Econometrica Supplementary Material

SUPPLEMENT TO "BERK–NASH EQUILIBRIUM: A FRAMEWORK FOR MODELING AGENTS WITH MISSPECIFIED MODELS"

(*Econometrica*, Vol. 84, No. 3, May 2016, 1093–1130)

BY IGNACIO ESPONDA AND DEMIAN POUZO

APPENDIX B: EXAMPLE: TRADING WITH ADVERSE SELECTION

IN THIS SECTION, we provide the formal details for the trading environment in Example 2.5. Let $p \in \Delta(\mathbb{A} \times \mathbb{V})$ be the true distribution; we use subscripts, such as p_A and $p_{V|A}$, to denote the corresponding marginal and conditional distributions. Let $\mathbb{Y} = \mathbb{A} \times \mathbb{V} \cup \{\square\}$ denote the space of observable consequences, where $\square$ will be a convenient way to represent the fact that there is no trade. We denote the random variable taking values in $\mathbb{V} \cup \{\square\}$ by $\hat{V}$. Notice that the state space in this example is $\Omega = \mathbb{A} \times \mathbb{V}$.

Partial feedback is represented by the function $f^P : \mathbb{X} \times \mathbb{A} \times \mathbb{V} \to \mathbb{Y}$ such that $f^P(x, a, v) = (a, v)$ if $a \leq x$ and $f^P(x, a, v) = (a, \square)$ if $a > x$. Full feedback is represented by $f^F(x, a, v) = (a, v)$. In all cases, payoffs are given by $\pi : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$, where $\pi(x, a, v) = v - x$ if $a \leq x$ and 0 otherwise. The objective distribution for the case of partial feedback, Q^P, is, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}$, $Q^P(a, v \mid x) = p(a, v)1_{\{x \geq a\}}(x)$, and, $\forall x \in \mathbb{X}$, $\forall a \in \mathbb{A}$, $Q^P(a, \square \mid x) = p_A(a)1_{\{x<a\}}(x)$. The objective distribution for the case of full feedback, Q^F, is, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}$, $Q^F(a, v \mid x) = p(a, v)$, and, $\forall x \in \mathbb{X}$, $\forall a \in \mathbb{A}$, $Q^F(a, \square \mid x) = 0$.

The buyer knows the environment except for the distribution $p \in \Delta(\mathbb{A} \times \mathbb{V})$. Then, for any distribution in the subjective model, Q_θ, the perceived expected profit from choosing $x \in \mathbb{X}$ is

$$E_{Q_\theta(\cdot|x)}\big[\pi(x, A, \hat{V})\big] = \sum_{(a,v)\in\mathbb{A}\times\mathbb{V}} 1_{\{x\geq a\}}(x)(v - x)Q_\theta(a, v \mid x). \tag{17}$$

The buyer has either one of two misspecifications over p captured by the parameter sets $\Theta_I = \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$ (i.e., independent beliefs) or $\Theta_A = \times_j \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$ (i.e., analogy-based beliefs) defined in the main text. Thus, combining feedback and parameter sets, we have four cases to consider, and, for each case, we write down the subjective model and wKLD function.

CURSED EQUILIBRIUM: Feedback is f^F and the parameter set is Θ_I. The subjective model is, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}$, $Q^C_\theta(a, v \mid x) = \theta_A(a)\theta_V(v)$, and,

 DOI: 10.3982/ECTA12609

$\forall x \in \mathbb{X}$, $\forall a \in \mathbb{A}$, $Q^C_\theta(a, \square \mid x) = 0$, where $\theta = (\theta_A, \theta_V) \in \Theta_I$.[47] This is an analogy-based game. From (17), the perceived expected profit from $x \in \mathbb{X}$ is

$$\Pr_{\theta_A}(A \leq x)\big(E_{\theta_V}[V] - x\big), \tag{18}$$

where $\Pr_{\theta_A}$ denotes probability with respect to θ_A and E_{θ_V} denotes expectation with respect to θ_V. Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is[48]

$$\begin{aligned} K^C(x, \theta) &= E_{Q^F(\cdot|x)}\left[\ln \frac{Q^F(A, \hat{V} \mid x)}{Q^C_\theta(A, \hat{V} \mid x)}\right] \\ &= \sum_{(a,v) \in \mathbb{A}\times\mathbb{V}} p(a, v) \ln \frac{p(a, v)}{\theta_A(a)\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_A(x), \theta_V(x)) \in \Theta_I = \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$, where $\theta_A(x) = p_A$ and $\theta_V(x) = p_V$ is the unique parameter value that minimizes $K^C(x, \cdot)$. Together with (18), we obtain equation Π^{CE} in the main text.

BEHAVIORAL EQUILIBRIUM (NAIVE VERSION): Feedback is f^P and the parameter set is Θ_I. The subjective model is, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}$, $Q^{\mathrm{BE}}_\theta(a, v \mid x) = \theta_A(a)\theta_V(v)1_{\{x \geq a\}}(x)$, and, $\forall x \in \mathbb{X}$, $\forall a \in \mathbb{A}$, $Q^{\mathrm{BE}}_\theta(a, \square \mid x) = \theta_A(a)1_{\{x<a\}}(x)$, where $\theta = (\theta_A, \theta_V) \in \Theta_I$. From (17), perceived expected profit from $x \in \mathbb{X}$ is as in equation (18). Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is

$$\begin{aligned} K^{\mathrm{BE}}(x, \theta) &= E_{Q^P(\cdot|x)}\left[\ln \frac{Q^P(A, \hat{V} \mid x)}{Q^{\mathrm{BE}}_\theta(A, \hat{V} \mid x)}\right] \\ &= \sum_{\{a \in \mathbb{A}: a > x\}} p_A(a) \ln \frac{p_A(a)}{\theta_A(a)} \\ &\quad + \sum_{\{(a,v) \in \mathbb{A}\times\mathbb{V}: a \leq x\}} p(a, v) \ln \frac{p(a, v)}{\theta_A(a)\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_A(x), \theta_V(x)) \in \Theta_I = \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$, where $\theta_A(x) = p_A$ and $\theta_V(x)(v) = p_{V|A}(v \mid A \leq x)$ $\forall v \in \mathbb{V}$ is the unique parameter value that minimizes $K^{\mathrm{BE}}(x, \cdot)$. Together with (18), we obtain equation Π^{BE} in the main text.

ANALOGY-BASED EXPECTATIONS EQUILIBRIUM: Feedback is f^F and the parameter set is Θ_A. The subjective model is, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}_j$, all

[47] In fact, the symbol $\square$ is not necessary for this example, but we keep it so that all feedback functions are defined over the same space of consequences.

[48] In all cases, the extension to mixed strategies is straightforward.

$j = 1, \dots, k$, $Q_\theta^{\text{ABEE}}(a, v \mid x) = \theta_j(a)\theta_V(v)$, and, $\forall x \in \mathbb{X}$, $\forall a \in \mathbb{A}$, $Q_\theta^{\text{ABEE}}(a, \square \mid x) = 0$, where $\theta = (\theta_1, \dots, \theta_k, \theta_V) \in \Theta_A$. This is an analogy-based game. From (17), perceived expected profit from $x \in \mathbb{X}$ is

$$\sum_{j=1}^{k} \Pr_{\theta_V}(V \in \mathbb{V}_j)\bigl\{\Pr_{\theta_j}(A \le x \mid V \in \mathbb{V}_j)\bigl(E_{\theta_V}[V \mid V \in \mathbb{V}_j] - x\bigr)\bigr\}. \tag{19}$$

Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is

$$\begin{aligned} K^{\text{ABEE}}(x, \theta) &= E_{Q^F(\cdot \mid x)}\left[\ln \frac{Q^F(A, \hat{V} \mid x)}{Q_\theta^{\text{ABEE}}(A, \hat{V} \mid x)}\right] \\ &= \sum_{j=1}^{k} \sum_{(a,v) \in \mathbb{A} \times \mathbb{V}_j} p(a, v) \ln \frac{p(a, v)}{\theta_j(a)\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_1(x), \dots, \theta_k(x), \theta_V(x)) \in \Theta_A = \bigtimes_j \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$, where $\theta_j(x)(a) = p_{A|V_j}(a \mid V \in \mathbb{V}_j)$ $\forall a \in \mathbb{A}$ and $\theta_V(x) = p_V$ is the unique parameter value that minimizes $K^{\text{ABEE}}(x, \cdot)$. Together with (19), we obtain equation Π^{ABEE} in the main text.

BEHAVIORAL EQUILIBRIUM (NAIVE VERSION) WITH ANALOGY CLASSES: It is natural to also consider a case, unexplored in the literature, where feedback f^P is partial and the subjective model is parameterized by Θ_A. Suppose that the buyer's behavior has stabilized to some price x^*. Due to the possible correlation across analogy classes, the buyer might now believe that deviating to a different price $x \neq x^*$ affects her valuation. In particular, the buyer might have multiple beliefs at x^*. To obtain a natural equilibrium refinement, we assume that the buyer also observes the analogy class that contains her realized valuation, whether she trades or not, and that $\Pr(V \in \mathbb{V}_j, A \le x) > 0$ for all $j = 1, \dots, k$ and $x \in \mathbb{X}$.[49] We denote this new feedback assumption by a function $f^{P^*} : \mathbb{X} \times \mathbb{A} \times \mathbb{V} \to \mathbb{Y}^*$, where $\mathbb{Y}^* = \mathbb{A} \times \mathbb{V} \cup \{1, \dots, k\}$ and $f^{P^*}(x, a, v) = (a, v)$ if $a \le x$ and $f^{P^*}(x, a, v) = (a, j)$ if $a > x$ and $v \in \mathbb{V}_j$. The objective distribution given this feedback function is, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}$, $Q^{P^*}(a, v \mid x) = p(a, v)1_{\{x \ge a\}}(x)$, and, $\forall x \in \mathbb{X}$, $\forall a \in \mathbb{A}$ and all $j = 1, \dots, k$, $Q^{P^*}(a, j \mid x) = p_{A|V_j}(a \mid V \in \mathbb{V}_j)p_V(\mathbb{V}_j)1_{\{x < a\}}(x)$. The subjective model is, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}_j$ and all $j = 1, \dots, k$, $Q_\theta^{\text{BEA}}(a, v \mid x) = \theta_j(a)\theta_V(v)1_{\{x \ge a\}}(x)$, and, $\forall x \in \mathbb{X}$, $\forall (a, v) \in \mathbb{A} \times \mathbb{V}_j$ and all $j = 1, \dots, k$, $Q_\theta^{\text{BEA}}(a, j \mid x) = \theta_j(a)(\sum_{v \in \mathbb{V}_j} \theta_V(v))1_{\{x < a\}}(x)$, where $\theta = (\theta_1, \theta_2, \dots, \theta_k, \theta_V) \in$

[49] Alternatively, and more naturally, we could require the equilibrium to be the limit of a sequence of mixed strategy equilibria with the property that all prices are chosen with positive probability.

Θ_A. In particular, from (17), perceived expected profit from $x \in \mathbb{X}$ is as in equation (19). Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is

$$\begin{aligned} K^{\mathrm{BEA}}(x,\theta) &= E_{Q^{P^*}(\cdot\mid x)}\left[\ln \frac{Q^{P^*}(A,\hat{V}\mid x)}{Q_\theta^{\mathrm{BEA}}(A,\hat{V}\mid x)}\right] \\ &= \sum_{j=1}^{k} \sum_{\{(a,v)\in\mathbb{A}\times\mathbb{V}_j : a\le x\}} p(a,v)\ln\frac{p(a,v)}{\theta_j(a)\theta_V(v)} \\ &\quad + \sum_{\{(a,j)\in\mathbb{A}\times\{1,\dots,k\}:a>x\}} p_{A\mid V_j}(a\mid V\in\mathbb{V}_j)\,p_V(\mathbb{V}_j) \\ &\quad \times \ln\frac{p_{A\mid V_j}(a\mid V\in\mathbb{V}_j)\,p_V(\mathbb{V}_j)}{\theta_j(a)\sum_{v\in\mathbb{V}_j}\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_1(x),\dots,\theta_k(x),\theta_V(x)) \in \Theta_A = \bigtimes_j \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$, where $\theta_j(x)(a) = p_{A\mid V_j}(a \mid V \in \mathbb{V}_j)$ $\forall a \in \mathbb{A}$ and $\theta_V(x)(v) = p_{V\mid A}(v \mid V \in \mathbb{V}_j, A \le x)p_V(\mathbb{V}_j)$ $\forall v \in \mathbb{V}_j$, all $j = 1,\dots,k$ is the unique parameter value that minimizes $K^{\mathrm{BEA}}(x,\cdot)$. Together with (19), we obtain $\Pi^{\mathrm{BEA}}(x,x^*) = \sum_{i=j}^{k} \Pr(V \in \mathbb{V}_j)\Pr(A \le x \mid V \in \mathbb{V}_j)(E[V \mid V \in \mathbb{V}_j, A \le x^*] - x)$.

APPENDIX C: PROOF OF CONVERSE RESULT: THEOREM 3

Let $(\bar{\mu}^i)_{i\in I}$ be a belief profile that supports σ as an equilibrium. Consider the following policy profile $\phi = (\phi_t^i)_{i,t}$: For all $i \in I$ and all t,

$$\begin{aligned} (\mu^i, s^i, \xi^i) &\mapsto \phi_t^i(\mu^i, s^i, \xi^i) \\ &\equiv \begin{cases} \varphi^i(\bar{\mu}^i, s^i, \xi^i) & \text{if } \max_{i\in I}\|\bar{Q}^i_{\mu^i} - \bar{Q}^i_{\bar{\mu}^i}\| \le \frac{1}{2C}\varepsilon_t, \\ \varphi^i(\mu^i, s^i, \xi^i) & \text{otherwise,} \end{cases} \end{aligned}$$

where φ^i is an arbitrary selection from Ψ^i, $C \equiv \max_I\{\#\mathbb{Y}^i \times \sup_{\mathbb{X}^i\times\mathbb{Y}^i}|\pi^i(x^i, y^i)|\} < \infty$, and the sequence $(\varepsilon_t)_t$ will be defined below. For all $i \in I$, fix any prior μ_0^i with full support on Θ^i such that $\mu_0^i(\cdot \mid \Theta^i(\sigma)) = \bar{\mu}^i$ (where for any $A \subset \Theta$ Borel, $\mu(\cdot \mid A)$ is the conditional probability given A).

We now show that if $\varepsilon_t \ge 0$ $\forall t$ and $\lim_{t\to\infty}\varepsilon_t = 0$, then ϕ is asymptotically optimal. Throughout this argument, we fix an arbitrary $i \in I$. Abusing notation, let $U^i(\mu^i, s^i, \xi^i, x^i) = E_{\bar{Q}_{\mu^i}(\cdot\mid s^i, x^i)}[\pi^i(x^i, Y^i)] + \xi^i(x^i)$. It suffices to show that

$$U^i(\mu^i, s^i, \xi^i, \phi_t^i(\mu^i, s^i, \xi^i)) \ge U^i(\mu^i, s^i, \xi^i, x^i) - \varepsilon_t \tag{20}$$

for all (i,t), all (μ^i, s^i, ξ^i), and all x^i. By construction of ϕ, equation (20) is satisfied if $\max_{i\in I}\|\bar{Q}^i_{\mu^i} - \bar{Q}^i_{\bar{\mu}^i}\| > \frac{1}{2C}\varepsilon_t$. If, instead, $\max_{i\in I}\|\bar{Q}^i_{\mu^i} - \bar{Q}^i_{\bar{\mu}^i}\| \le \frac{1}{2C}\varepsilon_t$, then

$$U^i\big(\bar{\mu}^i, s^i, \xi^i, \phi^i_t\big(\mu^i, s^i, \xi^i\big)\big) = U^i\big(\bar{\mu}^i, s^i, \xi^i, \varphi^i\big(\bar{\mu}^i, s^i, \xi^i\big)\big) \geq U^i\big(\bar{\mu}^i, s^i, \xi^i, x^i\big) \tag{21}$$

$\forall x^i \in \mathbb{X}^i$. Moreover, $\forall x^i$,

$$\begin{aligned}
&\big|U^i\big(\bar{\mu}^i, s^i, \xi^i, x^i\big) - U^i\big(\mu^i, s^i, \xi^i, x^i\big)\big| \\
&\quad= \Big|\sum_{y^i\in\mathbb{Y}^i} \pi\big(x^i, y^i\big)\big(\bar{Q}^i_{\bar{\mu}^i}\big(y^i \mid s^i, x^i\big) - \bar{Q}^i_{\mu^i}\big(y^i \mid s^i, x^i\big)\big)\Big| \\
&\quad\leq \sup_{\mathbb{X}^i\times\mathbb{Y}^i}\big|\pi^i\big(x^i, y^i\big)\big| \sum_{y^i\in\mathbb{Y}^i}\big|\big(\bar{Q}^i_{\bar{\mu}^i}\big(y^i \mid s^i, x^i\big) - \bar{Q}^i_{\mu^i}\big(y^i \mid s^i, x^i\big)\big)\big| \\
&\quad\leq \sup_{\mathbb{X}^i\times\mathbb{Y}^i}\big|\pi^i\big(x^i, y^i\big)\big| \times \#\mathbb{Y}^i \times \max_{y^i,x^i,s^i}\big|\bar{Q}^i_{\bar{\mu}^i}\big(y^i \mid s^i, x^i\big) - \bar{Q}^i_{\mu^i}\big(y^i \mid s^i, x^i\big)\big|,
\end{aligned}$$

so by our choice of C, $|U^i(\bar{\mu}^i, s^i, \xi^i, x^i) - U^i(\mu^i, s^i, \xi^i, x^i)| \leq 0.5\varepsilon_t\ \forall x^i$. Therefore, equation (21) implies equation (20); thus ϕ is asymptotically optimal if $\varepsilon_t \geq 0\ \forall t$ and $\lim_{t\to\infty}\varepsilon_t = 0$.

We now construct a sequence $(\varepsilon_t)_t$ such that $\varepsilon_t \geq 0\ \forall t$ and $\lim_{t\to\infty}\varepsilon_t = 0$. Let $\bar{\phi}^i = (\bar{\phi}^i_t)_t$ be such that $\bar{\phi}^i_t(\mu^i, \cdot, \cdot) = \varphi^i(\bar{\mu}^i, \cdot, \cdot)\ \forall \mu^i$; that is, $\bar{\phi}^i$ is a stationary policy that maximizes utility under the assumption that the belief is always $\bar{\mu}^i$. Let $\zeta^i(\mu^i) \equiv 2C\|\bar{Q}^i_{\mu^i} - \bar{Q}^i_{\bar{\mu}^i}\|$ and suppose (the proof is at the end) that

$$\mathbf{P}^{\mu_0,\bar{\phi}}\Big(\lim_{t\to\infty}\max_{i\in I}\big|\zeta^i\big(\mu^i_t(h)\big)\big| = 0\Big) = 1 \tag{22}$$

(recall that $\mathbf{P}^{\mu_0,\bar{\phi}}$ is the probability measure over $\mathbb{H}$ induced by the policy profile $\bar{\phi}$; by definition of $\bar{\phi}$, $\mathbf{P}^{\mu_0,\bar{\phi}}$ does not depend on μ_0). Then by the second Borel–Cantelli lemma (Billingsley (1995, pp. 59–60)), for any $\gamma > 0$, $\sum_t \mathbf{P}^{\mu_0,\bar{\phi}}(\max_{i\in I}|\zeta^i(\mu^i_t(h))| \geq \gamma) < \infty$. Hence, for any $a > 0$, there exists a sequence $(\tau(j))_j$ such that

$$\sum_{t\geq\tau(j)} \mathbf{P}^{\mu_0,\bar{\phi}}\Big(\max_{i\in I}\big|\zeta^i\big(\mu^i_t(h)\big)\big| \geq 1/j\Big) < \frac{3}{a}4^{-j} \tag{23}$$

and $\lim_{j\to\infty}\tau(j) = \infty$. For all $t \leq \tau(1)$, we set $\varepsilon_t = 3C$, and, for any $t > \tau(1)$, we set $\varepsilon_t \equiv 1/N(t)$, where $N(t) \equiv \sum_{j=1}^{\infty} 1\{\tau(j) \leq t\}$. Observe that, since $\lim_{j\to\infty}\tau(j) = \infty$, $N(t) \to \infty$ as $t \to \infty$ and thus $\varepsilon_t \to 0$.

Next, we show that $\mathbf{P}^{\mu_0,\phi}(\lim_{t\to\infty}\|\sigma_t(h^\infty) - \sigma\| = 0) = 1$, where $(\sigma_t)_t$ is the sequence of intended strategies given ϕ, that is, $\sigma^i_t(h)(x^i \mid s^i) = P_\xi(\xi^i : \phi^i_t(\mu^i_t(h), s^i, \xi^i) = x^i)$. Observe that, by definition, $\sigma^i(x^i \mid s^i) = P_\xi(\xi^i : x^i \in$

$\arg\max_{\hat{x}^i \in \mathbb{X}^i} E_{\bar{Q}_{\bar{\mu}^i}(\cdot|s^i,\hat{x}^i)}[\pi^i(\hat{x}^i, Y^i)] + \xi^i(\hat{x}^i))$. Since $\varphi^i \in \Psi^i$, it follows that we can write $\sigma^i(x^i \mid s^i) = P_\xi(\xi^i : \varphi^i(\bar{\mu}^i, s^i, \xi^i) = x^i)$. Let $H \equiv \{h\colon \|\sigma_t(h) - \sigma\| = 0$, for all $t\}$. It is sufficient to show that $\mathbf{P}^{\mu_0,\phi}(H) = 1$. To show this, observe that

$$\begin{aligned}
\mathbf{P}^{\mu_0,\phi}(H) \geq \mathbf{P}^{\mu_0,\phi}\left(\bigcap_t \left\{\max_i \zeta^i(\mu_t) \leq \varepsilon_t\right\}\right)
&= \prod_{t=\tau(1)+1}^{\infty} \mathbf{P}^{\mu_0,\phi}\left(\max_i \zeta^i(\mu_t) \leq \varepsilon_t \middle| \bigcap_{l<t}\left\{\max_i \zeta^i(\mu_l) \leq \varepsilon_l\right\}\right) \\
&= \prod_{t=\tau(1)+1}^{\infty} \mathbf{P}^{\mu_0,\bar{\phi}}\left(\max_i \zeta^i(\mu_t) \leq \varepsilon_t \middle| \bigcap_{l<t}\left\{\max_i \zeta^i(\mu_l) \leq \varepsilon_l\right\}\right) \\
&= \mathbf{P}^{\mu_0,\bar{\phi}}\left(\bigcap_{t>\tau(1)} \left\{\max_i \zeta^i(\mu_t) \leq \varepsilon_t\right\}\right),
\end{aligned}$$

where the second line omits the term $\mathbf{P}^{\mu_0,\phi}(\max_i \zeta^i(\mu_t) < \varepsilon_t \text{ for all } t \leq \tau(1))$ because it is equal to 1 (since $\varepsilon_t \geq 3C\ \forall t \leq \tau(1)$); the third line follows from the fact that $\phi^i_{t-1} = \bar{\phi}^i_{t-1}$ if $\zeta^i(\mu_{t-1}) \leq \varepsilon_{t-1}$, so the probability measure is equivalently given by $\mathbf{P}^{\mu_0,\bar{\phi}}$; and where the last line also uses the fact that $\mathbf{P}^{\mu_0,\bar{\phi}}(\max_i \zeta^i(\mu_t) < \varepsilon_t \text{ for all } t \leq \tau(1)) = 1$. In addition, $\forall a > 0$,

$$\begin{aligned}
&\mathbf{P}^{\mu_0,\bar{\phi}}\left(\bigcap_{t>\tau(1)} \left\{\max_i \zeta^i(\mu_t) \leq \varepsilon_t\right\}\right) \\
&= \mathbf{P}^{\mu_0,\bar{\phi}}\left(\bigcap_{n\in\{1,2,\dots\}} \bigcap_{\{t>\tau(1):N(t)=n\}} \left\{\max_i \zeta^i(\mu_t) \leq n^{-1}\right\}\right) \\
&\geq 1 - \sum_{n=1}^{\infty} \sum_{\{t:N(t)=n\}} \mathbf{P}^{\mu_0,\bar{\phi}}\left(\max_i \zeta^i(\mu_t) \geq n^{-1}\right) \\
&\geq 1 - \sum_{n=1}^{\infty} \frac{3}{a} 4^{-n} = 1 - \frac{1}{a},
\end{aligned}$$

where the last line follows from (23). Thus, we have shown that $\mathbf{P}^{\mu_0,\phi}(H) \geq 1 - 1/a\ \forall a > 0$; hence, $\mathbf{P}^{\mu_0,\phi}(H) = 1$.

We conclude the proof by showing that equation (22) indeed holds. Observe that σ is trivially stable under $\bar{\phi}$. By Lemma 2, $\forall i \in I$ and all open sets $U^i \supseteq \Theta^i(\sigma)$,

$$\lim_{t\to\infty} \mu^i_t(U^i) = 1 \tag{24}$$

a.s.-$\mathbf{P}^{\mu_0,\bar{\phi}}$ (over $\mathbb{H}$). Let $\mathcal{H}$ denote the set of histories such that $x_t^i(h) = x^i$ and $s_t^i(h) = s^i$ imply that $\sigma^i(x^i \mid s^i) > 0$. By definition of $\bar{\phi}$, $\mathbf{P}^{\mu_0,\bar{\phi}}(\mathcal{H}) = 1$. Thus, it suffices to show that $\lim_{t\to\infty} \max_{i\in I} |\zeta^i(\mu_t^i(h))| = 0$ a.s.-$\mathbf{P}^{\mu_0,\bar{\phi}}$ over $\mathcal{H}$. To do this, take any $A \subseteq \Theta$ that is closed. By equation (24), $\forall i \in I$, and almost all $h \in \mathcal{H}$,

$$\limsup_{t\to\infty} \int 1_A(\theta)\mu_{t+1}^i(d\theta) = \limsup_{t\to\infty} \int 1_{A\cap\Theta^i(\sigma)}(\theta)\mu_{t+1}^i(d\theta).$$

Moreover,

$$\begin{aligned}
&\int 1_{A\cap\Theta^i(\sigma)}(\theta)\mu_{t+1}^i(d\theta) \\
&\quad \le \int 1_{A\cap\Theta^i(\sigma)}(\theta) \left\{ \frac{\prod_{\tau=1}^{t} Q_\theta^i\big(y_\tau^i \mid s_\tau^i, x_\tau^i\big)\mu_0^i(d\theta)}{\int_{\Theta^i(\sigma)} \prod_{\tau=1}^{t} Q_\theta^i\big(y_\tau^i \mid s_\tau^i, x_\tau^i\big)\mu_0^i(d\theta)} \right\} \\
&\quad = \mu_0^i\big(A \mid \Theta^i(\sigma)\big) = \bar{\mu}^i(A),
\end{aligned}$$

where the first inequality follows from the fact that $\Theta^i(\sigma) \subseteq \Theta^i$; the first equality follows from the fact that, since $h \in \mathcal{H}$, the fact that the game is weakly identified given σ implies that $\prod_{\tau=1}^{t} Q_\theta^i(y_\tau^i \mid s_\tau^i, x_\tau^i)$ is constant with respect to θ $\forall\theta \in \Theta^i(\sigma)$, and the last equality follows from our choice of μ_0^i. Therefore, we established that a.s.-$\mathbf{P}^{\mu_0,\bar{\phi}}$ over $\mathcal{H}$, $\limsup_{t\to\infty} \mu_{t+1}^i(h)(A) \le \bar{\mu}^i(A)$ for A closed. By the portmanteau lemma, this implies that, a.s.-$\mathbf{P}^{\mu_0,\bar{\phi}}$ over $\mathcal{H}$, $\lim_{t\to\infty} \int_\Theta f(\theta)\mu_{t+1}^i(h)(d\theta) = \int_\Theta f(\theta)\bar{\mu}^i(d\theta)$ for any f real-valued, bounded, and continuous. Since, by assumption, $\theta \mapsto Q_\theta^i(y^i \mid s^i, x^i)$ is bounded and continuous, the previous result applies to $Q_\theta^i(y^i \mid s^i, x^i)$, and since y, s, x take a finite number of values, this result implies that $\lim_{t\to\infty} \|\bar{Q}_{\mu_t^i(h)}^i - \bar{Q}_{\bar{\mu}^i}^i\| = 0$ $\forall i \in I$ a.s.-$\mathbf{P}^{\mu_0,\bar{\phi}}$ over $\mathcal{H}$. *Q.E.D.*

APPENDIX D: NON-MYOPIC PLAYERS

In the main text, we proved the results for the case where players are myopic. Here, we assume that players maximize discounted expected payoffs, where $\delta^i \in [0, 1)$ is the discount factor of player i. In particular, players can be forward looking and decide to experiment. Players believe, however, that they face a stationary environment and, therefore, have no incentives to influence the future behavior of other players. We assume for simplicity that players know the distribution of their own payoff perturbations.

Because players believe that they face a stationary environment, they solve a (subjective) dynamic optimization problem that can be cast recursively as follows. By the Principle of Optimality, $V^i(\mu^i, s^i)$ denotes the maximum expected discounted payoffs (i.e., the value function) of player i who starts a period by observing signal s^i and by holding belief μ^i if and only if

$$V^i(\mu^i, s^i) = \int_{\Xi^i} \Big\{ \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}(\cdot|s^i,x^i)} \big[\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} \big[V^i(\hat{\mu}^i, S^i) \big] \big] \Big\} P_\xi(d\xi^i), \tag{25}$$

where $\hat{\mu}^i = B^i(\mu^i, s^i, x^i, Y^i)$ is the updated belief. For all (μ^i, s^i, ξ^i), let

$$\Phi^i(\mu^i, s^i, \xi^i) = \arg\max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}(\cdot|s^i,x^i)} \big[\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} \big[V^i(\hat{\mu}^i, S^i) \big] \big].$$

The proof of the next lemma relies on standard arguments and is, therefore, omitted.[50]

LEMMA 3: *There exists a unique solution V^i to the Bellman equation* (25); *this solution is bounded in $\Delta(\Theta^i) \times \mathbb{S}^i$ and continuous as a function of μ^i. Moreover, Φ^i is single-valued and continuous with respect to μ^i, a.s.-P_ξ.*

Because players believe they face a stationary environment with i.i.d. perturbations, it is without loss of generality to restrict behavior to depend on the state of the recursive problem. Optimality of a policy is defined as usual (with the requirement that $\phi^i_t \in \Phi^i \; \forall t$).

Lemma 2 implies that the *support* of posteriors converges, but posteriors need not converge. We can always find, however, a subsequence of posteriors that converges. By continuity of dynamic behavior in beliefs, the stable strategy profile is dynamically optimal (in the sense of solving the dynamic optimization problem) given this convergent posterior. For weakly identified games, the convergent posterior is a fixed point of the Bayesian operator. Thus, the players' limiting strategies will provide no new information. Since the value of experimentation is nonnegative, it follows that the stable strategy profile must also be myopically optimal (in the sense of solving the optimization problem that ignores the future), which is the definition of optimality used in the definition of Berk–Nash equilibrium. Thus, we obtain the following characterization of the set of stable strategy profiles when players follow optimal policies.

THEOREM 4: *Suppose that a strategy profile σ is stable under an optimal policy profile for a perturbed and weakly identified game. Then σ is a Berk–Nash equilibrium of the game.*

[50] Doraszelski and Escobar (2010) studied a similarly perturbed version of the Bellman equation.

PROOF: The first part of the proof is identical to the proof of Theorem 2. Here, we prove that, given that $\lim_{j\to\infty}\sigma_{t(j)}=\sigma$ and $\lim_{j\to\infty}\mu^i_{t(j)}=\mu^i_\infty\in\Delta(\Theta^i(\sigma))$ $\forall i$, then, $\forall i$, σ^i is optimal for the perturbed game given $\mu^i_\infty\in\Delta(\Theta^i)$, that is, $\forall(s^i,x^i)$,

$$\sigma^i(x^i\mid s^i)=P_\xi\big(\xi^i:\psi^i(\mu^i_\infty,s^i,\xi^i)=\{x^i\}\big), \tag{26}$$

where $\psi^i(\mu^i_\infty,s^i,\xi^i)\equiv\arg\max_{x^i\in\mathbb{X}^i}E_{\bar{Q}^i_{\mu^i_\infty}(\cdot\mid s^i,x^i)}[\pi^i(x^i,Y^i)]+\xi^i(x^i)$.

To establish (26), fix $i\in I$ and $s^i\in\mathbb{S}^i$. Then

$$\begin{aligned}\lim_{j\to\infty}\sigma^i_{t(j)}(h)\big(x^i\mid s^i\big)&=\lim_{j\to\infty}P_\xi\big(\xi^i:\phi^i_{t(j)}\big(\mu^i_{t(j)},s^i,\xi^i\big)=x^i\big)\\&=P_\xi\big(\xi^i:\Phi^i\big(\mu^i_\infty,s^i,\xi^i\big)=\{x^i\}\big),\end{aligned}$$

where the second line follows by optimality of ϕ^i and Lemma 3. This implies that $\sigma^i(x^i\mid s^i)=P_\xi(\xi^i:\Phi^i(\mu^i_\infty,s^i,\xi^i)=\{x^i\})$. Thus, it remains to show that

$$P_\xi\big(\xi^i:\Phi^i\big(\mu^i_\infty,s^i,\xi^i\big)=\{x^i\}\big)=P_\xi\big(\xi^i:\psi^i\big(\mu^i_\infty,s^i,\xi^i\big)=\{x^i\}\big) \tag{27}$$

$\forall x^i$ such that $P_\xi(\xi^i:\Phi^i(\mu^i_\infty,s^i,\xi^i)=\{x^i\})>0$. From now on, fix any such x^i. Since $\sigma^i(x^i\mid s^i)>0$, the assumption that the game is weakly identified implies that $Q^i_{\theta^i_1}(\cdot\mid x^i,s^i)=Q^i_{\theta^i_2}(\cdot\mid x^i,s^i)$ $\forall\theta^i_1,\theta^i_2\in\Theta(\sigma)$. The fact that $\mu^i_\infty\in\Delta(\Theta^i(\sigma))$ then implies that

$$B^i\big(\mu^i_\infty,s^i,x^i,y^i\big)=\mu^i_\infty \tag{28}$$

$\forall y^i\in\mathbb{Y}^i$. Thus, $\Phi^i(\mu^i_\infty,s^i,\xi^i)=\{x^i\}$ is equivalent to

$$\begin{aligned}&E_{\bar{Q}_{\mu^i_\infty}(\cdot\mid s^i,x^i)}\big[\pi^i\big(x^i,Y^i\big)+\xi^i\big(x^i\big)+\delta E_{p_{S^i}}\big[V^i\big(\mu^i_\infty,S^i\big)\big]\big]\\&\quad>E_{\bar{Q}_{\mu^i_\infty}(\cdot\mid s^i,\tilde{x}^i)}\big[\pi^i\big(\tilde{x}^i,Y^i\big)+\xi^i\big(\tilde{x}^i\big)\\&\qquad+\delta E_{p_{S^i}}\big[V^i\big(B^i\big(\mu^i_\infty,s^i,\tilde{x}^i,Y^i\big),S^i\big)\big]\big]\\&\quad\ge E_{\bar{Q}_{\mu^i_\infty}(\cdot\mid s^i,\tilde{x}^i)}\big[\pi^i\big(\tilde{x}^i,Y^i\big)+\xi^i\big(\tilde{x}^i\big)\big]\\&\qquad+\delta E_{p_{S^i}}\big[V^i\big(E_{\bar{Q}_{\mu^i_\infty}(\cdot\mid s^i,\tilde{x}^i)}\big[B^i\big(\mu^i_\infty,s^i,\tilde{x}^i,Y^i\big)\big],S^i\big)\big]\\&\quad=E_{\bar{Q}_{\mu^i_\infty}(\cdot\mid s^i,\tilde{x}^i)}\big[\pi^i\big(\tilde{x}^i,Y^i\big)+\xi^i\big(\tilde{x}^i\big)\big]+\delta E_{p_{S^i}}\big[V^i\big(\mu^i_\infty,S^i\big)\big]\end{aligned}$$

$\forall\tilde{x}^i\in\mathbb{X}^i$, where the first line follows by equation (28) and definition of Φ^i, the second line follows by the convexity[51] of V^i as a function of μ^i and Jensen's

[51]See, for example, Nyarko (1994), for a proof of convexity of the value function.

inequality, and the last line by the fact that Bayesian beliefs have the martingale property. In turn, the above expression is equivalent to $\psi(\mu^i_\infty, s^i, \xi^i) = \{x^i\}$. *Q.E.D.*

APPENDIX E: POPULATION MODELS

We discuss some variants of population models that differ in the matching technology and feedback. The right variant of population model will depend on the specific application.[52]

SINGLE PAIR MODEL: Each period, a single pair of players is randomly selected from each of the i populations to play the game. At the end of the period, the signals, actions, and outcomes of their own population are revealed to everyone.[53] Steady-state behavior in this case corresponds exactly to the notion of Berk–Nash equilibrium described in the paper.

RANDOM-MATCHING MODEL: Each period, all players are randomly matched and observe only feedback from their own match. We now modify the definition of Berk–Nash equilibrium to account for this random-matching setting. The idea is similar to Fudenberg and Levine's (1993) definition of a heterogeneous self-confirming equilibrium. Now each agent in population i can have different experiences and, hence, have different beliefs and play different strategies in steady state.

For all $i \in I$, define

$$\mathrm{BR}^i(\sigma^{-i}) = \{\sigma^i : \sigma^i \text{ is optimal given } \mu^i \in \Delta(\Theta^i(\sigma^i, \sigma^{-i}))\}.$$

Note that σ is a Berk–Nash equilibrium if and only if $\sigma^i \in \mathrm{BR}^i(\sigma^{-i})$ $\forall i \in I$.

DEFINITION 9: A strategy profile σ is a *heterogeneous Berk–Nash equilibrium* of game $\mathcal{G}$ if, for all $i \in I$, σ^i is in the convex hull of $\mathrm{BR}^i(\sigma^{-i})$.

Intuitively, a heterogeneous equilibrium strategy σ^i is the result of convex combinations of strategies that belong to $\mathrm{BR}^i(\sigma^{-i})$; the idea is that each of these strategies is followed by a segment of the population i.[54]

[52] In some cases, it may be unrealistic to assume that players are able to observe the private signals of previous generations, so some of these models might be better suited to cases with public, but not private, information.

[53] Alternatively, we can think of different incarnations of players born every period who are able to observe the history of previous generations.

[54] Unlike the case of heterogeneous self-confirming equilibrium, a definition where each action in the support of σ is supported by a (possibly different) belief would not be appropriate here. The reason is that $\mathrm{BR}^i(\sigma^{-i})$ might contain only mixed, but not pure, strategies (e.g., Example 1).

RANDOM-MATCHING MODEL WITH POPULATION FEEDBACK: Each period, all players are randomly matched; at the end of the period, each player in population i observes the signals, actions, and outcomes of their own population. Define

$$\bar{\mathrm{BR}}^i(\sigma^i, \sigma^{-i}) = \{\hat{\sigma}^i : \hat{\sigma}^i \text{ is optimal given } \mu^i \in \Delta(\Theta^i(\sigma^i, \sigma^{-i}))\}.$$

DEFINITION 10: A strategy profile σ is a *heterogeneous Berk–Nash equilibrium with population feedback* of game $\mathcal{G}$ if, for all $i \in I$, σ^i is in the convex hull of $\bar{\mathrm{BR}}^i(\sigma^i, \sigma^{-i})$.

The main difference when players receive population feedback is that their beliefs no longer depend on their own strategies but rather on the aggregate population strategies.

E.1. *Equilibrium Foundation*

Using arguments similar to the ones in the text, it is now straightforward to conclude that the definition of heterogeneous Berk–Nash equilibrium captures the steady state of a learning environment with a population of agents in the role of each player. To see the idea, let each population i be composed of a continuum of agents in the unit interval $K \equiv [0, 1]$. A strategy of agent ik (meaning agent $k \in K$ from population i) is denoted by σ^{ik}. The aggregate strategy of population (i.e., player) i is $\sigma^i = \int_K \sigma^{ik}\, dk$.

RANDOM-MATCHING MODEL: Suppose that each agent is optimizing and that, for all i, (σ_t^{ik}) converges to σ^{ik} a.s. in K, so that individual behavior stabilizes.[55] Then Lemma 2 says that the support of beliefs must eventually be $\Theta^i(\sigma^{ik}, \sigma^{-i})$ for agent ik. Next, for each ik, take a convergent subsequence of beliefs μ_t^{ik} and denote it μ_∞^{ik}. It follows that $\mu_\infty^{ik} \in \Delta(\Theta^i(\sigma^{ik}, \sigma^{-i}))$ and, by continuity of behavior in beliefs, σ^{ik} is optimal given μ_∞^{ik}. In particular, $\sigma^{ik} \in \mathrm{BR}^i(\sigma^{-i})$ for all ik and, since $\sigma^i = \int_K \sigma^{ik}\, dk$, it follows that σ^i is in the convex hull of $\mathrm{BR}^i(\sigma^{-i})$.

RANDOM-MATCHING MODEL WITH POPULATION FEEDBACK: Suppose that each agent is optimizing and that, for all i, $\sigma_t^i = \int_K \sigma_t^{ik}\, dk$ converges to σ^i. Then Lemma 2 says that the support of beliefs must eventually be $\Theta^i(\sigma^i, \sigma^{-i})$ for any agent in population i. Next, for each ik, take a convergent subsequence of beliefs μ_t^{ik} and denote it μ_∞^{ik}. It follows that $\mu_\infty^{ik} \in \Delta(\Theta^i(\sigma^i, \sigma^{-i}))$ and, by continuity of behavior in beliefs, σ^{ik} is optimal given μ_∞^{ik}. In particular, $\sigma^{ik} \in$

[55]We need individual behavior to stabilize; it is not enough that it stabilizes in the aggregate. This is natural, for example, if we believe that agents whose behavior is unstable will eventually realize they have a misspecified model.

$\bar{\mathrm{BR}}^i(\sigma^{-i})$ for all i, k and, since $\sigma^i = \int_K \sigma^{ik}\, dk$, it follows that σ^i is in the convex hull of $\bar{\mathrm{BR}}^i(\sigma^{-i})$.

APPENDIX F: LACK OF PAYOFF FEEDBACK

In the paper, players are assumed to observe their own payoffs. We now provide two alternatives to relax this assumption. In the first alternative, players observe no feedback about payoffs; in the second alternative, players may observe partial feedback.

No payoff feedback. In the paper, we had a single, deterministic payoff function $\pi^i : \mathbb{X}^i \times \mathbb{Y}^i \to \mathbb{R}$, which can be represented in vector form as an element $\pi^i \in \mathbb{R}^{\#(\mathbb{X}^i \times \mathbb{Y}^i)}$. We now generalize it to allow for uncertain payoffs. Player i is endowed with a probability distribution $P_{\pi^i} \in \Delta(\mathbb{R}^{\#(\mathbb{X}^i \times \mathbb{Y}^i)})$ over the possible payoff functions. In particular, the random variable π^i is independent of Y^i, and so there is nothing new to learn about payoffs from observing consequences. With random payoff functions, the results extend provided that optimality is defined as follows: A strategy σ^i for player i is *optimal* given $\mu^i \in \Delta(\Theta^i)$ if $\sigma^i(x^i \mid s^i) > 0$ implies that

$$x^i \in \arg\max_{\bar{x}^i \in \mathbb{X}^i} E_{P_{\pi^i}} E_{\bar{Q}^i_{\mu^i}(\cdot \mid s^i, \bar{x}^i)}\big[\pi^i\big(\bar{x}^i, Y^i\big)\big].$$

Note that by interchanging the order of integration, this notion of optimality is equivalent to the notion in the paper where the deterministic payoff function is given by $E_{P_{\pi^i}} \pi^i(\cdot, \cdot)$.

Partial payoff feedback. Suppose that player i knows her own consequence function $f^i : \mathbb{X} \times \Omega \to \mathbb{Y}^i$ and that her payoff function is now given by $\pi^i : \mathbb{X} \times \Omega \to \mathbb{R}$. In particular, player i may not observe her own payoff, but observing a consequence may provide partial information about (x^{-i}, ω) and, therefore, about payoffs. Unlike the case in the text where payoffs are observed, a belief $\mu^i \in \Delta(\Theta^i)$ may not uniquely determine expected payoffs. The reason is that the distribution over consequences implied by μ^i may be consistent with several distributions over $\mathbb{X}^{-i} \times \Omega$; that is, the distribution over $\mathbb{X}^{-i} \times \Omega$ is only partially identified. Define the set $\mathcal{M}_{\mu^i} \subseteq \Delta(\mathbb{X}^{-i} \times \Omega)^{\mathbb{S}^i \times \mathbb{X}^i}$ to be the set of conditional distributions over $\mathbb{X}^{-i} \times \Omega$ given $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$ that are consistent with belief $\mu^i \in \Delta(\Theta^i)$, that is, $m \in \mathcal{M}_{\mu^i}$ if and only if $\bar{Q}^i_{\mu^i}(y^i \mid s^i, x^i) = m(f^i(x^i, X^{-i}, W) = y^i \mid s^i, x^i)$ for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$ and $y^i \in \mathbb{Y}^i$. Then optimality should be defined as follows: A strategy σ^i for player i is *optimal* given $\mu^i \in \Delta(\Theta^i)$ if there exists $m_{\mu^i} \in \mathcal{M}_{\mu^i}$ such that $\sigma^i(x^i \mid s^i) > 0$ implies that

$$x^i \in \arg\max_{\bar{x}^i \in \mathbb{X}^i} E_{m_{\mu^i}(\cdot \mid s^i, \bar{x}^i)}\big[\pi^i\big(\bar{x}^i, X^{-i}, W\big)\big].$$

Finally, the definition of identification would also need to be changed to require not only that there is a unique distribution over consequences that matches the observed data, but also that this unique distribution implies a unique expected utility function.

APPENDIX G: GLOBAL STABILITY: EXAMPLE 2.1 (MONOPOLY WITH UNKNOWN DEMAND)

Theorem 3 says that all Berk–Nash equilibria can be approached with probability 1 provided we allow for vanishing optimization mistakes. In this appendix, we illustrate how to use the techniques of stochastic approximation theory to establish stability of equilibria under the assumption that players make no optimization mistakes. We present the explicit learning dynamics for the monopolist with unknown demand, Example 2.1, and show that the unique equilibrium in this example is globally stable. The intuition behind global stability is that switching from the equilibrium strategy to a strategy that puts more weight on a price of 2 changes beliefs in a way that makes the monopoly want to put less weight on a price of 2, and similarly for a deviation to a price of 10.

We first construct a perturbed version of the game. Then we show that the learning problem is characterized by a nonlinear stochastic system of difference equations and employ stochastic approximation methods for studying the asymptotic behavior of such system. Finally, we take the payoff perturbations to zero.

In order to simplify the exposition and thus better illustrate the mechanism driving the dynamics, we modify the subjective model slightly. We assume the monopolist only learns about the parameter $b \in \mathbb{R}$; that is, her beliefs about parameter a are degenerate at a point $a = 40 \neq a^0$ and thus are never updated. Therefore, beliefs μ are probability distributions over $\mathbb{R}$, that is, $\mu \in \Delta(\mathbb{R})$.

PERTURBED GAME: Let ξ be a real-valued random variable distributed according to P_ξ; we use F to denote the associated c.d.f. and f the p.d.f. The perturbed payoffs are given by $yx - \xi 1\{x = 10\}$. Thus, given beliefs $\mu \in \Delta(\mathbb{R})$, the probability of optimally playing $x = 10$ is

$$\sigma(\mu) = F\big(8a - 96E_\mu[B]\big).$$

Note that the only aspect of μ that matters for the decision of the monopolist is $E_\mu[B]$. Thus, letting $m = E_\mu[B]$ and slightly abusing notation, we use $\sigma(\mu) = \sigma(m)$ as the optimal strategy.

BAYESIAN UPDATING: We now derive the Bayesian updating procedure. We assume that the prior μ_0 is given by a Gaussian distribution with mean and variance m_0, τ_0^2.[56] It is possible to show that, given a realization (y, x) and a prior

[56]This choice of prior is standard in Gaussian settings like ours. As shown below, this choice simplifies the exposition considerably.

$N(m,\tau^2)$, the posterior is also Gaussian and the mean and variance evolve as follows: $m_{t+1}=m_t+(\frac{-(Y_{t+1}-a)}{X_{t+1}}-m_t)(\frac{X_{t+1}^2}{X_{t+1}^2+\tau_t^{-2}})$ and $\tau_{t+1}^2=\frac{1}{(X_{t+1}^2+\tau_t^{-2})}$.

NONLINEAR STOCHASTIC DIFFERENCE EQUATIONS AND STOCHASTIC APPROXIMATION: For simplicity, let $r_{t+1}\equiv\frac{1}{t+1}(\tau_t^{-2}+X_{t+1}^2)$ and note that the previous nonlinear system of stochastic difference equations can be written as

$$\begin{aligned}
m_{t+1}&=m_t+\frac{1}{t+1}\frac{X_{t+1}^2}{r_{t+1}}\left(\frac{-(Y_{t+1}-a)}{X_{t+1}}-m_t\right),\\
r_{t+1}&=r_t+\frac{1}{t+1}\big(X_{t+1}^2-r_t\big).
\end{aligned}$$

Let $\beta_t=(m_t,r_t)'$, $Z_t=(X_t,Y_t)$,

$$G(\beta_t,z_{t+1})=\begin{bmatrix}\frac{x_{t+1}^2}{r_{t+1}}\left(\frac{-(y_{t+1}-a)}{x_{t+1}}-m_t\right)\\ \big(x_{t+1}^2-r_t\big)\end{bmatrix},$$

and

$$\begin{aligned}
\mathbb{G}(\beta)=\begin{bmatrix}\mathbb{G}_1(\beta)\\ \mathbb{G}_2(\beta)\end{bmatrix}&=E_{P_\sigma}\big[G(\beta,Z_{t+1})\big]\\
&=\begin{bmatrix}D\\ \big(4+F(8a-96m)96-r\big)\end{bmatrix},
\end{aligned}$$

where

$$\begin{aligned}
D&=F(8a-96m)\frac{100}{r}\left(\frac{-(a_0-a-b_010)}{10}-m\right)\\
&\quad+\big(1-F(8a-96m)\big)\frac{4}{r}\left(\frac{-(a_0-a-b_02)}{2}-m\right),
\end{aligned}$$

and P_σ is the probability over Z induced by σ (and $y=a^0-b^0x+\omega$). Therefore, the dynamical system can be cast as

$$\beta_{t+1}=\beta_t+\frac{1}{t+1}\mathbb{G}(\beta_t)+\frac{1}{t+1}V_{t+1}$$

with $V_{t+1}=G(\beta_t,Z_{t+1})-\mathbb{G}(\beta_t)$. Stochastic approximation theory (e.g., Kushner and Yin (2003)) implies, roughly speaking, that in order to study the asymptotic behavior of $(\beta_t)_t$, it is enough to study the behavior of the orbits of the following ODE:

$$\dot{\beta}(t)=\mathbb{G}\big(\beta(t)\big).$$

CHARACTERIZATION OF THE STEADY STATES: In order to find the steady states of $(\beta_t)_t$, it is enough to find β^* such that $\mathbb{G}(\beta^*) = 0$. Let $H_1(m) \equiv F(8a - 96m)10(-(a_0 - a) + (b_0 - m)10) + (1 - F(8a - 96m))2(-(a_0 - a) + (b_0 - m)2)$. Observe that $\mathbb{G}_1(\beta) = r^{-1}H_1(m)$ and that H_1 is continuous and $\lim_{m\to-\infty} H_1(m) = \infty$ and $\lim_{m\to\infty} H_1(m) = -\infty$. Thus, there exists at least one solution $H_1(m) = 0$. Therefore, there exists at least one β^* such that $\mathbb{G}(\beta^*) = 0$.

Let $\bar{b} = b_0 - \frac{a_0 - a}{10} = 4 - \frac{1}{5} = \frac{19}{5}$ and $\underline{b} = b_0 - \frac{a_0 - a}{2} = 4 - \frac{42-40}{2} = 3$, $\bar{r} = 4 + F(8a - 96\underline{b})96$ and $\underline{r} = 4 + F(8a - 96\bar{b})96$, and $\mathbb{B} \equiv [\underline{b}, \bar{b}] \times [\underline{r}, \bar{r}]$. It follows that $H_1(m) < 0\ \forall m > \bar{b}$, and thus m^* must be such that $m^* \le \bar{b}$. It is also easy to see that $m^* \ge \underline{b}$. Moreover, $\frac{dH_1(m)}{dm} = 96f(8a - 96m)(8(a_0 - a) - (b_0 - m)96) - 4 - 96F(8a - 96m)$. Thus, for any $m \le \bar{b}$, $\frac{dH_1(m)}{dm} < 0$, because $m \le \bar{b}$ implies $8(a_0 - a) \le (b_0 - m)80 < (b_0 - m)96$.

Therefore, on the relevant domain $m \in [\underline{b}, \bar{b}]$, H_1 is decreasing, thus implying that there exists only one m^* such that $H_1(m^*) = 0$. Therefore, there exists only one β^* such that $\mathbb{G}(\beta^*) = 0$.

We are now interested in characterizing the limit of β^* as the perturbation vanishes, that is, as F converges to $1\{\xi \ge 0\}$. To do this, we introduce some notation. We consider a sequence $(F_n)_n$ that converges to $1\{\xi \ge 0\}$ and use β_n^* to denote the steady state associated to F_n. Finally, we use H_1^n to denote the H_1 associated to F_n.

We proceed as follows. First note that since $\beta_n^* \in \mathbb{B}\ \forall n$, the limit exists (going to a subsequence if needed). We show that $m^* \equiv \lim_{n\to\infty} m_n^* = \frac{8a}{96} = 8\frac{40}{96} = \frac{10}{3}$. Suppose not; in particular, suppose that $\lim_{n\to\infty} m_n^* < \frac{8a}{96} = \frac{10}{3}$ (the argument for the reverse inequality is analogous and thus omitted). In this case, $\lim_{n\to\infty} 8a - 96m_n^* > 0$, and thus $\lim_{n\to\infty} F_n(8a - 96m_n^*) = 1$. Therefore

$$\lim_{n\to\infty} H_1^n(\beta_n^*) = 10\big(-(a_0 - a) + (b_0 - m^*)10\big) \ge 10\left(-2 + \frac{2}{3}10\right) > 0.$$

But this implies that $\exists N$ such that $H_1^n(\beta_n^*) > 0\ \forall n \ge N$, which is a contradiction.

Moreover, define $\sigma_n^* = F_n(8a - 96m_n^*)$ and $\sigma^* = \lim_{n\to\infty} \sigma_n$. Since $H_1^n(m_n^*) = 0\ \forall n$ and $m^* = \frac{10}{3}$, it follows that

$$\sigma^* = \frac{-2\left(-2 + \left(4 - \frac{10}{3}\right)2\right)}{10\left(-2 + \left(4 - \frac{10}{3}\right)10\right) - 2\left(-2 + \left(4 - \frac{10}{3}\right)2\right)} = \frac{1}{36}.$$

GLOBAL CONVERGENCE TO THE STEADY STATE: In our example, it is in fact possible to establish that behavior converges with probability 1 to the unique equilibrium. By the results in Benaim (1999, Section 6.3), it is sufficient to establish the *global* asymptotic stability of β_n^* for any n, that is, the basin of attraction of β_n^* is all of $\mathbb{B}$.

In order to do this, let $L(\beta)=(\beta-\beta_n^*)'P(\beta-\beta_n^*)$ for all β; where $P\in\mathbb{R}^{2\times 2}$ is positive definite and *diagonal* and will be determined later. Note that $L(\beta)=0$ iff $\beta=\beta_n^*$. Also

$$\begin{aligned}\frac{dL(\beta(t))}{dt}&=\nabla L(\beta(t))'\mathbb{G}(\beta(t))\\&=2(\beta(t)-\beta_n^*)'P(\mathbb{G}(\beta(t)))\\&=2\{(m(t)-m_n^*)P_{[11]}\mathbb{G}_1(\beta(t))+(r(t)-r_n^*)P_{[22]}\mathbb{G}_2(\beta(t))\}.\end{aligned}$$

Since $\mathbb{G}(\beta_n^*)=0$,

$$\begin{aligned}\frac{dL(\beta(t))}{dt}&=2(\beta(t)-\beta_n^*)'P(\mathbb{G}(\beta(t))-\mathbb{G}(\beta_n^*))\\&=2(m(t)-m_n^*)P_{[11]}(\mathbb{G}_1(\beta(t))-\mathbb{G}_1(\beta_n^*))\\&\quad+2(r(t)-r_n^*)P_{[22]}(\mathbb{G}_2(\beta(t))-\mathbb{G}_2(\beta_n^*))\\&=2(m(t)-m_n^*)^2P_{[11]}\int_0^1\frac{\partial\mathbb{G}_1(m_n^*+s(m(t)-m_n^*),r_n^*)}{\partial m}\,ds\\&\quad+2(r(t)-r_n^*)^2P_{[22]}\int_0^1\frac{\partial\mathbb{G}_2(m_n^*,r_n^*+s(r(t)-r_n^*))}{\partial r}\,ds,\end{aligned}$$

where the last equality holds by the mean value theorem. Note that $d\mathbb{G}_2(m_n^*, r_n^*+s(r(t)-r_n^*))/dr=-1$ and $\int_0^1(d\mathbb{G}_1(m_n^*+s(m(t)-m_n^*),r_n^*)/dm)\,ds=\int_0^1(r_n^*)^{-1}(dH_1(m_n^*+s(m(t)-m_n^*))/dm)\,ds$. Since $r(t)>0$ and $r_n^*\geq 0$, the first term is positive, and we already established that $\frac{dH_1(m)}{dm}<0\ \forall m$ in the relevant domain. Thus, by choosing $P_{[11]}>0$ and $P_{[22]}>0$, it follows that $\frac{dL(\beta(t))}{dt}<0$.

Therefore, we show that L satisfies the following properties: is strictly positive $\forall\beta\neq\beta_n^*$ and $L(\beta_n^*)=0$, and $\frac{dL(\beta(t))}{dt}<0$. Thus, the function satisfies all the conditions of a Lyapunov function and, therefore, β_n^* is globally asymptotically stable $\forall n$ (see Hirsch, Smale, and Devaney (2004, p. 194)).

REFERENCES

BENAIM, M. (1999): "Dynamics of Stochastic Approximation Algorithms," in *Seminaire de Probabilites XXXIII*. Lecture Notes in Mathematics, Vol. 1709. Berlin: Springer, 1–68. [15]

BILLINGSLEY, P. (1995): *Probability and Measure*. New York: Wiley. [5]

DORASZELSKI, U., AND J. F. ESCOBAR (2010): "A Theory of Regular Markov Perfect Equilibria in Dynamic Stochastic Games: Genericity, Stability, and Purification," *Theoretical Economics*, 5 (3), 369–402. [8]

FUDENBERG, D., AND D. K. LEVINE (1993): "Self-Confirming Equilibrium," *Econometrica*, 523–545. [10]

HIRSCH, M. W., S. SMALE, AND R. L. DEVANEY (2004): *Differential Equations, Dynamical Systems and an Introduction to Chaos*. Amsterdam: Elsevier Academic Press. [16]

KUSHNER, H. J., AND G. G. YIN (2003): *Stochastic Approximation and Recursive Algorithms and Applications*. Berlin: Springer-Verlag. [14]

NYARKO, Y. (1994): "On the Convexity of the Value Function in Bayesian Optimal Control Problems," *Economic Theory*, 4 (2), 303–309. [9]

Olin Business School, Washington University in St. Louis, 1 Brookings Drive, Campus Box 1133, St. Louis, MO 63130, U.S.A.; iesponda@wustl.edu

and

Dept. of Economics, UC Berkeley, 530-1 Evans Hall #3880, Berkeley, CA 94720, U.S.A.; dpouzo@econ.berkeley.edu.

Co-editor Philippe Jehiel handled this manuscript.

Manuscript received June, 2014; final revision received February, 2016.