Social Learning and the Shadow of the Past

Yuval Heller and Erik Mohlin

Bar Ilan University, Lund University

28 April 2017

Social Learning and the Shadow of the Past*

Yuval Heller† and Erik Mohlin‡

June 21, 2018

Abstract

In various environments new agents may base their decisions on observations of actions taken by a few other agents in the past. In this paper we analyze a broad class of such social learning processes, and study under what circumstances the initial behavior of the population has a lasting effect. Our results show that this question strongly depends on the expected number of actions observed by new agents. Specifically, we show that if the expected number of observed actions is: (1) less than one, then the population converges to the same behavior independently of the initial state; (2) between one and two, then in some (but not all) environments there are decision rules for which the initial state has a lasting impact on future behavior; and (3) more than two, then in all environments there is a decision rule for which the initial state has a lasting impact.

Keywords: Social learning, steady state, unique limiting behavior, path dependence. **JEL Classification:** C73, D83.

Manuscript accepted for publication in *Journal of Economic Theory* (final pre-print).

1 Introduction

Agents must often make decisions without knowing the costs and benefits of the possible choices. In such situations an inexperienced (or "newborn") agent may learn from the experience of others, by basing his decision, on observations of a few actions taken by other agents in the past (see, e.g., the social learning models of Ellison & Fudenberg, 1993, 1995; Acemoglu *et al.*, 2011). In other environments, agents interact with random opponents, and an agent may base his choice of action on a few observations of how his current opponent behaved in the past (as first described in Rosenthal, 1979, and further developed and applied to various models of community enforcement in the Prisoner's Dilemma game in Nowak & Sigmund (1998); Takahashi (2010); Heller & Mohlin (forthcoming)).

When analyzing such dynamic situations two related important questions arise: (1) does the initial behavior of the population have a lasting influence on the population's behavior in the long run, and (2) does the model admit a unique prediction or multiple predictions for the long-run behavior? For concreteness, consider an

*Old versions of this paper were previously titled "When Is Social Learning Path-Dependent?" and "Unique Stationary Behavior." We thank Ron Peretz, Doron Ravid, Satoru Takahashi, Xiangqian Yang, Peyton Young, an associate editor, and a referee for valuable discussions and helpful suggestions. Yuval Heller is grateful to the European Research Council for its financial support (ERC starting grant #677057). Erik Mohlin is grateful to Handelsbankens forskningsstiftelser (grant #P2016-0079:1), the Swedish Research Council (grant #2015-01751), and the Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellowship #2016-0156) for their financial support.

†Affiliation: Department of Economics, Bar Ilan University, Israel. E-mail: yuval.heller@biu.ac.il.

‡Affiliation: Department of Economics, Lund University, Sweden. E-mail: erik.mohlin@nek.lu.se.

environment in which new agents face a choice between competing technologies with positive externalities, where the unknown state of nature determines which technology is superior (see, e.g., Banerjee & Fudenberg, 2004). A central issue when analyzing this environment is to characterize in which setups the population may be "stuck" in the long run with the inferior technology, due to the arbitrary behavior of a few early adopters, and in which setups the population will always converge to the superior technology regardless of the initial behavior.

The present paper analyzes a broad class of processes in which agents obtain information by sampling the behavior of other agents, and it shows that the above two questions strongly depend on the expected number of actions observed by new agents. Specifically, we show that: (1) if the mean sample size (expected number of actions observed by a new agent) is less than one, then the population converges to the same behavior independently of the initial state; (2) if the mean sample size is between one and two, then any environment allows for a rule with multiple steady states according to which agents learn from the experience of other agents, but only some environments allow for decision rules with multiple locally stable states; and (3) if the mean sample size is more than two, then all environments admit a decision rule with multiple locally stable states, and the initial state determines which steady states will prevail. It should be noted that none of our results in any way depend on specifying how payoff considerations enter the agent's choices.[1]

Overview of the Model We consider an infinite population of agents (a continuum of mass one). Time is discrete and in every period each agent is faced with a choice among a fixed set of alternatives. The *population state* is a distribution of actions describing the aggregate behavior of agents in the population. In each period a fixed share of the agents die and are replaced with new agents. Each new agent observes a finite sequence of actions (called a *sample*) of random size. The sample may consist either of past actions of random agents in the population (as in the social learning models mentioned above) or past actions of the current, randomly drawn, opponent (as in the community enforcement models mentioned above).

A sampling process is a tuple that specifies all the above components. A decision rule specifies the distribution of actions played by new agents as a function of the observed sample. The agent keeps playing the same action throughout his life. The sampling process and the decision rule jointly induce a mapping between population states that determines a new population state for each initial state. A population state is a *steady state* if it is a fixed point of this mapping.

Characterization of Multiple Steady States Theorem 2 fully characterizes for which sampling processes there exist decision rules that admit multiple steady states. Specifically, it shows that a sampling process allows for a decision rule that admits multiple steady states if and only if the mean sample size is strictly more than one (or if agents always observe a single action). In the opposite case, each decision rule admits a unique steady state, and, moreover, the population converges to the unique steady state from any initial state.

The intuition for the "only if" side is as follows. Consider two different initial population states of a sampling process with a mean sample size below one. The aggregate behavior of new agents may differ only to the extent in which they observe different distributions of samples. This implies that the distance between the distributions of actions played by new agents is bounded by the distance between the distributions of samples that they observe. One can show that this latter distance is bounded by the distance between the distributions

[1] This features is a main difference compared to other models of convergence of social learning such as Schlag (1998); Oyarzun & Ruf (2014).

of actions played by the incumbents multiplied by the mean sample size. Hence, if the mean sample size is less than one, the distance between the distributions of actions of new agents is strictly less than the distance between the distributions of actions of the incumbents, which implies that the mapping between population states is a contraction mapping.

The "if" side relies on constructing a specific decision rule, according to which agents play action a' if they observe action a' in their sample, and play action a'' otherwise. One can show that such a decision rule always admits two different steady states provided that the expected number of observed actions is greater than one. We demonstrate that this decision rule (as well as all other decision rules used in the other results in the paper) may be consistent with Bayesian inference and the agents using best replies to plausible payoff functions.

Characterization of Multiple Locally Stable States A steady state is locally stable if the population converges back to this state after any sufficiently small perturbation. Arguably, the initial state can be said to have a lasting effect only when there are multiple locally stable states. In particular, in the construction of the above result (Theorem 2), only one of the steady states is locally stable. Moreover, one can show that a population converges to this state from almost all initial states. Our remaining results characterize when there are decision rules that admit multiple locally stable states.

Theorem 3 shows that in any sampling process in which the mean sample size is larger than two it is possible to define a decision rule that admits multiple locally stable states. According to this decision rule, each new agent (1) plays action a' if he observes action a' at least twice in his sample, (2) plays action a'' if he never observes action a', and (3) mixes between the two actions if he observes action a' exactly once. We show that this decision rule (with a well-calibrated mixing probability) admits two locally stable states, one in which a' is never played, and the other in which it is played with a positive probability.

Our next two results show that when the mean sample size is between one and two, then some (but not all) sampling processes allow for decision rules that admit multiple locally stable states. Specifically, we show that if each new agent observes at most two actions, and there are two feasible actions, then any decision rule admits a unique locally stable state, and, moreover, the population converges to this state from almost all initial states. The intuition is that when the sample size is at most two, then the mapping induced by the learning process can be represented as a polynomial of degree two. Hence, the mapping can have at most two steady states, and it is relatively straightforward to show that at most one of these states can be locally stable.

Finally, we show that a sampling process with two feasible actions a', a'' in which some agents observe a single action, while others observe three actions, and each new agent chooses the frequently observed action in his sample, admits two locally stable states: one in which all agents choose action a', and another in which everyone chooses action a'' (in addition, to an unstable state in which half of the population plays each action).

Extensions Our results so far have not assumed anything about the agents' decision rules. Obviously, additional information on the decision rules, may allow us to achieve stronger results. Next, we present a simple notion that measures how responsive a decision rule is to different samples. For example, a decision rule might be relatively unresponsive due to new agents having strong priors about which action is best. We use this notion of responsiveness to define the effective sample size of a learning process (which is always weakly smaller than the mean sample size). Next, we apply the notion of effective sample size to derive a tighter upper bound for learning processes that admit unique steady states.

Finally, we extend our model and main results to (1) heterogeneous populations in which agents are endowed with different types, and the various types differ in their sample sizes and decision rules, (2) non-stationary sampling processes, in which the distribution of sample sizes and the agents' decision rules depend on calendar time, and (3) stochastic shocks that influence the decision rules of all agents (at the aggregate level).

Related Literature Various papers have studied different aspects of the question of when the initial behavior of the population has lasting effects on social learning processes. Most of this literature focuses on specific decision rules, according to which new (or revising) agents myopically best reply to the empirical frequency of the observed actions. Arthur (1989) (see related models and extensions in Arthur, 1994; Kaniovski & Young, 1995; Smith & Sorensen, 2014) studies games in which agents sequentially choose which competing technology to adopt, and he shows that social learning is path-dependent if the technologies have positive externalities.

Kandori *et al.* (1993) and Young (1993a) study models of finite large populations that are involved in a social learning process, and agents occasionally make mistakes (e.g., an agent adopts a technology that is not his myopic best reply to his sampled information). They show that the path dependency of the social learning process vanishes when infinite time horizons are considered. In many cases, when the probability of mistakes is sufficiently small the population spends almost all the time in a unique "stochastically stable state," which is independent of the initial state. A key difference between our model and theirs is that we model an infinite population, rather than a large finite population. In Section 6, we discuss the relations between the present paper and the literature on stochastic stability, and, in particular, the implications of our results for finite large populations.

Banerjee & Fudenberg (2004) study a model with a continuum of agents in which a fixed share of new agents in each period choose one of two technologies. There are two possible states of nature, and each technology has a higher quality in one of these states. Each agent, after he observes l past actions and a noisy signal about the quality of each technology, chooses the technology with the higher expected quality, conditional on the information that he has observed. Banerjee & Fudenberg show that when $l \geq 2$ the behavior of the population converges to everyone choosing the efficient technology, while if $l = 1$ the population converges to an inefficient state in which only some of the agents choose the (ex-post) better technology.

Sandholm (2001) shows that when each new agent observes k actions and the game admits a $\frac{1}{k}$-dominant action a^* (i.e., action a^* is the unique best reply against any mixed strategy assigning a mass of at least $\frac{1}{k}$ to a^*), then social learning converges to this action regardless of the initial state. Recently, Oyama *et al.* (2015) strengthened this result by extending it to iterated p-dominant actions, and by showing that global convergence is fast.

Our model differs from all the above-mentioned research in that we study general sampling processes and arbitrary decision rules. Specifically, we ask what properties of the agents' sampling procedures imply that *any* decision rule admits a unique steady state and global convergence to this state, whereas the existing literature focuses on the dynamic behavior induced by a specific decision rule (in most of the literature, the agents myopically best reply to specific payoffs, such as those induced by competing technologies with positive externalities).

Structure We present motivating examples in Section 2. The basic model is described in Section 3. Section 4 presents our main results. In Section 5 we define and apply the notion of a responsiveness of a decision rule.

We conclude in Section 6. Appendix A extends the basic model to heterogeneous populations, non-stationary processes, and common stochastic shocks. Technical proofs are presented in Appendix B.

2 Motivating Examples

In this section we present three motivating examples, which will be revisited further below to demonstrate the applicability of our model and the implications of our results. In all the examples the population is modeled as a continuum of mass one, and time is discrete. The first example deals with social learning with competing technologies, while the second example studies situations in which agents are randomly matched to play a two-player game.

Example 1 (Main Motivating Example: Competing Technologies with Positive Externalities[2]). Consider a population in which in each period a share $\beta \in (0,1)$ of the incumbent agents die, and are replaced with new agents. Each new agent chooses one of two competing technologies a' and a'', which he adopts for the rest of his life. A share of 99% of the new agents observe the technology followed by a single random incumbent, and then they choose to adopt this observed technology.

We consider two cases for what the remaining 1% of the new agents observe before they choose a technology (as summarized in Table 1):

1. They observe nothing, and in this case half of the new agents adopt technology a', and the other half adopt technology a''.

2. They observe the technologies adopted by three random incumbents, and in this case each new agent adopts the technology chosen by the majority of his sample.

Let $\alpha_1 \in [0,1]$ describe the share of agents who use technology a' initially (in the first period). One can show that in Case (1), in which the mean sample size of a new agent is slightly less than one, the population converges to a unique steady state in which half of the agents follow each technology. By contrast, in Case (2), in which the mean sample size is slightly more than one, the initial behavior of the population has a lasting effect. Specifically, the population converges to everyone following technology a' if initially a majority of the agents followed technology a' (i.e., if $\alpha_1 > 50\%$), and the population converges to everyone following technology a'' if $\alpha_1 < 50\%$.

Table 1: Summary of the Two Cases in Example 1

Case	Probability of Observing			Mean Sample Size	Convergence and steady states
	0 actions	1 action	3 actions		
1	1%	99%	-	0.99	Global convergence to 50%–50%
2	-	99%	1%	1.02	Convergence to a' if $\alpha_1 > 0.5$; convergence to a'' if $\alpha_1 < 0.5$.

We conclude this example by noting that the described behavior is consistent with each new agent playing a unique best reply, if we make the following further assumptions. Nature privately chooses the initial share of agents who follow each technology in the first period. In each later period, new agents have a symmetric

[2]The example is similar to the model of Banerjee & Fudenberg (2004), except that the technologies have positive externalities, rather than having unknown inherent different qualities.

common prior on this initial share.[3] The payoff of each new agent is increasing in the current share of agents who follow the same technology (i.e., the technologies have positive externalities), and hence agents have an incentive to play the action that they believe that the majority is playing. In addition, half of the agents have a weak preference for technology a', and the remaining half have a weak preference for technology a''.

For example, consider the case in which α_1 is uniformly distributed on $[0,1]$, and each revising agent observes, in addition to the sample of actions described above, a signal about which technology fits better the agent's idiosyncratic skill set. Half of the new agents observe signal s', and the remaining agents observe s''. Assume that the payoff function of a new agent that chooses action a and observes signal s at time $t+1$ is given by

$$U(a,s)=\begin{cases}100\cdot\alpha_t+1 & a=a',\ s=s'\\ 100\cdot\alpha_t-1 & a=a',\ s=s''\\ 100\cdot(1-\alpha_t)+1 & a=a'',\ s=s''\\ 100\cdot(1-\alpha_t)-1 & a=a'',\ s=s',\end{cases} \tag{1}$$

where α_t is the share of agents who follow technology a' at time t. Note that a (risk-neutral) agent who observes s' strictly prefers action a' iff $\boldsymbol{E}(\alpha_t)>49\%$. In what follows we show that the behavior of the new agents described above is consistent with each agent playing his unique best reply. A new agent who observes no actions has a symmetric posterior regarding α_t, and, thus, he chooses an action based on his idiosyncratic preferences. A new agent who joins the population in round 2 and observes a single action a' updates his posterior belief to be $\alpha_1\sim Beta(2,1)$,[4] and, thus, his expectation of α_1 is given by $\boldsymbol{E}(\alpha_1)=\frac{2}{3}$. This in turn implies that choosing action a' yields a higher expected payoff regardless of the agent's own idiosyncratic preferences. Similarly, a new agent who observes two a'-s in a sample of three actions in round 2 updates his posterior belief to be $\alpha_1\sim Beta(3,2)$, and, thus, his expectation of α_1 is given by $\boldsymbol{E}(\alpha_1)=\frac{3}{5}$. This implies that choosing action a' induces a higher expected payoff regardless of the agent's own idiosyncratic preferences. Furthermore, one can show in Case 2 that a new agent observing a majority of a'-s in his sample in any later round $t+1>2$ assigns a high expected value to α_t, and, thus, he chooses a' regardless of his idiosyncratic preferences.[5]

Example 2 (Community Enforcement in the Prisoner's Dilemma)**.** Consider a population such that in each round each agent is randomly matched with three opponents, and plays a Prisoner's Dilemma with each of them. In round one, each agent defects with probability α in each match. In any later round and match, with a probability of 95% each agent observes two actions played in the previous period by the current opponent (i.e., actions played by the current opponent against two of his three opponents in the previous period). With the remaining probability of 5% each agent observes k actions played by the current opponent in the previous period. We consider two cases: (1) $k=1$, and (2) $k=3$. All agents follow the same behavior (in both cases):

[3]As argued by Banerjee & Fudenberg (2004, p. 5), the aggregate uncertainty about the initial population state may reflect the choices of a group of "early adopters" whose preferences are uncertain even at the aggregate level.

[4]Recall that the density of a random variable with a Beta distribution $x\sim\text{Beta}(a,b)$ is given by $f(x)=\text{constant}\cdot x^{\alpha-1}\cdot(1-x)^{\beta-1}$ (where $supp(x)=[0,1]$, $\alpha,\beta>0$, and Beta$(1,1)$ is the uniform distribution), and its expectation is given by $\boldsymbol{E}(x)=\frac{\alpha}{\alpha+\beta}$. Further recall that if variable x is interpreted as the prior over the frequency of technology a' in the population, and if one observes independent random observations of c agents playing action a' and d agents playing action a'', then the posterior about the frequency of technology a' is given by Beta$(a+c,b+d)$.

[5]In Case 1, when $t>>1$ a new agent who joins the population in round t knows that the population has already converged to a small neighborhood around the unique steady state in which half of the agents follow each technology. As a result, the new agent will follow his idiosyncratic preferences and ignore the signal about the population (which will slightly accelerate the speed of convergence to the unique steady state).

(I) an agent defects if he observes his partner defecting more times than cooperating, (II) an agent cooperates if he observes his partner cooperating more times than defecting, and (III) an agent defects with probability 51% if he observes the partner defecting and cooperating an equal number of times.
One can show that in both cases the sampling process admits two steady states: one in which all agents cooperate, and one in which all agents defect. In the first case ($k = 1$, in which the mean sample size is slightly below 2), only the state in which all agents defect is locally stable. Specifically, the population converges to everyone defecting for any $\alpha > 0$. By contrast, in the second case ($k = 3$, in which the mean sample size is slightly above 2), both steady states are locally stable. In particular, one can show that the population converges to everyone defecting if $\alpha > 31\%$, and it converges to everyone cooperating if $\alpha < 31\%$.

3 Model

Throughout the paper we restrict attention to distributions with a finite support. Given a (possibly infinite) set X, let $\Delta(X)$ denote the set of distributions over this set that have a finite support. With slight abuse of notation we use $x \in X$ to denote the degenerate distribution $\mu \in \Delta(X)$ that assigns probability one to x (i.e., we write $\mu \equiv x$ if $\mu(x) = 1$). We use $\mathbb{N}$ to denote the set of natural numbers including zero.

Population state. Consider an infinite population of agents. More precisely, the population consists of a continuum of agents with mass one. Time is discrete and in every period (or "round") each agent is faced with a choice among a fixed set of alternatives A. Let A be a finite set of at least two actions (i.e., $|A| \geq 2$).

The *population state* (or *state* for short) is identified with the aggregate distribution of actions played in the population, denoted $\gamma \in \Delta(A)$. Let Γ denote the set of all population states.

New/Revising agents. In each period, a share of $0 < \beta \leq 1$ of the agents exit the population and are replaced with new agents, while the remaining share of $1 - \beta$ of the agents play the same action as they played in the past (see, e.g., Banerjee & Fudenberg, 2004). Each new agent chooses an action based on a sample of a few actions of incumbents. The agent then keeps playing this chosen action throughout his active life, possibly because the initial choice requires a substantial action-specific investment, and it is too costly for an agent to reinvest in a different action later on. The model can also be interpreted as describing a fixed population in which each agent reevaluates his action only every $\frac{1}{\beta}$ periods. [6]

Sample. Each new agent observes a finite sequence of actions (or *sample*). The size of the observed sample is a random variable with a distribution $\nu \in \Delta(\mathbb{N})$. Let M denote the set of all feasible samples, i.e., $M = \cup_{l \in supp(\nu)} A^l$, where $A^0 = \{\emptyset\}$ is a singleton consisting of the empty sample $\emptyset$.

Let $\bar{l} = \max(supp(\nu)) < \infty$ be the maximal sample size. Note that M is finite in virtue of the finite-support assumption. For each sample size $l \in \mathbb{N}$, let $\psi_l : \Gamma \to \Delta(A^l)$ denote the distribution of samples observed by each agent in the population (*or sampling rule* for short), conditional on the sample having size l, and given a state of the population γ. A typical sample of size l is represented by the vector $\overrightarrow{a} = (a_1, ..., a_l)$.

[6] Under the interpretation of a fixed population, one can assume (without affecting the learning process) that the first observed action after a revision (when the sample is non-empty) is the revising agent's own past action. The reason why the learning process is unaffected by having the first observed action be the revising agent's own past action (rather than an action of another random agent) is that the revising agent himself is chosen uniformly, and, thus, the aggregate behavior of the revising agents coincides with the aggregate behavior in the population.

We assume that each agent independently samples different agents, and observes a random action played by each of these agents. This kind of sampling is common in models of social learning (see, e.g., Ellison & Fudenberg, 1995; Banerjee & Fudenberg, 2004). Formally, for each positive sample size $l \in \mathbb{N} \setminus \{0\}$, each state $\gamma \in \Gamma$, and each sample $(a_1, ..., a_l)$, we define

$$\psi_{l,\gamma}(a_1, ..., a_l) = \prod_{1 \le i \le l} \gamma(a_i). \tag{2}$$

For $l = 0$ we define $\psi_{0,\gamma}(\emptyset) = 1$ for each state $\gamma \in \Gamma$.

Sampling process. A *sampling process* (or environment) is a tuple $E = (A, \beta, \nu)$ that includes the three components described above: a finite set of actions A, a fraction of new agents at each stage β, and a distribution of sample sizes ν.

Given sampling process $E = (A, \beta, \nu)$, let μ_l denote the *mean sample size*, i.e., the expected number of actions observed by a random agent in the population. Formally:

$$\mu_l = \sum_{l' \in supp(\nu)} \nu(l') \cdot l'.$$

Decision rule and learning process. Each new agent chooses his action in the new population state by following a stationary (i.e., time-independent) *decision rule* $\sigma : M \to \Delta(A)$. That is, a new agent who observes sample $m \in M$ plays action a with probability $\sigma_m(a)$. The remaining $1 - \beta$ agents play the same action as in the previous round.

Observe that a decision rule can (implicitly) capture the agents' payoffs. As demonstrated in Example 1 below, a decision rule can describe the aggregate behavior of new agents who choose the expected payoff-maximizing actions, given their beliefs, which they have formed through Bayesian inference, on the basis of the observed sample of actions and possibly additional information about the state of nature.

A *learning process* is a pair $P = (E, \sigma)$ consisting of a sampling process and a decision rule.

Population dynamics. An initial state and a learning process uniquely determine a new state. To see this note that since the sets of samples M, and actions A are finite, whereas the population is a continuum, the *probability* that an agent observes a sample m and switches to an action a is equal to the *fraction* of agents who observe m and switch to an action a. For this reason we say that the learning process is *deterministic*, despite the fact that the choice of an individual agent may be stochastic.[7]

Time is discrete in our model. Let $f_P : \Gamma \to \Gamma$ denote the mapping between states induced by a single step of the learning process P. That is, $f_P(\hat{\gamma})$ is the new state induced by a single step of the process P, given an initial state $\hat{\gamma}$. Similarly, for each $t > 1$, let $f_P^t(\hat{\gamma})$ denote the state induced after t steps of the learning process P, given an initial state $\hat{\gamma}$ (e.g., $f_P^2(\hat{\gamma}) = f_P(f_P(\hat{\gamma}))$, $f_P^3(\hat{\gamma}) = f_P(f_P(f_P(\hat{\gamma})))$, etc.).

[7]The formalization of the intuitive claim that the probability that each agent who chooses action a is equal to the fraction of agents who choose action a does raise various technical difficulties. For example, when the population is a continuum, and each agent independently chooses action a with probability 50%, then the set of agents who happen to choose action a in a particular realization may not be measurable. For brevity and clarity, we do not deal with these technical difficulties in the present paper (as they are not directly related to the main focus of the paper). We refer the interested reader to Duffie & Sun (2012) for details on how the above intuitive claim can be formalized in a closely related setup.

L_1-distance. Throughout the paper we measure distances with the L_1-distance (norm). Specifically, let the L_1-distance between two distributions of samples $\psi_{l,\gamma}, \psi_{l,\gamma'} \in \Delta\left(A^l\right)$ of size l, be defined as follows:

$$\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1 = \sum_{m \in A^l} \left|\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right|.$$

Similarly the L_1-distance between two distributions of actions $\gamma, \gamma' \in \Delta(A)$ is defined as follows:

$$\left\|\gamma - \gamma'\right\|_1 = \sum_{a \in A} \left|\gamma(a) - \gamma'(a)\right|.$$

Steady States and Stability We say that γ^* is a *steady state* with respect to the learning process P, if it is a fixed point of the induced mapping f_P, i.e., if $f_P(\gamma^*) = \gamma^*$.

Steady state $\gamma^*$is (asymptotically) locally stable if a population beginning near γ^* remains close to γ^*, and eventually converges to γ^* . Formally, for each $\epsilon > 0$ there exists $\delta > 0$ such that $\left\|\hat{\gamma} - \gamma^*\right\|_1 < \delta$ implies that:[8]

$$(1) \;\; \left\|f_P^t(\hat{\gamma}) - f_P^t(\gamma^*)\right\|_1 < \epsilon \;\; \forall t \geq 1, \quad (2) \;\; \lim_{t \longrightarrow \infty} f_P^t(\hat{\gamma}) = \gamma^*.$$

Steady state γ^* is an *(almost-)global attractor*, if the population converges to γ^* from any (interior) initial state, i.e., if

$$\lim_{t \longrightarrow \infty} f_P^t(\hat{\gamma}) = \gamma^* \text{ for all } \hat{\gamma} \in \Gamma \;\; (\hat{\gamma} \in \operatorname{Int}(\Gamma)),$$

where $\operatorname{Int}(\Gamma)$ denotes the set of totally mixed distributions of actions (distributions that assign positive probability to all actions).[9]

We conclude by demonstrating how the model captures our motivating examples.

Example 1 (Competing Technologies Revisited). We model the process by which agents adopt one of two competing technologies with positive externalities as a learning process $P = (\{a', a''\}, \beta, \nu, \sigma)$, where the sampling process consists of (i) the set of competing technologies $\{a', a''\}$, (ii) the share of new agents that join the population in each round, $\beta \in (0,1)$, and (iii) the distribution of sample size ν, which is defined as

$$\text{Case I: } \nu(l) = \begin{cases} 1\% & l = 0 \\ 99\% & l = 1 \end{cases} \qquad \text{Case II: } \nu(l) = \begin{cases} 1\% & l = 3 \\ 99\% & l = 1. \end{cases}$$

Note that the mean sample size (μ_l) is equal to 0.99 in Case I, and is equal to 1.02 in Case II.

[8] The notion of local stability is often called "asymptotic stability" (see, e.g., Weibull, 1995, Def. 6.5). A state that satisfies the first requirement of our definition is called Lyapunov stable.

[9] Weibull (1995, p. 101) uses the term "globally stable" to refer to an almost-global attractor. Sandholm (2011, Section 7.A.2) uses the term "globally attracting" to refer to a global attractor. All the results in the paper that show that a state is a "global attractor" (namely, Corollaries 1–7) can be strengthened in a straightforward way to show that the relevant state is "globally asymptotically stable" à la Sandholm (2011, Section 7.A.2) (i.e., to show that the global attractor also satisfies Lyapunov stability).

Finally, the decision rule of the new agents is defined as

$$\sigma\left(\overrightarrow{a}\right)=\begin{cases}0.5\cdot a'+0.5\cdot a'' & \overrightarrow{a}=\emptyset\\ a' & \overrightarrow{a}\in\left\{a',\left(a',a',a'\right),\left(a'',a',a'\right),\left(a',a'',a'\right),\left(a',a',a''\right)\right\}\\ a'' & \text{otherwise.}\end{cases}$$

The initial population state is given by $\hat{\gamma}\left(a'\right)=\alpha$.

Example. 2 (Prisoner's Dilemma Revisited). The process by which agents choose how to play the Prisoner's Dilemma is modeled by a learning process

$$P=\left(\left\{c,d\right\},\beta=1,\nu,\sigma\right),$$

where $\nu\left(2\right)=95\%$, and in Case (1) $\nu\left(1\right)=5\%$, while in Case (2) $\nu\left(3\right)=5\%$. In Case (1) the decision rule is given by:

$$\sigma\left(c,c\right)=\sigma\left(c\right)=c,\quad\sigma\left(d,d\right)=\sigma\left(d\right)=d,\quad\sigma\left(c,d\right)=\sigma\left(d,c\right)=51\%\cdot d+49\%\cdot c.$$

In Case (2) the decision rule is given by:

$$\sigma\left(c,c\right)=\sigma\left(c,c,c\right)=\sigma\left(c,c,d\right)=\sigma\left(c,d,c\right)=\sigma\left(d,c,c\right)=c,$$

$$\sigma\left(d,d\right)=\sigma\left(d,d,d\right)=\sigma\left(d,d,c\right)=\sigma\left(d,c,d\right)=\sigma\left(c,d,d\right)=d,$$

$$\sigma\left(c,d\right)=\sigma\left(d,c\right)=51\%\cdot d+49\%\cdot c.$$

Observe that $\mu_l=1.95$ in Case (1) , and $\mu_l=2.05$ in Case (2).

4 Main Results

4.1 Upper Bound on the Distance between New States

Our first result shows that for any two initial states $\gamma\neq\gamma'$, the distance between the corresponding states one period later is at most $1-\beta+\beta\cdot\mu_l$ times the distance between the two initial states. The intuition is as follows. Consider two different initial population states γ and γ'. The incumbents who do not die (a share of $1-\beta$) continue to behave as before, and hence the distance between the distributions of actions of those agents remains $\left\|\gamma-\gamma'\right\|_1$. The aggregate behavior of new agents may differ only to the extent that they face different distributions of samples. This implies that the distance between the distributions of actions played by new agents is bounded by the distance between the distributions of samples that they observe. By using the triangle inequality one can show that this latter distance is bounded by the distance between the distributions of actions played by the incumbents ($\left\|\gamma-\gamma'\right\|_1$) multiplied by the mean sample size (μ_l). Finally, due to another use of a triangle inequality, this implies that the distance between the new population states is at most $\left(1-\beta+\beta\cdot\mu_l\right)\cdot\left\|\gamma-\gamma'\right\|_1$. Formally,

Theorem 1. *Let $P=\left(A,\beta,\nu,\sigma\right)$ be a learning process, and let $\gamma\neq\gamma'\in\Gamma$ be two population states. Then:*

$$\left\|f_P\left(\gamma\right)-f_P\left(\gamma'\right)\right\|_1\leq\left(1-\beta+\beta\cdot\mu_l\right)\cdot\left\|\gamma-\gamma'\right\|_1,$$

with a strict inequality if $\nu(l) > 0$ for some $l > 1$.

Sketch of proof (Formal proof is presented for the more general result of Theorem 8 in Appendix B.1). The distance between the new population states is bounded as follows:

$$\left\|(f_P(\gamma)) - (f_P(\gamma'))\right\|_1 \leq \beta \cdot \sum_{l\in\mathbb{N}} \left(\nu(l) \cdot \left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1\right) + (1-\beta) \cdot \left\|\gamma - \gamma'\right\|_1. \tag{3}$$

The reason for this inequality is as follows. The first part of the RHS of Eq. (3) reflects the actions played by the β new agents. The social learning stage may induce different behaviors for new agents who observe samples of size l only if they observe different samples. Thus, taking the weighted average of the distances between samples yields the bound on how much the aggregate behaviors of the new agents may differ (i.e., $\sum_{l\in\mathbb{N}} \nu(l) \cdot \left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1$). Finally, the mixed average of this expression and the behavior of the incumbents, gives the total bound on the difference between the final population states.

Next, observe that the distance between distributions of samples is bounded by the sample size times the distance between the distributions of actions:

$$\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1 \leq l \cdot \left\|\gamma - \gamma'\right\|_1,$$

with a strict inequality if $l > 1$. This is so because the event that two samples of size l differ is (a non-disjoint) union of the l events: the first action in the samples differs, the second action in the samples differs, ..., the l^{th} action in the samples differ.

Substituting the second inequality in (3) yields:

$$\left\|(f_P(\gamma)) - (f_P(\gamma'))\right\|_1 \leq \beta \cdot \sum_{l\in\mathbb{N}} \nu(l) \cdot l \cdot \left\|\gamma - \gamma'\right\|_1 + (1-\beta) \cdot \left\|\gamma - \gamma'\right\|_1 =$$

$$\left(\beta \cdot \left(\sum_{l\in\mathbb{N}} \nu(l) \cdot l\right) + (1-\beta)\right) \cdot \left\|\gamma - \gamma'\right\|_1 = (\beta \cdot \mu_l + 1 - \beta) \cdot \left\|\gamma - \gamma'\right\| = (1 - \beta + \beta \cdot \mu_l) \cdot \left\|\gamma - \gamma'\right\|_1,$$

with a strict inequality if $\nu(l) > 0$ for some $l > 1$. □

Observe that $1-\beta+\beta\cdot\mu_l < 1$ iff $\mu_l < 1$. Recall that f is a *weak contraction* mapping if $\left\|(f(\gamma)) - (f(\gamma'))\right\|_1 < \left\|\gamma - \gamma'\right\|_1$ for each $\gamma \neq \gamma'$. Theorem 1 implies that f_P is a weak contraction mapping if either (1) $\mu_l < 1$, or (2) $\mu_l = 1$ and $\nu(1) < 1$.[10] The fact that the mapping f_P is a weak contraction mapping implies that f_p admits a global attractor.[11] Formally:

Corollary 1. *Let $P = (A, \beta, \nu, \sigma)$ be a learning process satisfying (1) $\mu_l < 1$, or (2) $\mu_l = 1$ and $\nu(1) < 1$. Then f_P is a weak contraction mapping, which implies that (1) f_P admits a unique steady state γ^*, and (2) this unique steady state γ^* is a global attractor (i.e., $\lim_{t\longrightarrow\infty} f_P^t(\hat{\gamma}) = \gamma^*$ for each $\hat{\gamma} \in \Gamma$).*

[10] Note that $\mu_l = 1$ and $\nu(1) < 1$ jointly imply that there exists $l > 1$ such that $\nu(l) > 0$.

[11] See Pata (2014, Theorem 1.7) for a formal proof that any weak contraction mapping on a compact metric space admits a global attractor (see also the sketch of the proof in Munkres, 2000, Section 28, Exercise 7). We thank Xiangqian Yang for kindly referring us to these proofs.

4.2 Full Characterization of Sampling Processes with Multiple Steady States

Our main result fully characterizes which sampling processes allow for decision rules for which the past casts a long shadow. Specifically, it shows that a sampling process allows a decision rule that admits multiple steady states iff $\mu_l > 1$ (alternatively if all agents sample exactly one action). In the opposite case ($\mu_l \le 1$) each decision rule admits a unique steady state, and, moreover, the population converges to the unique steady state from any initial state. Formally:

Theorem 2. *Let $E = (A, \beta, \nu)$ be a sampling process. The following two conditions are equivalent:*

1. $\mu_l > 1$*, or* $\nu(1) = 1$*.*
2. *There exists a decision rule σ^*, such that the learning process (E, σ^*) admits at least two different steady states.*

Proof. Corollary 1 immediately implies that $\neg 1 \Rightarrow \neg 2$. We are left with the task of showing that $1 \Rightarrow 2$.

Case A: Assume that $\nu(1) = 1$ (i.e., each new agent in the population observes a single action). Consider the decision rule in which each agent plays the action that he observed, i.e., $\sigma^*(a) = a$. Let γ be an arbitrary population state. Observe that γ is a steady state of the learning process (E, σ^*) because:

$$(f_P(\gamma))(a) = \gamma(a).$$

Case B: Assume that $\mu_l > 1$. Let a^* and a' be different actions ($a^* \neq a' \in A$). Let σ^* be a decision rule according to which each agent plays action a^* if he has observed action a^* at least once, and plays action a' otherwise, that is,

$$\sigma^*\left(a^l\right) = \begin{cases} a^* & \exists i, \text{ s.t., } a_i^l = a^* \\ a' & \text{otherwise.} \end{cases}$$

It is immediate that the population state in which all agents play action a' (i.e., $\gamma(a') = 1$) is a steady state of the learning process (E, σ^*). We denote this state γ^0 to reflect the fact that no agents play action a^* in this state. We now show that there exists an $x > 0$, such that the population state γ^x in which a share x of the agents play action a^*, and the remaining agents play action a' (i.e., $\gamma^x(a^*) = x$ and $\gamma^x(a') = 1 - x$), is another steady state of the learning process (E, σ^*). Observe that the state γ^x is consistent with the learning process (E, σ^*) if and only if

$$(f_P(\gamma^x))(a^*) = \sum_{l \in supp(\nu)} \nu(l) \cdot \frac{1}{|A|^l} \cdot \sum_{\vec{a} \in A^l} \mathbf{1}_{(\exists i \text{ s.t., } a_i = a^*)} = \sum_{l \in supp(\nu)} \nu(l) \cdot \left(1 - (1-x)^l\right) \equiv g(x). \qquad (4)$$

Observe that: (1) $g(x)$ (defined in (4) above) is continuous and differentiable, (2) the derivative of $g(x)$ is given by $g'(x) = \sum_{l \in supp(\nu)} \nu(l) \cdot l \cdot (1-x)^{l-1}$, (3) $g'(0) = \sum_{l \in supp(\nu)} \nu(l) \cdot l = \mu_l > 1$, (4) $g(0) = 0$, and (5) $g(1) \le 1$. These observations imply by the intermediate value theorem that there is $x^* > 0$ such that $g(x^*) = x^*$, and hence γ^{x^*} is an additional steady state of the learning process (E, σ^*). ◻

Remark 1. We note that the decision rules constructed in the proof above can be consistent with Bayesian inference and best-replying in plausible setups. The decision rule in Case A (playing the observed action) induces a Nash equilibrium in a setup with competing technologies with positive externalities and uncertainty about the initial population state, such as the setup presented in Example 1.

The decision rule in Case B induces a Nash equilibrium in the following setup of two competing technologies with uncertainty about their quality. There are two states of the world. In State 1 technology a^* has a higher quality, and in state 2 technology a' has a higher quality. The technology with the higher quality yields a payoff of one to an agent who follows it, and the technology with the lower quality yields a payoff of zero. State 1 has a prior probability of 60%. In state 1, 10% of the agents follow technology a^* in the first period, and the remaining agents follow technology a'. In state 2, all agents follow technology a' in period one (i.e., the setup has a payoff-determined initial popularity à la Banerjee & Fudenberg, 2004). Observe that the unique Nash equilibrium in this setup is for an agent to play a^* when observing a^* at least once (as in this case the agent knows for sure that action a^* has a higher quality), and to play a' otherwise (as in this case the posterior probability that action a' has a higher quality is at least 60%).

Similarly, one can design plausible setups, in which the decision rules presented in all other constructions in the paper are consistent with Bayesian inference and best-replying (omitted for brevity).

4.3 Any Sampling Process with $\mu_l > 2$ Admits Multiple Locally Stable States

Theorem 2 shows that any sampling process with a sample size larger than one admits multiple steady states, but it does not address the question of whether these steady states are locally stable. In particular, the decision rule presented in the proof of Theorem 2 (Case B) admits two steady states γ^0 and γ^{x^*}. It is relatively simple to see that the state γ^{x^*} is an almost global attractor: the population converges to γ^{x^*} from any initial state $\hat{\gamma}$ that assigns a positive probability to action a^* (see the related continuous-time analysis in Oyama *et al.* (2015, Sections 3.2 and 3.3)). Hence, γ^0 (the state where no one plays a^*) is not locally stable. In the next two sections we establish necessary and sufficient conditions for sampling processes to admit multiple locally stable states.

The following result shows that in any sampling process with a mean sample size larger than 2 it is possible to define a decision rule that admits multiple locally stable states.

Theorem 3. *Let $E = (A, \beta, \nu)$ be a sampling process satisfying $\mu_l > 2$. There exists a decision rule σ^*, such that the learning process (E, σ^*) admits two different locally stable states.*

The decision rule mentioned in the theorem is such that, each new agent (1) plays action a' if he observes action a' at least twice in his sample, (2) plays action a'' if he never observes action a', and (3) plays action a' with probability q and action a'' with probability $1 - q$, if he observes action a' exactly once.

The sketch of the proof is as follows (the formal proof is presented in Appendix B.2). If the incumbents play action a' with a frequency of $x << 1$, then the share of new agents who play action a' is $q \cdot \mu_l \cdot x + (1 - 2 \cdot q) \cdot O\left(x^2\right)$ (the first term reflects the fact that the probability that a new agent plays action a' is approximately q times the expected number of times in which action a' is observed, namely, $\mu_l \cdot x$; the second term "corrects" the fact that when a new agent observes action a' twice he plays action a' with probability one rather than with probability $2 \cdot q$). Choosing $q < \frac{1}{\mu_l}$ implies that a population in which very few agents play a' converges to no one playing a'. Choosing q sufficiently close to $\frac{1}{\mu_l} < \frac{1}{2}$ implies that a population in which a few more agents play action a' converges to a larger share of agents playing action a' (due to the second-order term, $(1 - 2 \cdot q) \cdot O\left(x^2\right)$, being positive).

4.4 Some Processes with $1 < \mu_l < 2$ Admit Multiple Locally Stable States

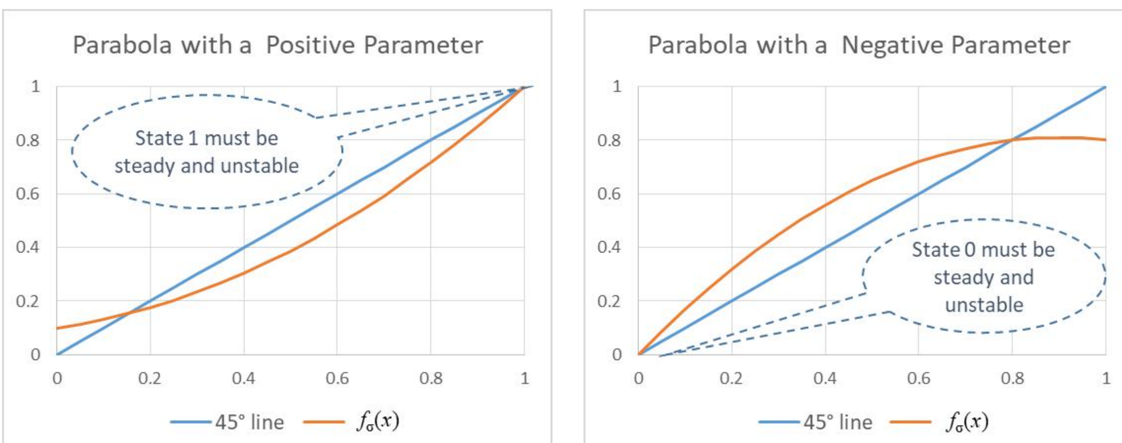
In this section we show that some (but not all) sampling processes in which the mean sample size is between one and two allow for a decision rule that admits multiple locally stable states.

Theorem 4 presents a family of sampling processes with a mean sample size of up to two, in which every decision rule admits at most one locally stable state. Specifically, we show that in any sampling process in which (1) there are two feasible actions ($|A| = 2$), and (2) each new agent observes at most 2 actions, any decision rule admits at most one locally stable state.

Theorem 4. *Let $E = (A = \{a', a''\}, \beta, \nu)$ be a sampling process. Assume that $\nu(l) = 0$ for each $l > 2$. Then for any decision rule σ, the learning process (E, σ) admits at most one locally stable state.*

The sketch of the proof of Theorem 4 is as follows (the formal proof is presented in Appendix B.3). In sampling processes with two actions, the state can be identified with a number $x \in [0,1]$ representing the frequency of agents playing the first action. Recall that any steady state is a solution to the equation $f_\sigma(x) = x$, where $f_\sigma(x)$ is the dynamic mapping induced by decision rule σ. The fact that the maximal sample size is two implies that $f_\sigma(x)$ is a polynomial of degree two. This implies that there are at most two steady states solving $f_\sigma(x) = x$. Simple geometric arguments regarding the intersection points of a parabola and the 45° line imply that at most one of these steady states can be locally stable (as illustrated in Figure 1).

Figure 1: Illustrations for the Intersections of a Parabola and the 45° Line



Theorem 5 presents a family of sampling processes (which extends Case (2) in Example 1) in which the mean sample size is between one and two, such that a simple "follow the majority" rule admits multiple locally stable states. Specifically, in these sampling processes (1) some agents observe a single action and the remaining agents observe three actions, and (2) each agent plays action a' if action a' has been observed in a majority of his sample, and he plays action a'' otherwise. It is relatively straightforward to see that this process admits two locally stable steady states: one in which all agents play action a', and one in which all agents play action a''. In addition, the state in which half of the agents play each action is an unstable steady state. The formal proof of Theorem 5 is given in Appendix B.4.

Theorem 5. *Let $E = (A, \beta, \nu)$ be a sampling process. Assume that $\nu(1) < 1$ and $\nu(1) + \nu(3) = 1$. Then there exists a decision rule σ^*, such that (E, σ^*) admits multiple locally stable states.*

4.5 Summary of Main Results

Combining the various results of this section shows that the sampling process's mean sample size has important implications for determining whether the initial behavior of the population has has a lasting influence on the long-run behavior of the population.

Corollary 2. *Let E be a sampling process with an expected sample size μ_l.*

1. *If $\mu_l < 1$ (or $\mu_l = 1$ and $\nu(1) \neq 1$), then any decision rule admits a unique steady state that is globally stable.*
2. *If $1 < \mu_l \leq 2$, then there exists a decision rule that admits multiple steady states. By contrast, the multiplicity of locally stable states depends on other details of the sampling process. That is, for each $1 < \mu_l \leq 2$ there exist sampling processes E' and E'', both with mean sample size μ_l, such that sampling process E' allows for a decision rule that admits multiple locally stable states, while sampling process E'' does not.*
3. *If $\mu_l > 2$, then there exists a decision rule that admits multiple locally stable states.*

5 Responsiveness and Effective Sample Size

In this section, we present simple notions of responsiveness and expected effective sample size, and use them to derive a (weakly) tighter upper bound for processes that admit global attractors (relative to the upper bound presented in Theorem 1).

5.1 Definitions

Fix a learning process $P = (A, \beta, \nu, \sigma)$. For each sample size $l \in supp(\nu)$, and each action $a \in A$, let $\underline{\sigma}_l(a)$ ($\overline{\sigma}_l(a)$) be the minimal (maximal) probability that decision rule σ assigns to action a after observing a sample of size l, i.e.,

$$\underline{\sigma}_l(a) = \min_{m \in A^l} \sigma_m(a) \qquad \left(\overline{\sigma}_l(a) = \max_{m \in A^l} \sigma_m(a)\right).$$

Let r_l denote the maximal *responsiveness* of new agents to changes in observed samples of size l, which is defined as follows:

$$r_l = \min\left(1,\ \frac{1}{2} \cdot \sum_{a \in A} \left(\overline{\sigma}_l(a) - \underline{\sigma}_l(a)\right)\right), \tag{5}$$

and let $r_0 = 0$. The responsiveness effectively limits the maximal influence of different samples of length l on the behavior of agents to be at most $r_l \leq 1$. One reason a decision rule may have limited responsiveness is that new agents might have strong priors about the best action, which are influenced only to a limited extent by their observed samples. Observe that when there are two actions, i.e., $A = \{a, b\}$, then l is simply the difference between the maximal and minimal probability assigned to each action, i.e.,

$$r_l = \overline{\sigma}_l(a) - \underline{\sigma}_l(a) = \overline{\sigma}_l(b) - \underline{\sigma}_l(b). \tag{6}$$

When there are more than two actions, $\frac{1}{2} \cdot \sum_{a \in A} (\overline{\sigma}_l(a) - \underline{\sigma}_l(a))$ may be larger than one. We bound r_l from above by one in Eq.(5) because, any change of sample cannot affect an agent's mixed behavior by more than one (as measured by the L_1-distance over the set of mixed actions).

We call the product of the sample size and the responsiveness, $r_l \cdot l$ the *effective sample size.* Let $\mu_l^e \in \mathbb{R}^+$ denote the *effective sample size*, i.e.,

$$\mu_l^e = \sum_{l \in supp(\nu)} \nu(l) \cdot r_l \cdot l.$$

It is immediate that the effective sample size is always weakly smaller than the mean sample size in the population; i.e., $\mu_l^e \leq \mu_l$.

5.2 A Tighter Bound on the Distance between New States

Our main result in this section shows that the distance between two new states is at most $(1 - \beta + \beta \cdot \mu_l^e)$ times the distance between the two initial states. This bound is (weakly) tighter than the one presented in Theorem 1, as we replace expected sample size μ_l with the (weakly) smaller effective sample size μ_l^e. Formally,

Theorem 6. *Let $P = (A, \beta, \nu, \sigma)$ be a learning process, and let $\gamma \neq \gamma' \in \Gamma$ be two population states. Then:*

$$\|f_P(\gamma) - f_P(\gamma')\|_1 \leq (1 - \beta + \beta \cdot \mu_l^e) \cdot \|\gamma - \gamma'\|_1,$$

where the inequality is strict if there exists an $l > 1$ such that $\nu(l) > 0$.

Proof. The key step of the proof is to show the following inequality:

$$\|(f_P(\gamma)) - (f_P(\gamma'))\|_1 \leq \beta \cdot \sum_{l \in \mathbb{N}} \nu(l) \cdot r_l \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1 - \beta) \cdot \|\gamma - \gamma'\|_1. \tag{7}$$

Inequality (7) is the same as (3) in the proof of Theorem 1, except for the factor of $r_l \leq 1$ on the RHS. All other arguments of the proof of Theorem 1 remain the same. We prove (7) in Lemma 6 in Appendix B.5.

Observe that $(1 - \beta + \beta \cdot \mu_l^e) < 1$ iff $\mu_l^e < 1$, and in this case f_P is a contraction mapping, which implies that f_P admits a global attractor. This allows us to strengthen Corollary 1 as follows. □

Corollary 3. *Let $P = (A, \beta, \nu, \sigma)$ be a learning process satisfying (1) $\mu_l^e < 1$, or (2) $\mu_l^e = 1$ and $\nu(1) < 1$. Then f_P is a contraction mapping, which implies that (1) f_P admits a unique steady state γ^*, and (2) this unique steady state γ^* is a global attractor (i.e., $lim_{t \longrightarrow \infty} f_P^t(\hat{\gamma}) = \gamma^*$ for each $\hat{\gamma} \in \Gamma$).*

We demonstrate the implications of Corollary 3 in the following example.

Example 3. Consider a population in which in each period a share $\beta \in (0, 1)$ of the incumbent agents die, and are replaced with new agents. A population state describes the share of agents who use each of two competing technologies, a_1 and a_2. Each new agent observes the technology followed by a single random incumbent. Assume that the decision rule used by the agents implies that each new agent plays (on average) action a_1 with a probability of $\bar{\alpha} \in [0, 1]$ after observing action a', and with a probability of $\underline{\alpha} < \bar{\alpha}$ after observing action a''. Observe that the effective number of observations, μ_l^e, is equal to:

$$\mu_l^e = r_{l=1} \cdot 1 = \frac{1}{2} \cdot \sum_{a \in A} (\overline{\sigma}_l(a) - \underline{\sigma}_l(a)) = \frac{1}{2} \cdot ((\overline{\alpha} - \underline{\alpha}) + ((1 - \underline{\alpha}) - (1 - \overline{\alpha}))) = \overline{\alpha} - \underline{\alpha},$$

which is strictly less than one if $\overline{\alpha} < 1$ or $\underline{\alpha} > 0$. Corollary 3 implies that the learning process converges to a global attractor (which is the unique steady state) whenever $\overline{\alpha} < 1$ or $\underline{\alpha} > 0$.[12]

Our final result demonstrates that our bound of the effective sample size being less than one is tight. Specifically, it shows that given any sampling process in which the expected sample size $\mu_l > 1$, and any number $1 < y \le \mu_l$, there is a decision rule with an effective sample size of $\mu_l^e = y$ with multiple steady states. Formally:

Theorem 7. *Let $E = (A, \beta, \nu)$ be a sampling process satisfying $\mu_l > 1$. Let $1 < y \le \mu_l$. Then there exists a decision rule σ, such that the learning process (E, σ) admits two different steady states, and satisfies $\mu_l^e = y$.*

Proof. Let a and a' be different actions ($a \neq a' \in A$). Let σ^* be a decision rule according to which each agent plays action a^* with a probability of $\frac{y}{\mu_l}$ if he has observed action a^* at least once, and plays action a' otherwise, that is,

$$\sigma^*\left(a^l\right) = \begin{cases} \frac{y}{\mu_l} \cdot a^* + \left(1 - \frac{y}{\mu_l}\right) \cdot a' & \text{if } \exists i, \text{ s.t., } a_i^l = a^* \\ a' & \text{otherwise.} \end{cases}$$

Observe that the responsiveness of (E, σ) is equal to y because:

$$\mu_l^e = \sum_{l' \in supp(\nu)} \nu(l') \cdot r_{l'} \cdot l' = \sum_{l' \in supp(\nu)} \nu(l') \cdot \frac{1}{2} \cdot \sum_{a \in A} \left(\overline{\sigma}_{l'}(a) - \underline{\sigma}_{l'}(a)\right) \cdot l' =$$

$$\sum_{l' \in supp(\nu)} \nu(l') \cdot \frac{1}{2} \cdot \left(\left(\frac{y}{\mu_l} - 0\right) + \left(1 - \left(1 - \frac{y}{\mu_l}\right) + 0 + \ldots + 0\right)\right) \cdot l' =$$

$$\sum_{l' \in supp(\nu)} \nu(l') \cdot \frac{y}{\mu_l} \cdot l' = \frac{y}{\mu_l} \cdot \sum_{l' \in supp(\nu)} \nu(l') \cdot l' = \frac{y}{\mu_l} \cdot \mu_l = y.$$

It is immediate that the population state in which all agents play action a' (i.e., $\gamma(a') = 1$) is a steady state of the learning process (E, σ^*). An analogous argument to the one presented in Case B of the proof of Theorem 2 shows that there exists $x > 0$ such that the population state γ^x in which all agents play action a^* with probability x, and play action a' with the remaining probability of $1 - x$, is another steady state of the learning process (E, σ^*). $\square$

6 Concluding Remarks

Extensions. The basic model assumes that all agents share the same distribution of sample sizes, and the same decision rule. In many applications the population might be heterogeneous, i.e., the population includes various groups that differ in their sampling procedures and decision rules (see, e.g., Ellison & Fudenberg, 1993; Young, 1993b; Munshi, 2004). In Appendix A.1 we formally extend our model and results to heterogeneous populations.

The basic model assumes that the decision rule is stationary. In Appendix A.2 we extend our model and results to time-dependent decision rules, and we characterize when a non-stationary sampling process admits a unique sequence of states, such that it converges to this sequence of states from any initial population state.

[12]One can show that in this global attractor a share $\frac{\underline{\alpha}}{1+\underline{\alpha}-\bar{\alpha}}$ of the agents play action a_1. If $\overline{\alpha} = 1$ and $\underline{\alpha} = 0$, then any population state is steady.

Finally, we further extend the model in Appendix A.3 to stochastic shocks that influence the decision rules of all agents (on the aggregate level), and we characterize when the initial population state may have a lasting effect in such sampling processes.

Repeated Interactions without Calendar Time. In many real-life situations agents are randomly matched within a community, and these interactions have been going on since time immemorial. Modeling such situations as repeated games with a definite starting point and strategies that can be conditioned on calendar time may be problematic, as it seems implausible that agents would be aware of the the exact time that has transpired since the starting point, and be aware of the very distant history of play of other agents. An alternative approach is to model behavior in such situations as steady states of environments without a calendar time (see, e.g., Rosenthal 1979; Okuno-Fujiwara & Postlewaite 1995; Heller & Mohlin forthcoming; and the working paper version of Phelan & Skrzypacz, 2006).

An interesting question about such environments is whether the distribution of strategies used by the players to choose their actions as a function of their observations is sufficient to uniquely determine the steady states, or whether the same distribution of rules may admit multiple steady states. Our main result shows that the former is true whenever the expected number of observed actions is less than one, while if the expected number of observed actions is more than one (two), then there is always a distribution of rules with multiple (locally stable) steady states.

Large Finite Populations. Our model studies infinite populations, and it is important to know what the implications of our results are for large finite populations. The key difference between an infinite and a finite population, is that in the former, the law of large numbers implies that the new state of the population is a deterministic function of the initial state and the decision rule. By contrast, in finite populations the new population state is a random variable. If the finite population is sufficiently large then we expect the resulting stochastic process to be close to the deterministic process over finite time horizons. However, when time goes to infinity, rare random events will occasionally take the population away from one (locally stable) steady state towards another steady state (see Sandholm, 2011 for a textbook overview of the deterministic approximation of stochastic evolutionary processes).

When dealing with large finite populations, one may therefore interpret our main result (Theorem 2) as follows. In sampling processes in which $\mu_l < 1$, all learning processes admit a unique globally stable state γ^*. The population is highly likely to quickly converge to state γ^*, and will almost always remain very close to this state. In the rare event that the realized observations of many agents substantially differ from their expected values, the population may temporarily move away from γ^*, but with a very high probability the population will quickly converge back to γ^*.

In sampling processes in which $\mu_l > 1$, there are decision rules that admit multiple steady states. The fact that the population is finite and that the new population state is a random variable will typically quickly take the population away from steady states that are not locally stable. If the sampling process admits multiple locally stable states, then the initial state is highly likely to determine which of these locally stable states the population converges to in the medium run. Moreover the population will likely stay there for a significant amount of time.[13]

[13]The literature on stochastic evolutionary game theory (starting with the pioneering works of Foster & Young, 1990; Kandori *et al.*, 1993; Young, 1993a; see Young, 2015, for a recent survey) studies situations the long-run behavior in environments with

Observations of Action Profiles. In Heller & Mohlin (forthcoming) we investigate sampling processes in which an agent may observe action profiles played in past interactions by the current opponent against her past opponents. All of our results can be extended to this setup, with relatively minor adjustments to the proofs. Specifically one should count an observation of an action profile (in a two-player game) as two actions when calculating the expected number of observed actions μ_l. Our main result still holds in this setup: a sampling process allows a profile of decision rules that admits multiple steady states, essentially, if and only if $\mu_l \leq 1$.

A Extensions

A.1 Heterogeneous Population

The basic model assumes that all agents share the same distribution of sample sizes, and the same decision rule. In many applications the population might be heterogeneous, i.e., the population includes various groups that differ in their sampling procedures and decision rules. A few examples of such models with heterogeneous populations can be found in: (1) Ellison & Fudenberg (1993), who study competing technologies where each technology is better for some of the players and these different tastes induce different decision rules (see also Munshi, 2004); (2) Young (1993b), who studies social learning in a bargaining model in which agents differ in the size of their samples; and (3) Heller & Mohlin (forthcoming), who analyze community enforcement in which the population includes several types of agents, and each type uses a different strategy.

A.1.1 Model with Heterogeneous Population

In what follows we introduce heterogeneous populations that include different types, and we redefine the notions of population state, sampling process, and learning process to deal with this heterogeneity.

Population state. Let Θ denote a finite set of types with a typical element θ. Let λ_θ denote the mass of agents of type θ (or θ-*agents*). For simplicity, we assume that λ has full support. We redefine a *population state* (or *state* for short) to be a vector $\gamma = (\gamma_\theta)_{\theta \in \Theta}$, where each $\gamma_\theta \in \Delta(A)$ denotes the aggregate distribution of actions played by θ-agents. Let $\bar{\gamma} \in \Delta(A)$ denote the average distribution of actions in the population (i.e., $\bar{\gamma}(a) = \sum_\theta \lambda_\theta \gamma_\theta(a)$ for each action $a \in A$). A population state is *uniform* if all types play the same aggregate distribution of actions, i.e., if $\gamma_\theta(a) = \bar{\gamma}(a)$ for each type $\theta \in \Theta$ and action $a \in A$. We redefine Γ to denote the set of all populations with heterogeneous types.

New/Revising agents. In each period, a share of $0 < \beta \leq 1$ of the agents of each type die and are replaced with new agents (or, alternatively, are randomly selected to reevaluate their choice), while the remaining share of $1 - \beta$ of the agents of each type play the same action as they played in the past.

Sample. Each new agent observes a finite sequence of actions (or *sample*). The size of the sample observed by type θ is a random variable with a distribution $\nu_\theta \in \Delta(\mathbb{N})$. Let M, the set of all feasible samples,

multiple locally stable states, and in which there is a small level of noise in the agents' behavior. We think that it would be interesting to extend the methodology of this literature in order to apply it to the setup analyzed in this paper. It might be that such future research can characterize various cases in which, if the population size is sufficiently large, in the long run the population will spend almost all of the time in one of these locally stable states.

be redefined as: $M = \cup_{\theta\in\Theta} \cup_{l\in supp(\nu_\theta)} A^l$. Let $\bar{l} = \max_{l\in} \left(\cup_{\theta\in\Theta} supp\left(\nu_\theta\right)\right) < \infty$ be the maximal sample size. For each sample size $l \in \mathbb{N}$, let $\psi_l : \Gamma \to \Delta\left(A^l\right)$ denote the distribution of samples observed by each agent in the population (or *sampling rule* for short), conditional on the sample having size l. A typical sample of size l is represented by the vector $\overrightarrow{a} = (a_1, ..., a_l)$.

We analyze two kinds of sampling methods in heterogeneous populations:

1. *Observing different random agents*: Each agent independently samples different agents, and observes a random action played by each of these agents. This kind of sampling is a common modeling choice in situations in which an agent's payoff depends not on the behavior of a specific sub-group of opponents, but on the agent's own action, the state of nature, and, possibly, the aggregate behavior of the population (see, e.g., Ellison & Fudenberg, 1995; Banerjee & Fudenberg, 2004). Formally, we define for each sample size $l \in \mathbb{N}$, each state $\gamma \in \Gamma$, and each sample $(a_1, ..., a_l)$,
$$\psi_{l,\gamma}\left(a_1, ..., a_l\right) = \prod_{1\le i\le l} \bar{\gamma}\left(a_i\right). \tag{8}$$

2. *Observing a single random type*: Each agent randomly draws a type $\bar{\theta}$, and then the agent samples different agents of type $\bar{\theta}$, and observes a random action played by each of these $\bar{\theta}$-agents. This kind of observation is relevant to models in which the agent is randomly matched with an opponent, and may sample some actions played in the previous period by agents with the same type as the opponent. Formally, we define for each size $l \in \mathbb{N}$, each state $\gamma \in \Gamma$, and each sample $(a_1, ..., a_l)$,
$$\psi_{l,\gamma}\left(a_1, ..., a_l\right) = \sum_{\theta\in\Theta} \lambda_\theta \cdot \prod_{1\le i\le l} \gamma_\theta\left(a_i\right). \tag{9}$$
In the case of $\beta = 1$, this sampling method has another interpretation that is common in models of strategic interactions among randomly matched agents (e.g., Rosenthal, 1979; Nowak & Sigmund, 1998; Heller & Mohlin, forthcoming). According to this interpretation, each agent is involved in $n \ge \bar{l}$ interactions in each period. In each of these interactions the agent is randomly matched with a different opponent, and the agent observes a sample of random actions played by the opponent in the previous round. The random type of the opponent is distributed according to λ_θ, and each of the actions played by the opponent of type θ in the previous round is distributed according to γ_θ.

Observe that both cases, i.e., (8) and (9), coincide in two special setups: (1) when the population state is uniform (as in the basic model), or (2) when agents observe at most one action (i.e., $\bar{l} = 1$).

Remark 2. Our results work also in a setup in which some types use the first sampling method, while other types use the second sampling method.

Sampling Process. We redefine a *sampling process* as a tuple
$$E = \left(A, \Theta, \beta, \psi_l, \left(\lambda_\theta, \nu_\theta\right)_{\theta\in\Theta}\right)$$
that includes the six components described above: a finite set of actions A, a finite set of types Θ, a fraction of new agents at each stage β, a sampling rule ψ_l (satisfying either (8) or (9)), a distribution over the set of types λ, and a profile of distributions of sample sizes $(\nu_\theta)_{\theta\in\Theta}$.

Given sampling process $E = \left(A, \Theta, \beta, \psi_l, (\lambda_\theta, \nu_\theta)_{\theta\in\Theta}\right)$, let μ_l, the *mean sample size,* be redefined as the expected number of actions observed by a random agent in the population. Formally:

$$\mu_l = \sum_{\theta\in\Theta} \lambda_\theta \sum_{l\in supp(\nu_\theta)} \nu_\theta (l) \cdot l.$$

Decision rule and learning process. Each new θ-agent chooses his action in the new population state by following a stationary (i.e., time-independent) *decision rule* $\sigma_\theta : M \to \Delta(A)$. That is, a new θ-agent who observes sample $m \in M$ plays action a with probability $\sigma_{\theta,m}(a)$. The remaining $1-\beta$ incumbent agents play the same action as in the previous round. A profile of decision rules $(\sigma_\theta)_{\theta\in\Theta}$ is *uniform* if all types use the same decision rule, i.e., if $\sigma_\theta = \sigma_{\theta'}$ for each type $\theta, \theta' \in \Theta$.

A learning process is a pair

$$P = \left(E, (\sigma_\theta)_{\theta\in\Theta}\right) = \left(A, \Theta, \beta, \psi_l, (\lambda_\theta, \nu_\theta, \sigma_\theta)_{\theta\in\Theta}\right),$$

consisting of a sampling process and a decision rule.

As in the basic model, let $f_P : \Gamma \to \Gamma$ denote the mapping between states induced by a single step of the learning process P.

L_1-distance. Each population state $\gamma \in \Gamma$ corresponds to a distribution $q_\gamma \in \Delta(\Theta \times A)$ as follows: $q_\gamma(\theta, a) = \lambda_\theta \cdot \gamma_\theta(a)$. We define the distance between two population states $\gamma, \gamma' \in \Gamma$ as the L_1-distance between the corresponding distributions $q_\gamma, q_{\gamma;} \in \Delta(\Theta \times A)$:

$$\|\gamma - \gamma'\|_1 = \|q_\gamma - q_{\gamma'}\|_1 = \sum_{\theta\in\Theta}\sum_{a\in A} |\lambda_\theta \cdot \gamma_\theta(a) - \lambda_\theta \cdot \gamma'_\theta(a)| = \sum_{\theta\in\Theta} \lambda_\theta \cdot \|\gamma_\theta - \gamma'_\theta\|_1.$$

A.1.2 Generalizing Results

In what follows we formally show how to generalize the first result (Theorem 1) to heterogeneous populations.

Theorem 8. *(Generalization of Theorem 1) Let $P = \left(A, \Theta, \beta, \psi_l, (\lambda_\theta, \nu_\theta, \sigma_\theta)_{\theta\in\Theta}\right)$ be a learning process, and let $\gamma \neq \gamma' \in \Gamma$ be two population states. Then:*

$$\|f_P(\gamma) - f_P(\gamma')\|_1 \leq (1 - \beta + \beta \cdot \mu_l) \cdot \|\gamma - \gamma'\|_1,$$

with a strict inequality if there exist a type θ and an $l > 1$ such that $\nu_\theta(l) > 0$.

The intuition is similar to Theorem 1. The proof is presented in Appendix B.

Similarly to the generalization of Theorem 1 above, one can generalize in a straightforward way all the other results of the paper to the setup of a heterogeneous population (proofs omitted for brevity).

A.2 Non-Stationary Learning Process

In this section we further extend the model to deal with non-stationary deterministic learning processes, in which the process explicitly depends on calendar time, and we show how to generalize our results to this setup.

Adaptations to the model. For each period $t \geq 1$, let $\beta^t \in [0,1]$ denote the random share of agents who revise their actions in period t. For each type $\theta \in \Theta$ and period $t \geq 1$, let $\nu_\theta^t \in \Delta(\mathbb{N})$ denote the distribution of sample sizes of type θ in period t. To simplify the notation we assume that the support of the sample sizes of each type is independent of the period, i.e., $supp\left(\nu_\theta^{t_1}\right) = supp\left(\nu_\theta^{t_2}\right) := supp\left(\nu_\theta\right)$ for each type $\theta \in \Theta$ and periods $t_1, t_2 \geq 1$. As in the basic model, let M denote the set of all feasible sample sizes. A *non-stationary sampling process* is a tuple

$$E = \left(A, \Theta, \left(\beta^t\right)_{t\in\mathbb{N}}, \psi_l, \left(\lambda_\theta\right)_{\theta\in\Theta}, \left(\nu_\theta^t\right)_{\theta\in\Theta, t\geq 1}\right).$$

Given a non-stationary sampling process, let μ_l^t denote the expected number of actions observed in period t, i.e., $\mu_l^t = \sum_{\theta\in\Theta}\lambda_\theta \sum_{l\in supp(\nu_\theta)} \nu_\theta^t(l)\cdot l$.

Given a non-stationary sampling process E, let $\bar{\mu}_l$ be the upper limit of the geometric mean of $1-\beta^t\cdot\left(1-\mu_l^t\right)$ as t goes to to infinity, i.e.,

$$\bar{\mu}_l = \text{limsup}_{\hat{t}\to\infty}\sqrt[\hat{t}]{\prod_{t\leq\hat{t}}\left(1-\beta^t+\beta^t\cdot\mu_l^t\right)}.$$

For each type $\theta \in \Theta$ and period $t \geq 1$, let $\sigma_\theta^t : M \to \Delta(A)$ denote the non-stationary decision rule of new θ-agents in period t. A *non-stationary learning process* is a pair consisting of a non-stationary sampling process and a non-stationary decision rule, i.e.,

$$P = \left(E, \left(\sigma_\theta^t\right)_{\theta\in\Theta, t\geq 1}\right) = \left(A, \Theta, \left(\beta^t\right)_{t\geq 1}, \psi_l, \left(\lambda_\theta\right)_{\theta\in\Theta}, \left(\nu_\theta^t, \sigma_\theta^t\right)_{\theta\in\Theta, t\geq 1}\right).$$

As in the basic model, a non-stationary learning process P and an initial state uniquely determine a new state in each period t. Let $f_p^t(\hat{\gamma}) \in \Gamma$ denote the state induced after t stages of the non-stationary learning process P.

A sequence of states $\left(\gamma_t^*\right)_{t\in\mathbb{N}}$ is a *global attractor* of the non-stationary learning process P, if

$$\lim_{t\longrightarrow\infty}\left\|f_P^t(\hat{\gamma}) - \gamma_t^*\right\|_1 = 0$$

for each initial state $\hat{\gamma} \in \Gamma$.

Adapted results. Minor adaptations to the proof of Theorem 8 and a simple inductive argument imply that the distance between two states at time t_o is at most $\prod_{t\leq t_0}\left(1-\beta^t+\beta^t\cdot\mu_l^t\right)$ times the initial distance.[14] Formally:

Corollary 4. *Let* $P = \left(A, \Theta, \left(\beta^t\right)_{t\geq 1}, \psi_l, \left(\lambda_\theta\right)_{\theta\in\Theta}, \left(\nu_\theta^t, \sigma_\theta^t\right)_{\theta\in\Theta, t\geq 1}\right)$ *be a non-stationary learning process, let* $\hat{\gamma}, \hat{\gamma}' \in \Gamma$ *be two population states, and let* $\hat{t} \geq 1$. *Then:*

$$\left\|f_p^{\hat{t}}(\hat{\gamma}) - f_p^{\hat{t}}(\hat{\gamma}')\right\|_1 \leq \left\|\hat{\gamma} - \hat{\gamma}'\right\|_1 \cdot \prod_{t\leq\hat{t}}\left(1-\beta^t+\beta^t\cdot\mu_l^t\right).$$

[14] Specifically, applying Theorem 8 to the first round implies that the distance at $t = 1$ is at most $\left(1-\beta^1+\beta^1\cdot\mu_l^1\right)$ times the initial distance. Applying Theorem 8 again, this time to the second round, implies that that the distance at $t = 2$ is at most $\left(1-\beta^1+\beta^1\cdot\mu_l^1\right)\cdot\left(1-\beta^2+\beta^2\cdot\mu_l^2\right)$ times the initial distance. By induction, the distance after t_0 rounds is at most $\prod_{t\leq t_0}\left(1-\beta^t+\beta^t\cdot\mu_l^t\right)$ the initial distance.

This, in turn, immediately implies that in any non-stationary sampling process in which $\bar{\mu}_l < 1$, any profile of non-stationary decision rules admits a global attractor. Formally:

Corollary 5. *Let* $E = \left(A, \Theta, (\beta^t)_{t\geq 1}, \psi_l, (\lambda_\theta)_{\theta\in\Theta}, (\nu_\theta^t)_{\theta\in\Theta, t\geq 1}\right)$ *be a non-stationary sampling process satisfying* $\bar{\mu}_l < 1$. *Then for any profile of non-stationary decision rules* $(\sigma_\theta^t)_{\theta\in\Theta, t\geq 1}$, *the non-stationary learning process* $P = \left(E, (\sigma_\theta^t)_{\theta\in\Theta, t\geq 1}\right)$ *admits a global attractor.*

The example presented in Case A of the proof of Theorem 2 demonstrates that the above bound of $\bar{\mu}_l < 1$ is binding in the sense that there is a sampling process with $\bar{\mu}_l = 1$ that admits a profile of decision rules with multiple steady states.

The adaptation of the remaining results to the time-dependent setup is similar (proof omitted for brevity).

A.3 Process with Common Shocks

In this section we further extend our model to deal also with common stochastic shocks to the decision rules.

Additional adaptations to the model. In what follows we further adapt the model of Section A.2 by allowing common stochastic shocks to the decision rules of the agents.

Let $(\Omega, \mathcal{F}, p)$ be an arbitrary probability space. Each element $\omega \in \Omega$ represents the state of nature, which determines the realizations of all common shocks to the decision rules in all periods. For each type $\theta \in \Theta$ and period $t \in \mathbb{N}$, let $\sigma_\theta^t : \Omega \times M \to \Delta(A)$ denote the state-dependent decision rule of new θ-agents in period t.

Our interpretation of the state-dependent decision rule σ_θ^t is as follows. The state of nature determines a distribution of noisy signals from which each new agent draws a signal. Based on this noisy signal as well as on the sample of past actions (and on information about calendar time), each new agent chooses an action. The choices of actions (which depend on the noisy signals) and the distribution of noisy signals jointly generate a distribution of actions that depend only on the state of nature (and on calendar time), which is captured by the state-dependent decision rule σ_θ^t.

A *learning process* with common shocks is a pair consisting of a non-stationary sampling process and a state-dependent decision rule, i.e., $P = \left(E, (\sigma_\theta^t)_{\theta\in\Theta, t\geq 1}\right) = \left(A, \Theta, (\beta^t)_{t\geq 1}, \psi_l, (\lambda_\theta)_{\theta\in\Theta}, (\nu_\theta^t, \sigma_\theta^t)_{\theta\in\Theta, t\geq 1}\right)$.

Learning processes with commons shocks are important in modeling situations in which there are stochastic factors that influence the decision rules of all new agents in period t. For example , Ellison & Fudenberg (1995) model a situation in which new agents in period t choose between two agricultural technologies, and each such new agent observes a noisy signal about the expected payoff of each technology conditional on the weather in period t (which is common to all agents), where the (unknown) state of nature determines the weather in all periods.

The state of nature, the learning process, and the initial population state uniquely determine the population state in each period. Let $f_p^t(\omega)(\hat{\gamma}) \in \Gamma$ denote the population state induced after t stages of the non-stationary learning process P, given an initial population state $\hat{\gamma}$, and state of nature $\omega \in \Omega$.

We say that a sequence of state-dependent population states $(\gamma_t^*)_{t\geq 1}$, where $\gamma_t^* : \Omega \to \Gamma$, is a *state-dependent global attractor* of the learning process with commons shocks P if, for each $\omega \in \Omega$, $\lim_{t\longrightarrow\infty} \left\| f_P^t(\omega)(\hat{\gamma}) - \gamma_t^*(\omega) \right\|_1 = 0$ for each initial state $\hat{\gamma} \in \Gamma$.

Example 4 below demonstrates how to apply the extended model to a social learning process with competing technologies with common shocks:

Example 4 (Competing Technologies with Common Shocks)**.** Consider a stochastic sampling process in which there are two possible regimes $\{1,2\}$. There are two technologies: a_1 and a_2. Technology a_1 is advantageous in regime 1, while technology a_2 is advantageous in regime 2. There is a uniform common prior about the regime in round 1. In each subsequent round, the regime is the same as in the previous round with probability 99%, and it is a new regime with probability 1%. In each round, a share of 25% of the incumbents die, and are replaced with new agents. Each new agent observes the action of a single random incumbent and a noisy signal about the current regime, and based on these observations, the agent chooses one of the two technologies. Assume that the decision rule used by the agents implies that each new agent plays action a_1:

1. with a probability of 95% after observing action a_1 in regime 1;
2. with a probability of 80% after observing action a_1 in regime 2;
3. with a probability of 20% after observing action a_2 in regime 1;
4. with a probability of 5% after observing action a_2 in regime 2.

One can show that the sampling process admits a unique steady state that is a state-dependent global attractor. The induced aggregate behavior of the population converges towards playing action a_1 with an average probability of 80% in regime 1, and it converges towards playing action a_1 with an average probability of 20% in regime 2.

This learning process with common shocks is modeled as

$$P=\left(\{a_1,a_2\},\{\theta\},\left(\beta^t\equiv 25\%\right)_{t\in\mathbb{N}},\psi_l,\lambda_\theta,\left(\nu_\theta^t\equiv 1,\sigma_\theta^t\right)_{t\geq 1}\right).$$

The set of states of nature $\Omega=\left\{(\omega_n)_{n\in\mathbb{N}}\right\}$ is the set of infinite binary sequences, where each $\omega_n\in\{1,2\}$ describes the regime in round n. The definition of $(\mathcal{F},p)$ is derived from the Markovian process determining the regime in each round in a standard way. Given state $\omega=(\omega_n)_{n\in\mathbb{N}}$, let the decision rule be defined as follows:

$$\sigma_\theta\left(a_1,\omega\right)=\begin{cases}95\% & a=a_1 \text{ and } \omega_t=1\\ 80\% & a=a_1 \text{ and } \omega_t=2\\ 20\% & a=a_2 \text{ and } \omega_t=1\\ 5\% & a=a_2 \text{ and } \omega_t=2.\end{cases}$$

Adapted Results. Minor adaptations to the proof of Theorem 8 imply that the distance between two states at time $\hat{t}$ is at most $\prod_{t\leq\hat{t}}\left(1-\beta^t+\beta^t\cdot\mu_l^t\right)$ the initial distance. Formally:

Corollary 6. *Let $P=\left(A,\Theta,\left(\beta^t\right)_{t\geq 1},\psi_l,\left(\lambda_\theta\right)_{\theta\in\Theta},\left(\nu_\theta^t,\sigma_\theta^t\right)_{\theta\in\Theta,t\geq 1}\right)$ be a learning process with commons shocks, let $\hat{\gamma},\hat{\gamma}'\in\Gamma$ be two population states, and let $\hat{t}\in\mathbb{N}$. Then, for each $\omega\in\Omega$,*

$$\left\|f_p^{\hat{t}}\left(\omega\right)\left(\hat{\gamma}\right)-f_p^{\hat{t}}\left(\omega\right)\left(\hat{\gamma}'\right)\right\|_1\leq\left\|\hat{\gamma}-\hat{\gamma}'\right\|_1\cdot\prod_{t\leq\hat{t}}\left(1-\beta^t+\beta^t\cdot\mu_l^t\right)\cdot$$

An immediate corollary is that any sampling process with common shocks in which $\bar{\mu}_l<1$, given any profile of decision rules, admits a state-dependent global attractor. That is, in the long run, the population's behavior depends only on the state of nature, but it is independent of the initial population state in time zero. Formally:

Corollary 7. *Let* $E = \left(A, \Theta, (\beta^t)_{t\geq 1}, \psi_l, (\lambda_\theta)_{\theta\in\Theta}, (\nu_\theta^t)_{\theta\in\Theta, t\geq 1}\right)$ *be a sampling process satisfying* $\bar{\mu}_l < 1$. *Then for any profile of stochastic decision rules* $(\sigma_\theta^t)_{\theta\in\Theta, t\geq 1}$*, the learning process with common shocks* $P = \left(E, (\sigma_\theta^t)_{\theta\in\Theta, t\geq 1}\right)$ *admits a state-dependent global attractor.*

The adaptation of the remaining results to the time-dependent setup is similar to the adaptation above (proof omitted for brevity).

B Formal Proofs

B.1 Proof of Theorem 8 (Upper Bound Result; Generalization of Theorem 1)

The distance between the new population states is bounded as follows (where the second inequality is strict if $\nu_\theta(l) > 0$ for some $\theta \in \Theta$ and $l \geq 2$):

$$\|(f_P(\gamma)) - (f_P(\gamma'))\|_1 \leq \beta \cdot \sum_{\theta\in\Theta} \lambda_\theta \cdot \sum_{l\in\mathbb{N}} \nu_\theta(l) \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1-\beta) \cdot \|\gamma - \gamma'\|_1 \leq$$

$$\beta \cdot \sum_{\theta\in\Theta} \lambda_\theta \cdot \sum_{l\in\mathbb{N}} \nu_\theta(l) \cdot l \cdot \|\gamma - \gamma'\|_1 + (1-\beta) \cdot \|\gamma - \gamma'\|_1 =$$

$$\left(\beta \cdot \left(\sum_{\theta\in\Theta} \lambda_\theta \cdot \sum_{l\in\mathbb{N}} \nu_\theta(l) \cdot l\right) + (1-\beta)\right) \cdot \|\gamma - \gamma'\|_1 = (\beta \cdot \mu_L + 1 - \beta) \cdot \|\gamma - \gamma'\|_1 = (1 - \beta \cdot (1 - \mu_l)) \cdot \|\gamma - \gamma'\|_1 .$$

The first inequality is proven in Lemma 1. The second inequality (which is strict if $\nu_\theta(l) > 0$ for some $\theta \in \Theta$ and $l \geq 2$) is implied by the inequality

$$\|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 \leq l \cdot \|\gamma - \gamma'\|_1 \text{ (with a strict inequality if } l \geq 2),$$

which is proven in Lemma 4.

Proofs of the various Lemmas used in the Proof of Theorem 8

Lemma 1. *For each sampling process* E *and states* $\gamma \neq \gamma' \in \Gamma$,

$$\|(f_P(\gamma)) - (f_P(\gamma'))\|_1 \leq \beta \cdot \sum_{\theta\in\Theta} \lambda_\theta \cdot \sum_{l\in\mathbb{N}} \nu_\theta(l) \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1-\beta) \cdot \|\gamma - \gamma'\|_1 .$$

Proof.

$$\|(f_P(\gamma)) - (f_P(\gamma'))\|_1 = \sum_{\theta\in\Theta} \lambda_\theta \cdot \|(f_P(\gamma))_\theta - (f_P(\gamma'))_\theta\|_1 \leq$$

$$\sum_{\theta\in\Theta} \lambda_\theta \cdot \left(\beta \cdot \sum_{l\in\mathbb{N}} \nu_\theta(l) \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1-\beta) \cdot \|\gamma_\theta - \gamma'_\theta\|_1\right) =$$

$$\beta \cdot \sum_{\theta\in\Theta} \lambda_\theta \cdot \sum_{l\in\mathbb{N}} \nu_\theta(l) \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1-\beta) \cdot \sum_{\theta\in\Theta} \lambda_\theta \cdot \|\gamma_\theta - \gamma'_{\theta'}\|_1 =$$

$$\beta \cdot \sum_{\theta \in \Theta} \lambda_\theta \cdot \sum_{l \in \mathbb{N}} \nu_\theta (l) \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1-\beta) \cdot \|\gamma - \gamma'\|_1 ,$$

where the inequality is due to Lemma 2. □

Lemma 2. *For each sampling process E, type $\theta \in \Theta$, and each two states $\gamma \neq \gamma' \in \Gamma$:*

$$\|(f_P(\gamma))_\theta - (f_P(\gamma'))_\theta\|_1 \le \beta \cdot \sum_{l \in \mathbb{N}} \nu_\theta (l) \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1-\beta) \cdot \|\gamma_\theta - \gamma'_\theta\|_1 .$$

Proof.

$$\|(f_P(\gamma))_\theta - (f_P(\gamma'))_\theta\|_1 = \sum_{a \in A} |(f_P(\gamma))_\theta (a) - (f_P(\gamma'))_\theta (a)| =$$

$$\sum_{a \in A} \left| \left(\sum_{l \in supp(\nu_\theta)} \beta \cdot \nu_\theta (l) \sum_{m \in A^l} \psi_{l,\gamma}(m) \cdot \sigma_{\theta,m} + (1-\beta) \cdot \gamma_\theta \right)(a) \right.$$

$$\left. - \left(\beta \cdot \sum_{l \in supp(\nu_\theta)} \nu_\theta (l) \cdot \sum_{m \in A^l} \psi_{l,\gamma'}(m) \cdot \sigma_{\theta,m} + (1-\beta) \cdot \gamma'_\theta \right)(a) \right| =$$

$$\sum_{a \in A} \left| \beta \cdot \sum_{l \in supp(\nu_\theta)} \nu_\theta (l) \cdot \sum_{m \in A^l} (\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)) \cdot \sigma_{\theta,m}(a) + (1-\beta) \cdot (\gamma_\theta(a) - \gamma'_\theta(a)) \right| \le \tag{10}$$

$$\sum_{a \in A} \left(\beta \cdot \sum_{l \in supp(\nu_\theta)} \nu_\theta (l) \cdot \left| \sum_{m \in A^l} (\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)) \cdot \sigma_{\theta,m}(a) \right| + (1-\beta) \cdot |\gamma_\theta(a) - \gamma'_\theta(a)| \right) =$$

$$\beta \cdot \sum_{l \in supp(\nu_\theta)} \nu_\theta (l) \cdot \sum_{a \in A} \left| \sum_{m \in A^l} (\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)) \cdot \sigma_{\theta,m}(a) \right| + (1-\beta) \cdot \sum_{a \in A} |\gamma_\theta(a) - \gamma'_\theta(a)| \le \tag{11}$$

$$\beta \cdot \sum_{l \in supp(\nu_\theta)} \nu_\theta (l) \cdot \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 + (1-\beta) \cdot \|\gamma_\theta - \gamma'_{\theta'}\|_1 ,$$

where the (10) is a triangle inequality, and (11) is due to Lemma 3. □

Lemma 3. *For each sampling process E, each size $l \in \mathbb{N}$, each type $\theta \in \Theta$, and any two states $\gamma \neq \gamma' \in \Gamma$:*

$$\sum_{a \in A} \left| \sum_{m \in A^l} (\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)) \cdot \sigma_{\theta,m}(a) \right| \le \|\psi_{l,\gamma} - \psi_{l,\gamma'}\|_1 .$$

Proof.

$$\begin{aligned}
\sum_{a \in A} \left| \sum_{m \in A^l} (\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)) \cdot \sigma_{\theta,m}(a) \right| &\le \sum_{a \in A} \sum_{m \in A^l} |\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)| \cdot \sigma_{\theta,m}(a) \\
&= \sum_{m \in A^l} |\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)| \cdot \sum_{a \in A} \sigma_{\theta,m}(a) \\
&= \sum_{m \in A^l} |\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)| \cdot 1,
\end{aligned}$$

where the inequality is a triangle inequality. $\square$

Lemma 4. *For each sampling process E, type $\theta \in \Theta$, sample size $l \in \mathbb{N}$, and states $\gamma \neq \gamma' \in \Gamma$*

$$\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1 \leq l \cdot \left\|\gamma - \gamma'\right\|_1,$$

with a strict inequality if $l > 1$.

Proof. Case I - Observing different random agents:

$$\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1 = \sum_{\vec{a} \in A^l} \left|\psi_{l,\gamma}\left(\vec{a}\right) - \psi_{l,\gamma'}\left(\vec{a}\right)\right| = \tag{12}$$

$$\sum_{\vec{a} \in A^l} \left|\prod_{1 \leq i \leq l} \bar{\gamma}\left(a_i\right) - \prod_{1 \leq i \leq l} \bar{\gamma}'\left(a_i\right)\right| = \tag{13}$$

$$\sum_{\vec{a} \in A^l} \left|\sum_{1 \leq i \leq l} \left(\bar{\gamma}\left(a_i\right) - \bar{\gamma}'\left(a_i\right)\right) \cdot \prod_{i < j \leq l} \bar{\gamma}\left(a_j\right) \cdot \prod_{1 \leq k < i} \bar{\gamma}'\left(a_k\right)\right| \leq \quad (< \text{if } l > 1) \tag{14}$$

$$\sum_{\vec{a} \in A^l} \left(\sum_{1 \leq i \leq l} \left|\bar{\gamma}\left(a_i\right) - \bar{\gamma}'\left(a_i\right)\right| \cdot \prod_{i < j \leq l} \bar{\gamma}\left(a_j\right) \cdot \prod_{1 \leq k < i} \bar{\gamma}'\left(a_k\right)\right) =$$

$$\sum_{1 \leq i \leq l} \left(\sum_{\vec{a} \in A^l} \left|\bar{\gamma}\left(a_i\right) - \bar{\gamma}'\left(a_i\right)\right| \cdot \prod_{i < j \leq l} \bar{\gamma}\left(a_j\right) \cdot \prod_{1 \leq k < i} \bar{\gamma}'\left(a_k\right)\right) =$$

$$\sum_{1 \leq i \leq l} \left(\sum_{a_i \in A} \left|\bar{\gamma}\left(a_i\right) - \bar{\gamma}'\left(a_i\right)\right|\right) \cdot \left(\sum_{\left(a_{i+1}, \ldots, a_l\right) \in A^{l-i}} \prod_{i < j \leq l} \bar{\gamma}\left(a_j\right)\right) \cdot \left(\sum_{\left(a_1, \ldots, a_{i-1}\right) \in A^{i-1}} \prod_{1 \leq k < i} \bar{\gamma}'\left(a_k\right)\right) = \tag{15}$$

$$\sum_{1 \leq i \leq l} \left(\sum_{a_i \in A} \left|\bar{\gamma}\left(a_i\right) - \bar{\gamma}'\left(a_i\right)\right|\right) \cdot 1 \cdot 1 = \sum_{1 \leq i \leq l} \left(\left\|\bar{\gamma} - \bar{\gamma}'\right\|_1\right) = l \cdot \left\|\bar{\gamma} - \bar{\gamma}'\right\|_1 \leq l \cdot \left\|\gamma - \gamma'\right\|.$$

Eq. (12) is due to the independence of different observations. Eq. (13) is implied by adding to the sum elements that cancel out. Specifically, let $b_i = \bar{\gamma}\left(a_i\right)$ and $c_i = \bar{\gamma}'\left(a_i\right)$; then due to a "telescoping series" argument (in which each new element appears once with a positive sign and once with a negative sign):[15]

$$\prod_{1 \leq i \leq l} \bar{\gamma}\left(a_i\right) - \prod_{1 \leq i \leq l} \bar{\gamma}'\left(a_i\right) = \prod_{1 \leq i \leq l} b_i - \prod_{1 \leq i \leq l} c_i =$$

$$\left(b_1 \cdot \ldots \cdot b_l - c_1 \cdot b_2 \cdot \ldots \cdot b_l\right) + \left(c_1 \cdot b_2 \cdot \ldots \cdot b_l - c_1 \cdot c_2 \cdot b_3 \cdot \ldots \cdot b_l\right) + \left(c_1 \cdot c_2 \cdot b_3 \cdot \ldots \cdot b_l - \ldots\right) + \ldots + \left(\ldots - c_1 \cdot \ldots \cdot c_l\right) =$$

$$\left(b_1 - c_1\right) \cdot \left(b_2 \cdot \ldots \cdot b_l\right) + \left(b_2 - c_2\right) \cdot \left(b_3 \cdot \ldots \cdot b_l \cdot c_1\right) + \left(b_3 - c_3\right) \cdot \left(b_4 \cdot \ldots \cdot b_l \cdot c_1 \cdot c_2\right) + \ldots + \left(b_l - c_l\right) \cdot \left(c_1 \cdot \ldots \cdot c_{l-1}\right) =$$

[15] We use the convention that a product of an empty set (e.g., $\prod_{1 \leq j < 1}$) is equal to one.

$$= \sum_{1\le i\le l}\left((b_i - c_i)\cdot\prod_{i<j\le l} b_j \cdot \prod_{1\le j<i} c_j\right) = \sum_{1\le i\le l}\left(\bar{\gamma}(a_i) - \bar{\gamma}'(a_i)\right)\cdot\prod_{i<j\le l}\bar{\gamma}(a_j)\cdot\prod_{1\le k<i}\bar{\gamma}'(a_k).$$

Eq. (14) is a triangle inequality, and it is strict if $l > 1$ because the sum inside the "$\Big|\ \Big|$" in (14) includes both positive and negative elements. Eq. (15) holds because each sum adds the probabilities of disjoint and exhausting events. The final inequality is implied by Lemma 5.

Case II - Observing a single random type:

$$\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1 = \sum_{\vec{a}\in A^l}\left|\psi_{l,\gamma}(\vec{a}) - \psi_{l,\gamma'}(\vec{a})\right| = \tag{16}$$

$$\sum_{\vec{a}\in A^l}\left|\sum_{\theta\in\Theta}\lambda_\theta\cdot\left(\prod_{1\le i\le l}\gamma_\theta(a_i) - \prod_{1\le i\le l}\gamma'_\theta(a_i)\right)\right| = \tag{17}$$

$$\sum_{\vec{a}\in A^l}\left|\sum_{\theta\in\Theta}\lambda_\theta\cdot\left(\sum_{1\le i\le l}\left(\gamma_\theta(a_i) - \gamma'_\theta(a_i)\right)\cdot\prod_{i<j\le l}\gamma_\theta(a_j)\cdot\prod_{1\le j<i}\gamma'_\theta(a_j)\right)\right| \le \quad (< \text{if } l>1) \tag{18}$$

$$\sum_{\vec{a}\in A^l}\sum_{\theta\in\Theta}\lambda_\theta\cdot\left(\sum_{1\le i\le l}\left|\gamma_\theta(a_i) - \gamma'_\theta(a_i)\right|\cdot\prod_{i<j\le l}\gamma_\theta(a_j)\cdot\prod_{1\le j<i}\gamma'_\theta(a_j)\right) =$$

$$\sum_{1\le i\le l}\sum_{\theta\in\Theta}\lambda_\theta\cdot\left(\sum_{\vec{a}\in A^l}\left|\gamma_\theta(a_i) - \gamma'_\theta(a_i)\right|\cdot\prod_{i<j\le l}\gamma_\theta(a_j)\cdot\prod_{1\le j<i}\gamma'_\theta(a_j)\right) =$$

$$\sum_{1\le i\le l}\sum_{\theta\in\Theta}\lambda_\theta\cdot\left(\sum_{a_i\in A}\left|\gamma_\theta(a_i) - \bar{\gamma}_\theta{}'(a_i)\right|\right)\cdot\left(\sum_{(a_{i+1},\ldots,a_l)\in A^{l-i}}\prod_{i<j\le l}\gamma_\theta(a_j)\right)\cdot\left(\sum_{(a_i,\ldots,a_{i-1})\in A^{i-1}}\prod_{1\le j<i}\gamma'_\theta(a_j)\right) = \tag{19}$$

$$\sum_{1\le i\le l}\sum_{\theta\in\Theta}\lambda_\theta\cdot\left(\sum_{a_i\in A}\left|\gamma_\theta(a_i) - \bar{\gamma}_\theta{}'(a_i)\right|\right)\cdot 1\cdot 1 = \sum_{1\le i\le l}\sum_{\theta\in\Theta}\lambda_\theta\cdot\left\|\gamma_\theta - \gamma'_\theta\right\|_1 = \sum_{1\le i\le l}\left\|\gamma - \gamma'\right\|_1 = l\cdot\left\|\gamma - \gamma'\right\|_1.$$

Eq. (16) is due to the different observations being independent conditional on the observed type θ. Eq. (17) is implied by adding to the sum elements that cancel out (i.e., a "telescoping series"). Eq. (18) is a triangle inequality, and it is strict if $l > 1$ because the sum inside the "$\Big|\ \Big|$" in (18) includes both positive and negative elements. Eq. (19) holds because each sum adds the probabilities of disjoint and exhausting events. □

Lemma 5. *$\left\|\bar{\gamma} - \bar{\gamma}'\right\|_1 \le \left\|\gamma - \gamma'\right\|_1$ for each two states $\gamma \ne \gamma' \in \Gamma$.*

Proof.

$$\left\|\gamma - \gamma'\right\|_1 = \sum_{\theta\in\Theta}\lambda_\theta\cdot\left\|\gamma_\theta - \gamma'_\theta\right\|_1 = \sum_{\theta\in\Theta}\lambda_\theta\cdot\sum_{a\in A}\left|\gamma_\theta(a) - \gamma'_\theta(a)\right| = \sum_{a\in A}\sum_{\theta\in\Theta}\lambda_\theta\cdot\left|\gamma_\theta(a) - \gamma'_\theta(a)\right| \ge$$

$$\sum_{a\in A}\left|\sum_{\theta\in\Theta}\lambda_\theta\left(\gamma_\theta(a) - \gamma'_\theta(a)\right)\right| = \sum_{a\in A}\left|\sum_{\theta\in\Theta}\lambda_\theta\gamma_\theta(a) - \sum_{\theta\in\Theta}\lambda_\theta\gamma'_\theta(a)\right| = \sum_{a\in A}\left|\bar{\gamma}(a) - \bar{\gamma}'(a)\right| = \left\|\bar{\gamma} - \bar{\gamma}'\right\|_1,$$

where the various equalities are immediately implied by the definitions on the L_1-norm and $\bar{\gamma}$, and the inequality is a triangle inequality. □

B.2 Proof of Theorem 3 ($\mu_l > 2$)

For each $0 < q < \frac{1}{\mu_l}$ define σ_q as the decision rule according to which each agent plays action a^* if he has observed action a^* at least twice, plays action a' if he has not observed action a^*, and he plays action a^* with probability q and action a' with the remaining probability $1-q$ if he has observed a^* exactly once; that is, for each $a^l \in A^l$,

$$\sigma_q\left(a^l\right) = \begin{cases} a^* & \left|\left\{i|a_i^l = a^*\right\}\right| \geq 2 \\ q \cdot a^* + (1-q) \cdot a' & \left|\left\{i|a_i^l = a^*\right\}\right| = 1 \\ a' & \left|\left\{i|a_i^l = a^*\right\}\right| = 0. \end{cases}$$

Observe that new agents play only a^* or a'. Further note that the probability that a new agent plays a^* depends only on the frequency with which the incumbents play action a^* (and not on any other aspect of the population state). Thus, with slight abuse of notation, we identify a state $\gamma \in \Delta(A)$ with the frequency x of agents who choose action a^*, i.e., $x := \gamma(a)$ (as the actions played by the remaining $1-x$ of the agents do not play any role in the dynamics, and in the long run each agent plays either action a^* or a'). The mapping induced by the learning process $P_q = (E, \sigma_q)$ is given by the function $f_q : [0,1] \to [0,1]$:

$$f_q(x) := f_{(P_q)}(x) = (1-\beta) \cdot x + \beta \cdot \left(q \cdot \mu_l \cdot x + (1 - 2 \cdot q) \cdot \sum_{2 \leq l \in supp(\nu)} \nu(l) \cdot \begin{pmatrix} l \\ 2 \end{pmatrix} \cdot x^2 + O\left(x^3\right)\right). \quad (20)$$

The argument for (20) is as follows. The term of $(1-\beta) \cdot x$ describes the behavior of incumbents who have not died. The terms multiplying β represent the behavior of new agents. The first of these terms $(q \cdot \mu_l \cdot x)$ derives from the fact that the first-order approximation for the probability that a new agent plays action a' is q times the expected number of times that action a' is observed $(\mu_l \cdot x)$, since action a' is almost always observed once in a sample. The second term multiplying β in (20) reflects the correction required to adjust the above first-order approximation due to the fact that an agent who observes action a' twice in the sample plays action a' with probability 1, rather than with probability $2 \cdot q$. Hence, the additional probability of playing a^* conditional on observing a^* twice is $(1 - 2 \cdot q)$. The probability of observing a^* twice in a random sample is $\sum_{2 \leq l \in supp(\nu)} \nu(l) \cdot \begin{pmatrix} l \\ 2 \end{pmatrix} \cdot x^2 + O\left(x^3\right)$. Finally, note that the probability of observing a^* three or more times is negligible (i.e., $O\left(x^3\right)$), so that the remaining adjustment required for (20) to coincide with the dynamic mapping induced by P_q is $O\left(x^3\right)$.

It is immediate from the definition of decision rule σ^* that $f_q(x)$ is strictly increasing in x. Recall that state x is a steady state iff $f_q(x) = x$. Observe that: $f_q(0) = 0$ and $f_q'(0) = q \cdot \mu_l < 1$ for each $q < \frac{1}{\mu_l}$. Fix $q < \frac{1}{\mu_l}$. The previous observations imply that there exists $\bar{x} > 0$ such that $f_q(x) < x$ and $f_q'(x) < 1$ for each $x \in (0, \bar{x})$. This implies that $f_q^t(x) < x$ and $\lim_{t \to \infty} f_q^t(x) = 0$ for each $x \in (0, \bar{x})$, and hence state 0 is locally stable.

Assume that $\nu(0) = \nu(1) = 0$. Then, the definition of decision rule σ^* implies that $f_q(1) = 1$, and that for each $\epsilon << 1$ it holds that

$$f_q(1-\epsilon) = (1-\beta) \cdot (1-\epsilon) + \beta \cdot \left(1 - \nu(2) \cdot 2 \cdot \epsilon \cdot (1-q) + O\left(\epsilon^2\right)\right).$$

This is so because when $x = 1 - \epsilon$ and $\epsilon << 1$, a new agent plays action a' with probability $1-q$ when observing action a^* once in a sample of size two (the probability of this event is given by $\nu(2)\cdot 2\cdot\epsilon$). Moreover, observing a^* once in a longer sample (or not observing a^* at all) is a rare event with a probability of $O(\epsilon^2)$. This implies that for each $q<0.5$, there is $\bar{\epsilon}>0$, such that for each $x>1-\bar{\epsilon}$: (1) $f_q^t(x)>x$ for each t, and (2) $\lim_{t\to\infty} f_q^t(x)=1$. This shows that the state 1 is locally stable.

We are left with the case in which $\nu(0)>0$ or $\nu(1)>0$. Observe that (1) $\lim_{q\longrightarrow\frac{1}{\mu_l}} f_q'(0)=1$ and (2) $f_q''(0)>0$ for each $q<\frac{1}{\mu_l}$. This implies (by a Taylor approximation around $x=0$) that there exists $(q^*,\hat{x})$ satisfying: (1) $0<\hat{x}<<1$, (2) $q^*<\frac{1}{\mu_l}$, (3) $f_{q^*}(\hat{x})=\hat{x}$, (4) $f_{q^*}(x)<x$ for each $x\in(0,\hat{x})$, and (5) $f_{q^*}'(\hat{x})>1$. This implies that $\hat{x}$ is an (unstable) steady state.

Next observe that $f_{q^*}(\hat{x})=\hat{x}$, $f_{q^*}'(\hat{x})>1$, and $f_{q^*}(1)<1$. These observations, due to the intermediate value theorem and standard arguments, imply that there exists $\hat{x}<x^*<1$, such that $f_{q^*}(x^*)=x^*$ and $f_{q^*}'(x^*)<1$. This, in turn, implies that there exists $\bar{\epsilon}>0$, such that for each $x\in(x^*-\bar{\epsilon},x^*+\bar{\epsilon})$: (1) $f_{q^*}(x)<x$ if $x<x^*$, (2) $f_{q^*}(x)>x$ if $x>x^*$, and (3) $f_{q^*}'(x)<1$. These observations imply (due to the monotonicity of f_{q^*}) that for each $x^*\neq x\in(x^*-\bar{\epsilon},x^*+\bar{\epsilon})$: (1) $f_q^t(x)$ is strictly between x and x^* for each t, and (2) $\lim_{t\to\infty} f_q^t(x)=x^*$. Hence, state x^* is locally stable.

B.3 Proof of Theorem 4 (Two Feasible Actions, $\mu(l)=0\ \forall l>2$)

Let $E=(A=\{a',a''\},\beta,\nu)$ be a sampling process satisfying $\nu(l)=0$ for each $l>2$. Let σ be an arbitrary decision rule. The fact that there are two feasible actions (i.e., $|A|=2$) implies that we can identify a population state with a number $x\in[0,1]$ representing the frequency of agents who play action a'. Let $f_\sigma(x)$ be the dynamic mapping induced by decision rule σ. The fact that the maximal length of the sample observed by new agents is two implies that $f_\sigma(x)$ is a polynomial of degree at most two. Specifically, the explicit formula for $f_\sigma(x)$ is given by:

$$
\begin{aligned}
f_\sigma(x) &= (1-\beta)\cdot x+\beta\cdot[\nu(0)\,\sigma(\emptyset)(a')+\nu(1)\cdot(x\cdot\sigma(a')(a')+(1-x)\cdot\sigma(a'')(a'))+\\
&\qquad \nu(2)\cdot\Big(x^2\cdot\sigma(a',a')(a')+(1-x)^2\cdot\sigma(a'',a'')(a')+x\cdot(1-x)\cdot(\sigma(a',a'')(a')+\sigma(a'',a')(a'))\Big)\Big]=\\
&= b\cdot x^2+c\cdot x+d,
\end{aligned}
$$

where $b,c,d\in\mathbb{R}$. Recall that x^* is a steady state iff $f(x^*)=x^*$. We conclude the proof by looking at three exhaustive cases. Cases 2 and 3 are illustrated in Figure 1 in Section 4.4.

1. Case 1: $b=0$ (recall, that b is the parameter multiplying x^2 in the formula for $f_\sigma(x)$). If $f_\sigma(x)\equiv x$ (i.e., if $f_\sigma(x)=x$ for each x), then any state is steady, but none is locally stable. Otherwise, the equation $f_\sigma(x)=x$ has at most one solution, and, hence, σ^* has at most one locally stable state.

2. Case 2: $b>0$. The equation $f_\sigma(x)=x$ has at most two solutions. Assume that it has two solutions in the interval $[0,1]$ (otherwise, it is immediate that σ admits at most one locally stable state). Simple geometric arguments (regarding the incidence points of a parabola satisfying $f_\sigma(1)\leq 1$ and the 45° line) imply that one of these solutions must be one (i.e., $f_\sigma(1)=1$), and that $f_\sigma'(1)>1$. By standard continuity arguments there exists a sufficiently small $\bar{\epsilon}>0$ such that $f_\sigma'(x)>1$ for each $x>1-\bar{\epsilon}$. This implies that for each $x>1-\bar{\epsilon}$: (1) $f_\sigma^t(x)<x$, and (2) if $\lim_{t\to\infty} f_\sigma^t(x)$ exists then it must satisfy $\lim_{t\to\infty} f_\sigma^t(x)<1-\bar{\epsilon}$. Hence, state 1 cannot be locally stable, and the decision rule σ admits at most one locally stable state.

3. Case 3: $b < 0$. Assume that the equation $f_\sigma(x) = x$ has two solutions in the interval $[0,1]$ (otherwise, it is immediate that σ admits at most one locally stable state). Simple geometric arguments (regarding the points of intersection of a parabola bounded with positive values and the 45° line) imply that one of these solutions must be zero (i.e., $f_\sigma(0) = 0$), and that $f'_\sigma(0) > 1$. By standard continuity arguments there exists a sufficiently small $\bar{\epsilon} > 0$ such that $f'_\sigma(x) > 1$ for each $x \in (0, \bar{\epsilon})$. This implies that for each $x \in (0, \bar{\epsilon})$: (1) $f^t_\sigma(x) > x$, and (2) if $\lim_{t\to\infty} f^t_\sigma(x)$ exists then it must satisfy $\lim_{t\to\infty} f^t_q(x) > \bar{\epsilon}$. This implies that state x^* cannot be locally stable, and hence decision rule σ admits at most one locally stable state.

B.4 Proof of Theorem 5 $(\nu(1) + (\nu(3) = 1)$, "Follow Majority" Rule

Let $E = (A, \beta, \nu)$ be a sampling process, such that $\nu(l) = 0$ for each $l \notin \{1,3\}$ and $\nu(1) < 1$. Let $a' \neq a'' \in A$. Let σ^* be the decision rule in which each new agent follows the frequently observed action in his sample, i.e.,

$$\sigma^*\left(a^l\right) = \begin{cases} a' & \left|\left\{i | a^l_i = a'\right\}\right| > \frac{l}{2} \\ a'' & \text{otherwise.} \end{cases}$$

Observe that new agents play only a' or a''. Further, note that the probability that a new agent plays a' depends only on the frequency with which the incumbents play action a' (and not on any other aspect of the population state). Thus, with slight abuse of notation, we identify a state $\gamma \in \Delta(A)$ with the frequency x of agents who choose action a', i.e., $x := \gamma(a')$ (as the actions played by the remaining $1-x$ of the agents do not play any role in the dynamics, and in the long run each agent plays either action a' or a''). Let $f_{\sigma^*}(x)$ be the dynamic mapping induced by decision rule σ^*. The explicit formula for $f_{\sigma^*}(x)$ is given by

$$f_{\sigma^*}(x) = (1-\beta)\cdot x + \beta \cdot \left(\nu(1)\cdot x + \nu(3)\cdot\left(x^3 + 3\cdot x^2\cdot(1-x)\right) = \nu(1)\cdot x + \nu(3)\cdot\left(3\cdot x^2 - 2\cdot x^3\right)\right),$$

and its derivative is given by

$$f'_{\sigma^*}(x) = (1-\beta) + \beta\cdot\nu(1) + \beta\cdot\nu(3)\cdot\left(6\cdot x - 6\cdot x^2\right) = (1-\beta) + \beta\cdot\nu(1) + \beta\cdot 6\cdot\nu(3)\cdot x\cdot(1-x).$$

Observe that: (1) $f_{\sigma^*}(x)$ is strictly increasing, (2) $f_{\sigma^*}(x^*) = x^*$ for three values of x: $0, 0.5, 1$, (3) $f'_{\sigma^*}(0) = f'_{\sigma^*}(1) = \nu(1) < 1$, and (4) $f'_{\sigma^*}(0.5) = \nu(1) + 1.5\cdot\nu(3) > 1$. These observations imply (by arguments analogous to those in the proof of B.3 above) that the process (E, σ^*) admits three steady states: two locally stable states, 0 and 1, and the locally unstable state 0.5.

B.5 Lemma Required for the Proof of Theorem 6 (Bound with Responsiveness)

Lemma 6. *For each sampling process E, each size $l \in \mathbb{N}$, and any two states $\gamma \neq \gamma' \in \Gamma$:*

$$\sum_{a\in A}\left|\sum_{m\in A^l}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right)\cdot\sigma_m(a)\right| \leq r_l \cdot \left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1.$$

Proof. We begin with a preliminary definition. Let $A^l_{\gamma>\gamma'} \subseteq A^l$ be the set of samples that have higher probabilities given state γ than given state γ', i.e.,

$$A^l_{\gamma>\gamma'} = \left\{m \in A^l | \psi_{l,\gamma}(m) > \psi_{l,\gamma'}(m)\right\}.$$

We now prove the lemma:

$$\sum_{a\in A}\left|\sum_{m\in A^l}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right)\cdot\sigma_m(a)\right| =$$

$$\sum_{a\in A}\left|\sum_{m\in A^l_{\gamma>\gamma'}}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right)\cdot\sigma_m(a) - \sum_{m\in A^l_{\gamma'>\gamma}}\left(\psi_{l,\gamma'}(m) - \psi_{l,\gamma}(m)\right)\cdot\sigma_m(a)\right| \leq$$

$$\sum_{a\in A}\left|\sum_{m\in A^l_{\gamma>\gamma'}}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right)\cdot\overline{\sigma}_l(a) - \sum_{m\in A^l_{\gamma'>\gamma}}\left(\psi_{l,\gamma'}(m) - \psi_{l,\gamma}(m)\right)\cdot\underline{\sigma}_l(a)\right| =$$

$$\sum_{a\in A}\left|\overline{\sigma}_l(a)\cdot\sum_{m\in A^l_{\gamma>\gamma'}}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right) - \underline{\sigma}_l(a)\cdot\sum_{m\in A^l_{\gamma'>\gamma}}\left(\psi_{l,\gamma'}(m) - \psi_{l,\gamma}(m)\right)\right| = \tag{21}$$

$$\sum_{a\in A}\left|\left(\overline{\sigma}_l(a) - \underline{\sigma}_l(a)\right)\cdot\sum_{m\in A^l_{\gamma>\gamma'}}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right)\right| =$$

$$\sum_{a\in A}\left(\overline{\sigma}_l(a) - \underline{\sigma}_l(a)\right)\cdot\sum_{m\in A^l_{\gamma>\gamma'}}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right) =$$

$$\sum_{a\in A}\left(\overline{\sigma}_l(a) - \underline{\sigma}_l(a)\right)\cdot 0.5\cdot\left(\sum_{m\in A^l}\left|\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right)\right|\right) =$$

$$0.5\cdot\sum_{a\in A}\left(\overline{\sigma}_l(a) - \underline{\sigma}_l(a)\right)\cdot\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1.$$

Equality (21) is implied by the fact that $\psi_{l,\gamma}$ and $\psi_{l,\gamma'}$ are both distributions, and the sum of the differences in the probabilities that they assign to samples of size l must be equal to zero. Thus we have shown that

$$\sum_{a\in A}\left|\sum_{m\in A^l}\left(\psi_{l,\gamma}(m) - \psi_{l,\gamma'}(m)\right)\cdot\sigma_m(a)\right| \leq 0.5\cdot\sum_{a\in A}\left(\overline{\sigma}_l(a) - \underline{\sigma}_l(a)\right)\cdot\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1, \tag{22}$$

which together with Lemma 3 implies that the LHS of (22) is weakly smaller than $r_l\cdot\left\|\psi_{l,\gamma} - \psi_{l,\gamma'}\right\|_1$. □

References

Acemoglu, Daron, Dahleh, Munther A., Lobel, Ilan, & Ozdaglar, Asuman. 2011. Bayesian learning in social networks. *The Review of Economic Studies*, **78**(4), 1201–1236.

Arthur, W. Brian. 1989. Competing technologies, increasing returns, and lock-in by historical events. *The economic journal*, **99**(394), 116–131.

Arthur, W. Brian. 1994. *Increasing Returns and Path Dependence in the Economy.* University of michigan Press.

Banerjee, Abhijit, & Fudenberg, Drew. 2004. Word-of-mouth learning. *Games and Economic Behavior*, **46**(1), 1–22.

Duffie, Darrell, & Sun, Yeneng. 2012. The exact law of large numbers for independent random matching. *Journal of Economic Theory*, **147**(3), 1105–1139.

Ellison, Glenn, & Fudenberg, Drew. 1993. Rules of thumb for social learning. *Journal of Political Economy*, **101**, 612–643.

Ellison, Glenn, & Fudenberg, Drew. 1995. Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, **110**(1), 93–125.

Foster, Dean, & Young, Peyton. 1990. Stochastic evolutionary game dynamics. *Theoretical Population Biology*, **38**(2), 219–232.

Heller, Yuval, & Mohlin, Erik. forthcoming. Observations on Cooperation. *the Review of Economic Studies.*

Kandori, Michihiro, Mailath, George J., & Rob, Rafael. 1993. Learning, mutation, and long run equilibria in games. *Econometrica*, **61**(1), 29–56.

Kaniovski, Yuri M., & Young, H. Peyton. 1995. Learning dynamics in games with stochastic perturbations. *Games and Economic Behavior*, **11**(2), 330–363.

Munkres, James R. 2000. *Topology.* 2nd edn. Prentice Hall.

Munshi, Kaivan. 2004. Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. *Journal of development Economics*, **73**(1), 185–213.

Nowak, Martin A., & Sigmund, Karl. 1998. Evolution of indirect reciprocity by image scoring. *Nature*, **393**(6685), 573–577.

Okuno-Fujiwara, Masahiro, & Postlewaite, Andrew. 1995. Social norms and random matching games. *Games and Economic Behavior*, **9**(1), 79–109.

Oyama, Daisuke, Sandholm, William H., & Tercieux, Olivier. 2015. Sampling best response dynamics and deterministic equilibrium selection. *Theoretical Economics*, **10**(1), 243–281.

Oyarzun, Carlos, & Ruf, Johannes. 2014. Convergence in models with bounded expected relative hazard rates. *Journal of Economic Theory*, **154**, 229–244.

Pata, Vittorino. 2014. *Fixed Point Theorems and Applications.* Mimeo.

Phelan, Christopher, & Skrzypacz, Andrzej. 2006. *Private monitoring with infinite histories.* Tech. rept. Federal Reserve Bank of Minneapolis.

Rosenthal, Robert W. 1979. Sequences of games with varying opponents. *Econometrica*, **47**(6), 1353–1366.

Sandholm, William H. 2001. Almost global convergence to p-dominant equilibrium. *International Journal of Game Theory*, **30**(1), 107–116.

Sandholm, William H. 2011. *Population Games and Evolutionary Dynamics*. Cambridge, MA: MIT Press.

Schlag, Karl H. 1998. Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, **78**(1), 130–156.

Smith, Lones, & Sorensen, Peter Norman. 2014. *Rational social learning by random sampling*. mimeo.

Takahashi, Satoru. 2010. Community enforcement when players observe partners' past play. *Journal of Economic Theory*, **145**(1), 42–62.

Weibull, Jörgen W. 1995. *Evolutionary game theory*. The MIT press.

Young, H. Peyton. 1993a. The evolution of conventions. *Econometrica*, **61**(1), 57–84.

Young, H. Peyton. 1993b. An evolutionary model of bargaining. *Journal of Economic Theory*, **59**(1), 145–168.

Young, H. Peyton. 2015. The evolution of social norms. *Annual Reviews of Economics*, **7**(1), 359–387.