Econometrica Supplementary Material

SUPPLEMENT TO "A THEORY OF NON-BAYESIAN SOCIAL LEARNING"

(*Econometrica*, Vol. 86, No. 2, March 2018, 445–490)

POOYA MOLAVI
Department of Economics, Massachusetts Institute of Technology

ALIREZA TAHBAZ-SALEHI
Kellogg School of Management, Northwestern University

ALI JADBABAIE
Institute for Data, Systems, and Society, Massachusetts Institute of Technology

THIS ONLINE SUPPLEMENT contains three parts. Appendix B contains several proofs omitted from the main body of the paper. Appendix C extends the concept of unanimity introduced in Definition 1 and provides a generalization to Theorem 4 in the paper. Appendix D illustrates that if signals are normally distributed and agents' prior beliefs are normal, Bayesian updating takes a log-linear form regardless of the structure of the underlying social network.

APPENDIX B: OMITTED PROOFS

This appendix contains the proofs of Lemmas A.3, A.5, A.9, A.10, A.14, A.16 and parts (b) and (c) of Theorem 7, omitted from the main body of the paper.

Proof of Lemma A.3

By definition, $a_{it,j\tau} = 0$ whenever $t < \tau$. Therefore, it is sufficient to restrict our attention to the case that $t \geq \tau$. Furthermore, (22) implies that $a_{it,j\tau} = a_{it-\tau,j0}$. Consequently, it is sufficient to establish that $a_{it,j0} = 0$ for all $j \notin N_i$ and all $t \geq 0$. We prove this statement by strong induction.

To establish the base case, note that (22) guarantees that $a_{l0,k0} = 0$ for all pairs of agents k and l such that $k \notin N_l$. As the induction hypothesis, fix a time instance t and suppose that $a_{lr,k0} = 0$ for all $r < t$ and all pairs of agents l and k such that $k \notin N_l$.

Consider two distinct agents i and j such that $j \notin N_i$, which is equivalent to assuming that $d(i,j) > 1$. By definition,

$$a_{it,j0} = -\sum_{r=0}^{t-1} \sum_{k:\substack{d(i,k)\leq t-r \\ d(k,j)\leq r}} a_{kr,j0} = -\sum_{r=0}^{t-2} \sum_{k:\substack{d(i,k)\leq t-r \\ d(k,j)\leq r}} a_{kr,j0} - \sum_{k:\substack{d(i,k)\leq 1 \\ d(k,j)\leq t-1}} a_{kt-1,j0}.$$

Pooya Molavi: mpooya@mit.edu
Alireza Tahbaz-Salehi: alirezat@kellogg.northwestern.edu
Ali Jadbabaie: jadbabai@mit.edu

 https://doi.org/10.3982/ECTA14613

By the induction hypothesis, $a_{kt-1,j0}$ is nonzero only if $j \in N_k$, in which case, $d(k,j) \le 1$. Hence,

$$a_{it,j0} = -\sum_{r=0}^{t-2} \sum_{k:\substack{d(i,k)\le t-r \\ d(k,j)\le r}} a_{kr,j0} - \sum_{k:\substack{d(i,k)\le 1 \\ d(k,j)\le 1}} a_{kt-1,j0} = -\sum_{r=0}^{t-2} \sum_{k:\substack{d(i,k)\le t-r \\ d(k,j)\le r}} a_{kr,j0} - \sum_{k:\substack{d(i,k)=1 \\ d(k,j)=1}} a_{kt-1,j0},$$

where the second equality is a consequence of the assumption that $d(i,j) > 1$. Fix a shortest path that connects agent j to i and let $l \in N_i \setminus \{i\}$ denote the agent in i's neighborhood that lies on this path. Note that the assumption that $d(i,j) > 1$ implies that $l \neq j$. Expanding the first term on the right-hand side above, we have

$$a_{it,j0} = -\sum_{r=0}^{t-2} \sum_{k:\substack{d(i,k)\le t-r \\ d(k,j)\le r \\ d(l,k)\le t-r-1}} a_{kr,j0} - \sum_{r=0}^{t-2} \sum_{k:\substack{d(i,k)\le t-r \\ d(k,j)\le r \\ d(l,k)> t-r-1}} a_{kr,j0} - \sum_{k:\substack{d(i,k)=1 \\ d(k,j)=1}} a_{kt-1,j0}.$$

The fact that $d(i,l) = 1$, together with the triangle inequality, implies that if $d(l,k) \le t-r-1$, then $d(i,k) \le t-r$. Therefore, by (22), the first term on the right-hand side above is equal to $a_{lt-1,j0}$. Hence,

$$a_{it,j0} = a_{lt-1,j0} - \sum_{r=0}^{t-2} \sum_{k:\substack{d(i,k)\le t-r \\ d(k,j)\le r \\ d(l,k)> t-r-1}} a_{kr,j0} - \sum_{k:\substack{d(i,k)=1 \\ d(k,j)=1}} a_{kt-1,j0}. \tag{B.1}$$

Next, we show that the first and last terms on the right-hand side of (B.1) cancel each other.

Consider two separate cases. First, suppose $d(i,j) > 2$. This implies that $d(l,j) > 1$, in which case, by the induction hypothesis, $a_{lt-1,j0} = 0$. Furthermore, the last term on the right-hand side of (B.1) is also equal to zero, as there exists no agent k such that $d(i,k) = d(k,j) = 1$. Next, suppose $d(i,j) = 2$. In this case, Assumption 1 guarantees that there exists a unique agent k such that $d(i,k) = d(k,j) = 1$. More specifically, if there are two distinct such agents, then there are at least two vertex-independent paths of length 2 from j to i, a configuration that violates Assumption 1. As a result, the last term on the right-hand side of (B.1) is equal to $-a_{lt-1,j0}$, which cancels out the first term.

Taken together, the above argument implies that

$$a_{it,j0} = -\sum_{r=0}^{t-2} \sum_{k:\substack{d(k,j)\le r \\ d(i,k)\le t-r \\ d(l,k)> t-r-1}} a_{kr,j0} = -\sum_{r=0}^{t-2}\Bigg(\sum_{k:\substack{1\le d(k,j)\le r \\ d(i,k)\le t-r \\ d(l,k)> t-r-1}} a_{kr,j0} + \sum_{k:\substack{d(i,j)\le t-r \\ d(l,j)> t-r-1}} a_{jr,j0}\Bigg).$$

The fact that l is on a shortest path from j to i guarantees that $d(i,j) = d(l,j)+1$, which in turn implies the two inequalities in the second term above are mutually inconsistent. Hence, the second term on the right-hand side of the above equality is equal to zero, leading to

$$a_{it,j0} = -\sum_{r=1}^{t-2} \sum_{k:\substack{d(k,j)=1 \\ d(i,k)\le t-r \\ d(l,k)> t-r-1}} a_{kr,j0},$$

where we are using the fact that, by the induction hypothesis, $a_{kr,j0} = 0$ unless $d(k, j) = 1$.

The proof is complete once we show that there is no k for which the restrictions imposed in the sum above are jointly satisfied. Suppose to the contrary that such a k exists. It is immediate that $k \neq l$. Furthermore, since $d(i, k) \leq t - r$ and $d(l, k) > t - r - 1$, the triangle inequality implies l is not on the shortest path from k to i. On the other hand, recall that, by Assumption 1, there cannot be two vertex-independent paths from j to i. Therefore, by Menger's theorem (McCuaig (1984)), there exists a vertex $m \neq i, j$ that lies on all directed paths from j to i. In particular, m lies on the following two paths: (i) the path that connects j to k and then follows a shortest path from k to i; and (ii) the shortest path from j to i, which, recall, by assumption also passes through l. But this implies that the shortest path from m to i and hence the shortest path from k to i also pass through l, which is a contradiction. *Q.E.D.*

Proof of Lemma A.5

We prove this claim by an inductive argument on k. Equation (38) establishes the claim for $k = 1$. As the induction hypothesis, suppose that (41) is satisfied for some integer $k - 1$ and any collection of arbitrary signals $(\omega_2, \dots, \omega_k) \in S^{k-1}$. Taking expectations from both sides of (37) implies that $\mathbb{E}^{\theta}_{-i,t+1}[\zeta_{it+1}(\hat{\theta})] = \sum_{j=1}^{n} a_{ijt+1}\zeta_{jt}(\hat{\theta})\mathbb{E}^{\theta}_{-i,t+1}[\ell^{\hat{\theta}}_j(\omega_{jt+1})/m_{jt}(\omega_{jt+1})]$, where $\mathbb{E}^{\theta}_{-i,t+1}$ denotes the expectation conditional on the σ-field generated by $(\{\omega_{j\tau}\}_{\tau\leq t, j\in N}, \{\omega_{jt+1}\}_{j\neq i})$. Subtracting this equation from (37) leads to

$$a_{iit+1}\zeta_{it}(\hat{\theta})\left(\frac{\ell^{\hat{\theta}}_i(\omega_{it+1})}{m_{it}(\omega_{it+1})} - \mathbb{E}^{\theta}_{-i,t+1}\left[\frac{\ell^{\hat{\theta}}_i(\omega_{it+1})}{m_{it}(\omega_{it+1})}\right]\right) = \zeta_{it+1}(\hat{\theta}) - \mathbb{E}^{\theta}_{-i,t+1}\big[\zeta_{it+1}(\hat{\theta})\big].$$

Pick an arbitrary sequence of signals $(\omega_2, \dots, \omega_k)$ of length $k - 1$, multiply both sides of the above equation by $\prod_{r=2}^{k}\ell^{\hat{\theta}}_i(\omega_r)/\ell^{\theta}_i(\omega_r)$, and sum over all $\hat{\theta} \in \Theta$ to obtain

$$\begin{aligned}
&a_{iit+1}\left(\sum_{\hat{\theta}\in\Theta}\zeta_{it}(\hat{\theta})\frac{\ell^{\hat{\theta}}_i(\omega_{it+1})}{\ell^{\theta}_i(\omega_{it+1})}\prod_{r=2}^{k}\frac{\ell^{\hat{\theta}}_i(\omega_r)}{\ell^{\theta}_i(\omega_r)} - 1\right) \\
&\quad = a_{iit+1}\sum_{\hat{\theta}\in\Theta}\zeta_{it}(\hat{\theta})\left(\frac{\ell^{\hat{\theta}}_i(\omega_{it+1})}{\ell^{\theta}_i(\omega_{it+1})} - \frac{\ell^{\hat{\theta}}_i(\omega_{it+1})}{m_{it}(\omega_{it+1})}\right)\prod_{r=2}^{k}\frac{\ell^{\hat{\theta}}_i(\omega_r)}{\ell^{\theta}_i(\omega_r)} \\
&\qquad - a_{iit+1}\sum_{\hat{\theta}\in\Theta}\zeta_{it}(\hat{\theta})\mathbb{E}^{\theta}_{-i,t+1}\left[1 - \frac{\ell^{\hat{\theta}}_i(\omega_{it+1})}{m_{it}(\omega_{it+1})}\right]\prod_{r=2}^{k}\frac{\ell^{\hat{\theta}}_i(\omega_r)}{\ell^{\theta}_i(\omega_r)} \\
&\qquad + a_{iit+1}\left(\sum_{\hat{\theta}\in\Theta}\zeta_{it}(\hat{\theta})\prod_{r=2}^{k}\frac{\ell^{\hat{\theta}}_i(\omega_r)}{\ell^{\theta}_i(\omega_r)} - 1\right) \\
&\qquad + \left(\sum_{\hat{\theta}\in\Theta}\zeta_{it+1}(\hat{\theta})\prod_{r=2}^{k}\frac{\ell^{\hat{\theta}}_i(\omega_r)}{\ell^{\theta}_i(\omega_r)} - 1\right) - \mathbb{E}^{\theta}_{-i,t+1}\left[\sum_{\hat{\theta}\in\Theta}\zeta_{it+1}(\hat{\theta})\prod_{r=2}^{k}\frac{\ell^{\hat{\theta}}_i(\omega_r)}{\ell^{\theta}_i(\omega_r)} - 1\right].
\end{aligned} \tag{B.2}$$

As our next step, we show that squaring both sides of the above equation, taking conditional expectations, summing both sides from $t = \tau$ to $t = \infty$, and then taking the limit as

$\tau \to \infty$ leads to

$$\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty} a_{iit+1}^2 \mathbb{E}_\tau^\theta\left[\sum_{\hat\theta\in\Theta}\zeta_{it}(\hat\theta)\frac{\ell_i^{\hat\theta}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})}\prod_{r=2}^{k}\frac{\ell_i^{\hat\theta}(\omega_r)}{\ell_i^{\theta}(\omega_r)}-1\right]^2=0 \tag{B.3}$$

$\mathbb{P}^\theta$-almost surely. To establish this, it is sufficient to show that applying the above procedure to each term on the right-hand side of (B.2) separately results in a limit of zero. In addition, given that the second and fifth terms on the right-hand side of (B.2) are simply conditional expectations of the first and fourth terms, respectively, it is sufficient to focus on the first, third, and fourth terms.[1] First consider the third and fourth terms on the right-hand of (B.2). By the induction hypothesis,

$$\begin{aligned}&\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty} a_{iit+1}^2 \mathbb{E}_\tau^\theta\left[\sum_{\hat\theta\in\Theta}\zeta_{it}(\hat\theta)\prod_{r=2}^{k}\frac{\ell_i^{\hat\theta}(\omega_r)}{\ell_i^{\theta}(\omega_r)}-1\right]^2\\&\quad=\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty}\mathbb{E}_\tau^\theta\left[\sum_{\hat\theta\in\Theta}\zeta_{it+1}(\hat\theta)\prod_{r=2}^{k}\frac{\ell_i^{\hat\theta}(\omega_r)}{\ell_i^{\theta}(\omega_r)}-1\right]^2=0\end{aligned} \tag{B.4}$$

$\mathbb{P}^\theta$-almost surely. Next, consider the first term on the right-hand side of (B.2). Note that

$$\frac{\ell_i^{\hat\theta}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})}-\frac{\ell_i^{\hat\theta}(\omega_{it+1})}{m_{it}(\omega_{it+1})}=\frac{\ell_i^{\hat\theta}(\omega_{it+1})}{m_{it}(\omega_{it+1})}\left(\sum_{\tilde\theta\in\Theta}\zeta_{it}(\tilde\theta)\frac{\ell_i^{\tilde\theta}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})}-1\right),$$

where we are using the fact that $m_{it}(\omega_{it+1})=\sum_{\tilde\theta\in\Theta}\zeta_{it}(\tilde\theta)\ell_i^{\tilde\theta}(\omega_{it+1})$. Consequently,

$$\begin{aligned}&\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty} a_{iit+1}^2 \mathbb{E}_\tau^\theta\left[\sum_{\hat\theta\in\Theta}\zeta_{it}(\hat\theta)\left(\frac{\ell_i^{\hat\theta}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})}-\frac{\ell_i^{\hat\theta}(\omega_{it+1})}{m_{it}(\omega_{it+1})}\right)\prod_{r=2}^{k}\frac{\ell_i^{\hat\theta}(\omega_r)}{\ell_i^{\theta}(\omega_r)}\right]^2\\&\quad\le c\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty}\mathbb{E}_\tau^\theta\left[\sum_{\tilde\theta\in\Theta}\zeta_{it}(\tilde\theta)\frac{\ell_i^{\tilde\theta}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})}-1\right]^2\end{aligned}$$

for some positive constant c. By the induction base, the right-hand side of the above inequality is equal to zero $\mathbb{P}^\theta$-almost surely, hence implying that the left-hand side is also equal to zero. This observation in juxtaposition with (B.4) thus establishes (B.3), which in turn implies that

$$\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty}\mathbb{E}_\tau^\theta\mathbb{E}_t^\theta\left[\sum_{\hat\theta\in\Theta}\zeta_{it}(\hat\theta)\frac{\ell_i^{\hat\theta}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})}\prod_{r=2}^{k}\frac{\ell_i^{\hat\theta}(\omega_r)}{\ell_i^{\theta}(\omega_r)}-1\right]^2=0,$$

[1]More specifically, the inequality $\mathbb{E}_{t+1}^2[x_t]\le\mathbb{E}_{t+1}[x_t^2]$ implies that $\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty}\mathbb{E}_\tau[\mathbb{E}_{t+1}^2[x_t]]=0$ whenever $\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty}\mathbb{E}_\tau[x_t^2]=0$.

where we are using the fact that a_{iit+1} is uniformly bounded away from zero. Consequently,

$$\sum_{\omega\in S}\ell_i^\theta(\omega)\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty}\mathbb{E}_\tau^\theta\left[\sum_{\hat\theta\in\Theta}\zeta_{it}(\hat\theta)\frac{\ell_i^{\hat\theta}(\omega)}{\ell_i^\theta(\omega)}\prod_{r=2}^{k}\frac{\ell_i^{\hat\theta}(\omega_r)}{\ell_i^\theta(\omega_r)}-1\right]^2=0$$

$\mathbb{P}^\theta$-almost surely. Finally, the fact that ℓ_i^θ has a full support over S implies that the above equality holds only if $\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty}\mathbb{E}_\tau^\theta[\sum_{\hat\theta\in\Theta}\zeta_{it}(\hat\theta)\frac{\ell_i^{\hat\theta}(\omega)}{\ell_i^\theta(\omega)}\prod_{r=2}^{k}\frac{\ell_i^{\hat\theta}(\omega_r)}{\ell_i^\theta(\omega_r)}-1]^2=0$ for all $\omega\in S$. The fact that the collection of signals $(\omega_2,\dots,\omega_k)$ were arbitrary completes the inductive argument. *Q.E.D.*

Proof of Lemma A.9

To prove the first statement, it is sufficient to show that

$$\sum_{k=1}^{n}\sum_{j=1}^{n}y_ky_jg^{(kj)}(x)\le 0$$

for all $x\in(0,\infty)^n$ and all $y\in\mathbb{R}^n$, where $g^{(kj)}(x)=\partial^2g(x)/\partial x_j\partial x_k$. Since g is homogeneous of degree 1, $g^{(k)}$ is homogeneous of degree zero for all k. Therefore, by Euler's theorem, $\sum_{j=1}^{n}x_jg^{(kj)}(x)=0$ for all $x\in(0,\infty)^n$, and as a result, $g^{(kk)}(x)=-(1/x_k)\sum_{j\neq k}x_jg^{(kj)}(x)$. Hence, for any given vector $y\in\mathbb{R}^n$,

$$\begin{aligned}\sum_{k=1}^{n}\sum_{j=1}^{n}y_ky_jg^{(kj)}(x)&=\sum_{k=1}^{n}\sum_{j\neq k}y_ky_jg^{(kj)}(x)-\sum_{k=1}^{n}\sum_{j\neq k}\frac{x_j}{x_k}y_k^2g^{(kj)}(x)\\&=-\frac{1}{2}\sum_{k=1}^{n}\sum_{j\neq k}\frac{1}{x_kx_j}(y_kx_j-y_jx_k)^2g^{(kj)}(x).\end{aligned}$$

Consequently,

$$\sum_{k=1}^{n}\sum_{j=1}^{n}y_ky_jg^{(kj)}(x)=-\frac{1}{2}\sum_{k=1}^{n}\sum_{j\neq k}\frac{1}{x_kx_j}(y_kx_j-y_jx_k)^2\big(1-\kappa_g^{(kj)}(x)\big)g^{(k)}(x)g^{(j)}(x)/g(x),$$

where $\kappa_g^{(kj)}(x)$ denotes g's logarithmic curvature as defined in equation (12) of the paper. The assumptions that g is increasing and $\kappa_g^{(kj)}(x)\le 1$ imply that the right-hand side of the above equation is nonpositive.

To prove the second statement, note that the concavity of g implies that

$$g(x)\le g(\mathbf{1})+\sum_{j=1}^{n}(x_j-1)g^{(j)}(\mathbf{1})$$

for all $x\in(0,\infty)^n$. Since g is homogeneous of degree 1, Euler's theorem implies that $\sum_{j=1}^{n}g^{(j)}(\mathbf{1})=g(\mathbf{1})$. Hence, $g(x)\le\sum_{j=1}^{n}x_jg^{(j)}(\mathbf{1})$. *Q.E.D.*

Proof of Lemma A.10

The monotonicity and homogeneity of ϕ_i are immediate consequences of the fact that ψ_i is monotonically increasing and homogeneous of degree 1. To prove the last statement, note that the logarithmic curvature of ϕ_i is equal to the negative of the logarithmic curvature of ψ_i, that is,

$$\left(\frac{\partial^2 \log \phi_i(x)}{\partial \log x_k \partial \log x_j}\right) \Big/ \left(\frac{\partial \log \phi_i(x)}{\partial \log x_k}\frac{\partial \log \phi_i(x)}{\partial \log x_j}\right)$$
$$= -\left(\frac{\partial^2 \log \psi_i(x)}{\partial \log x_k \partial \log x_j}\right) \Big/ \left(\frac{\partial \log \psi_i(x)}{\partial \log x_k}\frac{\partial \log \psi_i(x)}{\partial \log x_j}\right).$$

Therefore, the fact that $\kappa_{\psi_i}^{(kj)}(x) \geq -1$ guarantees that the logarithmic curvature of ϕ_i is less than or equal to 1. Finally, an immediate application of Lemma A.9 guarantees that ϕ_i is concave. *Q.E.D.*

Proof of Lemma A.14

We prove the lemma by relying on an inductive argument on k. Equation (53) of the paper implies that the lemma's statement holds for $k = 1$, establishing the induction base. Next, suppose that the lemma is satisfied for $k - 1$. To establish the result for k, recall from equation (50) that as $t \to \infty$,

$$\psi_i\big(\mu_{t+1}(\hat{\theta})\big) - \sum_{j=1}^{n} \psi_i^{(j)}(\mathbf{1})\mu_{jt+1}(\hat{\theta}) \to 0. \tag{B.5}$$

On the other hand, recall from equation (51) that

$$\mu_{it+1}(\hat{\theta}) = \psi_i\big(\mu_t(\hat{\theta})\big) + \left(\frac{\ell_i^{\hat{\theta}}(\omega_{it+1})}{m_{it}(\omega_{it+1})} - 1\right)\psi_i\big(\mu_t(\hat{\theta})\big). \tag{B.6}$$

Using (B.6) to substitute for $\mu_{jt+1}(\hat{\theta})$ in (B.5) implies that

$$\psi_i\big(\mu_{t+1}(\hat{\theta})\big) - \sum_{j=1}^{n} \psi_i^{(j)}(\mathbf{1})\psi_j\big(\mu_t(\hat{\theta})\big)\frac{\ell_j^{\hat{\theta}}(\omega_{jt+1})}{m_{jt}(\omega_{jt+1})} \to 0 \tag{B.7}$$

$\mathbb{P}^\theta$-almost surely for all $\hat{\theta} \in \Theta$. By the dominated convergence theorem for conditional expectation,

$$\mathbb{E}_{-i,t+1}^{\theta}\big[\psi_i\big(\mu_{t+1}(\hat{\theta})\big)\big] - \psi_i^{(i)}(\mathbf{1})\psi_i\big(\mu_t(\hat{\theta})\big)\mathbb{E}_{-i,t+1}^{\theta}\left[\frac{\ell_i^{\hat{\theta}}(\omega_{it+1})}{m_{it}(\omega_{it+1})}\right]$$
$$- \sum_{j\neq i}^{n} \psi_i^{(j)}(\mathbf{1})\psi_j\big(\mu_t(\hat{\theta})\big)\frac{\ell_j^{\hat{\theta}}(\omega_{jt+1})}{m_{jt}(\omega_{jt+1})} \to 0$$

$\mathbb{P}^\theta$-almost surely, where $\mathbb{E}_{-i,t+1}^{\theta}$ denotes the expectation operator conditional on the σ-field generated by $(\{\omega_{j\tau}\}_{1\leq\tau\leq t, j\in N}, \{\omega_{jt+1}\}_{j\neq i})$. Note that the induction base, alongside the dominated convergence theorem, guarantees that $\mathbb{E}^{\theta}_{-i,t+1}[\ell_i^{\hat{\theta}}(\omega_{it+1})/m_{it}(\omega_{it+1})] \to 1$ as $t \to \infty$, which means that

$$\mathbb{E}^{\theta}_{-i,t+1}\big[\psi_i\big(\mu_{t+1}(\hat{\theta})\big)\big] - \psi_i^{(i)}(\mathbf{1})\psi_i\big(\mu_t(\hat{\theta})\big) - \sum_{j\neq i}^{n} \psi_i^{(j)}(\mathbf{1})\psi_j\big(\mu_t(\hat{\theta})\big)\frac{\ell_j^{\hat{\theta}}(\omega_{jt+1})}{m_{jt}(\omega_{jt+1})} \to 0.$$

Subtracting (B.7) from the above equation results in

$$\mathbb{E}^{\theta}_{-i,t+1}\big[\psi_i\big(\mu_{t+1}(\hat{\theta})\big)\big] - \psi_i\big(\mu_{t+1}(\hat{\theta})\big) + \psi_i^{(i)}(\mathbf{1})\psi_i\big(\mu_t(\hat{\theta})\big)\left(\frac{\ell_i^{\hat{\theta}}(\omega_{it+1})}{m_{it}(\omega_{it+1})} - 1\right) \to 0$$

$\mathbb{P}^{\theta}$-almost surely for all $\hat{\theta} \in \Theta$.

Pick an arbitrary sequence of signals $(\omega_2, \dots, \omega_k)$ of length $k-1$. Multiplying both sides of the above equation by $\prod_{r=2}^{k} \ell_i^{\hat{\theta}}(\omega_r)/\ell_i^{\theta}(\omega_r)$ and summing over all $\hat{\theta} \in \Theta$ leads to

$$\mathbb{E}^{\theta}_{-i,t+1}\left[\sum_{\hat{\theta}\in\Theta} \psi_i\big(\mu_{t+1}(\hat{\theta})\big) \prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)}\right] - \sum_{\hat{\theta}\in\Theta} \psi_i\big(\mu_{t+1}(\hat{\theta})\big) \prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)}$$
$$+ \psi_i^{(i)}(\mathbf{1}) \sum_{\hat{\theta}\in\Theta} \left(\prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)}\right)\left(\frac{\ell_i^{\hat{\theta}}(\omega_{it+1})}{m_{it}(\omega_{it+1})} - 1\right)\psi_i\big(\mu_t(\hat{\theta})\big) \to 0.$$

By the induction hypothesis, the first two terms on the left-hand side of the above expression converge to 1 almost surely as $t \to \infty$. Thus, for any arbitrary collection of signals $(\omega_2, \dots, \omega_k) \in S^{k-1}$,

$$\sum_{\hat{\theta}\in\Theta} \psi_i\big(\mu_t(\hat{\theta})\big)\frac{\ell_i^{\hat{\theta}}(\omega_{it+1})}{m_{it}(\omega_{it+1})} \prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)} - \sum_{\hat{\theta}\in\Theta} \psi_i\big(\mu_t(\hat{\theta})\big) \prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)} \to 0,$$

with $\mathbb{P}^{\theta}$-probability 1 as $t \to \infty$, where we are using the observation that monotonicity requires that $\psi_i^{(i)}(\mathbf{1}) > 0$. The induction hypothesis once again implies that the last term on the left-hand side above converges to 1 asymptotically. Furthermore, recall that $\ell_i^{\theta}(\omega_{it+1})/m_{it}(\omega_{it+1}) \to 1$ almost surely. Therefore, it is immediate that

$$\sum_{\hat{\theta}\in\Theta} \psi_i\big(\mu_t(\hat{\theta})\big)\frac{\ell_i^{\hat{\theta}}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})} \prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)} - 1 \to 0$$

$\mathbb{P}^{\theta}$-almost surely. Hence, by the dominated convergence theorem,

$$\mathbb{E}^{\theta}_{t}\left|\sum_{\hat{\theta}\in\Theta} \psi_i\big(\mu_t(\hat{\theta})\big)\frac{\ell_i^{\hat{\theta}}(\omega_{it+1})}{\ell_i^{\theta}(\omega_{it+1})} \prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)} - 1\right| \to 0.$$

Expressing the conditional expectation as a sum over all possible realizations of ω_{it+1}, we have

$$\sum_{\omega\in S} \ell_i^{\theta}(\omega)\left|\sum_{\hat{\theta}\in\Theta} \psi_i\big(\mu_t(\hat{\theta})\big)\frac{\ell_i^{\hat{\theta}}(\omega)}{\ell_i^{\theta}(\omega)} \prod_{r=2}^{k} \frac{\ell_i^{\hat{\theta}}(\omega_r)}{\ell_i^{\theta}(\omega_r)} - 1\right| \to 0$$

$\mathbb{P}^\theta$-almost surely. Since ℓ_i^θ has full support over S, it is immediate that the expression in the absolute values above has to converge to zero almost surely for all possible signals $\omega \in S$. This observation, alongside the fact that the sequence $(\omega_2, \dots, \omega_k)$ was arbitrary, establishes the result. *Q.E.D.*

Proof of Lemma A.16

The first claim is a trivial consequence of the fact that $b_{t,\tau} \leq 1$.

To prove the second claim, note that, by the Weierstrass product inequality (Steele (2004, p. 190)), $b_{t,\tau} \leq \sum_{r=\tau}^{t-1} r^{-\alpha}$. Bounding the sum on the right-hand side of this inequality by an integral implies that $b_{t,\tau} \leq \int_{\tau-1}^{t-1} z^{-\alpha}\,dz$ for $\tau \geq 2$. Therefore,

$$\sum_{\tau=2}^{t-1} b_{t,\tau} \leq \frac{1}{\alpha-1}\sum_{\tau=2}^{t-1}(\tau-1)^{1-\alpha} - \frac{1}{\alpha-1}(t-2)(t-1)^{1-\alpha}.$$

Multiplying both sides by $t^{\alpha-2}$, upper bounding the sum with an integral, and taking limits leads to the following upper bound on the object of interest:

$$\lim_{t\to\infty} t^{\alpha-2}\sum_{\tau=1}^{t-1} b_{t,\tau} \leq \frac{1}{\alpha-1}\lim_{t\to\infty} t^{\alpha-2}\int_1^{t-2} z^{1-\alpha}\,dz - \frac{1}{\alpha-1} = \frac{1}{2-\alpha}, \tag{B.8}$$

where we are using the fact that $b_{t,1} = 1$. We next prove that this upper bound is tight by showing that $1/(2-\alpha)$ is also a lower bound for the expression on the left-hand side of (B.8). To this end, note that $\prod_{r=\tau}^{t-1}(1-r^{-\alpha}) \leq \exp(-\sum_{r=\tau}^{t-1} r^{-\alpha})$. This is due to the fact that $1-z \leq \exp(-z)$ for all z. Therefore, the observation that $1-\exp(-z) \geq z - z^2/2$ for all $z \geq 0$ implies that

$$b_{t,\tau} = 1 - \prod_{r=\tau}^{t-1}\big(1-r^{-\alpha}\big) \geq \sum_{r=\tau}^{t-1} r^{-\alpha} - \frac{1}{2}\left(\sum_{r=\tau}^{t-1} r^{-\alpha}\right)^2.$$

Bounding the sums on the right-hand side of the above inequality with integrals and summing over τ, we obtain

$$\sum_{\tau=2}^{t-1} b_{t,\tau} \geq \frac{1}{\alpha-1}\sum_{\tau=2}^{t-1}\tau^{1-\alpha} - \frac{1}{\alpha-1}(t-2)t^{1-\alpha} - \frac{1}{2(\alpha-1)^2}\sum_{\tau=2}^{t-1}\big((\tau-1)^{1-\alpha} - (t-1)^{1-\alpha}\big)^2.$$

Consequently, $\lim_{t\to\infty} t^{\alpha-2}\sum_{\tau=1}^{t-1} b_{t,\tau} \geq 1/(2-\alpha)$. The juxtaposition of this inequality with (B.8) completes the proof. *Q.E.D.*

Proof of Parts (b) and (c) of Theorems 7

We prove parts (b) and (c) of Theorem 7 by constructing a social network, a signal structure, and social learning rules with logarithmic curvatures outside the $[-1, 1]$ interval for which agents fail to learn the state asymptotically. Suppose agents interact over the complete social network depicted in the left panel of Figure 1 and rely on a common (weakly-separable) CES learning rule as in (13). Furthermore, suppose that there are only two states, labeled θ and $\hat{\theta}$, and two signals, also labeled θ and $\hat{\theta}$. Finally, suppose agent j receives the signal that matches the state with probability $p > 1/2$ (and the other signal with the complementary probability), whereas all other agents' signals are uninformative.

Let $\zeta_{it} = f(\mu_t)$ be the interim belief of agent i after observing her neighbor's time-t reports but before observing her private signal ω_{it+1}. Since agents $i \neq j$ observe no informative signals, $\mu_{it+1} = \zeta_{it}$ for all $i \neq j$ and all t. Therefore,

$$\frac{\zeta_{it+1}(\theta)}{\zeta_{it+1}(\hat{\theta})} = \frac{\psi\left(\zeta_{1t}(\theta), \dots, \dfrac{\zeta_{jt}(\theta)\ell_j^{\theta}(\omega_{jt+1})}{m_{jt}(\omega_{jt+1})}, \dots, \zeta_{nt}(\theta)\right)}{\psi\left(\zeta_{1t}(\hat{\theta}), \dots, \dfrac{\zeta_{jt}(\hat{\theta})\ell_j^{\hat{\theta}}(\omega_{jt+1})}{m_{jt}(\omega_{jt+1})}, \dots, \zeta_{nt}(\hat{\theta})\right)},$$

where $m_{jt}(\omega_{jt+1}) = \zeta_{jt}(\theta)\ell_j^{\theta}(\omega_{jt+1}) + \zeta_{jt}(\hat{\theta})\ell_j^{\hat{\theta}}(\omega_{jt+1})$. Note that since agents use identical social learning rules, their interim beliefs coincide at all times, that is, $\zeta_{it} = \zeta_{kt}$ for all i, k and all t. Therefore, we can divide both sides of the above expression by $\zeta_{it}(\theta)/\zeta_{it}(\hat{\theta})$ and use the assumption that ψ is homogeneous of degree 1 to obtain

$$\log\frac{\zeta_{it+1}(\theta)}{\zeta_{it+1}(\hat{\theta})} = \varphi\left(\log\frac{\ell_j^{\theta}(\omega_{jt+1})}{\ell_j^{\hat{\theta}}(\omega_{jt+1})}; \zeta_{it}(\theta)\right) + \log\frac{\zeta_{it}(\theta)}{\zeta_{it}(\hat{\theta})}, \tag{B.9}$$

where $\varphi(\log\lambda; x) = \log\psi(1, \dots, \frac{\lambda}{1+x(\lambda-1)}, \dots, 1) - \log\psi(1, \dots, \frac{1}{1+x(\lambda-1)}, \dots, 1)$. Note that, given the functional form assumption (13) on ψ, we have

$$\begin{aligned}\varphi(\log\lambda; x) = {} & \frac{1}{\xi}\log\left(1 - a_{ij} + \frac{a_{ij}\lambda^{\xi}}{\big(1 + x(\lambda-1)\big)^{\xi}}\right) \\ & - \frac{1}{\xi}\log\left(1 - a_{ij} + \frac{a_{ij}}{\big(1 + x(\lambda-1)\big)^{\xi}}\right).\end{aligned} \tag{B.10}$$

We have the following lemma.

LEMMA B.1: *Let $\varphi(\log\lambda; x)$ be defined as in* (B.10).

(a) *If $\xi < -1$ and $a_{ij} < 1 + 1/\xi$, then $\mathbb{E}^{\theta}\varphi(\log\lambda(\omega_{jt}); x) < 0$ in a neighborhood of $x = 0$ and $p = 1/2$.*

(b) *If $\xi > 1$ and $a_{ij} < 1 - 1/\xi$, then $\mathbb{E}^{\theta}\varphi(\log\lambda(\omega_{jt}); x) < 0$ in a neighborhood of $x = 1$ and $p = 1/2$.*

(c) *If $\xi < 0$, then $\varphi(\log\lambda; x)$ is increasing in x, whereas if $\xi > 0$, then $\varphi(\log\lambda; x)$ is decreasing in x.*

PROOF: To prove part (a), suppose $\xi < -1$ and $a_{ij} < 1 + 1/\xi$. It is easy to verify that when $p = 1/2$,

$$\mathbb{E}^{\theta}\varphi\big(\log\lambda(\omega_{jt}); 0\big) = \frac{d}{dp}\mathbb{E}^{\theta}\varphi\big(\log\lambda(\omega_{jt}); 0\big) = 0,$$

$$\frac{d^2}{dp^2}\mathbb{E}^{\theta}\varphi\big(\log\lambda(\omega_{jt}); 0\big) = 16a_{ij}\big(1 + (1 - a_{ij})\xi\big) < 0.$$

The statement then follows from the smoothness of $\mathbb{E}^\theta\varphi(\log\lambda(\omega_{jt}); x)$ in x and p.

To prove part (b), suppose that $\xi > 1$ and $a_{ij} < 1 - 1/\xi$. Once again, it is straightforward to verify that when $p = 1/2$,

$$\mathbb{E}^\theta\varphi\big(\log\lambda(\omega_{jt}); 1\big) = \frac{d}{dp}\mathbb{E}^\theta\varphi\big(\log\lambda(\omega_{jt}); 1\big) = 0,$$

$$\frac{d^2}{dp^2}\mathbb{E}^\theta\varphi\big(\log\lambda(\omega_{jt}); 1\big) = 16a_{ij}\big(1-(1-a_{ij})\xi\big) < 0.$$

Noting that $\mathbb{E}^\theta\varphi(\log\lambda(\omega_{jt}); x)$ is smooth in x and p completes the proof.

To prove part (c), we differentiate $\varphi(\log\lambda; x)$ with respect to x to get

$$\begin{aligned}&\frac{d}{dx}\varphi(\log\lambda; x)\\ &\quad= \frac{a_{ij}(1-a_{ij})\big(1+x(\lambda-1)\big)^{\xi}(1-\lambda)\big(\lambda^\xi-1\big)}{\big(1+x(\lambda-1)\big)\big(a_{ij}\lambda^\xi+(1-a_{ij})\big(1+x(\lambda-1)\big)^{\xi}\big)\big(a_{ij}+(1-a_{ij})\big(1+x(\lambda-1)^\xi\big)\big)}.\end{aligned}$$

The sign of the right-hand side of the expression coincides with that of $(1-\lambda)(\lambda^\xi-1)$. Therefore, $d\varphi(\log\lambda; x)/dx \geq 0$ whenever $\xi < 0$, whereas $d\varphi(\log\lambda; x)/dx \leq 0$ whenever $\xi > 0$. *Q.E.D.*

Proof of Theorem 7(b)

Let θ denote the underlying state of the world and suppose that $\xi < -1$. By Lemma B.1, there exists a triple $(\underline{a}_{ij}, \underline{x}, \underline{p})$ such that if $a_{ij} \in (0, \underline{a}_{ij})$ and $p \in (1/2, \underline{p})$, then $\mathbb{E}^\theta\varphi(\log\lambda(\omega_{jt}); x) < 0$ for all $x \leq \underline{x}$. In the rest of the proof, fix such a_{ij} and p.

Since $\mu_{it+1} = \mathrm{BU}(\zeta_{it}; \omega_{it+1})$, it is sufficient to show that $\zeta_{it}(\theta)$ converges to 0 with $\mathbb{P}^\theta$-positive probability. Equation (B.9) implies that for all $t \geq \tau$,

$$\log\frac{\zeta_{it}(\theta)}{\zeta_{it}(\hat\theta)} = \sum_{r=\tau+1}^{t}\varphi\big(\log\lambda(\omega_{jr}); \zeta_{ir-1}(\theta)\big) + \log\frac{\zeta_{i\tau}(\theta)}{\zeta_{i\tau}(\hat\theta)}. \tag{B.11}$$

We start by assuming that $\zeta_{i\tau}(\theta) \leq \underline{x}$ for some deterministic τ and a sequence of signals $(\omega_{j1}, \dots, \omega_{j\tau})$ that is realized with $\mathbb{P}^\theta$-positive probability—a claim we prove below. Fix such a τ and $(\omega_{j1}, \dots, \omega_{j\tau})$ and define $\{y_t\}_{t\geq\tau}$ recursively by setting $y_\tau = 0$ and $y_t = \max\{y_{t-1} + \varphi(\log\lambda(\omega_{jt}); \underline{x}), 0\}$ for $t > \tau$. Let $T = \inf\{t > \tau : y_t = 0\}$. Lemma B.1 and equation (B.11) imply that

$$\log\frac{\zeta_{it}(\theta)}{\zeta_{it}(\hat\theta)} \leq y_t + \log\frac{\zeta_{i\tau}(\theta)}{\zeta_{i\tau}(\hat\theta)} \tag{B.12}$$

for all $t \in [\tau, T)$. On the other hand, $\{y_t\}_{t\geq\tau}$ is a random walk with negative expected drift stopped at zero. Thus, it converges to $-\infty$ and, hence, $T = \infty$ with $\mathbb{P}^\theta$-positive probability. But both $\{y_t\}_{t\geq\tau}$ and T are measurable with respect to the σ-field generated by $(\omega_{j\tau+1}, \omega_{j\tau+2}, \dots)$. Therefore, since the signals observed by agent j in periods $t = 1, \dots, \tau$ are independent of the signals observed by her after period τ, (B.12) implies that $\log(\zeta_{it}(\theta)/\zeta_{it}(\hat\theta)) \to -\infty$ and thus $\zeta_{it}(\theta) \to 0$ with $\mathbb{P}^\theta$-positive probability.

The proof of this part is complete once we show that there exist some deterministic τ and a signal sequence $(\omega_{j1}, \dots, \omega_{j\tau})$ that is realized with positive probability such that $\zeta_{i\tau}(\theta) \le \underline{x}$ given such a realization of signals. But this is a simple consequence of the fact that, by Lemma B.1, $\varphi(\log \frac{1-p}{p}; x) < \varphi(\log \frac{1-p}{p}; 1) < 0$ for all x. Thus, whenever agent j observes signal $\omega_{jr} = \hat{\theta}$ in periods $r = 1, \dots, \tau$ for a sufficiently large τ, by (B.11), the interim beliefs satisfy $\zeta_{i\tau}(\theta) \le \underline{x}$ for all i. *Q.E.D.*

Proof of Theorem 7(c)

Let θ denote the underlying state and suppose that $\xi > 1$. By Lemma B.1, there exists a triple $(\underline{a}_{ij}, \overline{x}, \underline{p})$ such that if $a_{ij} \in (0, \underline{a}_{ij})$ and $p \in (1/2, \underline{p})$, then $\mathbb{E}^\theta \varphi(\log \lambda(\omega_{jt}); x) < 0$ for all $x \ge \overline{x}$. Fix such a_{ij} and p, and let $\underline{T}_t = \{\tau \le t : \zeta_{i\tau}(\theta) < \overline{x}\}$ and $\overline{T}_t = \{\tau \le t : \zeta_{i\tau}(\theta) \ge \overline{x}\}$. Evaluating (B.11) for $\tau = 0$ and using Lemma B.1 and the uniform prior assumption, we obtain

$$\begin{aligned} \log \frac{\zeta_{it}(\theta)}{\zeta_{it}(\hat{\theta})} \le & \sum_{\tau \in \underline{T}_t} \bigl[\varphi\bigl(\log \lambda(\omega_{j\tau}); \zeta_{i\tau-1}(\theta)\bigr) - \varphi\bigl(\log \lambda(\omega_{j\tau}); \overline{x}\bigr)\bigr] \\ & + \sum_{\tau=1}^{t} \varphi\bigl(\log \lambda(\omega_{j\tau}); \overline{x}\bigr). \end{aligned} \tag{B.13}$$

By the strong law of large numbers, as $t \to \infty$,

$$\frac{1}{t} \sum_{\tau=1}^{t} \varphi\bigl(\log \lambda(\omega_{j\tau}); \overline{x}\bigr) \to \mathbb{E}^\theta \varphi\bigl(\log \lambda(\omega_{j\tau}); \overline{x}\bigr) < 0 \tag{B.14}$$

$\mathbb{P}^\theta$-almost surely. In what follows, we show that $\zeta_{it}(\theta) \to 1$ and (B.14) cannot be satisfied simultaneously. This observation proves that $\zeta_{it}(\theta) \to 1$ with zero probability, thus completing the proof.

Suppose to the contrary that $\zeta_{it}(\theta) \to 1$ and (B.14) is satisfied. That $\zeta_{it}(\theta) \to 1$ implies that the left-hand side of (B.13) converges to $+\infty$. It also guarantees that the set $\underline{T}_t$ remains finite as $t \to \infty$, thus ensuring that the first two terms on the right-hand side of (B.13) remain finite as $t \to \infty$. But (B.14) implies that the last term on the right-hand side of (B.13) goes to $-\infty$, leading to a contradiction. *Q.E.D.*

APPENDIX C: UNANIMITY IN THE LIMIT

Theorem 4 in the main body of the paper establishes that log-linear learning leads to a complete aggregation of information as long as (i) agents rely on unanimous social learning rules (in the sense of Definition 1) and (ii) the rate of decay of the weights they assign to their neighbors' beliefs is slower than $1/t$. This appendix extends the concept of unanimous learning rules in Definition 1, followed by a generalization of Theorem 4.

DEFINITION C.1: Agent i's sequence of social learning rules $f_{it} : \Delta\Theta^n \to \Delta\Theta$ satisfying imperfect recall is *unanimous in the limit* if there exist $\rho_{it} > 0$ such that $f_{it}(\mu, \dots, \mu) = \rho_{it}\mu$ for all $\mu \in \Delta\Theta$ and all t and

$$\lim_{t\to\infty} \sum_{\tau=1}^{t} \prod_{r=\tau}^{t-1} \rho_{ir} = \infty, \tag{C.1}$$

$$\frac{\prod_{r=\tau}^{t-1}\rho_{ir}}{\sum_{\tau=1}^{t}\prod_{r=\tau}^{t-1}\rho_{ir}} \leq C_i t^{-\alpha_i} \quad \text{for all } \tau \leq t \tag{C.2}$$

for some $C_i, \alpha_i > 0$.

It is immediate to verify that Definition C.1 is a relaxation of Definition 1 in the paper: if agent i's learning rules are unanimous in the limit with $\rho_{it} = 1$ for all t, then her learning rules are unanimous at all times. More generally, conditions (C.1) and (C.2) guarantee that $\lim_{t\to\infty}\rho_{it} = 1$, whenever the limit exists—hence the name unanimity in the limit—though the converse is not necessarily true. Intuitively, (C.1) ensures that agents assign non-vanishing weights to the signals they observe early on, whereas (C.2) guarantees that no piece of information receives an outsized weight.

Our next result generalizes Theorem 4 to the class of log-linear learning rules (3) that are unanimous in the limit with $\rho_{it} = \rho_t$. As in the main body of the paper, we discipline the rate of decay of weights a_{ijt} that each agent i assigns to her neighbors by assuming that there exist a sequence $\lambda_t \in (0, 1)$ and constants $\underline{a}, \overline{a} \in (0, 1)$ such that $a_{ijt} \geq \lambda_t\rho_t\underline{a}$ and $\sum_{k\neq i} a_{ikt} \leq \lambda_t\rho_t\overline{a}$ for all t and all pairs of agents $i \neq j$ such that $j \in N_i$

THEOREM C.1: *Suppose agents follow the log-linear learning rule* (3) *with weights that decay at rate* λ_t. *If learning rules are unanimous in the limit with* $\rho_{it} = \rho_t$ *for all* i *and* $\lim_{t\to\infty} t^\alpha\lambda_t = \infty$ *for* α *in equation* (C.2), *then all agents learn the state almost surely.*

PROOF: The proof mirrors the proof of Theorem 4. Let θ denote the underlying state of the world and $A_t = [a_{ijt}]$ denote the matrix of weights that agents assign to their neighbors' beliefs at time t, with the convention that $a_{ijt} = 0$ if $j \notin N_i$. By assumption, $\sum_{j=1}^{n} a_{ijt} = \rho_t$. Hence, $A_t = \rho_t\tilde{A}_t$, where $\tilde{A}_t$ is a stochastic matrix such that $\tilde{a}_{ijt} \geq \lambda_t\underline{a}$ and $\sum_{k\neq i}\tilde{a}_{ikt} \leq \lambda_t\overline{a}$ for all t and all pairs of agents $i \neq j$ such that $j \in N_i$.

Given any state $\hat{\theta} \neq \theta$, equation (3) implies that

$$x_{t+1} = \rho_t\tilde{A}_t x_t + y_{t+1}(\omega_{t+1}),$$

where $x_{it} = \log(\mu_{it}(\theta)/\mu_{it}(\hat{\theta}))$ and $y_{it}(\omega_{it}) = \log(\ell_i^\theta(\omega_{it})/\ell_i^{\hat{\theta}}(\omega_{it}))$. Consequently,

$$x_t = y_t(\omega_t) + \sum_{\tau=1}^{t-1}\rho_{t-1}\ldots\rho_\tau\tilde{A}_{t-1}\ldots\tilde{A}_\tau y_\tau(\omega_\tau). \tag{C.3}$$

By Lemma A.1, there exists a sequence of uniformly lower-bounded probability vectors v_τ that jointly satisfy $v'_{t+1}\tilde{A}_t\ldots\tilde{A}_\tau = v'_\tau$ for all $t \geq \tau$. Therefore, pre-multiplying both sides of (C.3) by v'_t implies that

$$v'_t x_t = \sum_{\tau=1}^{t}\rho_{t,\tau}v'_\tau y_\tau(\omega_\tau),$$

where $\rho_{t,\tau} = \prod_{r=\tau}^{t-1} \rho_r$. Consequently,

$$\liminf_{t\to\infty} \frac{v_t' x_t}{\sum_{\tau=1}^{t} \rho_{t,\tau}} = \lim_{t\to\infty} \frac{\sum_{\tau=1}^{t} \rho_{t,\tau} v_\tau' \big(y_\tau(\omega_\tau) - h(\theta,\hat{\theta})\big)}{\sum_{\tau=1}^{t} \rho_{t,\tau}} + \liminf_{t\to\infty} \frac{\sum_{\tau=1}^{t} \rho_{t,\tau} v_\tau' h(\theta,\hat{\theta})}{\sum_{\tau=1}^{t} \rho_{t,\tau}}, \tag{C.4}$$

where $h_i(\theta,\hat{\theta}) = \mathbb{E}^\theta[y_{it}(\omega_{it})]$. Recall that agents' private signals are independently and identically distributed over time. Furthermore, condition (C.2) requires that $(\prod_{r=\tau}^{t-1} \rho_r)/(\sum_{\tau=1}^{t} \prod_{r=\tau}^{t-1} \rho_r) \le C t^{-\alpha}$ for all $\tau \le t$. Therefore, by Theorem 2 of Pruitt (1966), the first term on the right-hand side of (C.4) is equal to zero almost surely. On the other hand, Lemma A.1 guarantees that $\liminf_{t\to\infty} v_{it} > 0$ for all i, while the assumption that agents do not face a collective identification problem guarantees that there exists an agent i such that $h_i(\theta,\hat{\theta}) > 0$. Hence,

$$\liminf_{t\to\infty} \frac{v_t' x_t}{\sum_{\tau=1}^{t} \rho_{t,\tau}} > 0 \tag{C.5}$$

with probability 1.

With the above inequality in hand, it is sufficient to establish that, for any pair of agents i and j,

$$\lim_{t\to\infty} \frac{1}{\sum_{\tau=1}^{t} \rho_{t,\tau}} (x_{it} - x_{jt}) = 0 \tag{C.6}$$

almost surely. In particular, (C.5) and (C.6), together with the fact that v_t is a probability vector, imply that $\liminf_{t\to\infty} x_{it}/(\sum_{\tau=1}^{t} \rho_{t,\tau}) > 0$ almost surely for all agents i. Therefore, (C.1) implies that $\lim_{t\to\infty} x_{it} = \infty$ with probability 1, which subsequently guarantees that $\mu_{it}(\hat{\theta}) \to 0$ almost surely for all $\hat{\theta} \neq \theta$. In other words, all agents learn the underlying state with probability 1.

To establish (C.6), recall from equation (C.3) that $x_t = y_t(\omega_t) + \sum_{\tau=1}^{t-1} \rho_{t,\tau} \tilde{A}_{t-1} \ldots \tilde{A}_\tau y_\tau(\omega_\tau)$. Thus, by part (d) of Lemma A.2,

$$\begin{aligned} \max_i x_{it} - \min_i x_{it} \le{}& \max_i y_{it}(\omega_{it}) - \min_i y_{it}(\omega_{it}) \\ &+ \sum_{\tau=1}^{t-1} \rho_{t,\tau} \pi(\tilde{A}_{t-1} \ldots \tilde{A}_\tau) \Big(\max_i y_{i\tau}(\omega_{i\tau}) - \min_i y_{i\tau}(\omega_{i\tau}) \Big). \end{aligned}$$

On the other hand, as in the proof of Lemma A.1, we have $\tilde{A}_t = \mathbb{E}^*[\Lambda_t B_t]$, where Λ_t is a sequence of independent Bernoulli random variables that take value 1 with probability λ_t and B_t is a stochastic matrix whose nonzero elements are uniformly lower bounded by a constant $\eta \in (0,1)$ that is independent of t. Therefore,

$$\begin{aligned}\max_i x_{it} - \min_i x_{it} &\leq \sum_{\tau=1}^{t} \pi\Bigg(\mathbb{E}^* \prod_{\substack{r:\, \Lambda_r=1\\ \tau\leq r<t}} B_r\Bigg)\rho_{t,\tau}\Big(\max_i y_{i\tau}(\omega_{i\tau}) - \min_i y_{i\tau}(\omega_{i\tau})\Big)\\ &\leq \sum_{\tau=1}^{t} \rho_{t,\tau}\mathbb{E}^*\Bigg[\pi\Bigg(\prod_{\substack{r:\, \Lambda_r=1\\ \tau\leq r<t}} B_r\Bigg)\Bigg]\Big(\max_i y_{i\tau}(\omega_{i\tau}) - \min_i y_{i\tau}(\omega_{i\tau})\Big),\end{aligned}$$

where the expectation $\mathbb{E}^*$ is over the collection of random variables Λ_t and we are using the convexity of π, established in Lemma A.2. Since the set of signals S is finite, there exists a constant $c \geq 0$, independent of t, such that

$$\max_i x_{it} - \min_i x_{it} \leq c\sum_{\tau=1}^{t} \rho_{t,\tau}\mathbb{E}^*\Bigg[\pi\Bigg(\prod_{\substack{r:\, \Lambda_r=1\\ \tau\leq r<t}} B_r\Bigg)\Bigg].$$

All matrices in the matrix sequence B_t are irreducible, with nonzero elements that are uniformly lower bounded by η for all t. Therefore, any product of length n of these matrices is element-wise strictly positive, with elements that are lower bounded by η^n. Dividing the matrix product $\prod_{r:\, \substack{\tau\leq r<t\\ \Lambda_r=1}} B_r$ into groups of length n and using parts (b) and (c) of Lemma A.2 therefore implies that

$$\pi\Bigg(\prod_{\substack{r:\, \Lambda_r=1\\ \tau\leq r<t}} B_r\Bigg) \leq \big(1-\eta^n\big)^{\lfloor(\Lambda_\tau+\cdots+\Lambda_{t-1})/n\rfloor},$$

where $\lfloor z\rfloor$ denotes the integer part of z. Consequently,

$$\max_i x_{it} - \min_i x_{it} \leq \frac{c}{\beta^n}\sum_{\tau=1}^{t} \rho_{t,\tau}\mathbb{E}^*\big[\beta^{(\Lambda_\tau+\cdots+\Lambda_{t-1})}\big],$$

where $\beta = (1-\eta^n)^{1/n} < 1$. Since random variables Λ_t are independent, we have

$$\limsup_{t\to\infty} \frac{1}{\sum_{\tau=1}^{t}\rho_{t,\tau}}\Big(\max_i x_{it} - \min_i x_{it}\Big) \leq \frac{c}{\beta^n}\limsup_{t\to\infty}\frac{1}{\sum_{\tau=1}^{t}\rho_{t,\tau}}\sum_{\tau=1}^{t}\rho_{t,\tau}\big(1-(1-\beta)\underline{\lambda}_t\big)^{t-\tau},$$

where $\underline{\lambda}_t = \min_{1\leq r<t}\lambda_r$. Equation (C.2) then guarantees that

$$\limsup_{t\to\infty} \frac{1}{\sum_{\tau=1}^{t}\rho_{t,\tau}}\Big(\max_i x_{it} - \min_i x_{it}\Big) \leq \frac{cC}{\beta^n}\limsup_{t\to\infty} t^{-\alpha}\sum_{\tau=1}^{t}\big(1-(1-\beta)\underline{\lambda}_t\big)^{t-\tau},$$

and as a result,

$$\limsup_{t\to\infty}\frac{1}{t}\Big(\max_i x_{it}-\min_i x_{it}\Big)\le\frac{cC}{(1-\beta)\beta^n}\limsup_{t\to\infty}\frac{1}{t^{\alpha}\underline{\lambda}_t}.$$

The assumption that $\lim_{t\to\infty}t^{\alpha}\lambda_t=\infty$ guarantees that the right-hand side of the above inequality is equal to zero, hence establishing (C.6). *Q.E.D.*

APPENDIX D: BAYESIAN LEARNING WITH NORMAL SIGNALS

In this appendix, we show that if signals are normally distributed and agents' prior beliefs are normal, then Bayesian updating takes a log-linear form regardless of the structure of the underlying social network, that is,

$$\log\frac{\mu_{it}(\theta)}{\mu_{it}(\hat{\theta})}=\log\frac{\ell_{it}^{\theta}(\omega_{it})}{\ell_{it}^{\hat{\theta}}(\omega_{it})}+\sum_{\tau=1}^{t-1}\sum_{j\in N_i}a_{it,j\tau}\log\frac{\mu_{j\tau}(\theta)}{\mu_{j\tau}(\hat{\theta})}.\tag{D.1}$$

In contrast to the representation in Theorem 3(c), however, the weights $a_{it+1,j\tau}$ may depend on the signal precisions.

Consider an environment with $\Theta=\mathbb{R}$ and where the agents share an (improper) uniform prior at time $t=0$. Suppose agent i's time-t private signal is given by $\omega_{it}=\theta+\epsilon_{it}$, where θ is the realized state and ϵ_{it} is distributed independently of other random variables according to a zero-mean normal distribution with precision η_{it}. Moreover, suppose that η_{it}'s and the structure of the social network are common knowledge among all agents. In such an environment, agents' beliefs remain normal at all times with precisions that are independent of the realization of agents' private signals (Mossel, Olsman, and Tamuz (2016)). Let m_{it} and κ_{it} denote the mean and precision of agent i's belief at time t. For any given pair of states θ and $\hat{\theta}$, the log-likelihood ratio corresponding to i's belief is equal to

$$\log\frac{\mu_{it}(\theta)}{\mu_{it}(\hat{\theta})}=\frac{\kappa_{it}}{2}(\hat{\theta}-\theta)(\hat{\theta}+\theta-2m_{it}).\tag{D.2}$$

Similarly, the normality of private signals implies that

$$\log\frac{\ell_{it}^{\theta}(\omega_{it})}{\ell_{it}^{\hat{\theta}}(\omega_{it})}=\frac{\eta_{it}}{2}(\hat{\theta}-\theta)(\hat{\theta}+\theta-2\omega_{it}).\tag{D.3}$$

We use the two equations above and an inductive argument to prove that agents' belief dynamics follow (D.1) at all times.

First, note that equation (D.1) is trivially satisfied at $t=1$, when each agent only has access to a single private observation. As the induction hypothesis, suppose that (D.1) is satisfied for all $\tau\le t$ and any agent i. Recall that agent i's information set at time $t+1$ consists of the history of her private signals, $(\omega_{i1},\dots,\omega_{it+1})$, and that of her neighbors' reports, $\mu_i^t=(\mu_{j\tau})_{j\in N_i,0\le\tau\le t}$. The joint normality of these signals and beliefs implies that i's belief at time $t+1$ is normally distributed with a mean that linearly depends on her private signals and the means of her neighbors' beliefs in previous periods. Formally, there are constants $c_{i\tau}$ and $b_{it,j\tau}$, independent of the signal realizations and summing up to 1, such that

$$m_{it+1} = \sum_{\tau=1}^{t+1} c_{i\tau}\omega_{i\tau} + \sum_{\tau=1}^{t}\sum_{j\in N_i} b_{it+1,j\tau} m_{j\tau}.$$

Note that $c_{it+1} = \eta_{it+1}/\kappa_{it+1}$ because ω_{it+1} is conditionally independent of i's prior observations and her neighbors' reports. We can use the above expression for m_{it+1} to express the log-likelihood ratio corresponding to agent i's time-$t+1$ belief as a linear function of her private signals and the means of her neighbors' reports:

$$\log\frac{\mu_{it+1}(\theta)}{\mu_{it+1}(\hat{\theta})} = \frac{\kappa_{it+1}}{2}(\hat{\theta}-\theta)\left(\hat{\theta}+\theta-2\sum_{\tau=1}^{t}\sum_{j\in N_i} b_{it+1,j\tau}m_{j\tau} - 2\sum_{\tau=1}^{t+1} c_{i\tau}\omega_{i\tau}\right).$$

Using equations (D.2) and (D.3) to substitute for $m_{j\tau}$ and $\omega_{i\tau}$ with the corresponding likelihood ratios and invoking the fact that $c_{it+1} = \eta_{it+1}/\kappa_{it+1}$, we obtain

$$\begin{aligned}\log\frac{\mu_{it+1}(\theta)}{\mu_{it+1}(\hat{\theta})} &= \log\frac{\ell^{\theta}_{it+1}(\omega_{it+1})}{\ell^{\hat{\theta}}_{it+1}(\omega_{it+1})} + \sum_{\tau=1}^{t}\sum_{j\in N_i}\left(b_{it+1,j\tau}\frac{\kappa_{it+1}}{\kappa_{j\tau}}\right)\log\frac{\mu_{j\tau}(\theta)}{\mu_{j\tau}(\hat{\theta})} \\ &\quad + \sum_{\tau=1}^{t}\left(c_{i\tau}\frac{\kappa_{it+1}}{\eta_{i\tau}}\right)\log\frac{\ell^{\theta}_{i\tau}(\omega_{i\tau})}{\ell^{\hat{\theta}}_{i\tau}(\omega_{i\tau})}.\end{aligned}$$

On the other hand, recall that the induction hypothesis maintains that (D.1) is satisfied for all $\tau \le t$. Therefore, the last term on the right-hand side above is itself a linear function of the log-likelihood ratios $\log(\mu_{j\tau}(\theta)/\mu_{j\tau}(\hat{\theta}))$ for $j \in N_i$ and $\tau \le t$. Thus, equation (D.1) is satisfied at $t+1$ for a collection of constants $a_{it+1,j\tau}$ that is independent of signal realizations $\omega_{j\tau}$ but depends on the signal precisions $\eta_{j\tau}$.

REFERENCES

MCCUAIG, W. (1984): "A Simple Proof of Menger's Theorem," *Journal of Graph Theory*, 8, 427–429. [3]

MOSSEL, E., N. OLSMAN, AND O. TAMUZ (2016): "Efficient Bayesian Learning in Social Networks With Gaussian Estimators," in *Annual Allerton Conference on Communication, Control, and Computing*. Illinois: Allerton House, 425–432. [15]

PRUITT, W. E. (1966): "Summability of Independent Random Variables," *Journal of Mathematics and Mechanics*, 15, 769–776. [13]

STEELE, J. M. (2004): *The Cauchy–Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. New York, NY: Cambridge University Press. [8]

Co-editor Joel Sobel handled this manuscript.

Manuscript received 9 August, 2016; final version accepted 20 December, 2017; available online 20 December, 2017.