Econometrica Supplementary Material

SUPPLEMENT TO "STATIONARY SOCIAL LEARNING IN A CHANGING ENVIRONMENT"

(*Econometrica*, Vol. 92, No. 6, November 2024, 1939–1966)

RAPHAËL LEVY
Department of Economics and Decision Sciences, HEC Paris

MARCIN PĘSKI
Department of Economics, University of Toronto

NICOLAS VIEILLE
Department of Economics and Decision Sciences, HEC Paris

THIS SUPPLEMENTAL APPENDIX contains the analysis of some extensions of our baseline setup, as well as numerical evidence in the case $n = 3$. Section E discusses the extension of Banerjee and Fudenberg (2004)'s main result to the case of costly signals. In Section F, we provide numerical evidence on equilibrium behavior in the case of samples of size 3. Section G refines the proof of Theorem 3 when $n = 3$. It shows that the conclusion of Theorem 3 holds as soon as $n^2\pi(1-\pi) < 1$ (and c is small). In Section H, we investigate the case where agents can choose any statistical experiment. Section I extends Theorem 2 to a fully asymmetric setup. Section J discusses the case where samples are not drawn from the previous period, but from the further past (in the case $n = 2$).

APPENDIX E: BANERJEE–FUDENBERG WITH COSTLY SIGNALS

We prove here that the main insights of Banerjee and Fudenberg (2004) extend to the case of costly signals.

Consider the following sequential learning setup. The state $\theta \in \Theta$ is drawn uniformly at random at time 0, and kept fixed. In each period $t \geq 1$, each agent in a new continuum of agents sample n agents from the previous period, then decide whether to acquire information. Information has cost $c > 0$, and yields a private signal with cdf H_θ. Agents are aware of calendar time, with utility function $u(a, \theta) = 1_{a=\theta}$. Signals satisfy the usual symmetry assumption.

We denote by z_t the equilibrium fraction of agents in period t playing the correct action.[1]

PROPOSITION E.1: *Assume that* $n \geq 2$ *and that Assumption* 1 ($c < v(\frac{1}{2}) - u(\frac{1}{2})$) *holds. Then* $\lim_{t\to+\infty} z_t = 1$.

PROOF: The first period $t = 1$ is specific. Agents do not sample, acquire information, hold an interim belief equal to the prior belief on θ, hence $z_1 = v(\frac{1}{2})$.

Raphaël Levy: levyr@hec.fr
Marcin Pęski: marcin.peski@utoronto.ca
Nicolas Vieille: vieille@hec.fr

[1]Because of symmetry, it is the same in both states, so that the fraction x_t of agents playing action 1 is $x_t = z_t$ if $\theta = 1$, and $x_t = 1 - z_t$ if $\theta = 0$.

 https://doi.org/10.3982/ECTA20475

Let $t \geq 1$ be arbitrary, such that $z_t \geq v(\frac{1}{2})$. At equilibrium, agents acquire information with a balanced sample $k = n/2$ (if n is even), and possibly for other values of k as well. Therefore, the equilibrium payoff of a $t+1$-agent is *at least* z_{t+1} minus the probability of a balanced sample times (multiplied by c). Consider a $t+1$-agent who switches unilaterally to the strategy that acquires information when $k = n/2$, and plays the majority action for other values of k. Denote by y_{t+1} the probability that this agent's action is correct. Since the equilibrium strategy buys "more" often information than the alternative strategy, it must also be more often "correct": one has $z_{t+1} \geq y_{t+1}$.

On the other hand,

$$y_{t+1} = \sum_{k>n/2} \binom{n}{k} z_t^k (1-z_t)^{n-k} + \mathbf{1}_{n \text{ is even}} \times v\left(\frac{1}{2}\right) \binom{n}{\frac{n}{2}} z_t^{n/2} (1-z_t)^{n/2}.$$

Using the identity $z = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} z^k (1-z)^{n-k}$, it follows that

$$\begin{aligned}
z_{t+1} - z_t \geq{} & \sum_{k>n/2} \binom{n}{k} z_t^k (1-z_t)^{n-k} + \mathbf{1}_{n \text{ is even}} \times v\left(\frac{1}{2}\right) \binom{n}{\frac{n}{2}} z_t^{n/2} (1-z_t)^{n/2} \\
& - \sum_{k} \frac{k}{n} \binom{n}{k} z_t^k (1-z_t)^{n-k} \\
={} & \sum_{k>n/2} \binom{n}{k} \left(\frac{n-k}{n}\right) z_t^k (1-z_t)^{n-k} + \mathbf{1}_{n \text{ is even}} \times \left(v\left(\frac{1}{2}\right) - \frac{1}{2}\right) \binom{n}{\frac{n}{2}} z_t^{n/2} (1-z_t)^{n/2} \\
& - \sum_{k<\frac{n}{2}} \frac{k}{n} \binom{n}{k} z_t^k (1-z_t)^{n-k} \\
\geq{} & \sum_{k>n/2} \binom{n}{k} \left(\frac{n-k}{n}\right) \left[z_t^k (1-z_t)^{n-k} - z_t^{n-k} (1-z_t)^k\right] \\
={} & \sum_{k>n/2} \binom{n}{k} \left(\frac{n-k}{n}\right) \left(z_t(1-z_t)\right)^{n-k} \left[z_t^{2k-n} - (1-z_t)^{2k-n}\right] \\
\geq{} & \left(z_t(1-z_t)\right)^n \left[v\left(\frac{1}{2}\right)^n - \left(1 - v\left(\frac{1}{2}\right)\right)^n\right] > 0,
\end{aligned}$$

where the last inequality holds since $z_t \geq v(\frac{1}{2})$.

This implies that (z_t) is increasing, hence convergent, with a limit z that satisfies $(z(1-z))^n = 0$. Since $z > \frac{1}{2}$, one has $z = 1$, as desired. *Q.E.D.*

APPENDIX F: NUMERICAL EVIDENCE IN THE CASE $n = 3$

This section contains numerically based characterizations of equilibrium strategies for samples of size $n = 3$. We start with an informal summary. When $n = 3$, strategies can be summarized by three variables:

- the probability $\beta(0)$ of acquiring information with a unanimous sample,
- $\beta(1)$ of acquiring information with a balanced sample, and

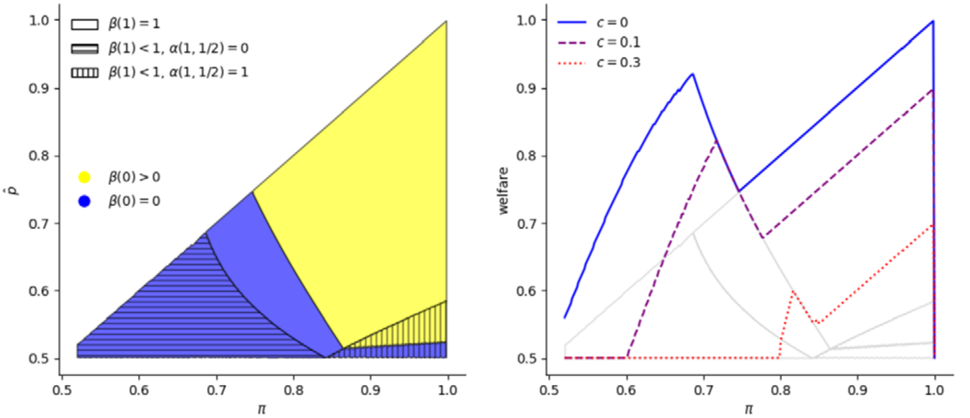

FIGURE F.1.—Equilibrium behavior (left) and welfare (right) for $n = 3$.

- the action $\alpha(1, \frac{1}{2})$ played when information is not acquired at balanced samples.[2]

For each possible profile of strategies, we adapt the proof techniques of Theorem 3 to derive estimates of the limit interim beliefs, and thus check for what values $(\pi, \hat{p})$ such a profile of strategies is an equilibrium.

The results are summarized in Figure F.1. The parameter region identified in Theorem 3 ($\beta(0) > 0$ and $\beta(1) = 1$) corresponds to the plain light (yellow) area in the upper-right corner. In addition, we notice the existence of an area (with vertical stripes in the lower- right corner) corresponding to high π and c where agents adopt a *contrarian* behavior when facing mixed evidence and not acquiring information($\alpha(1, \frac{1}{2}) = 1$). By fostering diversity, such contrarian behavior slows down convergence and helps the population respond to change.

The right panel of Figure F.1 depicts the highest equilibrium welfare as a function of π, for various values of c. We observe, for example, from the solid curve ($c = 0$), that the equilibrium welfare may exceed $\hat{p}$ for signals with low enough precision (corresponding to the dark/blue area in the left panel), but that learning is always incomplete, in accordance with our main message. We also note that for stronger signals, the equilibrium welfare is given by $\hat{p}$, irrespective of c.

Interestingly, we observe that welfare is both decreasing in π and independent of c as long as $(\pi, \hat{p})$ belongs to the plain dark (blue) region. There, agents never acquire information at unanimous samples, and always acquire otherwise. Therefore, equilibrium beliefs do not vary with c. When π increases, information acquisition is thus unchanged, but the correlation of actions within samples increases, lowering informativeness, and then decreasing welfare. This suggests that more precise signals hurt as long as they do not encourage more information acquisition. This is an instance of the *principle of countervailing adjustment* (Bikhchandani, Hirshleifer, Tamuz, and Welch (2024)): a favorable shift in information availability does not necessarily improve decision-making or welfare.[3]

[2] It is clear that one follows one's signal upon acquiring information, and one follows the crowd when drawing a unanimous sample and not acquiring information, since the consensus is positively correlated with the state.

[3] Relatedly, Dasaratha, Golub, and Hak (2023) underline the importance of having agents with sufficiently diverse signal distributions for information aggregation. While diversity of signals comes from different ex ante signal distributions in their case, in our case it arises when agents are more likely to have *ex post* different signal realizations, that is, less precise signals.

The reminder of the section contains a detailed justification of the above claims. Section F.1 contains the statements of formal results that we rely on in our numerical calculations. Section F.2 contains the proofs. Section F.3 discusses numerical computations and, in particular, explains how we drew Figure F.1.

F.1. *Equilibrium Characterization*

Let us first describe the additional notation we use, together with some minor changes relative to the paper. Notation that is not defined here is defined in the paper.

Given that signals are binary, players follow the signal whenever they acquire information, in any equilibrium. It is worthwhile to redefine a strategy $\sigma = (\alpha, \beta)$ so that $\alpha(k)$ is the probability of playing action 1 if no information is acquired at sample k. A symmetric strategy is defined by the four numbers $(\beta(0), \beta(1), \alpha(0), \alpha(1))$. The equilibrium conditions and $\hat{p} > \frac{1}{2}$ imply that $\alpha(k) \in \{0, 1\}$. Additionally, the consensus theorem (Theorem 2) implies that, for sufficiently small λ, $\alpha(0) = 0$.

We will use a different parametrization of strategies. We will encode an agent's behavior at sample $k = 0$ by a single parameter, $\beta \in [0, 1]$, and one's behavior at sample $k = 1$ by a another parameter, $\alpha \in [-1, 1]$, as follows:

- an agent sampling $k = 0$ or $k = n$ follows the sample with probability $1 - \beta$ and acquires information with probability β (hence $\beta = \beta(0)$),
- an agent sampling $k = 1$ or $k = n - 1$ acquires information with probability $1 - |\alpha|$. If the agent does not, it follows the majority action if $\alpha > 0$, and the minority action if $\alpha < 0$. That is, $|\alpha| = 1 - \beta(1)$ encodes the probability of acquiring information, and and the sign of α encodes one's behavior when not acquiring information.

With $\pi_\theta = \pi\theta + (1 - \pi)(1 - \theta)$ for $\theta \in \Theta$, the transitions induced by a strategy $(\alpha, \beta) \in [-1, 1] \times [0, 1]$ are given by

$$g_\theta(x|\alpha, \beta, \pi) = x^3(1 - \beta + \beta\pi_\theta) + 3x^2(1 - x)\big((1 - |\alpha|)\pi_\theta + \max(0, \alpha)\big)$$
$$+ 3x(1 - x)^2\big((1 - |\alpha|)\pi_\theta + \max(0, -\alpha)\big) + (1 - x)^3\beta\pi_\theta.$$

We denote by $\phi_\theta(\alpha, \beta, \pi) = g'_\theta(0|\alpha, \beta, \pi)$ the derivative at 0 of g_θ:

$$\phi_\theta(\alpha, \beta, \pi) = 3\big(\max(0, -\alpha) + \big(1 - |\alpha| - \beta\big)\pi_\theta\big).$$

Let $p_k^\lambda(\alpha, \beta)$ be the interim beliefs at a stationary distribution for (α, β).

REMARK F.1: To be precise, $p_k^\lambda(\alpha, \beta)$ are the beliefs deduced from some invariant distribution (that puts no mass on 0 and 1) given (α, β). In the following, statements about limits $\lim_\lambda p_k^\lambda(\alpha, \beta)$ should be interpreted as statements about limits of *all* possible sequences $p_k^\lambda(\alpha, \beta)$ consistent with (α, β).

Given primitives $(\lambda, \pi, \hat{p})$, we view an equilibrium as a 4-tuple $(\alpha, \beta, p_0, p_1)$ where (α, β) is optimal given the beliefs p_k for $k = 0, 1$, and p_0, p_1 are deduced from some invariant distribution for (α, β) (with no atom at 0 and 1). Let $\Gamma^\lambda(\pi, \hat{p})$ denote the set of all equilibrium tuples. From the proofs in Section B, $\Gamma^\lambda(\pi, \hat{p})$ is closed and nonempty. We set $\Gamma(\pi, \hat{p}) = \limsup_{\lambda\to 0} \Gamma^\lambda(\pi, \hat{p})$.

We know from the main paper that the equilibrium value of β vanishes as $\lambda \to 0$, hence the case $\beta = 0$ will have specific importance. In what follows, when we drop β from the notation, it means that $\beta = 0$.

The first result provides necessary conditions on elements of $\Gamma(\pi, \hat{p})$. For this purpose, we set

$$p_0^*(\alpha, \pi) := \frac{-\log\phi_0(\alpha, \pi)}{\log\phi_1(\alpha, \pi) - \log\phi_0(\alpha, \pi)},$$

and

$$p_1^{\max}(\alpha, \pi) = \max_x \frac{\sum_{i\in\mathbf{Z}} \psi_1\big(g_1^i(x|\alpha, \pi)\big)}{\sum_{i\in\mathbf{Z}} \psi_1\big(g_1^i(x|\alpha, \pi)\big) + \sum_{i\in\mathbf{Z}} \psi_2\big(g_1^i(x|\alpha, \pi)\big)},$$

$$p_1^{\min}(\alpha, \pi) = \min_x \frac{\sum_{i\in\mathbf{Z}} \psi_1\big(g_1^i(x|\alpha, \pi)\big)}{\sum_{i\in\mathbf{Z}} \psi_1\big(g_1^i(x|\alpha, \pi)\big) + \sum_{i\in\mathbf{Z}} \psi_2\big(g_1^i(x|\alpha, \pi)\big)}.$$

THEOREM F.1: *For each* $(\alpha, \beta, p_0, p_1) \in \Gamma(\pi, \hat{p})$, *we have* $\beta = 0$, *and one of* ***P1*** *or* ***P2*** *holds.*

P1 $p_1^{\max}(\alpha, \pi)$ *and* $p_1^{\min}(\alpha, \pi)$ *are well-defined (i.e., the sums* $\sum_{i\in\mathbf{Z}} \psi_k(g_1^i(x|\alpha, \pi))$ *are convergent and bounded),* $p_0 = \min(p_0^*(\alpha, \pi), 1 - \hat{p})$, *and either:*
1. $\alpha \in (-1, 0)$ and $\hat{p} \in (p_1^{\min}(\alpha; \pi), p_1^{\max}(\alpha; \pi))$,
2. $\alpha = 0$ and $1 - \hat{p} \le p_1^{\max}(0; \pi)$ *and* $p_1^{\min}(0, \pi) \le \hat{p}$,
3. $\alpha \in (0, 1)$ *and* $\hat{p} \in (1 - p_1^{\max}(\alpha; \pi), 1 - p_1^{\min}(\alpha; \pi))$, *or*

P2 $\alpha = 1 - \frac{1}{3\pi}$ *and* $p_0 = p_1 = 1 - \hat{p}$.

We are not able to disprove the existence of the second type of equilibrium ($\alpha = 1 - \frac{1}{3\pi}$). However, any such (limit) equilibrium has welfare equal to $\hat{p}$, since $p_0 = 1 - \hat{p}$.

We will focus on the first type of equilibrium. For each π, we let

$$\Gamma^*(\pi, \hat{p}) = \left\{ p_0 : \text{there exists } \alpha < 1 - \frac{1}{3\pi}, \beta, p_1 \text{ st. } (\alpha, \beta, p_0, p_1) \in \Gamma(\pi, \hat{p}) \right\}.$$

We have the following corollary to Theorem F.1.

COROLLARY F.1: *Let* $\pi > \frac{1}{2}$, $\hat{p} \in (\frac{1}{2}, \pi)$, *and* $p_0 \in \Gamma^*(\pi, \hat{p})$ be given. Then $p_0 = \min(p_0^*(\alpha, \pi), 1 - \hat{p})$ for some α such that either:
1. $\alpha < 0$ and $\hat{p} \in (p_1^{\min}(\alpha; \pi), p_1^{\max}(\alpha; \pi))$, or
2. $\alpha = 0$ *and* $\hat{p} \ge \max(p_1^{\min}(0; \pi), 1 - p_1^{\max}(0; \pi))$, *or*
3. $\alpha > 0$ *and* $\hat{p} \in (1 - p_1^{\max}(\alpha, \pi), 1 - p_1^{\min}(\alpha, \pi))$.

The second result is about sufficient conditions on elements of $\Gamma^*(\pi, \hat{p})$.

THEOREM F.2: *Let* $\pi > \frac{1}{2}$ *and* $\hat{p} \in (\frac{1}{2}, \pi)$ *be given. Let* $-1 \le \alpha_0 \le \alpha_1 \le 1$ *be s.t. that one of* ***C1***, ***C2***, *or* ***C3*** *below holds. Then there is* $\alpha \in [\alpha_0, \alpha_1]$ *such that* $\min(p_0^*(\alpha, \pi), 1 - \hat{p}) \in \Gamma^*(\pi, \hat{p})$.

C1 : $\alpha_0, \alpha_1 < 0$ *and* $\hat{p} \in (\min_i p_1^{\max}(\alpha_i, \pi), \max p_1^{\min}(\alpha_i, \pi))$,
C2 : $\alpha_0 = \alpha_1 = 0$ *and* $\hat{p} > \max(p_1^{\max}(0, \pi), 1 - p_1^{\min}(0, \pi))$, *or*
C3 : $0 < \alpha_0, \alpha_1 < 1 - \frac{1}{3\pi}$ *and* $\hat{p} \in (1 - \max_i p_1^{\min}(\alpha, \pi), 1 - \min_i p_1^{\max}(\alpha_i, \pi))$.

F.2. *Proofs*

F.2.1. *Preliminary Remarks*

The consensus theorem (Theorem 2 from the paper) implies that, as $\lambda \to 0$, the distribution of population states converges to the uniform distribution on $\{0, 1\}$. We list a few implications for equilibrium strategies $(\alpha^\lambda, \beta^\lambda)$:

- $\lim_{\lambda\to 0} F_0^\lambda(x) + F_1^\lambda(x) = \frac{1}{2}$ for each $x \in (0,1)$, and $\lim_{\lambda\to 0} p_0^\lambda(\alpha^\lambda, \beta^\lambda; \pi) = \lim_{\lambda\to 0} \frac{F_1^\lambda(x)}{F_0^\lambda(x)+F_1^\lambda(x)}$,
- $\limsup \phi_0(\alpha^\lambda, \pi) \leq 1 \leq \liminf \phi_1(\alpha, \pi)$.

Below, we collectively refer to these implications as "the consensus theorem." Because of the existence results, the proofs of Theorems F.1 and F.2 consider only strategies that satisfy the consensus theorem.

F.2.2. *Beliefs at Sample $k = 0$*

LEMMA F.1: *Let $(\alpha^\lambda, \beta^\lambda)_\lambda$ be any strategy indexed by λ. Suppose that $\beta^\lambda \to 0$ and $\liminf_{\lambda\to 0} \phi_1(\alpha^\lambda, \pi) > 1$. Then*

$$\limsup_{\lambda\to 0} p_0^\lambda(\alpha^\lambda, \beta^\lambda, \pi) \leq \limsup p_0^*(\alpha^\lambda, \pi).$$

PROOF: Fix α as a cluster point of sequence α^λ. For the most part of the proof, we fix λ and consider transitions g_θ and distributions F_θ associated with $(\alpha^\lambda, \beta^\lambda)$. In the last line, we take limits over λ.

First, we show that for each $\varepsilon > 0$, there exists $\delta > 0$ and m, n such that $\frac{n}{m} \leq (1 + \varepsilon)\frac{p_0^*(\alpha,\pi)}{1-p_0^*(\alpha,\pi)}$ and $h_1^n(x) \leq g_0^m(x)$ for each $x \in (0, \delta)$ (where h_1 is the generalized inverse of g_1, see the Appendix to the paper). For some $\varepsilon' > 0$ to be determined later, let $\phi_\theta = (1 - \varepsilon')\phi_\theta(\alpha, \pi)$. Find δ such that for each $x \in (0, \delta)$,

$$g_\theta(x) \geq \beta\pi_\theta + \phi_\theta x \geq \phi_\theta x.$$

Then, for each m, n such that $g_1^n(g_0^m(x)) \in (0, \delta)$, we have

$$g_1^n\big(g_0^m(x)\big) \geq \phi_1^n \phi_0^m x.$$

Choose n so that

$$n - 1 \leq m\frac{2\varepsilon' - \log\phi_0(\alpha, \pi)}{\log\phi_1(\alpha, \pi) - 2\varepsilon'} \leq n.$$

Then $\phi_1^n\phi_0^m \geq 1$, which implies that $h_1^n(x) \leq g_0^m(x)$. Moreover, for ε' sufficiently small and m sufficiently large,

$$\frac{n}{m} \leq \frac{2\varepsilon' - \log\phi_0(\alpha, \pi)}{\log\phi_1(\alpha, \pi) - 2\varepsilon'} + \frac{1}{m} \leq (1 + \varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)}.$$

Next, fix $x \in (0, \delta)$. By recursively applying stationary distribution equations, we obtain

$$\begin{aligned} F_0(x) - F_0\big(h_1^n(x)\big) &\geq F_0(x) - F_0\big(g_0^m(x)\big) \\ &= \sum_{i=0}^{m-1}\big(F_0\big(g_0^i(x)\big) - F_0\big(g_0^{i+1}(x)\big)\big) \end{aligned}$$

$$
\begin{aligned}
&= \lambda \sum_{i=0}^{m-1} \bigl(F_0\bigl(g_0^i(x)\bigr) - F_1\bigl(h_1\bigl(g_0^{i+1}(x)\bigr)\bigr)\bigr) \\
&\geq \lambda m \bigl(F_0\bigl(g_0^{m-1}(x)\bigr) - F_1\bigl(h_1\bigl(g_0(x)\bigr)\bigr)\bigr).
\end{aligned}
$$

Also,

$$
\begin{aligned}
F_1(x) - F_1\bigl(h_1^n(x)\bigr) &= \sum_{i=0}^{n-1} \bigl(F_1\bigl(h_1^i(x)\bigr) - F_1\bigl(h_1^{i+1}(x)\bigr)\bigr) \\
&= \lambda \sum_{i=0}^{n-1} \bigl(F_0\bigl(h_0\bigl(h_1^i(x)\bigr)\bigr) - F_1\bigl(h_1^{i+1}(x)\bigr)\bigr) \\
&\leq \lambda n \bigl(F_0\bigl(h_0(x)\bigr) - F_1\bigl(h_1^n(x)\bigr)\bigr).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&F_1(x) - F_1\bigl(h_1^n(x)\bigr) - (1+\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)}\bigl(F_0(x) - F_0\bigl(h_1^n(x)\bigr)\bigr) \\
&\quad \leq F_1(x) - F_1\bigl(h_1^n(x)\bigr) - \frac{n}{m}\bigl(F_0(x) - F_0\bigl(h_1^n(x)\bigr)\bigr) \\
&\quad \leq \lambda n \bigl[F_0\bigl(h_0(x)\bigr) - F_0\bigl(g_0^{m-1}(x)\bigr) + F_1(x) - F_1\bigl(h_1^n(x)\bigr)\bigr].
\end{aligned}
$$

Define a sequence $x_0 = x$ and $x_k = h_1^n(x_{k-1})$. Adding up the above formula for $x = x_k$ across all $k \geq 0$, using the fact that $F_\theta(0) = 0$ and that $h_0(x_k) \leq g_0^{m-1}(x_{k-1})$, we get

$$
F_1(x) - (1+\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)} F_0(x) \leq \lambda n \bigl(F_0\bigl(h_0(x)\bigr) + F_1(x)\bigr).
$$

In the limit $\lambda \to 0$, we get

$$
F_1(x) \leq (1+\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)} F_0(x).
$$

By the consensus theorem, we get

$$
\limsup_{\lambda \to 0} p_0^\lambda\bigl(\alpha, \beta^\lambda, \pi\bigr) = \lim_{x \to 0} \limsup_{\lambda \to 0} \frac{F_1(x)}{F_0(x) + F_1(x)} \leq (1+\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 + \varepsilon p_0^*(\alpha, \pi)}.
$$

Because $\varepsilon > 0$ is arbitrary, the result follows. *Q.E.D.*

LEMMA F.2: *Let* $(\alpha^\lambda, 0)_\lambda$ *be indexed by* λ. *If* $\lim_{\lambda \to 0} \phi_1(\alpha^\lambda, \pi) > 1$, *then* $\lim_{\lambda \to 0} p_0^\lambda(\alpha^\lambda, \pi) = \lim_{\lambda \to 0} p_0^*(\alpha^\lambda, \pi)$.

PROOF: Suppose that $\alpha^\lambda \to \alpha$ such that $\phi_1(\alpha, \pi) > 1$. Given Lemma F.1, it is enough to show that $\liminf_{\lambda \to 0} p_0^\lambda(\alpha^\lambda, \pi) \geq p_0^*(\alpha, \pi)$. Additionally, given Lemma F.1, it is enough to assume that $p_0^*(\alpha, \pi) > 0$, or that $\phi_0(\alpha, \pi) > 0$. The argument is analogous to the proof of Lemma F.1. We present it here for the sake of completeness. First, we show that for each $\varepsilon > 0$, there exists $\delta > 0$ and m, n, such that $\frac{n}{m} \geq (1-\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)}$ and $h_1^n(x) \geq g_0^m(x)$ for each $x \in (0, \delta)$. For some $\varepsilon' > 0$ to be determined later, let $\phi_k = (1+\varepsilon')\phi_k(\alpha, \pi)$. Find δ such that $g_k(x) \le \phi_k x$ for each $x \in (0, \delta)$. Then, for each m, n such that $g_1^n(g_0^m(x)) \in (0, \delta)$, we have

$$g_1^n\big(g_0^m(x)\big) \le \phi_1^n \phi_0^m x.$$

Choose n so that

$$n \le m \frac{-2\varepsilon' - \log \phi_0(\alpha, \pi)}{\log \phi_1(\alpha, \pi) + 2\varepsilon'} \le n+1.$$

Then $\phi_1^n \phi_0^m \le 1$, which implies that $h_1^n(x) \ge g_0^m(x)$. Moreover, for ε' sufficiently small and m sufficiently large,

$$\frac{n}{m} \ge \frac{2\varepsilon' - \log \phi_0(\alpha, \pi)}{\log \phi_1(\alpha, \pi) - 2\varepsilon'} - \frac{1}{m} \ge (1-\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)}.$$

Next, fix $x \in (0, \delta)$. By recursively applying stationary distribution equations, we obtain

$$\begin{aligned}
F_0(x) - F_0\big(g_0^m(x)\big) &= \sum_{i=0}^{m-1}\big(F_0\big(g_0^i(x)\big) - F_0\big(g_0^{i+1}(x)\big)\big) \\
&= \lambda \sum_{i=0}^{m-1}\big(F_0\big(g_0^i(x)\big) - F_1\big(h_1\big(g_0^{i+1}(x)\big)\big)\big) \\
&\le \lambda m\big(F_0(x) - F_1\big(h_1\big(g_0^m(x)\big)\big)\big).
\end{aligned}$$

Also,

$$\begin{aligned}
F_1(x) - F_1\big(g_0^m(x)\big) &\ge F_1(x) - F_1\big(h_1^n(x)\big) \\
&= \sum_{i=0}^{n-1}\big(F_1\big(h_1^i(x)\big) - F_1\big(h_1^{i+1}(x)\big)\big) \\
&= \lambda \sum_{i=0}^{n-1}\big(F_0\big(h_0\big(h_1^i(x)\big)\big) - F_1\big(h_1^{i+1}(x)\big)\big) \\
&\ge \lambda n\big(F_0\big(h_0\big(h_1^{n-1}(x)\big)\big) - F_1\big(h_1(x)\big)\big) \\
&\ge \lambda n\big(F_0\big(g_0^m(x)\big) - F_1\big(h_1(x)\big)\big)
\end{aligned}$$

due to $g_0^m(x) \le h_1^n(x) \le h_0(h_1^{n-1}(x))$. Hence,

$$\begin{aligned}
&(1-\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)}\big(F_0(x) - F_0\big(g_0^m(x)\big)\big) - \big(F_1(x) - F_1\big(g_0^m(x)\big)\big) \\
&\le \frac{n}{m}\big(F_0(x) - F_0\big(g_0^m(x)\big)\big) - \big(F_1(x) - F_1\big(g_0^m(x)\big)\big) \\
&\le \lambda n\big[F_0(x) - F_1\big(h_1\big(g_0^m(x)\big)\big) - \big(F_0\big(g_0^m(x)\big) - F_1\big(h_1(x)\big)\big)\big] \\
&\le \lambda n\big[F_0(x) - F_0\big(g_0^m(x)\big) + F_1\big(h_1(x)\big) - F_1\big(h_1\big(g_0^m(x)\big)\big)\big]
\end{aligned}$$

Adding up across $x_0 = x$ and $x_k = g_0^m(x_{k-1})$ for all values of $k \geq 0$, and using the fact that $h_0(x_k) = g_0^{m-1}(x_{k-1})$, we get

$$F_1(x) - (1-\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)} F_0(x) \geq -\lambda n\big(F_0(x) + F_1\big(h_1(x)\big)\big).$$

In the limit $\lambda \to 0$, we get

$$F_1(x) \geq (1-\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 - p_0^*(\alpha, \pi)} F_0(x).$$

By the consensus theorem, we get

$$\limsup_{\lambda\to 0} p_0^\lambda\big(\alpha^\lambda, \pi\big) = \lim_{x\to 0} \limsup_{\lambda\to 0} \frac{F_0(x)}{F_0(x) + F_1(x)} \geq (1-\varepsilon)\frac{p_0^*(\alpha, \pi)}{1 + \varepsilon p_0^*(\alpha, \pi)}.$$

Because $\varepsilon > 0$ is arbitrary, the result follows. *Q.E.D.*

LEMMA F.3: *Suppose that* $\alpha^\lambda \to \alpha$ *and* $\phi_1(\alpha, \pi) > 1$. *For each* $\hat{p}$ *such that* $1 - \hat{p} < \lim_\lambda p_0^*(\alpha^\lambda, \pi)$, *there exists* β^λ *such that* (*i*) $\lim_{\lambda\to 0} \beta^\lambda \lambda^{-k} = 0$ *for each* k, *and* (*ii*) $p_0^\lambda(\alpha^\lambda, \beta^\lambda, \pi) = 1 - \hat{p}$ *for each sufficiently small* λ. *Moreover, if* β^λ *is a sequence such that* $p_0^\lambda(\alpha^\lambda, \beta^\lambda, \pi) = 1 - \hat{p}$ *as* $\lambda \to 0$, *then* $\lim_{\lambda\to 0} \beta^\lambda \lambda^k = 0$.

PROOF: An argument identical to the one used in Lemma C.4 shows that if $\limsup_\lambda \beta^\lambda \lambda^{-k} > 0$, then $\lim_{\lambda\to 0} p_0^\lambda(\alpha, \beta^\lambda, \pi) = 0$. Given the second part of Lemma F.2, the existence of a sequence β^λ such that $\lim_{\lambda\to 0} \beta^\lambda \lambda^{-k} = 0$ and $p_0^\lambda(\alpha, \beta^\lambda, \pi) = 1 - \hat{p}$ follows by continuity. *Q.E.D.*

REMARK F.2: For each $\alpha \in [-1, 1]$, let

$$B^\lambda(\alpha; \hat{p}, \pi) = \begin{cases} \{0\} & p_0^\lambda(\alpha, 0, \pi) < 1 - \hat{p}, \\ \big\{\beta : p_0^\lambda(\alpha, \beta, \pi) = 1 - \hat{p}\big\} & \text{otherwise.} \end{cases}$$

The definition implies that if $\beta \in B^\lambda(\alpha; \hat{p}, \pi)$, then the strategy $(\alpha, \beta) \in [-1, 1] \times [0, 1]$ is a best response (to the stationary distribution induced by itself) at sample $k = 0$.

Clearly, $B^\lambda(\alpha; \hat{p}, \pi)$ is upper hemicontinuous. The continuity and the proof of Lemma F.3 implies that, if $\alpha^\lambda \to \alpha$ and $\phi_1(\alpha, \pi) > 1$, then $B^\lambda(\alpha^\lambda; \hat{p}, \pi) \neq \emptyset$ for λ sufficiently small, and $\lim \beta^\lambda \lambda^{-k} = 0$ for each $\beta^\lambda \in B^\lambda(\alpha^\lambda, \hat{p}, \pi)$.

F.2.3. *Beliefs at Sample* $k = 1$

LEMMA F.4: *Let* $(\alpha^\lambda, \beta^\lambda)_\lambda$ *be an arbitrary strategy indexed by* λ.
Assume $\liminf_{\lambda\to 0} \phi_1(\alpha^\lambda, \pi) > 1$ *and* $\lim_{\lambda\to 0} \beta^\lambda \lambda^4 = 0$.
If $\limsup_{\lambda\to 0} \phi_0(\alpha^\lambda, \pi) < 1$, *then* $p_1^{\min}(\alpha, \pi)$ *and* $p_1^{\max}(\alpha, \pi)$ *are well-defined and*

$$p_1^{\min}(\alpha, \pi) \leq \liminf_\lambda p_1^\lambda\big(\alpha, \beta^\lambda, \pi\big) \leq \limsup_\lambda p_1^\lambda\big(\alpha, \beta^\lambda, \pi\big) \leq p_1^{\max}(\alpha, \pi).$$

If $\limsup_{\lambda\to 0} \phi_0(\alpha^\lambda, \pi) = 1$, *then*

$$0 = \lim p_1\big(\alpha, \beta^\lambda, \pi\big) = p_1^{\max}(\alpha, \pi) = 0.$$

PROOF: The proof of the first part follows the same argument as in the proof of Proposition 2 from the paper (note, in particular, that the proof of Lemma A.1 applies *verbatim*). Suppose now that $\phi_0(\alpha^\lambda, \pi) \leq 1$ and $\lim_{\lambda\to 0} \phi_0(\alpha^\lambda, \pi) = 1$. Following the same ideas as in the proof of Proposition 2, it is enough to show that, for each $\varepsilon > 0$,

$$\lim_{\lambda\to 0} \frac{1}{\lambda} \int_0^\varepsilon x dF_0(x; \alpha^\lambda, \beta^\lambda, \pi) = +\infty.$$

Using the stationarity equations, we get

$$\begin{aligned}\frac{1}{\lambda} \int_0^\varepsilon x dF_0(x; \alpha^\lambda, \beta^\lambda, \pi) &\geq \frac{1}{\lambda} \sum_{i\geq 0} g_0^{i+1}(\varepsilon)\left[F_0\left(g_0^i(\varepsilon)\right) - F_0\left(g_0^{i+1}(\varepsilon)\right)\right] \\ &\geq \sum_{i\geq 0} g_0^{i+1}(\varepsilon)\left[F_0\left(g_0^i(\varepsilon)\right) - F_1\left(h_1\left(g_0^{i+1}(\varepsilon)\right)\right)\right].\end{aligned}$$

Because of Lemma F.2, as $\lambda \to 0$, for each $x > 0$, $F_0(x) \to 1$, and $F_1(x) \to 0$. Hence, the above is not smaller than

$$\liminf_{\lambda} \sum_{i\geq 0} \bar{g}_0^{i+1}(\varepsilon),$$

where $\bar{g}_0(x) := \lim_{\lambda\to 0} g_0(x)$. Using $\lim_{\lambda\to 0} \phi_0(\alpha^\lambda, \pi) = 3((1 - |\alpha|)\pi_0 + \max(0, -\alpha)) = 1$, we get

$$\begin{aligned}\bar{g}_0(x) &= x^3 + 3x^2(1-x)\big((1-|\alpha|)\pi_0 + \max(0,\alpha)\big) \\ &\quad + 3x(1-x)^2\big((1-|\alpha|)\pi_0 + \max(0,-\alpha)\big) \\ &= x\phi_0(\alpha,\pi) - x^2\big(6\big((1-|\alpha|)\pi_0 + \max(0,-\alpha)\big) \\ &\quad - 3\big[(1-|\alpha|)\pi_0 + \max(0,\alpha)\big]\big) + o\left(x^2\right) \\ &= x\phi_0(\alpha,\pi) - x^2\big(\phi_0(\alpha,\pi) - 3\big[\max(0,\alpha) - \max(0,-\alpha)\big]\big) + o\left(x^2\right) \\ &= x - x^2(1-3\alpha) + o\left(x^2\right).\end{aligned}$$

Let $c = 2(1-3\alpha)$. Then, for sufficiently small x,

$$\bar{g}_0(x) \geq x - cx^2.$$

Find indexes $i_k = \min\{i : g_0^{*i}(x) \leq \frac{1}{2^k}\varepsilon\}$. For sufficiently small $\varepsilon > 0$, we have $1 - c\varepsilon \geq \frac{3}{4}$, and

$$\frac{2}{3} \geq \frac{\frac{1}{2^{k+1}}\varepsilon}{\frac{1}{2^k}\varepsilon(1-c\varepsilon)} \geq \frac{\bar{g}_0^{i_{k+1}}(\varepsilon)}{\bar{g}_0^{i_k}(\varepsilon)} \geq \left(1 - c\bar{g}_0^{i_k}(\varepsilon)\right)^{i_{k+1}-i_k}.$$

Hence,

$$\sum_{i\geq 0} \bar{g}_0^{i+1}(\varepsilon) = \sum_{k=0}^{\infty} \sum_{i=i_k}^{i_{k+1}-1} \bar{g}_0^i(\varepsilon) \geq \sum_{k=0}^{\infty} \frac{1 - \left(1 - c\bar{g}_0^{i_k}(\varepsilon)\right)^{i_{k+1}-i_k}}{1 - \left(1 - c\bar{g}_0^{i_k}(\varepsilon)\right)} \bar{g}_0^{i_k}(\varepsilon)$$

$$\geq \sum_{k=0}^{\infty} \frac{\frac{1}{3}}{c\bar{g}_0^{i_k}(\varepsilon)} \bar{g}_0^{i_k}(\varepsilon) = \sum_{k=0}^{\infty} \frac{1}{3c} = +\infty,$$

and the series is unbounded. *Q.E.D.*

F.2.4. *Proof of Theorem F.1*

Suppose that $(\alpha, \beta, p_0, p_1) = \lim(\alpha^\lambda, \beta^\lambda, p_0(\alpha^\lambda, \beta^\lambda; \pi), p_1(\alpha^\lambda, \beta^\lambda; \pi)) \in \Gamma(\pi, \hat{p})$ for some $(\alpha^\lambda, \beta^\lambda, p_0(\alpha^\lambda, \beta^\lambda; \pi), p_1(\alpha^\lambda, \beta^\lambda; \pi)) \in \Gamma^\lambda(\pi, \hat{p})$.

The consensus theorem implies that $\beta = 0$, $p_0 \leq 1 - \hat{p}$, and $\phi_0(\alpha, \pi) \leq 1 \leq \phi_1(\alpha, \pi)$. We have:

- $\phi_0(\alpha, \pi) \leq 1$ implies that

$$\alpha \geq \alpha^{\min}(\pi) := \begin{cases} 1 - \dfrac{1}{3(1-\pi)} & \pi < \dfrac{2}{3}, \\ \dfrac{2}{3\pi} - 1 & \pi \geq \dfrac{2}{3}, \end{cases}$$

  which further implies that $\alpha > -1$,
- $\phi_1(\alpha, \pi) \geq 1$ implies that $\alpha \leq \alpha^{\max}(\pi) = 1 - \frac{1}{3\pi}$, which further implies that $\alpha < 1$,

Consider the case when $\phi_1(\alpha, \pi) > 1$. Lemmas F.1 and F.2 imply that either $\beta^\lambda = 0$ for sufficiently small λ and $p_0 = p_0^*(\alpha, \pi) \leq 1 - \hat{p}$, or $p_0 = 1 - \hat{p} \leq p_0^*(\alpha, \pi)$. Hence, $p_0 = \min(1 - \hat{p}, p_0^*(\alpha, \pi))$. The equilibrium conditions imply that

$$\begin{cases} \alpha \in (-1, 0) & \text{if } p_1 = 1 - \hat{p}, \\ \alpha = 0 & \text{if } p_1 \in [1 - \hat{p}, \hat{p}], \\ \alpha \in (0, 1) & \text{if } p_1 = \hat{p}. \end{cases}$$

Additionally, Lemma F.4 implies that

$$p_1^{\min}(\alpha, \pi) \leq p_1 \leq p_1^{\max}(\alpha, \pi).$$

These observations imply the characterization from the theorem.

F.2.5. *Proof of Theorem F.2*

Let $B^\lambda(\alpha, \hat{p}, \pi)$ sets be defined as in Remark F.2. We consider the three cases separately.

C1 The assumptions, Lemma F.4, and continuity (including Remark F.1) imply that, for sufficiently low λ, there exist $\alpha^\lambda \in (\alpha_0, \alpha_1)$ and $\beta^\lambda \in B(\alpha^\lambda, \hat{p}, \pi)$ such that $p_1^\lambda(\alpha^\lambda, \beta^\lambda, \pi) = 1 - \hat{p}$. Due to the definition of sets $B^\lambda(\alpha^\lambda, \hat{p}, \pi)$, the strategy $\alpha^\lambda, \beta^\lambda$ is an equilibrium for sufficiently low λ. By Lemma F.2 and the construction, the limit beliefs at sample $k = 0$ are equal to $\lim p_0^\lambda(\alpha^\lambda, \beta^\lambda, \pi) = \min(p_0^*(\alpha, \pi), 1 - \hat{p})$.

C2 Let $\beta^\lambda \in B^\lambda(0, \hat{p}, \pi)$. Lemma F.4 implies that $\lim p_1^\lambda(0, \beta^\lambda, \pi) = p_1^*(0, \pi) \in (1 - \hat{p}, \hat{p})$. Hence, strategy $\sigma(0, \beta^\lambda)$ is an equilibrium strategy for each sufficiently low λ. By Lemma F.2, the limit beliefs at sample $k = 0$ are equal to $\lim p_0^\lambda(0, \beta^\lambda, \pi) = \min(p_0^*(0, \pi), 1 - \hat{p})$.

C3 The same proof as in case (1).

TABLE F.I

PARAMETER VALUES

Parameter Name	Value
`first_pi`	0.51
`n_pis`	200
`n_alphas`	200
`n_xs`	101
`phat_step`	0.0001
δ	0.001
ε	0.0001

F.3. *Numerical Computations*

The code generates the set (a grid) of $(\pi, \hat{p}, p_0)$ that satisfy conditions of Corollary F.1 and Theorem F.2. The code can be found at https://github.com/marcinpeski/Python—Stationary-learning-simulation.

F.3.1. *Generate beliefs.py*

The purpose of this script is to generate numerical approximations to the set of all tuples $(\pi, \alpha, p_0^*(\alpha, \pi), p_1^{\min}(\alpha, \pi), p_1^{\max}(\alpha, \pi))$ where $\sigma(\alpha, 0)$ is a possible limit of an equilibrium strategy (which means that $\phi_0(\alpha, \pi) < 1$, $\phi_1(\alpha, \pi) > 1$, and $g_1(x|\alpha, \pi) > x$ for all x. We take the following parameters (Table F.I).

The script relies on the following functions:

1. generate_alphas: Generate set of potential (α, π) equilibrium pairs S:
   (a) Generate set generate a grid of `no_pis` πs between `first_pi` and 1,
   (b) for each π in the grid, generate a grid of `no_alphas` αs between $\alpha^{\min}(\pi) + \delta$ and $\alpha^{\max}(\pi) - \delta$,
   (c) for each π and α, let $n(\alpha, \pi)$ be the first iteration such that $g_1^n(\varepsilon|\alpha, \pi) \geq 1 - \varepsilon$. For some αs, such n does not exist (e.g., if $g_1(x|\alpha, \pi) = x$ for some x.) We drop αs such that for some $x \in (\varepsilon, 1 - \varepsilon)$, $g_1(x|\alpha, \pi) - x \leq 0.001\epsilon$.
2. compute_beliefs_for_alpha: For all $(\alpha, \pi) \in S$, compute beliefs $\bar{p}_0(\alpha, \pi)$ and the range $[\bar{p}_1^{\min}(\alpha, \pi), \bar{p}_1^{\min}(\alpha, \pi)]$
   (a) find a grid of `n_xs` xs between ε and $g_1(\varepsilon|\alpha, \pi)$,
   (b) For each x in the grid, compute $A_k(x) = \sum_{i=0}^{n(\alpha,\pi)-1} \psi_k(x) g_1^i(x|\alpha, \pi)$,
   (c) compute beliefs $\bar{p}_0(\alpha, \pi) = p_0^*(\alpha, \pi)$, and $\bar{p}_1^{\min}(\alpha, \pi) = \min_x \frac{A_1(x)}{A_1(x)+A_2(x)}$, $\bar{p}_1^{\max}(\alpha, \pi) = \max_x \frac{A_1(x)}{A_1(x)+A_2(x)}$.

The numerical computations show that for all α, π in set S,

$$1 \times 10^{-6} \leq \bar{p}_1^{\max}(\alpha, \pi) - \bar{p}_1^{\min}(\alpha, \pi) \leq 1 \times 10^{-3}.$$

F.3.2. *Figure F.1 superset.py*

Let S^* be the set of pairs of $(\alpha, \pi) \in S$ such that either:

1. $\alpha < 0$, $p^{\max}(\alpha, \pi) > \frac{1}{2}$ and $p^{\min}(\alpha, \pi) \leq \pi$, or
2. $\alpha = 0$, $p^{\min}(\alpha, \pi) \leq \frac{1}{2}$, or
3. $\alpha > 0$, $p^{\min}(\alpha, \pi) \leq \frac{1}{2}$ and $p^{\max}(\alpha, \pi) \geq 1 - \pi$.

File Figure F.1 superset.py draws Figure F.1. Its heart is function generate_equilibria, which uses set S to generate correspondence Γ^*:

1. First, restrict set S to the set S^* of tuples $(\alpha, \pi) \in S$ such that either:
   (a) $\alpha < 0$ and $p^{\max}(\alpha, \pi) > \frac{1}{2}$ and $p^{\min}(\alpha, \pi) \leq \pi$, or
   (b) $\alpha = 0$, $\max(p^{\min}(\alpha, \pi), 1 - p^{\max}(\alpha, \pi)) \leq \pi$, or
   (c) $\alpha > 0$, $p^{\min}(\alpha, \pi) \leq \frac{1}{2}$ and $1 - p^{\max}(\alpha, \pi) \leq \pi$,
2. Next, in each of the three cases, identify the range of $\hat{p}$ that are consistent with Corollary F.1. We also classify a profile with a label X that describes qualitative behavior at sample $k = 1$,:
   (a) $\hat{p} \in [\max(\frac{1}{2}, p^{\min}(\alpha, \pi)), \min(p^{\max}(\alpha, \pi), \pi)]$, label X = "−," which corresponds to a contrarian behavior when information is not acquired,
   (b) $\hat{p} \in [\max(p^{\min}(\alpha, \pi), 1 - p^{\max}(\alpha, \pi)), \pi]$, label X = "0," which corresponds to always acquiring information,
   (c) $\hat{p} \in [\max(\frac{1}{2}, 1 - p^{\max}(\alpha, \pi)), \min(1 - p^{\min}(\alpha, \pi), \pi)]$, label X = "+," which corresponds to a regular behavior when information is not acquired.

   We denote the interval boundaries as $\hat{p}^{\min}(\alpha, \pi)$, $\hat{p}^{\max}(\alpha, \pi)$. Label X where X = " − ."
3. For each $(\alpha, \pi) \in S^*$, find a grid $P(\alpha, \pi)$ of $\hat{p} \in [\hat{p}^{\min}(\alpha, \pi), \hat{p}^{\max}(\alpha, \pi)]$ separated by `phat_step`. For each $\hat{p} \in P(\alpha, \pi)$, compute $u(\alpha, \pi, \hat{p}) = \min(p_0^*(\alpha, \pi), 1 - \hat{p})$ and associate it with a label Y that describes qualitative behavior at sample $k = 0$:
   (a) label $Y = 0$ if $p_0^*(\alpha, \pi) < 1 - \hat{p}$, which corresponds to never acquiring information, and
   (b) label Y = "A" if $p_0^*(\alpha, \pi) \geq 1 - \hat{p}$, which corresponds to sometimes acquiring information.

Set Γ^*consists of tuples $(\pi, \hat{p}, u(\alpha, \pi, \hat{p}))$ for all $(\alpha, \pi) \in S^*$ and $\hat{p} \in P(\alpha, \pi)$. This set is drawn on the left panel of Figure F.1, with label Y illustrated with two colors and label $X$with hatching pattern.

APPENDIX G: THE CASE $n = 3$: TIGHTER BOUNDS FOR THEOREM 3

For $n = 3$, we here show that the conclusion of Theorem 3 holds whenever $n^2\pi(1-\pi) < 1$ and c is small: under these conditions, there is an equilibrium with welfare $\hat{p}$ in which all agents acquire information with strictly positive probability, irrespective of their sample. This follows from the proof of Theorem 3, and from the result below, where $l(p) := \frac{p}{1-p}$.

PROPOSITION G.1: *If π is such that $3^2\pi(1 - \pi) < 1$ and c is small, one has*

$$l(1 - \hat{p}) < \inf_{(x_i)_{i \in \mathbf{Z}}} \frac{\sum_{i \in \mathbf{Z}} \psi_2(x_i)}{\sum_{i \in \mathbf{Z}} \psi_1(x_i)} \leq \sup_{(x_i)_{i \in \mathbf{Z}}} \frac{\sum_{i \in \mathbf{Z}} \psi_2(x_i)}{\sum_{i \in \mathbf{Z}} \psi_1(x_i)} < l(\hat{p}),$$

where the infimum and supremum are taken over all (doubly infinite) orbits of $\bar{g}_1$.

We stress that we make no attempt below at getting optimal bounds. Irrespective of π, the proof works as long as $\hat{p} \geq 0.85$. Numerical results suggest that the result holds whenever $\hat{p} \geq 0.6$.

PROOF: The assumption on π implies $\pi \geq 0.87$. We fix $\hat{p}$ s.t. $0.85 \leq \hat{p} < \pi$.
We recall that $\bar{g}_1(x) = \pi + x^3(1 - \pi) - (1 - x)^3\pi$, and that $\psi_k(x) = x^k(1 - x)^{n-k}$.

Let (x_i) be an orbit of $\bar{g}_1$ and set

$$i_0 := \max\{i : x_i \leq 1 - \hat{p}\} \quad \text{and} \quad i_1 := \min\{i : x_i \geq \hat{p}\}.$$

We list an number of elementary observations, which we later combine to deliver the proof, and next prove the upper and lower bounds in turn. *Q.E.D.*

Preliminary observations

O.1 : $\bar{g}_1'(x) = 3\{x^2(1-\pi) + (1-x)^2\pi\}$ is convex with a minimum at π. Note also that, viewed as a function of π, $\bar{g}_1$ is increasing. The inequality $\bar{g}_1(\frac{5}{11}) < 0.84$ holds for $\pi = 1$, hence for all π. This implies that $\bar{g}_1(\frac{5}{11}) < \hat{p}$ irrespective of $\hat{p}$ and π, hence $x_{i_1-1} > \frac{5}{11}$.

O.2 : If $r > R > 1$ and $(y_i)_{i\in\mathbf{Z}}$ are such that $ry_i \leq y_{i+1} \leq Ry_i$ for each $i < h$, then

$$\frac{R}{R-1} y_h \leq \sum_{i\leq h} y_i \leq \frac{r}{r-1} y_h.$$

O.3 : Since $\bar{g}_1'$ is decreasing on $[0, 1-\hat{p}]$ (see **O.1**) and $\bar{g}_1(0) = 0$, one has

$$\bar{g}_1'(1-\hat{p})x_i \leq x_{i+1} \leq \bar{g}_1'(0)x_i \quad \text{for each } i < i_0,$$

hence by **O.2**,

$$\frac{\bar{g}_1'(0)x_{i_0}}{\bar{g}_1'(0)-1} \leq \sum_{i\leq i_0} x_i \leq \frac{\bar{g}_1'(1-\hat{p})x_{i_0}}{\bar{g}_1'(1-\hat{p})-1}.$$

Since $1-\hat{p} \leq x_{i_0+1} \leq \bar{g}_1'(0)x_{i_0}$ and $x_{i_0} \leq 1-\hat{p}$, this yields

$$\frac{(1-\hat{p})}{\bar{g}_1'(0)-1} \leq \sum_{i\leq i_0} x_i \leq \frac{\bar{g}_1'(1-\hat{p})(1-\hat{p})}{\bar{g}_1'(1-\hat{p})-1} \tag{G.1}$$

O.4 : For $i \geq i_1$ and since $\bar{g}_1(1) = 1$, one has

$$\left(\min_{[\hat{p},1]} \bar{g}_1'\right) \times (1-x_i) \leq 1 - x_{i+1} \leq \left(\max_{[\hat{p},1]} \bar{g}_1'\right) \times (1-x_i).$$

By **O.1**, this rewrites $\bar{g}_1'(\pi)(1-x_i) \leq 1 - x_{i+1} \leq G(\hat{p})(1-x_i)$, where $G(\hat{p}) := \max(\bar{g}_1'(\hat{p}), \bar{g}_1'(1))$. Since $1 - x_{i_1} \leq 1-\hat{p} \leq 1 - x_{i_1-1}$, one gets

$$\sum_{i\geq i_1}(1-x_i) \leq \frac{1-x_{i_1}}{1-G(\hat{p})} \leq \frac{1-\hat{p}}{1-G(\hat{p})} \tag{G.2}$$

and

$$\sum_{i>i_1}(1-x_i) \geq \frac{1-x_{i_1-1}}{1-\bar{g}_1'(\pi)} \geq \frac{1-\hat{p}}{1-\bar{g}_1'(\pi)}. \tag{G.3}$$

O.5 : ψ_2 is single-peaked with a maximum at $\frac{2}{3}$.

O.6 : For $i \leq i_0$, $\hat{p}^2 x_i \leq \psi_1(x_i) \leq x_i$ and $\psi_2(x_i) \leq x_i^2$.

O.7 : For $i \geq i_1$, one has $\psi_2(x_{i_1-1}) \geq \min(\psi_2(\frac{5}{11}), \psi_2(\hat{p}))$, using **O.5** and $x_{i_1-1} > \frac{5}{11}$. Since $\psi_2(0.85) < \psi_2(\frac{5}{11})$ and $\hat{p} \geq 0.85$, one has $\psi_2(\hat{p}) < \psi_2(\frac{5}{11})$, so that

$$\psi_2(x_{i_1-1}) > \psi_2(\hat{p}). \tag{G.4}$$

The upper bound

We set $N_0 := \sum_{i=i_0+1}^{i_1-1} \psi_2(x_i)$, $D_0 := \sum_{i=i_0+1}^{i_1-1} \psi_1(x_i)$, $N_1 := \sum_{i\leq i_0} \psi_2(x_i)$, $D_1 := \sum_{i\leq i_0} \psi_1(x_i)$, $N_2 := \sum_{i\geq i_1} \psi_2(x_i)$, and $D_2 := \sum_{i\geq i_1} \psi_1(x_i)$, so that

$$R := \frac{\sum_{i\in\mathbf{Z}} \psi_2(x_i)}{\sum_{i\in\mathbf{Z}} \psi_1(x_i)} = \frac{N_0+N_1+N_2}{D_0+D_1+D_2}.$$

We prove that $\frac{N_0+N_1+N_2}{D_0+D_1} < l(\hat{p})$.

- $\frac{N_0}{D_0}$. For $i_0 < i < i_1$, one has $x_i < \hat{p}$, hence $\frac{\psi_2(x_i)}{\psi_1(x_i)} = l(x_i) < l(\hat{p})$.
- $\frac{N_1}{D_1}$. For $i \leq i_0$, one has $x_i \leq 1-\hat{p}$, hence $\frac{\psi_2(x_i)}{\psi_1(x_i)} = l(x_i) \leq l(1-\hat{p})$.
- $\frac{N_2}{D_1}$. For $i \geq i_1$, $\psi_2(x_i) \leq 1-x_i$. For $i \leq i_0$, $\psi_1(x_i) \geq \hat{p}^2 x_i$ (**O.6**), hence

$$\frac{N_2}{D_1} \leq \frac{1}{\hat{p}^2} \times \frac{\sum_{i\geq i_1}(1-x_i)}{\sum_{i\leq i_0} x_i} < \frac{1}{\hat{p}^2} \times \frac{\bar{g}_1'(0)-1}{1-G(\hat{p})} := H(\hat{p}). \tag{G.5}$$

It follows that

$$\frac{N_0+N_1+N_2}{D_0+D_1} < \max\bigl(l(\hat{p}), l(1-\hat{p})+H(\hat{p})\bigr)$$

The claim below concludes the proof of the upper bound.

CLAIM G.1: *One has $l(1-\hat{p})+H(\hat{p}) < l(\hat{p})$.*

PROOF: $H(\hat{p}) = \frac{1}{\hat{p}^2} \times \max(\frac{3\pi-1}{1-\bar{g}_1'(1)}, \frac{3\pi-1}{1-\bar{g}_1'(\hat{p})})$. For fixed $\hat{p}$, the quantity $\frac{3\pi-1}{1-\bar{g}_1'(1)}$ is decreasing in π while $\frac{3\pi-1}{1-\bar{g}_1'(\hat{p})}$ is increasing in π. Therefore,

$$H(\hat{p}) \leq \frac{1}{\hat{p}^2} \times \max\left(\frac{3\hat{p}-1}{3\hat{p}-2}, \frac{2}{1-(1-\hat{p})^2}\right).$$

The result then follows from simple computations. *Q.E.D.*

The lower bound

We set $\tilde{N}_0 := \sum_{i=i_0+1}^{i_1-2} \psi_2(x_i)$ and $\tilde{D}_0 := \sum_{i=i_0+1}^{i_1-2} \psi_1(x_i)$, so that

$$R := \frac{\sum_{i\in\mathbf{Z}} \psi_2(x_i)}{\sum_{i\in\mathbf{Z}} \psi_1(x_i)} = \frac{\tilde{N}_0+\psi_2(x_{i_1-1})+N_1+N_2}{\tilde{D}_0+\psi_1(x_{i_1-1})+D_1+D_2}.$$

We prove that $\frac{\tilde{N}_0+\psi_2(x_{i_1-1})+N_2}{\tilde{D}_0+\psi_1(x_{i_1-1})+D_1+D_2} > l(1-\hat{p})$ (ignoring $\tilde{N}_0$ and $\tilde{D}_0$ if these sums are empty).

- $\frac{\psi_2(x_{i_1-1})}{\psi_1(x_{i_1-1})} = l(x_{i_1-1}) \geq l(\frac{5}{11}) = \frac{5}{6}$, using **O.1**.
- $\frac{N_2}{D_2}$. For $i \geq i_1$, $\frac{\psi_2(x_i)}{\psi_1(x_i)} = l(x_i) \geq l(\hat{p})$
- $\frac{\tilde{N}_0}{\tilde{D}_0}$. For $i_0 < i < i_1 - 1$, one has $x_i > 1-\hat{p}$, hence $\frac{\psi_2(x_i)}{\psi_1(x_i)} = l(x_i) > l(1-\hat{p})$.
- $\frac{\psi_2(x_{i_1-1})}{D_1}$. By **O.7**, $\psi_2(x_{i_1-1}) \geq \psi_2(\hat{p}) = \hat{p}^2(1-\hat{p})$. On the other hand, $D_1 \leq \sum_{i\leq i_0} x_i$. Using **O.3**, this implies

$$\frac{\psi_2(x_{i_1-1})}{D_1} > \hat{p}^2 \times \frac{\bar{g}_1'(1-\hat{p})-1}{\bar{g}_1'(1-\hat{p})}.$$

Define $\phi(p_*)$ by the equality $\frac{1}{\phi(\hat{p})} := \frac{6}{5} + \frac{\bar{g}_1'(1-\hat{p})}{\hat{p}^2(\bar{g}_1'(1-\hat{p})-1)}$. The last two inequalities imply

$$\frac{\psi_2(x_{i_1-1})}{\psi_1(x_{i_1-1})+D_1} \geq \phi(\hat{p}).$$

One obtains

$$\frac{\tilde{N}_0+\psi_2(x_{i_1-1})+N_2}{\tilde{D}_0+\psi_1(x_{i_1-1})+D_1+D_2} > \min\bigl(l(\hat{p}), l(1-\hat{p}), \phi(\hat{p})\bigr)$$

The result follows from the claim below.

CLAIM G.2: *One has* $\phi(\hat{p}) > l(1-\hat{p})$.

PROOF: One needs to prove that

$$\frac{6}{5} + \frac{\bar{g}_1'(1-\hat{p})}{\hat{p}^2\bigl(\bar{g}_1'(1-\hat{p})-1\bigr)} < l(\hat{p}).$$

For fixed $\hat{p}$, the left-hand side is decreasing in π and, therefore, highest for $\pi = \hat{p}$. So, the claim is equivalent to

$$\frac{6}{5} + \frac{1}{\hat{p}^2} \times \left(1 + \frac{1}{3\hat{p}^3 + 3\bigl(1-\hat{p}^3\bigr) - 1}\right) < l(\hat{p}),$$

which holds for $\hat{p} \geq 0.85$. *Q.E.D.*

APPENDIX H: ENDOGENOUS INFORMATION STRUCTURES

We assume in the baseline version that agents have access to a fixed-information source at a lumpy cost, but the choice of this source could be *endogenized*. Assume that agents, upon seeing their samples, can choose any statistical experiment μ, at a cost $C(\mu)$.

Given interim beliefs p, an agent optimally obtains

$$v(p) = \sup_{\mu}\bigl\{p\mu_1(1) + (1-p)\mu_0(0) - C(\mu)\bigr\},$$

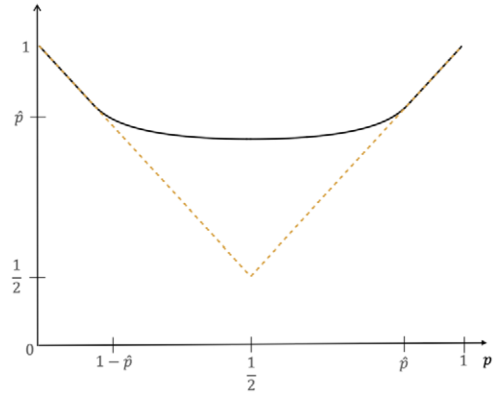

FIGURE H.1.—The value of information with endogenous information structures.

where $\mu_\theta(a)$ is the probability of (optimally choosing) action a in state θ under the information structure μ. As in the baseline model, there exists $\hat{p} \in [\frac{1}{2}, 1]$ such that $v(p) - c > u(p)$ if and only if $p \in (1-\hat{p}, \hat{p})$. The central question is whether $\hat{p} < 1$ or $\hat{p} = 1$. The answer depends on the functional $C(\cdot)$.

We follow Pomatto, Strack, and Tamuz (2023), who argue that an experiment μ may without loss be identified with the probabilities $\mu_\theta(a)$ of (optimally choosing) action a in state θ and use an axiomatic approach, and assume that the cost-of-information functional is given by $C(\mu) = \kappa \sum_{a \in A, \theta \in \Theta} \mu_\theta(a) \ln \frac{\mu_\theta(a)}{\mu_{1-\theta}(a)}$, for some κ. With such a functional, we show that there exists $\hat{p} < 1$ such that $v(p) = u(p)$ for each $p \notin [1-\hat{p}, \hat{p}]$.

This setup is therefore equivalent to having a single, exogenous source of information with bounded strength, which is implicitly defined by Figure H.1.[4]

PROOF: To simplify notation, we set $x = \mu_1(1)$ and $y = \mu_0(0)$. Elementary algebra yields

$$v(p) = \sup_{x,y}\big(px + (1-p)y - C(x,y)\big)$$
$$= \sup_{x,y}\left(px + (1-p)y - \kappa(x+y-1)\ln\frac{xy}{(1-x)(1-y)}\right).$$

We argue by contradiction and assume that there are interim beliefs arbitrarily close to 1 such that $v(p) > p$. For such p, we simply denote by x and y values such that the expression in the supremum in (H.6) exceeds p.

Since

$$px + (1-p)y - C(x,y) > p \tag{H.1}$$

for each p, it follows that $\lim_{p\to 1} x = 1$ and $\lim_{p\to 1} C(x,y) = 0$, which implies $\lim_{p\to 1} y = 0$. Up to a subsequence, we may assume that the limit $\alpha := \lim_{p\to 1} \frac{y}{1-x} \in [0, +\infty]$ exists.

We rewrite (H.1) as

$$-p + (1-p)\frac{y}{1-x} - \frac{C(x,y)}{1-x} > 0. \tag{H.2}$$

[4]Figure H.1 assumes $\kappa = 0.3$.

Assume first $\alpha \in (0, +\infty)$. In that case, $C(x, y) \sim_{p\to 1} (1-x)(\alpha-1)\ln\alpha$, hence (H.2) implies in the limit

$$-1 + (1-\alpha)\ln\alpha \geq 0,$$

which cannot possibly hold.

Assume next $\alpha = 0$. Inequality (H.2) implies $-p + (1-p)\frac{y}{1-x} > 0$ for p close to 1, which does not hold.

Assume finally $\alpha = +\infty$. Inequality (H.2) writes

$$-p + (1-p)\frac{y}{1-x} - \left(\frac{y}{1-x} - 1\right)\left(\ln\frac{y}{1-x} + \ln\frac{x}{1-y}\right) > 0,$$

which does not hold when p is close to 1 and $\frac{y}{1-x}$ is large enough. *Q.E.D.*

APPENDIX I: ASYMMETRIC VERSIONS

In this section, we drop all symmetry assumptions. Specifically, the invariant probability of state 1 is $p^* \in (0, 1)$, the utility is an arbitrary function $u : A \times \Theta \to [0, 1]$, and the (unconditional) distribution of private beliefs is an arbitrary distribution $H \in \Delta([0, 1])$ with expectation $\frac{1}{2}$, from which the state-conditional distributions H_θ are uniquely determined.

We assume that either signals are costly ($c > 0$), or that the distribution of beliefs has bounded precision.

ASSUMPTION I.1: *Either $c > 0$, or $H([\underline{q}, \bar{q}]) = 1$ for some $0 < \underline{q} < \bar{q} < 1$.*

Assumption I.1 implies the existence of $0 < p_0^* \leq p_1^* < 1$ such that (for $c > 0$) information acquisition is strictly optimal if and only if $p \in (p_0^*, p_1^*)$ and (for $c = 0$) signal are ignored unless $p \in (p_0^*, p_1^*)$. The cutoffs p_0^* and p_1^* play the role of $1 - \hat{p}$ and $\hat{p}$, respectively. To rule out trivialities, we assume that $p_0^* < p^* < p_1^*$. This is a joint assumption on all primitives of the model.

The notion of an equilibrium steady state in this asymmetric environment is the same as in the paper, with the symmetry requirement dropped.

I.1. *The Consensus Result*

THEOREM I.1—Asymptotic consensus: *Assume $n \geq 2$. There exists a constant $K < \infty$, depending only on the primitives $u(\cdot)$ and H such that the following holds. For every $\lambda > 0$ and every equilibrium steady state (μ, σ), one has*

$$\int_{\Theta\times[0,1]} x(1-x)\,d\mu(\theta, x) \leq K\lambda.$$

PROOF: Set

$$q_{\min} := \frac{1}{2}\min_\theta\left|p^* - p_\theta^*\right| \quad \text{and} \quad d := \min_{[p^*-q_{\min}, p^*+q_{\min}]} v - u.$$

Note that $d > 0$ since $v > u$ on $[p_0^*, p_1^*]$. We will prove that $\int_{\Theta\times[0,1]} x(1-x)\,d\mu(\theta, x) \leq \frac{1}{q_{\min} d}\lambda$ for each equilibrium steady state (μ, σ) of $G(\lambda, \rho)$, as soon as $\lambda \leq \lambda_* := (1 - q_{\min})\,d$.

Below, we assume $\lambda < \lambda_*$ and we write $f(\nu) := \int_{[0,1]} f\,d\nu$ for each $\nu \in \Delta([0,1])$ and measurable $f : \Delta([0,1]) \to \mathbf{R}$. Let an equilibrium steady state (μ, σ) be given, with equilibrium payoff v^*. Consider a thought experiment in which an agent observes the actions $a^{(1)}, \dots, a^{(n)}$ in her sample in some random order, and denote by $\mathcal{F}_m$ the information structure induced by the observation of the first m actions. We view $\mathcal{F}_m$ as a distribution over (interim) beliefs, hence $\mathcal{F}_m \in \Delta([0,1])$. With these notation, $v^* = v(\mathcal{F}_n)$. Because more information cannot hurt, we have

$$v^* = v(\mathcal{F}_n) \geq \cdots \geq v(\mathcal{F}_1) \quad \text{and} \quad v(\mathcal{F}_m) \geq u(\mathcal{F}_m). \tag{I.1}$$

We look more closely at the information structures $\mathcal{F}_0$, $\mathcal{F}_1$, and $\mathcal{F}_2$, and introduce some notation. We denote by $\nu := \int_{\Theta\times[0,1]} (1-x)\,d\mu(\theta, x)$ the probability of $a^{(1)} = 0$, and by $q_a = \mathbf{P}(\theta = 1 | a^{(1)} = a)$ the interim belief when first sampling $a \in A$. We also denote by ν_a be the conditional probability of next sampling $a^{(2)} = 0$ given $a^{(1)} = a$, and by $q_{aa'} := \mathbf{P}(\theta = 1 | (a^{(1)}, a^{(2)}) = (a, a'))$ the interim belief given the first two actions in the sample. Using the notation $(a_1)^{p_1} \cdots (a_k)^{p_k}$ to describe the finite-support probability distribution that assigns probability p_m to a_m, we thus have

$$\mathcal{F}_0 = \left(p^*\right)^1, \tag{I.2}$$

$$\mathcal{F}_1 = (q_0)^{\nu}(q_1)^{(1-\nu)}, \tag{I.3}$$

$$\mathcal{F}_2 = (q_{00})^{\nu\nu_0}(q_{01})^{\nu(1-\nu_0)}(q_{10})^{(1-\nu)\nu_1}(q_{11})^{(1-\nu)(1-\nu_1)}. \tag{I.4}$$

Because $a^{(1)}$ and $a^{(2)}$ are exchangeable, we have

$$q_{01} = q_{10} \quad \text{and} \quad \nu(1-\nu_0) = (1-\nu)\nu_1.$$

By the martingale property of beliefs, the expected belief in all three information structures is the same and equal to

$$p^* = \nu q_0 + (1-\nu)q_1 = \nu\nu_0 q_{00} + 2\nu(1-\nu_0)q_{01} + (1-\nu)(1-\nu_1)q_{11}. \tag{I.5}$$

The result relies on the fundamental observation below.

LEMMA I.1: *One has $v^* - \lambda \leq u(\mathcal{F}_1)$.*

PROOF: Consider a strategy of observing and replicating $a^{(1)}$. The payoff from such a strategy is $u(\mathcal{F}_1)$. On the other hand, apart from the possibility that the state has changed since the sampled individual took her action, the payoff from such a strategy is equal to the payoff of the sampled individual. The probability that the state has changed is equal to λ. But the expected payoff of the sampled individual is equal to v^*. *Q.E.D.*

Together with (I.1), Lemma I.1 readily implies

$$v(\mathcal{F}_1) - u(\mathcal{F}_1) \leq \lambda \quad \text{and} \quad v(\mathcal{F}_2) - v(\mathcal{F}_1) \leq \lambda. \tag{I.6}$$

For concreteness, we assume below that $q_{10} \geq p^*$.[5] Also, notice that $q_1 > q_0$—this holds since θ and a are positively correlated in any equilibrium steady state, and since states are persistent.

[5]This is w.l.o.g., since the argument below has a flip side otherwise.

LEMMA I.2: *One has* $\nu, 1-\nu \geq q_{\min}$.

PROOF: We prove that $\nu \geq q_{\min}$. A similar argument shows that $1-\nu \geq q_{\min}$. Thanks to (I.5), one has

$$\nu = \frac{q_1 - p^*}{q_1 - q_0}. \tag{I.7}$$

We argue by contradiction: if $\nu \leq q_{\min}$, (I.7) implies that $q_1 - p^* \leq q_{\min}$, or $q_1 \leq p^* + q_{\min}$. Using (I.6), it follows that

$$\begin{aligned}\lambda &\geq v(\mathcal{F}_1) - u(\mathcal{F}_1)\\ &= \nu\big[v(q_0) - u(q_0)\big] + (1-\nu)\big[v(q_1) - u(q_1)\big]\\ &\geq (1-\nu)\big[v(q_1) - u(q_1)\big] \geq (1-q_{\min})\, d = \lambda_*,\end{aligned}$$

where the last inequality follows from the choice of d, since $p^* \leq q_1 \leq p^* + q_{\min}$. This contradicts the assumption $\lambda < \lambda_*$. *Q.E.D.*

We collect two consequences in the claim below. In this statement, $\kappa : [0,1] \to \mathbf{R}$ is a tangent line to u at $p = q_0$, that is, $\kappa(\cdot)$ is affine, with $\kappa \leq u$ and $\kappa(q_0) = u(q_0)$. Note that $\kappa \leq v$ since $u \leq v$.

CLAIM I.1: *One has* $v(q_{01}) - \kappa(q_{01}) \geq d$ *and* $v(q_0) \leq \kappa(q_0) + \frac{1}{q_{\min}}\lambda$.

PROOF OF THE CLAIM: Start with the first inequality and note first that $q_0 \leq q_1$ implies $q_0 \leq p^*$ since p^* is an average of q_0 and of q_1. Therefore, $q_0 \leq p^* \leq q_{01}$. Since $\kappa \leq u \leq v$, and by convexity of v, $v - \kappa$ is nondecreasing on $[q_0, 1]$. In particular, $v(q_{01}) - \kappa(q_{01}) \geq v(p^*) - \kappa(p^*) \geq v(p^*) - u(p^*) \geq d$, where the last inequality follows from the definition of d.

The second inequality follows from (I.6) and Lemma I.2. Indeed,

$$\begin{aligned}\lambda &\geq v(\mathcal{F}_1) - u(\mathcal{F}_1)\\ &= \nu\big[v(q_0) - u(q_0)\big] + (1-\nu)\big[v(q_1) - u(q_1)\big].\\ &\geq q_{\min}\big[v(q_0) - u(q_0)\big] = q_{\min}\big[v(q_0) - \kappa(q_0)\big].\end{aligned}$$

Q.E.D.

We are now in a position to conclude. Starting from (I.6), one has the following sequence of inequalities:

$$\begin{aligned}\lambda &\geq v(\mathcal{F}_2) - v(\mathcal{F}_1)\\ &= \mathbf{E}_\mu\big[v(q_{a^{(1)}a^{(2)}}) - v(q_{(a^1)})\big]\\ &\geq \mathbf{P}\big(a^{(1)} = 0\big)\mathbf{E}\big[v(q_{a^{(1)}a^{(2)}}) - v(q_{(a^1)})\big| a^{(1)} = 0\big]\\ &\geq q_{\min}\big[\nu_0 v(q_{00}) + (1-\nu_0) v(q_{10}) - v(q_0)\big] &\text{(I.8)}\\ &\geq q_{\min}\big[\nu_0 \kappa(q_{00}) + (1-\nu_0)\kappa(q_{10}) - v(q_0)\big] + q_{\min}\, d(1-\nu_0) &\text{(I.9)}\\ &\geq q_{\min}\big[\nu_0 \kappa(q_{00}) + (1-\nu_0)\kappa(q_{10}) - \kappa(q_0)\big] - \lambda + (1-\nu_0)\, d q_{\min} &\text{(I.10)}\\ &\geq -\lambda + (1-\nu_0)\, d q_{\min}, &\text{(I.11)}\end{aligned}$$

where (I.8) holds since $\nu \geq q_{\min}$, (I.9) holds since $v \geq \kappa$ and $v(q_{01}) \geq \kappa(q_{01}) + d$, (I.10) holds using Claim I.1, and (I.11) holds since κ is affine, so that

$$\nu_0 \kappa(q_{00}) + (1-\nu_0)\kappa(q_{10}) = \kappa\big(\nu_0 q_{00} + (1-\nu_0) q_{10}\big) = \kappa(q_0).$$

The inequality (I.11) rewrites $1 - \nu_0 \leq \frac{2}{q_{\min} d}\lambda$. Since

$$\int_{\Theta\times[0,1]} x(1-x)\, d\mu(x) = \mathbf{P}\big(a^{(1)} = 0, a^{(2)} = 1\big) = \nu(1-\nu_0),$$

the result follows. *Q.E.D.*

APPENDIX J: SAMPLING FROM THE FURTHER PAST: THE CASE $n = 2$

Our analysis seamlessly accommodates situations where actions are sampled from the more distant past. We indicate here how, focusing for concreteness on Proposition 2. Specifically, assume that, in period t, past actions are each sampled from a random, possibly different, period $t - \tau$, where the lag $\tau \geq 1$ follows a geometric distribution with parameter $\rho < 1$, and that the vintages $t - \tau$ of the sampled actions are unobserved. For $\rho = 1$, actions are sampled from the previous period, as in baseline version. In the limit $\rho \to 0$, τ is uniformly distributed over the infinite past. This extension bends itself to an alternative interpretation. Under this equivalent narrative, a fraction ρ of the population is replaced in each period, agents are long-lived and act (only) when arriving. In such a extended setup, all our results still hold as stated.

When the vintage of sampled actions is drawn according to a geometric distribution with parameter ρ (or equivalently, when a fraction ρ of the population is replaced each period), the analysis of the case $n = 2$ requires minor adjustments, which we provide here. The notation is the same as in the main text.

Since agents who observe a balanced sample $k = 1$ hold the interim belief $p_1 = \frac{1}{2}$ and acquire information, the fraction of newborn agents choosing action 1 in period t is

$$\chi_t := x_{t-1}^2 + 2x_{t-1}(1 - x_{t_1})\phi_{\theta_t},$$

Hence, the fraction x_t of the period t population playing action 1 is given by

$$x_t = \sum_{m\geq 1} \rho(1-\rho)^{m-1}\chi_{t+1-m} = (1-\rho)x_{t-1} + \rho\chi_t.$$

Putting things together, the sequence (x_t) follows the recursive equation:

$$x_{t+1} = (1-\rho)x_t + \rho\big\{x_t^2 + 2x_t(1-x_t)\phi_{\theta_{t+1}}\big\}.$$

Set now $\phi_\theta^\rho := (1-\rho) + \rho \times 2\phi_\theta$ and observe that $\phi_0^\rho\phi_1^\rho < 1$. Indeed, $(\phi_0^\rho, \phi_1^\rho) \in \mathbf{R}^2$ is a convex combination of $(1, 1)$ and of $(2\phi_0, 2\phi_1)$ and, therefore, lies on the straight line with equation $y_0 + y_1 = 2$. This line is tangent to the (half-)hyperbola $\mathcal{C}$ of equation $y_0 y_1 = 1$ $(y_0, y_1 > 0)$ at the point $(1, 1)$, and strictly "below" $\mathcal{C}$, except at the point of tangency. Since $\rho > 0$, it follows that $\phi_0^\rho\phi_1^\rho < 1$.

The rest of the proof is identical.

REFERENCES

BANERJEE, ABHIJIT, AND DREW FUDENBERG (2004): "Word-of-Mouth Learning," *Games and economic behavior*, 46 (1), 1–22. [0001]

BIKHCHANDANI, SUSHIL, DAVID HIRSHLEIFER, OMER TAMUZ, AND IVO WELCH (2024): "Information Cascades and Social Learning," *Journal of Economic Literature*, 62 (3), 1040–1093. [0003]

DASARATHA, KRISHNA, BENJAMIN GOLUB, AND NIR HAK (2023): "Learning From Neighbours About a Changing State," *Review of Economic Studies*, 90 (5), 2326–2369. [0003]

POMATTO, LUCIANO, PHILIPP STRACK, AND OMER TAMUZ (2023): "The Cost of Information: The Case of Constant Marginal Costs," *American Economic Review*, 113 (5), 1360–1393. [0017]

Co-editor Asher Wolinsky handled this manuscript.

Manuscript received 7 January, 2022; final version accepted 13 August, 2024; available online 16 August, 2024.