Online Appendix for "Robust Misspecified Models"

Cuimin Ba*

June 15, 2025

C More Auxiliary Lemmas

Lemma 6. *Suppose the agent maximizes the discounted sum of payoffs under her current model with discount factor* δ. *For any* $\theta \in \Theta$, *the optimal action correspondence* $A^\theta : \Delta\Omega^\theta \rightrightarrows \mathcal{A}$ *is upper hemicontinuous in both the belief* π *and the discount factor* δ.

Proof. This follows from standard results in dynamic programming. Under continuity and boundedness of the payoff function and compactness of the action space $\mathcal{A}$, the optimal action correspondence is upper hemicontinuous in both belief and discount factor (see Blackwell, 1965; Maitra, 1968). □

Lemma 7. *If* σ *is a p-absorbing SCE, then for any* $\gamma \in (0,1)$ *and* $\epsilon > 0$, *there exists a full-support prior* π_0^θ *at which, with probability at least* γ, *a* θ*-modeler only plays actions in* $\text{supp}(\sigma)$ *and her belief stays within* $B_\epsilon(\Delta\Omega^\theta(\sigma))$ *for all periods.*

Proof. Suppose σ is a p-absorbing SCE under model θ, and consider the learning process of a θ-modeler starting from a full-support prior $\pi_0^\theta \in \Delta\Omega^\theta$. By definition, there exists π_0^θ such that with positive probability, she eventually only plays actions in $\text{supp}(\sigma)$ and each element of $\text{supp}(\sigma)$ is played infinitely often (this is without loss of generality). Denote those paths by $\tilde{H}$. Then by a similar argument as in the proof of Lemma 1, π_t^θ a.s. converges to a limit π_∞^θ on $\tilde{H}$, with $\text{supp}\left(\pi_\infty^\theta\right) \subseteq \Omega^\theta(\sigma) = \{\omega \in \Omega^\theta : q^*(\cdot|a) = q^\theta(\cdot|a,\omega), \forall a \in \text{supp}(\sigma)\}$.

This implies the existence of an integer $T > 0$ such that, with positive probability, we have $a_t \in \text{supp}(\sigma), \forall t \geq T$, and π_t^θ converges to a limit π_∞^θ with $\text{supp}(\pi_\infty^\theta) \subseteq \Omega^\theta(\sigma)$.

*University of Pittsburgh. Email: bacuimin@gmail.com.

Pick any $\epsilon > 0$. Since the learning process is Markov, we can find a posterior in one of those paths and use it as a new prior, $\tilde{\pi}_0^\theta \in B_\epsilon(\Delta\Omega^\theta(\sigma))$, under which, on a positive measure of histories, a θ-modeler behaves such that (1) $a_t \in \text{supp}(\sigma), \forall t \geq 0$, and (2) the posterior π_t^θ never leaves $B_\epsilon(\Delta\Omega^\theta(\sigma))$ for all $t \geq 0$. Let E denote the event that both (1) and (2) hold. Define a stopping time T as the first period in which either (1) or (2) is not met. Let $\tilde{a}_t = a_{\min\{T,t\}}$ and $\tilde{\pi}_t^\theta = \pi_{\min\{T,t\}}^\theta$. Then the probability of event E is the same as the probability of event E', defined by: (1') $\tilde{a}_t \in \text{supp}(\sigma), \forall t \geq 0$; and (2') $\tilde{\pi}_t^\theta$ never leaves $B_\epsilon(\Delta\Omega^\theta(\sigma))$, $\forall t \geq 0$.

Suppose, toward a contradiction, that no full-support prior exists such that $\mathbb{P}_D(E') > \gamma$. Denote the probability of E' given any π^θ by $\gamma(\pi^\theta)$, and define the supremum $\overline{\gamma} := \sup_{\pi_0^\theta \in \text{int}(\Delta\Omega^\theta)} \gamma(\pi^\theta)$. Then by assumption, $\overline{\gamma} < 1$. By definition, for any $\psi > 0$, there exists some prior $\pi_0^{\theta,\psi}$ such that $\overline{\gamma} \geq \gamma(\pi_0^{\theta,\psi}) > \overline{\gamma} - \psi$. But under this prior, with probability $1 - \gamma(\pi_0^{\theta,\psi})$, the dogmatic modeler eventually sees $\tilde{a}_t \notin \text{supp}(\sigma)$ or $\pi_t^\theta \notin B_\epsilon(\Delta\Omega^\theta(\sigma))$. So for sufficiently large t,

$$\mathbb{P}_D\left(\gamma(\pi_t^{\theta,\psi}) = 0\right) > 1 - \gamma(\pi_0^{\theta,\psi}) - \psi \geq 1 - \overline{\gamma} - \psi.$$

Now, consider the supremum probability that E is achieved if the agent starts with a prior that is equal to posterior $\pi_t^{\theta,\psi}$. Since $\gamma(\pi_0^{\theta,\psi}) = \mathbb{E}_{h_t \in H_t}^{\mathbb{P}_D} \gamma(\pi_t^{\theta,\psi})$ for all $t \geq 1$, we have

$$\sup_{h_t \in H_T} \gamma(\pi_T^{\theta,\psi}) \geq \frac{\gamma(\pi_0^{\theta,\psi})}{1 - \mathbb{P}_D\left(\gamma(\pi_t^{\theta,\psi}) = 0\right)} > \frac{\overline{\gamma} - \psi}{\overline{\gamma} + \psi}.$$

But notice that when ψ is sufficiently small, the last term is strictly larger than $\overline{\gamma}$, contradiction. □

Lemma 8. *Take any model $\theta \in \Theta$ and any $\omega \in \Omega^\theta$. There exists $\underline{\gamma} : (0,1) \to (0,1)$ such that given any set $Y \subset \mathcal{Y}$ such that $Q^\theta(Y|a,\omega) > \gamma$ with $\gamma \in (0,1)$, we have $Q^*(Y|a) > \underline{\gamma}(\gamma)$ and $\lim_{\gamma \to 1} \underline{\gamma}(\gamma) = 1$.*

Proof. If there does not exist $\underline{\gamma} : (0,1) \to (0,1)$ such that the statement holds, then there exists $\overline{\gamma} < 1$ such that for any $\eta \in (0,1)$, there exists $\gamma > \eta$ and $Y \subseteq \mathcal{Y}$ such that $Q^\theta(Y|a,\omega) > \gamma$ and yet $Q^*(Y|a) < \overline{\gamma}$. Let $\{\eta_n\}$ be a strictly increase sequence and $\lim_{n\to\infty} \eta_n = 1$. Then for each n, we can find a set $\check{Y}_n \subseteq \mathcal{Y}$ such that $Q^\theta(\check{Y}_n|a,\omega) < 1 - \eta_n$ and $Q^*(\check{Y}_n|a) > 1 - \overline{\gamma}$. Since $\lim_{n\to\infty} Q^\theta(\check{Y}_n|a,\omega) = 0$ and $q^\theta(\cdot|a,\omega)$ is positive, it must be that $\lim_{n\to\infty} \nu(\check{Y}_n) = 0$. Since Q^* is absolutely continuous w.r.t. ν, it follows that $\lim_{n\to\infty} Q^*(Y_n|a) = 0$, a contradiction. □

Lemma 9. *Fix model $\theta \in \Theta$ and $\omega \in \Omega^\theta$. For any $r > 0$ and $\gamma < 1$, there exists $\epsilon > 0$ such that, if model $\theta' \in \Theta$ and $\omega' \in \Omega^{\theta'}$ satisfy $d(Q^{\theta,\omega}, Q^{\theta',\omega'}) \leq \epsilon$, then letting $Y_{a,r} := \{y \in \mathcal{Y} : q^{\theta'}(y|a,\omega') \leq (1+r)q^\theta(y|a,\omega)\}$ we have $Q^*(Y_{a,r}|a) > \gamma$ for any $a \in \mathcal{A}$.*

Proof. We first show that the statement holds if we replace "$Q^*(Y_{a,r}|a) > \gamma$" with "$Q^\theta(Y_{a,r}|a,\omega) > \gamma$". Suppose the statement does not hold for some given $r > 0$ and $\gamma < 1$ for contradiction. Then for any $\epsilon > 0$ we can find a model θ' and ω' satisfying $d(Q^{\theta,\omega}, Q^{\theta',\omega'}) \leq \epsilon$ such that $Q^\theta(Y_{a,r}|a,\omega) \leq \gamma$ for some $a \in \mathcal{A}$. Note that this implies $Q^\theta(\mathcal{Y} \setminus Y_{a,r}|a,\omega) \geq 1-\gamma$. In addition,

$$
\begin{aligned}
Q^{\theta'}(\mathcal{Y} \setminus Y_{a,r}|a,\omega') &= \int_{\mathcal{Y}\setminus Y_{a,r}} q^{\theta'}(y|a,\omega')\nu(dy) \\
&> \int_{\mathcal{Y}\setminus Y_{a,r}} (1+r)q^\theta(y|a,\omega)\nu(dy) \\
&\geq Q^\theta(\mathcal{Y} \setminus Y_{a,r}|a,\omega) + r(1-\gamma)
\end{aligned}
$$

where the first inequality follows from the fact that $y \in \mathcal{Y} \setminus Y_{a,r}$ and the second follows from $Q^\theta(\mathcal{Y} \setminus Y_{a,r}|a,\omega) \geq 1-\gamma$.

On the other hand, since $d(Q^{\theta,\omega}, Q^{\theta',\omega'}) \leq \epsilon$, we know that for all $Y \subseteq \mathcal{Y}$, $Q^{\theta'}(Y|a,\omega') \leq Q^\theta(B_\epsilon(Y)|a,\omega) + \epsilon$. Let $Y = \mathcal{Y} \setminus Y_{a,r}$, then

$$
Q^{\theta'}(\mathcal{Y} \setminus Y_{a,r}|a,\omega') \leq Q^\theta(B_\epsilon(\mathcal{Y} \setminus Y_{a,r})|a,\omega) + \epsilon.
$$

However, when ϵ is sufficiently small, since q^θ is continuous, the right-hand side of the above inequality must be smaller than $Q^\theta(\mathcal{Y}\setminus Y_{a,r}|a,\omega)+r(1-\gamma)$. Since this contradicts the previous inequality, we must have $Q^\theta(Y_{a,r}|a,\omega) > \gamma$. Furthermore, by Lemma 8, we can choose ϵ sufficiently small and η sufficiently close to 1 such that $Q^\theta(Y_{a,r}|a,\omega) > \eta$ and $Q^*(Y_{a,r}|a) > \gamma$. □

D Omitted Examples in Section 4

D.1 A P-Absorbing Mixed SCE

Example 2. Consider a dogmatic modeler who holds model θ that contains three parameters $\Omega^\theta = \{\omega_1, \omega_2, \omega_3\} = \{1, 1.5, 2\}$. There are two actions $\mathcal{A} = \{1, 2\}$. The agent's payoff is simply the outcome y_t, with the true DGP being the normal distribution $N(0.25, 1)$ for all actions. Model θ is misspecified, predicting that $y_t \sim N((\omega - a_t)^2, 1)$.

Note that every mixed action is a self-confirming equilibrium, with the supporting belief assigning probability 1 to the parameter value $\omega_2^* = 1.5$. Here, every fully mixed SCE is p-absorbing since it is quasi-strict: its support contains all actions.

But her action sequence never converges. To see that, notice that the agent's optimal action is unique when her posterior belief assigns different probabilities to ω_1 and ω_3. In particular, her optimal action if 1 when $\pi_t^\theta(\omega_1) < \pi_t^\theta(\omega_3)$ and 2 when $\pi_t^\theta(\omega_1) > \pi_t^\theta(\omega_3)$. When playing $a = 2$, the agent anticipates the outcome to be distributed according to $y_t \sim N((\omega - 2)^2, 1)$. However, given the true distribution $N(0.25, 1)$, the agent eventually attaches a lower probability to ω_1 than ω_3, which then leads her to play $a = 1$. By a similar logic, the agent cannot settle on action $a = 1$ either. Therefore, the agent perpetually oscillates between the two actions, while her belief converges to a degenerate distribution at ω_2 since it outperforms the other two parameter values by fitting the data perfectly.

D.2 A SCE That Is Not P-Absorbing

Example 3. Consider a dogmatic modeler who holds model θ that contains three parameters $\Omega^\theta = \{1, 2, 3\}$. There are two actions $\mathcal{A} = \{1, 3\}$. The agent's payoff is the absolute value of the outcome, $|y_t|$. The true DGP of y_t given by a normal distribution $N(1, 1)$ for all actions. Model θ is misspecified and predicts that $y_t \sim N(\omega - a_t, 1)$. Note that θ admits a single self-confirming equilibrium in which the agent plays $a^* = 1$ with probability 1, supported by a belief that assigns probability 1 to $\omega^* = 2$. However, this SCE is not p-absorbing. To see that, notice that the agent is indifferent between the two actions when the parameter takes the value of 2. When the agent keeps playing $a = 1$, the parameters 1 and 3 fit the data equally well on average, so their log-posterior ratio is a random walk which a.s. crosses 1 infinitely often. However, the high action $a = 3$ is strictly optimal against any belief that assigns a higher probability to $\omega = 1$ than $\omega = 3$. Hence, the high action must be played infinitely often almost surely.

D.3 Another Type of Traps

Example 4. Consider an agent who chooses from $\mathcal{A} = \{a^1, a^2, a^3\}$ and observes outcomes from $\mathcal{Y} = \{0, 1\}$. The true DGP prescribes $y_t = 1$ with probability 0.5 for all actions. Given action a_t and a realized outcome y_t, the agent obtains a flow payoff of $y_t + h(a_t)$, where $h(a^1) = 0, h(a^2) = -0.3$, and $h(a^3) = 0.01$. The agent holds an initial model θ and considers a correctly specified competing model θ' same as the true model,

$q^{\theta}(1\|a,\omega)$	ω^1	ω^2
a^1	0.5	0.3
a^2	0.6	0.7
a^3	0.49	0.29

$q^{\theta'}(1\|a,\omega)$	ω^*
a^1	0.5
a^2	0.5
a^3	0.5

Table 1: Initial model θ and competing model θ' in Example 4.

as described in Table 1, and she employs a switching threshold of $\alpha = 3$. Under model θ, both a^1 and a^3 are optimal when $\pi_t^\theta(\omega_1) \geq 1/3$, and a^2 is optimal when $\pi_t^\theta(\omega^1) \leq 1/3$. Therefore, δ_{a^1} is a SCE with a supporting belief δ_{ω^1}, but it is not quasi-strict because a^2 is also optimal at this belief. Under model θ', a^3 is the uniquely optimal action at all beliefs. Suppose $\alpha = 1.2$. As illustrated below, action a^2 functions as a trap that prevents the "switcher" agent from ever playing a^1 under model θ.

Suppose the agent starts with a prior with $\pi_0^\theta(\omega^1) = 1/3$ such that she plays a^2 in period 0. In addition, suppose the agent adopts a pure policy under θ that prescribes a^1 for a countable set of beliefs A, where

$$A = \left\{ \pi \in \Delta\Omega^\theta : \pi(\omega^1) \geq \frac{1}{3} \text{ and } \frac{\pi(\omega^1)}{\pi(\omega^2)} = \frac{1}{2} \cdot \frac{4}{3} \cdot \left(\frac{5}{7}\right)^m \cdot \left(\frac{5}{3}\right)^n \text{ for some } m, n \in \mathbb{N} \right\}.$$

In period $t = 0$, the agent either (1) draws $y_0 = 0$ and then switches to model θ' since $\lambda_0 = 1.5 > \alpha$, followed by at least one period of playing a^3, or (2) draws $y_0 = 1$ and continues with model θ and a^2 in the next period. In scenario (1), the agent's belief π_2^θ is such that

$$\text{either } \frac{\pi_2^\theta(\omega^1)}{\pi_2^\theta(\omega^2)} = \frac{1}{2} \cdot \frac{4}{3} \cdot \frac{51}{71} \text{ or } \frac{\pi_2^\theta(\omega^1)}{\pi_2^\theta(\omega^2)} = \frac{1}{2} \cdot \frac{4}{3} \cdot \frac{49}{29},$$

depending on the outcome realization y_1. Therefore, the agent will never play a^1 in future periods. Meanwhile, in scenario (2), the agent's belief is such that

$$\text{either } \frac{\pi_2^\theta(\omega^1)}{\pi_2^\theta(\omega^2)} = \frac{1}{2} \cdot \frac{6}{7} \cdot \frac{6}{7} \text{ or } \frac{\pi_2^\theta(\omega^1)}{\pi_2^\theta(\omega^2)} = \frac{1}{2} \cdot \frac{6}{7} \cdot \frac{4}{3},$$

depending on the outcome realization y_1. Therefore, the agent's belief π_t^θ will remain outside of A and thus she will never play a^1 in future periods.

While the switcher never converges to the SCE δ_{a^1}, a θ-modeler converges to the SCE with positive probability. To see why, first notice that a θ-modeler also starts by playing a^2. However, upon drawing $y_0 = 0$, the agent's belief π_1^θ enters A and thus she chooses a^1 thereafter as long as her belief assigns probability weakly higher than 1/3 to state ω^1. Since playing a^1 is self-confirming with a supporting belief δ_{ω^1}, the previous event indeed occurs with positive probability. Therefore, the SCE is p-absorbing.

E More Details and Results for Applications in Section 5

E.1 Application 1: Over- and Underconfidence

Recall that Proposition 1 shows that there exists a sequence $\underline{b} = \beta_K < ... < \beta_1 < \beta_0 = b^*$ such that model θ is not locally robust if $\hat{b} \in (\beta_{2k}, \beta_{2k-1})$ for some $k \in \mathbb{N}_+$.

Proposition 4. *The total measure of* $\cup_{k=1}(\beta_{2k}, \beta_{2k-1})$ *is bounded below by a positive number for any discrete* $\mathcal{A}$*.*

Proof. Suppose without loss of generality that $g_{a\omega} > 0$ and $g_{ab} \leq 0$. Suppose $\mathcal{A} = \{a^{-N'}, ..., a^{-1}, a^0, a^1, ..., a^N\} \subset [\underline{a}, \overline{a}]$, where N, N' may be ∞, the actions are strictly increasing, and $a^0 = a^*$. Fix any $n \geq 0$. Denote the fundamental value at which the agent with $\hat{b} = b$ is indifferent between a^n and a^{n+1} as $\omega^n(b)$, that is,

$$g(a^n, b, \omega^n(b)) = g(a^{n+1}, b, \omega^n(b)). \tag{1}$$

If $\omega > \omega^n(b)$, then the agent strictly prefers a^{n+1} to a^n. No self-confirming equilibrium exists if $\hat{b} \in (b^{n,2}, b^{n,1})$, with $b^{n,1}$ and $b^{n,2}$ defined by

$$g(a^n, b^{n,1}, \omega^n(b^{n,1})) = g(a^n, b^*, \omega^*), \tag{2}$$

$$g(a^{n+1}, b^{n,2}, \omega^n(b^{n,2})) = g(a^{n+1}, b^*, \omega^*). \tag{3}$$

The fact that $b^{n,2} < b^{n,1}$ follows from the assumptions that $g_{ab} \leq 0, g_{a\omega} > 0$ and $g_b, g_\omega > 0$. Combining Eq. (1) and Eq. (2),

$$g(a^{n+1}, b^{n,1}, \omega^n(b^{n,1})) = g(a^n, b^*, \omega^*). \tag{4}$$

Subtracting both sides of Eq. (3) from Eq. (4), and summing across $n = 0, 1, ...,$

$$\sum_{n=0}^{N-1} \left[g(a^{n+1}, b^{n,1}, \omega^n(b^{n,1})) - g(a^{n+1}, b^{n,2}, \omega^n(b^{n,2})) \right] = g(a^0, b^*, \omega^*) - g(a^N, b^*, \omega^*). \quad (5)$$

Note that for any $n \geq 0$,

$$g(a^{n+1}, b^{n,1}, \omega^n(b^{n,1})) - g(a^{n+1}, b^{n,2}, \omega^n(b^{n,2})) \quad (6)$$

$$= (b^{n,1} - b^{n,2}) \left(g_b(a^{n+1}, b', \omega^n(b')) + g_\omega(a^{n+1}, b'', \omega^n(b'')) \frac{\partial \omega^n(b'')}{\partial b} \right), \quad (7)$$

where $b', b'' \in (b^{n,2}, b^{n,1})$. Since g is twice continuously differentiable over $[\underline{a}, \overline{a}] \times [\underline{b}, \overline{b}] \times [\underline{\omega}, \overline{\omega}]$, g_b and g_ω is bounded above and below by 0. Differentiate Eq. (1),

$$\frac{\partial \omega^n(b)}{\partial b} = \frac{g_b(a^n, b, \omega^n(b)) - g_b(a^{n+1}, b, \omega^n(b))}{g_\omega(a^{n+1}, b, \omega^n(b)) - g_\omega(a^n, b, \omega^n(b))}, \quad (8)$$

which is weakly positive and also bounded above. Together with Eq. (5), this implies there exists some positive $M > 0$ such that

$$\sum_{n=0}^{N-1} (b^{n,1} - b^{n,2}) \geq \frac{g(a^0, b^*, \omega^*) - g(a^N, b^*, \omega^*)}{M} > 0.$$

□

E.2 Application 2: Media Bias and Polarization

Micro-Foundation for Agent's Preference. Under both models, the agent is assumed to strictly prefer the media outlet whose leaning matches the state of the world. This preference can be micro-founded by the following utility function,

$$u(a, y, \omega) = \begin{cases} \omega & \text{if } a = a^L, \\ c & \text{if } a = a^M, \\ 1 - \omega & \text{if } a = a^R, \end{cases}$$

where $c \in (1/2, \delta)$ is a constant. Under either model, the agent strictly prefers a^L if the expected state $\mathbb{E}(\omega)$ exceeds c, a^R if $\mathbb{E}(\omega)$ is below c, and a^M otherwise. This implies that the agent strictly prefers a^L when $\pi_t^\theta(\omega^L)$ is sufficiently close to 1, a^R when $\pi_t^\theta(\omega^R)$

is sufficiently close to 1, and a^M otherwise. The same applies under models $\hat{\theta}$ and θ^*.

Thus, conditional on the true state being ω^M and given the misreporting strategies of the outlets, a^M is the unique SCE under both correctly specified models θ and θ^*, while a^L and a^R are two strict SCEs under the misspecified model $\hat{\theta}$. As noted in Section 3, the main results are unchanged if this component of the payoff is unobservable.

Proof of Proposition 2(ii). Below I prove the result that the probability that $\hat{\theta}$ eventually replaces θ converges to 1 as $\pi_0^\theta(\omega^L)$ or $\pi_0^\theta(\omega^R) \to 1$.

Proof. I show that if $\pi_0^\theta(\omega^L)$ is sufficiently close to 1, the probability that m_t converges to $\hat{\theta}$ can be arbitrarily close to 1. The argument for the case where $\pi_0^\theta(\omega^R)$ is close to 1 is symmetric. First suppose $\pi_0^\theta(\omega^L) = 1$. Then the agent chooses a^L so long as she remains under θ. Since $q^\theta(l|a^L, \omega^L) = \delta + (1-\delta)x > \frac{1}{2} + \frac{1}{2}x = q^*(l|a^L) = q^{\hat{\theta}}(l|a^L, \omega^L)$, the law of large numbers implies that the probability that m_t converges to θ is 0. If in addition, whenever the agent switches from θ to $\hat{\theta}$, the probability that the agent never switches back is bounded below by a positive constant, then the agent cannot switch between θ and $\hat{\theta}$ infinitely often with positive probability and hence m_t converge to $\hat{\theta}$. The next paragraph provides this argument.

Suppose the agent switches from θ to $\hat{\theta}$ at the end of period T, then $\lambda_T > \alpha$, which implies $\alpha \ell_T(\theta, \omega^L) < \ell_T(\hat{\theta}) = \pi_0^{\hat{\theta}}(\omega^L)\ell_T(\hat{\theta}, \omega^L) + \pi_0^{\hat{\theta}}(\omega^R)\ell_T(\hat{\theta}, \omega^R)$. By assumption, there exists $\bar{\epsilon} \in (0, 1/2)$ such that if $\epsilon \leq \bar{\epsilon}$, the agent strictly prefers a^L when $\pi_t^{\hat{\theta}}(\omega^L) \geq 1 - \epsilon$ and a^R when $\pi_t^{\hat{\theta}}(\omega^R) \geq 1 - \epsilon$. Thus, if $\frac{\pi_0^{\hat{\theta}}(\omega^L)\ell_T(\hat{\theta}, \omega^L)}{\pi_0^{\hat{\theta}}(\omega^R)\ell_T(\hat{\theta}, \omega^R)} \geq \frac{1-\epsilon}{\epsilon}$ (call this region 1), then the agent plays a^L upon switching to $\hat{\theta}$. Since ω^L predicts the true DGP under $\hat{\theta}$ and a^L, by Ville's maximal inequality for martingales, when ϵ is sufficiently small, the probability that $\pi_0^{\hat{\theta}}(\omega^L)\ell_T(\hat{\theta}, \omega^L) \geq \max\{\frac{1-\bar{\epsilon}}{\bar{\epsilon}}\pi_0^{\hat{\theta}}(\omega^R)\ell_T(\hat{\theta}, \omega^R), \ell_T(\theta, \omega^L)/\alpha\}, \forall t > T$, i.e. the agent remains under $\hat{\theta}$ and plays a^L forever, is bounded below by a positive constant. Similarly, if $\frac{\pi_0^{\hat{\theta}}(\omega^L)\ell_T(\hat{\theta}, \omega^L)}{\pi_0^{\hat{\theta}}(\omega^R)\ell_T(\hat{\theta}, \omega^R)} \leq \frac{\epsilon}{1-\epsilon}$ (call this region 2), then the probability that the agent remains under $\hat{\theta}$ and plays a^R forever is also bounded below by a positive constant. Suppose instead $\frac{\pi_0^{\hat{\theta}}(\omega^L)\ell_T(\hat{\theta}, \omega^L)}{\pi_0^{\hat{\theta}}(\omega^R)\ell_T(\hat{\theta}, \omega^R)} \in \left(\frac{\epsilon}{1-\epsilon}, \frac{1-\epsilon}{\epsilon}\right)$. If in addition $\alpha \ell_T(\theta, \omega^L) \leq \ell_T(\hat{\theta}, \omega^L)$, then by drawing r for $s > 0$ times, both $\frac{\ell_T(\hat{\theta}, \omega^R)}{\ell_T(\theta, \omega^L)}$ and $\frac{\ell_T(\hat{\theta}, \omega^R)}{\ell_T(\hat{\theta}, \omega^L)}$ increase, pushing the agent's posterior towards δ_{ω^R} while ensuring she does not switch away from $\hat{\theta}$. For a fixed ϵ, there exists a finite $s > 0$ such that the agent ends up in region 2. Therefore, the probability that the agent remains under $\hat{\theta}$ and eventually plays a^R forever is bounded below. Similarly, if $\alpha \ell_T(\theta, \omega^L) \leq \ell_T(\hat{\theta}, \omega^R)$, then by analogous reasoning, the probability that the agent remains under $\hat{\theta}$ and eventually plays a^L forever is bounded below.

In sum, when $\pi_0^\theta(\omega^L) = 1$, $\hat{\theta}$ eventually replaces θ with probability 1. Further, the agent converges to one of the two SCEs under $\hat{\theta}$, a^L or a^R. Hence, λ_t diverges to ∞. Therefore, for any $\eta \in (0,1), \epsilon \in (0,1)$, and $\overline{\lambda} \in \mathbb{R}_+$, there exists $T \in \mathbb{N}$ such that the probability that $\pi_t^{\hat{\theta}}(\omega^L) \in (0,\epsilon) \cap (1-\epsilon, 1)$ and $\lambda_t > \overline{\lambda}$ when $t \geq T$ is strictly larger than η. Denote such histories up to period T as $\hat{H}_T$. Since the optimal action correspondence is upper hemicontinuous in $\pi_0^\theta(\omega^L)$ and λ_t is continuous in $\pi_0^\theta(\omega^L)$ for any given t, when $\pi_0^\theta(\omega^L)$ is sufficiently close to 1, $\mathbb{P}(\hat{H}_T)$ is still strictly larger than η. Take any $h_T \in \hat{H}_T$ and suppose the agent plays a^L in the end. Note that $\lambda_t > \overline{\lambda} \frac{\sum_{\omega \in \Omega^\theta} \pi_0^\theta(\omega) \prod_{\tau=T+1}^t q^\theta(y_t|a_t,\omega)}{\prod_{\tau=T+1}^t q^{\hat{\theta}}(y_t|a_t,\omega)}$. Since ω^L predicts the true DGP under a^L, by Ville's maximal inequality, when ϵ is sufficiently small and $\overline{\lambda}$ is sufficiently large, the probability that $\pi_0^\theta(\omega^L)$ stays within $(1-\overline{\epsilon}, 1)$ and λ_t stays above $1/\alpha$ is close to 1. Therefore, $m_t \to \hat{\theta}$ with probability approaching 1 as $\pi_0^\theta(\omega^L) \to 1$. □

F More Results for Extensions in Section 6

F.1 The Likelihood Ratio Test (LRT)

Recall that under the LRT, the agent computes the likelihood ratio using maximum-likelihood estimates:

$$\lambda_t^{\max} = \frac{\ell_t^{\max}(\theta')}{\ell_t^{\max}(\theta)},$$

where

$$\ell_t^{\max}(\theta) = \max_{\omega \in \Omega^\theta} \ell_t(\theta, \omega), \quad \ell_t^{\max}(\theta') = \max_{\omega' \in \Omega^{\theta'}} \ell_t(\theta', \omega').$$

The agent switches to θ' if $\lambda_t^{\max}$ exceeds α and to θ if $\lambda_t^{\max}$ falls below $1/\alpha$. We have the following result.

Theorem 4. *Suppose the agent uses the LRT rule. Then, for any $\theta \in \Theta$, there exists $\overline{\alpha} > 1$ such that if $\alpha \in [1, \overline{\alpha})$, θ cannot be globally robust under any prior. Moreover, if θ does not have perfect asymptotic accuracy, if $\mathcal{Y}$ is a continuum, or if the competing model may consist of arbitrarily many DGPs, then $\overline{\alpha} = \infty$ for all $\theta \in \Theta$.*

Proof. Fix any $\theta \in \Theta$. Take $\theta' \in \Theta$ such that the set of DGPs included in θ is a subset of those in θ'. Consider an agent whose initial model is θ and competing model is θ'. Once she switches from θ to θ', she will never switch back, because

$$\ell_t^{\max}(\theta) = \max_{\omega \in \Omega^\theta} \ell_t(\theta, \omega) \leq \max_{\omega' \in \Omega^{\theta'}} \ell_t(\theta', \omega') = \ell_t^{\max}(\theta').$$

Therefore, it suffices to show that when α is sufficiently close to 1, we can construct θ' such that switching from θ to θ' occurs at least once.

The initial model θ and its prior determine the first action a_0, which in turn pins down θ's predictions for the first outcome, $\{q^\theta(\cdot|a_0,\omega)\}_{\omega\in\Omega^\theta}$. Define the function $q^{\max}: \mathcal{Y} \to \mathbb{R}$ as $q^{\max}(y) = \max_{\omega\in\Omega^\theta} q^\theta(y|a_0,\omega)$. By Assumption 2 and the finiteness of Ω^θ, there exists $\epsilon > 0$ and $\eta < 1$ such that $\int_{B_\epsilon(y)} q^{\max}(y)d\nu(y) < \eta$ for all $y \in \mathcal{Y}$. The finiteness also implies that there exists a compact set $\overline{Y} \subset \mathcal{Y}$ such that $\int_{\mathcal{Y}\setminus\overline{Y}} q^{\max}(y)d\nu(y) < \eta$. Since $\overline{Y}$ is compact, there exists a finite set $U \subset \overline{Y}$ such that $\mathcal{Y} \subset \bigcup_{y\in U} B_\eta(y) \cup (\mathcal{Y}\setminus\overline{Y})$. For each $y \in U$ and each $\omega \in \Omega^\theta$, construct a DGP $q_{y,\omega}$ satisfying Assumption 2 such that $q_{y,\omega}(y'|a_0) = q^\theta(y'|a_0,\omega)/\eta$ for all $y' \in B_\eta(y)$. For each $\omega \in \Omega^\theta$, construct a DGP $q_{\mathcal{Y}\setminus\overline{Y},\omega}$ satisfying Assumption 2 such that $q_{\mathcal{Y}\setminus\overline{Y},\omega}(y'|a_0) = q^\theta(y'|a_0,\omega)/\eta$ for all $y' \in \mathcal{Y}\setminus\overline{Y}$. Let θ' include all these DGPs. Then, regardless of the realization of y_0, we have $\ell_0^{\max}(\theta')/\ell_0^{\max}(\theta) \geq 1/\eta$. Let $\overline{\alpha} = 1/\eta$; then the agent must switch to θ' in the next period if $\alpha < \overline{\alpha}$.

If θ does not have perfect asymptotic accuracy, letting θ' be the true model, a similar argument as in the proof of Lemma 1 implies that a switch must eventually happen. Therefore, θ cannot be globally robust at any prior for any α, which implies $\overline{\alpha} = \infty$.

If $\mathcal{Y}$ is a continuum, the continuity of $q^\theta(\cdot|a_0,\omega)$ over $\mathcal{Y}$ for all $\omega \in \Omega^\theta$ allows the choice of η in the construction above to be made arbitrarily close to zero, implying $\overline{\alpha} = \infty$.

Finally, suppose θ has perfect asymptotic accuracy and $\mathcal{Y}$ is countable, but the competing model can consist of an arbitrary number of DGPs. Let $\mathcal{Y} = \{y^1, ..., y^n, ...\}$ and $\theta' = (\Delta\mathcal{Y})^{\mathcal{A}}$, meaning the competing model θ' consists of all possible data-generating processes. While this competing model may not satisfy Assumption 2 because some of its predicted DGPs may not have full-support, this can be addressed by slightly perturbing θ' and ensuring that Assumption 2 is satisfied. In the long term, since the empirical frequency of outcomes converges in probability to the true distribution, the maximized likelihood ratio between this perturbed model and θ will converge to the maximized likelihood ratio between θ' and θ, so the proof goes through.

Since θ' is correctly specified, a similar argument as in the proof of Lemma 1 implies that the agent's actions almost surely converge to the support of an SCE on the paths where the agent eventually settles on θ. For simplicity, suppose the agent's actions converge to a^* and she plays a^* starting from period 0. The argument is analogous for other cases but slightly heavier in notation.

Let $q_t \in \Delta(\mathcal{Y})$ denote the empirical frequency of the realized outcomes, $q_t(y) = \sum_{\tau=0}^{t} \mathbf{1}(y_\tau = y)/(t+1)$. Since θ has perfect asymptotic accuracy and θ' contains all possible DGPs, the likelihood ratio $\tilde{\lambda}_t$ is asymptotically bounded below by the likelihood ratio between the empirical frequency q_t and the true DGP, denoted by ξ_t. Note that

$$\ln \xi_t = \sum_{y^i \in \mathcal{Y}} q_t(y^i) t \left(\ln q_t(y^i) - \ln(q^*(y^i|a^*)) \right) = t D_{KL}(q_t \parallel q^*(\cdot|a^*)).$$

By the Law of Large Numbers, q_t converges almost surely to the true distribution $q^*(\cdot|a^*)$. Using Taylor's expansion,

$$t D_{KL}(q_t \parallel q^*(\cdot|a^*)) = \frac{1}{2} \sum_{y^i \in \mathcal{Y}} \frac{t(q_t(y^i) - q^*(y^i|a^*))^2}{q^*(y^i|a^*)} + o\left(t \parallel q_t - q^*(\cdot|a^*) \parallel^2 \right).$$

By the Central Limit Theorem, $\sqrt{t}(q_t - q^*(\cdot|a^*))$ converges in distribution to a multivariate normal distribution with time-invariant covariance. Therefore, $\ln \xi_t$ converges to a chi-squared distribution with $|\mathcal{Y}| - 1$ degrees of freedom (due to the constraint $\sum_{y^i \in \mathcal{Y}} q_t(y^i) = 1$). Since this is an unbounded distribution, $\ln \xi_t$ crosses any fixed threshold $\alpha > 1$ at least once almost surely. This further implies that $\tilde{\lambda}_t$ will exceed any fixed α almost surely. Therefore, $\overline{\alpha} = \infty$. □

F.2 The Min-Likelihood Ratio Test (Min-LRT)

Recall that under the Min-LRT, the agent computes the likelihood ratio using minimum likelihood estimates:

$$\lambda_t^{\min} = \frac{\ell_t^{\min}(\theta')}{\ell_t^{\min}(\theta)},$$

where

$$\ell_t^{\min}(\theta) = \min_{\omega \in \Omega^\theta} \ell_t(\theta, \omega), \quad \ell_t^{\min}(\theta') = \min_{\omega' \in \Omega^{\theta'}} \ell_t(\theta', \omega').$$

The agent switches to θ' if $\lambda_t^{\min}$ exceeds α and to θ if $\lambda_t^{\min}$ falls below $1/\alpha$. Theorem 5 states that only singleton models with perfect asymptotic accuracy can be globally robust. As an immediate corollary, any such model is globally robust at all priors.

Theorem 5. *Suppose the agent uses the Min-LRT rule and considers $\theta \in \Theta$ that satisfies the no-trap condition. Then, for any $\alpha \geq 1$, model θ is globally robust at any full-support prior π_0^θ if and only if θ is a singleton model and has perfect asymptotic accuracy, i.e., $\Omega^\theta = C^\theta$ and $|\Omega^\theta| = 1$.*

Proof. Take any $\theta \in \Theta$ that satisfies the no-trap condition. I first prove the "only if" direction. Suppose that θ is globally robust, but either not a singleton model or does not have perfect asymptotic accuracy. Take the competing model θ' be the true model, with $\Omega^{\theta'} = \omega^*$ and $q^{\theta'}(\cdot|a, \omega^*) = q^*(\cdot|a)$ for all $a \in \mathcal{A}$. Then the min-likelihood ratio is

$$\lambda_t^{\min} = \frac{\prod_{\tau=0}^{t} q^*(y_\tau|a_\tau)}{\min_{\omega \in \Omega^\theta} \prod_{\tau=0}^{t} q^\theta(y_\tau|a_\tau, \omega)}.$$

Taking the logarithm of both sides,

$$\ln \lambda_t^{\min} = \max_{\omega \in \Omega^\theta} \left[\sum_{\tau=1}^{t} \left(\ln q^*(y_\tau|a_\tau) - \ln q^\theta(y_\tau|a_\tau, \omega) \right) \right].$$

For every $\omega \in \Omega^\theta$, define $A^-(\omega) := \{a \in \mathcal{A} : q^\theta(\cdot|a, \omega) \neq q^*(\cdot|a)\}$. By the same argument used in the proof of Lemma 1, for $\lambda_t^{\min}$ to stay under α forever, every action in $A^-(\omega)$ must be played at most finite times for every $\omega \in \Omega^\theta$. This is only possible if $\hat{A} := A \setminus (\cup_{\omega \in \Omega^\theta} A^-(\omega)) \neq \emptyset$. By definition, any $a \in \hat{A}$ satisfies $q^\theta(\cdot|a, \omega) \equiv q^*(\cdot|a)$ for all $\omega \in \Omega^\theta$. Since the no-trap condition holds, there cannot be distinct $\omega, \omega' \in \Omega^\theta$ such that $q^\theta(\cdot|a, \omega) \equiv q^\theta(\cdot|a, \omega') \equiv q^*(\cdot|a)$ for any a. Therefore, Ω^θ must be a singleton, say $\{\hat{\omega}\}$, in which case $C^\theta = \emptyset$. It follows that the agent's belief within model θ does not change over time, and she takes the same action under θ, which must lie in $A^-(\hat{\omega})$. As a result, $\lambda_t^{\min}$ eventually exceeds α almost surely, contradicting the assumption that θ is globally robust.

Next, I prove the "if" direction. Suppose that θ is a singleton model and $\Omega^\theta = C^\theta = \{\hat{\omega}\}$. I now show that this model is globally robust. First, conditional on the agent not switching away from θ, the agent's belief $\pi_0^\theta = \mathbf{1}_{\hat{\omega}}$ stays unchanged, so she must take the same action, say $\hat{a}$. By definition, $q^\theta(\cdot|\hat{a}, \hat{\omega}) = q^*(\cdot|\hat{a})$. The min-likelihood ratio is then given by

$$\lambda_t^{\min} = \frac{\min_{\omega' \in \Omega^{\theta'}} \prod_{\tau=0}^{t} q^\theta(y_\tau|a_\tau, \omega')}{\prod_{\tau=0}^{t} q^*(y_\tau|a_\tau)} = \min_{\omega' \in \Omega^{\theta'}} \frac{\prod_{\tau=0}^{t} q^\theta(y_\tau|a_\tau, \omega')}{\prod_{\tau=0}^{t} q^*(y_\tau|a_\tau)}.$$

This is a supermartingale which, by Ville's maximal inequality, stays under α with positive probability. Therefore, model θ persists against any competing model θ' and is thus globally robust. $\square$

F.3 Multiple Competing Models

Let $\Theta' \subseteq \Theta$ denote the set of competing models that the agent considers simultaneously, and define $\Theta^\dagger := \Theta' \cup \{\theta\}$ as the full set of models considered, including the initial model θ. I assume that Θ' is finite, containing at most $K \geq 1$ distinct models. At the beginning of each period, the agent compares her current model to all alternatives in $\Theta^\dagger$ and switches to the most plausible one if it fits the data sufficiently better. Specifically, the agent calculates Bayes factors for each model in $\Theta^\dagger$ against the model used just now, represented by $\boldsymbol{\lambda}_t := (\lambda_t^{\theta'})_{\theta' \in \Theta^\dagger}$, where $\lambda_t^{\theta'} = \ell_t(\theta')/\ell_t(m_t)$. The agent switches if $\max_{\theta' \in \Theta^\dagger} \lambda_t^{\theta'} > \alpha$, adopting the model with the highest Bayes factor. Model θ is globally robust at prior π_0^θ if it persists against every $\Theta' \subseteq \Theta$ of size no larger than K at π_0^θ and each corresponding vector of priors $\pi_0^{\Theta'}$. The definition of local robustness is modified similarly.

Below is an example where the number of competing models K exceeds $1 + \alpha$, and the agent eventually switches away from the true model to one of the misspecified models and then stops switching with probability 1.

Example 5 (Overfitting)**.** Consider an agent who repeatedly chooses between two actions, $\mathcal{A} = \{a^1, a^2\}$. The true DGP prescribes a uniform distribution over K outcomes $\mathcal{Y} = \{1, ..., K\}$ for both actions. The agent incurs a loss of $-K$ for the outcome $y = 1$ while receiving a payoff of 0 from all other outcomes. The agent pays an additional cost $c > 0$ for playing a^1 and no cost if she plays a^2. Assuming that the agent's initial model θ is the true model θ^*, she optimally plays a^2 in the first period to avoid the cost. Suppose the agent evaluates K competing models that I describe below. Each model $\theta^k \in \{\theta^1, ..., \theta^K\}$ has a single parameter ω^k. When a^1 is played, model θ^k agrees with θ, correctly predicting a uniform outcome distribution. When a^2 is played, model θ^k diverges from θ. Specifically, for any $k > 1$, θ^k predicts

$$q^{\theta^k}(y|a^2, \omega^k) = \begin{cases} 1 - \frac{1}{K} - (K-1)\eta & \text{if } y = k, \\ \frac{1}{K} + \eta & \text{if } y = 1, \\ \eta & \text{if } y \in \mathcal{Y} \setminus \{1, k\}, \end{cases}$$

where η is a small positive constant. When $k = 1$, $q^{\theta^k}(\cdot|a^2, \omega^k)$ is given by

$$q^{\theta^1}(y|a^2, \omega^1) = \begin{cases} 1 - (K-1)\eta & \text{if } y = 1, \\ \eta & \text{if } y \in \mathcal{Y} \setminus \{1\}. \end{cases}$$

Note that model θ^k predicts that when a^2 is played, the outcome k is drawn with probability near 1. Given there is one such model for every possible outcome, the agent must switch to one of these competing models upon the first outcome realization, provided that η is sufficiently small.. In particular, if the realized outcome is k, the agent immediately switches to model θ^k when

$$\frac{\ell_0(\theta^k)}{\ell_0(\theta)} = \frac{1 - \frac{1}{K} - (K-1)\eta}{\frac{1}{K}} > \alpha.$$

Note that such η exists as $K > \alpha + 1$. Furthermore, since playing a^2 leads to the outcome $y = 1$ with probability larger than $1/K$ under every competing model, once the switch occurs, the agent finds it optimal to play a^1 to avoid the loss associated with outcome 1 when c is sufficiently small. However, since all models yield the same correct predictions under a^1, the Bayes factors $\boldsymbol{\lambda_t}$ remain constant thereafter. Hence, despite that the agent starts with the true model, the agent becomes permanently trapped with a misspecified model and chooses a suboptimal action.

Next, I show that if $\alpha > K$, perfect asymptotic accuracy is still both sufficient and necessary for global robustness, given prior flexibility.

Theorem 6. *Suppose that the agent considers at most K competing models and that $\alpha > K$. Model $\theta \in \Theta$ is locally and globally robust for at least one prior if and only if there exists a p-absorbing SCE under θ, i.e., $C^\theta \neq \emptyset$.*

Proof. It suffices to show that when $\alpha > K$, a model θ is globally robust for at least one full-support prior if θ admits a p-absorbing SCE. Without loss of generality, take any $\Theta' = \{\theta^1, ..., \theta^K\} \subseteq \Theta$ and define for each $k \in \{1, ..., K\}$ a process $\{S_t^k\}_t$ as follows,

$$S_t^k = \frac{\sum_{\omega' \in \Omega^{\theta^k}} \pi_0^{\theta^k}(\omega') \prod_{\tau=0}^t q^{\theta^k}(y_\tau | a_\tau, \omega')}{\prod_{\tau=0}^t q^*(y_\tau | a_\tau)}.$$

Then for any $\eta \in (1, \alpha)$, by Ville's maximal inequality we have

$$\mathbb{P}_D(S_t^k \leq \eta, \forall t \geq 0) \geq 1 - \frac{\mathbb{E}^{\mathbb{P}_D} S_0^k}{\eta} = 1 - \frac{1}{\eta}.$$

Hence, when η is sufficiently close to α,

$$\begin{aligned}
&\mathbb{P}_D(S_t^k \leq \eta, \forall t \geq 0, \forall k \in \{1, ..., K\}) \\
\geq &1 - \sum_{k=1}^{K} \mathbb{P}_D(S_t^k > \eta \text{ for some } t \geq 0) \\
\geq &1 - \frac{K}{\eta} > 0.
\end{aligned}$$

The rest of the argument is similar to Step 1 of the proof of Theorem 1 in Appendix B.2. □

F.4 Forward-Looking Agent

To formally introduce a forward-looking agent within each model, consider an agent who assumes that she will continue using her current model m_t and maximizes the expected discounted sum of payoffs under it. The optimal policy f^θ is a selection from the correspondence $A^\theta : \Delta\Omega^\theta \rightrightarrows \mathcal{A}$, which solves the following dynamic programming problem,

$$U^\theta\left(\pi_t^\theta\right) = \max_{a \in \mathcal{A}} \sum_{\omega \in \Omega^\theta} \pi_t^\theta(\omega) \int_{y \in \mathcal{Y}} \left[u(a, y) + \delta U^\theta\left(B^\theta\left(a, y, \pi_t^\theta\right)\right)\right] q^\theta(y|a, \omega)\, v(dy), \quad (9)$$

where $\delta \in (0, 1)$ is the discount factor.

Theorems 1 to 3 go through without changes as their proofs do not require the assumption that the agent is fully myopic (see the remark in the first paragraph of Appendix B). However, since experimentation motives make p-absorbingness harder to achieve, I provide stronger sufficient conditions in Lemma 10. In particular, any uniformly quasi-strict SCE is p-absorbing. An SCE σ with supporting belief π is *uniformly quasi-strict* if $\text{supp}(\sigma) = A_M^\theta(\pi)$ for every belief $\pi \in \Delta\Omega^\theta(\sigma)$.

Lemma 10. *Suppose a θ-modeler has discount factor $\delta \in (0, 1)$. Suppose σ is a uniformly quasi-strict SCE with supporting belief $\hat{\pi}$, then for any $\gamma \in (0, 1)$, there exists $\epsilon > 0$ such that starting from any prior $\pi_0^\theta \in B_\epsilon(\hat{\pi})$, the probability that the θ-modeler always plays actions in* $\text{supp}(\sigma)$ *for all periods is strictly larger than γ.*

Proof. Since σ is uniformly quasi-strict with supporting belief $\hat{\pi}$, $\text{supp}(\sigma)$ contains all actions that can be myopically optimal for any degenerate belief δ_ω at $\omega \in \text{supp}(\hat{\pi})$. This implies that other actions have no experimentation value at $\hat{\pi}$ and $\text{supp}(\sigma)$ is also dynamically optimal against $\hat{\pi}$. Further, since A^{θ} is upper hemicontinuous (by Lemma 6), there exists $\tilde{\epsilon} > 0$ small enough such that $\text{supp}\,(\sigma) = A^{\theta}\,(\tilde{\pi})$ for all $\tilde{\pi} \in B_{\tilde{\epsilon}}\,(\hat{\pi})$. The rest of the proof is identical to the proof of Lemma 5. □

In the case of a fully forward-looking agent who anticipates that her model might change in the future, the characterization of robust models is much more complicated and is beyond the scope of this paper.

F.5 Infinite Parameter Space

Consider an extension of the framework introduced in Section 3: expand the model universe Θ to include any model that consists of arbitrarily many—finite or infinite—of DGPs that satisfy Assumption 2. Even though the competing model could be more complex, perfect asymptotic accuracy remains a sufficient and necessary condition for any model θ to be globally and locally robust for at least one full-support prior. That is, Theorem 1 remains valid. However, if θ has infinitely many predictions, a full-support prior $\pi_0^{\theta} \in \Delta\Omega^{\theta}$ may not assign positive probability mass to C^{θ}, which is needed for local robustness at π_0^{θ}. Nevertheless, Theorem 2 still holds if we replace statement (ii) with the following: model θ is locally robust at prior π_0^{θ} if and only if $\pi_0^{\theta}(C^{\theta}) > 0$. Finally, Theorem 2(i) and Theorem 3 remain unchanged since $\pi_0^{\theta}(C^{\theta}) \geq 1/\alpha$ and $C^{\theta} = \Omega^{\theta}$ already imply this condition.

F.6 Alternative Persistence Definitions

The persistence definition in Section 3 requires that if the agent initially adopts a model, she will eventually settle on it with positive probability. This definition has a natural interpretation and helps predict whether a particular bias can persist stably in a large population. By relaxing or strengthening parts of this definition, we alternative formulations that are also worth exploring.

Almost sure eventual adoption. One possible modification strengthens persistence by requiring that the model is eventually adopted *almost surely*. However, this makes both global and local robustness impossible. In fact, for any model θ, we can construct a nearby competing model θ' that replaces θ permanently with positive probability. The key idea is that the agent may encounter a sequence of outcome realizations better explained by θ', leading to a stable switch. Since the two models' predictions are initially close and converge in the limit, the agent may not switch back. Thus, almost-sure persistence is too restrictive to be useful.

To see this, construct θ' to include all DGPs in θ plus one additional DGP that differs from all other DGPs for all actions. That is, let $\Omega^{\theta'} = \Omega^{\theta} \cup \{\hat{\omega}\}$, where $q^{\theta'}(\cdot|a,\omega) = q^{\theta}(\cdot|a,\omega)$ for shared parameters, and $q^{\theta'}(\cdot|a,\hat{\omega}) \neq q^{\theta}(\cdot|a,\omega)$ for all $\omega \in \Omega^{\theta}$ and all $a \in \mathcal{A}$. Setting the prior $\pi_0^{\theta'}$ to be proportional to π_0^{θ} for shared parameters, the Bayes factor λ_t is bounded below by $\pi_0^{\theta'}(\Omega^{\theta})$. Since $\hat{\omega}$ makes distinct predictions from θ, there is a positive probability that the agent finds θ' sufficiently compelling to switch and never returns if $\pi_0^{\theta'}(\Omega^{\theta}) > 1/\alpha$, which holds if $\alpha > 1$ and $\pi_0^{\theta'}(\Omega^{\theta})$ is close to 1.

No switch. The current persistence definition allows back-and-forth switching before the agent eventually settles on a model. A stricter definition would require the agent to never switch once adopting a model. The main results remain valid under this stricter condition. Intuitively, for the results to change, some initial models must *require* back-and-forth switching to be persistent. However, in Theorem 1, the prior can be concentrated around a p-absorbing SCE, ensuring the agent never needs to switch away from an asymptotically accurate model. In Theorem 2, the no-trap condition guarantees that the agent can draw outcomes and reach p-absorbing equilibria from any qualified prior without needing temporary switches.

References

Blackwell, D. (1965): "Discounted Dynamic Programming," *The Annals of Mathematical Statistics*, 36, 226–235.

Maitra, A. (1968): "Discounted Dynamic Programming on Compact Metric Spaces," *Sankhyā: The Indian Journal of Statistics, Series A*, 211–216.