ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT
Games and Economic Behavior 46 (2004) 1–22
GAMES and Economic Behavior
www.elsevier.com/locate/geb

Word-of-mouth learning

Abhijit Banerjee [a] and Drew Fudenberg [b,*]

[a] *Department of Economics, MIT, Cambridge, MA 02142-1347, USA*
[b] *Department of Economics, Harvard University, Cambridge, MA 02138, USA*

Received 17 December 2001

Abstract

This paper analyzes a model of rational word-of-mouth learning, in which successive generations of agents make once-and-for-all choices between two alternatives. Before making a decision, each new agent samples N old ones and asks them which choice they used and how satisfied they were with it. If (a) the sampling rule is "unbiased" in the sense that the samples are representative of the overall population, (b) each player samples two or more others, and (c) there is any information at all in the payoff observations, then in the long run every agent will choose the same thing. If in addition the payoff observation is sufficiently informative, the long-run outcome is efficient. We also investigate a range of biased sampling rules, such as those that over-represent popular or successful choices, and determine which ones favor global convergence towards efficiency.

JEL classification: C72; D83

Keywords: Social learning; Word-of-mouth; Herding; Imitation

1. Introduction

This paper introduces and analyzes a simple model of rational word-of-mouth learning, in which agents use information about the experiences of other agents to guide their own decisions. Such communication has long been known to be an important component of brand choice by consumers; it also seems to be relevant for the adoption of agricultural technologies and other production processes, and more generally to the spreading of fads, fashions, and ideas within society.

* Corresponding author.
E-mail address: dfudenberg@harvard.edu (D. Fudenberg).

0899-8256/$ – see front matter
doi:10.1016/S0899-8256(03)00048-4

Consider an example from everyday life. Someone in your department tells you that one of your colleagues is about to move to another university. She also tells you that she is sufficiently worried about it that she has started to make contingency plans. You decide that the rumor is probably not true and that contingency plans are unnecessary, but when you run into another colleague in the corridor, you pass on the story, adding that you do not think its worth worrying about. Because you are rushing off to a meeting, you do not actually tell him how you came about the story. And so the story spreads. . . .

The underlying model here is that there is piece of information that everyone would benefit from knowing, but the only source of information is word of mouth. No hard evidence is provided, and while people tell you their opinion, they do not give you all of their reasons, and you do not observe the entire process by which the story came to you. Moreover, you know that you must not believe everything everyone tells you, both because you know that other people also do not necessarily have hard evidence, and also because you know that people overlay their personal hopes and fears on what they report.

What is long-run outcome of such a process of information transmission? Does everyone learn the truth? Does everyone come around to the same view, be it right or wrong? Or does the diversity of views persist even in the long run?

This is a question that has been asked by others: Indeed, this is the subject of the entire literature on herd behavior/informational cascades.[1] The main difference between our model and the type of model studied in this literature comes from the fact that in our model people only learn by "word-of-mouth" communication with a few other agents, instead of observing the entire history leading up to them.[2] It is clear that in many real world situations, people do not get to find out what the whole world is doing: Often it is simply too costly to gather the information or the information is something that is naturally private.[3] And even when there is public information, for example about the popularity of different cars, this information tends to be about aggregate popularity of different choices, while agents want to know the popularities among those with similar preferences.

There is also an a priori reason to study the impact of introducing the word-of-mouth assumption into the models of social learning. As explained in Section 5, our intuition is that word-of-mouth learning makes herding less likely, as it reduces the correlation

1 See, for example, Banerjee (1992), Bhikchandani et al. (1992), Lee (1993), Smith and Sorenson (2000), and Vives (1997). Chamley and Gale (1994) and Caplin and Leahy (1994) apply related ideas to study the stability of the macroeconomy.

2 Ellison and Fudenberg (1993, 1995) study models of boundedly-rational word-of-mouth learning. Banerjee (1993) studies rational word-of-mouth learning in a setting that is not directly comparable to either this paper or the herding models. Bjonerstedt and Weibull (1995) and Schlag (1998) discuss how word-of-mouth processes of strategy revision in games can generate the "replicator dynamic" of evolutionary biology.

3 Udry and Conley (2001), for example, in their study of pineapple farmers in Ghana, find that most farmers only know about what a handful of other farmers are doing, but this group has a strong influence on their decisions. Duflo and Saez (2000), in their study of the decision to join a Tax Deferred Annuity plan (TDA) among employees of a large US university, also find that each person's choice is influenced by a small group of others. Finally, while Munshi and Myaux's (2000) study of contraception in Bangladesh has little explicit information about the size of the group that the women talk to, their presumption is clearly that in rural Bangladesh it is unreasonable to expect women to know every other villager's contraception practices.

between the observations of different agents. It is therefore interesting to see whether the herding results survive this change in assumptions.

In our model, at each point of time a new generation of agents has to make a once-and-for all choice between two alternatives, a and b.[4] The word-of-mouth element comes in because agents consult a sample of N others, and those consulted report what they themselves have chosen. The relative popularity of the choices in the agent's sample is a signal of their relative popularity in the whole of the "relevant" population. In addition to this information about relative popularity, agents may also receive signals that are correlated with the payoffs from the choices. This may arise, for example, if some of the sampled agents report not only their choice but also an indication of how satisfied they are with it. In the informal story above, the actions are "make plans" and "do not make plans," and the "relevant population" is people in your own department. The people who you ran into, including both those who told you the story and those who did not, are your sample. And what they told you about the strength of their feelings, were the signals.

In addition to the framing assumptions described above, we make a number of other modeling choices. First, we assume that all agents are ex-ante identical, so that one choice or the other is best for all of them, but there are unobserved idiosyncratic shocks to each agent's realized payoff, so that one person's report of a high payoff does not guarantee that his choice was the optimal one. Second, we assume that current decision makers do not observe the information that past decision makers used in making their decisions. Third, we assume that agents have a common prior on the mean difference in payoffs. Finally, we need to make an assumption about what is known about the initial conditions.

In general, we suppose that agents know the distributions, conditional on which choice is better, over the share of the population who chose a when the process started. A special case of this is *payoff-determined initial popularity*, meaning that the fractions of those who would initially choose a, both when a is the right choice and when b is the right choice, are common knowledge. As we will see, this case gives us the strongest results.

To interpret this assumption, we return to our initial example. When your colleague told you the bad news she might also have added that the rumor came from department Y. If you know the people in department Y very well, you may have a good idea of how the number of people there who would take the rumor seriously varies with whether the rumor is actually true; this corresponds to payoff-determined initial popularity. If you are less familiar with department Y, you may instead be uncertain just how many people there would start the rumor, so that your beliefs about this correspond to a nondegenerate probability distribution.

We focus on our model's long-run predictions. Our main result is a set of sufficient conditions for the long-run outcome to be "homogeneous" in the sense that all agents choose the same action (though this action may or may not be the correct one); in other words there is herding or an informational cascade. This conclusion follows if the samples are representative draws from the prevailing distribution in society (which we call "unbiased and proportional" sampling) and everyone samples at least two others.

[4] While the decisions in the examples we mention are not completely irreversible, the cost of making the wrong choice is either very substantial (in the case of contraception or switching to a new data base) or only knowable after a long delay (in the case of TDA), which make them more or less like a once-in-a-lifetime choice.

Our second main result provides a sufficient condition for the only stationary outcome to be the one where all agents make the correct choice, so that the system converges to efficiency. The result requires payoff-determined initial popularities, and also that the distribution of signals assigns positive probability to signals that are informative enough to outweigh the prior.

While the proportional sampling rule used in proving these results is a natural first choice for analysis, it seems that people can and do use other sampling rules. In Section 4 we show that the long-run outcome can be heterogeneous if the sampling rules are not proportional. Section 4.1 considers the effects of "perception biases" in the sampling process, by which we mean that the probability of sampling the user of a rare choice may be either greater or less than the population fraction currently using it. Section 4.2 investigates the effects of reporting biases: If samples are constructed from reports by others, and those with more extreme payoffs are more likely to report, the samples will be biased.[5] Under several of these alternative sampling rules, the first step in our convergence argument fails: It is no longer true that adopting the action of the first person contacted yields an expected payoff equal to the average payoff in the current population. More strongly, under some of these alternative sampling rules the efficient outcome is not even locally stable.

2. The model

Throughout the paper, we suppose there are two alternative choices, a and b, which we think of as representing brands or technologies. At every point in time, there is a continuum of agents of mass 1; a proportion x of these agents use choice a and all of the others use choice b. Each period, a representative fraction γ of consumers leaves the population and is replaced by newcomers, so for example mass γx of agents using choice a are replaced. These new agents have to make a once-and-for all choice of either adopting a or adopting b. If an agent chooses a, her payoff is the sum of a term u^a that is common to all agents and an individual specific noise term with zero mean. Similarly, if an agent chooses b, her payoff is the sum of a term u^b that is common to all agents and an individual-specific, mean-zero noise term.[6] We suppose that the noise terms are i.i.d. over time and across agents, so that the common terms u^a, u^b correspond to the "quality" of the two choices. Denote the difference in quality levels by $\Delta = u^a - u^b$. We suppose that agents do not know the value of Δ; for simplicity, we assume further that Δ has only 2 possible values, $\overline{\Delta} > 0 > \underline{\Delta}$. All agents assign common prior probability $q \geqslant 1/2$ to the event that $\Delta = \overline{\Delta}$. We suppose that $q\overline{\Delta} + (1-q)\underline{\Delta} > 0$, so that *ex-ante* a is better than b. As a normalization, we further specify that the "quality" of the inferior good is equal to 0, so that either $u^a = \overline{\Delta}$ and $u^b = 0$ or $u^a = 0$ and $u^b = -\underline{\Delta} > 0$.

[5] McKenna (1991) says that "A customer who has a good experience with a product will tell three other people. A customer who has a bad experience will tell ten other people."

[6] The individual noise terms do not enter explicitly in our analysis; they are needed to interpret the signals s introduced below as reports from past agents about their realized payoffs. Here and subsequently, we will speak loosely about a continuum of independent random variables. Since our analysis will only concern population aggregates, this looseness will not be important.

In each period, the change in the population fractions using each choice is determined by the distribution of responses in the new agents' samples and the decision rule of the new agents; we explain below how this is computed. To complete the description of the dynamical system we need to specify initial conditions. To this end, we suppose that there are $j+k$ states of the world denoted by $i \in I$. The fraction of the population that uses choice a at date 0 in state i is denoted $x_i(0)$; $\mathbf{x}(0)$ denotes the vector whose components are the $x_i(0)$. In states $i = 1, \ldots, j$, the quality difference is $\Delta = \overline{\Delta}$; we denote this event by $\overline{\theta}$. Similarly, the states $i = j+1, \ldots, j+k$ correspond to the event $\Delta = \underline{\Delta}$, which we denote $\underline{\theta}$. Agents have common prior distribution p over the states of the world, with p_i denoting the probability of state i. Thus the prior probability of the event $\overline{\theta}$ is $q = p(\overline{\theta}) = \sum_{i \in \overline{\theta}} q_i$.

The simplest version of our model has only two states, one for each value of Δ. We call this the case of *payoff-determined initial popularity*, as in this case $\overline{\theta}$ and $\underline{\theta}$ are singletons, and there is a deterministic map from payoffs to initial conditions. This case is of interest because it produces the sharpest results, and also because it or similar conditions are frequently used in this literature, but we will be at least as interested in the general case, with many states in the events $\overline{\theta}$ and $\underline{\theta}$, as this allows the relationship between payoff differences and initial popularity to be stochastic. One explanation for this aggregate uncertainty is that the initial condition reflects the choices of a group of "early adopters" whose preferences are uncertain even at the aggregate level; for example, the fraction of early adopters with a taste for novelty might be unknown.[7] It is important to note that we treat the distribution relating initial conditions and the payoffs of the choices as exogenous. Our choice to view this distribution as separate is consistent with Moore (1991), who argues that the very first adopters of a new technology do so for reasons that are very different from those that matter for most other adopters.

Turning to the mechanics of information gathering, the paper allows for different specifications of the rule by which players draw their samples. To accommodate this, for a fixed sample size N let Z denote the set of all pairs (α, β) with α and β both nonnegative; and $\alpha + \beta = N$; $\zeta \in Z$ is then a sample of N players, of whom α use a and β use b. Then $\mu(\zeta \mid x)$ is the probability of drawing sample ζ when fraction x of the population uses choice a. (Note that this probability depends on the state of the world only through the state's influence on the proportion of players using each choice.) Moreover, in the spirit of the law of large numbers, we will specify that the fraction of new agents who draw sample ζ exactly equals $\mu(\zeta \mid x)$.

In addition to observing the actions chosen, i.e., the sample ζ, players also receive a signal, denoted s, that may be correlated with the realized value of Δ. At this point we allow for the possibility that s is independent of Δ, in which case it is of no use to the agents, but we tend to think that the agents will typically have some sources of information

[7] To be more specific, suppose that early adopters receive private signals of the relative attractiveness of the alternatives according to known probability distributions, and choose the alternative that gives them the highest expected utility. The initial fractions using a and b will not be determined simply by the objective payoff difference, but will also depend on the unknown fraction of early adopters with a taste for novelty. Another possible story for multiple initial conditions is that an unknown fraction of the initial decision makers chose what they did because they were offered a large "introductory pricing" discount.

beyond the popularity itself, such as the reported satisfaction levels of the people they contact. On the other hand, we suppose that this information is not perfect, that is, that the signals s do not perfectly reveal the state of the world. We will assume that conditional on the state of the world, and receiving sample ζ the signal received by each agent is drawn independently from the same distribution. For convenience, we further suppose that this distribution is atomless, with density $f(s \mid \theta, \zeta)$; this is not important for our results but helps simplify a few details.

Definition. The signal *conveys some information* if for all samples $\zeta \in Z$ there is a positive probability of signals s such that $f(s \mid \bar{\theta}, \zeta) \neq f(s \mid \underline{\theta}, \zeta)$.

Most of our results suppose that this condition is satisfied. (Note that if the signals are simply the realized payoffs of the choices in the agent's sample, the condition is satisfied whenever at least one sampled agent reports his or her payoff in addition to their choice.) We use the standard "large numbers" convention that in each state θ the fraction of agents with samples ζ who see signal s exactly equals $f(s \mid \theta, \zeta)$. Finally we will assume that the structure of this process, including the rules that generate samples and signals, is common knowledge.

We assume that observed play corresponds to a pure-strategy Bayesian equilibrium of the game. (Since each player only moves once, and players are unconcerned about the actions of those who move either subsequently or simultaneously, it is easy to check that an equilibrium exists, as the equilibrium can be constructed by "rolling forward" from the initial period.) As a result we can assume that in an equilibrium, all the players know the functions that specify the fraction using choice a at the beginning of period t in each state of the world. We will represent these functions by a vector $\mathbf{x}(t) = (x_1(t), \ldots, x_I(t))$, where $x_i(t)$ is the fraction of a-users in state i.

Since the agent knows the fraction of a-users in all states of the world, the share of a-users in a player's sample is indirect evidence about the state of the world, and hence about which choice is better. Of course, the interpretation of this evidence depends on the way that the actual population fraction x influences the composition of the samples, and also on the correlation between the fraction x and the state θ.[8]

Once a player receives sample ζ he updates his prior beliefs as follows. First, the relative popularity of the two choices in the sample itself conveys information, so that the odds ratio after seeing the relative popularities in the sample is

$$\frac{p(\bar{\theta} \mid \zeta)}{p(\underline{\theta} \mid \zeta)} = \lambda_\zeta(\mathbf{x}). \tag{2.1}$$

Second, players also take account of the information conveyed by signal s. Combining this with the "interim" odds ratio in Eq. (2.1) yields the posterior odds ratio

$$\frac{p(\bar{\theta} \mid \zeta, s)}{p(\underline{\theta} \mid \zeta, s)} = \left(\frac{f(s \mid \bar{\theta}, \zeta)}{f(s \mid \underline{\theta}, \zeta)}\right)\lambda_\zeta(x). \tag{2.2}$$

[8] If, at time t, no two of the $x_t(t)$ are equal, then observing the actual share x reveals the state and so reveals the optimal choice. But, as we pointed out in the introduction, agents cannot directly observe the aggregate popularities.

Players will choose a when $p(\bar\theta \mid \zeta, s)\overline{\Delta} + p(\underline{\theta} \mid \zeta, s)\underline{\Delta} > 0$. This is equivalent to the posterior odds ratio in (2.2), $p(\bar\theta \mid \zeta, s)/p(\underline{\theta} \mid \zeta, s)$, being strictly greater than $-\underline{\Delta}/\overline{\Delta}$; we will denote this critical value by C. Players will choose b when the posterior odds ratio is strictly less than C; when the odds ratio is exactly C, players are indifferent. Note that our assumption that a is optimal under the prior beliefs implies that the prior odds ratio $q/(1-q)$ exceeds C.

Many of our results make use of the following assumption.

Definition. The system satisfies the *minimal informativeness condition* if in state $\underline{\theta}$, for all samples $\zeta \in Z$ there is positive probability of observations s such that

$$\left(\frac{f(s \mid \bar\theta, \zeta)}{f(s \mid \underline{\theta}, \zeta)}\right)\frac{q}{1-q} < C.$$

The assumption that the realized distribution of signals exactly equals the theoretical distribution that generates it allows us to compute, for each sample ζ, and each event θ, the fraction of the players that, after receiving the sample, strictly prefer a: This is the probability, under events $\bar\theta$ and $\underline{\theta}$, respectively, that the realization of s is such that the odds ratio in (2.2) exceeds C.[9] In a similar fashion, we can compute the fraction of agents that strictly prefer b. Our results will not depend on the way that agents choose when indifferent; by making an arbitrary selection here we arrive at the total fractions $\overline{\Phi}_\zeta(\mathbf{x}(t-1))$ and $\underline{\Phi}_\zeta(\mathbf{x}(t-1))$ of those who observe ζ that adopt a at date t in states $\bar\theta$ and $\underline{\theta}$, respectively.[10]

Thus the fraction of the population currently using a evolves according to

$$\begin{aligned}
&x_i(t) = (1-\gamma)x_i(t-1) + \gamma\left[\sum_\zeta \mu\big(\zeta \mid x_i(t-1)\big)\overline{\Phi}_\zeta\big(\mathbf{x}(t-1)\big)\right], \quad i = 1, \ldots, j,\\
&x_i(t) = (1-\gamma)x_i(t-1) + \gamma\left[\sum_\zeta \mu\big(\zeta \mid x_i(t-1)\big)\underline{\Phi}_\zeta\big(\mathbf{x}(t-1)\big)\right],\\
&\quad i = j+1, \ldots, j+k.
\end{aligned} \tag{2.3}$$

The remainder of the paper is devoted to analyzing the behavior of this deterministic dynamical system, and how it depends on the nature of the word-of-mouth process through the induced form of the functions Φ. Note that this is a function on the $(j+k)$-dimensional state vector $\mathbf{x}$, and that the "corners" of this state space (the points where every component of $\mathbf{x}$ is either 0 or 1) are the "herding points" where, in each state of the world, every agent is using the same action. The efficient outcome is the herding point where all agents use a in the states 1 throughly j (that is in $\bar\theta$) and all agents use b in states $j+1$ through $j+k$.

[9] These fractions depend on θ because it influences the distribution of realized payoffs in each sample, but since they depend only on the distribution of payoffs, they are the same for all states in a given event θ.

[10] In principle, the choice that agents make when indifferent could depend on calendar time as well as their sample; this possibility makes no difference to the results so we suppress it to lighten notation.

3. Proportional sampling

This section specializes the rule by which agents gather their information to "proportional" or "unbiased" sampling, which means that each agent sampled is an independent draw from the probability distribution $(x, 1-x)$ over a-users and b-users. Thus the distribution of samples is binomial, and the fraction of players who get a sample of m a-users and $N-m$ b-users is equal to $\binom{N}{m}x^m(1-x)^{N-m}$. This is the most commonly specified sampling rule, and perhaps the most natural; Section 4 discusses some plausible alternatives.

3.1. Some preliminaries

Definition. The population's average payoff is

$$\left(\sum_{i\in\overline{\theta}} p_i x_i(t)\right)\overline{\Delta} - \left(\sum_{i\in\underline{\theta}} p_i\bigl(1-x_i(t)\bigr)\right)\underline{\Delta} \equiv \overline{U}\bigl(\mathbf{x}(t)\bigr).$$

This is the expected payoff of a randomly drawn member of the population, where the expectation is taken with respect to the prior distribution.[11]

Lemma 1. (a) *For any sampling rule, $\overline{U}(\mathbf{x}(t))$ is strictly increasing over time whenever it is less than $q\overline{\Delta}$.*

(b) *With proportional sampling $\overline{U}(\mathbf{x}(t))$ is nondecreasing over time.*

(c) *With proportional sampling, $\overline{U}(\mathbf{x}(t+1)) = \overline{U}(\mathbf{x}(t))$ if and only if no decision rule can improve on the rule "copy the action of the first person in the sample."*

Proof. For (a), note that one feasible decision rule is to ignore the observations entirely and use a, which is the better choice with the prior beliefs; this yields payoff $q\overline{\Delta}$ so each generation of new players must get at least this much in expectation. For (b), note that under proportional sampling, the feasible decision rule "adopt the same action as that used by the first person in the sample" has expected payoff $\overline{U}(\mathbf{x}(t))$. Thus whatever strategy the new adopters use at date t must yield at least this high a payoff. For part (c) observe that if any strategy yields a higher payoff, then $\overline{U}(\mathbf{x}(t))$ must be increasing, while if no strategy yields a higher payoff, then the new adopters at date t must obtain exactly $\overline{U}(\mathbf{x}(t))$ in expected payoff. □

Remark. Parts (a) and (b) of this lemma apply either directly or with small modifications to several related models. For example, the lemma still holds if new agents "inherit" a choice of actions from their "parents," and must pay a cost if they switch to the other choice: If agents do not switch, their expected payoff is $\overline{U}(\mathbf{x}(t))$; if they do choose to switch, they must do at least as well. Also, in the "herd behavior" models where one agent moves in each period, and agents observe all previous choices, doing what the last person

[11] Recall that $-\underline{\Delta} > 0$ is the payoff to b in states where b is better.

did guarantees that you are as likely to make the right choice as he was. Therefore the equivalent of $\overline{U}(\mathbf{x}(t))$ cannot decrease.

3.2. The basic convergence result

The next two results give alternative sufficient conditions for the system to converge to "herding points" where the fraction using a is either 1 or 0 in each state of the world.

Theorem 1. *Assume that sampling is proportional, that $N > 1$, and that signals convey some information. Then*

(a) *if $\mathbf{x}(t)$ is such that at least one of its components, $x_i(t)$, is neither 0 nor 1, then $\overline{U}(\mathbf{x}(t+1))$ must be strictly greater than $\overline{U}(\mathbf{x}(t))$.*

Consequently,

(b) *the only stationary points of the system are those with "herding," in the sense that all agents are choosing the same action, and*
(c) *the system must converge to such a stationary point.*

Proof. (a) To show that $\overline{U}$ is strictly increasing over time everywhere but the specified "corner" points, we note first that in any state i such that x_i is neither 0 nor 1, every sample of size N has positive probability. We use this fact to show that the agents who choose at date t have a decision rule that yields strictly greater payoff than $\overline{U}(\mathbf{x}(t))$.

If there is no such rule, then part (c) of Lemma 1 implies that agents can do no better than to use the action of the first agent in their sample. Since each payoff realization in a sample is drawn from the same distribution, the order of the draw in each sample contains no information.[12] Therefore if agents are willing to choose a when a is the choice of the *first* person in their sample, they would also be willing to choose a when a is the choice of any other person in their sample, so a must be an optimal choice for all samples with at least one a and all possible payoff realizations. Likewise, b must be an optimal choice in all samples with at least one b. But then agents must be indifferent between a and b for all samples with at least one a and one b, and so for all such samples ζ there must be probability 1 that

$$\left(\frac{f(s \mid \overline{\theta}, \zeta)}{f(s \mid \underline{\theta}, \zeta)}\right)\lambda_\zeta(x) = C.$$

In other words, the likelihood ratio in such samples is the same for all values of s. This contradicts the assumption that signals convey some information.

This proves (a); (b) is an immediate corollary. To prove (c), note that because the right-hand side of Eq. (2.3) is continuous, the increase in $\overline{U}$ is bounded away from 0 in any region bounded away from the herding points. Since $\overline{U}$ is bounded above by the full-information

[12] That is, the draws are exchangeable.

payoff $q\overline{\Delta} - (1-q)\underline{\Delta}$, the state can only remain outside of any given neighborhood of the herding points for a finite number of periods. □

The next result shows that if $N > 2$, $\overline{U}(\mathbf{x}(t))$ is strictly increasing except perhaps at ‘corners,’ even when the signals s convey no information. We provide this result to clarify the structure of the system, and not because we believe that the case is of great independent interest.

Theorem 1′. *The conclusions of Theorem* 1 *hold if sampling is proportional and* $N > 2$, *even if signals convey no information.*

Proof. See Appendix A. □

3.3. Efficiency

The next step is to try to sharpen this conclusion: Given that the agents will all end up using the same action, will it be the action that is optimal under full information? Or more precisely, what are the probabilities of the efficient and inefficient outcomes? This will depend in part on the informativeness of the signals s.

Let r denote the maximum informativeness of the signals, which we assume (for notational convenience) is the same in samples of all a's and samples of all b's; we also maintain the symmetry assumption that $1/r$ is the common minimum of these two expressions. That is,

$$\max_{s,\zeta}\bigl(f(s \mid \overline{\theta}, \zeta)/f(s \mid \underline{\theta}, \zeta)\bigr) = r \quad \text{and} \quad \min_{s,\zeta}\bigl(f(s \mid \overline{\theta}, \zeta)/f(s \mid \underline{\theta}, \zeta)\bigr) = 1/r,$$

we allow for now that $r = \infty$.

A corner $\mathbf{x}$ is a stationary point if agents receiving a sample of all a's will ignore their payoff signal and choose a while agents observing all b's will choose b; this is the case if and only if

$$\frac{\sum_{i \in A^*(\mathbf{x}) \cap \overline{\theta}} p_i}{\sum_{i \in A^*(\mathbf{x}) \cap \underline{\theta}} p_i} \geqslant rC \quad \text{and} \quad \frac{\sum_{i \in B^*(\mathbf{x}) \cap \overline{\theta}} p_i}{\sum_{i \in B^*(\mathbf{x}) \cap \underline{\theta}} p_i} \leqslant \frac{C}{r} \tag{3.1}$$

where $A^*(x)$ and $B^*(x)$ are the sets of coordinates of the point $\mathbf{x}$ for which $x_i = 1$ and $x_i = 0$, respectively.[13] This gives us:

Theorem 2. *A corner* $\mathbf{x}$ *is a stationary point if and only if condition* (3.1) *holds.*

Note that this implies that the efficient point is a stationary point since at the efficient point $A(\mathbf{x}) \cap \underline{\theta}$ and $B(\mathbf{x}) \cap \overline{\theta}$ are both empty and therefore (3.1) holds. For future reference, note also that when payoffs determine the initial popularities, both $\overline{\theta}$ and $\underline{\theta}$ are singletons, so that under the minimum informativeness condition (see above), condition (3.1) can only

[13] When the denominator in either of the quotients in (3.1) is 0, set its value to be infinity. Note that these conditions are vacuous when $r = \infty$.

be satisfied at the efficient point. However, when at least one of $\overline{\theta}$ or $\underline{\theta}$ is not a singleton, for any fixed upper bound r on payoff informativeness, there exist prior beliefs that satisfy the minimum informativeness condition and that allow inefficient stationary points.[14] On the other hand, for fixed prior beliefs, there are no inefficient stationary points if the maximum informativeness r exceeds some finite lower bound.

Theorem 2 tells us only about the existence of inefficient stationary points, and not the degree of inefficiency, which (since all these stationary points are corners) is determined by the probability that all agents are making the wrong choice. An easy calculation shows that this probability is bounded above by $C/(C+r)$, so that it shrinks to 0 as the informativeness r grows. This bound holds for any prior beliefs p, no matter how extreme.

While inefficient stationary points exist, they are never stable, since the system must move away from the inefficient steady state if it starts at a nearby point which is more efficient. To see this, fix an inefficient steady state $\mathbf{x}^*$, and consider the hyperplane defined by the equation

$$\overline{\Delta}\sum_{i\in\overline{\theta}} p_i x_i - \underline{\Delta}\sum_{i\in\underline{\theta}} p_i(1-x_i) = \overline{U}(\mathbf{x}^*).$$

This hyperplane passes through $\mathbf{x}^*$ and divides the set of feasible points into two. Now consider any point near $\mathbf{x}^*$ on the side of this hyperplane which contains the efficient point. At this point the population's average payoff is higher than that at $\mathbf{x}^*$. Moreover, if the point is near enough to $\mathbf{x}^*$ it cannot be a corner, so by Theorem 1 a path starting at this point must move in the direction of increased average payoff, and this must move away from $\mathbf{x}^*$.

A related argument establishes that the efficient point is stable. To see this, let the population's average payoff at the *best inefficient steady state* be $\overline{U}^{**}$, and consider the hyperplane defined by

$$\overline{\Delta}\sum_{i\in\overline{\theta}} p_i x_i - \underline{\Delta}\sum_{i\in\underline{\theta}} p_i(1-x_i) = \overline{U}^{**}.$$

Any trajectory starting between this hyperplane and the efficient point must converge to the efficient point, since the population's average payoff is nondecreasing over time and (by construction) there no other stationary points on that side of the hyperplane.

Theorem 3. *Every inefficient stationary point is neither stable nor unstable. The efficient steady state is stable.*

Finally, recall that if there is payoff-determined initial popularity (i.e., only two states of the world) the minimal informativeness condition implies that (3.1) is only satisfied at the efficient point.

[14] The inefficient stationary points resemble the herding in Banerjee (1992) and Bhikchandani et al. (1992), with the difference that here the mistaken "herd" does not arise from the early movers having received misleading observations, but rather from uncertainty about the initial position of the system.

Lemma 2. *With only two states, and under any sampling rule, neither $(0, 1)$ nor $(0, 0)$ is a stationary point of system* (2.3)*, and $(1, 1)$ is not a stationary point if the minimal informativeness condition is satisfied. Moreover, with two states and proportional sampling the movement from each of these points results in an increase in $\overline{U}$.*

Theorem 4. *Under proportional sampling, if $N > 1$, the minimal informativeness condition is satisfied, signals convey some information, and there is payoff-determined initial popularity, then from any initial position the system converges to the efficient point $(1, 0)$. Moreover, the average efficiency $\overline{U}$ is strictly increasing along every trajectory.*

Proof. Lemmas 1 and 2 and Theorem 1 show that $\overline{U}$ is a strict Lyapunov function for the system. Since the system evolution equation is continuous, and the system variable lies in a compact space, the conclusion follows. □

The keys to this proof are that

(a) there is a simple strategy that new decision makers can use that yields as much as the average in the current population, and
(b) by using more information, the agent can obtain more than this average.

For expositional clarity, the version of the theorem stated is a bit weaker than necessary: Theorem 4 would be true under weaker conditions in the case where $N > 2$.[15] Also note that the convergence result obtains if the quality of the signals varies across the population, so long as each agent knows the rule that generates his observations.

3.4. The case $N = 1$

To conclude this section we analyze the case $N = 1$ in some detail, both to show why the hypothesis $N > 1$ is needed in Theorem 1, and to illustrate the workings of the model. For simplicity, we stick with the case of only two states, and we suppose that the payoff signal is the realized payoff of the person they contact. We further suppose that the payoffs to a and b in states $\overline{\theta}$ and $\underline{\theta}$ respectively are distributed according to the densities $\overline{f}_a$, $\overline{f}_b$, $\underline{f}_a$, and $\underline{f}_b$, with all the densities having support on the same interval.

Then the posterior odds ratio after sampling an a-user (so that $\zeta = a$) whose payoff is s is:

$$\frac{p(\overline{\theta} \mid s, a)}{p(\underline{\theta} \mid s, a)} = \frac{q x_1 \overline{f}_a(s)}{(1-q) x_2 \underline{f}_a(s)} = \lambda_a \frac{\overline{f}_a(s)}{\underline{f}_a(s)} \equiv \lambda_a \rho_a(s)$$

where λ is the interim odds ratio defined in (2.1) and $\rho_a(s)$ is the ratio of the likelihoods of the signal s in the two states. Likewise, the posterior odds ratio following a sample of b is $\lambda_b \rho_b(s)$, where $\rho_b(s) = \overline{f}_b(s)/\underline{f}_b(s)$. We assume that this latter ratio is bounded above

[15] It would suffice that signals in the sample $\zeta = (N, 0)$ satisfy minimal informativeness, even if the signals in other samples conveyed no information. This is allowed by our model but does not seem plausible.

and below, so that a single payoff signal will not change the decision of players who are otherwise very sure that they know the right choice.

With this sort of boundedly informative payoff information, all those who sample b will choose b if $(q(1-x_1)/((1-q)(1-x_2)))\rho_b(s) < C$ for all s, or

$$x_2 < 1 - \frac{rq(1-x_1)}{C(1-q)}. \tag{3.2}$$

Similarly, the condition that those who see an a *choose* a for all realizations of the payoff signal is

$$x_2 < \frac{qx_1}{rC(1-q)}. \tag{3.3}$$

Now that we have determined the agents' decision rules, we can plug them in to the equation of motion of the system to characterize the phase plane. Note first a key property of the $N = 1$ case with proportional sampling: if all agents adopt the action they see used, then the share of new agents who choose a exactly equals the current share using a, and the system is at a stationary point.

This is exactly what happens in the region satisfying both (3.2) and (3.3). In this region, anyone who sees an a chooses an a and likewise for b. Hence, no new information is incorporated into the x_i's, and every point in this region is a stationary point of the system. In particular, the system need not converge to a corner, and Theorem 1 fails, although these inefficient stationary points are consistent with Lemma 1.

Moreover, the dynamics of the system are easily characterized. When $r < q/(C(1-q))$, the herding point $(1, 1)$ satisfies (3.3), because the most favorable signal for choice b is not strong enough to overturn the prior belief that a is better. As the information bound r increases past this level, the boundary of the region where all those who see a choose a moves below the diagonal. This is the case in Fig. 1.

In the region where (3.3) is satisfied, but (3.2) is not, all those who see an a choose a, and some of those who see b choose a as well. Consequently

$$x_1(t+1) - x_i(t) \propto \bigl(x_i(t) + \vartheta_i\bigl(\mathbf{x}(t)\bigr)\bigr) - x_i(t)$$

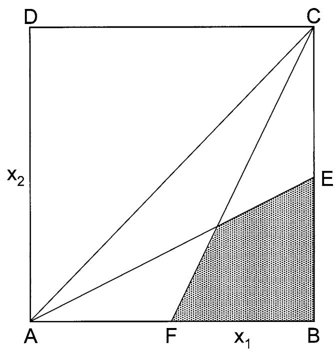

Fig. 1. $N = 1$. Below AE everyone who sees an A chooses A. To the right of CF everyone who sees a B chooses B. The shaded area BEF is the set of steady states.

for some functions $\vartheta_i(\mathbf{x}(t)) > 0$, and so the share of a is increasing in both states. Similarly, in the region where (3.2) is satisfied but (3.3) is not, the share of a is decreasing in both states.

In the region where neither (3.2) nor (3.3) are satisfied, it is strictly optimal for some agents who observe a to choose b, and for some who observe b to choose a. Therefore the population does strictly better than it would do if everyone copied the choice they observed. It follows from Lemma 1 that $\overline{U}(\mathbf{x}(t))$ must strictly increase in this region. Therefore the only stationary points are in the region satisfying (3.2) and (3.3), and the system must converge to this region.

Theorem 5. *With $N = 1$ and proportional sampling, the system always converges to some point in the closure of the region satisfying* (3.2) *and* (3.3)*. Moreover, once the system enters this region it stops, so that the* only *initial condition from which the system asymptotically converges to the efficient point is the efficient point itself.*

Proof. The fact that the system stops when it enters this region is an immediate consequence of the definitions of (3.1) and (3.2); Lemma 1 shows that the system has no steady states outside of this region. Finally, since payoff outside of the region is strictly increasing, there cannot be any cycles. Global convergence follows from the fact that in any region that is bounded away from the stationary points, $\overline{U}(\mathbf{x}(t))$ is increasing by an amount that is bounded away from zero. □

This observation further illuminates the relationship between our model and the models of herd behavior/cascades. In those models each agent observes the entire history of choices made in previous periods, so in any pair of agents, one will observe exactly the history observed by the other. As a result, certain histories may result in a situation where after a point, no agent can do better than to imitate the previous one. In other words, no agent receives a sample which forces them to make use of any information other than the social history, which is similar to the behavior of our system at the inefficient steady states when $N = 1$. By contrast, if N is greater than 1 in our framework, some agents receive a signal of the social history that is noisy enough that they will make use of additional information if any is available.

4. Alternative sampling rules

So far we have considered the “proportional” sampling rule, under which the odds of sampling an a-user exactly equal the share of the population using a. This section examines some alternative sampling rules, both because they may be equally plausible in some cases, and because this lets us identify the role that proportional sampling plays in our results. The first modification of the sampling rule allows for what we call “perception bias,” meaning that the probability $h(x)$ of sampling an a-user is independent of the payoff to a but need not equal the share x of a-users in the population. The second alternative we consider allows for “reporting bias,” in which people are more or less likely to talk about their experience with a choice if they were satisfied with it. In accordance with our rational-Bayesian methodology, we suppose that agents know the likelihood function generating their signals, so they are not "misled" by the biases in sampling. However, the biases can still alter the dynamics of social learning, because they change the information available to the agents. In particular, it is no longer true that adopting the action of the first person contacted yields an expected payoff equal to the average payoff in the current population, so Lemma 1 no longer applies, and we will see that the conclusions of our theorems fail to hold.

As the first example of a non-proportional rule, suppose that the numbers of a's and b's are the same in each agent's sample and at every point in time. In this "fixed-sample-composition" model, the relative popularity of the two choices conveys no information. Consequently, each generation of new agents faces exactly the same decision problem, namely to choose an action using only the information revealed by the signals, which need not reveal the true state. Hence in every period some fixed fraction of agents can choose a even in states of the world where b is better, so the long-run outcome can be heterogeneous; indeed the outcome will be heterogeneous unless the signals reveal the true state with probability one. This shows that the information provided by the in-sample popularity must play a key role in determining whether the system moves towards homogeneity.

4.1. Perception biases

Now consider a more general class of sampling rules that correspond to "perception bias." Specifically, let the probability of sampling an a user when fraction x of the population uses a be $h(x)$, with h a continuously differentiable function such that $h(0) = 0$, $h(1) = 1$, $h' \geqslant 0$, and $h(x) + h(1 - x) = 1$. As motivation, it seems plausible that a single person wearing black in a crowd of a thousand others wearing white is more likely to be in everyone's sample than would be warranted by unbiased sampling. Conversely, perhaps someone wearing light gray in a crowd of people wearing white may not get noticed as being different and therefore may be undersampled.[16] As before, each member of the sample is an i.i.d. draw—this precludes a conscious effort to have some of each choice in the sample. Then players who see a sample ζ consisting of m a's and $N - m$ b's will choose a if

$$\frac{\binom{N}{m}[h(x_1)]^m[1 - h(x_1)]^{N-m} q f(s \mid \bar{\theta}, \zeta)}{\binom{N}{m}[h(x_2)]^m[1 - h(x_2)]^{N-m} (1 - q) f(s \mid \underline{\theta}, \zeta)} > C. \tag{4.1}$$

The fixed-sample-composition example discussed above is a particular version of this kind of sampling rule with $h(x) = 1/2$ for all $x \neq 0, 1$. This example shows that with biased sampling inefficient outcomes can be globally stable. Our more general results here concern local as opposed to global stability.

Theorem 6. (a) *If $h(x)^N > x$ for x near* 0 (*severe oversampling of rare actions*), *then the efficient outcome is not even locally stable if payoff information has bounded informativeness.*

[16] This assumes that people do not choose how they are going to sample.

(b) *If $h(x)^N > x$ for x near 1, then the efficient outcome is locally stable for any N if payoff information has bounded informativeness.*

(c) *For any h function, and any $\varepsilon > 0$, every generation of new agents has expected payoff within ε of the full information outcome if N is sufficiently large.*[17]

Proof. See Appendix A. □

4.2. Reporting biases

Finally consider what may be the most plausible source of nonproportional sampling, over-reporting by agents with very high or very low payoffs.[18] This sort of reporting bias has two different effects. First, depending on how the sampling is modeled, there may be some agents who observe no other agents at all. An agent with such a sample will continue to hold the prior beliefs, and hence will chose a; this adds an impetus in the direction of "herding." The second effect arises from potential asymmetries in the probability of hearing from very satisfied and very dissatisfied agents. For example, if only very satisfied agents send signals, then in the neighborhood of the efficient point, the more popular choice will be oversampled relative to its frequency in the population.

Earlier versions of this paper present a specific example of a case where players are more likely to report if they have high payoffs. We will not present the details here, as all that is important for our general result is that the probability of sampling an a-user when the state is $i \in \overline{\theta}$ is $\overline{h}(x) > x$ for $0 < x < 1$, while the probability of sampling an a-user is less than its population share when the state is in $\underline{\theta}$.

Now in the neighborhood of the efficient outcome, oversampling the efficient choice is the same as oversampling the popular choice. However, it is easy to see that oversampling the efficient choice is not the same as oversampling the more prevalent choice in the region where all the x_i's are greater than $1/2$. As a result the two sampling rules have different consequences: Oversampling the popular choice can permit a steady state where the fraction of a-users is greater than $1/2$ in every state of the world, while as we now demonstrate, oversampling the efficient choice leads the system to globally converge to efficiency.

Theorem 7. *Suppose that the rule used for sampling is as follows: in state $i \in \overline{\theta}$, the probability of getting an a in any single drawing is $\overline{h}(x) > x,\ x \in (0,1)$. In state $i \in \underline{\theta}$, the probability of getting an a in any single drawing is $\underline{h}(x) < x,\ x \in (0,1)$. Then the system converges to the efficient outcome from any initial position.*

Proof. Suppose the probability of sampling an a-user is $\overline{h}(x) > x$ in $\overline{\theta}$, and $\underline{h}(x) < x$ in $\underline{\theta}$. If players adopt the action of the first agent in their sample, their expected payoff is $q\overline{h}(x_1(t))\overline{\Delta} - (1-q)(1-\underline{h}(x_2(t)))\underline{\Delta}$, and this exceeds $\overline{U}(\mathbf{x}(t))$ unless $x_1(t) = 1$ and $x_2(t) = 0$. Since expected payoffs are strictly increasing except at the efficient outcome, the efficient outcome is globally stable. □

[17] We thank David Levine for pointing this out to us.

[18] See the quotation from McKenna (1991) in footnote 4.

Note that this result holds even with multiple states of the world, and without a minimal informativeness condition. The reason is that here the simple rule "copy the first person you meet" guarantees that players do better than the current state unless the current state is fully efficient, while in the unbiased case this rule only guarantees a payoff equal to the current average. In contrast, if players are more likely to be sampled when they have low payoffs, we expect to find $\overline{h}(x) < x$ for $x \in (0, 1)$. Because near the efficient point the efficient choice is also the more prevalent choice, the dynamics of the system near the efficient point is going to be similar to that in the case with perception biases and undersampling of the prevalent choice, and so it is possible that the efficient point will be unstable.

5. Related work

Although the literature on social learning is now too large for us to give an exhaustive survey here, we should explain the paper's relationship to a few of the most closely related contributions.

We begin by comparing the model to the "herding" papers of Banerjee (1992) and Bhikchandani et al. (1992), in which there is a positive probability that agents will perpetually choose the wrong action. This contrasts with our finding that play converges to the full-information optimum when there are only two states of the world. We believe that the key to this difference in the results lies in the fact that in the present model, at any point of time, some people observe what we call uninformative histories, i.e., histories which force the decision-maker to rely on his own signal. To see why we think this is the key difference, consider a variant of the herding model, in which everyone decides in a fixed sequence, but every 5th decision-maker takes his decision without observing the history of other people's choices, and so makes a choice using only his own signal. The resulting system does not converge, since there is always the possibility of a long sequence of agents not observing the history and all making the same choice, thus switching the herd to that choice. However, the probability that someone taking a decision at a very late date who has observed the history will make the right choice should be close to one: Since everyone can tell when an agent failed to follow the herd, people will be able to look at the choices made by these people and figure out the truth.

This shows that the inefficient herding of the standard herding model does not occur if some agents are forced to use their own signals. In the model in the current paper, this happens endogenously, for example, when the sample is two A's and two B's. To introduce an analogous feature into the herding model, consider an environment where there are m side-by-side sequences of decision-makers. Each sequence is like the herding model in that each decision-maker gets a signal and then takes a decision, but unlike the herding model in that decision-makers in period t, instead of observing the entire history, observe a random sample of size 2 drawn from the m decisions made in period $t - 1$. Each of the m decision-makers in period t is assumed to observe an independently drawn sample. Our conjecture is that as m becomes large, in the long run most people in this world will make the correct choice. Intuitively, with m large, both choices will be present in the decisions made by the first round of decision-makers. Therefore, in the second round, a subset of the decision-makers will observe a sample with one of each choice, forcing them to use their own signals. This will continue over time and as the number of decisions based on signals rather than history increases, better and better decisions will be made on average.

Next we discuss Banerjee (1993)'s model of social learning with a continuum of agents, each of whom samples from the population of past decision-makers using proportional sampling. Unlike the symmetric choices in this paper, the choices in Banerjee (1993) are asymmetric: One option is to act and the second is to do nothing, and people remain unaware of the possibility of acting until they see someone who has already acted. Since the number of people who have already acted is a function of the state of the world, the date at which a player learns that the act is possible conveys information. Banerjee (1993) shows that the fraction of those who are currently choosing who make the right choice converges to 1, even though each person observes only one other. This contrasts with the finding in this paper, where inefficiency can persist when the sample size is one; the main source of this difference is that in the present model timing is not informative. Moreover, the interest in the convergence result in Banerjee (1993) is limited by the fact that the measure of the group that is making the choice goes to zero over time, so that even though asymptotically most active players make the right choice, most people in the population may make the wrong one.

Our results are also related to several papers that study the implications of exogenously given, boundedly rational learning rules in similar decision environments. In Ellison and Fudenberg (1995), each agent contacts N others (using proportional sampling) and observes their choice of action and their realized payoff.[19] Agents are assumed to use the following decision rule: If everyone in their sample uses a, the agent uses a; if everyone in the sample uses b, the agent uses b, and if there is at least 1 user of each choice, the agent chooses the action with the higher average payoff in the sample. Under this "must-see-to-adopt" rule, when $N = 1$ every distribution over actions is a stationary point; the interest is thus in larger sample sizes that allow players to receive "mixed" samples where both actions are used. In such cases, the decision rule specifies that players ignore the relative popularity and act as a Bayesian would if the odds ratio before seeing the payoffs was exactly 1. We now investigate the consequences of using that decision rule in the environment of this paper.

Fix a sample size N, and let $g > 1/2$ be the probability that a has a higher average payoff than b in a sample of size N, conditional on event $\overline{\theta}$. With the specified decision rule, g is also the fraction of agents with intermediate samples that choose a in event $\overline{\theta}$. Similarly, let $h < 1/2$ be the fraction of those with intermediate samples that chooses a in event $\underline{\theta}$.

If $N > 2$ and $g < (N-1)/N$, $h > 1/N$, then a computation shows that x_i is *decreasing* when it is near 1 and increasing whenever it is near 0. Consequently, the system cannot converge to efficiency, and moreover in the neighborhood of the efficient outcome $\overline{U}(\mathbf{x}(t))$ is decreasing. However, the simple rule above increases efficiency if the system is "far" from efficiency, and in the boundedly rational model agents do not try to use the correlation between the position of the system and the optimal choice in making their decision, either

[19] The decision environment differs slightly from that of this paper, as the mean payoff difference between the two choices is itself stochastic: each period, $u^a - u^b = 1$ with probability p, and $u^a - u^b = -1$ with probability $1 - p$, with the mean payoff difference in different periods being independent.

because they do not have the necessary information or they do not know how to process it. Thus the fact that the decision rule would be suboptimal if agents knew the way that the system variables behave is not necessarily an indication that the decision rule is implausibly suboptimal.

Schlag (1998) supposes that players observe the action and payoff of their "predecessor" and one (proportionally-sampled) individual under the same "must-see-to-adopt" restriction that is made by Ellison and Fudenberg; this corresponds to our model with $N = 2$. He characterizes the rules under which expected payoff is weakly increasing over time in all environments, and shows that the rule where the switch probability is a particular linear function of the payoff difference is "dominant" in a sense he makes precise. This rule leads to efficient long-run outcomes in large populations, which is consistent with our results.[20] Hofbauer and Schlag (1998) extend this work to larger sample sizes and consider the induced dynamics in a class of two-player two-action games, where the payoffs to the two actions depend on the distribution of actions in the opposing population. They find that the dynamics are very different with $N = 2$ and $N = 3$, even though both cases admit rules that tend to efficiency in the one-player case.[21]

Bjonerstedt and Weibull (1995) also consider $N = 2$ and "must-see-to-adopt," but suppose that players receive a noisy signal of the average payoff of the action they sample, instead of seeing the payoff realized by the particular player sampled. They show that average payoffs increase over time if the support of the noise is sufficiently large compared to the range of possible payoffs.

Smith and Sorenson (2000) develop a generalization of the one-agent-at-a-time model of Bhikchandani et al. (1992) in which there are several types of agents, each with different preferences. They show that learning in this model can converge to a situation in which the different types of agents make different choices, which they call "confounded learning." This confounded learning is superficially similar to the interior steady states of our model with $N = 1$, but the two in fact arise for quite different reasons: in Smith and Sorenson, confounded learning occurs when the history of past choices becomes uninformative, and agents respond only to their private signals. In contrast, at the interior steady states of our model agents ignore their private signals and base their decision only on the social history, as represented by their signal of popularity.

Finally, Bala and Goyal (1998) take the logic of word-of-mouth learning a step further: They argue, plausibly, that not only do people learn from a small number of people but also these people tend to be closer to them (in some sense) than the average person in the population—this is what they call "learning from neighbors." This contrasts with our model, where people learn from a random sample of others. We agree that this is in many ways a better assumption, at least in settings where the overall population is large. However, Bala and Goyal only look at the case where each person gets to observe an infinite number of independent draws from each of his neighbors: As a result the law of

[20] Schlag works with a finite population and shows that the long run is approximately efficient in the large-population limit.

[21] Once again, what they call "single sampling" corresponds to $N = 2$ in our model, since each player samples a predecessor and this is not counted in the sample size. In the case $N = 3$, they consider a particular sequential rule that does not allow for explicit popularity weighting.

large numbers can be applied and the learning from problem is substantially simplified. It would clearly be important to try to see what happens if we bring their assumption of learning from neighbors into our model—this is left for future research.

6. Concluding remarks

The broad aim of this paper was to identify the aspects of the economic environment that influence the long-run properties of learning by word-of-mouth communication. Lemma 1 shows that proportional sampling implies that the average expected payoff in the population never decreases over time. This weak monotonicity is automatic in single-agent models with perfect recall, and would be immediate if the agents choosing at time t had all of the information available at previous dates. Under proportional sampling, even though agents do not have all of the information used in the previous period, they have enough information to implement a decision rule that guarantees them at least the average payoff of the preceding period, while other sampling rules permit a form of “social forgetting” with average payoffs decreasing over time.

The paper also identifies situations where average payoffs are strictly increasing. Under proportional sampling, this is true at all positions except “corner states” if there is a sample size of at least two and any information at all in the signals, or if the sample size is three or more. Finally, the efficient outcome is globally stable if there is only one stationary point. This last condition in turn is satisfied if the signals agents receive are sufficiently informative relative to the number of states of the world and the prior beliefs. In particular, the global convergence holds under a weak informativeness condition if the only aggregate uncertainty concerns which of the two choices is better.

Weaker conditions are sufficient for the efficient point to be locally stable. For example, local stability obtains under proportional sampling with sample size at least two. For the case where there are only two states of the world, it also holds for a much larger class of sampling rules which allow for biases in the sampling process.

At the same time, it is important to emphasize that neither local nor global convergence is automatic. Local convergence fails even with proportional sampling when the sample size is one. It also can fail for sample sizes greater than one under some sorts of biased sampling. Global convergence can fail when there are more than two states of the world and the signals have limited informativeness.

Our results suggest a few tentative generalizations. First, proportional sampling seems relatively congenial to convergence to a homogeneous outcome, at least compared to relatively extreme forms of biased sampling. In the two-state model, it also leads to global convergence to the efficient point. A second tentative conclusion is that larger samples are more favorable to global convergence than smaller samples, both because larger samples increase the informativeness of the signals and also because a sample size of two or more allows “mixed” samples. These lead agents to respond to their observations of payoffs, so that new information can come into the system, while a sample size of one allows only “extreme” samples. This argument also suggests that it is important that agents in our model get independent samples. If everyone got the same sample (or a very similar sample as in the herding/cascades models) then there is no guarantee that some people will get a sample that makes them use their own information.

Acknowledgments

We thank Daron Acemoglu, Andres Almazan, Ramon Caminal, Glenn Ellison, David K. Levine, Eric Maskin, Peter Sorensen, Jeroen Swinkels, Jean Tirole, the associate editor, and the referee for helpful comments, and National Science Foundation grants SBR-9422942, SBR-9424013, and SES-0112018 for financial support.

Appendix A

Proof of Theorem 1′. If signals convey some information this is a special case of Theorem 1. Therefore the only interesting case is the one in which there is no payoff information at all. We will show that conclusion (a) of Theorem 1 continues to hold; this will imply (b) and (c) by the argument above.

As in the proof of Theorem 1, we will suppose that the agents cannot do better than imitating the first person in their samples and obtain a contradiction. As argued in that proof, if agents are willing to copy the first person in their sample, everyone who observes both an a and a b must be indifferent between choosing an a and choosing a b. Since there is no payoff information, this implies that $\lambda_\zeta(\mathbf{x}(t)) = C$ for every ζ that has at least one a and one b.

Now consider two alternative ζ's one of which has n a's and the other which has $n+1$ a's where $1 \leqslant n \leqslant N-2$. Call these, ζ and ζ', respectively. From above $\lambda_\zeta(\mathbf{x}(t)) = \lambda_\zeta(\mathbf{x}(t))$. A simple computation using the binomial formula shows that this implies $(1-x_2(t))/(1-x_1(t)) = x_2(t)/x_1(t)$ which implies $x_1(t) = x_2(t)$. But if $x_1(t) = x_2(t)$, $\lambda_\zeta(\mathbf{x}(t)) = \lambda_\zeta(\mathbf{x}(t)) = q/(1-q)$. But we have already assumed that $q/(1-q) > k$ which contradicts the assumption that $\lambda_\zeta(\mathbf{x}(t)) = k$ for all ζ that has at least one a and one b. □

Proof of Theorem 6. (a) In a neighborhood of the efficient point, $x_i \approx 0$ for $i \in \underline{\theta}$ and $x_i \approx 1$ for $i \in \overline{\theta}$. Since the h function is continuous and monotonically increasing, in a sufficiently small neighborhood of the efficient point, a sample ζ of all a's yields a likelihood ratio $\lambda_\zeta(\mathbf{x})$ that is close to infinity just as it does with proportional sampling.

Thus the assumption that the payoff signal is boundedly informative implies that everyone with a sample of "all a" will choose a regardless of their signals for all $\mathbf{x}$ in a sufficiently small neighborhood of the efficient point.

Therefore in a neighborhood of the efficient point, for $i \in \underline{\theta}$,

$$x_i(t) \geqslant (1-\gamma)x_i(t-1) + \gamma\big[h\big(x_i(t-1)\big)\big]^N > x_i(t-1).$$

(The last of these inequalities step makes use of the condition $h(x)^N > x$ for x near 0.)

This implies that for $i \in \underline{\theta}$, x_i is strictly increasing over time in a neighborhood of the efficient point. Therefore the efficient point is not stable.

(b) As argued above, when the system lies in a sufficiently small neighborhood of the efficient point, anyone who sees all a's or all b's will choose what they see irrespective of their payoff observation. Therefore in a neighborhood of the efficient point, for $i \in \overline{\theta}$,

$$x_i(t) \geqslant (1-\gamma)x_i(t-1) + \gamma\big[h\big(x_i(t-1)\big)\big]^N > x_i(t-1)$$

(the last step makes use of the condition $h(x)^N > x$ for x near 1) so that the fraction of those doing a increases in all the states where a is the correct choice. Also note that exactly the same argument can be used to prove that the fraction of those doing b increases in all the states where it is optimal to do b. Putting this together with the previous observation we have the result.

(c) Since payoff realizations are conditionally independent given the state of the world, when N is very large each new agent can guarantee himself approximately the efficient payoff by ignoring popularity completely and basing his choice only on the payoff observations. $\square$

References

Bala, V., Goyal, S., 1998. Learning from neighbours. Rev. Econ. Stud. 65, 595–621.

Banerjee, A., 1993. The economics of rumours. Rev. Econ. Stud. 19, 309–327.

Banerjee, A., 1992. A simple model of herd behavior. Quart. J. Econ. 107, 797–817.

Bhikchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and informational cascades. J. Polit. Econ. 100, 992–1026.

Bjonerstedt, J., Weibull, J., 1995. Nash equilibrium and evolution by imitation. In: Arrow, K., et al. (Eds.), The Rational Foundations of Economic Behavior. MacMillan.

Caplin, A., Leahy, J., 1994. Business as usual, market crashes, and wisdom after the fact. Amer. Econ. Rev. 84, 548–565.

Chamley, C., Gale, D., 1994. Information revelation and strategic delay in a model of investment. Econometrica 62, 1065–1086.

Duflo, E., Saez, E., 2000. Participation and investment decisions in a retirement plan: the influence of colleagues choices. Mimeo. MIT.

Ellison, G., Fudenberg, D., 1993. Rules of thumb for social learning. J. Polit. Econ. 101, 612–643.

Ellison, G., Fudenberg, D., 1995. Word-of mouth communication and social learning. Quart. J. Econ. 110, 95–126.

Hofbauer, J., Schlag, K., 1998. Sophisticated imitation in cyclic games. DP B-427. University of Bonn.

Lee, I., 1993. On the convergence of informational cascades. J. Econ. Theory 61, 395–411.

McKenna, R., 1991. Relationship Marketing: Successful Strategies for the Age of the Customer. Addison–Wesley, Reading, MA.

Moore, G., 1991. Crossing the Chasm: Marketing in High-Tech Industries. Harper Collins, New York.

Munshi, K., Myaux, J., 2000. Social norms, with an application of demographic transition in Bangladesh. Mimeo. University of Pennsylvania.

Schlag, K., 1998. Why imitate and if so, how? A boundedly rational approach to multi-armed bandits. J. Econ. Theory 78, 130–156.

Smith, L., Sorenson, P., 2000. Pathological models of observational learning. Econometrica 68, 371–398.

Udry, C., Conley, T., 2001. Social learning through networks: the adoption of new agricultural technologies in Ghana. Amer. J. Agricultural Econ. 83 (3), 668–673.

Vives, X., 1997. Learning from a welfare analysis. Games Econ. Behav. 20, 177–200.