Stationary social learning in a changing environment

Raphaël Lévy[*], Marcin Pęski[†], Nicolas Vieille[‡]

April 9, 2024

Abstract

We consider social learning in a changing world. With changing states, societies can be responsive only if agents regularly act upon fresh information, which significantly limits the value of observational learning. When the state is close to persistent, a consensus whereby most agents choose the same action typically emerges. However, the consensus action is not perfectly correlated with the state, because societies exhibit inertia following state changes. Phases of inertia may be longer when signals are more precise, even if agents draw large samples of past actions, as actions then become too correlated within samples, thereby reducing informativeness and welfare.

1 Introduction

The literature on social learning has extensively studied the extent to which agents learn from others' actions. In particular, it has been quite successful at understanding the possible emergence of informational cascades, and conditions under which the consensus that eventually forms over time is correct (see Bikhchandani et al. (2021) for a recent survey). However, little attention has been drawn to the possibility that the underlying state of nature might change over time.[1] Still, in several applications, e.g., technology adoption or investment decisions, the optimal course of action is likely to change over time, raising the question of whether information is efficiently aggregated.

The possibility of state changes provides new insights both from applied and theoretical perspectives. For instance, the dynamics of learning may shed light on how societies react to changes in the

[*]HEC Paris.

[†]University of Toronto

[‡]HEC Paris.

[1]Notable exceptions are Moscarini, Ottaviani and Smith (1998); Acemoglu, Nedic and Ozdaglar (2008); Frongillo, Schoenebeck and Tamuz (2011); Dasaratha, Golub and Hak (2023); Huang (2022). See below for a detailed account on how our work relates to these papers.

environment, and on how a dominant consensus may be replaced by a new one.[2] From a theoretical perspective, the possibility of state changes creates a tension between information aggregation and responsiveness to change. Indeed, while efficient aggregation supposes that some agents safely rely on their peers to act, it also requires that society reacts swiftly to a change in the environment, ruling out informational cascades. The goal of this paper is to evaluate how this tension shapes equilibrium welfare.

To address this question, we analyze steady-state equilibria in a model where (a) the state of nature follows a Markov chain and (b) in each period, a continuum of short-lived agents draw a finite sample of past actions and have access to a (possibly costly) informative signal. Though unnecessary for our results, allowing for costly information makes social learning even more desirable. Indeed, the potential welfare gains from learning from others result not only from better-informed decisions, but also from savings on information costs.

Two forces drive equilibrium welfare. The first one is reminiscent of the Grossman-Stiglitz paradox: while past actions must be informative at the steady state, they cannot be very informative, for otherwise agents would have no incentive to acquire further information. In a steady state, this logic imposes that some fresh information flows in every period, and thus puts a limit on equilibrium welfare. When agents sample at least two actions, a second (countervailing) force comes into play. In that case, some agents will sample conflicting evidence, and rely entirely on their private information, while others will obtain unambiguous evidence allowing them to possibly free-ride on the information acquired by the former. This creates a inter-temporal externality across samples which may be expected to lead to higher welfare. This is exactly what happens in Banerjee and Fudenberg (2004), who show in a fixed-state model that learning is eventually complete when agents sample as few as two actions. In sharp contrast, our chief finding is that allowing for even arbitrarily rare changes in the state typically results in incomplete learning.

We first analyze the case where agents sample at most two actions. In any equilibrium, *all* agents must acquire (or make use of) information with positive probability, regardless of their samples. Indeed, we establish that the dynamics of the population would otherwise eventually bring all agents to play the same action regardless of the state. In a changing world, such an irreversible consensus would be uninformative, and thus cannot be an equilibrium outcome. Thus, learning is incomplete: agents can never entirely rely on their peers, even when the state is arbitrarily persistent. Not only observing two actions does not ensure complete learning with changing states,

[2] Examples of a change in the dominant technology abound, ranging from the "war of the currents" in the late 19th century, the "quartz crisis" in watchmaking in the 1970s to Facebook overtaking MySpace as the dominant social network.

but actually does no better than observing only one, and typically worse.

For larger samples, we focus on the case where the state is highly persistent, and first show that a consensus prevails. That is, in a steady-state equilibrium, most likely most agents play the same action. However, this consensus must be fluctuating, and the population must oscillate over time following state changes. The dynamics following a state change is characterized by two phases. Agents first stick to the current consensus unless their sample conveys somewhat mixed evidence. But such conflicting evidence is unlikely in a society where one action dominates, hence there may be significant inertia in moving away from a consensus. Once the population displays some minimal dissent, the fraction of agents acting against the old consensus quickly takes off, and the population snowballs towards a new consensus. The efficiency of learning reflects how long it takes to move away from an outdated consensus following state changes.

We next show that, provided that (binary) signals are sufficiently precise and information not too costly, there exists an equilibrium in which welfare is the same as if agents sampled one action only (or even none). When signals are precise, agents are more likely to play the right action, and both the convergence towards a correct consensus and away from a wrong consensus are fast. However, it turns out that the *relative* speed of convergence towards a correct consensus is higher. As for samples of size two, agents observing even unanimous samples must then acquire information to make sure that society remains responsive to change and does not get stuck in an irreversible consensus. A complementary intuition is that, when signals are precise, the actions of agents acquiring information are highly correlated (to the state, hence among themselves), and convergence towards a consensus is quite fast. N: NOT SURE THIS IS WHAT YOU HAVE IN MIND.

The proof of this result involves a methodological innovation to get bounds on the beliefs for non-unanimous sample compositions. The idea is that, when state changes are rare, equilibrium beliefs can be approximated in terms of time-average values of the dynamical systems that describe the evolution of the population's behavior. Perhaps surprisingly, these belief estimates point to the presence of belief reversal: seeing one dissenting action within an otherwise unanimous sample should be taken as evidence that the minority action is the correct one.

We introduce the model in Section 2. Section 3 is devoted to small samples, and Section 4 provides results for larger samples in the persistent limit case.

Section 5 addresses robustness issues and extensions. Based on numerical evidence, we provide a complete description of equilibrium behavior and welfare in the case of samples of size three. When signals are weak, equilibrium welfare may be higher than in the no-sampling case, unlike for precise signals. Yet, learning is always incomplete, in line with the main message of the paper.

In addition, while our incomplete learning results rest on the premise that signals are bounded and/or costly, we recover a positive social learning result (complete learning at the persistent limit) if signals are *both* unbounded and free, irrespective of the sample size.

We discuss the relation to the literature in Section 6, and conclude in Section 7.

2 The model

2.1 States, actions and payoffs

We consider a social learning model in discrete time with an evolving, binary state of nature $\theta \in \Theta := \{0, 1\}$. In each period, a continuum of short-lived agents each choose an action from the action set $A := \{0, 1\}$. Agents obtain a utility of one when their action matches the current state, and of zero otherwise.

Successive states (θ_t) follow a symmetric Markov chain over Θ. The parameter $\lambda := \mathbf{P}(\theta_{t+1} \neq \theta \mid \theta_t = \theta)$ captures the degree of persistency. States are i.i.d. if $\lambda = \frac{1}{2}$, and fully persistent if $\lambda = 0$. We assume that $\lambda \in \left(0, \frac{1}{2}\right)$: the state is persistent, but not fully.

2.2 Timing, sampling and signals

At each date t, events unfold as follows. Each new-born agent (i) first observes a random sample of n past actions, (ii) next decides whether or not to acquire additional information about the current state θ_t at cost $c \geq 0$, (iii) finally picks an action $a \in A$.

We assume that sampled actions are drawn from the pool of actions played in the previous period, in proportion to their prevalence in the population (proportional sampling).That is, the sample composition at date t, measured by the count of ones, follows a binomial distribution $B(n, x_{t-1})$, where x_{t-1} is the fraction of agents playing action 1 in period $t-1$. Samples are independent across agents and private.

The additional information available to agents consists of a signal that is independent across periods and agents, conditional on the sequence of states. We denote by q the posterior belief assigned to $\theta = 1$ given the signal, under a uniform prior, and refer to q as a *private* belief.

We denote by H_θ the right-continuous cdf of q in state θ. Signal distributions are assumed symmetric across states. That is, the distribution of the posterior probability assigned to θ conditional on the state being θ, is the same for both states. This corresponds to $H_0(q)_- = 1 - H_1(1-q)$ for each $q \in [0, 1]$.[3] We rule out uninformative signals, and assume throughout that $H_1(q) < H_0(q)$ for some $q \in (0, 1)$. Finally, we denote by $\bar{q}$ the supremum of the support of the unconditional

[3]The asymmetric case is discussed in Section 5.5.

distribution $H := \frac{1}{2}H_0 + \frac{1}{2}H_1$. By symmetry, the infimum is $1-\bar{q}$. Following the usual terminology, signals are *unbounded* if $\bar{q} = 1$, and *bounded* if $\bar{q} < 1$.

2.3 Equilibrium concept

We focus on equilibrium steady states in which all agents across and within periods use the same decision rule σ.[4] The equilibrium notion requires that σ is optimal given beliefs, and that beliefs are derived from the invariant joint distribution of states and samples induced by σ. In addition, we restrict attention to symmetric equilibria, i.e., we require that the equilibrium is unchanged when relabelling actions and states so that the decisions given a sample $n-k$ are the mirror images of those made with sample k. A formal definition of equilibrium steady states is given in Section 4.1, together with an existence result. At this stage, we simply denote by p_k the (interim) probability that the current state is $\theta = 1$ conditional on seeing a sample composed of $k \in \{0, \ldots, n\}$ ones.

2.4 Information acquisition

Consider an agent who holds an interim belief p and contemplates acquiring information. If he does not, he plays the action most likely to be optimal, and the expected probability of matching the state is $u(p) := \max(p, 1-p)$. If he does, and receives a signal inducing a private belief q, the agent chooses action 1 whenever state 1 is the most likely state, i.e., if $pq \geq (1-p)(1-q) \Leftrightarrow q \geq 1-p$, with indifference if $q = 1-p$.

Accordingly, the probability of playing the correct action $a = \theta$ is given by

$$v(p) := p\,(1 - H_1(1-p)) + (1-p)H_0(1-p). \tag{1}$$

The function v is convex, increasing on $[\frac{1}{2}, 1]$ and symmetric: $v(p) = v(1-p)$ for all p. Of course, $v(p) \geq u(p)$ for all p since more information cannot hurt.

The net value of acquiring information $v(p) - u(p)$ is maximal when $p = \frac{1}{2}$. If $c > v(\frac{1}{2}) - u(\frac{1}{2})$, agents never acquire information and samples are not informative. We rule out this uninteresting case and assume throughout:

Assumption 1 $v(\frac{1}{2}) - u(\frac{1}{2}) \geq c$.

These properties imply the existence of a unique $\hat{p} \in [\frac{1}{2}, 1]$ such that $v(p) - c > u(p)$ iff $p \in (1-\hat{p}, \hat{p})$. As our analysis highlights, what ultimately matters is whether $\hat{p} < 1$ or not.

[4] This includes the decision whether to acquire information as a function of one's sample, and which action to choose.

When signals are bounded ($\overline{q} < 1$), one has $v(p) = u(p)$ for $p > \overline{q}$, hence $\hat{p} < 1$ irrespective of whether $c = 0$ or $c > 0$. When signals are unbounded, $v(p) > u(p)$ for all $p \in (0, 1)$, hence $\hat{p} = 1$ if $c = 0$ and $\hat{p} < 1$ if $c > 0$. In what follows, we maintain the assumption that $\hat{p} < 1$ (the richer case), and defer the discussion of the easier case $\hat{p} = 1$ to Section 5.2.

Assumption 2 $\hat{p} < 1$.

Figures 1 and 2 illustrate two typical cases. In Figure 1, signals are unbounded and $c > 0$; in Figure 2, signals are binary with precision $\frac{1}{2} < \pi < 1$ (bounded). In that case, private beliefs are either $1 - \pi$ or π, and $v(p) = \max(p, 1 - p, \pi)$, implying $\hat{p} = \pi - c$.[5]

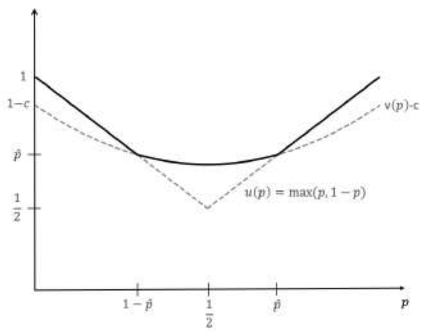

Figure 1: Signals with unbounded strength

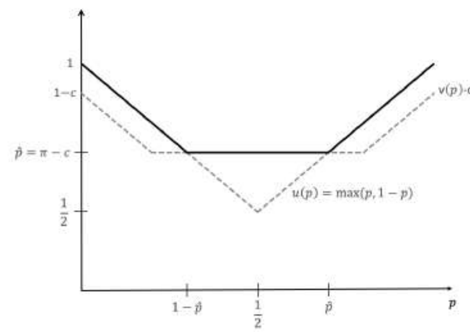

Figure 2: Binary signals with precision π

We conclude this section by noting that private signals are sometimes used in any equilibrium.

Lemma 1 *In any equilibrium steady state, one has* $1 - \hat{p} \leq p_k \leq \hat{p}$ *for some* k.

Proof. Assume instead that for some equilibrium and for each k, one has either $p_k < 1 - \hat{p}$ or $p_k > \hat{p}$. Signals are never used, hence samples are uninformative at the steady state and therefore, $p_k = \frac{1}{2}$ for each k. A contradiction. ■

Lemma 1 then rules out steady-state cascades. Still, one may have $p_k \notin [1 - \hat{p}, \hat{p}]$ for *some* sample k. Upon observing such a sample, agents play their perceived best action without [acquiring or using] further information.

3 Small samples

We discuss here small sample sizes ($n \leq 2$). The case $n = 0$ where agents do not sample serves as a no-social-learning benchmark. In that case, the belief of all agents at a steady state is given by the

[5]Note that $\hat{p} < 1$ even for $c = 0$, as discussed above.

invariant distribution of (θ_t), which is uniform over Θ. Agents acquire information, and obtain an expected payoff of $v(\frac{1}{2}) - c$.

3.1 Samples of size $n = 1$

We assume here that $n = 1$. Equilibrium symmetry implies that $p_1 = 1 - p_0$ and that all agents acquire information with the same probability β. Lemma 1 implies that $\beta > 0$. We set $\lambda_* := \frac{c}{2(\hat{p}+c)-1} < \frac{1}{2}$. Note that $\lambda_* = 0$ when $c = 0$.

Proposition 1 *If $n = 1$, there is a unique equilibrium steady state:*

- *If $\lambda < \lambda_*$, one has $p_1 = \hat{p}$ and $\beta = \lambda \dfrac{2\hat{p} - 1}{c(1 - 2\lambda)} \in (0, 1)$.*
- *If $\lambda \geq \lambda_*$, one has $p_1 \in (1 - \hat{p}, \hat{p})$ and $\beta = 1$.*

Proof. Consider an agent A in period t who samples the action a_{t-1} of some player B. In the steady state, both A and B hold either an interim belief p_1 or $p_0 = 1 - p_1$, where p_1 obeys the following equation:

$$\begin{aligned} p_1 &= \mathbf{P}(\theta_t = 1 \mid a_{t-1} = 1) \\ &= (1 - \lambda)\mathbf{P}(\theta_{t-1} = 1 \mid a_{t-1} = 1) + \lambda \mathbf{P}(\theta_{t-1} = 0 \mid a_{t-1} = 1). \end{aligned} \tag{2}$$

Either, with probability β, B acquired information and then played the right action with probability $v(p_1)(= v(p_0))$, or did not, and matched the state with probability $u(p_1)$.

If $0 < \beta < 1$, B's indifference condition imposes $p_1 = \hat{p}$. One therefore derives $\mathbf{P}(\theta_{t-1} = 1 \mid a_{t-1} = 1) = \beta v(\hat{p}) + (1 - \beta)u(\hat{p}) = \hat{p} + \beta c$, using the definition of $\hat{p}$. Substituting into (2) yields $\hat{p} = \lambda + (1 - 2\lambda)(\hat{p} + \beta c)$, hence $\beta = \lambda \dfrac{2\hat{p} - 1}{c(1 - 2\lambda)}$. $\beta < 1$ then requires $\lambda < \lambda_*$.

If $\beta = 1$, one has $\mathbf{P}(\theta_{t-1} = 1 \mid a_{t-1} = 1) = v(p_1)$ and (2) now reads

$$p_1 = (1 - \lambda)v(p_1) + \lambda(1 - v(p_1)). \tag{3}$$

Since $v'(p_1) < 1$ for $p_1 \leq \hat{p}$ and since $v(\frac{1}{2}) - c > \frac{1}{2}$, (3) has a (unique) solution in $[\frac{1}{2}, \hat{p}]$ if and only if $\lambda \geq \lambda_*$. ■

When the state changes frequently, *past* actions cannot possibly be very informative about the *current* state, and information is acquired with probability 1. As the state gets more persistent, past actions potentially become informative, and β decreases. In the persistent limit $\lambda \to 0$, agents acquire information with vanishing probability, hence most likely replicate the action they sample. Welfare increases from $v(\frac{1}{2}) - c$ to $\hat{p}$ as λ decreases from the i.i.d. case to the persistent limit. More generally, welfare is given as follows.

Corollary 1 *The equilibrium payoff is $v(p_1) - c$ for $\lambda \geq \lambda_*$, where p_1 is the solution of (3), and $p_1 = \hat{p}$ for $\lambda < \lambda_*$.*

If signals are binary with precision π, the equilibrium payoff is $\pi - c = \hat{p}$, for each $\lambda > 0$.

3.2 Samples of size $n = 2$

If $n = 1$, private signals are always interim valuable in equilibrium, which limits (and pins down) the informativeness of past actions. The situation is quite different with larger samples. With $n > 1$, the efficiency of social learning can be improved if agents with some sample k generate enough information (in a given period) that (later) agents with a different sample k' find it optimal to herd – that is, if there exists (k, k') such that $1 - \hat{p} < p_k < \hat{p} < p_{k'}$. Ultimately, equilibrium payoffs are determined by the magnitude of such information externalities across samples.

We here assume $n = 2$. By symmetry, $p_1 = \frac{1}{2}$: agents who sample conflicting actions are confused and acquire information. In line with the previous remark, the key question is whether $p_2 > \hat{p}$. Casual intuition suggests that this should be the case when the state is sufficiently persistent. Remarkably, this intuition is incorrect.

Proposition 2 *In any equilibrium, one has $p_2 \in [1 - \hat{p}, \hat{p}]$.*

Proposition 2 implies that agents are *always* willing to acquire information when $c > 0$. For $c = 0$, it implies that it can *never* be strictly optimal to ignore one's signal.

The logic works as follows. Assume that $p_2 > \hat{p}$, so that agents only acquire information when sampling conflicting evidence. By acquiring information, these agents are instrumental in moving towards a correct consensus, but also in moving away from this (outdated) consensus once the state changes. However, this arrival of fresh information does not suffice to compensate for state changes: the forces of imitation (herding) are so strong that sooner or later there will be too few agents sampling mixed evidence, and the population will stop being responsive.

The complete proof of Proposition 2 is in the Appendix. We provide a sketch below.

Proof Sketch. We argue by contradiction and assume that $p_2 > \hat{p}$ at some equilibrium. Since $p_1 = \frac{1}{2}$, agents with a balanced sample acquire information, and choose action 1 if their private belief exceeds $\frac{1}{2}$, which has probability $\phi_\theta := 1 - H_\theta(\frac{1}{2})$ in state θ.

Denoting by x_{t-1} the fraction of agents playing action 1 in period $t - 1$, the probability that a generic agent in period t plays action 1 is thus given by $x_t = \overline{g}_{\theta_t}(x_{t-1})$, where

$$\overline{g}_\theta(x) := x^2 + 2x(1 - x)\phi_\theta.$$

Since $\phi_1 > \frac{1}{2} > \phi_0$, one has $\overline{g}_1(x) > x > \overline{g}_0(x)$ for each $x \in (0, 1)$: the popularity x_t of action 1 increases over time when $\theta_t = 1$, and decreases otherwise, as shown in Figure 3.

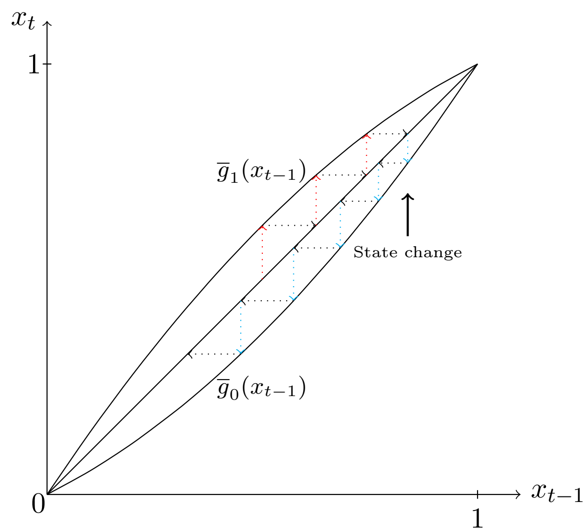

Figure 3: Population dynamics

For x close to 0, the ratio $\overline{g}_\theta(x)/x$ is approximately equal to $2\phi_\theta$. Thus, as long as x_t is close to zero, $\ln x_t$ increases by $\ln 2\phi_1$ in every period where $\theta_t = 1$, and decreases by $\ln 2\phi_0$ when $\theta_t = 0$. Since $4\phi_0\phi_1 = 4\phi_1(1-\phi_1) < 1$, one has $\ln 2\phi_1 < -\ln 2\phi_0$: step sizes are *higher* when $\ln x_t$ *decreases*. This means that (x_t) converges faster *towards* a consensus on action 0 when $\theta_t = 0$ than it moves *away* from this consensus when $\theta_t = 1$. Consequently, whenever (x_t) approaches either 0 or 1, there is a positive probability that the population will never bounce away from this consensus, even following state changes.

Combined with the observation, obvious from Figure 3, that (x_t) cannot stay bounded away from 0 and 1 over time, this implies that the sequence (x_t) converges to either 0 or 1, almost surely: the population converges to a *permanent* consensus, although the state of nature keeps changing. Thus, past actions are fully uninformative hence $p_2 = \frac{1}{2}$. A contradiction. ■

Corollary 2 *If signals are binary, the equilibrium payoff is $\hat{p}$ for all $n = 0, 1, 2$, all $\lambda > 0$ and all $c \geq 0$. If H_θ has support $[1 - \bar{q}, \bar{q}]$, the equilibrium payoff is strictly lower for $n = 2$ than for $n = 1$ if $\lambda \leq \lambda_*$.*

Proof. Interim equilibrium beliefs are $p_1 = \frac{1}{2}$ and $p_2 = 1 - p_0 \in \left[\frac{1}{2}, \hat{p}\right]$, hence the equilibrium payoff is a convex combination of $v(\frac{1}{2}) - c$ and of $v(p_2) - c$, with a positive weight on the former.

If signals are binary with precision π, $v(p) = \pi$ for all $p \in [1 - \hat{p}, \hat{p}]$, and the equilibrium payoff is then $\pi - c$ for every $\lambda > 0$ and $c \geq 0$, like in the cases $n = 0$ and $n = 1$.

If H_θ has support $[1 - \bar{q}, \bar{q}]$, v is strictly convex on $[1 - \hat{p}, \hat{p}]$, and the equilibrium payoff is below

$\hat{p} = v(\hat{p}) - c$. From Corollary 1, welfare with $n = 1$ is $\hat{p}$ as long as $\lambda \leq \lambda^*$. ■ If signals are binary, social learning arises in equilibrium (some agents do herd), yet there is no welfare gain over the benchmark case $n = 0$ where information is always acquired. With richer signals, the equilibrium welfare is non-monotonic in the sample size, at least for small λ and n. This is all the more surprising as this holds when the state is highly persistent, that is, when the logic of observational learning acts most forcefully.

3.3 Comparison to fixed state models

We here compare these findings to the literature with a fixed state and known calendar time. The relevant comparison is with Banerjee and Fudenberg (2004) (henceforth BF), who assume that in each period, a continuum of short-lived agents each draws a random sample of past actions and observe a signal for free.[6] If $n = 1$, the analysis of BF implies the existence of a continuum of steady states. In any such steady state, the fraction of agents who play the correct action is constant over time, and equal to $p_1 \geq \hat{p}$. In each period, all agents ignore their signals, and replicate the action they sample.[7] This multiplicity is ruled out with changing states, as it cannot be that agents systematically ignore their signals.

If $n = 2$, our results sharply contrast with BF. For $n \geq 2$, BF shows that, under minimal assumptions on signals (that are satisfied in our setup), learning is eventually complete: actions converge to the correct one. As soon as the state may change, our analysis instead shows that equilibrium payoffs do not exceed $\hat{p}$, even as $\lambda \to 0$. Accordingly, allowing for a changing state has a strong negative impact. In addition, the comparison between samples of size 1 and 2 with bounded signals suggests that, while observing more actions compensates for a limited signal quality when the state is fixed, this may exacerbate the inefficiency in a changing world.

4 The general case: equilibrium analysis

The discussion on small samples illustrates a key complication. When only one action is sampled, the interim belief at date t involves only the *expected* value of x_{t-1} in each state. That is, interim beliefs reflect how often on average past agents play the right action. This allows for a closed form analysis.

[6] Although BF considers free signals while we allow for costly signals, this distinction is irrelevant: on the one hand, our results apply to free signals without loss; on the other hand, the results of BF still hold with costly signals (we provide a proof in the supplementary material).

[7] Notice that since there is a continuum of agents, not all agents play the same action.

As soon as two actions are sampled, the interim belief p_k also involves the correlation of actions within samples.[8] Formally, the belief likelihood ratio is

$$\frac{\mathbf{P}\left(\theta_{t-1}=1 \mid k\right)}{\mathbf{P}\left(\theta_{t-1}=0 \mid k\right)}=\frac{\int_0^1 x^k(1-x)^{n-k} d\mu_1(x)}{\int_0^1 x^k(1-x)^{n-k} d\mu_0(x)} \tag{4}$$

where μ_θ is the distribution of x_{t-1} in state θ, so that the formula for p_k involves all l-th moments of x_{t-1} for $l \leq n$; in addition, the evolution of x_t^l over time involves even higher powers of x_t, as can be checked. Consequently, the equilibrium conditions involve the entire joint distribution of (θ_t, x_t), a complex object. [9]

For this reason, we focus on the persistent limit $\lambda \to 0$, which we view as the most relevant case for a comparison with the usual fixed-state setup. Strategies and equilibrium steady states are formally defined in Section 4.1. In Section 4.2, we provide an asymptotic consensus result. Section 4.3 discusses learning completeness, that is, the extent to which the prevailing consensus is correct.

4.1 Strategies and equilibrium

A strategy specifies whether or not to acquire information (if $c > 0$) and which action to choose. These decisions depend on the composition of one's sample, and (whenever relevant) on the signal. A strategy is thus a pair $\sigma = (\beta, \alpha)$ of (measurable) maps, with $\beta : \{0, \ldots, n\} \to [0, 1]$ and $\alpha : \{0, \ldots, n\} \times [0, 1] \to \Delta(\{0, 1\})$, with the understanding that $\beta(k)$ is the probability of acquiring information upon observing a sample composed of k ones, and $\alpha(k, q)$ is the probability of playing action 1 upon drawing sample k and observing a private belief $q \in [0, 1]$.[10] Not acquiring information is informationally equivalent to drawing a signal $q = \frac{1}{2}$ for sure: an agent with sample k who does not acquire information thus plays 1 with probability $\alpha(k, \frac{1}{2})$.

Conditional on the state being θ, an agent sampling k thus plays action 1 with probability

$$\phi_\theta(k) := \beta(k) \int_0^1 \alpha(k, q) dH_\theta(q) + (1 - \beta(k))\alpha(k, \frac{1}{2}). \tag{5}$$

The sample composition at date t follows a Binomial distribution $B(n, x_{t-1})$. It follows that the fraction of agents choosing action 1 in period t is

$$x_t = g_{\theta_t}(x_{t-1}) := \sum_{k=0}^{n} \binom{n}{k} x_{t-1}^k (1 - x_{t-1})^{n-k} \phi_{\theta_t}(k) \tag{6}$$

[8] Actions sampled in period t are independent conditional on x_{t-1}, but correlated *ex ante*.

[9] Even in the much simpler case where states are i.i.d. and $n = 1$, extensive work has focused on the properties of the steady-state distribution of x (Solomyak, 1995; Bhattacharya and Majumdar, 2007). This leaves little hope for a tractable analysis in our case.

[10] We recall that q is the belief derived from the signal, computed with a uniform prior.

For fixed σ, the pair (θ_t, x_t) follows a Markov chain over $\Theta \times [0,1]$. Over time, θ_t evolves independently of x_t, and $x_t = g_{\theta_t}(x_{t-1})$.

An *equilibrium steady state* is a pair (μ, σ) where $\mu \in \Delta(\Theta \times [0,1])$ is an invariant measure for (θ_t, x_t), and σ is optimal given μ. The optimality condition on σ reads as **C1** and **C2** below:

C1 $\beta(k) = 1$ if $p_k \in (1-\hat{p}, \hat{p})$ and $\beta(k) = 0$ if $p_k \notin [1-\hat{p}, \hat{p}]$;[11]

C2 $\alpha(k,q) = 1$ if $q > 1 - p_k$ and $\alpha(k,q) = 0$ if $q < 1 - p_k$.

We recall that the interim belief $p_k = \mathbf{P}(\theta_t = 1 \mid k)$ is the belief on the *current* state. It is related to the belief on the *previous* state through the equality $p_k = (1-\lambda)\mathbf{P}(\theta_{t-1} = 1 \mid k) + \lambda \mathbf{P}(\theta_{t-1} = 0 \mid k)$. The belief $\mathbf{P}(\theta_{t-1} = 1 \mid k)$ is itself related to the steady-state distribution μ through (4).

The condition that μ is an invariant distribution for (θ_t, x_t) reads

C3 $\mu(\theta, X) = (1-\lambda)\,\mu\left(\theta, g_\theta^{-1}(X)\right) + \lambda\mu\left(1-\theta, g_\theta^{-1}(X)\right)$ for all measurable $X \subset [0,1]$.

Our focus is on symmetric equilibria: we require in addition that μ and σ treat the two states and actions symmetrically. Formally:

C4 $\beta(k) = \beta(n-k)$ and $\alpha(k,q) = 1 - \alpha(n-k, 1-q)$ for each k and q.

C5 μ is invariant under the transformation $(\theta, x) \mapsto (1-\theta, 1-x)$.

We denote by $G(\lambda)$ the game with transition parameter λ.

Theorem 1 *The game $G(\lambda)$ has a symmetric equilibrium steady state.*

The proof relies on a fixed-point argument (see the supplementary material).

4.2 Aggregate behavior: a consensus result

We first derive a general result on the aggregate behavior in the population, and prove that the actions of agents become highly correlated as the state gets close to persistent.

Theorem 2 *Let $n \geq 2$ and let $(\mu_\lambda, \sigma_\lambda)$ be any equilibrium steady state of $G(\lambda)$, for $\lambda > 0$. As $\lambda \to 0$, the marginal of μ_λ over $x \in [0,1]$ converges to the uniform distribution over the two-point set $\{0,1\}$.*[12]

[11] If $c = 0$, condition **C1** can be omitted. Indeed, if $p_k > \hat{p}$, it is optimal to play $a = 1$ for all q so that $\alpha(k,q) = 1$. For such k, $\phi_\theta(k) = 1$ irrespective of $\beta(k)$. Similarly, when $p_k < 1-\hat{p}$, $\phi_\theta(k) = 0$ for any $\beta(k)$.

[12] Limits are understood in the sense of weak convergence of probability measures over $[0,1]$.

According to Theorem 2, most likely most agents play the same action in any given period. This consensus result implies in turn that most likely most agents draw a unanimous sample $k \in \{0, n\}$ consisting only of zeroes or of ones.

Proof. Let λ and an equilibrium (μ, σ) of $G(\lambda)$ be given, and denote by κ^* the (steady-state) average fraction of agents whose action matches the state. κ^* is weakly higher than the equilibrium payoff w^* because information acquisition costs are not accounted for.

Consider a generic agent in period t, and assume that he observes his sample in some random order $a^{(1)}, \ldots, a^{(n)}$. One available strategy σ_1 is to simply imitate $a^{(1)}$. The strategy σ_1 would yield a payoff of κ^* if the state were fixed. Accounting for state transitions, and assuming $a^{(1)} = 1$ for concreteness, σ_1 yields

$$w(\sigma_1) := \mathbf{P}(\theta_t = 1 \mid a^{(1)} = 1) = (1 - \lambda)\kappa^* + \lambda(1 - \kappa^*) \geq \kappa^* - \lambda,$$

and thus, $w(\sigma_1) \geq w^* - \lambda$. Since no strategy improves upon w^*, this implies that the marginal gain of observing $a^{(2)}$ is at most λ.[13]

In turn, this implies that $a^{(1)}$ and $a^{(2)}$ coincide with high probability when λ is small. Indeed, consider the alternative strategy σ_2 consisting in playing $a^{(1)}$ if the second action confirms the first one ($a^{(1)} = a^{(2)}$) and acquiring information otherwise ($a^{(1)} \neq a^{(2)}$). In the latter case, the agent's belief is $1/2$, hence the agent's payoff conditional on $a^{(1)} \neq a^{(2)}$ is $v(1/2) - c$. Therefore, the payoff $w(\sigma_2)$ is a convex combination of $v(1/2) - c$ and of $\mathbf{P}(\theta_t = 1 \mid a^{(1)} = a^{(2)} = 1)$, where the weights are the conditional probabilities of $a^{(2)} = 0$ and of $a^{(2)} = 1$ given $a^{(1)} = 1$.

On the other hand, the martingale property of beliefs ensures that $\mathbf{P}(\theta_t = 1 \mid a^{(1)} = 1)$ (which is also $w(\sigma_1)$) is a convex combination of the beliefs $1/2$ and of $\mathbf{P}(\theta_t = 1 \mid a^{(1)} = a^{(2)} = 1)$, with the *same* weights. Since $v(1/2) - c > u(1/2) = 1/2$, and since $w(\sigma_2) \leq w^* \leq w(\sigma_1) + \lambda$, it follows that the probability $\mathbf{P}(a^{(2)} = 0 \mid a^{(1)} = 1)$ that $a^{(2)}$ contradicts $a^{(1)}$ is at most of the order of λ.

To conclude, recall that $a^{(1)}$ and $a^{(2)}$ are independent draws from a Bernoulli distribution with parameter x, where x is first drawn according to μ. Since $a^{(1)}$ and $a^{(2)}$ coincide with high probability, it must be that x is quite close to 0 or to 1, with high μ-probability. ■

According to Theorem 2, the population is in consensus in a typical period, with x_t being close to either 0 or 1. At the same time, the consensus must evolve over time in response to changes in the state.[14] This implies that the population alternates between the two consensus, and that the transition time is vanishingly short as $\lambda \to 0$ relative to how long a given consensus prevails.

[13] If $c > 0$, this also implies that the steady-state expected fraction of agents who acquire information is at most of the order of λ.

[14] For otherwise, samples would not be informative, agents would always acquire information, and the population would adjust to new situations.

Let us briefly elaborate on the transition dynamics, assuming $n = 2$ and $c > 0$ for concreteness. We denote by $\beta > 0$ the equilibrium probability of acquiring information when sampling $k \in \{0, 2\}$. From Theorem 2, we know that $\beta \to 0$ as $\lambda \to 0$ (see Footnote 13). For λ small, interim beliefs are $p_2 = 1 - p_0 = \hat{p}$ and $p_1 = \frac{1}{2}$. Since agents either acquire information or herd, it can be checked that x_t evolves according to $x_t = g_{\theta_t}(x_{t-1})$, with

$$g_\theta(x) = \bar{g}_\theta(x) + \beta \left(\psi_\theta (1 - x)^2 - \psi_{1-\theta} x^2 \right), \tag{7}$$

where $\psi_\theta := 1 - H_\theta(\hat{p})$[15] and $\bar{g}_\theta(x) := x^2 + 2\phi_\theta x(1 - x)$ is the function on Figure 3.

Fix $\varepsilon > 0$. Assume that the state switches to $\theta_{t_0} = 1$ at a time where the population has settled on a near-consensus $x_{t_0 - 1} \simeq 0$. As long as $x_t < \varepsilon$, Eq (7) implies that $x_{t+1} \simeq 2\phi_1 x_t + \beta\psi_1$ increases at a speed that hinges on β. For β close to zero, many periods are required until $x_t > \varepsilon$. Since g_1 is bounded away from the diagonal $y = x$ on the interval $[\varepsilon, 1 - \varepsilon]$, it then takes only a finite number of stages, which does not depend on λ, until $x_t > 1 - \varepsilon$.

The transition dynamics from one consensus to the other thus involves two different phases. In a first phase (inertia), the old consensus persists despite the state change: most agents observe a unanimous sample and most likely herd, which slows down society's response. At some point, though, there is enough heterogeneity in the population, and sufficiently many agents draw a more balanced sample – enhancing information acquisition – and the society quickly switches to the new consensus: there is a *domino* effect whereby the popularity of the new action snowballs.

Since, as $\lambda \to 0$, society is almost always in consensus, and information is acquired with vanishing probability, equilibrium welfare is captured by the fraction of time spent in a wrong consensus; in other words, it is measured by the duration of the phase of inertia.

4.3 Equilibrium welfare

We now focus on equilibrium welfare, and examine how our incomplete learning result extends to larger samples. For simplicity, we assume here that signals are binary with precision $\pi > \frac{1}{2}$.

Theorem 3 *If $\hat{p} = \pi - c$ is high enough, and λ is small enough, there exists an equilibrium in which $p_k \in [1 - \hat{p}, \hat{p}]$ for all k : agents always acquire information.*

Proof insights:. We focus on strategies that always acquire information with a non-unanimous sample ($\beta(k) = 1$ for all $k \notin \{0, n\}$) and denote $\beta := \beta(0) = \beta(n)$ the probability of acquiring information with a unanimous sample.[16]

[15] ψ_θ is the probability of choosing action 1 when acquiring information with an interim belief of $1 - \hat{p}$, conditional on the current state being θ.

[16] If $c = 0$, β is the probability of following one's private signal.

We assume that π is high enough so that $n^2\pi(1-\pi) < 1$. As we show, this implies the existence of $\beta > 0$ s.t. $p_n = 1 - p_0 = \hat{p}$. We next show that, provided $\hat{p}$ is high enough, all other interim beliefs are contained in the interval $[1-\hat{p}, \hat{p}]$, which completes the proof.

Using (5) and (6) for such strategies and binary signals, (x_t) obeys $x_t = g_{\theta_t}(x_{t-1})$, with

$$g_\theta(x) = \pi_\theta + (1-\beta)\left\{x^n(1-\pi_\theta) - (1-x)^n \pi_\theta\right\}, \text{ where } \pi_1 = \pi = 1 - \pi_0.$$

If $\beta = 0$, the analysis of the case $n = 2$ (which holds verbatim) implies society is eventually trapped in an irreversible consensus and $p_n = p_0 = \frac{1}{2}$. If $\beta = \lambda^m$ for an arbitrary m, $p_n \to 1$ as $\lambda \to 0$. Indeed, whenever the state switches to $\theta = 1$, (x_t) jumps above $g_1(0) = \beta\pi$ and then increases at a rate of $n\pi > 1$. It thus escapes some fixed neighborhood of 0 in $\ln\frac{1}{\lambda}$ stages and next approaches 1 in boundedly many stages unless the state switches back to $\theta = 0$. For λ small, state changes occur on average every $\frac{1}{\lambda} \gg \ln\frac{1}{\lambda}$ stages; at the steady state, x_t and θ_t are then close to perfectly correlated. This contradicts $\beta > 0$. By a continuity argument, there is some β_λ such that $p_n = 1 - p_0 = \hat{p}$.

Such a β_λ is an equilibrium iff $p_k \in (1-\hat{p}, \hat{p})$ for each $k \notin \{0, n\}$. Our methodological contribution lies in providing estimates of such p_k as $\lambda \to 0$. Bayes rule writes $\dfrac{p_k}{p_{n-k}} = \dfrac{I_k}{I_{n-k}}$, with $I_k := \displaystyle\int_0^1 \psi_k(x)\, dF_1(x)$, where $\psi_k(x) = x^k(1-x)^{n-k}$ and F_θ is the invariant cdf of x conditional on θ.[17] The argument relies on approximating I_k with $\lambda \displaystyle\sum_{t\in} \psi_k(g_1^{(t)}(x))$, where $g_1^{(t)}$ is the t-th iterate of g_1 and $(g_1^{(t)}(x))_{t\in}$ is a doubly infinite orbit of g_1. This sum is similar to the time-average of ψ_k during a visit to state $\theta = 1$, hence this approximation is reminiscent of the law of large numbers.

The proof uses the following observation on F_1. As $\lambda \to 0$, the probability of a state switch in m periods is vanishingly small, hence the distribution F_1 around $x \in (0,1)$ coincides with the push-forward measure (given g_1) of F_1 around $g_1^m(x)$ for any fixed m. Specifically, we show that $\lim_{\lambda\to 0}\frac{1}{\lambda}(F_1(y) - F_1(x)) = \lim_{\lambda\to 0}\frac{1}{\lambda}\left(F_1\left(g_1^{(m)}(y)\right) - F_1\left(g_1^{(m)}(x)\right)\right)$ and is non-zero, for any $x, y \in (0,1)$. This yields $I_k \simeq \lambda\int_z^{g_1(z)} \sum_{t\in}\psi_k\left(g_1^{(t)}(x)\right) dF_1(x)$ for some z. In the limit $\lambda \to 0$, this leads to

$$\frac{p_k}{p_{n-k}} \le \sup_{x\in(0,1)} \frac{\sum_{t\in}\psi_k(\bar{g}_1^{(t)}(x))}{\sum_{t\in}\psi_{n-k}(\bar{g}_1^{(t)}(x))} \text{ for each } k \notin \{0, n\}, \tag{8}$$

where $\bar{g}_1(\cdot)$ is the function $g_1(\cdot)$ in the specific case where $\beta = 0$. It is straightforward to show that the RHS in (8) is uniformly bounded, which guarantees that $p_k \in (1-\hat{p}, \hat{p})$ for all $k \notin \{0, n\}$, provided $\hat{p}$ is high enough. ■

Since agents always acquire information with positive probability, equilibrium welfare equals $\pi - c = \hat{p}$, as when $n = 0, 1, 2$. Theorem 3 is reminiscent of the case $n = 2$ because information

[17] We recall that the cdf F_θ depends on the transition rate λ, and on the acquisition strategy β_λ.

is acquired following all sample realizations.The underlying reason is similar: if information was not acquired at unanimous samples, society would converge towards a consensus faster than away, making such consensus irreversible, hence uninformative. With $n > 2$, though, the chances of observing a dissenting action in one's sample are higher (all else equal), which lowers the relative speed of moving towards the correct consensus. If π is large, this convergence remains excessively fast, the intuition being that signals are then more correlated with the true state and among each other.[18] Actions are less diverse, accelerating convergence.

Interestingly, the upper and lower bounds on p_2/p_1 derived from (8) are quite close if $n = 3$, allowing for precise estimates of beliefs.[19] These equilibrium values of p_1 and p_2 are pictured on the right panel of Figure 4, as a function of π. Possibly surprisingly, there is some belief non-monotonicity, $p_1 > p_2$.

Some intuition can be found in the left and central panels of Figure 4, which assume $\pi = 0.99$ for concreteness. The left panel features the functions g_θ that capture the limit dynamics of the population in each state. Note that $g_0'(0)$ is close to 0, hence the population converges quickly towards a consensus on action 0 when $\theta = 0$. The dynamics *away* from 0 is not nearly as fast when $\theta = 1$. As a result, conditional on x being relatively low, society is more likely to be transitioning away from 0 than towards 0. Hence, the current state is more likely to be $\theta = 1$.[20] This is further illustrated by the central panel, where we plot the logs of (simulated) steady-state, equilibrium densities in the limit $\lambda \to 0$. We note that this belief reversal $p_1 - p_2$ increases with π, in line with the intuition that convergence to the correct consensus gets relatively faster.

4.4 The planner's problem

One key question is whether our incomplete learning result is an equilibrium feature or an inescapable feature of the environment. It turns out that a social planner who could dictate any strategy would attain a steady-state welfare of 1 in the persistent limit, so the learning failure is an equilibrium phenomenon. Consider *e.g.* a (symmetric) strategy in which the probability of acquiring information is of order λ for $k = 0$, is 1 for $k = \lfloor \frac{n}{2} \rfloor$ and increases linearly with $k \in \{0, \dots, \frac{n}{2}\}$. Such a strategy ensures that the duration of the phase of inertia is of order $\ln 1/\lambda$, and therefore negligible compared to $1/\lambda$. This ensures that the prevailing consensus is always correct in the limit $\lambda \to 0$. At the same

[18] This is why we require π to be large, unlike in Proposition 2. In addition, the assumption that c is low helps guarantee that $p_k \in (1 - \hat{p}, \hat{p}_k)$ for $k \notin \{0, n\}$. No such condition is required in Proposition 2 since $p_1 = \frac{1}{2}$.

[19] For larger values of n, the spread between the two bounds increases.

[20] This intuition is incomplete. Indeed, for $\lambda > 0$, agents with a unanimous sample do acquire information with positive probability, fostering the responsiveness of the population. It is not *ex ante* obvious that interim beliefs can be estimated *as if* such agents were *not* acquiring information, as (8) suggests.

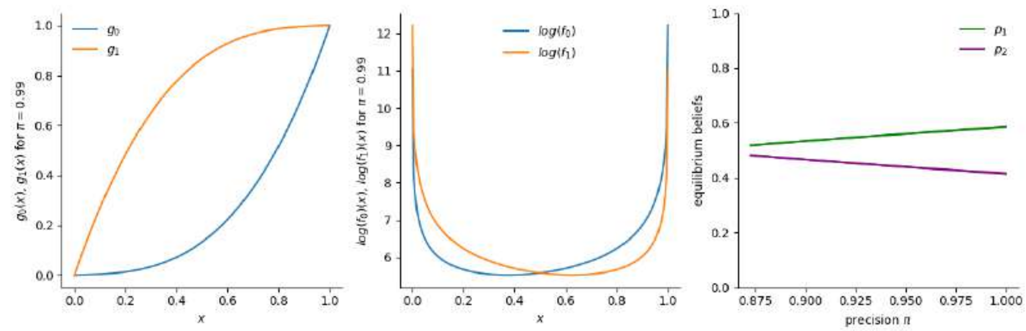

Figure 4: Transition functions, densities, and beliefs for $n = 3$

time, the expected fraction of agents acquiring information vanishes as $\lambda \to 0$.[21]

5 Discussion

5.1 Further evidence for the case $n = 3$: the case of weak signals

The formal analysis of the case $n = 3$ is beyond the scope of the paper, and we limit us to numerical evidence. When $n = 3$, strategies can be summarized by three variables: the probabilities $\beta(0)$ and $\beta(1)$ of acquiring information with a unanimous sample and with a more balanced sample respectively, and the action $\alpha(1, \frac{1}{2})$ played when information is not acquired at balanced samples. [22] For each type of strategy, the proof techniques of Theorem 3 can be adapted to get estimates of the limit interim beliefs, which can then be used to test equilibrium conditions,[23] as a function of $(\pi, \hat{p})$.

The results are summarized on Figure 5. For each strategy type, the left panel draws the region of $(\pi, \hat{p})$-values that are consistent with equilibrium conditions. The absence of overlap between these regions strongly suggests that equilibrium is unique. The equilibria identified in Theorem 3 correspond to the plain yellow area ($\beta(0) > 0$ and $\beta(1) = 1$) in the upper-right corner. The lower-left corner shows that for weak signals, agents herd with a unanimous sample, and follow the majority action whenever they choose not to acquire information ($\alpha(1, \frac{1}{2}) = 0$): there is no belief reversal for such π. By contrast, we note the existence of a shaded green area in the lower right corner (high π, high c), where agents herd with a unanimous sample, but adopt a *contrarian*

[21]See the working paper version (Lévy, Pęski and Vieille, 2022) for details.

[22]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.

[23]Details are in the supplementary material.

behavior when facing mixed evidence and not acquiring information ($\alpha(1, \frac{1}{2}) = 1$). This contrarian behavior helps the population to adjust to changes when signals are strong.

The right panel of Figure 5 draws the equilibrium welfare as a function of π, for various values of c. We observe, e.g. from the blue curve ($c = 0$), that the equilibrium welfare may exceed $\hat{p}$ for weak signals, but that learning is always incomplete, consistent with our main message. We also note that for stronger signals, the equilibrium welfare is given by $\hat{p}$, irrespective of c.

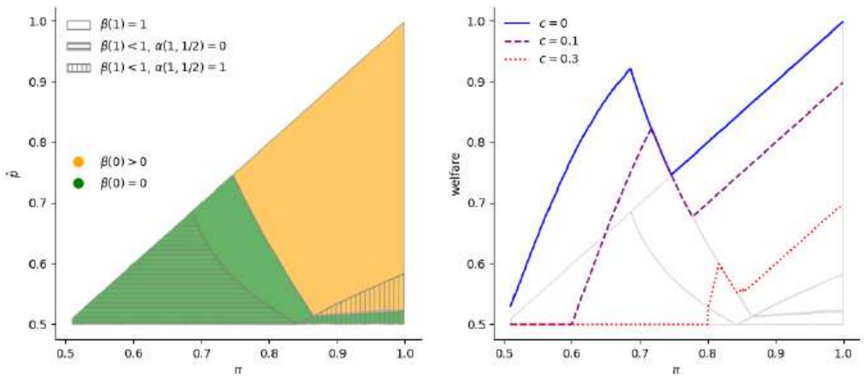

Figure 5: Equilibrium behavior (left) and welfare (right) for $n = 3$

Interestingly, we observe that welfare is both decreasing in π and independent of c as long as $(\pi, \hat{p})$ belongs to the plain green region. There, agents never acquire information at unanimous samples, and always acquire otherwise. Therefore, equilibrium beliefs do not vary with c. Instead, when π increases, information acquisition is unchanged, but the correlation of actions within samples increases, lowering informativeness and then decreasing welfare. This suggests that more precise signals always hurt as long as they do not encourage more information acquisition. This is an instance of the *principle of countervailing adjustment* (Bikhchandani et al., 2021): a favorable shift in information availability does not necessarily improve average decisions or welfare.[24]

5.2 The case $\hat{p} = 1$

We briefly review here the case $\hat{p} = 1$ where signals are free and unbounded. In such case, the prevailing consensus is asymptotically correct in the limit $\lambda \to 0$, as Proposition 3 shows. This

[24]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.

result is line with the fixed-state literature (see, e.g., Smith and Sørensen (2000); Banerjee and Fudenberg (2004)), in contrast to the case $\hat{p} < 1$.

Proposition 3 *Suppose $\hat{p} = 1$ and let $n \geq 1$ be arbitrary. For $\lambda > 0$, let any equilibrium steady state be given, with payoff w^*. Then $\lim_{\lambda \to 0} w^* = 1$.*

Proof. As before, we list the actions sampled by a random agent in a random order $a^{(1)}, \ldots, a^{(n)}$. Let $p^{(1)}$ be the interim belief formed on the basis of $a^{(1)}$ only. One strategy consists of replicating $a^{(1)}$, with a payoff of $u(p^{(1)})$. Since $a^{(1)}$ matches yesterday's state with probability w^*, this strategy yields a payoff $u(p^{(1)}) = (1-\lambda)w^* + \lambda(1-w^*) \geq w^* - \lambda$. Another strategy consists in ignoring all sampled actions except $a^{(1)}$ and using one's (free) signal optimally. This strategy yields a payoff $v(p^{(1)})$, which is no larger than the equilibrium payoff w^*. Thus,

$$u(p^{(1)}) \leq v(p^{(1)}) \leq w^* \leq u(p^{(1)}) + \lambda. \tag{9}$$

Since $\hat{p} = 1$, $v(p) > u(p)$ for all $p \in (0,1)$ and $v(p) = u(p) = 1$ for $p \in \{0,1\}$. Together with (9), and using the continuity of the functions u and v, this implies $\lim_{\lambda \to 0} p^{(1)} \in \{0,1\}$ and $\lim_{\lambda \to 0} w^* = 1$. ∎

5.3 Continuum of actions

While we have considered binary actions so far, we provide suggestive evidence that the discontinuity in information aggregation at the limit $\lambda \to 0$ could arise as well with a richer (infinite) action set.

For concreteness, assume $A = [0,1]$ and a square loss utility function $u(a,\theta) = 1 - (a-\theta)^2$. Note that, if $c = 0$, one has $\hat{p} = 1$ even with bounded signals because actions are responsive to any extra information; we thus assume $c > 0$ to stick to the case $\hat{p} < 1$.

In the fixed-state version where a continuum of new agents sample at least two actions from the past, the distribution of actions converges over time to the correct action (see *e.g.* (Lee, 1993).) For the evolving-state version, Proposition 4 below shows that the average dispersion of actions in the population vanishes as $\lambda \to 0$, thereby extending the consensus result of Theorem 2. In this statement, $a^{(1)}, \ldots, a^{(n)}$ are the actions sampled by a random agent, listed in a random order. The proof is in Appendix C.

Proposition 4 *Assume $n \geq 2$. At any equilibrium steady state, one has*

$$\mathbf{E}\left[(a^{(k)} - a^{(l)})^2\right] \leq \gamma\lambda \text{ for any } k, l,$$

where γ is independent of λ and of the equilibrium.

A full-blown analysis of equilibrium behavior and welfare is highly challenging, and beyond the scope of the paper anyway. Without aiming at generality, we discuss here the example of *perfect* signals, which yields results consistent with our incomplete learning result. If the state is fixed, all agents acquire information in the first period and convergence to the truth takes exactly one period. When the state is evolving, the population in period t is described by the pair (θ_t, x_t), where $x_t \in \Delta([0,1])$ is the distribution of actions in the population. Accordingly, an equilibrium steady state is a distribution $\mu \in \Delta\left(\Theta \times \Delta([0,1])\right)$.

We describe an equilibrium where all agents in period t choose the *same* action $a_t \in [0,1]$: actions are perfectly correlated, and inferences are independent of the sample size n.

In period $t+1$, agents then hold the interim belief $f(a_t) := (1-\lambda)a_t + \lambda(1-a_t)$. If $f(a_t) \notin [1-\hat{p}, \hat{p}]$, all agents choose the action $a_{t+1} = f(a_t)$. If instead $f(a_t) \in [1-\hat{p}, \hat{p}]$, all agents acquire information, learn the state and choose $a_{t+1} = \theta_{t+1} \in \{0,1\}$. Either way, the consensus is preserved in period $t+1$. In this equilibrium, agents acquire information periodically: when agents learn that the current state is, say, $\theta_t = 1$, their actions are $a_t = 1$, $a_{t+1} = f(1)$, etc., until they acquire information in period $t + M_\lambda$, where $M_\lambda := \left\lfloor \dfrac{\ln(2\hat{p}-1)}{\ln(1-2\lambda)} \right\rfloor$. The marginal $\mu_2 \in \Delta\left(\Delta([0,1])\right)$ over action distributions is uniform over the degenerate distributions $\delta_{f^m(\theta)}$ $(0 \leq m < M_\lambda,\ \theta \in \Theta)$.

As $\lambda \to 0$, μ_2 weakly converges to a distribution concentrated over degenerate distributions δ_a, $a \in [0,1]$: in the persistent limit, actions are always perfectly correlated within periods, and the consensus action a has a density, which is given by $\dfrac{1}{2a-1}\left(1_{[\hat{p},1]}(a) - 1_{[0,1-\hat{p}]}(a)\right)$:[25] there is no complete learning.

5.4 Endogenous information structures

We have assumed that agents have access to a fixed information source at a lumpy cost, but the choice of this source could be *endogenized.* For instance, 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

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

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 $\hat{p} < 1$ The answer depends on the functional $C(\cdot)$ one considers. In the supplementary material, we show that $\hat{p} < 1$ for the functional $C(\mu) =$

[25] Up to a normalization constant. See the supplementary material.

$\sum_{a\in A,\theta\in\Theta} \mu_\theta(a) \ln \frac{\mu_\theta(a)}{\mu_{1-\theta}(a)}$ that is axiomatically derived by Pomatto, Strack and Tamuz (2023).

5.5 Asymmetric case

The symmetry assumption in the paper is quite convenient, but plays no specific role. Consider a general setup where the invariant probability of state 1 is $p^* \in (0,1)$,[26] the utility is an arbitrary function $u : \{0,1\} \times \Theta \to [0,1]$, and the (unconditional) distribution of private beliefs is an arbitrary distribution $H \in \Delta([0,1])$ with expectation $\frac{1}{2}$.

There exist cutoffs $0 \le p_0^* \le p_1^* \le 1$ such that $v(p) > u(p)$ iff $p_0^* < p < p_1^*$. The results of the paper extend under the assumption that $0 < p_0^* < p^* < p_1^* < 1$. This is a joint assumption on all primitives of the model. The assumption that $p_0^*, p_1^* \in (0,1)$ states that signals are either bounded or costly (counterpart of Assumption 2). The assumption that $p_0^* < p^* < p_1^*$ ensures that any steady-state equilibrium entails some information acquisition (counterpart of Assumption 1). The detailed statements are in the supplementary material.

5.6 Sampling from the further past

Our analysis seamlessly accommodates situations where actions are sampled from the more distant past. Specifically, assume that in period t past actions are each sampled from a random, possibly different, period $t-\tau$, where the lag $\tau \ge 1$ 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 above. 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.[27]

6 Relation to the literature

Within the wide literature on observational learning, our paper more specifically connects to four themes.[28]

[26] The probability that the state changes from $\theta = 0$ to $\theta = 1$ (resp., from 1 to 0) is λp^* (resp., $\lambda(1-p^*)$). REMOVE?

[27] Note that one must then distinguish the prevalence of action 1 in the existing population at time t, from the fraction of new agents choosing action 1. See the working paper version (Lévy, Pęski and Vieille, 2022) for details.

[28] For more complete surveys of the literature, see Smith and Sørensen (2011) and Bikhchandani et al. (2021).

Random sampling. Banerjee and Fudenberg (2004) and Smith and Sørensen (2020) also assume that agents draw a random (finite) sample from a continuum of past actions and identify conditions under which learning is asymptotically complete when the state is fixed. As these papers note, models with a continuum of agents inherently exhibit better aggregative properties than those with a sequence of agents. Indeed, when agents observe a common set of predecessors in one-agent models, they eventually end up in a cascade if signals are bounded. With a continuum of agents instead, agents update beliefs considering all possible (mutually exclusive) histories, weighted by their chances (Smith and Sørensen, 2020) and cannot get stuck in a cascade. As soon as agents observe two or more actions, they will rely more on their private signals when their sample conveys mixed evidence, which fosters learning. With a fixed state, this logic guarantees complete learning even when signals are bounded and/or costly (Banerjee and Fudenberg, 2004), and our results thus come in stark contrast.

Costly information acquisition. Our modeling of costly information acquisition relates our work to Burguet and Vives (2000) and Ali (2018), who study within one-agent observational learning models whether costly private signals preclude (or not) complete learning when the state is fixed. Burguet and Vives (2000) endogenize the choice of precision, and argue that the completeness of learning is linked to the acquisition of information at beliefs close to certainly, in line with Section 5.4. Likewise, agents in Ali (2018) can choose from a set of experiments, at an idiosyncratic cost. When costs are bounded away from zero, it follows from their main result that learning is incomplete if signals are bounded. Unlike these papers however, we assume a continuum of agents. In this more favorable context, learning is complete when the state is fixed if $n \geq 2$, even if signals are costly (see our extension of Banerjee and Fudenberg (2004) to costly signals in the supplementary material), in contrast to our main results. The example we analyze in Section 5.3 of a model with a continuum of (responsive) actions, where learning is incomplete when the state is evolving, and complete if the state is fixed, reinforces our message: costly signals do not preclude complete learning with a fixed state, but concur to incomplete learning with changing states.

Stationary analyses of social learning. Dasaratha, Golub and Hak (2023) and Kabos and Meyer (2021) also develop stationary analyses of social learning. Dasaratha, Golub and Hak consider a Gaussian environment where agents in a network learn from their neighbors. They show that learning is improved when agents have heterogeneous neighbors who have access to signals of different precision. While we rule out such heterogeneity, our analysis also highlights the adverse welfare impact of an excessive correlation of actions between (symmetric) agents. Kabos and Meyer consider a Markovian environment where past actions may be misrecorded, and investigate whether agents put too much or too little weight on their private information, while we focus on providing bounds on equilibrium welfare.

Rate of learning and the Grossman-Stiglitz paradox. Our results echo the well-known Grossman-Stiglitz paradox (Grossman and Stiglitz, 1980) according to which agents would ignore their individual signals if information was fully aggregated, precluding information aggregation in the first place. In a fixed-state world when asymptotic learning is guaranteed, such a logic imposes that learning be necessarily slow, as shown in Vives (1993). In a social learning context, Harel et al. (2021) also establish slow social learning even if agents observe several actions, because the correlation in the agents' actions arising from social learning reduces the amount of information these actions reveal about the state.[29] While a high correlation reduces the speed of learning with a fixed state, it lowers responsiveness in our changing state environment.

7 Conclusion

We consider a general model of social learning with binary actions and states in which states change over time, information is possibly costly, and agents draw finite samples of past actions. We show that, under a wide range of situations, the possibility that the state changes drastically limits the extent of social learning at the steady state, in crisp contrast to what would happen in analogous fixed-state environments. Beyond this insight, the methods we develop could pave the way to address interesting questions on how a planner would optimally design the learning environment (sampling procedures, feedback given to players,...) to foster the welfare gains from social learning in such changing environments.

References

Acemoglu, Daron, Angelia Nedic, and Asuman Ozdaglar. 2008. "Convergence of rule-of-thumb learning rules in social networks." 1714–1720, IEEE.

Ali, S Nageeb. 2018. "Herding with costly information." *Journal of Economic Theory*, 175: 713–729.

Banerjee, Abhijit, and Drew Fudenberg. 2004. "Word-of-mouth learning." *Games and economic behavior*, 46(1): 1–22.

Bhattacharya, Rabi, and Mukul Majumdar. 2007. *Random dynamical systems: theory and applications.* Cambridge University Press.

[29]See also Huang, Strack and Tamuz (2024) on slow convergence.

Bikhchandani, Sushil, David Hirshleifer, Omer Tamuz, and Ivo Welch. 2021. "Information cascades and social learning." National Bureau of Economic Research.

Burguet, Roberto, and Xavier Vives. 2000. "Social learning and costly information acquisition." *Economic theory*, 15(1): 185–205.

Dasaratha, Krishna, Benjamin Golub, and Nir Hak. 2023. "Learning from neighbours about a changing state." *Review of Economic Studies*, 90(5): 2326–2369.

Frongillo, Rafael M, Grant Schoenebeck, and Omer Tamuz. 2011. "Social learning in a changing world." 146–157, Springer.

Grossman, Sanford J, and Joseph E Stiglitz. 1980. "On the impossibility of informationally efficient markets." *The American economic review*, 70(3): 393–408.

Harel, Matan, Elchanan Mossel, Philipp Strack, and Omer Tamuz. 2021. "Rational groupthink." *The Quarterly Journal of Economics*, 136(1): 621–668.

Huang, Wanying. 2022. "The Emergence of Fads in a Changing World." *arXiv preprint arXiv:2208.14570.*

Huang, Wanying, Philipp Strack, and Omer Tamuz. 2024. "Learning in repeated interactions on networks." *Econometrica*, 92(1): 1–27.

Kabos, Eszter, and Margaret Meyer. 2021. "A Welfare Analysis of a Steady-State Model of Observational Learning." working paper.

Lee, In Ho. 1993. "On the convergence of informational cascades." *Journal of Economic theory*, 61(2): 395–411.

Lévy, Raphaël, Marcin Pęski, and Nicolas Vieille. 2022. "Stationary social learning in a changing environment." arXiv:2201.02122.

Moscarini, Giuseppe, Marco Ottaviani, and Lones Smith. 1998. "Social learning in a changing world." *Economic Theory*, 11(3): 657–665.

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.

Smith, Lones, and Peter Norman Sørensen. 2011. "Observational learning." *The New Palgrave Dictionary of Economics Online Edition*, 29–52.

Smith, Lones, and Peter Norman Sørensen. 2020. "Rational Social Learning with Random Sampling." working paper.

Smith, Lones, and Peter Sørensen. 2000. "Pathological outcomes of observational learning." *Econometrica*, 68(2): 371–398.

Solomyak, Boris. 1995. "On the random series $\sum \pm\lambda^n$ (an Erdos problem)." *Annals of Mathematics*, 611–625.

Vives, Xavier. 1993. "How fast do rational agents learn?" *The Review of Economic Studies*, 60(2): 329–347.

A The case $n = 2$: Proposition 2

We argue by contradiction and assume that $p_2 > \hat{p}$ for some equilibrium steady state. Since agents who observe a balanced sample $k = 1$ hold the belief $p_1 = \frac{1}{2}$ and acquire information, the fraction of agents choosing action 1 in period $t+1$ reads

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

where $\phi_\theta = 1 - H_\theta(\frac{1}{2})$ is the probability of playing action 1 in state θ when holding an interim belief $\frac{1}{2}$. Note that $\phi_1 = 1 - \phi_0 > \frac{1}{2}$.

Lemma 2 *The sequence (x_t) converges a.s., with $\lim_{t\to+\infty} x_t \in \{0, 1\}$.*

Proof. Choose $\tilde{\phi}_\theta > \phi_\theta$ such that $\tilde{\phi}_0\tilde{\phi}_1 < \frac{1}{4}$, and $\varepsilon_0 > 0$ such that $2\phi_\theta + \varepsilon_0 < 2\tilde{\phi}_\theta$ for each θ. Note that $g_\theta(x) \leq 2\tilde{\phi}_\theta x$ for $x < \varepsilon_0$.

Let $\varepsilon < \varepsilon_0$ be arbitrary. We define two increasing and interlacing sequences $(\tau_m^{\text{in}})_m$ and $(\tau_m^{\text{out}})_m$ of possibly infinite stopping times. We first set

$$\tau_1^{\text{out}} := \inf\{t \geq 0 : x_t < \varepsilon \text{ and } \theta_t = 0, \text{ or } x_t > 1 - \varepsilon \text{ and } \theta_t = 1\},$$

and $\tau_1^{\text{in}} := \inf\{t \geq \tau_1^{\text{out}} : x_t \in [\varepsilon, 1 - \varepsilon]\}$, with $\inf \emptyset = +\infty$.

These are the first exit and entry times in $[\varepsilon, 1 - \varepsilon]$.[30] For $m \geq 1$, we set

$$\tau_{m+1}^{\text{out}} := \inf\{t \geq \tau_m^{\text{in}} : x_t < \varepsilon \text{ and } \theta_t = 0, \text{ or } x_t > 1 - \varepsilon \text{ and } \theta_t = 1\},$$

and $\tau_{m+1}^{\text{in}} := \inf\{t \geq \tau_{m+1}^{\text{out}} : x_t \in [\varepsilon, 1 - \varepsilon]\}$.

Below, we show that (x_t) cannot remain indefinitely in the interval $[\varepsilon, 1 - \varepsilon]$..

[30] Except for the extra condition on the exit state in the definition of τ_1^{out}, which is for convenience.

Claim 4 *One has* $\mathbf{P}(\tau_{m+1}^{\text{out}} < +\infty \mid \tau_m^{\text{in}} < +\infty) = 1$ *for each* m.

Proof of the claim. For $x \in [\varepsilon, 1-\varepsilon]$, one has

$$g_1(x) - x = x - g_0(x) = (2\phi_1 - 1)(x - x^2) \geq (2\phi_1 - 1)\,\varepsilon(1-\varepsilon) : \tag{A.1}$$

the difference $x_{t+1} - x_t$ is bounded away from zero. With $N := \lceil \frac{1}{\varepsilon((1-\varepsilon)2\phi_1 - 1)} \rceil$ and using (A.1), it follows that $x_t \in [\varepsilon, 1-\varepsilon]$ implies $x_{t+N} \notin [\varepsilon, 1-\varepsilon]$ as soon as $\theta_{t+1} = \cdots = \theta_{t+N}$, which has probability $(1-\lambda)^N$. This implies

$$\mathbf{P}\left(\tau_{m+1}^{\text{out}} \leq t + N \mid \tau_m^{\text{in}} \leq t < \tau_{m+1}^{\text{out}}\right) \geq (1-\lambda)^N,$$

and therefore,

$$\mathbf{P}\left(\tau_{m+1}^{\text{out}} \geq t + jN \mid \tau_m^{\text{in}} \leq t < \tau_{m+1}^{\text{out}}\right) \leq \left(1 - (1-\lambda)^N\right)^j$$

for each j. The result follows when $j \to +\infty$. ■

We show that the probability that x_t ever re-enters the interval $[\varepsilon, 1-\varepsilon]$ after it exits from it, is bounded away from 1. In the next statement, $(\mathcal{H}_t)_t$ is the natural filtration of $(\theta_t, x_t)_t$ and $\mathcal{H}_{\tau_m^{\text{out}}}$ is the stopped filtration at time τ_m^{out}.

Claim 5 *There exists* $a > 0$ *such that for each* m, $\mathbf{P}\left(\tau_m^{\text{in}} = +\infty \mid \mathcal{H}_{\tau_m^{\text{out}}}\right) \geq a$, *w.p. 1 on the event* $\tau_m^{\text{out}} < +\infty$.

Proof of the claim. Consider the event $\tau_m^{\text{out}} = t$. We assume for concreteness that $x_t < \varepsilon$ and $\theta_t = 0$. By the Markov property, we may assume w.l.o.g. that $t = 0$ and $m = 1$.

We define an auxiliary sequence (W_t) of random variables by $W_0 = 0$ and $W_{t+1} = W_t + \ln 2\tilde{\phi}_{\theta_{t+1}}$ for $t \geq 1$. Since $x_{t+1} \leq 2x_t\tilde{\phi}_{\theta_{t+1}}$, one has $W_t \geq \ln x_t - \ln x_0$ for each t. This implies that $\tau_1^{\text{in}} \geq \inf\{t \geq 1 : W_t \geq 0\}$, and $\mathbf{P}(\tau_1^{\text{in}} < +\infty) \leq \mathbf{P}\left(\sup_{t\geq 1} W_t \geq 0\right)$. Introduce the successive state changes $\psi_0 = 0$ and $\psi_{m+1} := \inf\{t > \psi_m : \theta_t \neq \theta_{t-1}\}$. Assuming $\theta_0 = 0$, $\theta_t = 1$ whenever $\psi_{2m-1} \leq t < \psi_{2m}$ for some m, and $\theta_t = 0$ otherwise.[31]

For $i \in \mathbf{N}$, set $X_i := W_{\psi_{2i+2}} - W_{\psi_{2i}}$. Observe that $W_t - W_{t-1} > 0$ iff $\theta_t = 1$, hence

$$\sup_t W_t \geq 0 \Leftrightarrow \sup_i\,(X_0 + \cdots + X_i) \geq 0.$$

The r.v.'s (X_i) are *i.i.d.* with $\mathbf{E}[X_1] = \frac{1}{\lambda}\left(\ln 2\tilde{\phi}_1 + \ln 2\tilde{\phi}_0\right) < 0$. The sequence $(X_0 + \cdots + X_j)_j$ is therefore a simple random walk with negative drift, which implies

$$\mathbf{P}\left(\sup_j\,(X_1 + \cdots + X_j) \geq 0\right) \leq 1 - a \text{ for some } a > 0.$$

[31]If $\theta_0 = 1$, odd and even phases should be switched.

■

Claims 4 and 5 yield $\mathbf{P}\left(\tau_m^{\text{out}} < +\infty \text{ and } \tau_{m+1}^{\text{in}} = +\infty \text{ for some } m\right) = 1$. Hence there is a.s. finite random time T_0 such that either $x_t < \varepsilon$ for all $t \geq T_0$, or $x_t > 1 - \varepsilon$ for all $t \geq T_0$. ■

Lemma 3 *The only symmetric invariant measure for (θ_t, x_t) is uniform over $\Theta \times \{0, 1\}$.*

Proof. By Lemma 2, any invariant measure is concentrated on $\Theta \times \{0, 1\}$.[32] Since the sets $\{x = 0\}$ and $\{x = 1\}$ are absorbing for (θ_t, x_t), one has for $a \in \{0, 1\}$

$$\begin{aligned}
\mu(0, a) &= \mathbf{P}\left((\theta_{t+1}, x_{t+1}) = (0, a)\right) \\
&= (1 - \lambda)\mathbf{P}\left((\theta_t, x_t) = (0, a)\right) + \lambda \mathbf{P}\left((\theta_t, x_t) = (1, a)\right) \\
&= (1 - \lambda)\mu(0, a) + \lambda\mu(1, a).
\end{aligned}$$

Hence $\mu(0, a) = \mu(1, a)$: μ is a product distribution. Since μ is symmetric, it is uniform. ■

The result follows. By Lemma 3, $p_2 = \frac{1}{2}$ – a contradiction.

B No Social Learning: Theorem 3

B.1 Notation and Preliminaries

We let $n \geq 3$, $c \geq 0$ and $\pi > \frac{1}{2}$ be given and assume throughout that $n^2\pi(1 - \pi) < 1$. We consider strategies such that agents (i) acquire information w.p β when sampling $k = 0$ or $k = n$ and w.p. 1 otherwise, (ii) follow their signal when they acquire information and (iii) follow the consensus action when they don't. Given β, the population state evolves according to $x_{t+1} = g^\beta_{\theta_{t+1}}(x_t)$, where

$$g^\beta_\theta(x) = \pi_\theta + (1 - \beta)\left\{x^n\left(1 - \pi_\theta\right) - (1 - x)^n \pi_\theta\right\} \qquad \text{(with } \pi_1 = \pi = 1 - \pi_0\text{)}.$$

We note that g^β_θ is an increasing bijection from $[0, 1]$ to $[\pi_\theta\beta, \pi_\theta + (1 - \beta)(1 - \pi_\theta)]$, and that there is a unique $\bar{x}^\beta_\theta$ such that $g^\beta_\theta(\bar{x}^\beta_\theta) = \bar{x}^\beta_\theta$. In addition, g^β_1 is concave on $[0, 1]$, $(g^\beta_1)'$ is convex on $[0, 1]$ and $g^\beta_1(x) - x$ is decreasing on $[\frac{1}{2}, 1]$.

We denote by $h^\beta_\theta : [g_\theta(\bar{x}^\beta_0), g_\theta(\bar{x}^\beta_1)] \to [\bar{x}^\beta_0, \bar{x}^\beta_1]$ the inverse of g^β_θ on $[g_\theta(\bar{x}^\beta_0), g_\theta(\bar{x}^\beta_1)]$. We will view h^β_θ as a function defined on $[\bar{x}^\beta_0, \bar{x}^\beta_1]$ by setting

$$h^\beta_0(x) = \bar{x}^\beta_1 \text{ for } x > g^\beta_0(\bar{x}^\beta_1) \text{ and } h^\beta_1(x) = \bar{x}^\beta_0 \text{ for } x < g^\beta_1(\bar{x}^\beta_0).$$

Given λ, any invariant measure $\mu_{\lambda,\beta}$ for the strategy β is concentrated on $\Theta \times (\bar{x}^\beta_0, \bar{x}^\beta_1)$. We denote by $F^{\lambda,\beta}_\theta(x) := 2\mu_{\lambda,\beta}(\{\theta\} \times [0, x])$ the cdf of the population state given θ.

[32] Indeed, by the invariance property, $\mu(\Theta \times [\varepsilon, 1 - \varepsilon]) = \mathbf{P}\left(x_t \in [\varepsilon, 1 - \varepsilon]\right)$ for each t. By Lemma 2, the RHS converges to zero as $t \to +\infty$ for each $\varepsilon > 0$.

To avoid clumsy notation, we henceforth omit the superscripts β and λ and simply write $\bar{x}_\theta, \mu, F_\theta, g_\theta$ and h_θ. Similarly, all interim beliefs p_k will depend on β and λ, although the notation does not show. The m-th iterate of g_θ is denoted g_θ^m.

Given a (symmetric) invariant measure μ for β, time invariance that for $x \in [\bar{x}_0, \bar{x}_1]$,

$$F_\theta(x) = (1-\lambda)F_\theta(h_\theta(x)) + \lambda F_{1-\theta}(h_{1-\theta}(x)) \text{ for each } \theta \in \Theta. \tag{C.1}$$

Lemma 4 *For each $x \in [\bar{x}_0, \bar{x}_1]$ and $\theta \in \Theta$, one has*

$$|F_\theta(x) - F_\theta(h_\theta(x))| \leq \lambda \text{ and } |F_\theta(x) - F_\theta(g_\theta(x))| \leq \lambda$$

Proof. The first claim is a rewriting of (C.1). The second follows from the first when applied to $g_\theta(x)$. ■

Lemma 5 *Let $\varepsilon > 0$ be given. There exists k such that for each β small enough, the following holds. For each $x \geq \varepsilon$, $y \in [x, g_1(x)]$, $\lambda > 0$ and $m \in \mathbf{N}$, one has*

$$\left| \frac{1}{\lambda} \left(F_1\left(g_1^m(y)\right) - F_1\left(g_1^m(x)\right) \right) - \frac{1}{\lambda}\left(F_1(y) - F_1(x)\right) \right| \leq mk\lambda.$$

In addition, for each $x, y \in (0,1)$, one has

$$\lim_{\lambda \to 0} \left(F_1(y) - F_1(x)\right) = 0.$$

Proof. Let $\varepsilon > 0$, and assume that $n\pi\beta \leq \frac{1}{2}\varepsilon$. There exists k s.t. $g_0^k(g_1(x)) \leq x$, hence $h_0(g_1(x)) \leq h_0^{k+1}(x)$, for each $x \geq \varepsilon$. By Lemma 4,

$$F_0\left(h_0(y)\right) \leq F_0(h_0\left(g_1(x)\right) \leq F_0\left(h_0^{k+1}(x)\right) \leq F_0(h_0(x)) + k\lambda$$

hence

$$F_0\left(h_0(y)\right) - F_0(h_0(x)) \leq k\lambda. \tag{C.2}$$

Applying (C.1) repeatedly, we obtain:

$$\begin{aligned}
&F_1\left(g_1^m(y)\right) - F_1\left(g_1^m(x)\right) \\
=&(1-\lambda)\left(F_1\left(g_1^{m-1}(y)\right) - F_1\left(g_1^{m-1}(x)\right)\right) + \lambda\left(F_0\left(h_0\left(g_1^m(y)\right)\right) - F_0\left(h_0\left(g_1^m(x)\right)\right)\right) \\
=&(1-\lambda)^m\left(F_1(y) - F_1(x)\right) + \lambda\sum_{i=1}^{m}(1-\lambda)^l\left(F_0\left(h_0\left(g_1^i(y)\right)\right) - F_0\left(h_0\left(g_1^i(x)\right)\right)\right) \\
\leq&F_1(y) - F_1(x) + mk\lambda^2,
\end{aligned}$$

where the last inequality follows from (C.2), since $g_1^i(y) \in [g_1^i(x), g_1\left(g_1^i(x)\right)]$ for each i.

Under the same conditions on β, for given x and y there exists m such that $y \leq g_1^m(x)$. Using Lemma 4, one has $F_1(x) \leq F_1(y) \leq F_1(x) + k\lambda$ for each $\lambda > 0$. ■

B.2 The choice of β

Given λ small, we prove here the existence of a strategy β_λ, such that agents with a unanimous sample are indifferent whether to acquire information.[33]

Proposition 5 *For any $m > 0$ and λ small enough, there exists $\beta_\lambda \leq \lambda^m$, and an invariant measure for β_λ, such that $p_n = 1 - p_0 = \hat{p}$.*

The result follows from Lemmas 6, 7 and 9, and from the fact that the set of symmetric invariant measures is convex-valued and upper hemi-continuous as a function of $\beta \in [0, 1]$.

Lemma 6 *Let $\beta = 0$. Then $p_n = p_0 = \frac{1}{2}$, for each λ.*

Proof. Since $n^2\pi(1-\pi) < 1$, the proof in Appendix A. In particular, the statements of Lemmas 2 and 3 and Claims 4 and 5 from the Appendix remain true verbatim. The only (symmetric) invariant measure is the uniform distribution over $\Theta \times \{0, 1\}$. ■

Lemma 7 *Let $m > 0$ be given. For $\lambda > 0$, set $\beta = \lambda^m$. Then $\lim_{\lambda \to 0} p_n = 1$.*

Proof. Observe that $g_1(\bar{x}_0) \geq \beta\pi = \lambda^m\pi$. On the other hand, since $g_1'(0) = (1-\beta)n\pi > 1$, there exists $\varepsilon > 0$ and $a_1 > 1$ such that $g_1(x) \geq a_1 x$ for every $x \in [\bar{x}_0, \varepsilon]$ and λ small enough. Hence there exists M[34] such that $g_1^i(\bar{x}_0) > \varepsilon$ for each $i \geq M \ln \frac{1}{\lambda}$. Using Lemma 4, it follows that

$$F_1(\varepsilon) \leq F_1(\bar{x}_0) + \sum_{i=0}^{\lfloor M \ln \frac{1}{\lambda} \rfloor} F_1\left(g_1^{i+1}(\bar{x}_0)\right) - F_1\left(g_1^i(\bar{x}_0)\right) \leq M\lambda \ln \frac{1}{\lambda}$$

hence $\lim_\lambda F_1(\varepsilon) = 0$. Using Lemma 5, this implies $\lim_\lambda F_1(x) = 0$ for each $x < 1$ and by symmetry, $\lim_\lambda F_0(x) = 1$ for each $x > 0$. The result follows. ■

B.3 Estimates on F

From now on, we set $\beta = \beta_\lambda$. We here derive further estimates on the invariant measure.

Lemma 8 *For each $x \in (0, 1)$, one has*

$$\lim_{\lambda \to 0} \frac{1}{\lambda} \left(F_1(g_1(x)) - F_1(x)\right) = 2\hat{p} - 1.$$

[33] The existence result requires no assumption on $\hat{p}$, beyond $\hat{p} \in (\frac{1}{2}, 1)$, but the value of β_λ of course depends on $\hat{p}$.

[34] independent of λ

Proof. Let $\varepsilon > 0$ be given. By Lemma 4, one has

$$\lim_{\lambda} \left(\int (1-x)^n dF_\theta(x) - F_\theta(\varepsilon) \right) = 0.$$

Since $\dfrac{\int (1-x)^n dF_1(x)}{\int (1-x)^n dF_0(x)} = \dfrac{p_0}{1-p_0} = \dfrac{1-\hat{p}}{\hat{p}}$, this implies that $F_\theta(\varepsilon)$ is bounded away from zero as $\lambda \to 0$, with $\lim \frac{F_1(\varepsilon)}{F_0(\varepsilon)} = \frac{1-\hat{p}}{\hat{p}}$.

On the other hand, symmetry and Lemma 4 imply that $\lim_\lambda F_1(\varepsilon) = 1 - \lim_\lambda F_0(\varepsilon))$. Thus, $\lim_\lambda F_1(\varepsilon) = \lim_\lambda F_1(x) = 1 - \hat{p}$ for each $x \in (0,1)$, and $\lim_\lambda F_0(\varepsilon) = \lim_\lambda F_0(x) = \hat{p}$ for each $x \in (0,1)$. It follows that

$$\lim_{\lambda \to 0} \frac{1}{\lambda} \left(F_1\left(g_1(x)\right) - F_1\left(x\right) \right) = \lim_{\lambda} (F_0(h_0(x)) - F_1(h_1(x)) = 2\hat{p} - 1.$$

■

Lemma 9 *For each $\varepsilon > 0$ and θ, one has* $\limsup_{\lambda \to 0} \left(\lambda\varepsilon + \lambda^2 \right)^{-1} \int_0^\varepsilon x dF_\theta(x) < \infty$.

Proof. Fix $\varepsilon > 0$, and assume for concreteness $\theta = 0$. The other case follows a similar logic. Note first that $\int_0^\varepsilon x dF_0(x) \leq \beta_\lambda \lambda^{-1} + \int_{\beta_\lambda \lambda^{-1}}^\varepsilon x dF_0(x)$. Since $\beta_\lambda \leq \lambda^3$, the first integral is at most λ^2 for λ small.

On the other hand, note that $g_0(x) \leq \bar{x}_0 + \left(\max_{[\bar{x}_0, x]} g_0' \right) (x - \bar{x}_0)$ for each $x \in [\bar{x}_0, \bar{x}_1]$.[35] Since $g'(0) = (1-\beta_\lambda)n(1-\pi) < 1$, and provided ε is small enough, there exists $a_0 < 1$[36] such that $g_0(x) < a_0 x$ for each $x \in [\bar{x}_0, \varepsilon]$ and λ small. Using Lemma 4, one therefore has

$$\begin{aligned} \int_{\beta_\lambda \lambda^{-1}}^\varepsilon x dF_0(x) &\leq \sum_{m : g_0^m(\varepsilon) > \beta_\lambda \lambda^{-1}} g_0^m\left(\varepsilon\right) \left[F_0\left(g_0^m\left(\varepsilon\right)\right) - F_0\left(g_0^{m+1}\left(\varepsilon\right)\right) \right] \\ &\leq \lambda \sum_{m : g_0^m(\varepsilon) > \beta_\lambda \lambda^{-1}} g_0^m\left(\varepsilon\right) \leq \lambda \sum_m a_0^m \varepsilon \leq \frac{1}{1-a_0} \lambda\varepsilon. \end{aligned}$$

■

B.4 Estimates on interim beliefs

Proposition 6 below is the central step in the proof.

For $k \in \{1, \ldots, n-1\}$, we set $\psi_k(x) = x^k(1-x)^{n-k}$, and denote by $\bar{g}_\theta$ the function g_θ^β in the limit case $\beta = 0$. Proposition 6 relates likelihood ratios to the average value of ψ_k and ψ_{n-k} along doubly infinite orbits of $\bar{g}_1$.

[35] In the case $\theta = 1$, there is k such that $g_1^k(x) \geq \varepsilon$ for each $x \geq \beta\lambda^{-1}$.

[36] Independent of λ.

Proposition 6 *For each $k \in \{1, \ldots, n-1\}$, one has*

$$\limsup_{\lambda \to 0} \frac{p_k}{p_{n-k}} \leq \sup_{x \in (0,1)} \frac{\sum_{i \in} \psi_k \left(\bar{g}_1^i (x)\right)}{\sum_{i \in} \psi_{n-k} \left(\bar{g}_1^i (x)\right)}. \tag{C.3}$$

We emphasize that in this statement, the interim belief p_k depends on the transition rate λ and the information acquisition strategy β_λ.

Exchanging p_k and p_{n-k} yields the inequality $\liminf_{\lambda \to 0} \frac{p_k}{p_{n-k}} \geq \inf_{x \in (0,1)} \frac{\sum_{i \in} \psi_k \left(\bar{g}_1^i (x)\right)}{\sum_{i \in} \psi_{n-k} \left(\bar{g}_1^i (x)\right)}$.

Proof. Let $\varepsilon > 0$ and $k \in \{1, \ldots, n-1\}$ be given. Given λ, we choose m (independent of λ) such that $g_1^{m+1} (\varepsilon) > 1 - \varepsilon$ and set $\varepsilon' = 1 - g_1^{m+1} (\varepsilon) \leq \varepsilon$. Note that

$$\frac{p_k}{p_{n-k}} = \frac{\int_0^1 \psi_k (x) \, dF_1(x)}{\int_0^1 \psi_{n-k} (x) \, dF_1(x)}. \tag{C.4}$$

We write the latter numerator as

$$N_k := \int_0^\varepsilon \psi_k (x) \, dF_1 (x) + \int_\varepsilon^{1-\varepsilon'} \psi_k (x) \, dF_1 (x) + \int_{1-\varepsilon'}^1 \psi_k (x) \, dF_1 (x) \, .$$

Because of Lemma 9 and since $\int_{1-\varepsilon'}^1 \psi_k (x) \, dF_1 (x) = \int_0^{\varepsilon'} \psi_{n-k} (x) \, dF_0 (x)$ by symmetry, there exists $C_0 < +\infty$ such that for every λ small, one has

$$N_k = c_{\lambda,\varepsilon} \lambda \varepsilon + \sum_{i=0}^m \int_{g_1^i(\varepsilon)}^{g_1^{i+1}(\varepsilon)} \psi_k (x) \, dF_1 (x) \, , \tag{C.5}$$

for some $c_{\lambda,\varepsilon} \in [-C_0, C_0]$.

Let $i_0 = \max \left\{ i : g^i (\varepsilon) \leq \frac{1}{2} \right\}$ and $z = g^{i_0} (\varepsilon)$.[37] For each i, the change-of-variable formula for Stieltjes integrals yields

$$\int_{g_1^i(\varepsilon)}^{g_1^{i+1}(\varepsilon)} \psi_k (x) \, dF_1 (x) = \int_z^{g_1(z)} \psi_k \left(g_1^{i-i_0}(x)\right) dF_1 \left(g_1^{i-i_0}(x)\right).$$

Lemma 5 and the continuous differentiability of ψ_k and of $g_1^{i-i_0}$ imply the existence of $C_1 < \infty$ such that

$$\int_z^{g_1(z)} \psi_k \left(g_1^{i-i_0}(x)\right) dF_1 \left(g_1^{i-i_0}(x)\right) = \int_z^{g_1(z)} \psi_k \left(g_1^{i-i_0}(x)\right) dF_1 (x) + d_{\lambda,\varepsilon} mk\lambda^2$$

for some $d_{\lambda,\varepsilon} \in [-C_1, C_1]$. Plugging into (C.5), we get

$$\frac{1}{\lambda} N_k = c_{\lambda,\varepsilon} \varepsilon + d_{\lambda,\varepsilon} mk\lambda + \sum_{i=-i_0}^{m-i_0} \frac{1}{\lambda} \int_z^{g_1(z)} \psi_k \left(g_1^i(x)\right) dF_1 (x) \, . \tag{C.6}$$

[37] Note that for λ small, i_0 does not depend on λ.

For fixed ε, we deduce from (C.6) and its counterpart for N_{n-k} that

$$\frac{p_k}{p_{n-k}} = \frac{c_{\lambda,\varepsilon}\varepsilon + d_{\lambda,\varepsilon} mk\lambda + \frac{1}{\lambda}\int_z^{g_1(z)} \left[\sum_{i=-i_0}^{m-i_0} \psi_k\left(g_1^i(z)\right)\right] dF_1(x)}{c'_{\lambda,\varepsilon}\varepsilon + d'_{\lambda,\varepsilon} mk\lambda + \frac{1}{\lambda}\int_z^{g_1(z)} \left[\sum_{i=-i_0}^{m-i_0} \psi_{n-k}\left(g_1^i(z)\right)\right] dF_1(x)} \tag{C.7}$$

for suitable constants $c'_{\lambda,\varepsilon}$ and $d'_{\lambda,\varepsilon}$.

We note that Lemma 8 implies that for $a \in \{k, n-k\}$, the expression

$$\frac{1}{\lambda}\int_z^{g_1(z)} \psi_a(x)\, dF_1(x) \geq \left(\min_{x\in[\varepsilon, g_1(1-\varepsilon)]} \psi_a(x)\right) \times \min_{z\in[\varepsilon, g_1(\varepsilon)]} \frac{1}{\lambda}\left(F_1(g_1(z)) - F_1(x)\right)$$

is bounded away from zero as $\lambda \to 0$.

It thus follows from (C.7) that

$$\frac{p_k}{p_{n-k}} \leq \frac{\int_z^{g_1(z)} \left[\sum_{i=-i_0}^{m-i_0} \psi_k\left(g_1^i(z)\right)\right] dF_1(x)}{\int_z^{g_1(z)} \left[\sum_{i=-i_0}^{m-i_0} \psi_{n-k}\left(g_1^i(z)\right)\right] dF_1(x)} + C\left(\varepsilon + mk\lambda\right) \tag{C.8}$$

for some constant C. We now note that

$$\frac{\int_z^{g_1(z)} \sum_{i=-i_0}^{m-i_0} \psi_k(g_1^i(x)) dF_1(x)}{\int_z^{g_1(z)} \sum_{i=-i_0}^{m-i_0} \psi_{n-k}(g_1^i(x)) dF_1(x)} \leq \sup_{x\in[z, g_1(z)]} \frac{\sum_{i=-i_0}^{m-i_0} \psi_k\left(g_1^i(x)\right)}{\sum_{i=-i_0}^{m-i_0} \psi_{n-k}\left(g_1^i(x)\right)}. \tag{C.9}$$

For fixed ε (and hence fixed m and i_0), the right-hand side of (C.9) converges to

$$\sup_{x\in[z,\bar{g}_1(z)]} \frac{\sum_{i=-i_0}^{m-i_0} \psi_k\left(\bar{g}_1^i(x)\right)}{\sum_{i=-i_0}^{m-i_0} \psi_{n-k}\left(\bar{g}_1^i(x)\right)} \leq \sup_x \frac{\sum_{i\in} \psi_k\left(\bar{g}_1^i(x)\right)}{\sum_{i\in} \psi_{n-k}\left(\bar{g}_1^i(x)\right)},$$

as $\lambda \to 0$. Taking first the limit $\lambda \to 0$, then $\varepsilon \to 0$ in (C.8) yields the result. ■

B.5 Conclusion

We have shown that given $\hat{p}$, there is λ_0 such that for $\lambda < \lambda_0$, there is a strategy β_λ for which $p_n = 1 - p_0 = \hat{p}$. To complete the proof, we need to show that provided $\hat{p}$ is high enough, all interim beliefs p_k ($k = 1, \ldots, n-1$) are s.t. $p_k \leq \hat{p}$ for λ small. This follows from Proposition 6 and Lemma 10 below.

Lemma 10 *One has* $\sup_x \sum_{i\in} \psi_k(\bar{g}_1^i(x)) < +\infty$ *and* $\inf_x \sum_{i\in} \psi_k(\bar{g}_1^i(x)) > 0$.

Proof. For the lower bound, note that $\inf_x \sum_{i\in} \psi_k(\bar{g}_1^i(x)) \geq \min_{x\in\left[\frac{1}{2}, \bar{g}_1\left(\frac{1}{2}\right)\right]} \psi_k(x) > 0$.

For the upper bound, let $\varepsilon > 0$ be such that $a_1 := \min_{[0,\varepsilon]} \bar{g}'_1(x) > 1$ and $a_0 := \max_{[0,\varepsilon]} \bar{g}'_0(x) < 1$, and let m be s.t. $\bar{g}_1^m(\varepsilon) \geq 1 - \varepsilon$.

Then, for each $x \in (0,1)$, one has

$$\begin{aligned}\sum_{i\in} \psi_k\left(\bar{g}_1^i(x)\right) &\le \sum_{i:\bar{g}_1^i(x)\le\varepsilon} \psi_k\left(\bar{g}_1^i(x)\right) + \sum_{i:\bar{g}_1^i(x)\in(\varepsilon,1-\varepsilon)} \psi_k\left(\bar{g}_1^i(x)\right) + \sum_{i:\bar{g}_1^i(x)\ge 1-\varepsilon} \psi_k\left(\bar{g}_1^i(x)\right) \\ &\le \sum_{i:\bar{g}_1^i(x)\le\varepsilon} \bar{g}_1^i(x) + m + \sum_{i:\bar{g}_1^i(x)\ge 1-\varepsilon} \left(1-\bar{g}_1^i(x)\right) \\ &\le \frac{a_1}{a_1-1}\varepsilon + k + \frac{1}{1-a_0}(1-\varepsilon) = C < \infty.\end{aligned}$$

■

C Continuum of actions: Proposition 4

Denote by $p_{a^{(1)}}$ and $p_{a^{(1)}a^{(2)}}$ the interim beliefs after sampling one or two actions. The action $a^{(1)}$ coincides with the posterior belief of the sampled agent, hence $\mathbf{E}[u(a^{(1)})]$ is an upper bound on the welfare and $\mathbf{E}[u(p_{a^{(1)}a^{(2)}})] \le \mathbf{E}[u(a^{(1)})]$. On the other hand, taking transitions into account, one has

$$p_{a^{(1)}} = (1-\lambda)a^{(1)} + \lambda(1-a^{(1)}), \tag{C.10}$$

and therefore, $\mathbf{E}[u(p_{a^{(1)}a^{(2)}})] \le \mathbf{E}[u(p_{a^{(1)}})] + \lambda$.

Since u is quadratic, $u(p) = u(q) + (p-q)u'(q) + (p-q)^2$ for each p and q, so that

$$\mathbf{E}[u(p_{a^{(1)}a^{(2)}})] - \mathbf{E}[u(p_{a^{(1)}})] = \mathbf{E}\left[\left(p_{a^{(1)}a^{(2)}} - p_{a^{(1)}}\right)\left(2p_{a^{(1)}}-1\right)\right] + \mathbf{E}\left[\left(p_{a^{(1)}a^{(2)}} - p_{a^{(1)}}\right)^2\right].$$

By iterated conditional expectations, the first expectation on the RHS is zero, so that $\mathbf{E}\left[\left(p_{a^{(1)}a^{(2)}} - p_{a^{(1)}}\right)^2\right] \le \lambda$ and similarly, $\mathbf{E}\left[\left(p_{a^{(1)}a^{(2)}} - p_{a^{(2)}}\right)^2\right] \le \lambda$. Using the inequality $(x-y)^2 \le 2(x^2+y^2)$, we get $\mathbf{E}\left[\left(p_{a^{(1)}} - p_{a^{(2)}}\right)^2\right] \le 2\lambda$. The result follows, using (C.10).