Supplemental Material to
Misinterpreting Others and the Fragility of Social Learning

By

Mira Frick, Ryota Iijima, and Yuhta Ishii

January 2019

COWLES FOUNDATION DISCUSSION PAPER NO. 2160S

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Supplementary Appendix to "Misinterpreting Others and the Fragility of Social Learning"

Mira Frick, Ryota Iijima, and Yuhta Ishii

D Proof of Proposition C.1

D.1 Preliminary Results

The proof of Proposition C.1 in the next subsection makes use of the following three lemmas. Given any pair of states ω^k and ω^ℓ and perception $\hat{F}$, let $\delta_t(\omega^\ell, \omega^k, \hat{F}) := q_t(\omega^\ell, \hat{F}) - q_t(\omega^k, \hat{F})$ denote the difference between agents' perceived period t fraction of action 0 in states ω^ℓ and ω^k. By the same inductive argument as in Lemma A.1, $q_t(\omega, \hat{F})$ is strictly decreasing in ω at all t. Hence, $\delta_t(\omega^\ell, \omega^k, \hat{F}) < 0$ whenever $\ell > k$, as we have relabeled states $\omega^i \in \Omega_n$ to be increasing in i.

For all perceptions $\hat{F}$ in a small ball around the true distribution F, our first lemma provides a bound on the perceived fraction of agents that assign high probability to state ω^ℓ when the true state is ω^k, where this bound depends on $\hat{F}$ only through $\delta_t(\omega^\ell, \omega^k, \hat{F})$.

Lemma D.1. *Fix any $F \in \mathcal{F}$ and $0 < \xi < \min\{F(\theta^*(\omega^n)), 1 - F(\theta^*(\omega^1))\}$. There exists a weakly decreasing map $I : (0,1) \to (0,1)$ such that for any $\hat{F} \in B(F, \xi)$, $\omega^\ell \neq \omega^k$, $\underline{s} \leq \overline{s}$ and $\kappa \in (0,1)$, we have*

$$\Pr\left(\max_{s \in [\underline{s}, \overline{s}]} H_t(\omega^\ell | a^{t-1}, s, \hat{F}) \geq \kappa \mid \omega^k, \hat{F}\right) \leq \left(\frac{C}{\kappa}\right)^{\frac{1}{2}} \prod_{\tau=1}^{t-1} I(|\delta_\tau(\omega^\ell, \omega^k, \hat{F})|)$$

for all t, where $C := \max_{s \in [\underline{s}, \overline{s}]} \frac{\phi(s|\omega^\ell)}{\phi(s|\omega^k)}$.

Proof. Let $m := \min\{F(\theta^*(\omega^n)) - \xi, 1 - F(\theta^*(\omega^1)) - \xi\} \in (0,1)$. Note that $1 - m \geq F(\theta^*(\omega^1)) + \xi > F(\theta^*(\omega^n)) - \xi \geq m$. For any $\alpha \in (0, 1-2m]$, define

$$I(\alpha) := \max\left\{\left((1-q)(1-q')\right)^{\frac{1}{2}} + \left(qq'\right)^{\frac{1}{2}} : q, q' \in [m, 1-m] \text{ and } |q - q'| \geq \alpha\right\}$$

and define $I(\alpha) := I(1-2m)$ for all $\alpha \in (1-2m, 1]$. By compactness of the domain of q, q', $I(\cdot)$ is well-defined. Moreover, $I(\cdot)$ is clearly weakly decreasing. Finally, for all $q, q' \in (0,1)$ with $q \neq q'$, Jensen's inequality implies

$$\left((1-q)(1-q')\right)^{\frac{1}{2}} + \left(qq'\right)^{\frac{1}{2}} = q\left(\frac{q'}{q}\right)^{\frac{1}{2}} + (1-q)\left(\frac{1-q'}{1-q}\right)^{\frac{1}{2}} < \left(q\frac{q'}{q} + (1-q)\frac{1-q'}{1-q}\right)^{\frac{1}{2}} = 1.$$

Hence, $I(\alpha) \in (0,1)$ for all α.

Observe that for any $\hat{F} \in B(F,\xi)$ and t, we have

$$\begin{aligned}
&\Pr\left(\max_{s\in[\underline{s},\overline{s}]} H_t(\omega^\ell|a^{t-1},s,\hat{F}) \geq \kappa \mid \omega^k,\hat{F}\right) \leq \Pr\left(\max_{s\in[\underline{s},\overline{s}]} \frac{H_t(\omega^\ell|a^{t-1},s,\hat{F})}{H_t(\omega^k|a^{t-1},s,\hat{F})} \geq \kappa \mid \omega^k,\hat{F}\right) \\
&= \Pr\left(C\prod_{\tau=1}^{t-1}\left(a_\tau \frac{1-q_\tau(\omega^\ell,\hat{F})}{1-q_\tau(\omega^k,\hat{F})} + (1-a_\tau)\frac{q_\tau(\omega^\ell,\hat{F})}{q_\tau(\omega^k,\hat{F})}\right) \geq \kappa \mid \omega^k,\hat{F}\right) \\
&= \Pr\left(\prod_{\tau=1}^{t-1}\left(a_\tau \frac{1-q_\tau(\omega^\ell,\hat{F})}{1-q_\tau(\omega^k,\hat{F})} + (1-a_\tau)\frac{q_\tau(\omega^\ell,\hat{F})}{q_\tau(\omega^k,\hat{F})}\right)^{\frac{1}{2}} \geq (\kappa/C)^{\frac{1}{2}} \mid \omega^k,\hat{F}\right) \\
&\leq \left(\frac{C}{\kappa}\right)^{\frac{1}{2}} \mathbb{E}\left[\prod_{\tau=1}^{t-1}\left(a_\tau \frac{1-q_\tau(\omega^\ell,\hat{F})}{1-q_\tau(\omega^k,\hat{F})} + (1-a_\tau)\frac{q_\tau(\omega^\ell,\hat{F})}{q_\tau(\omega^k,\hat{F})}\right)^{\frac{1}{2}} \mid \omega^k,\hat{F}\right] \\
&= \left(\frac{C}{\kappa}\right)^{\frac{1}{2}} \prod_{\tau=1}^{t-1}\mathbb{E}\left[\left(a_\tau \frac{1-q_\tau(\omega^\ell,\hat{F})}{1-q_\tau(\omega^k,\hat{F})} + (1-a_\tau)\frac{q_\tau(\omega^\ell,\hat{F})}{q_\tau(\omega^k,\hat{F})}\right)^{\frac{1}{2}} \mid \omega^k,\hat{F}\right] \\
&= \left(\frac{C}{\kappa}\right)^{\frac{1}{2}} \prod_{\tau=1}^{t-1}\left(\Big((1-q_\tau(\omega^k,\hat{F})(1-q_\tau(\omega^\ell,\hat{F}))\Big)^{\frac{1}{2}} + \Big(q_\tau(\omega^k,\hat{F})q_\tau(\omega^\ell,\hat{F})\Big)^{\frac{1}{2}}\right),
\end{aligned} \tag{11}$$

where the second inequality holds by Markov's inequality, the penultimate equality holds by independence of $(a_1,\ldots,a_t)$, and the final equality holds because conditional on ω^k and $\hat{F}$, a_τ takes value 1 with probability $1-q_\tau(\omega^k,\hat{F})$ and 0 with probability $q_\tau(\omega^k,\hat{F})$.

For every τ and $\hat{F} \in B(F,\xi)$, we have $q_\tau(\omega^k,\hat{F}), q_\tau(\omega^\ell,\hat{F}) \in [\hat{F}(\theta^*(\omega^n)), \hat{F}(\theta^*(\omega^1))] \subseteq [F(\theta^*(\omega^n)) - \xi, F(\theta^*(\omega^1)) + \xi]$. Hence, by choice of m, $q_\tau(\omega^k,\hat{F}), q_\tau(\omega^\ell,\hat{F}) \in [m, 1-m]$. Moreover, $|q_\tau(\omega^k,\hat{F}) - q_\tau(\omega^\ell,\hat{F})| = |\delta_\tau(\omega^\ell,\omega^k,\hat{F})|$. Thus, combining (11) with the definition of $I(\cdot)$ yields the desired conclusion. □

Second, Lemma D.2 shows that for sufficiently large t, $\delta_t(\omega^{k+1},\omega^k,\hat{F})$ is bounded away from zero uniformly for all perceptions $\hat{F}$ in a small ball around the true distribution F.

Lemma D.2. *Fix any $F \in \mathcal{F}$. There exists $\xi > 0$ such that for all $k = 1,2,\ldots,n-1$,*

$$\limsup_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} \delta_t(\omega^{k+1},\omega^k,\hat{F}) < 0.$$

Proof. See Appendix D.3. □

Finally, Lemma D.3 uses Lemmas D.1 and D.2 to show that the perceived fraction of agents who learn the true state converges to 1 at a uniform rate under all perceptions $\hat{F}$ in a small ball around F.

Lemma D.3. *Fix any $F \in \mathcal{F}$. There exists $\xi > 0$ such that for all $k = 1,2,\ldots,n$, $\lambda \in (0,1)$, and all $\underline{s} \leq \overline{s}$,*

$$\lim_{t\to\infty} \inf_{\hat{F}\in B(F,\xi)} \Pr\left(\min_{s\in[\underline{s},\overline{s}]} H_t(\omega^k|a^{t-1},s,\hat{F}) \geq \lambda \mid \omega^k,\hat{F}\right) = 1.$$

Proof. By Lemma D.2, there exists $\xi > 0$ such that for every $k = 1, 2, \ldots, n-1$

$$g_{k+1,k}(\xi) := \limsup_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} \delta_t(\omega^{k+1}, \omega^k, \hat{F}) < 0.$$

Since $g_{k+1,k}(\cdot)$ is weakly increasing, by picking ξ sufficiently small, we can assume additionally that $\xi < \min\{F(\theta^*(\omega^n)), 1 - F(\theta^*(\omega^1))\}$. Let $g := \max\{g_{2,1}(\xi), \ldots, g_{n,n-1}(\xi)\} < 0$.

To prove the result, fix any $\kappa > 0$, $\ell \neq k$, and $\underline{s} \leq \overline{s}$. Note that if $\ell > k$, then $|\delta_t(\omega^\ell, \omega^k, \hat{F})| = |\delta_t(\omega^\ell, \omega^{k+1}, \hat{F})| + |\delta_t(\omega^{k+1}, \omega^k, \hat{F})| \geq |\delta_t(\omega^{k+1}, \omega^k, \hat{F})|$, and likewise if $\ell < k$, then $|\delta_t(\omega^\ell, \omega^k, \hat{F})| \geq |\delta_t(\omega^{\ell+1}, \omega^\ell, \hat{F})|$. Hence, by choice of ξ and m, there exists t^* such that for all $t \geq t^*$ and $\hat{F} \in B(F, \xi)$, we have $|\delta_t(\omega^\ell, \omega^k, \hat{F})| \geq |g| > 0$. Let $C := \max_{s\in[\underline{s},\overline{s}]} \frac{\phi(s|\omega^\ell)}{\phi(s|\omega^k)} > 0$ and let $I(\cdot)$ be as in Lemma D.1. Then

$$\lim_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} \Pr\left(\max_{s\in[\underline{s},\overline{s}]} H_t(\omega^\ell|a^{t-1}, s, \hat{F}) \geq \kappa \mid \omega^k, \hat{F}\right) \leq \lim_{t\to\infty} \left(\frac{C}{\kappa}\right)^{\frac{1}{2}} (I(|g|))^{t-t^*} = 0,$$

where the inequality holds by Lemma D.1 and since $I(\cdot) \in (0,1)$ is weakly decreasing. □

D.2 Completing the Proof of Proposition C.1

We now prove Proposition C.1. By Lemma D.3, there exists some $\xi^* > 0$ such that for every $\lambda \in (0,1)$, $k = 1, \ldots, n$, and $\underline{s} \leq \overline{s}$

$$\lim_{t\to\infty} \inf_{\hat{F}\in B(F,\xi^*)} \Pr\left(\min_{s\in[\underline{s},\overline{s}]} H_t(\omega^k|a^{t-1}, s, \hat{F}) \geq \lambda \mid \omega^k, \hat{F}\right) = 1. \tag{12}$$

Hence, for every λ, k, and $\underline{s} \leq \overline{s}$, we have

$$\lim_{t\to\infty} \sup_{\hat{F}\in B(F,\xi^*)} \Pr\left(\min_{s\in[\underline{s},\overline{s}]} H_t(\omega^k|a^{t-1}, s, \hat{F}) < \lambda \mid \omega^k, \hat{F}\right) = 0. \tag{13}$$

We will show that ξ^* is as required. Under any (true or perceived) type distribution G, define $\gamma^0(k, \lambda, G)$ to be the (true or perceived) share of types that strictly prefer to play action 0 (respectively, action 1) whenever their posterior assigns probability at least λ to state ω^k.[63] Note that $\gamma^0(k, 1, F) = F(\theta^*(\omega^k)) = 1 - \gamma^1(k, 1, F)$.

Then for every $\hat{F} \in B(F, \xi^*)$, λ, k, $\underline{s} \leq \overline{s}$, and t, we have the following lower and upper bounds

[63] Formally, let $\gamma^0(k, \lambda, G) := G(\underline{\theta}_{k,\lambda})$, where $\underline{\theta}_{k,\lambda}$ satisfies $\lambda u(\underline{\theta}_{k,\lambda}, \omega^k) + (1-\lambda) u(\underline{\theta}_{k,\lambda}, \omega^n) = 0$, and let $\gamma^1(k, \lambda, G) := 1 - G(\overline{\theta}_{k,\lambda})$, where $\overline{\theta}_{k,\lambda}$ satisfies $\lambda u(\overline{\theta}_{k,\lambda}, \omega^k) + (1-\lambda) u(\overline{\theta}_{k,\lambda}, \omega^1) = 0$.

for agents' perceived fraction $q_t(\omega^k, \hat{F})$ of action 0 in state ω^k:

$$
\begin{aligned}
&\Phi([\underline{s}, \overline{s}] \mid \omega^k) \Pr\left(\min_{s\in[\underline{s},\overline{s}]} H_t(\omega^k|a^{t-1}, s, \hat{F}) \geq \lambda \mid \omega^k, \hat{F}\right) \gamma^0(k, \lambda, \hat{F}) \\
&\leq q_t(\omega^k, \hat{F}) \\
&\leq (1 - \Phi([\underline{s}, \overline{s}] \mid \omega^k)) + \Phi([\underline{s}, \overline{s}] \mid \omega^k) \Pr\left(\min_{s\in[\underline{s},\overline{s}]} H_t(\omega^k|a^{t-1}, s, \hat{F}) \geq \lambda \mid \omega^k, \hat{F}\right) (1 - \gamma^1(k, \lambda, \hat{F})) \\
&+ \Phi([\underline{s}, \overline{s}] \mid \omega^k) \Pr\left(\min_{s\in[\underline{s},\overline{s}]} H_t(\omega^k|a^{t-1}, s, \hat{F}) < \lambda \mid \omega^k, \hat{F}\right).
\end{aligned}
$$

In particular, this holds for $\underline{s} \leq \overline{s}$ such that $\Phi([\underline{s}, \overline{s}] \mid \omega^k)$ is arbitrarily close to 1, so combining this with (12) and (13) yields for every λ and k that

$$
\begin{aligned}
&\inf_{\hat{F}\in B(F,\xi^*)} \gamma^0(k, \lambda, \hat{F}) \leq \lim_{t\to\infty} \inf_{\hat{F}\in B(F,\xi^*)} q_t(\omega^k, \hat{F}) \\
&\leq \lim_{t\to\infty} \sup_{\hat{F}\in B(F,\xi^*)} q_t(\omega^k, \hat{F}) \leq \sup_{\hat{F}\in B(F,\xi^*)} (1 - \gamma^1(k, \lambda, \hat{F})).
\end{aligned} \tag{14}
$$

Hence, for every λ, k, and every $\xi \leq \xi^*$,

$$
\begin{aligned}
&\lim_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} q_t(\omega^k, \hat{F}) - F(\theta^*(\omega^k)) = \lim_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} q_t(\omega^k, \hat{F}) - \gamma^0(k, 1, F) \\
&\leq \sup_{\hat{F}\in B(F,\xi)} (1 - \gamma^1(k, \lambda, \hat{F})) - \gamma^0(k, 1, F) \leq 1 - \gamma^1(k, \lambda, F) - \gamma^0(k, 1, F) + \xi,
\end{aligned}
$$

where the first inequality holds by (14) and the second inequality holds for any $\hat{F} \in B(F, \xi)$. Since this holds for all $\lambda \in (0, 1)$ and $\lim_{\lambda\to 1} \gamma^1(k, \lambda, F) = 1 - \gamma^0(k, 1, F)$, this yields

$$
\lim_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} \left(q_t(\omega^k, \hat{F}) - F(\theta^*(\omega^k)\right) \leq \xi. \tag{15}
$$

Analogously, for every λ, k, and every $\xi \leq \xi^*$,

$$
\begin{aligned}
&\lim_{t\to\infty} \inf_{\hat{F}\in B(F,\xi)} q_t(\omega^k, \hat{F}) - F(\theta^*(\omega^k)) = \lim_{t\to\infty} \inf_{\hat{F}\in B(F,\xi)} q_t(\omega^k, \hat{F}) - \gamma^0(k, 1, F) \\
&\geq \gamma^0(k, \lambda, \hat{F}) - \gamma^0(k, 1, F) \geq -\xi + \gamma^0(k, \lambda, F) - \gamma^0(k, 1, F).
\end{aligned}
$$

Again, since $\lim_{\lambda\to 1} \gamma^0(k, \lambda, F) = \gamma^0(k, 1, F)$, this yields

$$
\lim_{t\to\infty} \inf_{\hat{F}\in B(F,\xi)} \left(q_t(\omega^k, \hat{F}) - F(\theta^*(\omega^k))\right) \geq -\xi. \tag{16}
$$

Combining (15) and (16), for every $k = 1, \ldots, n$ and $\xi \leq \xi^*$,

$$
\lim_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} \left|q_t(\omega^k, \hat{F}) - F(\theta^*(\omega^k))\right| \leq \xi,
$$

as required. ∎

D.3 Proof of Lemma D.2

It remains to prove Lemma D.2. Fix any $F \in \mathcal{F}$. For any $\ell, k = 1, \ldots, n$ and $\xi > 0$, define

$$g_{\ell,k}(\xi) := \limsup_{t\to\infty} \sup_{\hat{F}\in B(F,\xi)} \delta_t(\omega^\ell, \omega^k, \hat{F}).$$

We will establish the following claim: For any $i = 1, \ldots, n-1$,

$$\exists\, \xi > 0 \text{ such that } g_{k+n-i,k}(\xi) < 0 \text{ for all } k = 1, \ldots, i. \tag{17}$$

Note that Lemma D.2 corresponds to claim (17) when $i = n - 1$. Below we prove the claim by induction on i. Our proof repeatedly applies the following lemma:

Lemma D.4. *Let $F \in \mathcal{F}$ and $1 \le \underline{k} < \overline{k} \le n$. Consider a sequence of times t_ℓ and perceptions $\hat{F}_\ell \in \mathcal{F}$ with $\|\hat{F}_\ell - F\| \to 0$. Suppose that for every $\kappa \in (0, 1)$ and s,*

$$1 = \lim_{\ell\to\infty} \Pr\left(H_{t_\ell}([\omega^1, \omega^{\underline{k}-1}]|s, a^{t_\ell-1}, \hat{F}_\ell) > \kappa \mid \omega^j, \hat{F}_\ell\right) \qquad \forall j < \underline{k}, \tag{18}$$

$$1 = \lim_{\ell\to\infty} \Pr\left(H_{t_\ell}([\omega^{\overline{k}+1}, \omega^n]|s, a^{t_\ell-1}, \hat{F}_\ell) > \kappa \mid \omega^j, \hat{F}_\ell\right) \qquad \forall j > \overline{k}, \tag{19}$$

$$1 = \lim_{\ell\to\infty} \Pr\left(H_{t_\ell}([\omega^{\underline{k}}, \omega^{\overline{k}}]|s, a^{t_\ell-1}, \hat{F}_\ell) > \kappa \mid \omega^j, \hat{F}_\ell\right) \qquad \forall j \in \left\{\underline{k}, \ldots, \overline{k}\right\}. \tag{20}$$

Then $\limsup_{\ell\to\infty} \delta_{t_\ell}(\omega^{\overline{k}}, \omega^{\underline{k}}, \hat{F}_\ell) < 0$. Moreover, if $\overline{k} < n$, $\limsup_{\ell\to\infty} \delta_{t_\ell}(\omega^{\overline{k}+1}, \omega^{\overline{k}}, \hat{F}_\ell) < 0$, and if $\underline{k} > 1$, $\limsup_{\ell\to\infty} \delta_{t_\ell}(\omega^{\underline{k}}, \omega^{\underline{k}-1}, \hat{F}_\ell) < 0$.

Proof. See Section D.3.1. ∎

We now begin our inductive proof of claim (17). For the base case $i = 1$, we need to show that there exists $\xi > 0$ such that $g_{n,1}(\xi) < 0$. Suppose to the contrary that $g_{n,1}(\xi) = 0$ for all ξ. Then we can find a sequence of times t_ℓ and perceptions $\hat{F}_\ell$ with $\|\hat{F}_\ell - F\| \to 0$ such that $\delta_{t_\ell}(\omega^n, \omega^1, \hat{F}_\ell) \to 0$. But this contradicts Lemma D.4 applied with $\overline{k} = n$ and $\underline{k} = 1$, as in this case (18)–(20) hold trivially.

For the inductive step, suppose claim (17) holds for some $i \in \{1, \ldots, n-2\}$; that is, there exists $\xi^* > 0$ such that $g := \max_{k=1}^{i} g_{k+n-i,k}(\xi^*) < 0$. We will show that (17) holds at $i+1$. For this it suffices to show that for each $k = 1, \ldots, i+1$, there exists $\xi_k > 0$ such that $g_{k+n-(i+1),k}(\xi_k) < 0$, because then setting $\bar{\xi} := \min_{k=1}^{i+1}\{\xi_k\}$, we have $\max_{k=1}^{i+1} g_{k+n-(i+1),k}(\bar{\xi}) < 0$ as required.

Suppose for a contradiction that for some $k \in \{1, \ldots, i+1\}$, we have $g_{k+n-(i+1),k}(\xi_k) = 0$ for all ξ_k. Then we can find some sequence of times t_ℓ and perceptions $\hat{F}_\ell$ such that $\|\hat{F}_\ell - F\| \to 0$ and $\delta_{t_\ell}(\omega^{k+n-(i+1)}, \omega^k, \hat{F}_\ell) \to 0$. Without loss, assume that $\hat{F}_\ell \in B(F, \xi^*)$ for all ℓ. Below we consider separately the case when $k \in \{2, \ldots, i\}$ and when $k \in \{1, i+1\}$.

Case 1: $k \in \{2,\ldots,i\}$. By choice of ξ^* and g, we have $g_{k+n-i,k}(\xi^*), g_{k+n-i-1,k-1}(\xi^*) \le g < 0$. Since for any τ and $\hat{F}$,

$$\delta_\tau(\omega^{k+n-i},\omega^k,\hat{F}) = \sum_{j=k}^{k+n-i-1} \delta_\tau(\omega^{j+1},\omega^j,\hat{F})$$
$$\delta_\tau(\omega^{k+n-i-1},\omega^{k-1},\hat{F}) = \sum_{j=k-1}^{k+n-i-2} \delta_\tau(\omega^{j+1},\omega^j,\hat{F}),$$

we must then have

$$\limsup_{\tau\to\infty} \sup_{\hat{F}\in B(F,\xi^*)} \min_{j=k,\ldots,k+n-i-1} \delta_\tau(\omega^{j+1},\omega^j,\hat{F}) \le \frac{g}{n-i} < \frac{g}{2(n-i)} < 0,$$
$$\limsup_{\tau\to\infty} \sup_{\hat{F}\in B(F,\xi^*)} \min_{j=k-1,\ldots,k+n-i-2} \delta_\tau(\omega^{j+1},\omega^j,\hat{F}) \le \frac{g}{n-i} < \frac{g}{2(n-i)} < 0.$$

Thus, there exists T^* such that for all $\tau \ge T^*$ and ℓ, we have

$$\min_{j=k,\ldots,k+n-i-1} \delta_\tau(\omega^{j+1},\omega^j,\hat{F}_\ell) \le \frac{g}{2(n-i)},$$
$$\min_{j=k-1,\ldots,k+n-i-2} \delta_\tau(\omega^{j+1},\omega^j,\hat{F}_\ell) \le \frac{g}{2(n-i)}.$$

Note that j can take only $n-i$ possible values above. Thus, by the pigeonhole principle, for each ℓ, there must exist $\overline{j}_\ell \in \{k,\ldots,k+n-i-1\}$ and $\underline{j}_\ell \in \{k-1,\ldots,k+n-i-2\}$ such that

$$\#\left\{\tau \le t_\ell : \delta_\tau(\omega^{\overline{j}_\ell+1},\omega^{\overline{j}_\ell},\hat{F}_\ell) \le \frac{g}{2(n-i)}\right\} \ge \frac{t_\ell - T^*}{n-i},$$
$$\#\left\{\tau \le t_\ell : \delta_\tau(\omega^{\underline{j}_\ell+1},\omega^{\underline{j}_\ell},\hat{F}_\ell) \le \frac{g}{2(n-i)}\right\} \ge \frac{t_\ell - T^*}{n-i}.$$

Observe that either $\overline{j}_\ell = k+n-i-1$ and $\underline{j}_\ell = k-1$, or else we can assume that $k \le \underline{j}_\ell = \overline{j}_\ell \le k+n-i-2$. Moreover, since both sequences $(\overline{j}_\ell)$ and $(\underline{j}_\ell)$ can take only finitely many values, we can restrict to a subsequence such that $\overline{j}_{\ell_m}$ and $\underline{j}_{\ell_m}$ are constant, say $\overline{j}_{\ell_m} = \overline{j}$ and $\underline{j}_{\ell_m} = \underline{j}$ for all m. Again, we either have $\overline{j} = k+n-i-1$ and $\underline{j} = k-1$, or else we can assume that $k \le \underline{j} = \overline{j} \le k+n-i-2$.

We will derive a contradiction by applying Lemma D.4 along the subsequence ℓ_m. To do so, consider any $\kappa \in (0,1)$ and s. If either $j > \overline{j} \ge \ell$ or $j \le \overline{j} < \ell$, we have

$$\begin{aligned} &\lim_{m\to\infty} \Pr\left(H_{t_{\ell_m}}(\omega^\ell | a^{t_{\ell_m}-1}, s, \hat{F}_{\ell_m}) \ge \kappa \mid \omega^j, \hat{F}_{\ell_m}\right) \\ &\le \lim_{m\to\infty} \left(\frac{C}{\kappa}\right)^{\frac{1}{2}} \prod_{\tau=1}^{t_{\ell_m}-1} I(|\delta_\tau(\omega^\ell,\omega^j,\hat{F}_{\ell_m})|) \le \lim_{m\to\infty} \left(\frac{C}{\kappa}\right)^{\frac{1}{2}} \left(I\left(\left|\frac{g}{2(n-i)}\right|\right)\right)^{\frac{t_{\ell_m}-T^*}{n-i}} = 0, \end{aligned} \tag{21}$$

where $C := \frac{\phi(s|\omega^\ell)}{\phi(s|\omega^j)}$. Here the first inequality holds by Lemma D.1, and the second holds because

$I(\cdot) \in (0,1)$ is decreasing, $|\delta_\tau(\omega^\ell, \omega^j, \hat{F}_{\ell_m})| \geq |\delta_\tau(\omega^{\overline{j}+1}, \omega^{\overline{j}}, \hat{F}_{\ell_m})|$ and by choice of $\overline{j}$ the latter exceeds $\left|\frac{g}{2(n-i)}\right|$ at more than $\frac{t_{\ell_m}-T^*}{n-i}$ periods $\tau \leq t_{\ell_m}$.

By an analogous argument using Lemma D.1 and the choice of $\underline{j}$, if either $j > \underline{j} \geq \ell$ or $j \leq \underline{j} < \ell$, we also have

$$\lim_{m\to\infty} \Pr\left(H_{t_{\ell_m}}(\omega^\ell | a^{t_{\ell_m}-1}, s, \hat{F}_{\ell_m}) \geq \kappa \mid \omega^j, \hat{F}_{\ell_m}\right) = 0. \tag{22}$$

To apply Lemma D.4, suppose first that $k \leq \underline{j} = \overline{j} \leq k+n-i-2$. In this case, set $\underline{k} = 1$ and $\overline{k} = \overline{j} > 1$. Then (21) implies that conditions (19) and (20) hold along the subsequence ℓ_m; moreover, (18) holds trivially. Thus, by Lemma D.4, $\limsup_{m\to\infty} \delta_{t_{\ell_m}}(\omega^{\overline{k}+1}, \omega^{\overline{k}}, \hat{F}_{\ell_m}) < 0$. Since $k \leq \overline{k} < \overline{k}+1 \leq k+n-(i+1)$, this implies $\limsup_{m\to\infty} \delta_{t_{\ell_m}}(\omega^{k+n-(i+1)}, \omega^k, \hat{F}_{\ell_m}) < 0$, contradicting the assumption that $\lim_{\ell\to\infty} \delta_{t_\ell}(\omega^{k+n-(i+1)}, \omega^k, \hat{F}_\ell) = 0$.

The remaining possibility is that $\underline{j} = k-1$ and $\overline{j} = k+n-i-1$. In this case, set $\underline{k} = \underline{j}+1$ and $\overline{k} = \overline{j}$. Then (21) and (22) together imply that conditions (18)–(20) hold along the subsequence ℓ_m. Thus, by Lemma D.4, $0 > \limsup_{m\to\infty} \delta_{t_{\ell_m}}(\omega^{\overline{k}}, \omega^{\underline{k}}, \hat{F}_{\ell_m}) = \limsup_{m\to\infty} \delta_{t_{\ell_m}}(\omega^{k+n-(i+1)}, \omega^k, \hat{F}_{\ell_m})$, a contradiction. This completes Case 1.

Case 2: $k = 1, i+1$. The proof when $k=1$ or $k=i+1$ follows similar lines. We briefly illustrate the case $k=1$. As in Case 1, the choice of ξ^* and g implies that $g_{k+n-i,k}(\xi^*) \leq g < 0$. Thus, following exactly the same steps as in the first four displayed equations in Case 1 (but ignoring the second line of each equation), we obtain a T^* and for each ℓ some $\overline{j}_\ell \in \{k, \ldots, k+n-i-1\}$ such that

$$\#\left\{\tau \leq t_\ell : \delta_\tau(\omega^{\overline{j}_\ell+1}, \omega^{\overline{j}_\ell}, \hat{F}_\ell) \leq \frac{g}{2(n-i)}\right\} \geq \frac{t_\ell - T^*}{n-i}.$$

As in Case 1, we can restrict to a subsequence such that $\overline{j}_{\ell_m}$ is constant, say $\overline{j}_{\ell_m} = \overline{j}$ for all m.

Consider any $\kappa \in (0,1)$ and s. Just as in Case 1, we can show using Lemma D.1 that whenever either $j > \overline{j} \geq \ell$ or $j \leq \overline{j} < \ell$, we have

$$\lim_{m\to\infty} \Pr\left(H_{t_{\ell_m}}(\omega^\ell | a^{t_{\ell_m}-1}, s, \hat{F}_{\ell_m}) \geq \kappa \mid \omega^j, \hat{F}_{\ell_m}\right) = 0. \tag{23}$$

To apply Lemma D.4, there are two possibilities to consider. Suppose first that $\overline{j} > k = 1$. Then setting $\underline{k} = 1$ and $\overline{k} = \overline{j}$, (23) implies that (19) and (20) hold along the subsequence ℓ_m; moreover, (18) holds trivially. Thus, Lemma D.4 implies $\limsup_{m\to\infty} \delta_{t_{\ell_m}}(\omega^{\overline{k}}, \omega^{\underline{k}}, \hat{F}_{\ell_m}) < 0$, which contradicts $\lim_{\ell\to\infty} \delta_{t_\ell}(\omega^{k+n-(i+1)}, \omega^k, \hat{F}_\ell) = 0$, as $1 = \underline{k} < \overline{k} \leq k+n-(i+1)$.

The only remaining possibility is that $\overline{j} = 1 = k$. In this case, set $\underline{k} = 2$ and $\overline{k} = n$. Then (23) implies that (18) and (20) hold along ℓ_m; moreover, (19) holds trivially. Thus, Lemma D.4 implies $\limsup_{m\to\infty} \delta_{t_{\ell_m}}(\omega^{\underline{k}}, \omega^{\underline{k}-1}, \hat{F}_{\ell_m}) < 0$, which again contradicts $\lim_{\ell\to\infty} \delta_{t_\ell}(\omega^{k+n-(i+1)}, \omega^k, \hat{F}_\ell) = 0$, as $k = \underline{k}-1 < \underline{k} = 2 \leq k+n-(i+1)$. □

D.3.1 Proof of Lemma D.4

Finally, we prove Lemma D.4. Consider any F, $\underline{k} < \overline{k}$, and sequences (t_ℓ) and $(\hat{F}_\ell)$ as in the statement of the lemma. Let f and $\hat{f}_\ell$ denote the densities of F and $\hat{F}_\ell$ for each ℓ.

We first prove the second and third claims. For the second claim, suppose that $\overline{k} \leq n-1$. Then note that by (19),

$$\limsup_{\ell} q_{t_\ell}(\omega^{\overline{k}+1}, \hat{F}_\ell) \leq \lim_{\kappa \to 1} \sup_{H \in \Delta(\Omega)} \left\{ F(\theta^*(H)) : H([\omega^{\overline{k}+1}, \omega^n]) > \kappa \right\} \leq F(\theta^*(\omega^{\overline{k}+1})),$$

while by (20),

$$\liminf_{\ell} q_{t_\ell}(\omega^{\overline{k}}, \hat{F}_\ell) \geq \lim_{\kappa \to 1} \inf_{H \in \Delta(\Omega)} \left\{ F(\theta^*(H)) : H([\omega^{\underline{k}}, \omega^{\overline{k}}]) > \kappa \right\} \geq F(\theta^*(\omega^{\overline{k}})).$$

Combining these two observations yields the desired conclusion that $\lim_{\ell \to \infty} \delta_{t_\ell}(\omega^{\overline{k}+1}, \omega^{\overline{k}}, \hat{F}_\ell) \leq F(\theta^*(\omega^{\overline{k}+1})) - F(\theta^*(\omega^{\overline{k}})) < 0$. For the third claim, suppose $\underline{k} > 1$. Then an analogous argument using (18) and (20) yields that $\lim_{\ell \to \infty} \delta_{t_\ell}(\omega^{\underline{k}}, \omega^{\underline{k}-1}, \hat{F}_\ell) \leq F(\theta^*(\omega^{\underline{k}})) - F(\theta^*(\omega^{\underline{k}-1})) < 0$.

It remains to prove the first claim. Suppose for a contradiction that $\limsup_{\ell \to \infty} \delta_{t_\ell}(\omega^{\overline{k}}, \omega^{\underline{k}}, \hat{F}_\ell) = 0$. Restricting to an appropriate subsequence, we can assume that $\lim_{\ell \to \infty} \delta_{t_\ell}(\omega^{\overline{k}}, \omega^{\underline{k}}, \hat{F}_\ell) = 0$.

For any s, θ, and $\kappa \in [0,1)$, consider an agent of type θ with private signal s. Let $\mathcal{H}^{\ell,0}(s,\theta,\kappa)$ (resp. $\mathcal{H}^{\ell,1}(s,\theta,\kappa)$) denote the event that at time t_ℓ, the agent's interim optimal action is 0 (resp. 1) and that he assigns probability greater than κ to states in $[\omega^{\underline{k}}, \omega^{\overline{k}}]$ given perception $\hat{F}_\ell$:

$$\mathcal{H}^{\ell,0}(s,\theta,\kappa) := \left\{ a^{t_\ell - 1} : \int u(\theta,\omega) dH_{t_\ell}(\omega | s, a^{t_\ell-1}, \hat{F}_\ell) \leq 0, H_{t_\ell}([\omega^{\underline{k}}, \omega^{\overline{k}}] | s, a^{t_\ell-1}, \hat{F}_\ell) > \kappa \right\},$$
$$\mathcal{H}^{\ell,1}(s,\theta,\kappa) := \left\{ a^{t_\ell - 1} : \int u(\theta,\omega) dH_{t_\ell}(\omega | s, a^{t_\ell-1}, \hat{F}_\ell) \geq 0, H_{t_\ell}([\omega^{\underline{k}}, \omega^{\overline{k}}] | s, a^{t_\ell-1}, \hat{F}_\ell) > \kappa \right\}.$$

Define the probabilities of these events conditional on each state ω^j and perception $\hat{F}_\ell$:

$$q_{t_\ell}(\omega^j, s, \theta, \kappa) = \Pr\left(\mathcal{H}^{\ell,0}(s,\theta,\kappa) \mid \omega^j, \hat{F}_\ell \right), r_{t_\ell}(\omega^j, s, \theta, \kappa) = \Pr\left(\mathcal{H}^{\ell,1}(s,\theta,\kappa) \mid \omega^j, \hat{F}_\ell \right).$$

The remainder of the proof proceeds in two steps. The first step is to establish that these probabilities become "flat" in states ω and signals s as $\ell \to \infty$. The second step obtains a contradiction by considering agents' ex-ante expected payoffs conditional on the events $\mathcal{H}^{\ell,0}(s,\theta,\kappa), \mathcal{H}^{\ell,1}(s,\theta,\kappa)$.

Step 1. For the remainder of the proof, fix any $\kappa \in (0,1)$. Observe that for almost every s and θ,

$$0 = \lim_{\ell \to \infty} r_{t_\ell}(\omega^j, s, \theta, \kappa) = \lim_{\ell \to \infty} q_{t_\ell}(\omega^j, s, \theta, \kappa) \ \forall j \notin \left\{ \underline{k}, \dots, \overline{k} \right\}, \tag{24}$$

$$1 = \lim_{\ell \to \infty} r_{t_\ell}(\omega^j, s, \theta, \kappa) + q_{t_\ell}(\omega^j, s, \theta, \kappa) \ \forall j \in \left\{ \underline{k}, \dots, \overline{k} \right\}, \tag{25}$$

$$0 = \lim_{\ell \to \infty} \left| q_{t_\ell}(\omega^j, s, \theta, 0) - q_{t_\ell}(\omega^j, s, \theta, \kappa) \right| \ \ \forall j \in \left\{ \underline{k} \dots, \overline{k} \right\}. \tag{26}$$

where the first line follows from (18) and (19), and the second and third lines follow from (20).

By restricting to an appropriate subsequence, we can assume that for each $\omega \in \Omega_n$, $q_{t_\ell}(\omega,\cdot,\cdot,\kappa)$ converges in the weak-star topology to some L^∞ function $q_\infty(\omega,\cdot,\cdot)$.[64] Indeed, each $q_{t_\ell}(\omega,\cdot,\cdot,\kappa)$ is an L^∞ function of s and θ, and by the Banach-Alaoglu theorem, the unit ball in L^∞ is compact under the weak-star topology. Likewise, we can assume that $q_{t_\ell}(\omega,\cdot,\cdot,0)$ is weakly convergent.

For all $j \notin \{\underline{k},\dots,\overline{k}\}$, (24) implies that $q_\infty(\omega^j,s,\theta) = 0$ for almost every s and θ. For all $j \in \{\underline{k},\dots,\overline{k}\}$, (26) ensures that $q_\infty(\omega^j,\cdot,\cdot)$ coincides with the weak limit of $q_{t_\ell}(\omega,\cdot,\cdot,0)$. Hence, since $q_{t_\ell}(\omega^j,s,\theta,0)$ is weakly decreasing in s, θ, and j for each ℓ, it follows that $q_\infty(\omega^j,s,\theta)$ is weakly decreasing in s, θ, and $j \in \{\underline{k},\dots,\overline{k}\}$.

We now claim that for almost every θ, $q_\infty(\omega^j,s,\theta)$ is constant in s and $j \in \{\underline{k},\dots,\overline{k}\}$. To see this, note that

$$\begin{aligned}
\lim_{\ell\to\infty} \delta_{t_\ell}(\omega^{\overline{k}},\omega^{\underline{k}},\hat{F}_\ell) &= \lim_{\ell\to\infty} \int\int \left(q_{t_\ell}(\omega^{\overline{k}},s,\theta,0)\phi(s \mid \omega^{\overline{k}}) - q_{t_\ell}(\omega^{\underline{k}},s,\theta,0)\phi(s \mid \omega^{\underline{k}})\right) \hat{f}_\ell(\theta)\,ds\,d\theta \\
&= \int\int \left(q_\infty(\omega^{\overline{k}},s,\theta)\phi(s \mid \omega^{\overline{k}}) - q^0_\infty(\omega^{\underline{k}},s,\theta)\phi(s \mid \omega^{\underline{k}})\right) f(\theta)\,ds\,d\theta \\
&\leq \int\int \left(q_\infty(\omega^{\overline{k}},s,\theta) - q_\infty(\omega^{\underline{k}},s,\theta)\right)\phi(s \mid \omega^{\underline{k}}) f(\theta)\,ds\,d\theta \leq 0
\end{aligned}$$

where the second equality holds by weak-star convergence of q_{t_ℓ} to q_∞ and because $\|\hat{F}_\ell - F\| \to 0$ by assumption, and the inequalities follow from the fact that $q_\infty(\omega^j,s,\theta)$ is decreasing in s and j. Since we assumed that $\lim_{\ell\to\infty}\delta_{t_\ell}(\omega^{\overline{k}},\omega^{\underline{k}},\hat{F}_\ell) = 0$, the above inequalities hold with equality and thus $q_\infty(\omega^j,s,\theta)$ is constant in $j \in \{\underline{k},\dots,\overline{k}\}$ for almost every (s,θ). As a result,

$$0 = \int\int q_\infty(\omega^{\underline{k}},s,\theta)\left(\phi(s \mid \omega^{\overline{k}}) - \phi(s \mid \omega^{\underline{k}})\right) f(\theta) ds d\theta.$$

Hence, for almost all θ, there exists some $q_\infty(\theta)$ such that $q_\infty(\omega^j,s,\theta)$ for all $j \in \{\underline{k},\dots,\overline{k}\}$ and almost all s.

Step 2. Pick signals $\overline{s} > \underline{s}$ such that there exists some type θ^* that conditional on the event that $\omega \in \{\omega^{\underline{k}},\dots,\omega^{\overline{k}}\}$ prefers action 1 following signal $\overline{s}$, but prefers action 0 following signal $\underline{s}$. Thus, $\int_{\omega^{\underline{k}}}^{\omega^{\overline{k}}} u(\theta^*,\omega)\,dH(\omega|\overline{s}) > 0 > \int_{\omega^{\underline{k}}}^{\omega^{\overline{k}}} u(\theta^*,\omega)\,dH(\omega|\underline{s})$, where $H(\cdot|\overline{s})$ and $H(\cdot|\underline{s})$ denote the Bayesian updates of the prior following $\overline{s}$ and $\underline{s}$. Then picking $\alpha > 0$ sufficiently small, we can assume that the set of types

$$\bar{\Theta} := \left\{\theta : \int_{\underline{k}}^{\overline{k}} u(\theta,\omega)\,dH(\omega|\overline{s}) > \alpha > -\alpha > \int_{\underline{k}}^{\overline{k}} u(\theta,\omega)\,dH(\omega|\underline{s})\right\}$$

satisfies $\int_{\bar{\Theta}} f(\theta)d\theta > 0$. We will derive a contradiction of this with Step 1.

Consider any $\theta \in \bar{\Theta}$, $s \in [\overline{s},\overline{s}{+}1]$ and $a^{t_\ell-1} \in \mathcal{H}^{\ell,0}(s,\theta,\kappa)$. By definition, $0 \geq \int u(\theta,\omega)\,dH(\omega|s,a^{t_\ell-1},\hat{F}_\ell)$; i.e., action 0 is interim optimal for θ following s and $a^{t_\ell-1}$. But then, given signal s, playing action

[64] That is, $\int\int q_{t_\ell}(\omega,s,\theta,\kappa)h(s,\theta)\,ds\,d\theta \to \int\int q_\infty(\omega^j,s,\theta)h(s,\theta)\,ds\,d\theta$ for any L^1 function $h(s,\theta)$.

0 following each $a^{t_\ell-1} \in \mathcal{H}^{\ell,0}(s,\theta,\kappa)$ must also yield a higher ex ante payoff than playing action 1; that is, $0 \geq \int \Pr[a^{t_\ell-1}|\omega, \hat{F}_\ell]u(\theta,\omega)\, dH(\omega|s)$. Summing over all $a^{t_\ell-1}$ in $\mathcal{H}^{\ell,0}(s,\theta,\kappa)$, this implies $0 \geq \int q_{t_\ell}(\omega,s,\theta,\kappa)u(\theta,\omega)\, dH(\omega|s)$. Integrating across all $\theta \in \bar{\Theta}$ and $s \in [\overline{s}, \overline{s}+1]$ yields

$$0 \geq \int_{\bar{\Theta}} \int_{\overline{s}}^{\overline{s}+1} \int_{\Omega^n} q_{t_\ell}(\omega,s,\theta,\kappa)u(\theta,\omega)\, dH(\omega|s) f(\theta)\, d\theta\, ds.$$

Taking the limit as $\ell \to \infty$ and using the fact that by Step 1, for almost all θ and s, $q_\infty(\omega^j,s,\theta) = q_\infty(\theta)$ if $j \in \{\underline{k},\dots,\overline{k}\}$ and $q_\infty(\omega^j,s,\theta) = 0$ if $j \notin \{\underline{k},\dots,\overline{k}\}$, this implies

$$0 \geq \int\limits_{\bar{\Theta}} q_\infty(\theta) \int\limits_{\overline{s}}^{\overline{s}+1} \int_{\underline{k}}^{\overline{k}} u(\theta,\omega)\, dH(\omega|s) f(\theta)\, d\theta\, ds \geq \alpha \int\limits_{\bar{\Theta}} q_\infty(\theta) f(\theta)\, d\theta, \tag{27}$$

where the second inequality holds by definition of $\bar{\Theta}$.

Considering $\theta \in \bar{\Theta}$, $s \in [\underline{s}-1, \underline{s}]$ and $a^{t_\ell-1} \in \mathcal{H}^{\ell,1}(s,\theta,\kappa)$ and proceeding in an analogous manner to the previous paragraph yields

$$0 \leq \int_{\bar{\Theta}} \int_{\underline{s}-1}^{\underline{s}} \int_{\Omega^n} r_{t_\ell}(\omega,s,\theta,\kappa)u(\theta,\omega)\, dH(\omega|s) f(\theta)\, d\theta\, ds.$$

Taking the limit as $\ell \to \infty$ and using Step 1 along with (24)–(25) and the definition of $\bar{\Theta}$, we obtain

$$0 \leq \int\limits_{\bar{\Theta}} (1-q_\infty(\theta)) \int\limits_{\underline{s}-1}^{\underline{s}} \int_{\underline{k}}^{\overline{k}} u(\theta,\omega)\, dH(\omega|s) f(\theta)\, d\theta\, ds \leq -\alpha \int\limits_{\bar{\Theta}} (1-q_\infty(\theta)) f(\theta)\, d\theta. \tag{28}$$

Combining (27) and (28) implies

$$\int\limits_{\bar{\Theta}} f(\theta)\, d\theta \leq \int\limits_{\bar{\Theta}} q_\infty(\theta) f(\theta)\, d\theta \leq 0,$$

which is a contradiction. This concludes the proof of the first claim. □

E Proofs for Section 6

E.1 Proof of Proposition 1

Fix any $\alpha = \hat{\alpha} < 1$. We proved part 1 of Proposition 1 in Appendix A. To show the second part, define for each $F, \hat{F} \in \mathcal{F}$ and $\omega \in \Omega$ the set of steady states

$$\mathrm{SS}(F,\hat{F},\omega) := \{\hat{\omega}_\infty \in \Omega : \hat{\omega}_\infty \in \operatorname*{argmin}_{\hat{\omega}\in\Omega} \mathrm{KL}\left(\alpha F(\theta^*(\hat{\omega}_\infty)) + (1-\alpha)F(\theta^*(\omega)), \hat{F}(\theta^*(\hat{\omega}))\right). \tag{29}$$

The following lemma shows that whenever $\mathrm{SS}(F, \hat{F}, \omega)$ is finite, incorrect agents' long-run beliefs correspond to steady states.

Lemma E.1. *Fix any $F, \hat{F}$ such that $\mathrm{SS}(F, \hat{F}, \omega)$ is finite for each ω. Then in all states ω, there exists some state $\hat{\omega}_\infty(\omega) \in \mathrm{SS}(F, \hat{F}, \omega)$ such that almost all incorrect agents' beliefs converge to a point mass on $\hat{\omega}_\infty(\omega)$.*

Proof. Since Lemma B.2 continues to characterize incorrect agents' inferences from observed actions, the proof proceeds in an analogous manner to that of Proposition B.1. Specifically, let $q_t^C(\omega)$, $q_t^I(\omega) \in [0,1]$ denote the actual fraction of action 0 among correct and incorrect agents in period t and state ω, and let $\overline{q}_t^C(\omega) := \frac{1}{t}\sum_{\tau=1}^t q_\tau^C(\omega)$ and $\overline{q}_t^I(\omega) := \frac{1}{t}\sum_{\tau=1}^t q_\tau^I(\omega)$ denote the corresponding time averages.

Note that since by the first part of Proposition 1 almost all correct agents learn the true state as $t \to \infty$, it follows that $\lim_{t\to\infty} \overline{q}_t^C(\omega) = \lim_{t\to\infty} q_t^C(\omega) = F(\theta^*(\omega))$ for all ω. Moreover, since $\mathrm{SS}(F, \hat{F}, \omega, \alpha)$ is finite, we can follow the same argument as in the proof of Lemma B.3 to show (using Lemma B.2) that the limit $R^I(\omega) := \lim_{t\to\infty} \overline{q}_t^I(\omega)$ exists for all ω.

For each ω, let

$$\hat{\omega}_\infty(\omega) := \operatorname*{argmin}_{\hat{\omega}\in\Omega} \mathrm{KL}\left(\alpha R^I(\omega) + (1-\alpha)F(\theta^*(\omega)), \hat{F}(\theta^*(\hat{\omega}))\right).$$

Then by the same argument as in the proof of Proposition B.1, we obtain that conditional on each state ω, almost all incorrect agents' beliefs converge to a point mass on $\hat{\omega}_\infty(\omega)$. But then $R^I(\omega) = F(\theta^*(\hat{\omega}_\infty(\omega)))$, whence $\hat{\omega}_\infty(\omega) \in \mathrm{SS}(F, \hat{F}, \omega)$. □

Combined with Lemma E.1, the following lemma completes the proof of the proposition.

Lemma E.2. *Fix any analytic $F \in \mathcal{F}$ and $\delta > 0$. There exists $\varepsilon > 0$ such that for any analytic $\hat{F} \neq F$ with $\|F - \hat{F}\| < \varepsilon$ and every $\omega \in \Omega$:*

1. *$\mathrm{SS}(F, \hat{F}, \omega)$ is finite.*
2. *$|\omega - \hat{\omega}| < \delta$ for every $\hat{\omega} \in \mathrm{SS}(F, \hat{F}, \omega)$.*

Proof. Fix any analytic $F \in \mathcal{F}$ and $\delta > 0$, where we can assume that $\delta < \frac{\overline{\omega}-\underline{\omega}}{2}$. Choose $\varepsilon > 0$ sufficiently small such that $\frac{\varepsilon}{1-\alpha} < |F(\theta^*(\omega)) - F(\theta^*(\omega'))|$ for any pair of states ω, ω' with $|\omega - \omega'| \geq \delta$.

Consider any analytic $\hat{F} \neq F$ with $\|F - \hat{F}\| < \varepsilon$ and any ω. By (29), each $\hat{\omega} \in \mathrm{SS}(F, \hat{F}, \omega)$ satisfies one of the following three cases:

1. $\hat{\omega} \in (\underline{\omega}, \overline{\omega})$ and $\alpha F(\theta^*(\hat{\omega})) + (1-\alpha)F(\theta^*(\omega)) = \hat{F}(\theta^*(\hat{\omega}))$
2. $\hat{\omega} = \overline{\omega}$ and $\alpha F(\theta^*(\overline{\omega})) + (1-\alpha)F(\theta^*(\omega)) \leq \hat{F}(\theta^*(\overline{\omega}))$
3. $\hat{\omega} = \underline{\omega}$ and $\alpha F(\theta^*(\underline{\omega})) + (1-\alpha)F(\theta^*(\omega)) \geq \hat{F}(\theta^*(\underline{\omega}))$.

We first show that $|\omega - \hat{\omega}| < \delta$ for all $\hat{\omega} \in \mathrm{SS}(F, \hat{F}, \omega)$. We consider only the first case, as the remaining cases are analogous. Note that

$$\alpha F(\theta^*(\hat{\omega})) + (1-\alpha) F(\theta^*(\omega)) = \hat{F}(\theta^*(\hat{\omega})) \quad \Leftrightarrow \quad F(\theta^*(\omega)) - \hat{F}(\theta^*(\hat{\omega})) = \frac{\alpha}{1-\alpha}(\hat{F}(\theta^*(\hat{\omega})) - F(\theta^*(\hat{\omega}))),$$

so that $|F(\theta^*(\omega)) - \hat{F}(\theta^*(\hat{\omega}))| \leq \frac{\alpha}{1-\alpha}\varepsilon$. Thus,

$$|F(\theta^*(\omega)) - F(\theta^*(\hat{\omega}))| \leq |F(\theta^*(\omega)) - \hat{F}(\theta^*(\hat{\omega}))| + |\hat{F}(\theta^*(\hat{\omega})) - F(\theta^*(\hat{\omega}))| \leq \frac{\alpha}{1-\alpha}\varepsilon + \varepsilon = \frac{\varepsilon}{1-\alpha}.$$

By choice of ε, this implies $|\omega - \hat{\omega}| < \delta$.

To show that $\mathrm{SS}(F, \hat{F}, \omega)$ is finite, it suffices to show that the equality $\alpha F(\theta^*(\hat{\omega})) + (1 - \alpha) F(\theta^*(\omega)) = \hat{F}(\theta^*(\hat{\omega}))$ admits at most finitely many solutions $\hat{\omega} \in [\underline{\omega}, \overline{\omega}]$. Since F and $\hat{F}$ are analytic and $[\underline{\omega}, \overline{\omega}]$ is compact, if this equality admits infinitely many solutions, then $\alpha F(\theta^*(\hat{\omega})) + (1 - \alpha) F(\theta^*(\omega)) = \hat{F}(\theta^*(\hat{\omega}))$ holds for all $\hat{\omega} \in [\underline{\omega}, \overline{\omega}]$. But the latter is impossible since we have shown that $|\omega - \hat{\omega}| < \delta < \frac{\overline{\omega} - \underline{\omega}}{2}$ holds for any solution $\hat{\omega}$. ◻

E.2 Proof of Proposition 2

Fix any $F \in \mathcal{F}$, $\hat{\omega} \in \Omega$, $\hat{\alpha}, \alpha > 0$ with $\hat{\alpha} \neq \alpha$ and $\varepsilon > 0$. If $\hat{\alpha} < \alpha$, take $\hat{F} \in \mathcal{F}$ such that $\hat{F} - F$ crosses zero only once at $\theta^*(\hat{\omega})$ from below. If $\hat{\alpha} > \alpha$, take $\hat{F} \in \mathcal{F}$ such that $\hat{F} - F$ crosses zero only once at $\theta^*(\hat{\omega})$ from above. In either case we can additionally require that $\|F - \hat{F}\| < \varepsilon$, as in the proof of Theorem 1. In addition, we can take $\hat{F}$ sufficiently close to F such that the inverse function $F \circ \hat{F}^{-1}$ has a Lipschitz constant less than $\frac{1}{\hat{\alpha}}$.

Let $\hat{q}_t^I(\omega)$ and $\hat{q}_t^C(\omega)$ denote incorrect and quasi-correct agents' perceived population fractions of action 0 in period t and state ω. The proof of Lemma 1 applied to incorrect agents' perceptions implies that $\hat{q}_t^I(\omega)$ is strictly decreasing in ω with $\hat{q}_\infty^I(\omega) := \lim_{t\to\infty} \hat{q}_t^I(\omega) = \hat{F}(\theta^*(\omega))$. Likewise, the proof of Proposition 1 applied to quasi-correct agents' perceptions implies that $\hat{q}_\infty^C(\omega) := \lim_{t\to\infty} \hat{q}_t^C(\omega)$ exists, is strictly decreasing, and satisfies

$$\hat{q}_\infty^C(\omega) = \hat{\alpha} F(\theta^*(\hat{\omega}_\omega)) + (1-\hat{\alpha}) F(\theta^*(\omega)) \quad \text{where } \hat{\omega}_\omega = \operatorname*{argmin}_{\hat{\omega}'} \mathrm{KL}\left(\hat{q}_\infty^C(\omega), \hat{F}(\theta^*(\hat{\omega}'))\right). \tag{30}$$

Lemma E.3. *If $\hat{\alpha} < \alpha$ (resp. $\hat{\alpha} > \alpha$), then $\hat{F}(\theta^*(\omega)) - q_\infty^C(\omega)$ crosses zero only once from below (resp. above) at $\omega = \hat{\omega}$.*

Proof. Note that since by construction of $\hat{F}$ the Lipschitz constant of the the RHS of (30) is less than 1, there is a unique solution $\hat{q}_\infty^C(\omega)$ to (30). Given this, we have $\hat{q}_\infty^C(\hat{\omega}) = \hat{F}(\theta^*(\hat{\omega}))$ as $F(\theta^*(\hat{\omega})) = \hat{F}(\theta^*(\hat{\omega}))$. For the remaining claim, we focus on the case $\hat{\alpha} < \alpha$ as the case $\hat{\alpha} > \alpha$ follows a symmetric argument.

Take any $\omega < \hat{\omega}$. Then $\hat{q}_\infty^C(\omega) > \hat{q}_\infty^C(\hat{\omega}) = \hat{F}(\theta^*(\hat{\omega}))$, so that $\hat{\omega}_\omega = \operatorname{argmin}_{\hat{\omega}'} \mathrm{KL}\left(\hat{q}_\infty^C(\omega), \hat{F}(\theta^*(\hat{\omega}'))\right)$ must satisfy $\hat{\omega}_\omega < \omega$ and $\hat{F}(\theta^*(\hat{\omega}_\omega)) \leq \hat{q}_\infty^C(\omega)$. But since $F(\theta) < \hat{F}(\theta)$ for all $\theta > \theta^*(\hat{\omega})$, this implies $F(\theta^*(\hat{\omega}_\omega)) \in (F(\theta^*(\hat{\omega})), \hat{q}_\infty^C(\omega))$. Since by (30), $\hat{q}_\infty^C(\omega) = \hat{\alpha} F(\theta^*(\hat{\omega}_\omega)) + (1-\hat{\alpha}) F(\theta^*(\omega))$, this implies

$F(\theta^*(\hat{\omega}_\omega)) < \hat{q}^C_\infty(\omega) < F(\theta^*(\omega)) < \hat{F}(\theta^*(\omega))$, as required. Likewise if $\omega > \hat{\omega}$, then an analogous argument shows $\hat{q}^C_\infty(\omega) > \hat{F}(\theta^*(\omega))$. □

Let $q_t(\omega)$ denote the actual population fraction of action 0 in period t at state ω, and let $\bar{q}_t(\omega) := \frac{1}{t}\sum_{\tau=1}^t q_\tau(\omega)$ be its time average. The following lemma uses a similar argument as in Lemma B.3 to show that $\bar{q}_t$ converges to $F(\theta^*(\hat{\omega}))$.

Lemma E.4. *For every ω, $\lim_{t\to\infty} \bar{q}_t(\omega) = F(\theta^*(\hat{\omega}))$.*

Proof. Fix any ω. Let $\overline{R}(\omega) := \limsup_{t\to\infty} \bar{q}_t(\omega)$ and $\underline{R}(\omega) := \liminf_{t\to\infty} \bar{q}_t(\omega)$. Suppose for a contradiction that either $\overline{R}(\omega) > F(\theta^*(\hat{\omega}))$ or $\underline{R}(\omega) < F(\theta^*(\hat{\omega}))$. We consider only the first case, as the second case is analogous.

Consider any $R \in (F(\theta^*(\hat{\omega}), \overline{R}(\omega)]$. We first claim that in state ω and any period t if (i) almost all incorrect agents' beliefs assign probability 1 to $\hat{\omega}^I := \text{argmin}_{\hat{\omega}'} \text{KL}(R, \hat{F}(\theta^*(\hat{\omega}')))$ and (ii) almost all quasi-correct agents' beliefs assign probability 1 to $\hat{\omega}^C := \text{argmin}_{\hat{\omega}'} \text{KL}(R, \hat{q}^C_\infty(\hat{\omega}'))$, then $q_t(\omega) < R$.

To show this claim, we consider only the case $\hat{\alpha} < \alpha$, as the case $\hat{\alpha} > \alpha$ is analogous. By Lemma E.3, $\hat{q}^C_\infty(\omega) > \hat{F}(\theta^*(\hat{\omega}))$ iff $\omega < \hat{\omega}$. Hence, we have $\hat{\omega}^C < \hat{\omega}$ since $R > F(\theta^*(\hat{\omega})) = \hat{F}(\theta^*(\hat{\omega}))$. Likewise, $\hat{\omega}^I < \hat{\omega}$. Thus, since $\hat{F}(\theta^*(\omega)) > \hat{q}^C_\infty(\omega)$ for all $\omega < \hat{\omega}$, it follows that $\hat{\omega} > \hat{\omega}^I > \hat{\omega}^C$.

By definition of $\hat{\omega}^C$, this leaves two cases to consider:

1. $R = \hat{q}^C_\infty(\hat{\omega}^C)$
2. $R > \hat{q}^C_\infty(\hat{\omega}^C)$ and $\hat{\omega}^C = \underline{\omega}$.

In either case, $q_t(\omega) = \alpha F(\theta^*(\hat{\omega}^I)) + (1-\alpha)F(\theta^*(\hat{\omega}^C))$. Moreover, in case 1, (30) implies $R = \hat{\alpha}F(\theta^*(\hat{\omega}^I)) + (1-\hat{\alpha})F(\theta^*(\hat{\omega}^C))$, so that $R > q_t(\omega)$ because $\hat{\alpha} < \alpha$ and $\hat{\omega}^I > \hat{\omega}^C$.

For case 2, we can extend the domain of function $\hat{q}^C_\infty$ from Ω to $\mathbb{R}$ by first extending the domain of function θ^* from Ω to $\mathbb{R}$ (in such a way that θ^* is still continuous, strictly decreasing, and has full range) and then defining $\hat{q}^C_\infty$ by (30) on the whole of $\mathbb{R}$. It is easy to show (using the same argument as above) that the extended $\hat{q}^C_\infty$ continues to satisfy Lemma E.3. Choosing $\tilde{\omega}^C < \overline{\omega}$ such that $R = \hat{q}^C_\infty(\tilde{\omega}^C)$ yields

$$R = \hat{\alpha}F(\theta^*(\hat{\omega}^I)) + (1-\hat{\alpha})F(\theta^*(\tilde{\omega}^C)) > \hat{\alpha}F(\theta^*(\hat{\omega}^I)) + (1-\hat{\alpha})F(\theta^*(\hat{\omega}^C)),$$

where the equality holds by (30) and the inequality holds since $\hat{\omega}^C = \underline{\omega}$. Thus, we again have $R > q_t(\omega)$ because $\hat{\alpha} < \alpha$ and $\hat{\omega}^I > \hat{\omega}^C$.

As a result, by continuity of u and F, there exist signals $\underline{s} < \overline{s}$, intervals of states $E^I \ni \hat{\omega}^I, E^C \ni \hat{\omega}^C$ with non-empty interior, and $\gamma > 0$ such that in state ω and any period t if (i') at least fraction $1-\gamma$ of incorrect agents with private signals $s \in [\underline{s}, \overline{s}]$ hold beliefs such that $H_t(E^I|a^{t-1}, s) \geq 1-\gamma$ and (ii') at least fraction $1-\gamma$ of quasi-correct agents with private signals $s \in [\underline{s}, \overline{s}]$ hold beliefs such that $H_t(E^C|a^{t-1}, s) \geq 1-\gamma$, then $q_t(\omega) < R - \gamma$.

To complete the proof, we consider separately the case where $\overline{R}(\omega) > \underline{R}(\omega)$ and the case where $\overline{R}(\omega) = \underline{R}(\omega)$. In the former case, we can choose $R \in (F(\theta^*(\hat{\omega}), \overline{R}(\omega))$ that additionally satisfies

$R > \underline{R}(\omega)$. Then following a similar argument as in the proof of Lemma B.3 leads to a contradiction. Specifically, for any sufficiently small $\eta > 0$, by definition of $\overline{R}(\omega), \underline{R}(\omega)$ and since $|\bar{q}_t(\omega) - \bar{q}_{t-1}(\omega)| < \eta$ for all large enough t, we can find an infinite sequence of times t_k such that $R - \frac{\eta}{2} \leq \bar{q}_{t_k-1}(\omega) \leq R + \frac{\eta}{2} < \bar{q}_{t_k}(\omega)$. Moreover, by choosing η small enough, the law of large numbers together with Lemma B.2 implies that for all large enough t_k hypotheses (i)' and (ii)' are satisfied. But then $q_{t_k}(\omega) < R - \gamma < R + \frac{\eta}{2}$, so that $\bar{q}_{t_k}(\omega) = \frac{t_k - 1}{t_k}\bar{q}_{t_k-1}(\omega) + \frac{1}{t_k} q_{t_k}(\omega) < R + \frac{\eta}{2}$, a contradiction.

Finally, if $\overline{R}(\omega) = \underline{R}(\omega)$, then we choose $R = \overline{R}(\omega) = \underline{R}(\omega) > F(\theta^*(\hat{\omega}))$. In this case, by the law of large numbers and Lemma B.2, almost all incorrect agents' beliefs converge to a point-mass on $\hat{\omega}^I := \operatorname{argmin}_{\hat{\omega}'} \mathrm{KL}(R, \hat{F}(\theta^*(\hat{\omega}')))$, and almost all quasi-correct agents' beliefs converge to a point-mass on $\hat{\omega}^C := \operatorname{argmin}_{\hat{\omega}'} \mathrm{KL}(R, \hat{q}^C_\infty(\hat{\omega}'))$. Thus, hypotheses (i') and (ii') are satisfied for all large enough t, whence $\lim_{t\to\infty} q_t(\omega) \leq R - \gamma$. This contradicts $\lim_{t\to\infty} \bar{q}_t(\omega) = R$. □

To complete the proof of Proposition 2, let $\hat{\omega}^I := \operatorname{argmin}_{\hat{\omega}'} \mathrm{KL}(F(\theta^*(\hat{\omega})), \hat{q}^I_\infty(\hat{\omega}'))$ and $\hat{\omega}^C := \operatorname{argmin}_{\hat{\omega}'} \mathrm{KL}(F(\theta^*(\hat{\omega})), \hat{q}^C_\infty(\hat{\omega}'))$. Then Lemmas B.2 and E.4 imply that almost all incorrect agents' beliefs converge to a point-mass on $\hat{\omega}^I$ and almost all quasi-correct agents' beliefs converge to a point-mass on $\hat{\omega}^C$. Moreover, since $\hat{q}^I_\infty(\cdot) = \hat{F}(\theta^*(\cdot))$ and $\hat{F}(\theta^*(\hat{\omega})) = F(\theta^*(\hat{\omega}))$ by construction, we must have $\hat{\omega}^I = \hat{\omega}$. Likewise, by Lemma E.3, $\hat{q}^C_\infty(\theta^*(\hat{\omega})) = \hat{F}(\theta^*(\hat{\omega})) = F(\theta^*(\hat{\omega}))$, so that $\hat{\omega}^C = \hat{\omega}$.

F Robustness of Single-Agent Active Learning

Consider the active learning model defined in Section 4.3. We measure the amount of misperception by a "bias" parameter $b \in \mathbb{R}$. Specifically, we write $\hat{q}(a, \omega) = r(a, \omega, b)$ for some C^1 function r that is strictly decreasing in (a, ω) such that $q(a, \omega) = r(a, \omega, 0)$. We also assume that $a^*(\cdot) := \operatorname{argmax}_{a \in A} u(a, \cdot)$ is C^1.

Proposition F.1. *Fix any $\varepsilon > 0$. There exists $\bar{b} > 0$ such that if $|b| < \bar{b}$, then at each $\omega \in \Omega$, process (2) admits a unique steady state $\hat{\omega}_\infty(\omega)$; moreover, $\hat{\omega}_\infty(\omega) \in [\omega - \varepsilon, \omega + \varepsilon]$ and is globally stable.*

Proof. We first show that there exists $\bar{b} > 0$ such that at each $\omega \in \Omega$, process (2) satisfies $\hat{\omega}_t \in [\omega - \varepsilon, \omega + \varepsilon]$ for all $t \geq 2$ whenever $|b| \leq \bar{b}$. To see this, consider identity

$$r(a, \omega, 0) = r(a, \hat{\omega}, b) \tag{31}$$

as a function of $\hat{\omega}$. If $b = 0$, then for any a and ω, (31) admits $\hat{\omega} = \omega$ as the unique solution. Thus, by the implicit function theorem, $\frac{d\hat{\omega}}{db} = \frac{-\frac{\partial}{\partial b} r(a,\hat{\omega},b)}{\frac{\partial}{\partial \hat{\omega}} r(a,\hat{\omega},b)}$ holds at $b = 0$ and $\hat{\omega} = \omega$. But since r is C^1 and $A \times \Omega = [0,1] \times [\underline{\omega}, \overline{\omega}]$ is compact, $\max_{(a,\omega) \in A \times \Omega} \left| \frac{-\frac{\partial}{\partial b} r(a,\omega,0)}{\frac{\partial}{\partial \omega} r(a,\omega,0)} \right| < \infty$. Hence, there exists $\bar{b} > 0$ such that for every $b \in [-\bar{b}, \bar{b}]$, a, and ω, (31) admits a unique solution $\hat{\omega} \in [\omega - \varepsilon, \omega + \varepsilon]$; that is, process (2) satisfies $\hat{\omega}_t \in [\omega - \varepsilon, \omega + \varepsilon]$ for all $t \geq 2$ from any initial point $\hat{\omega}_1$.

Finally, applying the implicit function theorem to $r(a^*(\hat{\omega}_t), \omega, 0) = r(a^*(\hat{\omega}_t), \hat{\omega}_{t+1}, b)$, we obtain $\frac{d\hat{\omega}_{t+1}}{d\hat{\omega}_t} = -\frac{a^{*\prime}(\hat{\omega}_t)\left(\frac{\partial r(a^*(\hat{\omega}_t),\omega,0)}{\partial a^*} - \frac{\partial r(a^*(\hat{\omega}_t),\hat{\omega}_t,b)}{\partial a^*}\right)}{\frac{\partial r(a^*(\hat{\omega}_t),\hat{\omega}_{t+1},b)}{\partial \hat{\omega}_{t+1}}}$. By uniform continuity of the derivatives (which holds by compactness of the domain $A \times \Omega$), we can choose $\bar{b}$ sufficiently small such that for all $|b| \leq \bar{b}$ and ω, the right hand side is strictly less than 1 in absolute value at all $t \geq 2$. This guarantees that process (2) is a contraction on $[\omega - \varepsilon, \omega + \varepsilon]$. Hence, it admits a unique steady state $\hat{\omega}_\infty(\omega) \in [\omega - \varepsilon, \omega + \varepsilon]$, to which it converges from any initial point. □

G Continuous Actions

This section considers a continuous action space version of our model. We perform steady state analysis (under the limit model) to illustrate why our main insights do not rely on a finite action space. Throughout, we assume that the action space is an interval $A = [\underline{a}, \overline{a}] \subseteq \mathbb{R}$, with $-\infty \leq \underline{a} < \overline{a} \leq \infty$. Let $u(a, \theta, \omega)$ denote type θ's utility to choosing action a in state ω. We assume that for every type $\theta \in \mathbb{R}$ and state $\omega \in \Omega := [\underline{\omega}, \overline{\omega}]$, there exists a unique optimal action $a^*(\theta, \omega) := \text{argmax}_{a \in A}\, u(a, \theta, \omega)$ which is continuous and strictly increasing in (θ, ω) and such that $a^*(\cdot, \omega)$ has full range for all ω.

Given any true and perceived type distributions $F, \hat{F} \in \mathcal{F}$, we briefly analyze the set of steady states $\text{SS}(F, \hat{F})$ of this model. For each state ω, let $G(\cdot, \omega) \in \Delta(A)$ denote the true cdf over actions in the population when (almost all) agents assign probability 1 to state ω and let $g(\cdot, \omega)$ denote the corresponding density. Likewise, let $\hat{G}(\cdot, \omega)$ and $\hat{g}(\cdot, \omega)$ denote the corresponding perceived action distribution and density when agents assign probability 1 to ω. Note that $G(a, \omega) = F(\theta^*(a, \omega))$ and $\hat{G}(a, \omega) = \hat{F}(\theta^*(a, \omega))$, where $\theta^*(a, \omega)$ satisfies $a = a^*(\theta^*(a, \omega), \omega)$. Let $\text{KL}(H, \hat{H}) := \int \log \left[\frac{h(a)}{\hat{h}(a)}\right] h(a)\, da$ denote the KL divergence between continuous distributions H and $\hat{H}$ with densities h and $\hat{h}$. As in the binary action space setting, we define a steady state $\hat{\omega}^*$ to be a solution to

$$\hat{\omega}^* \in \operatorname*{argmin}_{\hat{\omega}} \text{KL}(G(\cdot, \hat{\omega}^*), \hat{G}(\cdot, \hat{\omega})).$$

Thus, as before, in a steady state agents assign probability 1 to a state that minimizes the KL divergence between the corresponding observed action distribution and agents' perceived action distribution. At interior steady states $\hat{\omega}^*$, the first-order condition yields

$$\int \frac{g(a, \hat{\omega}^*)}{\hat{g}(a, \hat{\omega}^*)} \frac{\partial \hat{g}(a, \hat{\omega}^*)}{\partial \hat{\omega}}\, da = 0. \tag{32}$$

Thus, the set of steady states $\text{SS}(F, \hat{F})$ is finite whenever there are at most finitely many $\hat{\omega}^*$ that satisfy (32). A sufficient condition for this is that the left-hand side of (32) is analytic in $\hat{\omega}^*$ and not constantly equal to 0; similar to the logic behind Theorem 2, this is ensured if $F \neq \hat{F}$ are analytic and $\theta^*(a, \cdot)$ is analytic. Moreover, similar to the logic behind Theorem 1, it is easy to construct examples where $\hat{F}$ is arbitrarily close to F but there is only a single (state-independent) steady state, as the following illustrates:

Example 5 (Gaussian type distributions)**.** Consider the quadratic-loss utility $u(a,\theta,\omega) = -(a-\theta-\omega)^2$, which implies that the optimal action takes the form $a^*(\theta,\omega) = \theta+\omega$. Suppose that F and $\hat{F}$ are cdfs of the Gaussian distributions $N(\mu,\sigma^2)$ and $N(\hat{\mu},\hat{\sigma}^2)$. Then the left-hand side of (32) is given by $\int \frac{\hat{\mu}-\theta}{\hat{\sigma}^2} \frac{\exp[-\frac{(\theta-\mu)^2}{2\sigma^2}]}{\sqrt{2\pi\sigma^2}} d\theta = \frac{\hat{\mu}-\mu}{\hat{\sigma}^2}$. Thus, there is no interior steady state, and whenever $\mu > \hat{\mu}$ (respectively, $\mu < \hat{\mu}$), the unique steady state is given by $\overline{\omega}$ (respectively, $\underline{\omega}$), paralleling Example 1 in the binary action setting. □

Finally, we have focused only on steady state analysis in this section, without considering belief dynamics and establishing convergence to steady states. However, we note that in specific settings such as the Gaussian environment from Example 5 (assuming additionally that states and signals are normally distributed), the evolution of agents' beliefs in every period t admits a simple characterization in terms of difference equations, based on which convergence to the above steady states can be readily established. Details are available on request.