreover, an S-preferred equilibrium outcome (p, s) exists such that p is a cutoff policy.

PROOF: Because v_M is quasiconvex, v_M is either nonincreasing on $[0, \theta_0]$ or nondecreasing on $[\theta_0, 1]$. Suppose the latter holds without loss. Because uninformative communication is a cutoff policy with cutoff quantile 0 or 1, the result is immediate if $s = v(\mu_0)$, so we may assume $s > v(\mu_0)$.

That (ii) implies (i) follows directly from Theorem 1. Now we suppose (i) holds and show (ii) does as well. The nonempty (because s is securable) compact sets $\Theta_L := \{\theta \in [0, \theta_0] : v_M(\theta) \geq s\}$ and $\Theta_R := \{\theta \in [\theta_0, 1] : v_M(\theta) \geq s\}$ both exclude θ_0 because $s > v(\mu_0)$. Let $\theta_L := \max \Theta_L$ and $\theta_R := \min \Theta_R$. By Theorem 1 and Lemma 4, a Bayes-plausible information policy p exists that barely secures s, which then implies $p \circ E^{-1}\{\theta_L, \theta_R\} = 1$. That is, some $\hat{q} \in (0, 1)$, $p_L \in \Delta[E^{-1}(\theta_L)]$, $p_R \in \Delta[E^{-1}(\theta_R)]$ exist such that $p = \hat{q} p_L + (1 - \hat{q}) p_R$. But Fact 1 implies $\theta_-^{\hat{q}} \leq \theta_L$ and $\theta_+^{\hat{q}} \geq \theta_R$. Because $\theta_-^{\hat{q}} \leq \theta_L < \theta_0 = \theta_-^1$, the intermediate value theorem (and Berge's theorem, which tells us from Fact 1 that $\theta_-^{(\cdot)}$ is continuous) delivers some $q_2 \in [\hat{q}, 1)$ such that $\theta_-^{q_2} = \theta_L$. Similarly, some $q_1 \in (0, \hat{q}]$ exists such that $\theta_+^{q_1} = \theta_R$. Now, because $\theta_-^{q_2} = \theta_L$, $\theta_+^{q_2} \geq \theta_R$, and $v_M|_{[\theta_0,1]}$ is nondecreasing, it follows that p^{q_2} secures s.

To prove the "moreover" part, we specialize to the case in which $s = \bar{v}(\mu_0)$. Let $Q := [q_1, q_2]$, $Q_+ := \{q \in Q : v_M(\theta_+^q) = s\}$, and $Q_- := \{q \in Q : v_M(\theta_-^q) = s\}$. That no value strictly above s is securable implies $Q = Q_+ \cup Q_-$. Therefore, the union of the closures has the same property: $\bar{Q}_+ \cup \bar{Q}_- = Q$. Because Q is connected (because v_M is monotone on each side of $\hat{q}$), some $q \in \bar{Q}_+ \cap \bar{Q}_-$ must then exist. That V is upper hemicontinuous (together with Lemma 1) then implies the q-cutoff policy, paired with payoff s, is an equilibrium outcome. *Q.E.D.*

Although not directly relevant to the broker example, we briefly note one can apply Claim 5 to simplify the one-dimensional model even when v_M is not quasiconvex. We do so in Corollary 5 below.

[39]This description is correct as stated in the case in which μ_0 is atomless; if the cutoff is itself a state with positive prior probability, S's message may need to be random conditional on the cutoff state itself occurring.

COROLLARY 5: *Suppose $\Theta \subseteq \mathbb{R}$ and $V_M : \mathrm{co}\,\Theta \rightrightarrows \mathbb{R}$ is such that $V = V_M \circ E$. Then, for any equilibrium sender payoff s, an equilibrium outcome of the form (p, s) exists, such that p is a garbling of a cutoff policy (with at most two supported posterior beliefs).*

PROOF: We have nothing to show for $s \in V(\mu_0)$. We now focus on the case of $s > v(\mu_0)$, the alternative case being symmetric.

Define the correspondence $\tilde{V}_M : [0, 1] \rightrightarrows \mathbb{R}$ by letting $\tilde{V}_M(\theta) := V_M(\mathrm{co}\{\theta, \theta_0\})$ for every $\theta \in [0, 1]$. Appealing to Lemma 3, $\tilde{V}_M$ is a Kakutani correspondence, so that $\tilde{V} := \tilde{V}_M \circ E : \Delta\Theta \rightrightarrows \mathbb{R}$ is as well. We can therefore apply the mathematical results of Claim 5, letting $\tilde{v}_M := \max \tilde{V}_M$ (which is quasiconvex and minimized at θ_0) replace v_M to find a cutoff $q \in [0, 1]$ such that $\tilde{v}_M(\theta^q_-), \tilde{v}_M(\theta^q_+) \geq s$. But, by definition of $\tilde{V}_M$, some two-message garbling p' of p^q exists that secures s in the original game, that is, has $p'\{v \geq s\} = 1$. Finally, Lemma 4 delivers a further two-message garbling p of p' such that (p, s) is an equilibrium outcome. *Q.E.D.*

B.2.2. *The Investor's Payoff (Equation (1))*

Suppose (p, s) is an equilibrium outcome of the broker example (Example 2), and let r be R's associated payoff. Then,

$$\begin{aligned}
&r + \mathbb{V}_{\theta\sim\mu_0}(\theta) \\
&= \int \left\{ \frac{1}{2} \int \left(a_0^2 - 2a_0\theta + \theta^2\right) \mathrm{d}\mu_0 \right. \\
&\quad \left. - \frac{1}{2} \int \left[a^*(\mu)^2 - 2a^*(\mu)\theta + \theta^2\right] \mathrm{d}\mu - u_S\big(a^*(\mu)\big) \right\} \mathrm{d}p(\mu) \\
&= \int \left\{ \int \left(\frac{1}{2}\theta^2 - a_0\theta\right)\left[\mathrm{d}\mu_0(\theta) - \mathrm{d}\mu(\theta)\right] \right. \\
&\quad \left. + \left[a^*(\mu) - a_0\right]E\mu + \frac{1}{2}\left[a_0^2 - a^*(\mu)^2\right] - s \right\} \mathrm{d}p(\mu) \\
&= 0 + \int \left[a^*(\mu) - a_0\right] \left\{ E\mu - \frac{1}{2}\left[a_0 + a^*(\mu)\right] \right\} \mathrm{d}p(\mu) - s \\
&= \int \left[a^*(\mu) - a_0\right] \left\{ \left[E\mu - a^*(\mu)\right] + \frac{1}{2}\left[a^*(\mu) - a_0\right] \right\} \mathrm{d}p(\mu) - s \\
&= \int \left\{ \left[a^*(\mu) - a_0\right]\left[E\mu - a^*(\mu)\right] + \frac{1}{2}\left[a^*(\mu) - a_0\right]^2 \right\} \mathrm{d}p(\mu) - s \\
&= \int \left[s + \frac{1}{2}\left(\frac{s}{\phi}\right)^2 \right] \mathrm{d}p(\mu) - s \\
&= \frac{1}{2\phi^2} s^2,
\end{aligned}$$

where the second to last equality follows from separately analyzing the case $a^*(\mu) = a_0$ and the complementary case.

B.3. *Proofs for Section 5.3: The Salesperson*

We begin by providing an alternative version of Kamenica and Gentzkow's (2011) Proposition 6 (which generalizes their Proposition 3). Their proposition shows an S-beneficial equilibrium exists whenever S's value function is a transformation of R's estimate of a finite-dimensional statistic, said transformation disagrees with its concave envelope, and the state is finite. We show that with sufficient continuity, the same conclusion holds when the state is infinite.

LEMMA 6: *Suppose some $N \in \mathbb{N}$ admits continuous $T : \Theta \to \mathbb{R}^N$ and continuous $G : \operatorname{co} T(\Theta) \to \mathbb{R}$ such that $v(\mu) = G(\int T\,\mathrm{d}\mu)$ for all beliefs $\mu \in \Delta\Theta$. If the concave envelope $\hat{G}$ of G satisfies $\hat{G}(\int T\,\mathrm{d}\mu_0) > G(\int T\,\mathrm{d}\mu_0)$, then $\hat{v}(\mu_0) > v(\mu_0)$.*

PROOF: Let $X := \operatorname{co} T(\Theta)$ and $x_0 := \int T\mu_0$, which is in the relative interior of X. By Carathéodory's theorem, that $\hat{G}(x_0) > G(x_0)$ means some $\tilde{p} \in \Delta X$ exists with affinely independent support $\{x^1, \dots, x^K\}$ such that $\int x\,\mathrm{d}\tilde{p}(x) = x_0$ and $\int G\,\mathrm{d}\tilde{p} > v(\mu_0)$. As G is continuous, we may assume without loss that $\{x^1, \dots, x^K\}$ has X in its affine hull.[40] For sufficiently small convex neighborhood $Y = \prod_{k=1}^K Y_k$ of $(x^1, \dots, x^K)$ in X^K, every $\vec{y} = (y^1, \dots, y^K) \in Y$ has $y^1, \dots, y^K$ affinely independent with x_0 in their convex hull, and so admits a unique $\tilde{p}_{\vec{y}} \in \Delta\{y^1, \dots, y^K\}$ such that $\int x\,\mathrm{d}\tilde{p}_{\vec{y}} = x_0$. Observe $\vec{y} \mapsto \tilde{p}_{\vec{y}}$ is continuous because $\vec{y} \mapsto \tilde{p}_{\vec{y}}(y^k)$ is an affine function of its finite-dimensional argument for each $k \in \{1, \dots, K\}$. Moreover, making Y smaller if necessary, we may assume $\int G\,\mathrm{d}\tilde{p}_{\vec{y}} > v(\mu_0)$ for every $\vec{y} \in Y$, because G is continuous.

Observe now that $D := \prod_{k=1}^K \{\mu_k \in \Delta\Theta : \int T\,\mathrm{d}\mu_k \in Y_k\}$ is a nonempty open subset of $(\Delta\Theta)^K$ such that every $\vec{\mu} \in D$ admits some $p_{\vec{\mu}} \in \Delta\{\mu_k\}_{k=1}^K$ with $\int(\int T\,\mathrm{d}\mu)\,\mathrm{d}p_{\vec{\mu}}(\mu) = x_0$ and $\int v\,\mathrm{d}p_{\vec{\mu}} > v(\mu_0)$. Indeed, D is open because Y is and T is continuous; D is nonempty because $Y \subseteq X^K$ is, and because every $x \in X$ admits a $\mu \in \Delta\Theta$ with $x = \int T\,\mathrm{d}\mu$; and $p_{\vec{\mu}}$ can be taken to be $\sum_{k=1}^K \tilde{p}_{(\int T\,\mathrm{d}\mu_1, \dots, \int T\,\mathrm{d}\mu_K)}(\int T\,\mathrm{d}\mu_k)\delta_{\mu_k}$.

Finally, Lemma 2 of Lipnowski and Mathevet (2018) says that the set of all $\mu \in \Delta\Theta$ such that $\tilde{\epsilon}\mu \le \mu_0$ for some $\tilde{\epsilon} > 0$ is dense. Therefore, D being open and nonempty delivers $\vec{\mu} \in D$ and $\epsilon > 0$ such that $\epsilon \sum_{k=1}^K \mu_k \le \mu_0$. Then, defining $\mu^* := \frac{1}{1-\epsilon}[\mu_0 - \epsilon \int \mu\,\mathrm{d}p_{\vec{\mu}}(\mu)]$ and $p^* := (1-\epsilon)\delta_{\mu^*} + \epsilon p^* \in \mathcal{I}(\mu_0)$, see that $\int v\,\mathrm{d}p^* - v(\mu_0) = \epsilon[\int v\,\mathrm{d}p_{\vec{\mu}} - v(\mu_0)] > 0$. *Q.E.D.*

With this lemma in hand, we readily complete the proof of Claim 3.

PROOF: In the main text, we demonstrated that $\hat{G}(t_0^*) > G(t_0^*)$ is necessary for commitment to strictly benefit the seller. To see it is sufficient, apply Lemma 6 with $N = 1$: The seller gets a value strictly higher than $\bar{v}(\mu_0) = G(t_0^*)$ by telling the buyer which product is best and by further revealing some (well-chosen) information about the value of the best product. *Q.E.D.*

[40] Indeed, observe K must lie in $\{1, \dots, N+1\}$, so let $\tilde{p}$ be chosen to make K as large as possible. Assume for a contradiction that $x^{K+1} \in X$ is outside the affine hull of $\{x^1, \dots, x^K\}$. As x_0 is relatively interior and $x_0 - x^{K+1}$ is a convex combination of $\{x^i - x^{K+1}\}_{i=1}^K$, some $i \in \{1, \dots, K\}$ has $x^i + \epsilon(x^i - x^{K+1}) \in X$ for sufficiently small $\epsilon > 0$. Then, consider $\tilde{p}_\epsilon := \tilde{p} + \tilde{p}(x^i)[\frac{1}{1+\epsilon}\delta_{x^i+\epsilon(x^i - x^{K+1})} + \frac{\epsilon}{1+\epsilon}\delta_{x^{K+1}} - \delta_{x^i}] \in \Delta X$. This measure has $\int x\,\mathrm{d}\tilde{p}_\epsilon(x) = x_0$ by construction and, converging to $\tilde{p}$ as $\epsilon > 0$, has $\int G\,\mathrm{d}\tilde{p}_\epsilon > v(\mu_0)$ when ϵ is sufficiently small, contradicting the maximality of K.

Now, en route to Claim 4, we prove the following slightly more general result about when commitment is valuable for a CDF G admitting a single-peaked continuous density g. For this purpose, let

$$\varphi_G : [0,1] \to \mathbb{R},$$
$$t \mapsto G(t) - G(0) - tg(t) = -\int_0^t \tilde{t}\,\mathrm{d}g(\tilde{t}),$$

where the equality follows from integration by parts.

LEMMA 7: *Suppose G admits a continuous, weakly quasiconcave density g. Let $t_M := \min[\arg\max_{t\in[0,1]} g(t)]$. Then, $\hat{G}(t_0^*) = G(t_0^*)$ if and only if $t_0^* \geq t_M$ and $\varphi_G(t_0^*) \geq 0$.*

PROOF: First, we show $\varphi_G(t_0^*) \geq 0$ is necessary for no commitment gap to exist. To that end, suppose $\varphi_G(t_0^*) < 0$. Recall that full support of μ_0 implies $t_0^* \in (0,1)$. Then, letting $\epsilon \in (0, 1-t_0^*]$, we have

$$\begin{aligned}\frac{t_0^*+\epsilon}{\epsilon}\big[\hat{G}\big(t_0^*\big) - G\big(t_0^*\big)\big] &\geq \frac{t_0^*+\epsilon}{\epsilon}\left[\frac{t_0^*}{t_0^*+\epsilon}G\big(t_0^*+\epsilon\big) + \frac{\epsilon}{t_0^*+\epsilon}G(0) - G\big({t^*}_0\big)\right] \\ &= t_0^*\frac{G\big(t_0^*+\epsilon\big) - G\big(t_0^*\big)}{\epsilon} - \big[G\big(t_0^*\big) - G(0)\big],\end{aligned}$$

which tends to $-\varphi_G(t_0^*) > 0$, as $\epsilon \to 0$. Therefore, $\frac{t_0^*+\epsilon}{\epsilon}[\hat{G}(t_0^*) - G(t_0^*)] > 0$ when $\epsilon > 0$ is sufficiently small, so that $\hat{G}(t_0^*) > G(t_0^*)$.

Now, we verify that $t_0^* \geq t_M$ is necessary for no commitment gap to exist. Suppose $t_0^* < t_M$. Then $g|_{[0,t_M]}$ is continuous, weakly increasing, and nonconstant. Therefore, $G|_{[0,t_M]}$ is weakly convex and not affine, implying

$$\hat{G}\big(t_0^*\big) \geq \frac{t_M - t_0^*}{t_M}G(0) + \frac{t_0^*}{t_M}G(t_M) > G\big(t_0^*\big).$$

Conversely, suppose $t_0^* \geq t_M$ and $\varphi_G(t_0^*) \geq 0$. Below, we construct a continuous concave function, G^*, that majorizes G and agrees with it at t_0^*. It follows $G^*(t_0^*) \geq \hat{G}(t_0^*) \geq G(t_0^*) = G^*(t_0^*)$, that is, there is no commitment gap.

Toward finding such a G^*, observe first φ_G decreases on $[0, t_M]$ (because $g|_{[0,t_M]}$ is increasing) and $\varphi_G(0) = 0$. Therefore, $\varphi_G(t_M) \leq 0 \leq \varphi_G(t_0^*)$. Because φ_G is continuous, the intermediate value theorem delivers a $t_* \in [t_M, t_0^*]$ with $\varphi_G(t_*) = 0$. We now use t_* to construct G^*. To do so, note

$$G(t_*) = G(0) + \int_0^{t_*} g(t)\,\mathrm{d}t = G(0) + t_*g(t_*) - \varphi_G(t_*) = G(0) + t_*g(t_*),$$

meaning

$$G^* : [0,1] \to \mathbb{R},$$
$$t \to \begin{cases} G(0) + tg(t_*) & : t \leq t_*, \\ G(t) & : t \geq t_*, \end{cases}$$

is well-defined and continuously differentiable. We now claim G^* satisfies the desired properties. Observe first $G^*(t_0^*) = G(t_0^*)$, because $t_0^* \geq t_*$. Second, because $t_* \geq t_M$, g is decreasing on $[t_*, 1]$, meaning G^* has a decreasing derivative, that is, G^* is concave. Thus, it remains to show G^* majorizes G. Because $G^*(t) = G(t)$ for all $t \geq t_*$ by construction, it remains to show $G^*(t) \geq G(t)$ for all $t < t_*$. For $t \in [t_M, t_*)$, observe $g|_{[t_M, t_0^*]}$ is decreasing, and so

$$G^*(t) - G(t) = \left[G^*(t) - G(t)\right] - \left[G^*(t_*) - G(t_*)\right] = \int_t^{t_*} \left[g(\tilde{t}) - g\left(t^*\right)\right] \mathrm{d}\tilde{t} \geq 0.$$

For $t \in [0, t_M)$, observe $G^*(0) = G(0)$, $G^*(t_M) \geq G(t_M)$, G is convex, and G^* is concave. Therefore,

$$G^*(t) \geq \frac{t}{t_M} G(t_M) + \frac{t_M - t}{t_M} G(0) \geq G(t).$$

The proof is now complete. *Q.E.D.*

From this, we can prove Claim 4 easily.

PROOF OF CLAIM 4: First, suppose g is weakly decreasing. Then, $t_0^* \geq 0 = t_M$ and $G(t_0^*) - G(0) = \int_0^{t_0^*} g(t)\,\mathrm{d}t \geq t_0^* g(t_0^*)$, and Lemma 7 applies.

Second, suppose g is nonconstant and increasing. If $t_0^* < t_M$, then $\hat{G}(t_0^*) > G(t^*)$ by Lemma 7. If $t_0^* \geq t_M$, then $g(t_0^*) \geq g(t_M) > g(0)$, implying $g|_{[0, t_0^*]}$ is continuous, nonconstant, and increasing. So $g|_{[0, t_0^*]}$ is below $g(t_0^*)$ everywhere, and strictly below it for some nondegenerate interval. Therefore, $t_0^* g(t_0^*) > \int_0^{t_0^*} g(t)\,\mathrm{d}t = G(t_0^*) - G(0)$, and Lemma 7 applies.

Third, suppose g is strictly quasiconcave. For any $\tilde{t} \in (0, t_M]$, the function g is continuous and strictly increasing on $[0, \tilde{t}]$. This tells us φ_G is nonconstant and decreasing on $[0, \tilde{t}]$, implying $\varphi_G(\tilde{t}) < \varphi_G(0) = 0$. Therefore, if $\varphi_G(t_0^*) \geq 0$, then $t_0^* \neq \tilde{t}$. Because $\tilde{t} \in (0, t_M]$ was arbitrary, we now know that if $\varphi_G(t_0^*) \geq 0$, then $t_0^* \geq t_M$. Thus, by Lemma 7, a commitment gap exists if and only if

$$0 > \varphi_G\left(t_0^*\right) = G\left(t_0^*\right) - G(0) - t_0^* g\left(t_0^*\right) = \int_0^{t_0^*} g(t)\,\mathrm{d}t - t_0^* g\left(t_0^*\right).$$

The claim follows. *Q.E.D.*

REFERENCES

ALIPRANTIS, C. D., AND K. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Science & Business Media. [1649,1652]

ALONSO, R., AND O. CÂMARA (2016): "Persuading Voters," *American Economic Review*, 106 (11), 3590–3605. [1632]

AUMANN, R. J., AND S. HART (1986): "Bi-Convexity and bi-Martingales," *Israel Journal of Mathematics*, 54 (2), 159–180. [1647]

——— (2003): "Long Cheap Talk," *Econometrica*, 71 (6), 1619–1660. [1632,1636,1647,1649]

AUMANN, R. J., AND M. MASCHLER (1966): "Game Theoretic Aspects of Gradual Disarmament," *Report of the US Arms Control and Disarmament Agency*, 80, 1–55. [1639]

——— (1995): *Repeated Games With Incomplete Information*. MIT Press. [1632,1636,1639]

BATTAGLINI, M. (2002): "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70 (4), 1379–1401. [1632]

BENOÎT, J.-P., AND J. DUBRA (2011): "Apparent Overconfidence," *Econometrica*, 79 (5), 1591–1625. [1636, 1650]

BEST, J., AND D. QUIGLEY (2020): "Persuasion for the Long Run." [1631]

CHAKRABORTY, A., AND R. HARBAUGH (2007): "Comparative Cheap Talk," *Journal of Economic Theory*, 132 (1), 70–94. [1632]

——— (2010): "Persuasion by Cheap Talk," *American Economic Review*, 100 (5), 2361–2382. [1631,1632, 1634,1636-1638,1644-1646,1648]

——— (2014): "Persuasive Puffery," *Marketing Science*, 33 (3), 382–400. [1644]

CHE, Y. K., W. DESSEIN, AND N. KARTIK (2013): "Pandering to Persuade," *American Economic Review*, 103 (1), 47–79. [1635]

CHUNG, W., AND R. HARBAUGH (2019): "Biased Recommendations From Biased and Unbiased Experts," *Journal of Economics & Management Strategy*, 28 (3), 520–540. [1631,1635]

CRAWFORD, V. P., AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50 (6), 1431–1451. [1631]

DE CLIPPEL, G. (2008): "An Axiomatization of the Inner Core Using Appropriate Reduced Games," *Journal of Mathematical Economics*, 44 (3), 316–323. [1649]

DWORCZAK, P., AND G. MARTINI (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127 (5), 1993–2048. [1654]

ELY, J. C. (2017): "Beeps," *American Economic Review*, 107 (1), 31–53. [1632]

FORGES, F. (1990): "Equilibria With Communication in a Job Market Example," *The Quarterly Journal of Economics*, 105 (2), 375–398. [1647]

GENTZKOW, M., AND E. KAMENICA (2016): "A Rothschild–Stiglitz Approach to Bayesian Persuasion," *American Economic Review*, 106 (5), 597–601. [1654]

GLAZER, J., AND A. RUBINSTEIN (2004): "On Optimal Rules of Persuasion," *Econometrica*, 72 (6), 1715–1736. [1633]

——— (2006): "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach," *Theoretical Economics*, 1 (4), 395–410. [1633]

GREEN, J. R., AND N. L. STOKEY (2007): "A Two-Person Game of Information Transmission," *Journal of Economic Theory*, 135 (1), 90–104. [1631]

HART, S. (1985): "Nonzero-sum Two-Person Repeated Games With Incomplete Information," *Mathematics of Operations Research*, 10 (1), 117–153. [1647]

HART, S., I. KREMER, AND M. PERRY (2017): "Evidence Games: Truth and Commitment," *American Economic Review*, 107 (3), 690–713. [1633]

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615. [1632,1633,1636,1639,1645,1650,1657]

KRISHNA, V., AND J. MORGAN (2004): "The Art of Conversation: Eliciting Information From Experts Through Multi-Stage Communication," *Journal of Economic theory*, 117 (2), 147–179. [1647]

LIPNOWSKI, E., AND L. MATHEVET (2018): "Disclosure to a Psychological Audience," *American Economic Journal: Microeconomics*, 10 (4), 67–93. [1657]

LIPNOWSKI, E., AND D. RAVID (2020): "Supplement to 'Cheap Talk With Transparent Motives'," *Econometrica Supplemental Material*, 88, https://doi.org/10.3982/ECTA15674. [1646]

MARGARIA, C., AND A. SMOLIN (2018): "Dynamic Communication With Biased Senders," *Games and Economic Behavior*, 110, 330–339. [1631]

PHELPS, R. R. (2001): *Lectures on Choquet's Theorem*. Springer Science & Business Media. [1646,1648,1652]

RAPPOPORT, D. (2020): "Evidence and Skepticism in Verifiable Disclosure Games." [1633]

SCHNAKENBERG, K. E. (2015): "Expert Advice to a Voting Body," *Journal of Economic Theory*, 160, 102–113. [1631,1639]

Co-editor Joel Sobel handled this manuscript.

Manuscript received 11 September, 2017; final version accepted 23 March, 2020; available online 30 March, 2020.

SUPPLEMENT TO "CHEAP TALK WITH TRANSPARENT MOTIVES"

(*Econometrica*, Vol. 88, No. 4, July 2020, 1631–1660)

ELLIOT LIPNOWSKI
Department of Economics, Columbia University

DORON RAVID
Department of Economics, University of Chicago

APPENDIX C

IN THIS APPENDIX, we elaborate on the results mentioned in Section 6 of "Cheap Talk With Transparent Motives" and discuss some additional relevant results.

C.1. *Proof of Proposition 1: Effective Communication*

We now operationalize Chakraborty and Harbaugh's (2010) insight of using fixed-point reasoning to show effective communication is possible, proving Proposition 1. We begin by representing the prior as an average of three posterior beliefs, μ_1, μ_2, and μ_3, such that the three induced estimates of the statistic are noncollinear; one can always find such beliefs because the statistic is itself multivariate. Next, we find a circle of beliefs around the prior within the convex hull of $\{\mu_1, \mu_2, \mu_3\}$. By construction, each belief on said circle yields a different estimate of the statistic. We then document a generalization of the one-dimensional Borsuk–Ulam theorem, which yields an antipodal pair of beliefs μ and μ' on the circle such that $V(\mu) \cap V(\mu')$ is nonempty. Therefore, we can split the prior across μ and μ' to obtain an equilibrium information policy.

In what follows, define the circle $\mathbb{S} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, and let $T\mu$ denote the estimate $\int T \,\mathrm{d}\mu$ of statistic T for any belief $\mu \in \Delta\Theta$.

LEMMA 8: *Let T be a multivariate statistic. Then, a continuous $\varphi : \mathbb{S} \to \Delta\Theta$ exists such that every $z \in \mathbb{S}$ has*:

1. $\frac{1}{2}\varphi(z) + \frac{1}{2}\varphi(-z) = \mu_0$;
2. $T(\varphi(z)) \neq T(\varphi(\hat{z}))$ *for every* $\hat{z} \in \mathbb{S} \setminus \{z\}$;
3. $2\varphi(z) - \mu_0 \in \Delta\Theta$.

PROOF: By assumption, $T(\Theta)$ is noncollinear, and so $T\mu_0 \notin \mathrm{co}\{T\theta_1, T\theta_2\}$ for some distinct $\theta_1, \theta_2 \in \Theta$. Because μ_0 has full support, both $\mu_0(N_1) > 0$ and $\mu_0(N_2) > 0$ for any open neighborhoods N_1 of θ_1 and N_2 of θ_2. We can then define the conditional distribution $\mu_i(\cdot) := \frac{\mu_0(N_i \cap (\cdot))}{\mu_0(N_i)}$ for $i \in \{1, 2\}$. Letting N_1, N_2 be sufficiently small neighborhoods, we may assume $N_1 \cap N_2 = \emptyset$, $T\mu_0 \notin \mathrm{co}\{T\mu_1, T\mu_2\}$, and $\mu(N_1 \cup N_2) < 1$. Therefore, letting $\mu_3(\cdot) := \frac{\mu_0((\cdot)\setminus(N_1 \cup N_2))}{1-\mu_0(N_1 \cup N_2)}$, we know that $\mu_0 \in \mathrm{co}\{\mu_1, \mu_2, \mu_3\}$, that μ_0 is not in the convex hull any two of $\{\mu_1, \mu_2, \mu_3\}$, and that the three points $\{T\mu_1, T\mu_2, T\mu_3\}$ are affinely independent. So $\mu_0 = \sum_{i=1}^{3} \lambda_i \mu_i$ for some $\mu_1, \mu_2, \mu_3 \in \Delta\Theta$ and $\lambda_1, \lambda_2, \lambda_3 \in (0, 1)$. Therefore,

Elliot Lipnowski: e.lipnowski@columbia.edu
Doron Ravid: dravid@uchicago.edu

Available at https://www.econometricsociety.org. https://doi.org/10.3982/ECTA15674

letting $\epsilon := \frac{1}{2}\min\{\lambda_1, \lambda_2, \lambda_3\}$, define the map

$$\varphi : \mathbb{S} \to \Delta\Theta,$$
$$(x, y) \mapsto (\lambda_1 + \epsilon x)\mu_1 + (\lambda_2 + \epsilon y)\mu_2 + \big[\lambda_3 - \epsilon(x + y)\big]\mu_3.$$

Affine independence of $T\mu_1, T\mu_2, T\mu_3$ ensures $T \circ \varphi$ is injective, and the other desiderata for φ are obviously satisfied. *Q.E.D.*

Next, we document a generalization of the one-dimensional Borsuk–Ulam theorem.

LEMMA 9: *Suppose $f : \mathbb{S} \to \mathbb{R}$ is upper semicontinuous, and every $z \in \mathbb{S}$ has* $\max\{f(z), f(-z)\} \geq 0$. *Then, some $z \in \mathbb{S}$ exists such that* $\min\{f(z), f(-z)\} \geq 0$.

PROOF: Define $\tilde{f} : \mathbb{S} \to \mathbb{R}$ by letting $\tilde{f}(z) := f(-z)$. By hypothesis, both f and $\tilde{f}$ are upper semicontinuous and $\{\tilde{f} < 0\} \subseteq \{f \geq 0\}$. Assume for a contradiction that the lemma fails, so that $\{\tilde{f} \geq 0\} \subseteq \{f < 0\}$. Because $\{\tilde{f} < 0\} \cup \{\tilde{f} \geq 0\} = \mathbb{S}$ and $\{f \geq 0\} \cap \{f < 0\} = \emptyset$, these containments in fact imply $\{\tilde{f} < 0\} = \{f \geq 0\}$ and $\{\tilde{f} \geq 0\} = \{f < 0\}$. But (given the definition of $\tilde{f}$) the two sets would both be empty if either were, and so would fail to cover $\mathbb{S}$. Therefore, the set $\{f \geq 0\}$ is a nonempty clopen proper subset of the connected space $\mathbb{S}$, a contradiction. *Q.E.D.*

We now complete the proof of the generalization of Chakraborty and Harbaugh's (2010) Theorem 1.

PROOF OF PROPOSITION 1: First, let $\varphi : \mathbb{S} \to \mathbb{R}$ be as delivered by Lemma 8. Next, define the function

$$f : \mathbb{S} \to \mathbb{R},$$
$$z \mapsto \max V\big(\varphi(z)\big) - \min V\big(\varphi(-z)\big).$$

Two properties of f are immediate. First, f is upper semicontinuous because V is upper hemicontinuous. Second, any $z \in \mathbb{S}$ satisfies $f(z) + f(-z) \geq 0$ because $\max V \geq \min V$. Therefore, Lemma 9 delivers $z \in \mathbb{S}$ with $f(z), f(-z) \geq 0$. That is, $\max V(\varphi(z)) \geq \min V(\varphi(-z))$ and $\max V(\varphi(-z)) \geq \min V(\varphi(z))$. Said differently (recall V is convex-valued), $V(\varphi(z)) \cap V(\varphi(-z)) \neq \emptyset$. Lemma 1 then guarantees the existence of an equilibrium that generates information policy $p = \frac{1}{2}\delta_{\varphi(z)} + \frac{1}{2}\delta_{\varphi(-z)}$. In particular, $T\mu$ is not $p(\mu)$-a.s. constant. *Q.E.D.*

Just as Proposition 1 generalizes Chakraborty and Harbaugh's (2010) Theorem 1, the following result generalizes their Theorem 2.

COROLLARY 6: *Let T be any statistic, and suppose $\tilde{u} : \overline{\text{co}}T(\Theta) \to \mathbb{R}$ is a strictly quasiconvex function such that $v(\mu) = \tilde{u}(T\mu)$ for every $\mu \in \Delta\Theta$. If T is multivariate, an S-beneficial equilibrium exists.*

Before proving this result, we note the result follows immediately from Proposition 1 under the additional hypothesis that R has a unique best response to every belief—as assumed in Chakraborty and Harbaugh (2010). Indeed, following Chakraborty and Harbaugh's (2010) argument, strict quasiconvexity of $\tilde{u}$ would imply the binary-message equilibrium constructed above is S-beneficial. The below proof for the general case is similar in spirit, although one additional step is needed.

PROOF OF COROLLARY 6: Again, let $\varphi : \mathbb{S} \to \mathbb{R}$ be as delivered by Lemma 8. Now, define $f := v \circ \varphi - v(\mu_0) : \mathbb{S} \to \mathbb{R}$, which is upper semicontinuous because v is. Moreover, for any $z \in \mathbb{S}$, the distinct estimates $T\varphi(z)$ and $T\varphi(-z)$ have $T\mu_0$ as their midpoint, and so $\max\{f(z), f(-z)\} \geq 0$ by quasiconvexity of $\tilde{u}$. Applying Lemma 9 to f then delivers a $z \in \mathbb{S}$ such that $v \circ \varphi(z), v \circ \varphi(-z) \geq v(\mu_0)$.

By Lemma 8 Part 3, both $\mu := 2\varphi(z) - \mu_0$ and $\mu' := 2\varphi(-z) - \mu_0$ are in $\Delta\Theta$. Because $T\varphi(z) = \frac{1}{2}T\mu + \frac{1}{2}T\mu_0$, strict quasiconvexity of $\tilde{u}$ delivers the following inequality chain:

$$v(\mu_0) \leq v \circ \varphi(z) = \tilde{u}\big(T\varphi(z)\big) < \max\big\{\tilde{u}(T\mu), \tilde{u}(T\mu_0)\big\} = \max\big\{v(\mu), v(\mu_0)\big\}.$$

It follows $v(\mu) > v(\mu_0)$. By the same argument, $v(\mu') > v(\mu_0)$. Thus, the information policy $p = \frac{1}{2}\delta_\mu + \frac{1}{2}\delta_{\mu'}$ secures $\min\{v(\mu), v(\mu')\} > v(\mu_0)$. The result then follows from Theorem 1. *Q.E.D.*

C.2. *The Equilibrium Payoff Set*

In this subsection, we briefly comment on how our tools, and the belief-based approach more broadly, can generate a more complete picture of the world of cheap talk with state-independent S preferences. As will be clear, the results outlined herein are all straightforward to derive given earlier results in the paper.

C.2.1. *Other Sender Payoffs*

Following the recent literature on communication with S commitment, our focus has largely been on high equilibrium S values, that is, those providing payoffs at least as high as those attainable under uninformative communication. However, the tools developed in our paper work equally well to characterize bad sender payoffs. Indeed, the proof of Lemma 1 used no special features of V other than it being a Kakutani correspondence, which $-V$ is as well. Therefore, our game has the same equilibrium distributions over $A \times \Theta$ as the game with S objective $-u_S$. To deliver the mirror-image versions of our main results, define the value function from S-adversarial tiebreaking, $w := \min V : \Delta\Theta \to \mathbb{R}$.

Theorem 1 implies a sender payoff $s \leq w(\mu_0)$ is an equilibrium payoff if and only if some $p \in \mathcal{I}(\mu_0)$ exists such that $p\{w \leq s\} = 1$. Combining this observation with the original statement of the securability theorem tells us $s \in \mathbb{R}$ is an equilibrium S payoff if and only if $p_+, p_- \in \mathcal{I}(\mu_0)$ exist such that $p_+\{v \geq s\} = p_-\{w \leq s\} = 1$. An easy consequence is that the equilibrium S payoff set is convex, which we document in Corollary 3. Corollary 1 has a mirror image as well, telling us the set of S equilibrium payoffs is exactly

$$\Big[\min_{p \in \mathcal{I}(\mu_0)} \sup w(\operatorname{supp} p),\ \max_{p \in \mathcal{I}(\mu_0)} \inf v(\operatorname{supp} p)\Big].$$

Note convexity of the set of attainable S payoffs is special to the case in which S's payoffs are state independent; indeed, the leading example of Crawford and Sobel (1982) does not share this feature.

The mirrored counterpart of our geometric Theorem 2 is that the lowest S payoff attainable in equilibrium is $\underline{w}(\mu_0)$, where $\underline{w}$ is the quasiconvex envelope of w, that is, the pointwise highest quasiconvex and lower semicontinuous function that minorizes w. Therefore, we can geometrically characterize S's equilibrium payoff set as $[\underline{w}(\mu_0), \bar{v}(\mu_0)]$.

C.2.2. *Receiver Payoffs*

Our most powerful tools (the securability theorem and its descendants) pertain to S payoffs. However, the belief-based approach (i.e., Lemma 1) can be used to describe R payoffs as well. Indeed, let $v_R : \Delta\Theta \to \mathbb{R}$ be R's value function, given by $v_R(\mu) := \max_{a\in A} \int_\Theta u_R(a, \cdot)\,\mathrm{d}\mu$. It follows from R's interim rationality that any equilibrium that generates outcome (p, s) will deliver a payoff of $r = \int_{\Delta\Theta} v_R\,\mathrm{d}p$ to R.

Given equilibrium S payoff s, we can then more explicitly derive the set of equilibrium R payoffs compatible with an equilibrium in which S gets payoff s. Let

$$B_s := \{w \leq s \leq v\} = \left\{\mu \in \Delta\Theta : \exists a_+, a_- \in \arg\max_{a\in A} \int_\Theta u_R(a, \cdot)\,\mathrm{d}\mu \text{ s.t. } u_S(a_-) \leq s \leq u_S(a_+)\right\}.$$

Then, (s, r) is an equilibrium payoff profile if and only if $r = \int_{\Delta\Theta} v_R\,\mathrm{d}p$ for some $p \in \mathcal{I}(\mu_0) \cap \Delta(B_s)$. The best such R payoff (given s) is given by $\widehat{v_R^s}(\mu_0)$, where $v_R^s : B_s \to \mathbb{R}$ is the restriction of v_R and $\widehat{v_R^s} : \overline{\text{co}}B_s \to \mathbb{R}$ is the concave envelope of v_R^s.

C.2.3. *Implementing Equilibrium Payoffs*

In addition to their role in proving Theorem 1, barely securing policies generate a straightforward way of implementing any equilibrium S payoff.[41] If S could commit, we could apply the revelation principle[42] to implement any S commitment payoff with a commitment protocol in which S makes a pure action recommendation to R, and R always complies. Using barely securing policies, we can show a similar result holds with cheap talk, with one important caveat: R must be allowed to mix. To state this result, for any S strategy σ, define $\mathcal{M}_\sigma$ as the set of messages in σ's support.[43]

PROPOSITION 2: *Fix some S payoff s. Then, the following are equivalent*:
1. *s is generated by an equilibrium*.
2. *s is generated by an equilibrium with $\mathcal{M}_\sigma \subseteq \Delta A$ and $\rho(\alpha) = \alpha\ \forall \alpha \in \mathcal{M}_\sigma$*.
3. *s is generated by an equilibrium with $\mathcal{M}_\sigma \subseteq A$ and $\rho(a|a) > 0\ \forall a \in \mathcal{M}_\sigma$*.

The proposition suggests two ways in which one can implement a payoff of s via incentive-compatible recommendations. The first way has S giving R a *mixed* action recommendation that R *always* follows. The second way has S giving R a *pure* action recommendation that R *sometimes* follows. Both ways can result in R mixing.

That 1 implies 2 follows from standard revelation principle logic. To prove 1 implies 3,[44] we start with a minimally informative information policy that secures s. Because p is minimally informative, it must barely secure s, meaning (p, s) is an equilibrium. Let $\mathcal{E}$

---

[41]For S payoffs $s \leq \min V(\mu_0)$, we use the mirror image of barely securing policies, that is, information policies p such that $\{\min V(\cdot) \leq s\} \cap \text{co}\{\mu, \mu_0\} = \{\mu\}$ holds for p-a.e. μ.

[42]See, for example, Myerson (1986), Kamenica and Gentzkow (2011), and Bergemann and Morris (2016).

[43]That is, let $\mathcal{M}_\sigma = \bigcup_{\theta\in\Theta} \text{supp}\,\sigma(\cdot|\theta)$.

[44]The equivalence between 1 and 3 echoes an important result of Bester and Strausz (2001), who studied a mechanism-design setting with one agent, finitely many types, and partial commitment by the principal.

be Part 2's implementation of (p, s), and take $\mathbf{a}(\mu)$ to be some S-preferred action among all those that R plays in $\mathcal{E}$ at belief μ. By minimality of p, $\mathbf{a}(\cdot)$ must be p-essentially one-to-one, because pooling any posteriors that induce the same $\mathbf{a}(\cdot)$ value would yield an even less informative policy that secures s. Thus, $\mathbf{a}(\cdot)$ takes distinct beliefs to distinct (on-path) actions: R can infer μ from $\mathbf{a}(\mu)$. One can then conclude the proof by having S recommend $\mathbf{a}(\mu)$ and R respond to $\mathbf{a}(\mu)$ as he would have responded to μ under $\mathcal{E}$.

The formal proof is below.

PROOF OF PROPOSITION 2: Because 2 and 3 each immediately imply 1, we show the converses.

Suppose s is an equilibrium S payoff. Now take some $p \in \mathcal{I}(\mu_0)$ Blackwell-minimal among all policies securing payoff s, and let $D := \text{supp}(p) \subseteq \Delta\Theta$.[45] Lemma 4 guarantees (p, s) is an equilibrium outcome, say, witnessed by equilibrium $\mathcal{E}_1 = (\sigma_1, \rho_1, \beta_1)$. Letting $\alpha = \alpha_s : D \to \Delta A$ be as delivered by Lemma 2, we may assume $\rho_1(\cdot|m) = \alpha(\cdot|\beta(m))$. In particular, ρ_1 specifies finite-support play for every message.

Let $\mathbb{M} := \text{marg}_M \mathbb{P}_{\mathcal{E}_1}$ and $X := \text{supp}[\mathbb{M} \circ \hat{\rho}^{-1}] \subseteq \Delta A$, and fix arbitrary $(\hat{\alpha}, \hat{\mu}) \in \text{supp}[\mathbb{M} \circ (\rho_1, \beta_1)^{-1}]$; in particular, $\hat{\alpha} \in X$. By continuity of u_R and receiver incentive compatibility, $\hat{\alpha} \in \arg\max_{\alpha \in \Delta A} u_R(\alpha \otimes \hat{\mu})$. Defining $\rho' : M \to \Delta A$ (resp. $\beta' : M \to \Delta\Theta$) to agree with ρ_1 (β_1) on path and take value $\hat{\alpha}$ ($\hat{\mu}$) off path, an equilibrium $\mathcal{E}' = (\sigma_1, \rho', \beta')$ exists such that $\mathbb{P}_{\mathcal{E}'} = \mathbb{P}_{\mathcal{E}_1}$ and $\rho'(\cdot|m) \in X$ for every $m \in M$.

Now define

$$
\begin{aligned}
\sigma_2 : \Theta &\to \Delta X \subseteq \Delta M, \\
\theta &\mapsto \sigma_1(\cdot|\theta) \circ \rho'^{-1}, \\
\rho_2 : M &\to X \subseteq \Delta A, \\
m &\mapsto \begin{cases} m & : m \in X, \\ \hat{\alpha} & : m \notin X, \end{cases} \\
\beta_2 : M &\to \Delta\Theta, \\
m &\mapsto \begin{cases} \mathbb{E}_{m \sim \mathbb{M}}\big[\beta(m)|\rho(m)\big] & : m \in X, \\ \hat{\mu} & : m \notin X. \end{cases}
\end{aligned}
$$

By construction, $(\sigma_2, \rho_2, \beta_2)$ is an equilibrium that generates outcome (p, s),[46] proving 1 implies 2.

Now define the (A- and D-valued, respectively) random variables $\mathbf{a}, \boldsymbol{\mu}$ on $\langle D, p\rangle$ by letting $\mathbf{a}(\mu) := \arg\max_{a \in \text{supp}\,\alpha(\mu)} u_S(a)$ and $\boldsymbol{\mu}(\mu) := \mu$ for $\mu \in D$. Next define the conditional expectation $\mathbf{f} := \mathbb{E}_p[\boldsymbol{\mu}|\mathbf{a}] : D \to D$, which is defined only up to a.e.-p equivalence. By construction, the distribution of $\boldsymbol{\mu}$ is a mean-preserving spread of the distribution of $\mathbf{f}$. That is, p is weakly more informative than $p \circ \mathbf{f}^{-1}$. By hypothesis, $\mathbf{a}(\mu)$ is incentive compatible

Applying a graph-theoretic argument, they showed one can restrict attention to direct mechanisms in which the agent reports truthfully with positive probability. Although the proof techniques are quite different, a common lesson emerges. Agent mixing helps circumvent limited commitment by the principal: in Bester and Strausz's (2001) setting, by limiting the principal's information, and in ours, by limiting her control.

[45]Some policy secures s if s is an equilibrium payoff. The set of such policies is closed (and so compact) because v is upper semicontinuous. Therefore, because the Blackwell order is closed-continuous, a Blackwell-minimal such policy exists.

[46]It generates $(\tilde{p}, s)$ for some garbling $\tilde{p}$ of p. Minimality of p then implies $\tilde{p} = p$.

for R at every $\mu \in D$. But $D = \text{supp}(p \circ \mathbf{f}^{-1})$, which implies $p \circ \mathbf{f}^{-1}$ secures s. But minimality of p implies $p \circ \mathbf{f}^{-1} = p$. So $\mathbf{f} = \mathbb{E}_p[\boldsymbol{\mu}|\mathbf{a}]$ and $\boldsymbol{\mu}$ have the same distribution, which implies $\mathbf{f} = \boldsymbol{\mu}$ a.s.-p. By definition, $\mathbf{f}$ is $\mathbf{a}$-measurable, so that Doob–Dynkin delivers some measurable $\mathbf{b} : A \to D$ such that $\mathbf{f} = \mathbf{b} \circ \mathbf{a}$.

Summing up, we have some measurable $\mathbf{b} : A \to D$ such that $\mathbf{b} \circ \mathbf{a} =_{\text{a.e.}-p} \boldsymbol{\mu}$. Now define

$$\begin{aligned}
\sigma_3 : \Theta &\to \Delta A \subseteq \Delta M, \\
&\theta \mapsto \sigma_2(\cdot|\theta) \circ (\mathbf{a} \circ \beta_2)^{-1}, \\
\rho_3 : M &\to X \subseteq \Delta A, \\
&m \mapsto \begin{cases} \alpha\big(\mathbf{b}(m)\big) & : m \in A, \\ \hat{\alpha} & : m \notin A, \end{cases} \\
\beta_3 : M &\to \Delta\Theta, \\
&m \mapsto \begin{cases} \mathbf{b}(m) & : m \in A, \\ \hat{\mu} & : m \notin A. \end{cases}
\end{aligned}$$

By construction, $(\sigma_3, \rho_3, \beta_3)$ is an equilibrium that generates outcome (p, s), proving 1 implies 3. *Q.E.D.*

Proposition 2 shows some forms of communication are without loss as far as S payoffs are concerned. First, any S equilibrium payoff is attainable in an equilibrium in which S recommends mixed actions that are (on path) followed exactly. This equivalence extends to equilibrium payoff *pairs*, with the same argument: Pooling messages that lead to the same R behavior relaxes incentive constraints and generates the same joint distribution over actions and states, preserving payoffs. Second, any S equilibrium payoff is attainable in an equilibrium in which S recommends pure actions that are followed with positive probability. Whether this result holds in general for payoff pairs is an open question. It is easy to see why, at least, our argument does not go through as stated. The proof begins by considering an information policy that gives no "extraneous" information to R, subject to securing the relevant S value. But taking information away from R in this way can result in a payoff loss.

Still, we can leverage Lemma 1 to show a result of a similar spirit: To implement an equilibrium payoff profile, it is sufficient for R to only use binary mixed actions, the support of which is S's message.

PROPOSITION 3: *Fix some payoff profile (s, r). Then, the following are equivalent*:

1. *(s, r) is generated by an equilibrium.*
2. *(s, r) is generated by an equilibrium with $\mathcal{M}_\sigma \subseteq \Delta A$ and $\rho(\alpha) = \alpha$ $\forall \alpha \in \mathcal{M}_\sigma$.*
3. *(s, r) is generated by an equilibrium with $\mathcal{M}_\sigma \subseteq \{\frac{1}{2}\delta_a + \frac{1}{2}\delta_{a'} : a, \hat{a} \in A\}$ and* $\text{supp}[\rho(\alpha)] = \text{supp}(\alpha)$ $\forall \alpha \in \mathcal{M}_\sigma$.

We can interpret 3 as describing equilibria in which S tells R, "Play a or $\hat{a}$," for some pair of actions, but does not suggest mixing probabilities.

To see the equivalence between 1 and 3, Lemma 2 from the Appendix can be used to show equilibrium payoff profile (s, r) can be implemented with an equilibrium in which R only ever uses pure actions or binary-support mixtures, with the latter only being used when S is not indifferent between the two supported actions. Without loss, say such equilibrium is as in 2, with S suggesting an incentive-compatible mixture to R. But S rationality implies no two on-path recommendations can have the same support, because then S would have an incentive to deviate to the one putting a higher probability on the preferred action. Therefore, the same behavior could be induced by having every message replaced with a uniform distribution over its (at most binary) support, and the result follows.

With finitely many actions, Proposition 3 yields an a priori upper bound on the number of distinct messages required in equilibrium, similar to Proposition 2. Still, the upper bound of Proposition 2 is significantly smaller: Whereas Proposition 2 says no more than $n := |A|$ messages are required to span the set of equilibrium S values, Proposition 3 guarantees any equilibrium payoff pair can be attained with at most $\frac{n(n-1)}{2}$ messages.

C.3. *Long Cheap Talk*

Let us define the long-cheap-talk game. In addition to the objects in our model section, R has some message space $\tilde{M}$, which we assume is compact metrizable. Let $\mathcal{H}_{<\infty} := \bigsqcup_{t=0}^{\infty}(M \times \tilde{M})^t$, $\mathcal{H}_\infty := (M \times \tilde{M})^{\mathbb{N}}$, and $\Omega := \mathcal{H}_\infty \times A \times \Theta$. In a long-cheap-talk game, S first sees the state $\theta \in \Theta$. Then, at each time $t \in \mathbb{Z}_+$, players send simultaneous messages: S sends $m_t \in M$ and R sends $\tilde{m}_t \in \tilde{M}$. Finally, after seeing the sequence of messages, R chooses an action $a \in A$. Formally, a (behavior) strategy for S is a measurable function $\sigma : \Theta \times \mathcal{H}_{<\infty} \to \Delta M$, and a strategy for R is a pair of measurable functions $(\tilde{\sigma}, \rho)$, where $\tilde{\sigma} : \mathcal{H}_{<\infty} \to \Delta\tilde{M}$ and $\rho : \mathcal{H}_\infty \to \Delta A$. These maps induce (together with the prior μ_0) a unique distribution, $\mathbb{P}_{\sigma,\tilde{\sigma},\rho} \in \Delta\Omega$, which induces payoff $u_S(\text{marg}_A \mathbb{P}_{\sigma,\tilde{\sigma},\rho})$ and $u_R(\text{marg}_{A\times\Theta} \mathbb{P}_{\sigma,\tilde{\sigma},\rho})$ for S and R, respectively.

C.3.1. *Extra Rounds Cannot Help the Sender*

Below, we use our Theorem 1 to show that any S payoff attainable under long cheap talk is also attainable under one-shot communication.[47]

PROPOSITION 4: *Every sender payoff attainable in a Nash equilibrium of the long-cheap-talk game is also attainable in a perfect Bayesian equilibrium of the one-shot cheap-talk game.*

To prove the proposition, fix a payoff s^* that S cannot attain in the one-shot game, and use our securability theorem to construct a continuous biconvex function on $\Delta\Theta \times \mathbb{R}$ that is strictly positive at (μ_0, s^*) and zero on V's graph. Mimicking Appendix A.3 of Aumann and Hart (2003), we then take an arbitrary equilibrium of the long-cheap-talk game, and construct a bimartingale $\{\boldsymbol{\mu}_k, \mathbf{s}_k\}_k$, that is, a martingale over the graph of V such that only one coordinate ever moves at a time.[48] The bimartingale converges to a measure over V's graph and has a time-zero value of $(\boldsymbol{\mu}_0, \mathbf{s}_0) = (\mu_0, s_0)$, where s_0 is S's payoff in said equilibrium. We then follow the easy direction of Aumann and Hart's (1986) characterization of the bi-span of a set, noting the expectations of continuous biconvex functions of a bimartingale grow over time, and so the function constructed at the beginning of the

[47]To ease notational overhead, we employ Nash equilibrium as our solution concept in studying long cheap talk, and so have no need to define a belief map for the receiver. We therefore obtain a stronger result, because any perfect Bayesian equilibrium is also Bayes Nash.

[48]Although the bimartingale we construct is related to the stochastic process of pairs of R beliefs and S payoffs, the two processes are not the same: Each round of communication corresponds to two periods under the bimartingale. Aumann and Hart (2003) used the same construction.

proof assigns (μ_0, s_0) a weakly negative value. It follows that $(\mu_0, s_0) \neq (\mu_0, s^*)$. Because the chosen long-cheap-talk equilibrium was arbitrary, no such equilibrium can yield S a payoff of s^*.

Other than our construction of a biconvex function, the proof follows the logic presented in Aumann and Hart (2003) and Aumann and Hart (1986). Because both papers assume a finite state space, the results of Aumann and Hart (1986) and Aumann and Hart (2003) do not apply directly. We therefore provide a self-contained proof below.

PROOF OF PROPOSITION 4: Take any $s_* \in \mathbb{R}$ that is not an equilibrium payoff for prior μ_0 in the one-shot cheap-talk game. In particular, $s_* \notin V(\mu_0)$. Focus on the case of $s_* > v^*(\mu_0)$, the mirror-image case being analogous. Fix some payoff $s' \in (v^*(\mu_0), s_*)$. Letting B be the closed convex hull of $v^{-1}[s', \infty)$, Theorem 1 tells us $\mu_0 \notin B$. Hahn–Banach then gives an affine continuous $\varphi : \Delta\Theta \to \mathbb{R}$ such that $\varphi(\mu_0) > \max\varphi(B)$. Now define the function[49]

$$F : \Delta\Theta \times \mathbb{R} \to \mathbb{R}_+,$$
$$(\mu, s) \mapsto \big[\varphi(\mu) - \max\varphi(B)\big]_+\big[s - s'\big]_+.$$

Observe that F is biconvex and continuous. Moreover, $F(\mu, s) = 0$ whenever $s \in V(\mu)$: either $s < s'$ because $\mu \notin B$, or $\mu \in B$ and so $\varphi(\mu) \leq \max\varphi(B)$.

Now consider any Nash equilibrium $(\sigma, (\tilde{\sigma}, \rho))$ of the long-cheap-talk game. Let us define several random variables on the Borel probability space $\langle \Omega, \mathbb{P}_{\sigma,\tilde{\sigma},\rho}\rangle$. For $\omega = ((m_t, \tilde{m}_t)_{t=0}^{\infty}, a, \theta) \in \Omega$, let $\boldsymbol{\theta}(\omega) := \theta$ and $\mathbf{a}(\omega) := a$; and, for $t \in \mathbb{Z}_+$, let $\mathbf{m}_{2t}(\omega) := m_t$ and $\mathbf{m}_{2t+1}(\omega) := \tilde{m}_t$. From these, we define a filtration $(\mathcal{F}_k)_{k\in K}$ with index set $K = \mathbb{Z}_+ \cup \{\infty\}$ by letting each $\mathcal{F}_k$ be the σ-algebra generated by $\{\mathbf{m}_\ell\}_{\ell\in\mathbb{Z}_+,\ell<k}$. Finally, for each $k \in K$, define the ($\Delta\Theta$-valued and $\mathbb{R}$-valued, respectively) random variables $\boldsymbol{\mu}_k := \mathbb{E}[\delta_{\boldsymbol{\theta}}|\mathcal{F}_k]$ and $\mathbf{s}_k := \mathbb{E}[u_S(\mathbf{a})|\mathcal{F}_k]$; and let $P_k \in \Delta(\Delta\Theta \times \mathbb{R})$ denote the distribution of $(\boldsymbol{\mu}_k, \mathbf{s}_k)$. Note that, by construction, P_0 has a distribution $\delta_{(\mu_0,s_0)}$ for some $s_0 \in \mathbb{R}$. Our task is to show $s_0 \neq s_*$.

In what follows, take any statements about the stochastic processes $(\boldsymbol{\mu}_k)_{k\in K}$ and $(\mathbf{s}_k)_{k\in K}$ to hold $\mathbb{P}_{\sigma,\tilde{\sigma},\rho}$-almost surely. By construction, $\boldsymbol{\mu}_{2t+2} = \boldsymbol{\mu}_{2t+1}$ for every $t \in \mathbb{Z}_+$, and both $(\boldsymbol{\mu}_k)_{k\in K}$ and $(\mathbf{s}_k)_{k\in K}$ are martingales. By S rationality, $\mathbf{s}_{2t} = \mathbb{E}[\mathbf{s}_{2t+1}|\mathcal{F}_{2t+1}] = \mathbf{s}_{2t+1}$ for every $t \in \mathbb{Z}_+$. Because F is biconvex and continuous, $\int F\,\mathrm{d}P_0 \leq \int F\,\mathrm{d}P_1 \leq \cdots$. In particular, $\int F\,\mathrm{d}P_k \geq \int F\,\mathrm{d}P_0 = F(\mu_0, s_0)$ for every $k \in \mathbb{Z}_+$. By the martingale convergence theorem, $\mathbf{s}_k$ converges to $\mathbf{s}_\infty$. By the same, every continuous $g : \Theta \to \mathbb{R}$ has $\int_\Theta g\,\mathrm{d}\boldsymbol{\mu}_k$ converging to $\int_\Theta g\,\mathrm{d}\boldsymbol{\mu}_\infty$; so $\boldsymbol{\mu}_k$ converges (weak*) to $\boldsymbol{\mu}_\infty$. But P_k converges (weak*) to P_∞. Therefore, $\int F\,\mathrm{d}P_\infty = \lim_{k\to\infty}\int F\,\mathrm{d}P_k \geq F(\mu_0, s_0)$. By R rationality, $\mathbf{s}_\infty \in V(\boldsymbol{\mu}_\infty)$, implying $F(\boldsymbol{\mu}_\infty, \mathbf{s}_\infty) = 0$, so that $\int F\,\mathrm{d}P_\infty = 0$, too. Therefore, $F(\mu_0, s_0) \leq 0 < F(\mu_0, s_*)$. So $s_0 \neq s_*$, as required. *Q.E.D.*

C.3.2. *Extra Rounds Can Help the Receiver*

Unlike S, R may benefit from long cheap talk when S's preferences are state independent. To see this, consider the following example, which we describe informally. Let $\Theta = \{0, 1\}$; $\mu_0(1) = \frac{1}{8}$; $A = \{\ell, b, t, r\}$; $u_S(b) = 0$, $u_S(\ell) = 1$, $u_S(t) = u_S(r) = 2$; and $u_R(a, \theta) = -(z_a - \theta)^2$, where $z_\ell = 0$, $z_r = 1$, and $z_b = z_t = \frac{1}{2}$. The associated value correspondence V and prior belief μ_0 are depicted in Figure 3.

[49]Recall that $[\cdot]_+ := \max\{\cdot, 0\}$.

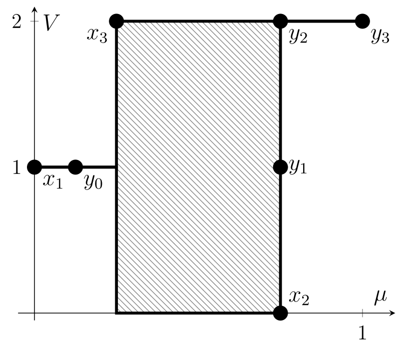


FIGURE 3.—S's value correspondence in an example where R strictly benefits from long cheap talk.

Because every $\mu \in \Delta\Theta$ with $\mu(1) \leq \mu_0(1)$ has $V(\mu) = \{1\}$, Lemma 1 immediately implies every equilibrium outcome (p, s) of the one-shot cheap-talk game has $s = 1$ and $p\{\mu : \mu(1) \leq \frac{3}{4}\} = 1$. In particular, every equilibrium of the long-cheap-talk game generates a "mean outcome" of y_0, as depicted in the figure.

Given the above observations, an equilibrium exists with one round of communication with R beliefs supported on $\{0, \frac{3}{4}\}$, and every other one-shot equilibrium generates less information (in a Blackwell sense) for R; we can depict this equilibrium as generating support $\{x_1, y_1\}$ in the figure. But now, with a jointly controlled lottery, this y_1 can be split in the next round to $\{x_2, y_2\}$.[50] Finally, S can provide additional information in the next round to split y_2 into $\{x_3, y_3\}$. Because action t is optimal for R at belief $\frac{3}{4}$ (i.e., that associated with y_2) but not at belief 1 (i.e., that associated with y_3), this additional information is instrumental to R. Therefore, our equilibrium is strictly better for R than any one-round equilibrium.

Thus, although additional rounds of communication do not change S's equilibrium payoff set, the static and long-cheap-talk models are economically distinct, even under state-independent S preferences.

C.4. *Optimality of Full Revelation*

This section presents formal results discussed in Section 6.4. This section's main result is Proposition 5, which shows two things when v is nowhere quasiconcave: First, full revelation is an S-favorite equilibrium; and second, every S-favorite equilibrium entails full revelation if the state is binary or R's best response is unique for every belief. We also demonstrate, via an example, that nowhere quasiconcavity is insufficient for full revelation to be uniquely S-optimal. The example also illustrates S-unfavorable tie breaking can create a benefit from commitment even when full revelation is both S's favorite equilibrium and S's favorite commitment policy. We conclude the section by discussing conditions under which v is nowhere quasiconcave. In particular, we show a strictly quasiconvex v is nowhere quasiconcave if and only if it is nowhere quasiconcave on each of the simplex's one-dimensional extreme subsets (Corollary 7).

[50]Informally, following Aumann and Hart (2003), each player could toss a fair coin (independent of the state for S) and announce its outcome. Then, the players move to x_2 if the coins come up the same, and y_2 otherwise. Such jointly controlled randomization could be done simultaneously with the information that S initially conveys, so that our three-round example can be converted into a slightly more complicated two-round example.

The next few lemmas serve as preliminary steps toward Proposition 5. Lemma 10 provides a way of constructing a measurable correspondence. Using this lemma, we show every non-full revelation commitment policy can be improved upon when v is nowhere concave, by splitting non-extreme beliefs. Similarly, one can split such beliefs to weakly increase a policy's secured value whenever v is nowhere quasiconcave (Lemma 11). An immediate consequence is that under nowhere quasiconcavity, full revelation secures S's highest equilibrium value (Lemma 12). Nowhere quasiconcavity also implies S can do better than no information at every non-extreme belief (Lemma 13). We then combine these lemmas with the observation that the payoff secured by full revelation depends only on the prior's support to show full revelation barely secures S's highest equilibrium payoff.

We now proceed with proving Lemma 10. This lemma is based on Aliprantis and Border's (2006) discussion concerning measurability of correspondences. All measurability statements are made with respect to the appropriate Borel σ-algebras.

LEMMA 10: *Let X and Y be compact metrizable spaces, $\Xi : X \to \mathbb{R}$ upper semicontinuous, and $\Upsilon : Y \to \mathbb{R}$ measurable. Then,*

$$\Gamma : Y \rightrightarrows X, \quad y \mapsto \Xi^{-1}[\Upsilon(y), \infty),$$

is weakly measurable.

PROOF: Recall that a nonempty-compact-valued correspondence into X is weakly measurable if and only if it is measurable when viewed as a $\mathcal{K}_X$-valued function (Theorem 18.10 from Aliprantis and Border (2006)).[51] We now proceed with proving the lemma. To begin, let $\bar{z} = \max \Xi(X)$, and observe that

$$\Lambda : (-\infty, \bar{z}] \rightrightarrows X, \quad z \mapsto \Xi^{-1}[z, \infty\big) = \{\Xi \geq z\},$$

is nonempty-compact-valued because Ξ is upper semicontinuous. We claim below that Ξ is weakly measurable. It follows that $y \mapsto \Lambda \circ \Upsilon(y)$ is a measurable function from Y into $\mathcal{K}_X$, and so is weakly measurable when viewed as a correspondence. Noting $\Gamma = \Lambda \circ \Upsilon$ completes the proof.

We now argue Ξ is weakly measurable. To do so, consider any open $G \subseteq X$. The lower inverse image of G under Λ is

$$\begin{aligned} \Lambda^l(G) &= \big\{z \leq \bar{z} : \Lambda(z) \cap G \neq \emptyset\big\} \\ &= \big\{z \leq \bar{z} : \{\Xi \geq z\} \cap G \neq \emptyset\big\} \\ &= \big\{z \leq \bar{z} : \Xi(G) \nsubseteq (-\infty, z)\big\}, \end{aligned}$$

which is an interval. *Q.E.D.*

When v is nowhere (quasi)concave, Lemma 10 gives a splitting of each non-extreme belief that increases v's expected (secured) value. We present this result below.

[51] $\mathcal{K}_X$ denotes all nonempty compact subsets of X, equipped with the Hausdorff metric.

LEMMA 11: *Suppose v is nowhere (quasi)concave. Then, a measurable selector r of $\mathcal{I} : \Delta\Theta \rightrightarrows \Delta\Delta\Theta$ exists such that $\int v \, \mathrm{d}r(\mu) > v(\mu)$ $(\inf v(\operatorname{supp} r(\mu)) > v(\mu))$ for all $\mu \in \Delta\Theta \backslash \{\delta_\theta\}_{\theta\in\Theta}$.*

PROOF: Observe that $\hat{v}(\cdot)$ $(\bar{v}(\cdot))$ is upper semicontinuous and therefore measurable. Moreover, $p \mapsto \int v \, \mathrm{d}p$ $(p \mapsto \inf v(\operatorname{supp} p))$ is an upper semicontinuous function from $\Delta\Delta\Theta$ to $\mathbb{R}$. Therefore, Lemma 10 implies $\mu \mapsto \{p \in \Delta\Delta\Theta : \int v \, \mathrm{d}p \geq \hat{v}(\mu)\}$ $(\mu \mapsto \{p \in \Delta\Delta\Theta : \inf v(\operatorname{supp} p) \geq \bar{v}(\mu)\})$ is weakly measurable. Noting $\mathcal{I}$ is also weakly measurable (by upper hemicontinuity) implies

$$\mu \mapsto \mathcal{I}(\mu) \cap \left\{ p \in \Delta\Delta\Theta : \int v \, \mathrm{d}p \geq \hat{v}(\mu) \right\}$$

$$\big(\mu \mapsto \mathcal{I}(\mu) \cap \big\{ p \in \Delta\Delta\Theta : \inf v(\operatorname{supp} p) \geq \bar{v}(\mu) \big\}\big)$$

is weakly measurable. Because the latter correspondence is nonempty-valued, it admits a measurable selector, r, by the Kuratowski and Ryll–Nardzewski selection theorem (Theorem 18.13 from Aliprantis and Border (2006)). The result follows from noting $\hat{v}(\mu) > v(\mu)$ $(\bar{v}(\mu) > v(\mu))$ holds for all $\mu \in \Delta\Theta \backslash \{\delta_\theta\}_{\theta\in\Theta}$ whenever v is nowhere (quasi)concave (appealing to Corollary 1). *Q.E.D.*

Lemma 11 above immediately implies full revelation is S's uniquely optimal commitment protocol whenever v is nowhere concave. The reason is that any other information policy can be strictly improved upon via the splitting generated by the lemma. Lemma 11 also implies that when v is nowhere quasiconcave, full revelation secures S's maximal equilibrium. We prove the latter result in the lemma below.

LEMMA 12: *If v is nowhere quasiconcave, $\bar{v}(\mu) = \inf_{\theta\in\operatorname{supp}(\mu)} v(\delta_\theta)$ for all $\mu \in \Delta\Theta$; that is, full information secures S's maximal equilibrium value.*

PROOF: Fix $\mu \in \Delta\Theta$. A unique $p^F \in \mathcal{I}(\mu)$ exists with $p^F\{\delta_\theta\}_{\theta\in\Theta} = 1$; clearly, p^F has support $\{\delta_\theta\}_{\theta\in\operatorname{supp}(\mu)}$. By Corollary 1, we know $\bar{v}(\mu)$ is the highest securable value at prior μ. Thus, letting $\mathcal{P} := \{p \in \mathcal{I}(\mu) : p \text{ secures } \bar{v}(\mu)\}$, our aim is to show $p^F \in \mathcal{P}$. Corollary 1 tells us $\mathcal{P}$ is nonempty, and upper semicontinuity of v implies $\mathcal{P}$ is closed. The mean-preserving spread order being closed-continuous, $\mathcal{P}$ contains some maximal element, p, with respect to this order. Letting r be as delivered by Lemma 11, the policy $\int r \, \mathrm{d}p$ belongs to $\mathcal{P}$ as well.[52] But maximality of p requires that $p = \int r \, \mathrm{d}p$, implying $p = p^F$. *Q.E.D.*

The next lemma establishes that under nowhere quasiconcavity, S can always benefit from cheap talk.

LEMMA 13: *If v is nowhere quasiconcave, $\bar{v}(\mu) > v(\mu)$ for all $\mu \in \Delta\Theta \backslash \{\delta_\theta\}_{\theta\in\Theta}$.*

PROOF: Fix any $\mu \in \Delta\Theta \backslash \{\delta_\theta\}_{\theta\in\Theta}$. By hypothesis, $\mu', \mu'' \in \Delta\Theta$ and $\lambda \in (0, 1)$ exist such that $\mu = \lambda\mu' + (1 - \lambda)\mu''$ and $v(\mu) < v(\mu'), v(\mu'')$. Therefore, $p = \lambda\delta_{\mu'} + (1 - \lambda)\delta_{\mu''} \in \mathcal{I}(\mu)$ secures a value strictly above $v(\mu)$, and so $\bar{v}(\mu) > v(\mu)$ by Theorem 1. *Q.E.D.*

We now prove our main result regarding nowhere quasiconcavity.

[52]Here, $\int r \, \mathrm{d}p \in \mathcal{I}(\mu)$ is given by $[\int r \, \mathrm{d}p](D) := \int r(D|\cdot) \, \mathrm{d}p$ for Borel $D \subseteq \Delta\Theta$.

PROPOSITION 5: *Suppose v is nowhere quasiconcave. Then,*

1. *Some S-preferred equilibrium entails full information.*
2. *If Θ is binary, or if R has a unique best response to every belief, every S-preferred equilibrium entails full information.*

PROOF: We begin by showing full revelation barely secures $\bar{v}(\mu_0)$. Fix some $\theta \in \operatorname{supp}\mu_0$. Consider any $\mu \in \operatorname{co}\{\delta_\theta, \mu_0\}\backslash\{\delta_\theta\}$. We argue $\bar{v}(\mu_0) > v(\mu)$, and so $v^{-1}[\bar{v}(\mu_0), \infty) \cap \operatorname{co}\{\delta_\theta, \mu_0\} = \{\delta_\theta\}$, as required. Because the support of μ and μ_0 is the same, full revelation secures the same value for both beliefs. Therefore, Lemma 12 and Lemma 13 yield

$$v(\mu) < \bar{v}(\mu) = \inf\sup v\big(\{\delta_\theta\}_{\theta\in\operatorname{supp}\mu_0}\big) = \bar{v}(\mu_0).$$

In other words, full revelation barely secures $\bar{v}(\mu_0)$. The securability theorem (more precisely, Lemma 4) then delivers the first point.

To show the second part, we claim below $\bar{v}(\mu) \leq \bar{v}(\mu_0)$ for each $\mu \in \Delta\Theta\backslash\{\delta_\theta\}_{\theta\in\Theta}$. Lemma 13 then implies $v(\mu) < \bar{v}(\mu) \leq \bar{v}(\mu_0)$ for all $\mu \in \Delta\Theta\backslash\{\delta_\theta\}_{\theta\in\Theta}$. As such, $p \in \mathcal{I}(\mu_0)$ secures $\bar{v}(\mu_0)$ only if $\operatorname{supp} p \subseteq \{\delta_\theta\}_{\theta\in\Theta}$, that is, p provides full information. To conclude the proof, we note $(p, \bar{v}(\mu_0))$ is an equilibrium outcome only if p secures $\bar{v}(\mu_0)$, meaning no p other than full revelation can yield S a payoff of $\bar{v}(\mu_0)$.

All that remains is to show $\bar{v}(\mu) \leq \bar{v}(\mu_0)$ for all $\mu \in \Delta\Theta\backslash\{\delta_\theta\}_{\theta\in\Theta}$. When $|\Theta| = 2$, this inequality holds with equality by Lemma 12. If R's best response is unique, v is continuous, and so every $\theta \in \Theta$ has

$$v(\delta_\theta) = \lim_{n\to\infty} v\left(\frac{n-1}{n}\delta_\theta + \frac{1}{n}\mu_0\right) \leq \lim_{n\to\infty} \bar{v}\left(\frac{n-1}{n}\delta_\theta + \frac{1}{n}\mu_0\right) = \bar{v}(\mu_0),$$

where the last equality follows from Lemma 12. The same lemma then implies $\bar{v}(\mu) = \inf v(\{\delta_\theta\}_{\theta\in\operatorname{supp}\mu}) \leq \bar{v}(\mu_0)$, as required. *Q.E.D.*

We now provide an example that witnesses two properties. First, it shows nowhere quasiconcavity alone is insufficient for uniqueness of full revelation as an S-favorite equilibrium. Second, it is possible for S to benefit from commitment despite full revelation being best for S both with and without commitment.

EXAMPLE 4: Let $\Theta := \{-1, 0, 1\}$, $A := \{0, 1\} \times \Delta\Theta$, $\mu^* := \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, $\mu_0 := \frac{1}{2}\delta_0 + \frac{1}{2}\mu^*$, and $H : \Delta\Theta \to \mathbb{R}_+$ a continuous and strictly concave function with $H(\delta_\theta) = 0\ \forall\theta \in \Theta$. Let players utilities $u_S : A \to \mathbb{R}$ and $u_R : A \times \Theta \to \mathbb{R}$ be given by

$$u_S\big((x, \hat{\mu})\big) := xH\big(\mu^*\big) - H(\hat{\mu})$$

and

$$u_R\big((x, \hat{\mu}), \theta\big) := -\sum_{\tilde{\theta}\in\Theta}\big[\hat{\mu}(\tilde{\theta}) - \mathbf{1}_{\tilde{\theta}=\theta}\big]^2 - x\big(1 - \theta^2\big).$$

Observe $(x, \hat{\mu})$ is a best response to R belief μ if and only if $\hat{\mu} = \mu$ and $x\mu(0) = 0$. Therefore, the value function is given by $v(\mu) = H(\mu^*)\mathbf{1}_{\mu(0)=0} - H(\mu)$. By construction, this function is strictly quasiconvex because $-H$ is. Appealing to Corollary 7 (see below), the value function is then nowhere quasiconcave, and so full information is an S-preferred equilibrium, yielding S payoff $\min\{H(\mu^*), 0\} = 0$.

Observe that, in an S-preferred equilibrium, R breaks indifferences against S when the state is nonzero. Therefore, S gets a payoff strictly lower than her commitment value of $\frac{1}{2}H(\mu^*)$. Moreover, full information is not the only S-preferred equilibrium information policy, because Lemma 1 implies $(\frac{1}{2}\delta_{\delta_0} + \frac{1}{2}\delta_{\mu^*}, 0)$ is an equilibrium outcome.

We conclude this section with sufficient conditions for v to be nowhere quasiconcave. In particular, we show a strictly quasiconvex v is nowhere quasiconcave if and only if it is nowhere quasiconcave on every one-dimensional extreme subset of $\Delta\Theta$.

COROLLARY 7: *Let v be strictly quasiconvex. The following are equivalent*:
(i) *v is nowhere quasiconcave*.
(ii) *$v|_{\Delta\{\theta,\theta'\}}$ is nowhere quasiconcave for every $\theta, \theta' \in \Theta$.*

PROOF: Clearly, (i) implies (ii). That (ii) implies (i) follows from applying Corollary 6 with $T(\theta) := \delta_\theta$. Indeed, for any prior $\mu \in \Delta\Theta$ with $|\operatorname{supp}\mu| \geq 3$, Corollary 6's proof delivers a pair of beliefs μ', μ'' with μ as their midpoint such that $v(\mu) < v(\mu'), v(\mu'')$. Therefore, the definition of nowhere quasiconcavity need only be verified at binary-support beliefs whenever v is strictly quasiconvex. *Q.E.D.*

REFERENCES

ALIPRANTIS, C. D., AND K. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Science & Business Media. [10,11]
AUMANN, R. J., AND S. HART (1986): "Bi-Convexity and bi-Martingales," *Israel Journal of Mathematics*, 54 (2), 159–180. [7,8]
——— (2003): "Long Cheap Talk," *Econometrica*, 71 (6), 1619–1660. [7-9]
BERGEMANN, D., AND S. MORRIS (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11 (2), 487–522. [4]
BESTER, H., AND R. STRAUSZ (2001): "Contracting With Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica*, 69 (4), 1077–1098. [4,5]
CHAKRABORTY, A., AND R. HARBAUGH (2010): "Persuasion by Cheap Talk," *American Economic Review*, 100 (5), 2361–2382. [1-3]
CRAWFORD, V. P., AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50 (6), 1431–1451. [3]
KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615. [4]
MYERSON, R. (1986): "Multistage Games With Communication," *Econometrica*, 54 (2), 323–358. [4]

Co-editor Joel Sobel handled this manuscript.

Manuscript received 11 September, 2017; final version accepted 23 March, 2020; available online 30 March, 2020.