"PERSUASION: THE ART OF CHANGING WORLDVIEWS"

BY SIMONE GALPERTI

B1. Lemma 1

Lemma 1. $V^* = V(\sigma)$. *In (2), it suffices to restrict attention to τs with $|\mathbf{supp}\,\tau| \leq |\Omega|$.*

Proof. For $q \in \Delta(\Omega)$, we have $\hat{v}(q) > -\infty$ by continuity of v and compactness of A. For all $q \in \mathbb{R}^{|\Omega|-1} \setminus \Delta(\Omega)$, define $\hat{v}(q) = -\infty$. By Carathéodory's Theorem (see Rockafellar (1997), Corollary 17.1.5),

$$V(\sigma) = \sup_{\mathcal{T}_\sigma} \sum_i \tau_i \hat{v}(q_i),$$

where

$$\mathcal{T}_\sigma = \left\{ (q_1, \tau_1; \ldots; q_{|\Omega|}, \tau_{|\Omega|}) : \begin{array}{c} \sum_{i=1}^{|\Omega|} \tau_i q_i = \sigma, \sum_{i=1}^{|\Omega|} \tau_i = 1, \tau_i \geq 0, \\ \text{and } q_i \in \Delta(\Omega) \text{ for all } i \end{array} \right\}.$$

Since $\mathcal{T}_\sigma \subseteq \mathcal{T}$, it follows that $V(\sigma) \leq V^*$. By definition of V^*, for every $\gamma > 0$ there exists $\tau_\gamma \in \mathcal{T}$ such that $\mathbb{E}_{\tau_\gamma}[\hat{v}(q)] \geq V^* - \gamma$. However, $\mathbb{E}_{\tau_\gamma}[\hat{v}(q)] \in \{\xi : (\sigma, \xi) \in \text{co(hyp}\,v)\}$, and hence $\mathbb{E}_{\tau_\gamma}[\hat{v}(q)] \leq V(\sigma)$. Thus, for every $\gamma > 0$, $V(\sigma) \geq V^* - \gamma$, which implies that $V(\sigma) \geq V^*$. $\qquad\square$

B2. Lemma 2

Lemma 2. *If $V^c(\sigma) < V(\sigma)$, there exists $\tau \in \mathcal{T}$ such that $\mathbb{E}_\tau[\hat{v}(q)] > V^c(\sigma)$. Moreover, if $\mathbb{E}_\tau[\hat{v}(q)] > V^c(\sigma)$, then the probability of disproving ρ is strictly positive.*

Proof. The first part follows from Lemma 1. For the second part, note that by the same argument as in the proof of Lemma 1,

$$V^c(\sigma) = \sup_{\mathcal{T}_\sigma^c} \sum_i \tau_i \hat{v}(q_i),$$

where

$$\mathcal{T}_\sigma^c = \left\{ (q_1, \tau_1; \ldots; q_N, \tau_N) : \begin{array}{c} N \geq 1, \sum_{i=1}^N \tau_i q_i = \sigma, \sum_{i=1}^N \tau_i = 1, \tau_i \geq 0, \\ \text{and } q_i \in \mathcal{C}_\varepsilon \text{ for all } i \end{array} \right\}.$$

Suppose $\mathbb{E}_\tau[\hat{v}(q)] > V^c(\sigma)$, but the probability of disproving Receiver is zero.

Since $\tau \in \mathcal{T}$, $|\mathbf{supp}\,\tau| = N$ for some finite N, and hence $\mathbf{supp}\,\tau \cap \mathcal{D} = \varnothing$. Therefore, $\tau \in \mathcal{T}_\sigma^c$, and hence $\mathbb{E}_\tau[\hat{v}(q)] \leq V^c(\sigma)$, a contradiction.

$\square$

B3. *Introduction Example with Common Priors*

Consider the example in the Introduction, with the only difference that both Sender and Receiver have prior σ as in Table 1 and Receiver is Bayesian as in KG. Proposition 2 in KG implies that π^{KG} is fully uninformative if $x > 2$ and at least partially informative if $1 < x < 2$. Focusing on the latter case, the following claims establish that the optimal π^{KG} satisfies

$\pi^{KG}(s\|\omega)$	s_T	s_H
h	$\frac{4}{7}$	$\frac{3}{7}$
H	0	1
T	$\frac{8.7}{9}$	$\frac{0.3}{9}$

Claim 7. If π is optimal, then $\pi(s_H|H) = 1$, $\pi(s_H|T) = \frac{0.3}{9}$, and $\pi(s_H|h) = \frac{3}{7}$, where s_H induces a_H.

Proof. We know that we can focus on πs which produce at most three signals.

Suppose no signal induces a_H. Then, in state H, π must always produce some s such that $\frac{\pi(s|\omega)\sigma(\omega)}{\pi(s|H)\sigma(H)} \geq 1$ for $\omega \neq H$. Given this, consider a new π' with the properties that $\pi'(\cdot|\omega) = \pi(\cdot|\omega)$ for all $\omega \neq H$, $\pi'(s|H) = 0$ if $\pi(s|H) > 0$, and $\pi'(s'|H) = 1$ where $\pi(s'|\omega) = 0$ for all other ω. Then, π' strictly dominates π. First, for every $\omega \neq H$, the signal produced by ω under π' induces the same action as the signal produced by ω under π and with the same probabilities as before. This is because, given $\omega \neq H$, if $\frac{\pi(s|\omega)\sigma(\omega)}{\pi(s|\omega')\sigma(\omega')} \geq 1$ for every ω', then by construction $\frac{\pi'(s|\omega)\sigma(\omega)}{\pi'(s|\omega')\sigma(\omega')} \geq 1$ for every ω', since π' removes the probability $\pi(s|H)$. Second, π' induces a_H and payoff 2 with strictly higher probability than does π. This argument also implies that $\pi(s_H|H) < 1$ cannot be optimal.

For the second part, by Proposition 5 in KG, Receiver must be indifferent between a_H and another action when choosing a_H. This means that $p(H|s_H, \pi) = p(\omega'|s_H, \pi)$ for some $\omega' \neq H$. Since $\pi(s_H|H) = 1$, in order to have this, we need

$$\frac{\pi(s_H|\omega')\sigma(\omega')}{\sigma(H)} = 1$$

and hence $\sigma(\omega') \geq \sigma(H)$. Given σ, this is possible for both $\omega' = h$ and $\omega' = T$. Thus, we get that $\pi(s_H|h) = \frac{\sigma(H)}{\sigma(h)} = \frac{3}{7}$ and $\pi(s_H|T) = \frac{\sigma(H)}{\sigma(T)} = \frac{0.3}{9}$. $\square$

Claim 8. If π is optimal, then $\pi(s_T|T) > 0$, where s_T induces a_T.

Proof. We know that $\pi(s_H|H) = 1$, $\pi(s_H|T) < 1$, and $\pi(s_H|h) < 1$. Suppose π does not induce a_T. Without loss of generality, π must then induce at most one signal s' other than s_H given σ. But, in this case, we have

$$\frac{p(h|s',\pi)}{p(T|s',\pi)} = \frac{4/7}{8.7/9} \times \frac{\sigma(h)}{\sigma(T)} = \frac{4/7}{8.7/9} \times \frac{0.07}{0.9} = \frac{0.04}{0.87} < 1.$$

Therefore, s' would induce a_T, a contradiction. $\square$

Claim 9. It π is optimal, then $\pi(s_T|T) = \frac{8.7}{9}$ and $\pi(s_T|h) = \frac{4}{7}$.

Proof. To be optimal, π cannot induce a_h. Suppose that this is not true and let s_h induces a_h. By Proposition 4 in KG, $\pi(s_h|h) > 0$ and $\pi(s_h|\omega) = 0$ for $\omega = T, H$. This already implies that $\pi(s_T|T) = \frac{8.7}{9}$. Given the previous claims, modify π by removing s_h and shifting $\pi(s_h|h)$ to increase $\pi(s_T|h)$. Even if $\pi(s_T|h) = \frac{4}{7}$, given σ, the signal s_T will still induce Receiver to choose a_T, but now with higher ex-ante probability. Receiver's choices for the other signals do not change. Hence, Sender's payoff is strictly higher. $\square$

B4.   *Multiple Alternative Worldviews*

This section extends the baseline model in the main paper by letting Receiver adopt different worldviews after unexpected evidence. It shows that multiple alternative worldviews can help Sender, but can also hurt her. Besides this, the main message of the paper does not change.

Lexicographic Belief Systems. One way to model a Receiver who may adopt alternative worldviews is to borrow Kreps and Wilson's (1982) idea of a "sequence of hypotheses." Imagine that Receiver treats $\rho^0 = \rho_\varepsilon$ as his "primary" worldview, which he always tries to apply first in order to process evidence. If he deems the evidence unexpected under ρ^0, he tries a second worldview, ρ^1, then a third, ρ^2, and so on, stopping at the first in the sequence for which the evidence is no longer unexpected.[29]

To formalize this, we can let each hypothesis ρ^i assign some small probability $\varepsilon^i \geq 0$ to a subset of states and adopt across hypotheses the same definition of unexpected evidence as in Definition 1. We will focus on the case of $\varepsilon^i = 0$.

Assumption 4 (A4: Lexicographic-Belief-System (LBS) Model). *Receiver is described by a finite sequence ρ^0, ρ^1, ..., $\rho^N \in \Delta(\Omega)$ such that* $\mathbf{supp}\,\rho^i \not\subset \mathbf{supp}\,\rho^j$ *if $i > j$, and for each $\omega \in \Omega$, $\rho^i(\omega) > 0$ for some $i \in \{0,\ldots,N\}$. Given (s,π), he computes $p(\cdot|s,\pi)$ by applying Bayes' rule to the ρ^i of lowest index i that satisfies $\sum_{\omega \in \Omega} \pi(s|\omega)\rho^i(\omega) > 0$.*

---

[29]The lexicographic belief systems in Blume et al. (1991*a*; 1991*b*) work in a fundamentally different way: The agent always takes into account, though lexicographically, all the beliefs of his system.

With regard to the feasible distributions over posteriors, a new function $\mathbf{p}_L$ maps Sender's posterior to Receiver's posterior, which generalizes $\mathbf{p}$ in Proposition 1 while maintaining its key properties in Corollary 1. This involves two changes. First, the function $\mathbf{r}$ is replaced by a new function $\mathbf{r}_L$ that, given LBS $(\rho^0, \ldots, \rho^N)$, says which prior Receiver updates for every $q \in \Delta(\Omega)$. By A4, for every (s, π) this is the ρ^i with lowest index that satisfies $\rho^i(\mathbf{supp}\, q(\cdot|s, \pi)) > 0$; therefore, for every $q \in \Delta(\Omega)$,

$$\mathbf{r}_L(q) = \rho^{i(q)}, \text{ where } i(q) = \min\{i : \rho^i(\mathbf{supp}\, q) > 0\}.$$

The second change is in the range of Receiver's feasible posteriors, which shrinks even further: $\mathbf{p}_L(\Delta(\Omega)) \subsetneq \mathbf{p}(\Delta(\Omega))$. To describe this, for $i = 1, \ldots, N$ let $\mathcal{P}^i$ contain all the states which are deemed possible by ρ^i but impossible by all ρ^j of lower index:

$$\mathcal{P}^i = \mathbf{supp}\, \rho^i \setminus (\cup_{j<i}\, \mathbf{supp}\, \rho^j).$$

Also, let $\mathcal{R} = \cup_{i=0}^N \Delta(\mathcal{P}^i)$, where $\mathcal{P}^0 = \mathcal{P}$. Then, under A4, the new function $\mathbf{p}_L$ from Sender's to Receiver's posterior has range $\mathcal{R}$ rather than $\Delta(\mathcal{P}) \cup \Delta(\mathcal{I})$. To see this, note first that if $\mathbf{r}_L(q) = \rho^i$, then $\mathbf{supp}\, q \cap (\cup_{j<i}\, \mathbf{supp}\, \rho^j) = \varnothing$, and hence $\mathbf{supp}\, \mathbf{p}(q) = \mathbf{supp}\, q \cap \mathbf{supp}\, \mathbf{r}_L(q) \subseteq \mathcal{P}^i$. Conversely, for any $i \in \{0, \ldots, N\}$ and $q \in \Delta(\mathcal{P}^i)$, we have $\mathbf{r}_L(q) = \rho^i$, and hence for all $\omega \in \mathcal{P}^i$,

$$\mathbf{p}_L(\omega; q) = q(\omega)\frac{\rho^i(\omega|\mathcal{P}^i)}{\sigma(\omega|\mathcal{P}^i)}\left[\sum_{\omega' \in \mathcal{P}^i} q(\omega')\frac{\rho^i(\omega'|\mathcal{P}^i)}{\sigma(\omega'|\mathcal{P}^i)}\right]^{-1}.$$

Then Proposition 1 applied to priors $\rho^i(\omega|\mathcal{P}^i)$ and $\sigma(\omega|\mathcal{P}^i)$ implies that by varying $q \in \Delta(\mathcal{P}^i)$ we obtain every $p \in \Delta(\mathcal{P}^i)$. The function $\mathbf{p}_L$ continues to have the three key properties highlighted after Proposition 1, which drive the analysis in Section III.

With regard to optimal experiments, consider any LBS $(\rho^0, \ldots, \rho^N)$ with $N \geq 2$. For every $\omega \in \mathcal{I}$, let $i(\omega) = \max\{i : \omega \notin \mathbf{supp}\, \rho^i\}$. Then for every $i \leq i(\omega)$ Sender can conceal ω in any q that satisfies $\mathbf{r}_L(q) = \rho^i$. Since the range of $\mathbf{p}_L$ is $\mathcal{R} = \cup_{i=0}^N \Delta(\mathcal{P}^i)$, for $i = 0, \ldots, N$ let A^i be the set of all actions Receiver could choose if he applied ρ^i; that is, $A^i = \cup_{p \in \Delta(\mathcal{P}^i)}\{\arg\max_{a \in A} \mathbb{E}_p[u(a, \omega)]\}$. Sender's best payoff from concealing each $\omega \in \mathcal{I}$ across *all ways* allowed by Receiver's LBS is then

$$k_*(\omega) = \max_{a \in \cup_{i=0}^{i(\omega)} A^i} v(a, \omega).$$

Her expected payoff from concealment is given by the function $\hat{k}_* : \mathcal{D} \to \mathbb{R}$ that satisfies

$$\hat{k}_*(q) = \mathbb{E}_q[k_*(\omega)], \quad q \in \mathcal{D}.$$

Now define $\hat{m}_* : \mathcal{D} \to \mathbb{R}$ by $\hat{m}_* = \max\{\hat{v}_*, \hat{k}_*\}$, where $\hat{v}_* : \mathcal{D} \to \mathbb{R}$ is the smallest

upper-semicontinuous (u.s.c.) function that satisfies $\hat{v}_*(q) \geq \hat{v}(q)$ for all $q \in \mathcal{D}$.[30] Let V_* and M_* denote the concavifications of $\hat{v}_*$ and $\hat{m}_*$, respectively. The reason for this slight modification of $\hat{v}$ is that, if $\hat{v}$ is u.s.c. over $\mathcal{D}$, then V is continuous over $\mathcal{D}$—by Lemma 3 below—and $V(q)$ is achieved for all $q \in \mathcal{D}$. But $\hat{v}$ can fail to be u.s.c. when $\mathbf{p}_L$ is discontinuous on $\mathcal{D}$. However, we can render $\hat{v}$ u.s.c. over $\mathcal{D}$ using $\hat{v}_*$ without loss of generality (see the next result).

The next result characterizes the payoff from an optimal π that possibly exploits Receiver's multiple worldviews.

Proposition 8. *Sender's expected payoff for $\mathcal{I}$-states—$M(\sigma(\cdot|\mathcal{I}))$ in (5)—satisfies*

$$\text{(B1)} \qquad M(\sigma(\cdot|\mathcal{I})) = M_*(\sigma(\cdot|\mathcal{I})) = \max_{\alpha, q^h, q^s} \{\alpha \hat{k}_*(q^h) + (1-\alpha)V_*(q^s)\},$$

subject to $\alpha \in [0,1]$, $q^h, q^s \in \mathcal{D}$, and $\alpha q^h + (1-\alpha)q^s = \sigma(\cdot|\mathcal{I})$. Given an optimal α^*, Sender conceals states with ex-ante probability $\alpha^* \sigma(\mathcal{I})$.

Proof of Proposition 8. We first need to establish the following lemma.

Lemma 3. *The function $V_* : \mathcal{D} \to \mathbb{R}$ has the following properties:*
(i) For every $q \in \mathcal{D}$, there exists $\tau \in \Delta(\mathcal{D})$ such that $V_*(q) = \mathbb{E}_\tau[\hat{v}_*(q')]$, where $q = \mathbb{E}_\tau[q']$ and $|\mathbf{supp}\,\tau| \leq |\mathcal{I}|$.
(ii) $V \leq V_*$ over $\mathcal{D}$, with equality over $int\,\mathcal{D}$.
(iii) $V_* = \text{cl}V_*$ and hence is continuous.

Proof. **Part (i):** By Corollary 17.1.5 in Rockafellar (1997),

$$V_*(q) = \sup_{\mathcal{T}(q)} \sum_{i=1}^{|\mathcal{I}|} \hat{v}_*(q_i)\tau_i,$$

where

$$\mathcal{T}(q) = \left\{ (q_1, \tau_1; \ldots; q_{|\mathcal{I}|}, \tau_{|\mathcal{I}|}) : \begin{array}{l} \sum_{i=1}^{|\mathcal{I}|} q_i\tau_i = q, \sum_{i=1}^{|\mathcal{I}|} \tau_i = 1, \tau_i \geq 0, \\ \text{and } q_i \in \Delta(\mathcal{I}) \text{ for all } i \end{array} \right\}.$$

Since $\hat{v}_*$ is u.s.c. and $\mathcal{T}(q)$ is compact, by standard arguments $V_*(q)$ is achieved for every $q \in \mathcal{D}$.
Part (ii): Given a function $f : \mathcal{D} \to \mathbb{R}$, let $\text{hyp}f$ be the hypograph of f: $\text{hyp}f = \{(q, \xi) : q \in \mathcal{D}, \xi \in \mathbb{R}, \xi \leq f(q)\}$. Note that $\text{hyp}\,\hat{v}_* = \overline{\text{hyp}\,\hat{v}}$ for $\hat{v}$ restricted to $\mathcal{D}$. Therefore, for all $q \in \mathcal{D}$,

$$V(q) \;=\; \sup\{\xi : (q, \xi) \in \text{co}(\text{hyp}\,\hat{v})\} \leq \sup\{\xi : (q, \xi) \in \text{co}(\overline{\text{hyp}\,\hat{v}})\} = V_*(q).$$

---

[30]Note that $\hat{v}$ is always u.s.c. over $\mathcal{D}$ for LBSs with only one alternative prior as in the baseline model. Also, since $\hat{v}_*$ and $\hat{k}_*$ are u.s.c., so is $\hat{m}_*$. Indeed, a function is u.s.c. if and only if its hypograph is closed (Theorem 7.1, Rockafellar [1997]). In our case, $\text{hyp}\,\hat{m}_* = \text{hyp}\,\hat{k}_* \cup \text{hyp}\,\hat{v}_*$.

Now consider the closure of V over $\mathcal{D}$, cl$V$, which is the unique continuous extension of V to $\mathcal{D}$ by Theorem 10.3 in Rockafellar (1997), is concave, and satisfies cl$V \geq V \geq \hat{v}$ over $\mathcal{D}$. Thus, for every $q \in \mathcal{D}$,

$$\hat{v}_*(q) = \lim_{q' \to q} \sup \hat{v}(q') \leq \lim_{q' \to q} \sup \mathrm{cl}V(q') = \mathrm{cl}V(q).$$

Hence, cl$V$ is a concave function majorizing $\hat{v}_*$. Since V_* is the smallest of such functions, cl$V \geq V_*$. Finally, since cl$V = V$ over $int\,\mathcal{D}$, property (ii) follows.

Part (iii): We already know that $V_* = \mathrm{cl}V_*$ over $int\,\mathcal{D}$. By definition, hyp cl$V_* = \overline{\mathrm{hyp}\,V_*}$. If hyp V_* is closed, then hyp $V_* = \mathrm{hyp}\,\mathrm{cl}V_*$ and we are done. Indeed, by definition, $V_* \leq \mathrm{cl}V_*$, so suppose there exists $q \in \partial\mathcal{D}$ such that $V_*(q) < \mathrm{cl}V_*(q)$. Then there exists $\xi \in \mathbb{R}$ such that $V_*(q) < \xi \leq \mathrm{cl}V_*$, which is a contradiction. Hence, all that remains is to prove that hyp V_* is closed.

First, for every $q \in \mathcal{D}$, by property (i) we have $V_*(q) = \max\{\xi : (q, \xi) \in \mathrm{co}(\mathrm{hyp}\,\hat{v}_*)\}$, and therefore hyp $V_* = \mathrm{co}(\mathrm{hyp}\,\hat{v}_*)$. Second, define $\underline{v}_* = \inf_{q \in \mathcal{D}} \hat{v}_*(q)$, so that we can express hyp $\hat{v}_*$ as $\mathcal{G} \cup \mathcal{H}$, where

$$\mathcal{G} = \{(q, \xi) : q \in \mathcal{D}, \underline{v}_* - 1 \leq \xi \leq \hat{v}_*(q)\} \quad \text{and} \quad \mathcal{H} = \{(q, \xi) : q \in \mathcal{D}, \xi \leq \underline{v}_* - 1\}.$$

Now we will show that $\mathrm{co}(\mathrm{hyp}\,\hat{v}_*) = (\mathrm{co}\mathcal{G}) \cup (\mathrm{co}\mathcal{H}) = (\mathrm{co}\mathcal{G}) \cup \mathcal{H}$. The inclusion from right to left is trivial, so consider $(q, \xi) \in \mathrm{co}(\mathrm{hyp}\,\hat{v}_*)$. Then, by Theorem 2.3 in Rockafellar (1997), (q, ξ) is a convex combination of points (q_n, ξ_n) in hyp $\hat{v}_*$. Therefore, $q \in \mathcal{D}$, as the latter is a convex set, and $\xi = \sum_n \alpha_n \xi_n \leq \sum_n \alpha_n \hat{v}_*(q_n)$ since $\alpha_n \geq 0$ for all n. But $(\mathrm{co}\mathcal{G}) \cup \mathcal{H}$ contains all convex combinations of points in hyp $\hat{v}_*$ that satisfy this property, proving the inclusion from left to right. Finally, note that $\mathcal{H}$ is closed, and that $\mathcal{G}$ is bounded and closed, since $\hat{v}_*$ is upper semicontinuous. Therefore, $\mathrm{co}(\mathcal{G})$ is also closed by Theorem 17.2 in Rockafellar (1997). We conclude that $\mathrm{co}(\mathrm{hyp}\,\hat{v}_*) = (\mathrm{co}\mathcal{G}) \cup \mathcal{H}$ is closed, as desired. $\square$

Using the properties of V_* in Lemma 3, one can extend the proof of the expression for $V(\sigma)$ in Proposition 4 to the present case. We then show that $M(\sigma(\cdot|\mathcal{I})) = M_*(\sigma(\cdot|\mathcal{I}))$—the second equality in (B1) follows from the definition of $M_*(\sigma(\cdot|\mathcal{I}))$. Since $k_* \geq k$ and hence $\hat{m}_* \geq \hat{m}$, it suffices to show that $M_*(\sigma(\cdot|\mathcal{I})) \leq M(\sigma(\cdot|\mathcal{I}))$. Recall that

$$M(\sigma(\cdot|\mathcal{I})) \geq \sup_{\{\tau : \mathbb{E}_\tau[q] = \sigma(\cdot|\mathcal{I})\}} \mathbb{E}_\tau[\max\{\hat{k}(q), \hat{v}(q)\}].$$

By the same logic as in the proof of Proposition 2, without loss of generality each $q \in \mathcal{D}$ conceals at most one $\omega \in \mathcal{I}$ and hence can be written as $q = (1 - q(\omega))q(\cdot|\mathcal{P}^i) + q(\omega)\delta_\omega$, where $\mathcal{P}^i = \mathbf{supp}\,q \cap \mathbf{supp}\,\mathbf{r}_L(q)$; moreover, Sender can always conceal ω in some $q \in \mathcal{D}$ that satisfies $v(a(\mathbf{p}_L(q)), \omega) = k_*(\omega)$. Let $\mathcal{T}^h_{\sigma(\cdot|\mathcal{I})}$ be the subset of $\{\tau : \mathbb{E}_\tau[q] = \sigma(\cdot|\mathcal{I})\}$ that contains only the $\tau \in \Delta(\mathcal{D})$ for which

she conceals states in this way. Then for every $\tau \in \mathcal{T}^h_{\sigma(\cdot|\mathcal{I})}$,

$$\mathbb{E}_\tau[\max\{\hat{k}(q), \hat{v}(q)\}] = \mathbb{E}_\tau[\max\{\hat{k}(q), (1 - q(\omega))\hat{v}(q(\cdot|\mathcal{P}^i)) + q(\omega)\delta_\omega \hat{k}_*(\delta_\omega)\}]$$
$$= \mathbb{E}_{\tau'}[\max\{\hat{k}_*(q'), \hat{v}(q')\}]$$

for some $\tau' \in \{\tau : \mathbb{E}_\tau[q] = \sigma(\cdot|\mathcal{I})\}$. It follows that

$$M(\sigma(\cdot|\mathcal{I})) \geq \sup_{\{\tau:\mathbb{E}_\tau[q]=\sigma(\cdot|\mathcal{I})\}} \mathbb{E}_\tau[\max\{\hat{k}_*(q), \hat{v}(q)\}].$$

Now let $\hat{m}^d = \max\{\hat{k}_*, \hat{v}\}$ over $\mathcal{D}$. Note that $\hat{m}^d \leq \max\{\hat{k}_*, \hat{v}_*\} = \hat{m}_*$ by the definition of $\hat{v}_*$, and that $\hat{m}_*$ is the smallest u.s.c. function that pointwise dominates $\hat{m}^d$. Therefore, $M^d \leq M_*$; moreover, by the same logic as in the proof of Lemma 3, we must have $M^d(\sigma(\cdot|\mathcal{I})) = M_*(\sigma(\cdot|\mathcal{I}))$ because $\sigma(\cdot|\mathcal{I}) \in int\,\mathcal{D}$. This implies that $M(\sigma(\cdot|\mathcal{I})) \geq M_*(\sigma(\cdot|\mathcal{I}))$ and completes the proof.

In Proposition 8, $\mathbf{supp}\,q^h$ tells us which states are concealed, while $k_*(\omega)$ tells us which evidence *and* worldview are used to conceal ω—hence, how deeply Receiver is induced to think. Similarly, if $\omega \in \mathbf{supp}\,q^s \cap \mathcal{P}^i$, it means that he is induced to think up to (at least) level i in his LBS. Proposition 8 also allows us to identify which states Sender will conceal, before characterizing her entire experiment.

Corollary 4. *Fix $\omega \in \mathcal{I}$. If $\hat{v}(\delta_\omega) \geq k_*(\omega)$, then Sender never conceals ω, that is, $\omega \notin \mathbf{supp}\,q^h$ in (B1). If $\hat{v}(\delta_\omega) < k_*(\omega)$ and $\hat{v}$ is convex over $\mathcal{D}$, then Sender always conceals ω, that is, $\omega \notin \mathbf{supp}\,q^s$ in (B1).*[31]

Even if $\hat{v}(\delta_\omega) \geq k_*(\omega)$, Sender need not fully reveal ω. If $\hat{v}$ is not convex, even if $\hat{v}(\delta_\omega) < k_*(\omega)$, she may reveal $\omega \in \mathcal{I}$ with positive probability (even whenever ω occurs).

To conclude, the next example illustrates the new constraints and opportunities introduced by a multiplicity of Receiver's worldviews, and why they may help or hurt Sender.
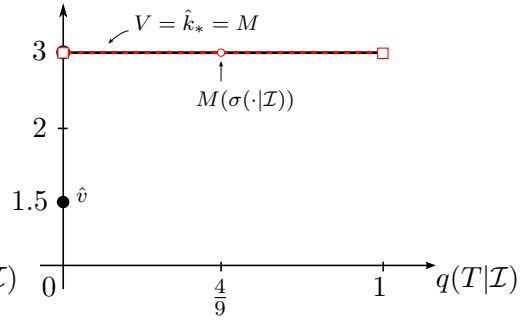
Example 4. Modify Example 3 by letting Rick have an intermediate worldview which rules out t or T. That is, $\mathrm{LBS}_T = (\rho^0, \hat{\rho}^1, \sigma)$ and $\mathrm{LBS}_t = (\rho^0, \tilde{\rho}^1, \sigma)$, where $\rho^0 = \sigma(\cdot|\mathcal{P})$ and $\hat{\rho}^1(T) = \tilde{\rho}^1(t) = 0$.

Under LBS_T, Susan can conceal T in two ways. The first uses evidence confirming ρ^0 and induces a_H; the second uses evidence disproving ρ^0 but consistent with ω_t and hence $\hat{\rho}^1$, so that Rick assigns to ω_t probability 1 and chooses a_t. Since Susan prefers a_T to a_H, she never conceals T as before. However, she can no longer pool t and T in the same evidence to raise her chances of getting a_T—any such evidence would result in a_t. Therefore, she always conceals t, inducing a_H. Figure B1 shows the optimal π for t and T. Relative to the original example,

---

[31]The proof of Corollary 4 follows the same logic as that of Corollary 2 and is therefore omitted.

Rick's reluctance to conceive the good state for Susan, T, causes her to conceal the bad state, t, with higher probability; it is easy to see that this harms her expected profits relative to the original example.

Under LBS_t, Susan can conceal t in two ways. The first is as before and induces a_H; the second uses evidence disproving ρ^0 but consistent with T and hence $\tilde{\rho}^1$, so that Rick chooses a_T. Concealing t is now even more attractive. Also, if Susan pools T and t in the same evidence, Rick is surprised, but thinks that T obtained and always chooses a_T. Therefore, now she always conceals t and induces a_T, which leads to higher expected profits than in the original example. Figure B2 shows the optimal π for LBS_t and LBS_T.
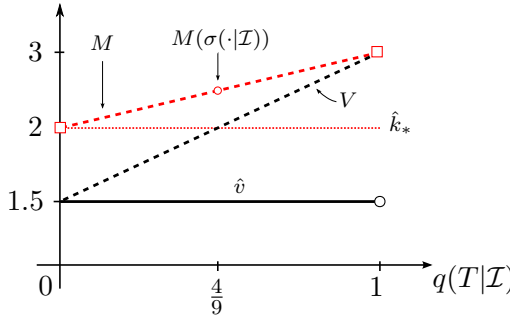
FIGURE B1. CONCEALING WITH LBS_T FIGURE B2. CONCEALING WITH LBS_t

Consistent with Kreps and Wilson (1982), one interpretation of A4 is that Receiver has a "meta-prior" over worldviews which assigns to higher-ranked elements of his LBS a higher likelihood of being correct. Receiver always adopts the worldview with highest likelihood under his meta-prior, after removing those disproven by the evidence. That is, the ranking of surviving worldviews is never updated. The next model relaxes this property.

Ortoleva's (2012) Model of Worldview Change. Ortoleva (2012) offers another way to model a Receiver with multiple alternative worldviews. Unlike in the LBS model, now he fully updates his meta-prior before adopting a new worldview. This model results from behavior-based notions of coherence and dynamic consistency—besides standard axioms of subjective expected utility. Coherence says that Receiver should have the same posterior after seeing pieces of evidence with the same content. Dynamic consistency is self-explanatory.

Formally, let $\mu \in \Delta(\Delta(\Omega))$ be such that $\mathbf{supp}\,\mu$ is finite, $\rho_\varepsilon = \arg\max_\rho \mu(\rho)$, and for every $\omega \in \Omega$ there exists $\rho \in \mathbf{supp}\,\mu$ with $\rho(\omega) > 0$. For every (s, π), let

$$(\text{B2}) \qquad \mu'(\rho|s,\pi) = \frac{\left[\sum_{\omega \in \Omega} \pi(s|\omega)\rho(\omega)\right]\mu(\rho)}{\sum_{\tilde{\rho} \in \mathbf{supp}\,\mu}\left[\sum_{\omega \in \Omega} \pi(s|\omega)\tilde{\rho}(\omega)\right]\mu(\tilde{\rho})}, \quad \rho \in \Delta(\Omega).$$

Since the maximizer of $\mu'(\cdot|s,\pi)$ need not be unique, following Ortoleva (2012) we endow Receiver with a strict linear order $\succ$ over priors, which he applies when the maximum-likelihood criterion is inconclusive.

Assumption 5 (A5: Hypothesis-Testing Model). *Receiver's initial prior is ρ_ε. If (s,π) is expected, he updates ρ_ε via Bayes' rule; otherwise, he updates $\rho(s,\pi)$ via Bayes' rule, where $\rho(s,\pi)$ is $\succ$-maximal in $\arg\max_\rho \mu'(\rho|s,\pi)$.*[32]

The notion of (un)expected evidence in A5 is as in Definition 1.

Under A5, the characterization of feasible posteriors is similar to Proposition 6. Sender's q again pins down which prior Receiver updates, denoted by $\mathbf{r}_\mu(q)$. If $q \in \mathcal{C}_\varepsilon$, $\mathbf{r}_\mu(q) = \rho_\varepsilon$. If $q \in \mathcal{D}_\varepsilon$, express $\pi(s|\omega)$ in terms of $q(\omega|s,\pi)$ using AC's logic and write (B2) as

$$\mu'(\rho; q) = \frac{\left[\sum_{\omega \in \Omega} q(\omega)\frac{\rho(\omega)}{\sigma(\omega)}\right]\mu(\rho)}{\sum_{\tilde\rho \in \mathbf{supp}\,\mu}\left[\sum_{\omega \in \Omega} q(\omega)\frac{\tilde\rho(\omega)}{\sigma(\omega)}\right]\mu(\tilde\rho)}, \quad \rho \in \Delta(\Omega).$$

Then, $\mathbf{r}_\mu(q)$ equals the $\succ$-maximal element in $\arg\max_\rho \mu'(\rho; q)$. For all q and $\mathbf{r}_\mu(q)$, we get Receiver's posterior using again AC's logic, thereby obtaining a new function $\mathbf{p}_\mu : \Delta(\Omega) \to \Delta(\Omega)$. This function retains the properties in Corollary 1. Depending on μ, $\mathbf{p}_\mu$ can allow Sender to conceal states after disproving ρ_ε as under A4 (see below).

Given $\mathbf{p}_\mu$, the insights of Section III remain valid. Under A5, finding the concealment payoff for $\mathcal{I}$-states can be more intricate: Now Receiver may discard a new ρ' that deems ω impossible for one that does not when the evidence is *sufficiently* against ρ'. Given those payoffs, the characterization of optimal persuasion for the LBS model applies unchanged.

Example 5 (Discontinuity of Receiver's Posterior over $\mathcal{D}$ under A5). Let $\varepsilon = 0$. Given $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and $\sigma = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, let $\mathbf{supp}\,\mu = (\rho^0, \rho^1, \rho^2)$ with $\mu(\rho^0) = \frac{1}{2}$, $\mu(\rho^1) = \mu(\rho^2) = \frac{1}{4}$, $\mathcal{P} = \{\omega_1\}$, $\rho^1 = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0)$, and $\rho^2 = \sigma$. Consider Sender's posterior $q_z = (0, \frac{1-z}{2}, \frac{1-z}{2}, z) \in \mathcal{D}$ for $z \in (0,1)$. Then

$$\mu'(\rho^1; q_z) = \frac{\sum_{\omega \in \Omega} q_z(\omega)\rho^1(\omega)}{\sum_{\omega \in \Omega} q_z(\omega)\rho^1(\omega) + \sum_{\omega \in \Omega} q_z(\omega)\rho^2(\omega)} = \frac{1}{1 + \frac{2}{3(1-z)}},$$

and

$$\mu'(\rho^2; q_z) = 1 - \mu'(\rho^1; q_z).$$

Hence, $\mu'(\rho^1; q_z) \geq \mu'(\rho^2; q_z)$ if and only if $z \leq \frac{1}{3}$. For $z = \frac{1}{3}$, Receiver will choose either ρ^1 or ρ^2, depending on how he ranks them under $\succ$.

---

[32]The analysis does not change if Receiver chooses worldviews using any other criterion that continues to rely only on the priors μ and μ'. For example, he may minimize the expectation under μ (resp. μ') of some loss function that depends on his choice and the "true" ρ.

Using AC's logic, we can compute Receiver's posteriors, starting with ρ^1 and ρ^2, when Sender has posterior q_z. Focusing on ω_3, we have

$$\mathbf{p}_1(\omega_3; q_z) = \frac{(1-z)\frac{1}{2}}{(1-z)\frac{1}{2} + (1-z)\frac{1}{4}} = \frac{2}{3},$$

$$\mathbf{p}_2(\omega_3; q_z) = \frac{(1-z)\frac{1}{4}}{(1-z)\frac{1}{4} + (1-z)\frac{1}{4} + 2z\frac{1}{4}} = \frac{1-z}{2}.$$

Thus, Receiver's posterior must vary discontinuously in q_z at $z = \frac{1}{3}$. Note also that for z with $0 < z < \frac{1}{3}$ we have $\mathbf{p}_1(\omega_4; q_z) = 0$, even though $q_z(\omega_4) > 0$. Therefore, Sender is hiding ω_4 in posterior q_z after disproving ρ^0.

B5. Is this Model Just the Limit of a Standard Model?

Let $\Omega = \{\omega_1, \omega_2\}$, $\sigma = (\frac{1}{2}, \frac{1}{2})$, and $\rho = (1-\varepsilon, \varepsilon)$ for $\varepsilon > 0$. Let $A = \{a, b, c\}$, and let the payoff functions be

v	a	b	c
ω_1	1	0	-1
ω_2	1	0	-1

u	a	b	c
ω_1	1	2	-1
ω_2	1	-1	2

Abusing notation, let p and q be the posterior probabilities Receiver and Sender assign to ω_2, so that $\mathbf{b}(q) = \frac{q\varepsilon}{q\varepsilon + (1-q)(1-\varepsilon)}$. She can induce him to choose a if $p \in [\frac{1}{3}, \frac{2}{3}]$; he will choose b if $p < \frac{1}{3}$ and c if $p > \frac{2}{3}$. Note that $\mathbf{b}(q) \in [\frac{1}{3}, \frac{2}{3}]$ if and only if $q \in [\underline{q}_\varepsilon, \bar{q}_\varepsilon]$, where $\underline{q}_\varepsilon = \frac{1-\varepsilon}{1+\varepsilon}$ and $\bar{q}_\varepsilon = \frac{2(1-\varepsilon)}{2-\varepsilon}$. Also, $\underline{q}_\varepsilon$ and $\bar{q}_\varepsilon$ satisfy $0 < \underline{q}_\varepsilon < \bar{q}_\varepsilon < 1$, are strictly decreasing, and converge to 1 as $\varepsilon \to 0$. Therefore, no matter how small ε is, Sender can always persuade Receiver to choose a with positive probability. Focusing on the case of $\underline{q}_\varepsilon > \frac{1}{2}$, it is easy to check that her optimal τ generates two posteriors, $q = 0$ and $q = \underline{q}_\varepsilon$, with $\tau(\underline{q}_\varepsilon) = \frac{1}{2\underline{q}_\varepsilon}$. This τ always induces Receiver to choose both a and b with positive probability and yields expected payoff $\frac{1}{2\underline{q}_\varepsilon}$, which converges to $\frac{1}{2}$ as $\varepsilon \to 0$. When ε reaches 0, we obtain a case of $\varepsilon = 0$ and $\mathcal{P} = \{\omega_1\}$ in the baseline model of Section I. Now Sender cannot induce any $p \in [\frac{1}{3}, \frac{2}{3}]$ and hence action a. Since she strictly prefers b to c, her best experiment is fully uninformative ($\tau = \delta_\sigma$), yielding payoff 0. Proposition 7 confirms that this substantive difference between the two models extends beyond the case of $\varepsilon = 0$.

B6. Omitted Proofs

Proof of Proposition 5

First, let's derive the function $\hat{v}(q) = \mathbb{E}_q[v(\mathbb{E}_{\mathbf{p}(q)}[\omega], \omega)]$:

$$\hat{v}(q) = -(\mathbb{E}_{\mathbf{p}(q)}[\omega])^2 + 2\kappa\mathbb{E}_{\mathbf{p}(q)}[\omega]\mathbb{E}_q[\omega] + 2b(\mathbb{E}_{\mathbf{p}(q)}[\omega] - \kappa\mathbb{E}_q[\omega]) - \kappa^2\mathbb{E}_q[\omega^2] - b^2.$$

To apply Proposition 4, we only need $\hat{v}(q)$ for $q \in \Delta(\mathcal{P})$ and $q \in \mathcal{D} = \Delta(\mathcal{I})$. Since $\rho = \sigma(\cdot|\mathcal{P})$ and $\rho^1 = \sigma$, we have $\mathbf{p}(q) = q$ for every $q \in \Delta(\mathcal{P}) \cup \Delta(\mathcal{I})$.[33] This yields

$$\hat{v}(q) = (2\kappa - 1)(\mathbb{E}_q[\omega])^2 + 2b(1 - \kappa)\mathbb{E}_q[\omega] - \kappa^2\mathbb{E}_q[\omega^2] - b^2, \quad q \in \Delta(\mathcal{P}) \cup \Delta(\mathcal{I}),$$

which is strictly convex (concave) if and only if $\kappa > (<) \frac{1}{2}$. If $\kappa \geq \frac{1}{2}$, $V(\sigma(\cdot|\mathcal{P}))$ and $V(\sigma(\cdot|\mathcal{I}))$ are then achieved with a fully informative π (uniquely if $\kappa > \frac{1}{2}$); if $\kappa < \frac{1}{2}$, they are achieved only with a fully uninformative π. Note that the range of policies Receiver can be induced to choose under his worldview is $A^c = [\omega_{\underline{n}}, \omega_{\overline{n}}]$, and that for every $\omega \in \mathcal{I}$,

$$k(\omega) = \begin{cases} -(\omega_{\overline{n}} - \beta(\omega))^2 & \text{if } \beta(\omega) > \omega_{\overline{n}} \\ -(\omega_{\underline{n}} - \beta(\omega))^2 & \text{if } \beta(\omega) < \omega_{\underline{n}} \\ 0 & \text{if } \omega_{\underline{n}} \leq \beta(\omega) \leq \omega_{\overline{n}}. \end{cases}$$

Case 1: $\kappa \geq \frac{1}{2}$. The properties stated in the proposition follows from the next claim.

Claim 10. If $\kappa \geq \frac{1}{2}$, Sender's optimal π has the following properties:
- π fully reveals ω_i for all $i \leq \overline{n}$;
- there exist $b^*(\kappa) \leq 0$ and $i^*(b, \kappa)$ such that, if $b \geq b^*(\kappa)$, π fully reveals all states, while if $b < b^*(\kappa)$, π always conceals ω_i if $\overline{n} < i < i^*(b, \kappa)$ and fully reveals ω_i if $i \geq i^*(b, \kappa)$;
- $b^*(\kappa)$ decreases in κ (strictly if negative) and $i^*(b, \kappa)$ decreases in κ and b.

Proof. We already know that the lobbyist fully reveals every $\omega \leq \omega_{\overline{n}}$. For $i > \overline{n}$, since $\hat{v}$ is convex over $\mathcal{D}$, by Corollary 2 we have that ω_i is concealed with probability 1 if $\hat{v}(\delta_{\omega_i}) < \hat{k}(\delta_{\omega_i})$ and is never concealed otherwise. So fix $i > \overline{n}$. For each value of κ there exists a value $b_i(\kappa)$ such that $\hat{v}(\delta_{\omega_i}) \geq \hat{k}(\delta_{\omega_i})$ if and only if $b \geq b_i(\kappa)$: This threshold is given by

$$b_i(\kappa) = \min\left\{\left(\frac{1}{2} - \kappa\right)\omega_i + \frac{1}{2}\omega_{\overline{n}}, 0\right\}.$$

---

[33]More generally, one could apply the techniques developed by AC for cases with different common-support priors to characterize $V(\sigma(\cdot|\mathcal{P}))$ and $M(\sigma(\cdot|\mathcal{I}))$.

Each $b_i(\kappa)$ is decreasing in κ (strictly when negative), and $b_{i+1}(\kappa) \le b_i(\kappa)$ (with $<$ if either threshold is negative). If $b \ge b^*(\kappa) \equiv b_{\overline{n}+1}(\kappa)$, we have that $V(\sigma(\cdot|\mathcal{I})) = M(\sigma(\cdot|\mathcal{I}))$, and hence $\tau^d = \sigma(\mathcal{I})$ by Proposition 4. On the other hand, if $b < b^*(\kappa)$, let $i^*(b, \kappa) = \min\{i > \overline{n} : b \ge b_i(\kappa)\}$, which is non-increasing in both κ and b. Thus it is optimal to conceal with probability 1 every ω_i with $\overline{n} < i < i^*(b, \kappa)$ and to fully reveal all other states in $\mathcal{I}$. $\qquad\square$

Case 2: $\kappa < \frac{1}{2}$**.** The properties stated in the proposition follows from the next claim.

Claim 11. If $\kappa < \frac{1}{2}$, Sender's expected payoff from her optimal π equals

$$-\tau^d \mathbb{E}_{q^d}(a^d - \omega)^2 + (1 - \tau^d) \sum_{\omega > \omega_{\overline{n}}} k(\omega)q^c(\omega) - \sum_{\omega \in \mathcal{P}} (\mathbb{E}_\sigma[\omega|\mathcal{P}] - \omega)^2 \sigma(\omega),$$

where $\tau^d > 0$, $\tau^d q^d + (1 - \tau^d)q^c = \sigma$, and $a^d = \mathbb{E}_{q^d}[\omega]$. Either $\mathbb{E}_\sigma[\omega|\omega < \omega_{\underline{n}}] \le a^d < \omega_{\underline{n}}$ or $a^d > \omega_{\overline{n}}$. If $\mathbb{E}_\sigma[\omega|\mathcal{I}] \le \omega_{\overline{n}}$, π conceals some $\omega > \omega_{\overline{n}}$ with positive probability and $a^d < \omega_{\underline{n}}$.

Proof. By Proposition 5, $M(\sigma(\cdot|\mathcal{I}))$ is given by

$$\max_{\gamma \in [0,1], q_1, q_2 \in \mathcal{D}} \left\{ \gamma \mathbb{E}_{q_1}[k(\omega)] + (1 - \gamma)\mathbb{E}_{q_2}[-(\mathbb{E}_{q_2}[\omega] - \beta(\omega))^2] \right\},$$

subject to $\gamma q_1 + (1 - \gamma)q_2 = \sigma(\cdot|\mathcal{I})$. By continuity of $\hat{k}$ and $\hat{v}$ over $\mathcal{D}$, a solution (γ, q_1, q_2) to this problem exists. Recall that $\omega_i \notin \operatorname{supp} q_1$ for all $i < \underline{n}$. Suppose that (γ, q_1, q_2) implies $\omega_{\underline{n}} \le \mathbb{E}_{q_2}[\omega] \le \omega_{\overline{n}}$. We will show that there exists a feasible (γ', q_1', q_2') which strictly dominates (γ, q_1, q_2). Since β is strictly increasing, we must have $\omega_j \in \operatorname{\mathbf{supp}} q_2$ for some $j > \overline{n}$. Suppose first that $\beta(\omega) > \omega_{\underline{n}}$ for some $\omega > \omega_{\overline{n}}$. Then, for any $\xi \in [\omega_{\underline{n}}, \omega_{\overline{n}}]$,

$$\begin{aligned}
-\mathbb{E}_{q_2}[(\xi - \omega)^2] \quad &< \quad -\mathbb{E}_{q_2}[(\omega_{\underline{n}} - \beta(\omega))^2 \mathbf{1}\{\beta(\omega) < \omega_{\underline{n}}\}] \\
&\qquad -\mathbb{E}_{q_2}[(\omega_{\overline{n}} - \beta(\omega))^2 \mathbf{1}\{\beta(\omega) > \omega_{\overline{n}}\}] \\
&= \quad \mathbb{E}_{q_2}[k(\omega)].
\end{aligned}$$

This means that $\gamma' = 1$ and $q_1' = \sigma(\cdot|\mathcal{I})$ strictly dominates (γ, q_1, q_2). Now, suppose that $\beta(\omega) \le \omega_{\underline{n}}$ for all $\omega > \omega_{\overline{n}}$. If $\omega_{\underline{n}} < \mathbb{E}_{q_2}[\omega] \le \omega_{\overline{n}}$, then again $\gamma' = 1$ and $q_1' = \sigma(\cdot|\mathcal{I})$ strictly dominates (γ, q_1, q_2). If $\mathbb{E}_{q_2}[\omega] = \omega_{\underline{n}}$, then $M(\sigma(\cdot|\mathcal{I})) = \hat{k}(\sigma(\cdot|\mathcal{I}))$. But we know that always concealing all states in $\mathcal{I}$ is not optimal: $V(\sigma) > V^c(\sigma)$ since $\hat{v}(\delta_{\omega_1}) > \hat{k}(\delta_{\omega_1})$. Therefore, (γ, q_1, q_2) is again strictly dominated.

Finally, if σ is such that $\omega_{\underline{n}} \le \mathbb{E}_{\sigma(\cdot|\mathcal{I})}[\omega] \le \omega_{\overline{n}}$, then $\tau^d = \sigma(\mathcal{I})$ implies that $q^d = \sigma(\cdot|\mathcal{I})$, and hence $\omega_{\underline{n}} \le \mathbb{E}_{q^d}[\omega] \le \omega_{\overline{n}}$, which cannot be optimal as we have just argued. Moreover, concealing can occur only for $\omega > \omega_{\overline{n}}$ and never occurs

for $\omega < \omega_{\underline{n}}$, so $\mathbb{E}_{q^d}[\omega]$ must be lower than $\omega_{\underline{n}}$. A fortiori, the same holds if $\mathbb{E}_{\sigma(\cdot|\mathcal{I})}[\omega] < \omega_{\underline{n}}$. $\qquad\qquad\square$

PROOF OF PROPOSITION 6

Since $\mathbf{supp}\,\rho_\varepsilon = \mathbf{supp}\,\sigma$, by AC logic there exists an homeomorphism from Sender's to Receiver's posteriors that describes the Bayes' updating of ρ_ε corresponding to the (s, π) inducing each q. Since the condition $p_\varepsilon(\mathcal{I}|q) \geq \theta(\varepsilon)$ partitions the set of Receiver's posteriors $\Delta(\Omega)$, it implies (via the homeomorphism) the partition of Sender's posterior given by

$$\mathcal{D}_\varepsilon = \{q : p_\varepsilon(\mathcal{I}|q) \geq \theta(\varepsilon)\} \quad \text{and} \quad \mathcal{C}_\varepsilon = \{q : p_\varepsilon(\mathcal{I}|q) < \theta(\varepsilon)\} = \Delta(\Omega) \setminus \mathcal{D}_\varepsilon.$$

As ε falls, $\mathcal{D}_\varepsilon$ shrinks in the sense of set inclusion. Indeed,

$$p_\varepsilon(\mathcal{I}|q) = \frac{\sum_{\omega \in \mathcal{I}} \frac{q(\omega)}{\sigma(\omega)} \rho_\varepsilon(\omega)}{\sum_{\omega' \in \Omega} \frac{q(\omega')}{\sigma(\omega')} \rho_\varepsilon(\omega')} \geq \theta(\varepsilon) \quad \text{iff} \quad \frac{\sum_{\omega \in \mathcal{P}} \frac{q(\omega)}{\sigma(\omega)} \frac{\rho_\varepsilon(\omega)}{\rho_\varepsilon(\mathcal{P})}}{\sum_{\omega' \in \mathcal{I}} \frac{q(\omega')}{\sigma(\omega')} \frac{\rho_\varepsilon(\omega')}{\rho_\varepsilon(\mathcal{I})}} \leq \frac{\varepsilon(1 - \theta(\varepsilon))}{\theta(\varepsilon)(1 - \varepsilon)};$$

also, by A1 the left-hand side of the second inequality is independent of ε, and by A2 the right-hand side decreases to zero as $\varepsilon \to 0$. Using continuity of $p_\varepsilon(\mathcal{I}|q)$ in q, we get $\mathcal{D}_{\varepsilon'} \subset \mathcal{D}_\varepsilon$ for all $\varepsilon' < \varepsilon$. In the limit $q \in \Delta(\mathcal{I})$ if and only if $q(\omega) = 0$ for all $\omega \in \mathcal{P}$. Hence, $\lim_{\varepsilon \to 0} \mathcal{D}_\varepsilon = \cap_{\varepsilon \geq 0} \mathcal{D}_\varepsilon = \Delta(\mathcal{I})$. Finally, since $\mathcal{C}_\varepsilon = \Delta(\Omega) \setminus \mathcal{D}_\varepsilon$, $\mathcal{C}_\varepsilon$ grows in the sense of set inclusion as ε falls. Also, $\lim_{\varepsilon \to 0} \mathcal{C}_\varepsilon = \Delta(\Omega) \setminus (\lim_{\varepsilon \to 0} \mathcal{D}_\varepsilon) = \Delta(\Omega) \setminus \Delta(\mathcal{I})$. The rest of the result follows immediately.

PROOF OF COROLLARY 3

Part (1): Note that $\mathcal{D}_\varepsilon$ is closed. Let $\hat{q} \in \mathcal{D}_\varepsilon \cap \text{cl}\mathcal{C}_\varepsilon$, where $\text{cl}\mathcal{C}_\varepsilon$ is the closure of $\mathcal{C}_\varepsilon$. Let $\{q^n\} \subset \mathcal{C}_\varepsilon$ be any sequence that satisfies $q^n \to \hat{q}$ as $n \to \infty$. For all $\omega \in \Omega$,

$$\lim_{n \to \infty} \mathbf{p}_\varepsilon(\omega; q^n) = \frac{\frac{\hat{q}(\omega)}{\sigma(\omega)} \rho_\varepsilon(\omega)}{\sum_{\omega' \in \Omega} \frac{\hat{q}(\omega')}{\sigma(\omega')} \rho_\varepsilon(\omega')} \quad \text{and} \quad \mathbf{p}_\varepsilon(\omega; \hat{q}) = \frac{\frac{\hat{q}(\omega)}{\sigma(\omega)} \rho^1(\omega)}{\sum_{\omega' \in \Omega} \frac{\hat{q}(\omega')}{\sigma(\omega')} \rho^1(\omega')}.$$

Since at $\hat{q}$ we have

$$\frac{\sum_{\omega \in \mathcal{P}} \frac{\hat{q}(\omega)}{\sigma(\omega)} \frac{\rho_\varepsilon(\omega)}{\rho_\varepsilon(\mathcal{P})}}{\sum_{\omega' \in \mathcal{I}} \frac{\hat{q}(\omega')}{\sigma(\omega')} \frac{\rho_\varepsilon(\omega')}{\rho_\varepsilon(\mathcal{I})}} = \frac{\varepsilon(1 - \theta(\varepsilon))}{\theta(\varepsilon)(1 - \varepsilon)} > 0,$$

it follows that $(\mathbf{supp}\,\hat{q}) \cap \mathcal{P} \equiv \mathcal{P}_{\hat{q}} \neq \varnothing$ and $(\mathbf{supp}\,\hat{q}) \cap \mathcal{I} \equiv \mathcal{I}_{\hat{q}} \neq \varnothing$. Thus, for all $\hat{\omega}, \omega' \in \mathcal{P}_{\hat{q}} \cup \mathcal{I}_{\hat{q}}$, the following quantities are well defined:

$$\lim_{n \to \infty} \frac{\mathbf{p}_\varepsilon(\hat{\omega}; q^n)}{\mathbf{p}_\varepsilon(\omega'; q^n)} = \frac{\frac{\hat{q}(\hat{\omega})}{\sigma(\hat{\omega})} \rho_\varepsilon(\hat{\omega})}{\frac{\hat{q}(\omega')}{\sigma(\omega')} \rho_\varepsilon(\omega')} \quad \text{and} \quad \frac{\mathbf{p}_\varepsilon(\hat{\omega}; \hat{q})}{\mathbf{p}_\varepsilon(\omega'; \hat{q})} = \frac{\frac{\hat{q}(\hat{\omega})}{\sigma(\hat{\omega})} \rho^1(\hat{\omega})}{\frac{\hat{q}(\omega')}{\sigma(\omega')} \rho^1(\omega')}.$$

If $\mathbf{p}_\varepsilon$ is continuous, the limit on the left must equal the quantity on the right for all pairs of states in $\mathcal{P}_{\hat{q}} \cup \mathcal{I}_{\hat{q}}$. This implies $\rho^1(\omega) = \frac{\rho^1(\omega')}{\rho_\varepsilon(\omega')}\rho_\varepsilon(\omega)$ for all $\omega \in \mathcal{P}_{\hat{q}} \cup \mathcal{I}_{\hat{q}}$ and some fixed $\omega' \in \mathcal{P}_{\hat{q}} \cup \mathcal{I}_{\hat{q}}$. Using this, we get

$$\frac{\rho^1(\mathcal{P}_{\hat{q}})}{\rho^1(\mathcal{I}_{\hat{q}})} = \frac{\rho_\varepsilon(\mathcal{P}_{\hat{q}})}{\rho_\varepsilon(\mathcal{I}_{\hat{q}})} = \frac{\rho_\varepsilon(\mathcal{P}_{\hat{q}})/\rho_\varepsilon(\mathcal{P})}{\rho_\varepsilon(\mathcal{I}_{\hat{q}})/\rho_\varepsilon(\mathcal{I})} \times \frac{1-\varepsilon}{\varepsilon}.$$

By A1, $\frac{\rho_\varepsilon(\mathcal{P}_{\hat{q}})/\rho_\varepsilon(\mathcal{P})}{\rho_\varepsilon(\mathcal{I}_{\hat{q}})/\rho_\varepsilon(\mathcal{I})}$ is independent of ε and can only take finitely many, strictly positive, values, since Ω is finite. Also, the properties of ρ^1 imply that $\frac{\rho^1(\mathcal{P}_{\hat{q}})}{\rho^1(\mathcal{I}_{\hat{q}})} \leq \frac{1-\rho^1}{\rho^1}$ for $\hat{q} \in \mathcal{D}_\varepsilon \cap \mathrm{cl}\mathcal{C}_\varepsilon$. Thus, there exists $\bar{\varepsilon} > 0$ such that, if $\varepsilon < \bar{\varepsilon}$, then $\frac{\rho^1(\mathcal{P}_{\hat{q}})}{\rho^1(\mathcal{I}_{\hat{q}})} < \frac{\rho_\varepsilon(\mathcal{P}_{\hat{q}})}{\rho_\varepsilon(\mathcal{I}_{\hat{q}})}$ for all $\hat{q} \in \mathcal{D}_\varepsilon \cap \mathrm{cl}\mathcal{C}_\varepsilon$ and hence $\mathbf{p}_\varepsilon$ cannot be continuous over this set.

Part (2): Consider $\mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)$. This set contains p if and only if p results from expected evidence, that is, $p(\mathcal{I}) < \theta(\varepsilon)$. By A2, this condition becomes more stringent as ε decreases. As a result, $\mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)$ shrinks, which also implies that $\sup\{p(\mathcal{I}) : p \in \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)\}$ decreases. By A2, $\lim_{\varepsilon \to 0} \theta(\varepsilon) = 0$. Therefore, $p \in \lim_{\varepsilon \to 0} \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon) = \cap_{\varepsilon \geq 0}\mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)$ implies that $p(\mathcal{I}) = 0$, which means that $p \in \Delta(\mathcal{P})$. Now suppose $p' \in \Delta(\mathcal{P})$. Since $p'(\mathcal{I}) = 0$, the evidence inducing p' rules out all $\omega \in \mathcal{I}$ and is therefore expected. Since $\theta(\varepsilon) > \varepsilon$, we have that $p'(\mathcal{I}) < \theta(\varepsilon)$ for all $\varepsilon \geq 0$. Therefore, $p' \in \cap_{\varepsilon \geq 0}\mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)$. This proves that $\cap_{\varepsilon \geq 0}\mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon) = \Delta(\mathcal{P})$.

Now consider $\mathbf{p}_\varepsilon(\mathcal{D}_\varepsilon)$. Each p in this set satisfies

$$\mathbf{p}_\varepsilon(\omega; q) = \frac{q(\omega)\frac{\rho^1(\omega)}{\sigma(\omega)}}{\sum_{\omega' \in \Omega} q(\omega')\frac{\rho^1(\omega')}{\sigma(\omega')}}, \quad \omega \in \Omega.$$

Since $\mathbf{supp}\, \rho^1 = \Omega$, viewed as a function over the entire set $\Delta(\Omega)$, this function is an homeomorphism (see Corollary 1). Therefore, since $\mathcal{D}_\varepsilon \subset \mathcal{D}_{\varepsilon'}$ for $\varepsilon < \varepsilon'$, we have $\mathbf{p}_\varepsilon(\mathcal{D}_\varepsilon) \subsetneq \mathbf{p}_\varepsilon(\mathcal{D}_{\varepsilon'})$ and hence $\sup\{p(\mathcal{P}) : p \in \mathbf{p}_\varepsilon(\mathcal{D}_\varepsilon)\}$ decreases as ε decreases. Moreover, $\lim_{\varepsilon \to 0} \mathbf{p}_\varepsilon(\mathcal{D}_\varepsilon) = \cap_{\varepsilon \geq 0}\mathbf{p}_\varepsilon(\mathcal{D}_\varepsilon) = \mathbf{p}_0(\Delta(\mathcal{I}))$. Restricted to $\Delta(\mathcal{I})$, $\mathbf{p}_0$ can be written as

$$\mathbf{p}_0(\omega; q) = \frac{q(\omega)\frac{\rho^1(\omega|\mathcal{I})}{\sigma(\omega|\mathcal{I})}}{\sum_{\omega' \in \mathcal{I}} q(\omega')\frac{\rho^1(\omega'|\mathcal{I})}{\sigma(\omega'|\mathcal{I})}}, \quad \omega \in \mathcal{I}.$$

Applying the first part of Corollary 1 for priors $\rho^1(\cdot|\mathcal{I})$ and $\sigma(\cdot|\mathcal{I})$, we conclude that $\mathbf{p}_0 : \Delta(\mathcal{I}) \to \Delta(\mathcal{I})$ is onto and hence $\mathbf{p}_0(\Delta(\mathcal{I})) = \Delta(\mathcal{I})$.

Part (3): Given q_η and the corresponding $q' \in \Delta(\mathcal{P})$ and $\hat{\omega} \in \mathcal{I}$, we have

$$
\begin{aligned}
\frac{1 - p_\varepsilon(\mathcal{I}|q_\eta)}{p_\varepsilon(\mathcal{I}|q_\eta)} &= \frac{(1-\eta)\sum_{\omega \in \mathcal{P}} \frac{q'(\omega)}{\sigma(\omega)}\rho_\varepsilon(\omega)}{\eta \frac{\rho_\varepsilon(\hat{\omega})}{\sigma(\hat{\omega})}} \\
&= \frac{(1-\eta)(1-\varepsilon)}{\eta\varepsilon} \times \frac{\frac{1}{\sigma(\mathcal{P})}\sum_{\omega \in \mathcal{P}} \frac{q'(\omega)}{\sigma(\omega|\mathcal{P})}\rho_\varepsilon(\omega|\mathcal{P})}{\frac{\rho_\varepsilon(\hat{\omega}|\mathcal{I})}{\sigma(\hat{\omega})}} \\
&\geq \frac{(1-\eta)(1-\varepsilon)}{\eta\varepsilon} \times \frac{\min_{\omega \in \mathcal{I}}\sigma(\omega)}{\sigma(\mathcal{P})},
\end{aligned}
$$

where the inequality uses $\sum_{\omega \in \mathcal{P}} \frac{q'(\omega)}{\sigma(\omega|\mathcal{P})}\rho_\varepsilon(\omega|\mathcal{P}) = 1$ for $q' \in \Delta(\mathcal{P})$. By A1, the quantity multiplying $\frac{(1-\eta)(1-\varepsilon)}{\eta\varepsilon}$ is independent of ε. Thus, for all $\varepsilon > 0$, $p_\varepsilon(\mathcal{I}|q_\eta) < \theta(\varepsilon)$ for all q' and $\hat{\omega}$ if and only if $\eta < \eta_\varepsilon$ for some unique $\eta_\varepsilon < 1$. Since $\frac{\varepsilon(1-\theta(\varepsilon))}{\theta(\varepsilon)(1-\varepsilon)} \to 0$ monotonically by A2, η_ε increases monotonically to 1.

Now suppose $\hat{q}, \tilde{q} \in \mathcal{C}_\varepsilon$ satisfy $\hat{q}(\cdot|\mathcal{P}) = \tilde{q}(\cdot|\mathcal{P}) = q'$. We can write $\hat{q} = (1 - \hat{q}(\mathcal{I}))q' + \hat{q}(\mathcal{I})\hat{q}'$ and $\tilde{q} = (1 - \tilde{q}(\mathcal{I}))q' + \tilde{q}(\mathcal{I})\tilde{q}'$, where $\hat{q}', \tilde{q}' \in \Delta(\mathcal{I})$. Letting $\hat{\mathbf{p}} = \mathbf{p}_\varepsilon(\hat{q})$ and $\tilde{\mathbf{p}} = \mathbf{p}_\varepsilon(\tilde{q})$, we have

$$
\begin{aligned}
||\hat{\mathbf{p}} - \tilde{\mathbf{p}}||^2 &= \sum_{\omega \in \Omega} [\hat{\mathbf{p}}(\mathcal{P})\hat{\mathbf{p}}(\omega|\mathcal{P}) + \hat{\mathbf{p}}(\mathcal{I})\hat{\mathbf{p}}(\omega|\mathcal{I}) - \tilde{\mathbf{p}}(\mathcal{P})\tilde{\mathbf{p}}(\omega|\mathcal{P}) - \tilde{\mathbf{p}}(\mathcal{I})\tilde{\mathbf{p}}(\omega|\mathcal{I})]^2 \\
&= \sum_{\omega \in \mathcal{P}} [\hat{\mathbf{p}}(\mathcal{P})\hat{\mathbf{p}}(\omega|\mathcal{P}) - \tilde{\mathbf{p}}(\mathcal{P})\tilde{\mathbf{p}}(\omega|\mathcal{P})]^2 + \sum_{\omega \in \mathcal{I}} [\hat{\mathbf{p}}(\mathcal{I})\hat{\mathbf{p}}(\omega|\mathcal{I}) - \tilde{\mathbf{p}}(\mathcal{I})\tilde{\mathbf{p}}(\omega|\mathcal{I})]^2 \\
&= \sum_{\omega \in \mathcal{P}} [\mathbf{p}'(\omega|\mathcal{P})]^2 [\hat{\mathbf{p}}(\mathcal{I}) - \tilde{\mathbf{p}}(\mathcal{I})]^2 + \sum_{\omega \in \mathcal{I}} [\hat{\mathbf{p}}(\mathcal{I})\hat{\mathbf{p}}(\omega|\mathcal{I}) - \tilde{\mathbf{p}}(\mathcal{I})\tilde{\mathbf{p}}(\omega|\mathcal{I})]^2 \\
&\leq |\mathcal{P}| [\hat{\mathbf{p}}(\mathcal{I}) - \tilde{\mathbf{p}}(\mathcal{I})]^2 + 2|\mathcal{I}| [\max\{\hat{\mathbf{p}}(\mathcal{I}), \tilde{\mathbf{p}}(\mathcal{I})\}]^2
\end{aligned}
$$

where the third step uses the fact that, for all $\omega \in \mathcal{P}$,

$$
\mathbf{p}'(\omega|\mathcal{P}) \equiv \frac{q'(\omega)\frac{\rho_\varepsilon(\omega)}{\sigma(\omega)}}{\sum_{\omega' \in \mathcal{P}} q'(\omega')\frac{\rho_\varepsilon(\omega')}{\sigma(\omega')}} = \frac{\hat{\mathbf{p}}(\omega)}{\hat{\mathbf{p}}(\mathcal{P})} = \frac{\tilde{\mathbf{p}}(\omega)}{\tilde{\mathbf{p}}(\mathcal{P})},
$$

and therefore $\hat{\mathbf{p}}(\omega|\mathcal{P}) = \tilde{\mathbf{p}}(\omega|\mathcal{P})$. Since $\sup\{p(\mathcal{I}) : p \in \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)\} \to 0$ as $\varepsilon \to 0$ by Part (2), $\max\{\hat{\mathbf{p}}(\mathcal{I}), \tilde{\mathbf{p}}(\mathcal{I})\} \to 0$ and $|\hat{\mathbf{p}}(\mathcal{I}) - \tilde{\mathbf{p}}(\mathcal{I})| \to 0$.

PROOF OF PROPOSITION 7

Part I: $\lim_{\varepsilon \to 0} V_\varepsilon^c(\sigma) = V_0^c(\sigma)$. The proof consists of the following three claims.

Claim 12. There exists $\bar{\varepsilon} > 0$ such that, if $\varepsilon < \bar{\varepsilon}$, then $a(\mathbf{p}_\varepsilon(q)) \in A^c$ for all $q \in \mathcal{C}_\varepsilon$.

Proof. For every $p \in \Delta(\Omega)$, let $a(p) \in \arg\max_{a \in A} \mathbb{E}_p[u(a, \omega)]$. Also, define

$$\zeta = \max_{a \notin A^c, \hat{a} \in A^c} \left\{ \max_{\omega \in \mathcal{I}} u(a, \omega) - \min_{\omega \in \mathcal{I}} u(\hat{a}, \omega) \right\},$$

which is bounded by finiteness of $\mathcal{I}$ and A, and

$$\chi = \max_{p \in \Delta(\mathcal{P})} \left\{ \max_{\tilde{a} \notin A^c} \mathbb{E}_p[u(\tilde{a}, \omega)] - \mathbb{E}_p[u(a(p), \omega)] \right\},$$

which is strictly negative because A is finite. Indeed, otherwise, there would be $p \in \Delta(\mathcal{P})$ and $\tilde{a} \notin A^c$ such that $\mathbb{E}_p[u(\tilde{a}, \omega)] = \mathbb{E}_p[u(a(p), \omega)]$, which contradicts the premise that $\tilde{a} \notin A^c$. Consider any $a' \notin A^c$ and $p \in \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)$. We have

$$
\begin{aligned}
\mathbb{E}_p[u(a', \omega)] \ = \ & \mathbb{E}_p[u(a(p(\cdot|\mathcal{P})), \omega)] \\
& + p(\mathcal{P}) \left\{ \mathbb{E}_{p(\cdot|\mathcal{P})}[u(a', \omega)] - \mathbb{E}_{p(\cdot|\mathcal{P})}[u(a(p(\cdot|\mathcal{P})), \omega)] \right\} \\
& + p(\mathcal{I}) \left\{ \mathbb{E}_{p(\cdot|\mathcal{I})}[u(a', \omega)] - \mathbb{E}_{p(\cdot|\mathcal{I})}[u(a(p(\cdot|\mathcal{P})), \omega)] \right\} \\
\leq \ & \mathbb{E}_p[u(a(p(\cdot|\mathcal{P})), \omega)] \\
& + p(\mathcal{P}) \left\{ \mathbb{E}_{p(\cdot|\mathcal{P})}[u(a', \omega)] - \mathbb{E}_{p(\cdot|\mathcal{P})}[u(a(p(\cdot|\mathcal{P})), \omega)] \right\} \\
& + p(\mathcal{I}) \left\{ \max_{\omega \in \mathcal{I}} u(a', \omega) - \min_{\omega \in \mathcal{I}} u(a(p(\cdot|\mathcal{P})), \omega) \right\} \\
\leq \ & \mathbb{E}_p[u(a(p(\cdot|\mathcal{P})), \omega)] \\
& + p(\mathcal{P}) \left\{ \mathbb{E}_{p(\cdot|\mathcal{P})}[u(a', \omega)] - \mathbb{E}_{p(\cdot|\mathcal{P})}[u(a(p(\cdot|\mathcal{P})), \omega)] \right\} + p(\mathcal{I})\zeta \\
\leq \ & \mathbb{E}_p[u(a(p(\cdot|\mathcal{P})), \omega)] \\
& + p(\mathcal{P}) \left\{ \max_{\tilde{a} \notin A^c} \mathbb{E}_{p(\cdot|\mathcal{P})}[u(\tilde{a}, \omega)] - \mathbb{E}_{p(\cdot|\mathcal{P})}[u(a(p(\cdot|\mathcal{P})), \omega)] \right\} + p(\mathcal{I})\zeta \\
\leq \ & \mathbb{E}_p[u(a(p(\cdot|\mathcal{P})), \omega)] + \chi + p(\mathcal{I})[\zeta - \chi],
\end{aligned}
$$

Now recall that $\sup_{p \in \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)} p(\mathcal{I}) \to 0$ as $\varepsilon \to 0$ by Corollary 3. Thus, there exists $\varepsilon_{a'} > 0$ so that if $\varepsilon < \varepsilon_{a'}$, then $\mathbb{E}_p[u(a', \omega)] < \mathbb{E}_p[u(a(p(\cdot|\mathcal{P})), \omega)]$ for every $p \in \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)$. Let $\bar{\varepsilon} = \min_{a' \in A^c} \varepsilon_{a'} > 0$. It follows that, if $\varepsilon < \bar{\varepsilon}$, then $\mathbb{E}_p[u(a', \omega)] < \mathbb{E}_p[u(a(p(\cdot|\mathcal{P})), \omega)]$ for every $a \notin A^c$ and $p \in \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)$. $\square$

By A3, if $a \in A^c$, there exists $p_a \in \Delta(\mathcal{P})$ such that $\mathbb{E}_{p_a}[u(a, \omega)] - \gamma \geq \mathbb{E}_{p_a}[u(a', \omega)]$ for some $\gamma > 0$ and all other $a' \in A^c$. Since $\sup_{p \in \mathbf{p}_\varepsilon(\mathcal{C}_\varepsilon)} p(\mathcal{I}) \to 0$ as $\varepsilon \to 0$, it follows that if $p \in \Delta(\Omega)$ is such that $p(\cdot|\mathcal{P}) = p_a$, then $a = \arg\max_{a' \in A} \mathbb{E}_p[u(a', \omega)]$ for ε small enough. If we let a^ω be a best action in A^c for $\omega \in \mathcal{I}$, then any $p \in \Delta(\Omega)$ that satisfies $p(\cdot|\mathcal{P}) = p_{a^\omega}$ guarantees that Receiver chooses a^ω provided that ε is smaller than some ε_{a^ω}. Since $\mathcal{I}$ is finite, if we let $\underline{\varepsilon} = \min_{\omega \in \mathcal{I}} \varepsilon_{a^\omega}$, then the previous conclusions holds for every $\omega \in \mathcal{I}$ provided that $\varepsilon < \underline{\varepsilon}$. Hereafter, $\varepsilon < \underline{\varepsilon}$.

Claim 13. $\liminf_{\varepsilon \to 0} V_\varepsilon^c(\sigma) \geq V_0^c(\sigma)$.

Proof. Consider any $\tau \in \mathcal{T}$ that satisfies properties (1)–(3) of Proposition 2 and $\mathbf{supp}\,\tau \subseteq \mathcal{C}_\varepsilon$. Then, for every $q^\omega \in \mathbf{supp}\,\tau$ that conceals an $\omega \in \mathcal{I}$, we can write q^ω as

$$q^\omega = (1 - q^\omega(\omega))p_{a^\omega} + q^\omega(\omega)\delta_\omega,$$

letting $q^\omega(\omega) = \frac{1}{1+\varepsilon}\eta_\varepsilon$ for all $\omega \in \mathcal{I}$, where η_ε is as in Corollary 3. It is also the case that $q^\omega(\omega)\tau(q^\omega) = \sigma(\omega)$. So, by the same argument as in Part IV of the proof of Proposition 2,

$$\mathbb{E}_\tau[\hat{v}_\varepsilon(q)] \leq \sigma(\mathcal{P})V_\varepsilon(\sigma(\cdot|\mathcal{P})) + \sum_{\omega\in\mathcal{I}} k(\omega)\sigma(\omega) = V_0^c(\sigma),$$

where the equality follows from observing $\mathbf{p}_\varepsilon$ is independent of ε for all $q \in \Delta(\mathcal{P})$: Indeed, for $q \in \Delta(\mathcal{P})$,

$$\mathbf{p}_\varepsilon(\omega; q) = \frac{\frac{q(\omega)}{\sigma(\omega)}\rho_\varepsilon(\omega)}{\sum_{\omega'\in\mathcal{P}} \frac{q(\omega')}{\sigma(\omega')}\rho_\varepsilon(\omega')} = \frac{\frac{q(\omega)}{\sigma(\omega|\mathcal{P})}\rho_\varepsilon(\omega|\mathcal{P})}{\sum_{\omega'\in\mathcal{P}} \frac{q(\omega')}{\sigma(\omega'|\mathcal{P})}\rho_\varepsilon(\omega'|\mathcal{P})},$$

which does not depend on ε by A1. This implies that

$$
\begin{aligned}
V_\varepsilon(\sigma(\cdot|\mathcal{P})) &= \sup_{\{\tau:\mathbb{E}_\tau[q]=\sigma(\cdot|\mathcal{P})\}} \mathbb{E}_\tau[\mathbb{E}_q[v(a(\mathbf{p}_\varepsilon(q)),\omega)]] \\
&= \sup_{\{\tau:\mathbb{E}_\tau[q]=\sigma(\cdot|\mathcal{P})\}} \mathbb{E}_\tau[\mathbb{E}_q[v(a(\mathbf{p}_0(q)),\omega)]] = V_0(\sigma(\cdot|\mathcal{P})).
\end{aligned}
$$

Now note that by the same observation

$$\mathbb{E}_\tau[\hat{v}_\varepsilon(q)] = \sigma(\mathcal{P})\mathbb{E}_{\tau_{\Delta(\mathcal{P})}}[\hat{v}_0(q)] + \sum_{\omega\in\mathcal{I}} k(\omega)\sigma(\omega),$$

where $\tau_{\Delta(\mathcal{P})}$ is the restriction of τ to $\Delta(\mathcal{P})$ which belongs to the set $\{\tau : \mathbb{E}_\tau[q] = \sigma(\cdot|\mathcal{P})\}$. Thus, if we construct τ so that $\mathbb{E}_{\tau_{\Delta(\mathcal{P})}}[\hat{v}_0(q)] \to V_0(\sigma(\cdot|\mathcal{P}))$ as in Part IV of the proof of Proposition 2, we get that $\mathbb{E}_\tau[\hat{v}_\varepsilon(q)] \to V_0(\sigma)$ from below.

We conclude that, for every sequence $\{\varepsilon_n\}$ converging to 0, we can construct a sequence $\{\tau_{\varepsilon_n}\}$ such that each $\tau_{\varepsilon_n} \in \mathcal{T}$ and satisfies $\mathbf{supp}\,\tau_{\varepsilon_n} \subseteq \mathcal{C}_{\varepsilon_n}$, $\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq V_{\varepsilon_n}^c(\sigma)$ (by definition), $\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq V_0^c(\sigma)$ (by construction), and $\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \to V_0^c(\sigma)$. The claim follows. $\square$

Claim 14. $\limsup_{\varepsilon\to 0} V_\varepsilon^c(\sigma) \leq V_0^c(\sigma)$.

Proof. The proof is by contradiction. Suppose $\limsup_{\varepsilon\to 0} V_\varepsilon^c(\sigma) = \overline{V} > V_0^c(\sigma)$. Fix $\gamma > 0$ so that $\overline{V} - \gamma > V_0^c(\sigma)$. There must exists a sequence $\{\varepsilon_n\}$ with $\varepsilon_n \to 0$ and a corresponding sequence $\{\tau_{\varepsilon_n}\}$ with each $\tau_{\varepsilon_n} \in \mathcal{T}$ satisfying $\mathbf{supp}\,\tau_{\varepsilon_n} \subseteq \mathcal{C}_{\varepsilon_n}$ such that $\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \in [\overline{V} - \gamma, \overline{V} + \gamma]$ for all $n \geq N$ sufficiently large.

Note that without loss of generality we can assume that, for every τ_{ε_n} in the sequence, there must exists $q \in \mathbf{supp}\,\tau_{\varepsilon_n}$ such that $a(\mathbf{p}_{\varepsilon_n}(q)) \notin A(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})))$—

otherwise, we would have $\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq V_0^c(\sigma)$. To see this, fix ε_n and note that

$$
\begin{aligned}
\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \;=\; & \sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega) q(\omega) \right. \\
& \left. + \sum_{\omega \in \mathcal{I}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega) q(\omega) \right] \tau_{\varepsilon_n}(q) \\
\leq\; & \sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega) q(\omega) + \sum_{\omega \in \mathcal{I}} k(\omega) q(\omega) \right] \tau_{\varepsilon_n}(q) \\
=\; & \sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega) q(\omega) \right] \tau_{\varepsilon_n}(q) \\
& + \sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{I}} k(\omega) q(\omega) \right] \tau_{\varepsilon_n}(q) \\
=\; & \sigma(\mathcal{P}) \sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega) q(\omega|\mathcal{P}) \right] \frac{q(\mathcal{P})\tau_{\varepsilon_n}(q)}{\sigma(\mathcal{P})} \\
& + \sum_{\omega \in \mathcal{I}} k(\omega) \sigma(\omega),
\end{aligned}
$$

where we used that $\sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} q(\omega)\tau(q) = \sigma(\omega)$ for every $\omega \in \Omega$. If we define $\tau'(q(\cdot|\mathcal{P})) = \frac{q(\mathcal{P})\tau_{\varepsilon_n}(q)}{\sigma(\mathcal{P})}$ for every $q(\cdot|\mathcal{P})$ corresponding to a $q \in \mathbf{supp}\,\tau_{\varepsilon_n}$, then τ' defines a distribution over $\Delta(\mathcal{P})$ with

$$
\mathbb{E}_{\tau'}[q(\omega)] = \frac{1}{\sigma(\mathcal{P})} \sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} q(\omega)\tau_{\varepsilon_n}(q) = \sigma(\omega|\mathcal{P}), \quad \omega \in \mathcal{P}.
$$

Recall that for ε small enough, $a(\mathbf{p}_{\varepsilon_n}(q)) \in A^c$ for all $q \in \mathbf{supp}\,\tau$. Hence, if $a(\mathbf{p}_{\varepsilon_n}(q)) \in A(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})))$ for all $q \in \mathbf{supp}\,\tau'$, we would have

$$
\begin{aligned}
\sum_{q \in \mathbf{supp}\,\tau'} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega) q(\omega|\mathcal{P}) \right] \tau'(q(\cdot|\mathcal{P})) \;\leq\; & \sum_{q \in \mathbf{supp}\,\tau'} \hat{v}_{\varepsilon_n}(q(\cdot|\mathcal{P}))\tau'(q(\cdot|\mathcal{P})) \\
\leq\; & V_{\varepsilon_n}(\sigma(\cdot|\mathcal{P})) = V_0(\sigma(\cdot|\mathcal{P})).
\end{aligned}
$$

We will proceed by constructing for every τ_{ε_n} as 'twin' distribution $\hat{\tau}_{\varepsilon_n}$ and by showing that $\lim_{n \to \infty} \left| \mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] - \mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \right| = 0$ and $\lim_{n \to \infty} \mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq V_0^c(\sigma)$, which delivers the desired contradiction.

Focussing on qs such that $a(\mathbf{p}_{\varepsilon_n}(q)) \notin A(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})))$, consider the following ob-

servations. By optimality $a(\mathbf{p}_{\varepsilon_n}(q))$ of for Receiver, $\frac{\mathbf{p}_{\varepsilon_n}(\mathcal{I};q)}{\mathbf{p}_{\varepsilon_n}(\mathcal{P};q)}\mathbb{E}_{\mathbf{p}_{\varepsilon_n}(\cdot|\mathcal{I};q)}[u(a(\mathbf{p}_{\varepsilon_n}(q)),\omega)-u(a(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))),\omega)]$ must exceed $\mathbb{E}_{\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))}[u(a(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))),\omega)-u(a(\mathbf{p}_{\varepsilon_n}(q)),\omega)]$, where we used

$$\mathbf{p}_{\varepsilon_n}(\omega|\mathcal{P};q) = \frac{\frac{q(\omega)}{\sigma(\omega)}\rho_o(\omega)}{\sum_{\omega'\in\mathcal{P}}\frac{q(\omega')}{\sigma(\omega')}\rho_o(\omega')} = \frac{\frac{q(\omega|\mathcal{P})}{\sigma(\omega)}\rho_o(\omega)}{\sum_{\omega'\in\mathcal{P}}\frac{q(\omega'|\mathcal{P})}{\sigma(\omega')}\rho_o(\omega')} = \mathbf{p}_{\varepsilon_n}(\omega;q(\cdot|\mathcal{P})).$$

Letting $M = \max_{\omega\in\mathcal{I}}[\max_{a\in A} u(a,\omega) - \min_{a\in A} u(a,\omega)] > 0$, which is bounded, it follows that

$$\begin{aligned}
B_{\varepsilon_n} &= M \sup_{p\in\mathbf{p}_{\varepsilon_n}(\mathcal{C}_{\varepsilon_n})} \frac{p(\mathcal{I})}{1-p(\mathcal{I})} \\
&\geq \mathbb{E}_{\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))}[u(a(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))),\omega) - u(a(\mathbf{p}_{\varepsilon_n}(q)),\omega)] > 0.
\end{aligned}$$

Recall that $\sup_{p\in\mathbf{p}_{\varepsilon_n}(\mathcal{C}_{\varepsilon_n})} \frac{p(\mathcal{I})}{1-p(\mathcal{I})} \to 0$ as $\varepsilon_n \to 0$. Therefore, for every ε_n and $q \in \mathcal{C}_{\varepsilon_n}$, there is a uniform upper abound $B_{\varepsilon_n} > 0$ in the payoff difference for Receiver between $a(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})))$ and $a(\mathbf{p}_{\varepsilon_n}(q))$, which both belong to A^c, and $B_{\varepsilon_n} \to 0$ as $\varepsilon_n \to 0$.

Next, for every $a \in A^c$, let $\Delta^a(\mathcal{P}) = \{p \in \Delta(\mathcal{P}) : a \in \arg\max_{a\in A}\mathbb{E}_p[u(a,\omega)]\}$. Clearly, $\Delta^a(\mathcal{P})$ is closed and convex and $int\Delta^a(\mathcal{P}) \neq \varnothing$ for every $a \in A^c$ by A3. Therefore, if for every $a \in A^c$ we fix $p_a \in int\Delta^a(\mathcal{P})$ so that $\min_{a'\in A}\mathbb{E}_{p_a}[u(a,\omega) - u(a',\omega)] > 0$ and define

$$p^\beta(a(\mathbf{p}_{\varepsilon_n}(q))) = \beta\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})) + (1-\beta)p_{a(\mathbf{p}_{\varepsilon_n}(q))},$$

then there exists a $\tilde{\beta}_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q))) < 1$ such that $\mathbb{E}_{p^\beta(a(\mathbf{p}_{\varepsilon_n}(q)))}[u(a(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))),\omega)-u(a(\mathbf{p}_{\varepsilon_n}(q)),\omega)] \geq 0$ if and only if $\beta \geq \tilde{\beta}_{\varepsilon_n}$ and with equality if and only if $\beta = \tilde{\beta}_{\varepsilon_n}$. Note that, by definition,

$$\begin{aligned}
\frac{\tilde{\beta}_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q)))}{1-\tilde{\beta}_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q)))} &= \frac{\mathbb{E}_{p_{a(\mathbf{p}_{\varepsilon_n}(q))}}[u(a(\mathbf{p}_{\varepsilon_n}(q)),\omega) - u(a(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))),\omega)]}{\mathbb{E}_{\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))}[u(a(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P}))),\omega) - u(a(\mathbf{p}_{\varepsilon_n}(q)),\omega)]}. \\
&\geq \frac{\min_{a\in A}\mathbb{E}_{p_{a(\mathbf{p}_{\varepsilon_n}(q))}}[u(a(\mathbf{p}_{\varepsilon_n}(q)),\omega) - u(a,\omega)]}{B_{\varepsilon_n}} \geq \frac{K}{B_{\varepsilon_n}}
\end{aligned}$$

where $K = \min_{a'\in A^c}\left\{\min_{a\in A}\mathbb{E}_{p_{a'}}[u(a',\omega) - u(a,\omega)]\right\} > 0$. This implies that

$$\tilde{\beta}_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q))) \geq \frac{K}{K+B_{\varepsilon_n}}.$$

Now let $\beta_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q))) = \frac{K}{(1+\varepsilon_n)(K+B_{\varepsilon_n})}$. By construction,

$$\text{(B3)} \qquad ||\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})) - p^{\beta_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q)))}(a(\mathbf{p}_{\varepsilon_n}(q)))||$$

$$= \ (1 - \beta_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q)))||\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})) - p_{a(\mathbf{p}_{\varepsilon_n}(q))}||$$

$$(\text{B4}) \qquad \leq \ \left[1 - \frac{K}{(1 + \varepsilon_n)(K + B_{\varepsilon_n})}\right]|\Omega|.$$

$\square$

Since $\mathbf{p}_{\varepsilon_n}$ is an homeomorphism independent of ε_n when restricted to $\Delta(\mathcal{P})$, there exists a unique $q'(a(\mathbf{p}_{\varepsilon_n}(q))) \in \Delta(\mathcal{P})$ such that

$$\mathbf{p}_{\varepsilon_n}(q'(a(\mathbf{p}_{\varepsilon_n}(q)))) = p^{\beta_{\varepsilon_n}(a(\mathbf{p}_{\varepsilon_n}(q)))}(a(\mathbf{p}_{\varepsilon_n}(q))).$$

Using $q'(a(\mathbf{p}_{\varepsilon_n}(q)))$, it is possible to construct $\hat{q}(a(\mathbf{p}_{\varepsilon_n}(q))) \in \Delta(\Omega)$ in the form $\hat{q}(a(\mathbf{p}_{\varepsilon_n}(q))) = q(\mathcal{P})q'(a(\mathbf{p}_{\varepsilon_n}(q))) + (1 - q(\mathcal{P}))q(\cdot|\mathcal{I})$. Now, note that

$$\mathbf{p}_{\varepsilon_n}(\cdot|\mathcal{I}; \hat{q}(a(\mathbf{p}_{\varepsilon_n}(q)))) = \mathbf{p}_{\varepsilon_n}(\cdot|\mathcal{I}; q)$$

by A1 and

$$\frac{\mathbf{p}_{\varepsilon_n}(\mathcal{P}; \hat{q})}{\mathbf{p}_{\varepsilon_n}(\mathcal{I}; \hat{q})} = \frac{\mathbf{p}_{\varepsilon_n}(\mathcal{P}; q)}{\mathbf{p}_{\varepsilon_n}(\mathcal{I}; q)}.$$

It follows that $a(\mathbf{p}_{\varepsilon_n}(q)) = \arg\max_{a \in A} \mathbb{E}_{\mathbf{p}_{\varepsilon_n}(\hat{q})}[u(a, \omega)]$ and $\hat{q}(a(\mathbf{p}_{\varepsilon_n}(q))) \in \mathcal{C}_{\varepsilon_n}$. This is because Receiver's expected payoff conditional on $\mathcal{I}$ is the same for $a(\mathbf{p}_{\varepsilon_n}(q))$ and every other strictly worse action under both $\mathbf{p}_{\varepsilon_n}(q)$ and $\mathbf{p}_{\varepsilon_n}(\hat{q})$, his expected payoff conditional on $\mathcal{P}$ strictly favors $a(\mathbf{p}_{\varepsilon_n}(q))$ to every other action under $\mathbf{p}_{\varepsilon_n}(\hat{q})$ compared to under $\mathbf{p}_{\varepsilon_n}(q)$, and the relative likelihood of $\mathcal{P}$ and $\mathcal{I}$ is the same under both beliefs.

We can now construct a twin distribution $\hat{\tau}_{\varepsilon_n}$ for every τ_{ε_n} in the sequence as follows. For every $q \in \mathbf{supp}\,\tau_{\varepsilon_n}$ such that $a(\mathbf{p}_{\varepsilon_n}(q)) \notin A(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})))$, construct $\hat{q}(a(\mathbf{p}_{\varepsilon_n}(q)))$ as before and assign to it probability $\hat{\tau}_{\varepsilon_n} = \tau_{\varepsilon_n}(q)$; if $a(\mathbf{p}_{\varepsilon_n}(q)) \in A(\mathbf{p}_{\varepsilon_n}(q(\cdot|\mathcal{P})))$, simply let $\hat{q}(a(\mathbf{p}_{\varepsilon_n}(q))) = q$ and assign to it probability $\hat{\tau}_{\varepsilon_n} = \tau_{\varepsilon_n}(q)$. In general, $\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{q}] \neq \sigma$. As we will show shortly, however, $\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{q}] \to \sigma$ as $n \to \infty$. Thus, since $\sigma \in int\Delta(\Omega)$ and $\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{q}] \in \Delta(\Omega)$ for every ε_n, we can always augment each $\hat{\tau}_{\varepsilon_n}$ by adding a realization $\tilde{q} \in \Delta(\Omega)$ so that $\tilde{q}\hat{\tau}_{\varepsilon_n}(\tilde{q}) + (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q}))\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{q}] = \sigma$ and choose $\tilde{q}$ so that $\hat{\tau}_{\varepsilon_n}(\tilde{q}) \to 0$ as $n \to \infty$.

Now, $\left|\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] - \mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)]\right|$ is bounded above by

$$\left|\sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega)[q(\omega|\mathcal{P}) - (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q}))q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q)))]\right]q(\mathcal{P})\tau_{\varepsilon_n}(q)\right|$$

$$+ \hat{\tau}_{\varepsilon_n}(\tilde{q})\left|\sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}}\left[\sum_{\omega \in \mathcal{I}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega)q(\omega|\mathcal{I})\right]q(\mathcal{I})\tau_{\varepsilon_n}(\hat{q}) - \sum_{\omega \in \Omega} v(a(\mathbf{p}_{\varepsilon_n}(\tilde{q})), \omega)\tilde{q}(\omega)\right|$$

$$\leq \sum_{q \in \mathbf{supp}\,\tau_{\varepsilon_n}} \left[M'\sum_{\omega \in \mathcal{P}}|q(\omega|\mathcal{P}) - (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q}))q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q)))|\right]q(\mathcal{P})\tau_{\varepsilon_n}(q) + \hat{\tau}_{\varepsilon_n}(\tilde{q})M''$$

$$\leq \overline{M} \left\{ \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \sum_{\omega \in \mathcal{P}} \left[|q(\omega|\mathcal{P}) - q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q)))| + \hat{\tau}_{\varepsilon_n}(\tilde{q}) q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q))) \right] q(\mathcal{P}) \tau_{\varepsilon_n}(q) + \hat{\tau}_{\varepsilon_n}(\tilde{q}) \right\}$$

$$\leq \overline{M} \left\{ \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \left[\max_{\omega \in \mathcal{P}} |q(\omega|\mathcal{P}) - q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q)))| \, q(\mathcal{P}) \tau_{\varepsilon_n}(q) \right] + \hat{\tau}_{\varepsilon_n}(\tilde{q}) [\sigma(\mathcal{P}) + 1] \right\}$$

$$\leq \overline{M} \left\{ \max_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \left[\max_{\omega \in \mathcal{P}} |q(\omega|\mathcal{P}) - q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q)))| \right] + \hat{\tau}_{\varepsilon_n}(\tilde{q}) [\sigma(\mathcal{P}) + 1] \right\},$$

where $\overline{M} = \max\{M', M''\}$ is bounded. Given this, consider the program

$$\Phi_{\varepsilon_n} = \max_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \max_{\omega \in \mathcal{P}} \left| q(\omega|\mathcal{P}) - q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q))) \right|$$

subject to

$$\|\mathbf{p}_0(q(\cdot|\mathcal{P})) - \mathbf{p}_0(q'(a(\mathbf{p}_{\varepsilon_n}(q))))\| \leq \left[1 - \frac{K}{(1 + \varepsilon_n)(K + B_{\varepsilon_n})} \right] |\Omega|,$$

where we replaced $\mathbf{p}_{\varepsilon_n}$ with $\mathbf{p}_0$ because $q'(a(\mathbf{p}_{\varepsilon_n}(q))) \in \Delta(\mathcal{P})$ for all considered qs. For each ε_n, we have $\Phi_{\varepsilon_n} \leq \hat{\Phi}_{\varepsilon_n}$, where

$$\hat{\Phi}_{\varepsilon_n} = \max_{q,q' \in \Delta(\mathcal{P})} \max_{\omega \in \mathcal{P}} \left| q(\omega) - q'(\omega) \right|$$

subject to

$$\|\mathbf{p}_0(q) - \mathbf{p}_0(q')\| \leq \left[1 - \frac{K}{(1 + \varepsilon_n)(K + B_{\varepsilon_n})} \right] |\Omega|.$$

By continuity of $\mathbf{p}_0$ and (B4), it follows that $\hat{\Phi}_{\varepsilon_n} \to 0$ as $\varepsilon_n \to 0$ and hence $\Phi_{\varepsilon_n} \to 0$.

Finally, note that

$$
\begin{aligned}
\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \;=\; & (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q})) \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(\hat{q})), \omega) \hat{q}(\omega|\mathcal{P}) \right] q(\mathcal{P}) \tau_{\varepsilon_n}(\hat{q}) \\
& + (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q})) \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{I}} v(a(\mathbf{p}_{\varepsilon_n}(\hat{q})), \omega) \hat{q}(\omega|\mathcal{I}) \right] q(\mathcal{I}) \tau_{\varepsilon_n}(\hat{q}) \\
& + \hat{\tau}_{\varepsilon_n}(\tilde{q}) \sum_{\omega \in \Omega} v(a(\mathbf{p}_{\varepsilon_n}(\tilde{q})), \omega) \tilde{q}(\omega) \\
\leq \;& (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q})) \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(\hat{q})), \omega) \hat{q}(\omega|\mathcal{P}) \right] q(\mathcal{P}) \tau_{\varepsilon_n}(\hat{q}) \\
& + (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q})) \sum_{\omega \in \mathcal{I}} k(\omega) \sigma(\omega)
\end{aligned}
$$

$$+\hat{\tau}_{\varepsilon_n}(\tilde{q}) \sum_{\omega \in \Omega} v(a(\mathbf{p}_{\varepsilon_n}(\tilde{q})), \omega)\tilde{q}(\omega).$$

Thus, in order to show that $\lim_{n\to\infty} \mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq V_0^c(\sigma)$, it remains to argue that

$$\sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \left[\sum_{\omega \in \mathcal{P}} v(a(\mathbf{p}_{\varepsilon_n}(q)), \omega)q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q))) \right] (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q}))q(\mathcal{P})\tau_{\varepsilon_n}(q)$$
$$\to V \leq V_0(\sigma(\cdot|\mathcal{P})).$$

Note that $\sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q}))q(\mathcal{P})\tau_{\varepsilon_n}(q) = (1 - \hat{\tau}_{\varepsilon_n}(\tilde{q}))\sigma(\mathcal{P})$ and therefore, if we define τ'_{ε_n} by $\tau'_{\varepsilon_n}(q) = \frac{(1-\hat{\tau}_{\varepsilon_n}(\tilde{q}))q(\mathcal{P})\tau_{\varepsilon_n}(q)}{(1-\hat{\tau}_{\varepsilon_n}(\tilde{q}))\sigma(\mathcal{P})}$, we have that τ'_{ε_n} is a distribution over $\Delta(\mathcal{P})$. Define $\mu_{\varepsilon_n} = \mathbb{E}_{\tau'_{\varepsilon_n}}[q]$. Then, recalling that $a(\mathbf{p}_{\varepsilon_n}(q)) = a(\mathbf{p}_{\varepsilon_n}(q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q)))))$, we have

$$\sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} \hat{v}_0(q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q))))\frac{q(\mathcal{P})\tau_{\varepsilon_n}(q)}{\sigma(\mathcal{P})} \leq V_0(\mu_{\varepsilon_n}).$$

If we prove that $\mu_{\varepsilon_n} \to \sigma(\cdot|\mathcal{P})$, we are done.[34] For every $\omega \in \mathcal{P}$, we have

$$|\mathbb{E}_{\tau'_{\varepsilon_n}}[q'(\omega)] - \sigma(\omega|\mathcal{P})| = \left| \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} [q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q))) - q(\omega|\mathcal{P})]\frac{q(\mathcal{P})\tau_{\varepsilon_n}(q)}{\sigma(\mathcal{P})} \right|$$
$$\leq \Phi_{\varepsilon_n},$$

which converges to zero as $\varepsilon_n \to 0$ as we saw before.

To conclude, we show that $\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{q}] \to \sigma$ as $n \to \infty$. Note that, for every $\omega \in \Omega$,

$$\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{q}(\omega)] = \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} [q(\mathcal{P})q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q))) + q(\mathcal{I})q(\omega|\mathcal{I})]\tau_{\varepsilon_n}(q)$$
$$= \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} q'(\omega; a(\mathbf{p}_{\varepsilon_n}(q)))q(\mathcal{P})\tau_{\varepsilon_n}(q)$$
$$+ \sum_{q \in \mathbf{supp}\, \tau_{\varepsilon_n}} q(\omega|\mathcal{I})q(\mathcal{I})\tau_{\varepsilon_n}(q)$$
$$= \sigma(\mathcal{P})\mu_{\varepsilon_n}(\omega) + \sigma(\mathcal{I})\sigma(\omega|\mathcal{I}).$$

Since we established that $\mu_{\varepsilon_n}(\omega) \to \sigma(\omega|\mathcal{P})$ for every $\omega \in \mathcal{P}$, it follows that $\mathbb{E}_{\hat{\tau}_{\varepsilon_n}}[\hat{q}(\omega)] \to \sigma(\omega)$ for every $\omega \in \Omega$.

Part II: $\lim_{\varepsilon \to 0} V_\varepsilon(\sigma) = V_0(\sigma)$.

Claim 15. $\liminf_{\varepsilon \to 0} V_\varepsilon(\sigma) \geq V_0(\sigma)$.

---

[34]Continuity of V_0 over $\Delta(\mathcal{P})$ can be established along the lines of Lemma 3 in Online Appendix B.B4.

Proof. By Claim 5, we can write $V(\sigma) = \tau^c V^c(q^c) + (1 - \tau^c) V(q^d)$, where $q^c \in \mathcal{C}$, $q^d \in \mathcal{D}$, $\sigma(\mathcal{P}) \leq \tau^c \leq 1$, and $\sigma = \tau^c q^c + (1 - \tau^c) q^d$. Also, since $\tau^c \in [\sigma(\mathcal{P}), 1]$, we must have $q^c(\omega) = \frac{1}{\tau^c}\sigma(\omega) \geq \sigma(\omega)$ for all $\omega \in \mathcal{P}$. This implies that for every ε,

$$\mathbf{p}_\varepsilon(\mathcal{P}; q^c) = \sum_{\omega \in \mathcal{P}} \mathbf{p}_\varepsilon(\omega; q^c) = \frac{\sum_{\omega \in \mathcal{P}} \frac{q^c(\omega)}{\sigma(\omega)} \rho_0(\omega)}{\sum_{\omega' \in \Omega} \frac{q^c(\omega')}{\sigma(\omega')} \rho_0(\omega')} \geq \frac{\rho_0(\mathcal{P})}{\sum_{\omega' \in \Omega} \frac{q^c(\omega')}{\sigma(\omega')} \rho_0(\omega')} \geq \rho_0(\mathcal{P}),$$

because $\sum_{\omega' \in \Omega} \frac{q^c(\omega')}{\sigma(\omega')} \rho_0(\omega') \leq 1$. It follows that $\mathbf{p}_\varepsilon(\mathcal{I}; q^c) \leq \rho_0(\mathcal{I}) \leq \varepsilon < \theta(\varepsilon)$. We conclude that $q^c \in \mathcal{C}_\varepsilon$ for all ε.

Therefore, for the distribution $\tau \in \mathcal{T}$ defined by τ^c, q^c, and q^d, for all ε we have that

$$\mathbb{E}_\tau[\hat{v}_\varepsilon(q)] = \tau^c \hat{v}(q^c) + (1 - \tau^c)\hat{v}(q^d) \leq \tau^c V_\varepsilon^c(q^c) + (1 - \tau^c) V_\varepsilon(q^d) \leq V_\varepsilon(\sigma).$$

Since $q^d \in \Delta(\mathcal{I}) \subseteq \mathcal{D}_\varepsilon$ for all ε, $V_\varepsilon(q^d) = V_0(q^d)$ for all ε. Now consider $V_\varepsilon^c(q^c)$. If $q^c \in \Delta(\mathcal{P})$, then $q^c = \sigma(\cdot|\mathcal{P})$ and hence $V_\varepsilon^c(q^c) = V_\varepsilon(\sigma(\cdot|\mathcal{P})) = V_0(\sigma(\cdot|\mathcal{P}))$. If instead $q^c \notin \Delta(\mathcal{P})$, then $\mathbf{supp}\, q^c \cap \mathcal{I} \neq \varnothing$. Letting $\Omega' = \mathbf{supp}\, q^c$, $\sigma' = q^c$, and $\rho_0' = \rho_0(\cdot|\Omega')$, we obtain a fictitious model equivalent to the main model, except for being characterized by $\varepsilon'(\varepsilon) < \varepsilon$ for every ε with $\varepsilon'(\varepsilon) \to 0$ as $\varepsilon \to 0$. Therefore, $V_\varepsilon^c(q^c) = V_{\varepsilon'(\varepsilon)}^c(q^c)$ for all ε. Since $V_{\varepsilon'(\varepsilon)}^c(q^c)$ must satisfy $\lim_{\varepsilon \to 0} V_{\varepsilon'(\varepsilon)}^c(q^c) = V_0^c(q^c)$ in the fictitious model with Sender's prior q^c, we have that $\lim_{\varepsilon \to 0} V_\varepsilon^c(q^c) = V_0^c(q^c)$.

Given this, we have

$$\liminf_{\varepsilon \to 0} V_\varepsilon(\sigma) \geq \liminf_{\varepsilon \to 0} [\tau^c V_\varepsilon^c(q^c) + (1 - \tau^c) V_\varepsilon(q^d)]$$
$$= \tau^c \lim_{\varepsilon \to 0} V_\varepsilon^c(q^c) + (1 - \tau^c) V_0(q^d) = V_0(\sigma).$$

$\square$

Claim 16. $\limsup_{\varepsilon \to 0} V_\varepsilon(\sigma) \leq V_0(\sigma)$.

Proof. The proof is by contradiction. Suppose that $\limsup_{\varepsilon \to 0} V_\varepsilon(\sigma) = \overline{V} > V_0(\sigma)$. Fix $\gamma > 0$ so that $\overline{V} - \gamma > V_0^c(\sigma)$. There must exists a sequence $\{\varepsilon_n\}$ with $\varepsilon_n \to 0$ and a corresponding sequence $\{\tau_{\varepsilon_n}\}$ with each $\tau_{\varepsilon_n} \in \mathcal{T}^{|\Omega|}$ satisfying $\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \in [\overline{V} - \gamma, \overline{V} + \gamma]$ for all $n \geq N$ sufficiently large, where $\mathcal{T}^{|\Omega|} = \{\tau \in \mathcal{T} : |\mathbf{supp}\, \tau_{\varepsilon_n}| = |\Omega|\}$ (see Lemma 1). Since $\lim_{n \to \infty} V_{\varepsilon_n}^c(\sigma) = V_0^c(\sigma) \leq V_0(\sigma)$. There exists N' such that $V_{\varepsilon_n}^c(\sigma) < \mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq V_{\varepsilon_n}(\sigma)$ for all $n \geq N'$. Therefore, for all $n \geq N'$, τ_{ε_n} involves disproving ρ_0 and $\sup_{n \geq N'} \tau_{\varepsilon_n}^c = \overline{\tau}^c < 1$.

For every ε_n we can write

$$\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] = \tau_{\varepsilon_n}^c \mathbb{E}_{\tau_{\varepsilon_n}(\cdot|\mathcal{C}^{\tau_{\varepsilon_n}})}[\hat{v}_{\varepsilon_n}(q)] + (1 - \tau_{\varepsilon_n}^c)\mathbb{E}_{\tau_{\varepsilon_n}(\cdot|\mathcal{D}^{\tau_{\varepsilon_n}})}[\hat{v}_{\varepsilon_n}(q)],$$

where $\mathcal{C}^{\tau_{\varepsilon_n}} = \mathcal{C}_{\varepsilon_n} \cap \mathbf{supp}\, \tau_{\varepsilon_n}$ and $\mathcal{D}^{\tau_{\varepsilon_n}} = \mathcal{D}_{\varepsilon_n} \cap \mathbf{supp}\, \tau_{\varepsilon_n}$, which is non-empty for

all $n \geq N'$. Let $q_{\varepsilon_n}^c = \mathbb{E}_{\tau_{\varepsilon_n}(\cdot | \mathcal{C}^{\tau_{\varepsilon_n}})}[q]$ and $q_{\varepsilon_n}^d = \mathbb{E}_{\tau_{\varepsilon_n}(\cdot | \mathcal{D}^{\tau_{\varepsilon_n}})}[q]$. Therefore, for all ε_n, we have

$$\mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq \tau_{\varepsilon_n}^c \mathbb{E}_{\tau_{\varepsilon_n}(\cdot | \mathcal{C}^{\tau_{\varepsilon_n}})}[\hat{v}_{\varepsilon_n}(q)] + (1 - \tau_{\varepsilon_n}^c) V_{\varepsilon_n}(q_{\varepsilon_n}^d),$$

where

$$V_{\varepsilon_n}(q_{\varepsilon_n}^d) = \max_{\{\tau : \mathbb{E}_\tau[q] = q_{\varepsilon_n}^d, \mathbf{supp}\, \tau \subseteq \mathcal{D}_{\varepsilon_n}\}} \mathbb{E}_\tau[\hat{v}_{\varepsilon_n}(q)].$$

Note that in the last expression $\hat{v}_{\varepsilon_n}(\cdot)$ is independent of ε_n (since Receiver always updates ρ^1 for all $q \in \mathcal{D}_{\varepsilon_n}$ and all ε_n) and u.s.c. (which explains the "max"). In particular, for every ε_n, we can replace $\hat{v}_{\varepsilon_n}$ with the u.s.c. function $\hat{w}$ defined by

$$\hat{w}(q) = \max_{a \in A(\mathbf{p}^*(q))} \mathbb{E}_q[v(a, \omega)], \quad q \in \Delta(\Omega),$$

where $\mathbf{p}^*$ is the function in Proposition 1 with priors σ and ρ^1.

Since $\mathcal{T}^{|\Omega|}$ is compact, $\{\tau_{\varepsilon_n}\}$ must have a converging subsequence. To avoid complicating notation, we will continue to use $\{\tau_{\varepsilon_n}\}$ for the subsequence. Denote its limit by τ_∞ with the corresponding objects τ_∞^c, q_∞^c, and q_∞^d. Note that $\tau_\infty^c \leq \overline{\tau}^c < 1$, $q_\infty^d \in \Delta(\mathcal{I})$, and $q_\infty^c \in \Delta(\Omega) \setminus \Delta(\mathcal{I})$. The latter property holds because, if $q_\infty^c \in \Delta(\mathcal{I})$, then $\tau_\infty^c q_\infty^c + (1 - \tau_\infty^c) q_\infty^d \neq \sigma$.

Consider the sequence of values $V_{\varepsilon_n}(q_{\varepsilon_n}^d)$. For every n,

$$V_{\varepsilon_n}(q_{\varepsilon_n}^d) \leq \max_{\{\tau : \mathbb{E}_\tau[q] = q_{\varepsilon_n}^d, \mathbf{supp}\, \tau \subseteq \Delta(\Omega)\}} \mathbb{E}_\tau[\hat{w}(q)] = W(q_{\varepsilon_n}^d).$$

By an argument similar to that establishing Lemma 3, one can conclude that W is continuous. Thus, $\lim_{n \to \infty} W(q_{\varepsilon_n}^d) = W(q_\infty^d)$. But since $q_\infty^d \in \Delta(\mathcal{I})$, we have $W(q_\infty^d) = V_0(q_\infty^d)$.

Now consider the sequence of values $\mathbb{E}_{\tau_{\varepsilon_n}(\cdot | \mathcal{C}^{\tau_{\varepsilon_n}})}[\hat{v}_{\varepsilon_n}(q)]$. Using an argument similar to that establishing Claim 14, we can construct a 'twin' sequence $\{\tau'_{\varepsilon_n}\}$, where $\mathbf{supp}\, \tau'_{\varepsilon_n} \subseteq \mathcal{C}_{\varepsilon_n}$ for all n, which satisfies $\lim_{n \to \infty} \left| \mathbb{E}_{\tau_{\varepsilon_n}(\cdot | \mathcal{C}^{\tau_{\varepsilon_n}})}[\hat{v}_{\varepsilon_n}(q)] - \mathbb{E}_{\tau'_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \right| = 0$ and $\lim_{n \to \infty} \mathbb{E}_{\tau'_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq V_0^c(q_\infty^c)$.

Combining these observations, we have that

$$\lim_{n \to \infty} \mathbb{E}_{\tau_{\varepsilon_n}}[\hat{v}_{\varepsilon_n}(q)] \leq \tau_\infty^c V_0^c(q_\infty^c) + (1 - \tau_\infty^c) V_0(q_\infty^d) \leq V_0(\sigma),$$

which delivers the desired contradiction. $\qquad\square$