Supplementary Appendices

SA.1. Belief Convergence

This section elaborates on Remark 3. Our discussion in this section focuses on deterministic networks.

One may expect the social belief to be eventually close to the stationary set with high probability: after all, when an agent's social belief is not close to the stationary set, her private information gives her a welfare improvement bounded away from zero; expanding observations should propagate these improvements, which implies (since utility is bounded) that they must eventually vanish. However, the following is a counterexample.[31]

Example SA.1. Consider binary states with a uniform prior, binary signals with symmetric precision (less than 1), and binary actions with simple utility. The network is as follows: agents 1 and 2 observe no one; for odd $n \geq 3$, agent n observes agent $n-2$; for even $n > 3$, agent n observes agent $n-1$ and agent 2. So there is expanding observations. In this network, the odd agents form an immediate-predecessor network and there is an equilibrium where a cascade along this subsequence starts from agent 3.

Now consider even agents. Consider the positive-probability event in which agents 1 and 2 take different actions. An even agent $n > 3$ observes agents $n-1$ and 2, which, given the equilibrium behavior of odd agents, is equivalent to observing agents 1 and 2. So the social belief of every even agent $n > 3$ equals the prior, which is bounded away from the stationary set.[32] □

The "problem" in Example SA.1 is that even though each of the even agents ($n > 3$) is getting a welfare improvement bounded away from zero, these improvements are not passed on to any future agents, and all future even agents continue

[31] Absent expanding observations, there are trivial counterexamples using the empty network.

[32] The example illustrates that with positive probability social beliefs may not eventually converge to the set of stationary beliefs. But the point also holds for posterior beliefs, not just social beliefs. For simplicity, consider the same example but with an additional signal that is uninformative. Call the two actions a and b. Consider an equilibrium in which the first agent plays a upon receiving the uninformative signal, while the second agent plays b upon receiving the uninformative signal. Then, in the event that the first agent plays b and the second agent plays a, the path of social beliefs for agents $n \geq 3$ is identical to the example above: odd agents are in a cascade, while even agents' social belief is just the prior. With positive probability, an even agent will now receive an uninformative signal, whereafter her posterior belief lies outside the stationary set.

to have social beliefs bounded away from the stationary set. In other words, expanding observations is not enough to validate the intuition described before the example. The following proposition identifies a reasonable condition on the network that is sufficient.

Proposition SA.1. *Assume there exist finitely many subsequences of agents* $\{n_{k,j}\}_{k=1}^{N_j}$ *(*$j = 1, \ldots, J < \infty$, $1 \leq N_j \leq \infty$*) such that agent* $n_{k,j}$ *observes* $n_{k-1,j}$*, and every agent in society is in at least one of the subsequences. Then, for all* $\varepsilon > 0$, $\lim_{n\to\infty} \varphi_n(\mu_n \in S^\varepsilon) = 1$.

The proposition's assumption encompasses canonical examples like the complete network and the k-immediate-predecessor networks (i.e., every agent observes the last k agents) for any $k \geq 1$. But it rules out any network in which infinite number of agents are not observed by any of their successors, which explains why it does not apply to Example SA.1.

Proof of Proposition SA.1. Along any subsequence j, $u(\varphi_{n_{k,j}}) \geq u(\varphi_{n_{k-1,j}}) + I(\varphi_{n_{k-1,j}})$ by the improvement principle, given that $n_{k-1,j}$ is observable to $n_{k,j}$. It follows that $\sum_{k=1}^{N_j} I(\varphi_{n_{k,j}}) \leq 2\overline{u}$. Hence, society's total improvement is bounded: $\sum_n I(\varphi_n) \leq 2\overline{u}J$.

Now fix any $\varepsilon, \delta > 0$. Consider $V_{\delta/2}$ defined in Lemma 3. The lemma established that $V_{\delta/2}$ is compact and $\varphi(\mu \notin V_{\delta/2}) < \delta/2, \forall \varphi \in \Phi^{BP}$. Since S^ε is open, $K := (S^\varepsilon)^c \cap V_{\delta/2}$ is compact. Next we argue $\mathbb{P}(\mu_n \in K\ i.o.) = 0$. Suppose, to contradiction, $\mathbb{P}(\mu_n \in K\ i.o.) > 0$. Then $\sum_n \mathbb{P}(\mu_n \in K) = \infty$ by the Borel-Cantelli lemma. Since K is compact and $I(\cdot) > 0$ on K, $I(\cdot)$ achieves its minimum in K at some $\underline{\mu} \in K$ with $I(\underline{\mu}) > 0$. So the total improvement is $\sum_n I(\varphi_n) \geq I(\underline{\mu}) \sum_n \mathbb{P}(\mu_n \in K) = \infty$, which contradicts $\sum_n I(\varphi_n) \leq 2\overline{u}J$.

Observe that $\mathbb{P}(\mu_n \in K\ i.o.) = 0$ implies $\varphi_n(\mu_n \in K) < \delta/2$ for all large n. Therefore, for all large n, $\varphi_n(\mu_n \in (S^\varepsilon)^c) \leq \varphi_n(\mu_n \in K) + \varphi_n(\mu_n \notin V_{\delta/2}) < \delta$. We conclude that for all $\varepsilon > 0$, $\lim_{n\to\infty} \varphi_n(\mu_n \in S^\varepsilon) = 1$. *Q.E.D.*

Remark 6. If $\Delta\Omega$ is compact (e.g., Ω itself is compact), we can replace $V_{\delta/2}$ in the proof with $\Delta\Omega$, so that $K = (S^\varepsilon)^c$. Then the argument in the proof's second paragraph shows that $\mathbb{P}(\mu_n \in (S^\varepsilon)^c\ i.o.) = 0$, i.e., the social belief converges to the stationary set almost surely rather than only in probability.

SA.2. ε-Excludability

This section elaborates on Remark 4. Say that for any $\varepsilon \in (0, 1/2)$ a set of states Ω' is *ε-distinguishable* from Ω'' if for any $\mu \in \Delta(\Omega' \cup \Omega'')$ with $\mu(\Omega') > \varepsilon$, there is a positive-measure set of signals S' such that $\mu(\Omega'|s) > 1 - \varepsilon$ for all $s \in S'$. A utility function and an information structure jointly satisfy *ε-excludability* if Ω_{a_1,a_2} and Ω_{a_2,a_1} are ε-distinguishable from each other, for any pair of actions a_1, a_2. Note that ε-excludability implies ε'-excludability for all $\varepsilon' > \varepsilon$, and excludability is equivalent to ε-excludability for all $\varepsilon > 0$.

Proposition SA.2. *Let Ω be finite. For all $\varepsilon \in (0, 1/2)$, ε-excludability implies that in any equilibrium σ, $\liminf_n \mathbb{E}_{\sigma,\mu_0}[u_n] \geq u^*(\mu_0) - 2\overline{u}\frac{\varepsilon}{1-\varepsilon}|\Omega|$.*

Before proving Proposition SA.2, we give an example illustrating the result's use.

Example SA.2. There are three states, $\omega \in \{1, 2, 3\}$, SCD preferences, and Laplace information:

$$f(s|\omega) = \frac{1}{2b} \exp\left(-\frac{|s - \omega|}{b}\right),$$

where $b > 0$ is a scale parameter; a smaller b corresponds to more precise information.

It is straightforward to verify that no two states can be distinguished from each other.[33] Therefore, not every stationary belief has adequate knowledge (so long as preferences are nontrivial), and by Theorem 1 there is inadequate learning.

Nonetheless, we claim that ε-excludability holds for any ε such that $\varepsilon > \frac{1}{1+\exp(\frac{1}{2b})}$. To see this, observe that since the information structure has MLRP and preferences satisfy SCD, we can focus on ε-distinguishing state 3 from 2 (or, equally, 2 from 1).[34] When $\varepsilon > \frac{1}{1+\exp(\frac{1}{2b})}$, we have $\frac{\varepsilon}{1-\varepsilon}\exp(1/b) > \frac{1-\varepsilon}{\varepsilon}$, so there exist signals that move the prior $(0, 1-\varepsilon, \varepsilon)$ to a posterior of at least $1 - \varepsilon$ on state 3, which implies ε-distinguishability of state 3 from 2.

Proposition SA.2 implies that in any equilibrium, $\liminf \mathbb{E}_{\sigma,\mu_0}[u_n] \geq u^*(\mu_0) - 6\overline{u}\exp(-\frac{1}{2b})$. This quantitative welfare bound yields, in particular, convergence to the full-information utility $u^*(\mu_0)$ as $b \to 0$. $\square$

[33] For any pair of states $\omega \neq \omega'$, and any signal s, the likelihood ratio $f(s|\omega)/f(s|\omega') \leq \exp(2/b)$.

[34] By MLRP, only arbitrarily large signals can distinguish a state from a lower state, and for large s the likelihood ratio $f(s|3)/f(s|2) < f(s|3)/f(s|1)$, so considering adjacent states is sufficient for ε-excludability.

Proof of Proposition SA.2. Take any stationary belief μ, and let a be an optimal action at belief μ. For each state ω, take any $a_\omega \in c(\omega)$, and consider $\mu_\omega(\cdot) := \mu(\cdot|\{\omega\} \cup \Omega_{a,a_\omega})$. If $\mu_\omega(\omega) \le \varepsilon$, then $\mu(\omega) \le \varepsilon$, so $(u(a_\omega,\omega) - u(a,\omega))\mu(\omega) \le 2\overline{u}\varepsilon$.

Consider the other case of $\mu_\omega(\omega) > \varepsilon$. For any $s \in S$, because $u(a,\omega') - u(a_\omega,\omega') \le 0$ for each $\omega' \notin \Omega_{a,a_\omega}$, and μ is stationary,

$$\sum_{\omega' \in \{\omega\} \cup \Omega_{a,a_\omega}} (u(a,\omega') - u(a_\omega,\omega'))\mu(\omega'|s) \ge \sum_{\omega' \in \Omega} (u(a,\omega') - u(a_\omega,\omega'))\mu(\omega'|s) \ge 0.$$

Then,

$$\begin{aligned}(u(a_\omega,\omega) - u(a,\omega))\mu_\omega(\omega|s) &\le \sum_{\omega' \in \Omega_{a,a_\omega}} (u(a,\omega') - u(a_\omega,\omega'))\mu_\omega(\omega'|s) \\ &\le 2\overline{u}\left(\sum_{\omega' \in \Omega_{a,a_\omega}} \mu_\omega(\omega'|s)\right) = 2\overline{u}(1 - \mu_\omega(\omega|s)).\end{aligned}$$

By ε-excludability, there exists a positive-measure set of signals S' such that, for any $s \in S'$, $\mu_\omega(\omega|s) > 1 - \varepsilon$, which implies that $u(a_\omega,\omega) - u(a,\omega) \le 2\overline{u}\frac{\varepsilon}{1-\varepsilon}$.

In either case ($\mu_\omega(\omega) \le \varepsilon$ or $\mu_\omega(\omega) > \varepsilon$), we have $(u(a_\omega,\omega) - u(a,\omega))\mu(\omega) \le 2\overline{u}\frac{\varepsilon}{1-\varepsilon}$. Since Ω is finite,

$$\sum_{\omega \in \Omega} (u(a_\omega,\omega) - u(a,\omega))\mu(\omega) \le 2\overline{u}\frac{\varepsilon}{1-\varepsilon}|\Omega|.$$

Namely, the utility gap $u^*(\mu) - u(\mu) \le 2\overline{u}\frac{\varepsilon}{1-\varepsilon}|\Omega|$, for any stationary belief μ.

Finally, for any $\varphi \in \Phi^S$,

$$u^*(\mu_0) - u(\varphi) = \mathbb{E}_\varphi[u^*(\mu) - u(\mu)] \le 2\overline{u}\frac{\varepsilon}{1-\varepsilon}|\Omega|.$$

By taking infimum of $u(\varphi)$ across $\varphi \in \Phi^S$, we obtain $u_*(\mu_0) \ge u^*(\mu_0) - 2\overline{u}\frac{\varepsilon}{1-\varepsilon}|\Omega|$, and subsequently by invoking Theorem 3, we conclude that in any equilibrium σ, $\liminf_n \mathbb{E}_{\sigma,\mu_0}[u_n] \ge u^*(\mu_0) - 2\overline{u}\frac{\varepsilon}{1-\varepsilon}|\Omega|$. *Q.E.D.*

SA.3. Details on Example 2

For Example 2, we show here how to construct a full-support prior such that the posterior probability is uniformly bounded away from 1 across signals and states. Take any prior μ such that for some $c > 0$, $\min\left\{\frac{\mu(n-1)}{\mu(n)}, \frac{\mu(n+1)}{\mu(n)}\right\} > c$ for all n

(e.g., a double-sided geometric distribution). Denoting the posterior after signal s by μ_s, the posterior likelihood ratio satisfies

$$\frac{\mu_s(\{n-1,n+1\})}{\mu_s(n)} = \frac{f(s|n-1)}{f(s|n)}\frac{\mu(n-1)}{\mu(n)} + \frac{f(s|n+1)}{f(s|n)}\frac{\mu(n+1)}{\mu(n)} > c\left(\frac{f(s|n-1)}{f(s|n)} + \frac{f(s|n+1)}{f(s|n)}\right).$$

As the last expression is the sum of a strictly positive decreasing function of s and a strictly positive increasing function of s, it is bounded away from 0 in s. The bound is independent of n because normal information is a location-shift family of distributions. Therefore, the posterior likelihood ratio is uniformly bounded away from 0, and hence, the posterior $\mu_s(n)$ is uniformly bounded away from 1.