Persuasion by Cheap Talk - Online Appendix

By Archishman Chakraborty and Rick Harbaugh

Online appendix to "Persuasion by Cheap Talk," American Economic Review

Our results in the main text concern the case where the expert's preferences $U(a)$ over the decision maker's actions are common knowledge and so independent of the state θ. But often the decision maker might believe that he has a good sense of the expert's preferences, and yet still face some uncertainty over the exact preferences. In this online Appendix we extend our results to allow the decision maker to face some limited uncertainty about the expert's preferences, including the possibility that preferences are correlated with the state θ.

Suppose that the expert's preferences are given by a function $U(a, t)$, continuous in a, where $t \in \mathbf{T}$ is the type of the expert. The expert knows her own type t (in addition to θ), but the decision maker only has a prior Φ on t. Let $F(\theta|t)$ summarize the conditional distribution of θ given $t \in \mathbf{T}$. This approach allows t to be independent of θ or to be correlated with θ. It also covers the Crawford-Sobel model where $t = \theta$. More importantly, it allows us to conceptually separate uncertainty about the expert's motives t from that about the decision maker's ideal course of action θ. Notice that the expert's type t is fully specified by the pair $U(., t)$ and $F(.|t)$.

We analyze three extensions in this Appendix which show the robustness of our main results. The first extension allows for arbitrary Φ but supposes that T is small. This captures situations where the decision maker attaches positive probability to only a few possible expert types. We find that if there are only a finite number of different types of expert preferences, then an informative equilibrium exists as long as there are a larger number of dimensions of interest to the decision maker(s).

The second concerns the case where the prior Φ is concentrated on a particular type $t^* \in \mathbf{T} = \{1, ..., T\}$. This captures situations where the decision maker is almost certain that the expert has preferences $U(a, t^*)$, although there are potentially many other unlikely possibilities. We find that when the expert's preferences are sufficiently likely to be from one type, then under mild regularity conditions an equilibrium exists in which the decision maker listens to the expert and adjusts his use of the expert's information to reflect the fact that low probability types always recommend the same action.

The third extension relaxes the equilibrium notion from cheap talk to epsilon cheap talk and focuses on the case where the expert's type space is rich, i.e., $t = \theta$, but her type has a limited effect on (differences in) her utility. In particular we consider Euclidean preferences which are state-dependent but that converge uniformly to state-independent preferences as the bias in each dimension increases. We show that any incentive to deviate from the cheap talk equilibria of the limiting preferences goes to zero.

A. More Issues Than Motives

We first consider the case where there is a larger number of issues N relative to the number of types T. For instance, suppose a car magazine is biased toward one of several car manufacturers and the reader is unsure of which manufacturer is favored, so that there are several types of expert preferences T. If the magazine has information on multiple models for some manufacturers and/or on multiple attributes of some models, then the dimensionality of information N is larger than T and it would seem that comparative statements about different models or about the relative strengths of particular models should be credible. Application of the Borsuk-Ulam theorem confirms this intuition quite generally.

PROPOSITION 5: *Suppose $N > T$. Then an influential cheap talk equilibrium exists for all* **T** *and* Φ.

Proof: The arguments are identical to that for Theorem 1. The only difference is that now we look for an s that simultaneously sets the $T \leq N - 1$ maps $\Delta(s, c, t) = U(a^-(h_{s,c}), t) - U(a^+(h_{s,c}), t) = 0$, one for each $t \in$ **T**. ∎

When $N > T$, it is possible to find a two-message partition of Θ that is an equilibrium for every type $t \in$ **T**. In other words, it is possible to find an informative communication strategy that induces actions in $\mathbb{R}^N$ that the T possible types of the expert agree on. This is not surprising for linear $U(., t)$ since satisfying all the T experts simultaneously still leaves $N - 1 - T$ degrees of freedom. Proposition 5 shows that it suffices simply to count equations and unknowns in general for all continuous preferences. Hence even in an environment where an expert's motives are unclear, an expert can often gain credibility via knowledge of a large number of decision-relevant issues. In the next section we consider the case where N is instead small relative to T but only one type is very likely.

B.   Almost Certain Motives

Now suppose that the prior $\Phi = (\phi_1, ..., \phi_T)$ is close to the degenerate distribution Φ^* on type $t^* \in$ **T** (i.e., $\phi_{t^*}^* = 1$). We use the implicit function theorem to look for equilibria in the neighborhood of the equilibria identified by Theorem 1 for the degenerate case, for general preferences and conditional distributions of θ. In such equilibria the decision maker anticipates that the low probability types will not be indifferent so they will always offer the same advice. For instance, the decision maker thinks that the expert is probably unbiased across dimensions, $U = a_1 + a_2$, but there is some chance that the expert has a relatively extreme slant, $U = 4a_1 + a_2$, in which case she will offer the same advice regardless of the state θ.

To apply the implicit function theorem we assume that $U(a, t)$ is continuously differentiable in a for each $t \in$ **T** and, for simplicity, consider only the case $N = 2$. We also suppose that the types $t \in$ **T** have different preferences from each other in the following sense: for any two action profiles a, a' with $a \neq a'$, if $U(a, t^*) = U(a', t^*)$, then $U(a, t) \neq U(a', t)$ for all $t \neq t^*$, $t \in$ **T**. We call this condition (S). Notice that it will hold if, for instance, the indifference curves of the different types satisfy a single-crossing property in $\mathbb{R}^2$.[1]

PROPOSITION 6: *Suppose U satisfies (S) and $N = 2$. Generically in $U(., t^*)$, $F(.|t^*)$, there exists $\varepsilon > 0$ such that for each Φ with $||\Phi - \Phi^*|| < \varepsilon$ an influential cheap talk equilibrium exists.*

Proof: For any fixed $c \in int(\Theta)$, let s^* be the orientation of an equilibrium hyperplane through c with corresponding actions $a^+(h_{s^*,c})$ and $a^-(h_{s^*,c})$ when priors are degenerate on t^*, i.e., given by Φ^*. Such s^* exists by Theorem 1 and we have $U(a^+(h_{s^*,c}), t^*) = U(a^-(h_{s^*,c}), t^*)$. By condition (S), for all $t \neq t^*$, $U(a^+, t) \neq U(a^-, t)$. Wlog, rename types so that $U(a^+, t) > U(a^-, t)$ for all $t > t^*$ and $U(a^+, t) < U(a^-, t)$ for all $t < t^*$ and consider the actual priors Φ. Pick a hyperplane of arbitrary orientation s through c (the same c as above) and, using usual notation, let the actions or expected values corresponding to each message be

$$\alpha^+(h_{s,c}; \Phi) = \Pr[t^*|m^+]E[\theta|t^*, m^+] + \sum_{t > t^*} \Pr[t|m^+]E[\theta|t] \tag{A1}$$

$$\alpha^-(h_{s,c}; \Phi) = \Pr[t^*|m^-]E[\theta|t^*, m^-] + \sum_{t < t^*} \Pr[t|m^-]E[\theta|t] \tag{A2}$$

---

[1]Proposition 5 provides insight on how the following result can be extended to the case where $N > 2$. In brief, one applies the implicit function theorem to the equilibrium of Proposition 5 constructed with $N - 1$ types (including t^*) disclosing a partition and the remaining types sending only one message. Condition (S) then has to be suitably amended.

That is, we assume type t^* discloses the partition of Θ associated with $h_{s,c}$ truthfully, while all types $t > t^*$ (resp., $t < t^*$) send only message m^+ (resp., m^-). Thus,

$$\text{(A3)} \qquad \Pr[t^*|m^+] = \frac{\Pr[\theta \in m^+|t^*]\phi_{t^*}}{\Pr[\theta \in m^+|t^*]\phi_{t^*} + \sum_{t > t^*} \phi_t}$$

and

$$\text{(A4)} \qquad \Pr[t|m^+] = \frac{\phi_t}{\Pr[\theta \in m^+|t^*]\phi_{t^*} + \sum_{t < t^*} \phi_t}$$

if $t > t^*$ and is 0 otherwise, and similarly for the message m^-. Let

$$\text{(A5)} \qquad \Delta(s, \Phi, t) = U(\alpha^+(h_{s,c}; \Phi), t) - U(\alpha^-(h_{s,c}; \Phi), t)$$

which is a continuously differentiable function of s. We know that $\Delta(s^*, \Phi^*, t^*) = 0$. We wish to show via the implicit function theorem that there exists $\varepsilon > 0$, such that for $||\Phi - \Phi^*|| < \varepsilon$, there exists $s(\Phi)$ close to s^* for which $\Delta(s(\Phi); \Phi, t^*) = 0$. This is enough to show the result, since when $s(\phi)$ is close to s^* the corresponding actions $\alpha^+(s(\Phi); \phi)$ and $\alpha^-(s(\phi); \phi)$ are close to a^+ and a^- respectively, so that $\Delta(s^*, \Phi^*, t) < 0$ if $t < t^*$ and $\Delta(s^*, \Phi^*, t) > 0$ if $t > t^*$. Then type t^* has the right incentives to disclose the partition of Θ associated with $h_{s(\Phi),c}$ truthfully, while all types $t > t^*$ (resp., $t < t^*$) send only message m^+ (resp., m^-).

We use the fact that the circle is locally like the line. That is, we set $s_1(z) = z \in [-1, 1]$ and, when $s_2^* = \sqrt{1 - s_1^{*2}}$ set $s_2(z) = \sqrt{1 - z^2}$ (and, similarly, when $s_2^* = -\sqrt{1 - s_1^{*2}}$ set $s_2(z) = -\sqrt{1 - z^2}$). We then consider the function $\Delta(s(z), \Phi, t^*)$ as a function of z in a neighborhood of s^*. To apply the implicit function theorem we have to show that $\partial\Delta(s(z), \Phi^*, t^*)/\partial z \neq 0$. It is easy to see that this derivative consists of terms involving the derivative of $U(., t^*)$ with respect to the actions (that do not depend on $F(.|t^*)$) and terms involving the expected actions α^+ and α^- (that depend on $F(.|t^*)$ but not on $U(., t^*)$). Since we can vary $F(.|t^*)$ and $U(., t^*)$ independently, $\frac{\partial\Delta(s(z), \Phi^*, t^*)}{\partial z}$ can vanish only in non-generic cases, establishing the result. ∎

In the influential equilibrium of Proposition 6, type t^* discloses a two-message partition of Θ similar to the equilibria of Theorem 1. By condition (S), no other type can be indifferent between two induced actions and so will send one of the two messages with probability one. Since Φ is close to Φ^*, the induced actions (and the equilibrium partition) are close to the equilibrium of the case where the expert's likely type t^* is common knowledge. The decision maker essentially ignores the implications of messages from unlikely types $t \neq t^*$ in determining his action.[2]

C. Epsilon Cheap Talk

For our third robustness test we consider preferences that are highly state-dependent, but that converge to state-independent preferences. As a notable example of such preferences, we consider Euclidean preferences where the expert's utility is based on the distance between the ex-

---

[2]Since unlikely types always influence the decision maker in their preferred direction, this provides a multidimensional perspective on the finding by Stefano DellaVigna and Ethan Kaplan (2007) that, in an environment where few viewers expected a news network to have a conservative bias, Fox News had a large influence on voting behavior by viewers.

pert's ideal action and the decision maker's ideal action,

$$(A6) \qquad U(a; \theta) = -d(a, \theta + b) = -\left(\sum_{i=1}^{N} (a_i - (\theta_i + b_i))^2 \right)^{1/2}$$

where $d(\cdot, \cdot)$ is the Euclidean distance function and $b = (b_1, ..., b_N) \in \mathbb{R}^N$ is the vector of known biases representing the distance between the decision maker's ideal action θ, and the expert's ideal action $\theta + b$. Distance preferences are used in the leading example from Crawford and Sobel and in a wide variety of applications.

As the expert's bias in each dimension increases, Euclidean preferences converge uniformly to state-independent linear preferences with known biases across dimensions equal to the ratios of these biases within dimensions. More precisely, if we write $b = \rho B$ for some vector $\rho \in \mathbb{R}^N$, $\rho \neq 0$, and real number $B \geq 0$, then for any θ, as B increases without bound the expert's ideal point $\theta + b$ becomes more and more distant from θ, and the circular indifference curves for Euclidean preferences become straighter, and converge to those of known linear preferences of the form $\rho \cdot a$ given by (1).[3]

To use this convergence, we modify the game so that the expert's payoff from any action a and message m given θ is $U(a, m; \theta) = -d(a, \theta + b)$ less an arbitrarily small cost $\varepsilon > 0$ of lying if the message m is not consistent with θ.[4] We study if influential equilibria exist in the modified game with distance preferences (A6) and an arbitrarily small cost of lying. We say that a communication strategy is an ε-*cheap talk equilibrium for large biases* of the game with distance preferences if and only if for each $\varepsilon > 0$ there exists $\overline{B}$ such that for all $B > \overline{B}$ and any θ, the incentive to lie for an expert is at most ε. Our next result shows an equivalence between such equilibria and the cheap talk equilibria for linear preferences characterized by Theorem 3.[5, 6]

PROPOSITION 7: *Suppose U is Euclidean. Then for all F and all k, a communication strategy is a k-message ε–cheap talk equilibrium for large biases if and only if it is a cheap talk equilibrium for the limiting linear U with slant ρ.*

Proof: We show that for all F and all k, a communication strategy is a k-message ε–cheap talk equilibrium for Euclidean U with large B if and only if it is a cheap talk equilibrium for the limiting linear preferences $\rho \cdot a$. To do this first pick an arbitrary hyperplane $h_{s,c}$ of orientation $s \in \mathbb{S}^{N-1}$ passing through $c \in int(\Theta)$ and let L be the line joining the corresponding actions $a^+ = a^+(h_{s,c})$ and $a^- = a^-(h_{s,c})$. Pick any $\theta \in \Theta$ and let $p(\theta)$ be the point where the

---

[3]All arguments go through unchanged if each b_i is of the form $b_i = \kappa_i + \rho_i B$, for constants κ_i. In such cases, as B becomes large the expert becomes infinitely biased in dimensions where $\rho_i \neq 0$ but has finite (possibly, no) bias in dimensions where $\rho_i = 0$.

[4]Since the meaning of a message is derived from a (candidate) equilibrium communication strategy, the notion of what constitutes a "lie" is endogenous here. Therefore this equilibrium notion is distinct from that of an "almost cheap talk" equilibrium (Navin Kartik, 2008) or a "costly talk" equilibrium (Navin Kartik, Marco Ottaviani, and Francesco Squintani, 2007) in which the sender's reports have an exogenous meaning corresponding to the true value of the state and any deviation from this true value is costly. Our notion corresponds to that of an ε-equilibrium of a cheap talk game.

[5]In an extension of their analysis of lexicographic preferences, Levy and Razin (2007) show for Euclidean preferences that even slight asymmetries in distributions can preclude pure cheap talk for sufficiently large but finite B. Our result shows that for sufficiently large B any incentive to deceive the decision maker is arbitrarily close to 0.

[6]Epsilon equilibria are not invariant to monotonic transformations of the underlying preferences. For instance, with a quadratic variant of the Euclidean specification which drops the square root term in (A6), the difference in the sender's utilities from actions a and a' is unbounded in B, implying that our equivalence result obtains only if the cost of lying also increases in the unit of payoffs B, e.g., if it is equal to εB for any $\varepsilon > 0$. Note however that indifference curves corresponding to such quadratic preferences are also linear *at* the limit of infinite biases.

perpendicular from $\theta + B\rho$ on to L meets L. Then

$$(A7) \qquad\qquad p(\theta) = q(\theta)a^+ + (1 - q(\theta))a^-$$

where $q(\theta) \in \mathbb{R}$ is given by

$$(A8) \qquad\qquad q(\theta) = \frac{(\theta - a^-) \cdot (a^+ - a^-) + B\rho \cdot (a^+ - a^-)}{(a^+ - a^-) \cdot (a^+ - a^-)}.$$

Notice that this is well-defined since $a^+ \neq a^-$. Notice next that

$$
\begin{aligned}
d(a^+, \theta + b) - d(a^-, \theta + b) &= \frac{d^2(a^+, \theta + b) - d^2(a^-, \theta + b)}{d(a^+, \theta + b) + d(a^-, \theta + b)} \qquad (A9)\\
&= \frac{d^2(a^+, p(\theta)) - d^2(a^-, p(\theta))}{d(a^+, \theta + b) + d(a^-, \theta + b)}.
\end{aligned}
$$

For the if part consider first a two-message cheap talk equilibrium with induced actions a^+ and a^- when U is given by (1), so that $\rho \cdot (a^+ - a^-) = 0$. Then $q(\theta)$ and so $p(\theta)$ do not depend on B. Furthermore, using (A9),

$$
\begin{aligned}
\left|U(a^+; \theta + b) - U(a^-; \theta + b)\right| &= \left|d(a^+, \theta + b) - d(a^-, \theta + b)\right| \qquad (A10)\\
&= \left|\frac{d^2(a^+, p(\theta)) - d^2(a^-, p(\theta))}{d(a^+, \theta + b) + d(a^-, \theta + b)}\right|\\
&\leq \max_{\theta \in \Theta} \left|\frac{d^2(a^+, p(\theta)) - d^2(a^-, p(\theta))}{d(a^+, \theta + B\rho) + d(a^-, \theta + B\rho)}\right|.
\end{aligned}
$$

Let θ_B be the solution to the last maximization problem. As B rises, θ_B stays bounded in the compact set Θ, so that $p(\theta_B)$ stays bounded as well, implying that the numerator stays bounded. However the denominator becomes arbitrarily large. It follows that for any $\varepsilon > 0$, for B large enough, $\left|U(a^+, \theta + b) - U(a^-, \theta + b)\right| < \varepsilon$ for all θ. An analogous argument obtains for the k-message equilibria if we consider pairs of equilibrium actions that must all lie on the same line L and use the logic above.

For the only if part, suppose that two actions a^+ and a^- do not constitute a two-message cheap talk equilibrium when U is given by (1). Wlog, suppose that $\rho \cdot a^+ < \rho \cdot a^-$. Consider type $\theta = a^+$ and observe via (A9) that

$$
\begin{aligned}
\lim_{B \to \infty} \left[U(a^+; a^+ + b) - U(a^-; a^+ + b)\right] &= \lim_{B \to \infty} \left[\frac{d^2(a^+, p(a^+)) - d^2(a^-, p(a^+))}{d(a^+, a^+ + B\rho) + d(a^-, a^+ + B\rho)}\right]\\
&= \frac{\rho \cdot (a^- - a^+)}{\sqrt{\rho \cdot \rho}} \qquad (A11)
\end{aligned}
$$

where we have used (A7) and (A8) in the last line. It follows that when $\varepsilon < \rho \cdot (a^- - a^+)/\sqrt{\rho \cdot \rho}$, and B is large enough, type $\theta = a^+$ would gain by more than ε from lying (i.e., by inducing the decision maker to choose the action a^- instead of a^+), implying in turn that h is not an ε-cheap talk equilibrium for large B when U is given by (A6). An identical argument obtains for the k-message equilibria of Theorem 3, $k \geq 2$ and finite, if we consider some pair of actions for which $\rho \cdot a^+ \neq \rho \cdot a^-$. ∎

This convergence result implies that a strong incentive to exaggerate in each dimension can

be captured by our simple model of linear preferences. Moreover, since the expert's biases or slant across dimensions converges to that of the ratios of the biases within dimensions, the result formally links the idea of an expert being biased towards a higher action as developed in the Crawford and Sobel model, and the idea of an expert being biased across dimensions as emphasized in this paper.