Rational Social Learning with Random Sampling[*] [†]

Lones Smith[‡]
Department of Economics
University of Wisconsin

Peter Norman Sørensen[§]
Department of Economics
University of Copenhagen

October 19, 2014

Abstract

This paper explores rational social learning in which everyone only sees unordered random samples from the action history. In this model, herds need not occur when the distant past can be sampled. If private signal strengths are unbounded and *the past is not over-sampled* — not forever affected by any individual — there is complete learning and a correct *proportionate herd*. With recursive sampling, welfare almost surely converges under the new proviso that *the recent past is not over-sampled*. In this case, there is almost surely complete learning with unbounded beliefs and unit sample sizes. The sampling noise in this Polya urn model induces a path-dependent structure, so that re-running the model with identical signals generally produces different outcome.s

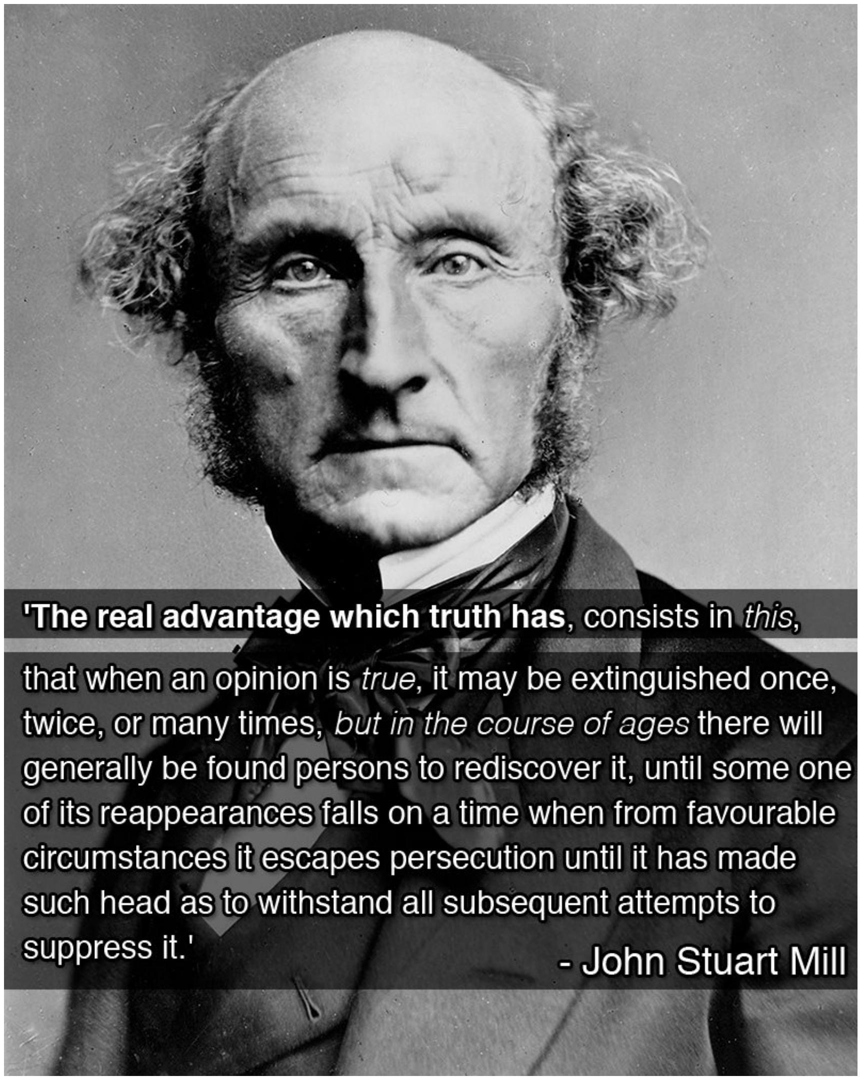
---

[*] *While the paper has been completely reworked, clarified, and polished, with two minor exceptions, all results and proofs in this paper first appeared as chapter three ("Rational Social Learning with Random Sampling") in Peter's 1996 MIT PhD thesis,* supervised by Lones and Abhijit Banerjee. The results were first presented at the MIT Theory Lunch in 1995, and some later documented in Chamley's 2004 text. We have removed new results on Polya urns with sample sizes more than one that appeared in subsequent working papers in the last decade, since we hit intractable open problems. The only new results are a comparative static in Lemma 2 and an observation in Corollary 2.

[†]The paper profited from presentations at theory lunches at Nuffield, Stockholm School of Economics, and theory seminars at MIT-Harvard, Oxford Séminaire Roy (Paris), and finally at CEPR's ESSET 1997 conference at Studienzentrum Gerzensee, and the 1997 Warwick Summer Research Workshop. Bruno Jullien, Jozsef Sakovics and Xavier Vives made suggestions. Josh Cherry provided MATLAB assistance. All errors remain our responsibility. Lones acknowledges financial support from the National Science Foundation throughout this project, and Peter the Danish Social Sciences Research Council.

[‡]e-mail address: `lones@ssc.wisc.edu`

[§]e-mail address: `peter.sorensen@econ.ku.dk`



1 INTRODUCTION

An infinite ordered sequence of individuals with identical preferences is faced with a one-shot action choice from a finite menu with uncertain payoffs. Decisions optimally reflect a private signal and the perfect knowledge of what all predecessors have done. In the unique equilibrium, everyone eventually settles upon one action, possibly an unwise one. This classic informational herding fable owes to Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992), whose working paper offered the metaphor of the blind leading the blind. Smith and Sørensen (2000) (hereafter, SS) later found that incorrect action herds happen with positive probability *exactly when* the individuals' signals are uniformly bounded in informativeness. Otherwise, a herd on the correct action must occur.

This paper pursues a different theory of rational social learning grounded on *anonymity*, i.e. the assumption at the heart of economics. We ask how well people learn when everyone knows her own ordinal rank but only sees how many in her sample took each action. We revisit claims in Surowiecki (2004) that "crowds" smartly aggregate information. We hereby introduce a new metaphor for social learning: In a *Polya urn*, a ball is drawn randomly from an urn containing black and white balls; its color is observed, and then it and a new like-colored ball is replaced; the urn composition evolves as the process repeats.

This setting subsumes previous social learning models: observing aggregate action statistics,[1] random samples,[2] or simply the immediate predecessor (Celen and Kariv (2004)). For motivation, consider how SS derived their results by exploiting a "dynamic discontinuity" in the herding model: By the *overturning principle*, a single violation of a would-be herd can overturn the weight of arbitrarily many predecessors' recorded actions. So a herd occurs since the absence of one precludes belief convergence, but this is a contradiction, since a bounded martingale converges; meanwhile with unboundedly strong private signals, any incorrect would-be herd that starts is eventually violated by a well-enough informed successor. But this dynamic discontinuity is counterfactual for most applied settings. For observing perfectly ordered actions histories is intuitively rare, and overwhelmingly memory intensive. Second, individual choices are often lost in the crowd.[3]

Random sampling radically alters social learning. Anonymity subtracts enough information from the sequential history that the overturning principle fails. For instance, an initial action sequence of AB is only seen as "either AB or BA" by #3. Whereas either sequence may push a standard herding model into one cascade or another, the anonymity assumption leaves #3 with no useful inference. So early actors are not as pivotal. Next consider what happens if #3 chooses action B, after seeing AA. In the original herding model, this effects a radical shift favoring B, by the overturning principle. But with random sampling, an *opposite message* emerges: the history of AAB still favors A at this stage. In other words, the very last individual is no longer pivotal, but instead sees his impact muted by the crowd. All told, the dynamic discontinuity disappears, and settled message of history herding model with anonymity is far less responsive to individual choices.

Our first finding is simple but notable: With random sampling from the entire action

[1] The earliest example is a continuum agent learning model in Smith (1991).

[2] The first published example we know is Banerjee and Fudenberg (2004) (hereafter, BF). Early mimeos of BF and Sørensen (1996) were contemporaneous at MIT. Still, we profited from seeing their paper.

[3] SS also considered a variant on the herding model with noisy choices in which the overturning principle failed because deviations from herds are eventually attributed mostly to error. That model is observationally distinct, as the noisy choices remained constant. Here, contrarian choices vanish with time.

history, a herd is impossible, unless it starts at the outset. For early deviants are forever later sampled and able to mislead. While just an example, it reveals how herding is no longer synonymous with information aggregation. In its stead, we introduce a new benchmark of *proportional herds*, or convergent fractions of choices. Proposition 1 derives the iff conditions for correct proportionate herds and *complete learning*, namely, that samples eventually reveal the truth: the unbounded informativeness condition on private signals of SS, and a new proviso: *the distant past must not be "over-sampled"* — not reliant on anyone with boundedly positive chance. For instance, observing either the most recent predecessor or a randomly drawn person from the past meet our sampling condition.

A key idea is that one does as well on average as a typical sampled predecessor (Lemma 2), and more so with better quality signals. Mere imitation of a randomly drawn action from the sample guarantees this. Corollary 1 finds that welfare is monotone and so converges with *recursive sampling*, an intuitive special case where past sample chances are discounted at a fixed rate. Corollary 2 leverages this and deduces that the best possible sampling protocol with unit sample sizes is to observe the immediate predecessor.

We next turn to a deeper novelty that uniquely arises with random sampling. Social learning from samples proceeds by comparing the chances of a sample in the two states of the world. But does a sample of AB occur half the time because the urn is equally full of A's and B's, or is it entirely A's or entirely B's with equal chance? These radically different outcomes generate similar messages. We must therefore understand not just what is the average behavior of the urn, but more strongly what is the realized path. The theory of Polya urns is not yet advanced enough to allow a sharp conclusion here, but in the special case of unit sample sizes, Proposition 2 deduces an almost sure limit. In this case, we find strong convergence of the action fractions with sample size one under one additional assumption: *the recent past must not be over-sampled.* This assumption rules out sampling just the immediate predecessor, and offers a big picture insight into the oscillation found by Celen and Kariv (2004). In summary, too much weight on the distant past precludes complete learning even in the presence of unbounded beliefs, while too much weight on the recent past can lead to divergent behavior, such as oscillations.

Our analysis relies on a key technical innovation. The proof logic in SS exploited the martingale character of *public beliefs* and their implied likelihood ratios. But here, the pre-history of any two individuals is not commonly observed; thus, the information set is not "growing" (i.e. a filtration), and a "public" belief is not a meaningful notion: Rather, everyone recursively computes the chance of a sample in each state, and thereby deduces *sample beliefs* via likelihood odds. We have nonetheless rescued a martingale analysis that should prove generally useful in social learning. We exploit the fact that the cumulative sums of one-period look-ahead forecast errors for any stochastic process is a martingale.

The social learning literature is large, but two papers stand out for relevance. We will return in §B.2 to a detailed discussion of the elegant model of BF, since the continuum model comparison is technical. We claim that not modeling the stochastic path dependence makes it a poor approximation of the finite agent world. The continuum world essentially secures its tractability by *averaging over mutually exclusive histories of finite agent models.*

Also closely related is Acemoglu, Dahleh, Lobel, and Ozdaglar (2008) (ADLO), recently published in this journal. The original herding models assumed the simplest network structure; following on work by Gale and Kariv (2003), ADLO characterizes complete learning in social networks. Since people know precisely the identity of sampled individuals, it is far from our paper exploring anonymity. But its complete learning characterization is related to ours that the past not be over-sampled; we defer a careful comparison until §B.1.

Arthur and Lane (1994) also uses urn theory, but agents learn from past outputs, and not actions. Building on our (1996) working paper, Celen and Kariv (2004) is an insightful study of the predecessor sampling model, whose key divergence insight we revisit. Chamley (2004) offers a gentle treatment of random sampling work, including our own, in §5.1. Monzon and Rapp (2011) relax our informational assumption, denying individuals the knowledge of their current decision rank; they derive similar complete learning conclusions.

Section 2 sets up the model. Section 3 develops two key lemmas for learning from signals and samples. In sections 4 and 5, we study the expected and the stochastic evolution of the model. The Appendix offers cautionary insights to help guide work on random sampling.

2 THE MODEL

A. Private Signals. A probability space $(\Omega, \mathcal{E}, \nu)$ underlies all randomness. There are two *states* (of the world): $\theta = H$ ('high') and $\theta = L$ ('low'). So this partitions the background state space Ω into events H and L, with common prior $\nu(H)=\nu(L)=1/2$.

An infinite sequence of exogenously ordered individuals $n = 1, 2, \ldots$ sequentially acts. Each initially sees a private signal $\sigma \in \Sigma$ about the state — assumed i.i.d. across individuals conditional on the state. The signal is distributed according to the probability measure μ^θ in states $\theta = H, L$. Some signals are informative, so that $\mu^H \neq \mu^L$, but no signal perfectly reveals the state, so that μ^H and μ^L are mutually absolutely continuous. Thus, there exists a positive and finite Radon-Nikodym derivative $g = d\mu^L/d\mu^H : \Sigma \to (0, \infty)$ of μ^L w.r.t. μ^H.

Using Bayes' rule, the individual computes his *private belief* $p(\sigma) = 1/[g(\sigma)+1] \in (0,1)$ that the state is H. Conditional on the state, private beliefs are i.i.d. across individuals because signals are. In states $\theta = H, L$, the private belief p has distribution F^θ on $(0,1)$, where F^H and F^L have a common support supp(F). By construction, the convex hull is $\text{co}(\text{supp}(F)) \equiv [\underline{p}, \bar{p}] \subseteq [0,1]$ with $0 \leq \underline{p} < 1/2 < \bar{p} \leq 1$, since μ^L and μ^H are distinct. We call the private beliefs *bounded* if $0 < \underline{p} < \bar{p} < 1$; if $\text{co}(\text{supp}(F)) = [0,1]$, private beliefs are *unbounded.* To exhaust all possibilities we should also consider supports that are bounded above and not below, and conversely, but this tedious exercise sheds no additional insights.

A restricted symmetric class of signals affords some tighter results. As with a weather forecast, imagine first drawing a *signal quality* q from a distribution F over $(0,1)$, and then learning one of two possible statistically true statements "with chance q, the state is high/low". Then F is symmetric, i.e. $F(p) = 1 - F(1-p)$, since a 70% or more chance of rain is a 30% or less chance of sun. Given a flat prior, one's private belief is $q/[q+(1-q)] = q$ after learning that $\theta = H$ has chance q, and is $1 - q$ if told that $\theta = L$ with chance q. So the private belief cdf obeys $dF^H(p) = p dF(p)$ in state H, and $dF^L(p) = (1-p)dF(p)$ in state L. These cdf's inherit symmetry: $F^H(p) = \int_0^p dF^H(r) = \int_{1-p}^1 dF^L(r) = 1 - F^L(1-p)$.

For example, assume signal quality has a uniform distribution on $[0,1]$. Then $F(p) = p$, and so private beliefs are unbounded with cdf's $F^H(p) = p^2$ and $F^L(p) = 2p - p^2$ on $[0,1]$.

B. Action Choices. Everyone chooses among actions a, b, seeking to take the action that maximizes his expected payoff. We assume that a is a safe action, and b is a risky action, better in state H. (The restriction to two actions is not crucial to our results, except in §A.) Actions have common vNM payoffs $u^H(a) = u^L(a) = 0$, $u^H(b) = 2u$, and $u^L(b) = -2$. In other words, *a perfectly revealing signal yields expected payoff* u (and so is our complete learning benchmark, for state H). Action b is best iff the posterior belief r obeys $ru-(1-r)\geq 0$, i.e., exceeds the threshold $1/(1+u)$. To avoid trivialities, $1/(1+u)$ lies strictly inside the support of private beliefs — we do not start out in a cascade.

C. Random Sampling. Every individual n observes an *unordered* sample of actions $s=(s_a,s_b)\in\mathcal{S}=\{0,1,2,\ldots\}^2$ drawn from the pool of history, namely, the numbers s_a and s_b of sampled predecessors who took the two actions. Two steps removed from the standard herding model, an individual is neither apprised of the action sequence, nor what samples were seen by predecessors. A sampling process Σ defines for each individual n the chance of drawing each (possibly empty) subset $J\subseteq\{1,2,\ldots,n-1\}$ of predecessors. The samples and signals sampled by $n\neq m$ are independent, and the sampling process is independent of the state. For a given sampling process Σ, we can derive the chance that n draws a sample of size $j<n$; we assume that nonempty samples occur with positive chance. We can also compute the chance $\tau(n,m)$ that agent n samples predecessor $m<n$. We let $\tau(n,0)$ be the chance that individual n samples the empty sample, so $\sum_{m=0}^{n-1}\tau(n,m)=1$.

For a salient special case, assume that pairs of individuals sample the common past in the same way: in other words, $\tau(n+1,m)/\tau(n,m)=1-\pi_n$ constant for all $m<n$. This sampling process is *recursive* since the induced measure over $\{1,\ldots,n\}$ consists of some weight π_n on n, and remaining weight $1-\pi_n$ on the previous distribution over $\{1,\ldots,n-1\}$.

The only *stationary* recursive sampling process involves geometric weighting, where individual n samples $m<n$ with frequency $\tau(n,m)\propto\pi_{m+1}\Pi_{m+2}^{n}(1-\pi_i)=\pi(1-\pi)^{n-m-1}$, where $\pi<1$, and the limit *proportional sampling* world, where $\tau(n+1,m)=1/n$ for all $m\leq n$, and thus $\pi_n=1/n$. Think of $1-\pi$ as a *decay factor* on the information in old choices. This special case of our model in turn subsumes other studied social learning models. In the limit $\pi\to 1$, all information decays, as only the immediate predecessor is sampled, as in Celen and Kariv (2004). For fixed $\pi>0$, recent predecessors are more heavily weighted, and the past is discounted. The limit $\pi\to 0$ includes the cases where one sees all predecessors' actions without order (Smith (1991)), as well as BF's proportional finite sampling model, in which all predecessors are sampled with equal chance.[4]

D. The Dynamic Behavior of Sample Beliefs. Individuals $n=1,2,\ldots$ play a Bayes-Nash equilibrium. Each learns from history by drawing samples of predecessors' actions. For any given sampling process, the observation by individual n of any sample depends stochastically on the true state and the realized history of length n. In equilibrium, sampling actions allows one to make imperfect inferences about the sampled individuals'

[4]Since they assume a positive death rate, they too have a decay factor $1-\pi<1$ on old information.

private signals. Since one can calculate the probabilities of making that observation in either state, individual n can then form his *sample belief* q_n in state H. This would be his posterior belief in state H had he a purely neutral private belief.

As private signals are random, sample beliefs $\langle q_n \rangle_{n=1}^{\infty}$ are a stochastic process. Then individual n forms his posterior belief r_n from the sample belief q_n and the private belief p_n using Bayes' rule:

$$r_n = \frac{p_n q_n}{p_n q_n + (1-p_n)(1-q_n)}. \tag{1}$$

3 LEARNING FROM SIGNALS AND FROM SAMPLES

After seeing one's private signal, further Bayesian updating must be in vain, but this only holds if $(dF^H/dF^L)(p) = p/(1-p)$. This *no introspection condition* from SS quantifies how much more proportionately strongly signals in favor of a state occur in that state.

Lemma 1 (Signal Tails are Informative) *The inequalities* $(1-p)F^H(p) \leq pF^L(p)$ *and* $(1-p)(1-F^H(p)) \geq p(1-F^L(p))$ *obtain for all* $p \in [0,1]$. *Moreover, the first inequality is strict when* $p > \underline{p}$, *and the second is strict when* $p < \bar{p}$.

Proof: As a standard inequality for the monotone likelihood ratio property, we arrive at:

$$\frac{F^H(p)}{F^L(p)} \leq \frac{dF^H(p)}{dF^L} \leq \frac{1-F^H(p)}{1-F^L(p)}$$

The no introspection condition finishes the inequality. Strictness follows since $(1-p)F^H(p) - pF^L(p)$ and $(1-p)(1-F^H(p)) - p(1-F^L(p))$ strictly decrease on the support of p. □

We next develop a useful insight that individuals can use their observed sample of predecessors to obtain an expected welfare above that of the average sampled predecessor. Given an observed sample of actions, an individual could randomly mimic one of the sampled actions. We build on this nice insight from Lemma 1 in the deterministic model of Banerjee and Fudenberg (2004), but additionally for our stochastic setting, we can precisely measure the welfare improvement in terms of the private signal distribution.

The sampling process defines the chance that any subset of predecessors is drawn, as well as the equilibrium chances of the actions taken, and thereby the chance P_s^θ of any sample $s \in \mathcal{S}$ in state θ. Also, let $\beta(s)$ be the chance that a uniformly drawn individual from s took action b.[5] The probability that an average sampled predecessor takes action b

[5]So if j out of ℓ individuals in s choose b, then $\beta(s) = j/\ell$. With sample size zero, we let $\beta(s)$ be the chance that an individual with posterior $1/2$ chooses b.

is therefore

$$R^{\theta} = \sum_{s \in \mathcal{S}} \beta(s) P_s^{\theta}. \tag{2}$$

This equivalently expresses the expected proportion of the predecessors who took action b. Consequently, $uR^H - R^L$ is the expected welfare of the average sampled population. The *sample belief* resulting from the sample s is $q(s) = P_s^H/[P_s^H + P_s^L]$. Outside of *cascade*, we have $\underline{p} < P_s^L/(uP_s^H + P_s^L) < \bar{p}$ for some $s \in \mathcal{S}$. The sample and private beliefs produce the individual's posterior r, as in (1). Individual n takes action b exactly when $ru - (1-r) \geq 0$, i.e. exactly when the private belief p exceeds the threshold $(1-q)/[uq + (1-q)]$. The probability that individual n chooses action b in state θ is then

$$Q^{\theta} = \sum_{s \in \mathcal{S}} P_s^{\theta} \left[1 - F^{\theta} \left(\frac{P_s^L}{uP_s^H + P_s^L} \right) \right]. \tag{3}$$

Lemma 2 (Welfare Improvement) *One's expected welfare exceeds one's average sampled predecessor's, and strictly so outside of a cascade. In the signal quality paradigm, the improvement is greater given a mean-preserving spread of the signal quality distribution F.*

Proof: Given $(1-F) - \beta \equiv (1-\beta)(1-F) - \beta F$, we can regroup terms in (2) and (3):

$$\begin{aligned}
& uQ_n^H - Q_n^L - uR_n^H + R_n^L \\
&= \sum_{s \in \mathcal{S}} \beta(s) \left\{ P_s^L F^L \left(\frac{P_s^L}{uP_s^H + P_s^L} \right) - uP_s^H F^H \left(\frac{P_s^L}{uP_s^H + P_s^L} \right) \right\} \qquad (4) \\
&+ \sum_{s \in \mathcal{S}} (1 - \beta(s)) \left\{ uP_s^H \left[1 - F^H \left(\frac{P_s^L}{uP_s^H + P_s^L} \right) \right] - P_s^L \left[1 - F^L \left(\frac{P_s^L}{uP_s^H + P_s^L} \right) \right] \right\}. \\
&= \sum_{s \in \mathcal{S}} \left\{ P_s^L F^L \left(\frac{P_s^L}{uP_s^H + P_s^L} \right) - uP_s^H F^H \left(\frac{P_s^L}{uP_s^H + P_s^L} \right) \right\} + \sum_{s \in \mathcal{S}} (1 - \beta(s))(uP_s^H - P_s^L)
\end{aligned}$$

The desired inequality $uR^H - R^L \leq uQ^H - Q^L$, and its strict version, owes to Lemma 1, for $p = P_s^L/(uP_s^H + P_s^L)$. Finally, in the signal quality world, integration by parts yields:

$$pF^L(p) - (1-p)F^H(p) = p \int_0^p (1-t) dF(t) - (1-p) \int_0^p t dF(t) = \int_0^p F(p)$$

This increment rises with a mean-preserving spread in F (higher quality more likely). □

In words, because signals are informative, the private belief tails favor the corresponding correct state, and push individuals stochastically towards the correct action for each state.

4 PROPORTIONATE HERDS AND MEAN CONVERGENCE

We begin with a fundamental way that random sampling overturns the signature informational herding message — namely, that herds may eventually start. To see this, assume proportional sampling and a positive probability of boundedly finite sample sizes $|S| \leq 2k-1$, some $k > 0$. *Unless a herd starts by period k, there is a positive chance that an infinite subsequence of individuals chooses a contrary action.* To see why, assume that a herd starts after period one with positive probability. Now, the initial k deviants will almost surely be sampled by infinitely many successors in samples of size $2k-1$ or less.[6] But the herd persists only if the sample belief from history eventually overwhelms all private beliefs.[7] Perversely, this means that these early deviants carry tremendous weight when sampled late enough. Nearly anyone drawing such a sample will eventually mimic it, and choose the same suboptimal action. So, even if the chance that people take any given action converges to one, an infinite subsequence takes the contrary action.

Abandoning hope of stochastic regularities for the total numbers choosing actions, we focus instead on the strongest form of action convergence that we can hope for — namely, *proportionate herds*, when the fraction of individuals taking an action converges to one. So the share of b-takers converges to 1 in state H and to 0 in state L. Call learning *complete* if sample beliefs eventually focus on the correct state. But this happens iff samples do not mislead, and thus iff a correct proportionate herd arises. Define the respective expected welfare V_n of individual n and W_n of his average sampled predecessor. Then $\lim_{n\to\infty} V_n = \lim_{n\to\infty} W_n = u$ is both necessary and sufficient for complete learning.

Intuitively, learning is complete when the sampling mechanism casts a wide enough net among sufficiently informed individuals. Social learning must be neither too forgetful (seeing empty samples) nor too non-acquisitive (seeing just early deciders).

Definition *The sampling process* Σ does not over-sample the past *if for all* $m \in \mathbb{N}$ *and* $\varepsilon > 0$*, there exists* $M > m$ *such that* $\tau(n,m) < \varepsilon$ *and* $\tau(n,0) < \varepsilon$ *for all* $n \geq M$*.*

In particular, a recursive sampling process described by (π_k) does not over-sample the past when the chance of sampling in $\{2,\ldots,n\}$ vanishes as $n \to \infty$. By the independence of the samples, this chance is $\Pi_{k=2}^{n}(1-\pi_k)$, which vanishes when $\Pi_{k=2}^{\infty}(1-\pi_k) = 0$.

We next claim that no over-sampling the past and unbounded private beliefs are jointly sufficient for complete learning, and that learning is incomplete if either fails.

[6] For instance, with sample size one, each individual $n+1 > N$ samples N with probability $1/n$. Since $\sum_{n=N+1}^{\infty} 1/n = \infty$, the Second Borel-Cantelli Lemma implies that individual N is sampled infinitely often.

[7] So the analog of a limit cascade happens, as SS deduce for the standard observational learning model.

Proposition 1 (Complete Learning) *(a) If Σ does not over-sample the past and private beliefs are unbounded, learning is complete and a correct proportionate herd occurs.* *(b) Learning is incomplete if Σ over-samples the past:* $\liminf_{n\to\infty} V_n, \liminf_{n\to\infty} W_n < u$. *(c) Learning is incomplete for bounded private beliefs and almost surely non-empty samples.*

Proof of (b). When the past is over-sampled, fix m and $\varepsilon > 0$ and a subsequence n_k of individuals observing either m or nothing with chance at least ε. Ex ante, individual m errs with some fixed positive chance since it is based on less than m private signals. As samples are unordered, those including m (or no one) cannot achieve maximal welfare. So W_{n_k} is bounded below u. From (2) it follows that there is positive probability of sample beliefs bounded away from 0 and 1. Then (3) implies that V_{n_k} is bounded away from u.

Proof of (c). Assume bounded beliefs, no zero-size samples, and the past not over-sampled. If #1 chooses action b with chance one, then so must all successors, since the sample beliefs are unchanged. Complete learning cannot obtain in this case. Assume next that #1 takes either action with positive probability. If there is complete learning, then the probability Q_n^H with which individual n takes the correct action b in state H converges to 1. Then there exists N so large that later individuals who observe pure b samples are sufficiently convinced about the true state of the world, and so ignore their own signal. In state L, the first N individuals all take action b with positive probability. Since this history remains pure after individual N, Q_n^H cannot converge to 1 — contradiction. □

The delicate contradiction proof of part (a) is in the appendix. For a direct intuition, observe that — paraphrasing Newton — no over-sampling the past ensures that everyone stands on the informational shoulders of giants, namely, sampling from those who have sampled from predecessors, etc. in longer and longer chains. For instance, the definition easily yields a threshold $M_\varepsilon(m)$ such that with chance at least $\varepsilon > 0$, *no individuals* in $\{1, 2, \ldots, m\}$ are sampled by any given $n > M_\varepsilon(m)$. Define any sequence by $m_1 \in N$, and recursively $m_{k+1} = M_\varepsilon(m_k)$ for $k = 1, 2, \ldots$. Then, for instance, with chance at least $1-\varepsilon$, individual $m_3 + 1$ samples among $\{m_2 + 1, \ldots, m_3\}$, each of whom with chance at least $1-\varepsilon$, sampled among $\{m_1+1, \ldots, m_2\}$, each of whom sampled a predecessor with chance at least $1 - \varepsilon$. To wit, complete learning is possible given the accumulation of signals.

Let us see the role of the non-empty samples proviso in Proposition 1 (c). Inspired by Sgroi (2002), let individuals $2^0, 2^1, 2^2, \ldots$ be sacrificial lambs, unable to view predecessors' actions. Assume that n observes the unordered sample containing every $2^k < n$ before acting. This sample consists of conditionally iid realizations, and thus its informativeness explodes as kn does, even though individuals are sampled without order. So there is complete learning by all individuals except powers of 2, and thus by a fraction tending to 1. Indeed, as Surowiecki (2004) writes in his popular book: "One key to successful group decisions is getting people to pay much less attention to what everyone else is saying."

This example speaks to the importance of how well the sampling process "mixes" over histories. If groups of individuals just sample among themselves, or in one direction only (as in the last example), then each group might well achieve different outcomes. But when everyone is sampled the same proportionately, as with recursive sampling, a common limit expected welfare emerges — irrespective of private beliefs. Since by Lemma 2, the expected welfare of agent n exceeds the average sampled welfare among $1, \ldots, n-1$, we conclude:

Corollary 1 (Increasing Welfare) *Assume recursive sampling. For all private signals:*
(a) *The expected welfare* $uR_n^H - R_n^L$ *of the average sampled population weakly increases, and so converges. The limit welfare is less than* u *in state* H *with bounded beliefs.*
(c) *Welfare strictly rises when not in a cascade.*
(d) *In the signal quality paradigm, welfare rises more with a mean-preserving quality spread.*

This asserts that welfare converges, but not that a proportionate herd arises, or even that the b-sampling chances (R_n^H, R_n^L) separately converge. Also, our proof fails for nonrecursive sampling mechanisms. For instance, assume a very weak signal. Then if everyone samples his two immediate predecessors, the welfare sequence $(0, 3, 2, 2.8, \ldots)$ is consistent with each individual beating his average sample. Individual 4's average sampled welfare is 2.5 but 5's is only 2.4. This roughly captures the logic of the welfare monotonicity failure.

Corollary 1 identifies the best recursive sampling mechanism with unit sample sizes.

Corollary 2 (Efficient Recursive Sampling) *Sampling the immediate predecessor is the most efficient recursive sampling mechanism with unit size.*

The proof is instructive. Consider a recursive sampling mechanism in which individual n samples the predecessor with chance $\pi_n \in (0, 1)$. By Corollary 1, we know that welfare is monotone: $V_1 < V_2 < V_3 < \cdots$ when not in a cascade. Let us call the values when observing the predecessor $\hat{V}_1 < \hat{V}_2 < \hat{V}_3 < \cdots$. Hereby, we assume that there is no cascade, so that these inequalities are strict, for otherwise, values have converged, and the proof is trivial. Suppose that individual n is given an additional signal about his sample, indicating whether he is sampling his predecessor. Since the signal can be ignored, it weakly raises the expected payoff. That $V_n < \hat{V}_n$ follows from:

$$V_n \leq \pi_n \hat{V}_n + (1 - \pi_n) V_{n-1} < \pi_n \hat{V}_n + (1 - \pi_n) V_n$$

5 ALMOST SURE CONVERGENCE VIA URNS

So far we have explored the unconditional properties of the model. But in any finite agent stochastic learning model, convergence in mean possibly conceals complex patterns reflecting the path dependence of the urn model. In fact, we argue that it does. For a foretaste, Celen and Kariv (2004) found oscillations with sample size one. Also, as noted in the introduction, opposing purification herds are unconditionally indistinguishable from a fully mixed outcome. We now tackle head-on the problems of path dependence.

To this end, consider the Polya urn model. Eggenberger and Polya (1923) created it as a model contagion, and we use it for "informational contagion." At each stage, a randomly-chosen ball is examined, and another of the same color (black or white) is added. Starting with B_0 black and W_0 white balls, the limit fraction of white balls W_n/n converges to a beta distribution $\beta(W_0, B_0)$.[8] Balls here are individuals, and colors the chosen actions.

The recursive sampling model with unit sample sizes therefore falls prey to methods in Arthur, Ermoliev, and Kaniovski (1986) (henceforth AEK). They explore the evolution of generalized Polya urns containing balls having a finite number of colors. Under stringent conditions, AEK describe the limit distribution of balls in the urn.

We focused in §4 on the (time-0) *expected* chance $R_n^\theta = P_n^\theta(b)$ of sampling b among $\{1, 2, \ldots, n-1\}$ in state θ. We now turn to the *realized* chances X_n, in other words, the chance given the history up to n's predecessor; these reflect the realized urn composition. There is *almost surely complete learning* if $X_n^H \to 1$ a.s. and $X_n^L \to 0$ a.s.[9]

Fix the state $\theta = H$. Let $i_n = 1$ if individual n takes action b, and otherwise $i_n = 0$. As a function of the current action proportion x, the chance $\chi_n^H(x)$ that $i_n = 1$ is

$$\chi_n^H(x) = x\left[1 - F^H\left(\frac{R_n^L}{uR_n^H + R_n^L}\right)\right] + (1-x)\left[1 - F^H\left(\frac{1-R_n^L}{u(1-R_n^H) + (1-R_n^L)}\right)\right].$$

Sampling b is more likely if $\theta = H$, and thus $R_n^H > R_n^L$. Thus, $x \mapsto \chi_n^H(x)$ is a positively sloped linear function and a contraction. Since people eventually follow their sample with large chance, this curve tends to the diagonal — where AEK's theory has no bite.

Assume state H. Given recursive sampling with unit sized samples, the process (X_n) obeys $X_{n+1} = (1-\pi_n)X_n + \pi_n i_n$. So the *forecast error* $\epsilon_n^H(X_n^H) = i_n - \chi_n^H(X_n^H)$ obeys the recursion:

$$X_{n+1}^H - X_n^H = \pi_n\left[\chi_n^H(X_n^H) - X_n^H + \epsilon_n^H(X_n^H)\right]. \tag{5}$$

[8]See Freedman (1965) and more recently, §3.2 of the book Mahmoud (2009).

[9]We let X_n^θ be the process X_n conditional on state θ.

Since $E[\epsilon_n^H(X_n)|X_n] = 0$, its drift is $\pi_n(\chi_n^H(x) - x)$. So a deterministic analogue of (5) is

$$R_{n+1}^H - R_n^H = \pi_n(\chi_n^H(R_n^H) - R_n^H) \tag{6}$$

swapping the argument of χ_n^H. When the functions χ_n^H are constant, AEK prove that X_n^H converges, and hence so does R_n^H. The approach here is more subtle in our Bayesian world because χ_n^H depends on R_n^L and R_n^H. Our next result finds a.s. convergence of the "unexpected motion" $X_n^H - R_n^H$ if both the past *and* recent past are not over-sampled.

Proposition 2 (Unit Sample Sizes) *Assume everyone samples one predecessor, and that sampling is recursive, with weights satisfying* $\sum_1^\infty \pi_n = \infty$ *and* $\sum_1^\infty \pi_n^2 < \infty$.
(a) *The forecast error* $X_n^\theta - R_n^\theta$ *converges a.s. in states* $\theta = L, H$*;*
(b) *There is a.s. complete learning if the private beliefs are unbounded.*

Part (a) says that if the population mean converges, its composition almost surely does. This premise is true, eg., for all private beliefs in the symmetric binary model found in Appendix A, and always holds with unbounded beliefs, by Proposition 1 — hence part (b).

The twin premises of Proposition 2 capture the dynamic tension needed for almost sure convergence. The weight π_n on the last individual must be high enough that information accumulates ($\sum \pi_n = \infty$), but low enough that the lesson of history can "settle down" ($\sum \pi_n^2 < \infty$). The first condition demands that Σ sample the recent past enough (large π_n), or as we have put it, not over-sample the past.[10] The second asks that Σ *not over-sample the recent past.* Because over-sampling the past ($\sum_{n=1}^\infty \pi_n < \infty$) is a stronger condition than not over-sampling the recent past, some recursive sampling regimes will exhibit mean convergence but not almost sure convergence, like observing one's immediate predecessor ($\pi_n = 1$). Proportional sampling ($\pi_n = 1/n$) satisfies both conditions in Proposition 2, as does any recursive sampling with a decay factor $1 - \pi_n \le 1 - \underline{\pi} < 1$ bounded below one.

To see the necessity of not over-sampling the recent past, consider how Celen and Kariv (2004) found that oscillations could arise when sampling one's immediate predecessor. For example, assume uniform quality signals, so that $F^H(p) = p^2$. If individual n sees action $j = a, b$ in state θ with chance $P_n^\theta(j)$, then he copies any sampled action j when his private beliefs are at least $\psi_n(j) \equiv P_n^L(j)/(P_n^L(j) + P_n^H(j))$. Symmetry yields $P_n^H(b) =$

[10] Assume $\sum_1^\infty \pi_n = \infty$. Since $1 - x \le e^{-x}$ for all x, the limit as $N \to \infty$ of $\Pi_1^N(1 - \pi_n) \le e^{-\sum_1^N \pi_n}$ is zero. Conversely, assume $\sum_{n=1}^\infty \pi_n < \infty$. Then $\sum_{n=N}^\infty \pi_n < 1$ for large N. Since one can prove (by induction) that $\Pi_{n=1}^N(1 - \pi_n) > 1 - \sum_{n=1}^N \pi_i$ for all N, $\Pi_{n=1}^\infty(1 - \pi_n) = \Pi_{n=1}^{N-1}(1 - \pi_n)\Pi_{n=N}^\infty(1 - \pi_n) > 0$.

$1 - P_n^L(b) = 1 - \psi_n(b) = \psi_n(a)$. Then individual n chooses action b in state H with chance

$$P_{n+1}^H(b) = P_n^H(b)[1 - \psi_n(b)^2] + (1 - P_n^H(b))[1 - \psi_n(a)^2] = P_n^H(b) + [1 - P_n^H(b)]^2 \quad (7)$$

given (3). This is obviously also the chance that individual $n+1$ samples action b.

When $P_n^H(b)$ is near the limit 1, the difference equation (7) is well-approximated by the differential equation $dP = (1-P)^2 dt$. To wit, $1 - P_n^H(b) = O(1/n)$. Since $\sum 1/n = \infty$, action a will a.s. be observed infinitely often in state H, by the Second Borel-Cantelli Lemma (and independence). Since the mimicking chance is bounded away from 0, the realized actions switch infinitely often from b to a, and of course, back again as $P_n^H(b) \to 1$.

Proof of Proposition 2: We first claim that the *cumulative forecast error process* $\mu_n = \sum_{k=1}^n \pi_k \epsilon_k^H(X_k)$ is a martingale with respect to the σ-algebra generated by $(X_1, \ldots, X_n)$. To see this, check that $E[\epsilon_n^H(X_n)|X_n, H] = 0$. Since $\epsilon_n^H(X_n)$ and $\epsilon_m^H(X_m)$ are uncorrelated for $m \neq n$ and any X_n and X_m, the variance of μ_n is $\sum_{k=1}^n \pi_k^2 \left[\mathrm{Var}(\epsilon_k^H(X_k))\right] \leq \sum_{k=1}^\infty \pi_k^2 < \infty$. Having verified this, we can apply the Martingale Convergence Theorem for bounded variance random variables (Theorem 5.14 in Breiman (1968)) to deduce that μ_n converges a.s. to a random limit μ_∞. Since the cumulative tail forecast errors $\sum_{k=n}^\infty \pi_k \epsilon_k^H(X_k)$ vanish, the Appendix proves that the drift of the system fixes its evolution. □

An omitted part (c) might have considered bounded private beliefs. With unit sample sizes and recursive sampling, everyone eventually mimics their sample — private information is eventually ignored since welfare W_n converges, by Proposition 1. This behavior corresponds to adding a same colored ball as the one sampled in Polya's urn. This yields an easy insight: the limit outcome cannot be a correct purifying herd in a cascade, since samples would identify the state, and mimicking one's sample would be optimal. But that induces the Polya urn, whose long-run proportion of balls of each color — a beta distribution with full support on $[0, 1]$ — entails a mixed population. But in general, a cascade never starts, and action proportions randomly evolve en route to the mimicking limit.

Assume weights are not unbounded as in Proposition 2. With unbounded beliefs, eventually there are individuals with arbitrarily strong and incorrect beliefs.

6 CONCLUSION

Random sampling dramatically changes both the predictions and analysis of the social learning paradigm. For it induces a richer *path-dependent process*, in which randomness owes not only to the variability of individual signals, but also to the vagaries of who samples whom. Accordingly, re-running the model with the same private signals and different realizations of the sampling process can induce a radically different outcome.

In this rich framework, we have found that if early deviants can be re-sampled, then unlike the sequential social learning models, only proportionate herds can emerge. When the past is not over-sampled, the condition in SS that private signals have unbounded strength secures complete learning in mean, while learning is incomplete with bounded beliefs. Under recursive sampling, we deduce monotone convergence of welfare, finding that it is faster with better quality signals. Moreover, sampling the predecessor is the best unit sampling mechanism. More strongly, when the recent past is not over-sampled, as when one learns from a uniformly-drawn random predecessor, almost sure convergence also obtains. Otherwise, realized and mean outcomes might unpredictably diverge.

The analysis of bounded beliefs remains a challenging and important open problem. Equally important and even tougher is the analysis of random sampling with larger sample sizes. We have long struggled on this point, since the Polya urn literature is less helpful. Path dependent social learning is an important direction for future economic analysis.

A LEARNING FROM SAMPLES: CAUTIONARY INSIGHTS

Our social learning model defeats some many common sense insights.

1. Learning from More Informed Individuals. Lemma 2 showed that in the signal quality model, it is better to sample from a more informed individual. *But learning from better informed individuals does not always yield a better signal.* Assume two actions and a posterior belief threshold $\bar{r} = 2/3$. Assume one observes an agent with posteriors $(.3, .7)$ having chances $(.5, .5)$. Then the action reveals the belief, and the observer gets the same belief distribution from action observation. Change the distribution by a mean preserving spread (MPS) to $(.3, .6, .8)$ with chances $(.5, .25, .25)$, reflecting a sufficiency improvement. Then the observer's belief is $(.4, .8)$ with chances $(.75, .25)$, i.e. not a MPS.

In special cases, it is better to sample from more informed individuals. For instance, assume the symmetric binary model, whose sampling chances $P_n^H(s), P_n^L(s)$ obey $P_n^H(s) = 1 - P_n^L(s)$. So each action occurs with ex ante probability 1/2, and the chance of taking action a in state L and action b in state H rises with a MPS in $P_n^H(s), P_n^L(s)$. A stronger belief in state L arises from action a, and a stronger belief in state H from action b.

Recall that with bounded beliefs and recursive sampling, the welfare $W_n = uR_n^H - R_n^L$

monotonically converges to a limit below u. We could not also conclude convergence of the action proportions R_n^θ, but this now follows from symmetry, since $R_n^H = 1 - R_n^L$. With sample size one, Proposition 2 (a) would then imply a.s. convergence of X_n in each state.

2. Social Learning with Larger Samples. *Welfare might fall when individuals learn from larger samples.* Using $F^H(p) = (5p-2)/(2p)$ and $F^L(p) = (5p-2)(p+2)/(8p^2)$ for beliefs $p \in (2/5, 2/3)$, computer simulations reveal that individuals after $n = 84$ are better off if everyone samples 49 instead of 50 predecessors.[11] Assume that everyone has been better off in the (sample size) 50-model so far. Sampling 50 actions in the 50-model can still be less informative than sampling 49 actions in the 49-model because the former actions are more strongly correlated, having relied less on their own private signals.[12]

3. Social Learning from Words or Polls and not Actions? The vehicle for transmission of social learning in the literature is action observation. Yet the imprimatur of rational social learning is learning from coarse signal of predecessors' beliefs.[13] Verbal discourse has this form (see Shiller (1995)). Survey data often works this way, as individuals indicate their strength of feeling on a 1–5 scale. How special is learning from actions?

Assume uniform sampling of one predecessor. The chance encounter yields one of *two posterior belief reports* ρ^H and ρ^L arise with chances $\psi(p)$ and $1 - \psi(p)$, respectively, if the sampled agent has posterior p. The *misperception function* $\psi : [0,1] \to [0,1]$ is weakly increasing. With action observations, ψ is a step function. Finally, posit both a symmetric signal $F^H(p) = 1 - F^L(1-p)$ and misperception function $\psi(p) = 1 - \psi(1-p)$.

Let $\phi^\theta(q)$ be the chance in state θ that someone who samples a predecessor who himself faced the social belief q will receive an encouraging report ρ_H. Then:

$$\phi^\theta(q) = \int_0^1 \psi\left(\frac{pq}{pq + (1-p)(1-q)}\right) dF^\theta(p). \tag{8}$$

By symmetry, the chance P_n^θ in state θ that agent $n+1$ sees report ρ_H obeys $P_n^H = 1 - P_n^L$.

If n observes ρ, he forms the sample belief $q_n(\rho)$, where $q_n(\rho^L) = 1 - P_n^H$ and $q_n(\rho^H) = P_n^H$. In state H, individual $n+1$ observes ρ^H with chance P_n^H, resulting in a chance

[11] We assume that all predecessors' (unordered) actions are observed by individuals $1, 2, \ldots, 51$. Computer simulations ran to 200 individuals. The welfare ordering may reverse again for later individuals.

[12] Vives (1993) explains a related phenomenon in a Gaussian social learning environment.

[13] While one might imagine that individuals observe information about the payoff realizations of predecessors (as in BF), we argue in §B.2 that it formally a standard private signal.

$\phi^H(q_n(\rho^H))$ that the report from individual n will be ρ^H. Then

$$P_{n+1}^H = \frac{n}{n+1} P_n^H + \frac{n}{n+1} (P_n^H \phi^H(P_n^H) + \left(1 - P_n^H\right) \phi^H(1 - P_n^H)). \tag{9}$$

For an example, assume the unbounded uniform quality example with $F^H(p) = p^2$. Assume the piecewise linear misperception function, based on a parameter $a \in [0, 1/2]$:

$$\psi(p) = \begin{cases} 0 & \text{for } p \leq a, \\ (p-a)/(1-2a) & \text{for } a < p < 1-a, \\ 1 & \text{for } p \geq 1-a. \end{cases}$$

For $a = 1/2$, this is the action observation step function, and so complete learning obtains. But for $a < 1/2$, the function ϕ^H only correctly conveys private beliefs near 0 and 1. For (9) has a fixed point $P^*(a) \in (1/2, 1)$, where $P^*(1/2) = 1$ and $P^*(a)$ strictly falls on $[1/2, 1]$. A little misperception for intermediate beliefs then suffices to wreck the complete learning outcome that arises with unbounded beliefs for action observations.

B COMPARISONS AMONG SOCIAL LEARNING MODELS

B.1 Social Learning In Networks

ADLO assume a network structure independently drawn at time zero, at which point individuals observe those in their neighborhood. While our samples our drawn as play progresses, from an individual's perspective, this timing difference is moot (by the assumed independence). Their network structure occasions a vast difference in theoretical analysis, but their final conclusion for complete learning is related. Their condition is that beliefs are unbounded and that the network has "expanding observations" — namely, in the limit, rarely does one observe just the first K people who acted. This is a weaker condition than our requirement of "not over-sampling the past": For we ask that in the limit almost no individual observes *any* of the first K people who acted. In other words, ADLO requires that *exclusively* sampling from a given finite set of agents occurs with a vanishing chance, whereas we ask that *any* sampling from a given finite set of agents occurs with a vanishing chance. It is instructive to see why a weaker condition suffices. As a network model, *they assume sampling with names preserved, but we assume unordered anonymous samples.* As noted earlier, one early individual sampled in an anonymous model can be poisonous to a Bayesian inference, since one might worry that he is the last to choose. But in a network setting, his identity is known, and he can be discounted as an early ill-informed choice. All told, the papers are very complementary.

B.2 Discrete vs. Continuum Agent Learning

BF pursue an elegant continuum agent approach to rational social learning by random sampling that eliminates aggregate randomness and renders moot strong convergence issues. They find complete learning (only correct herds) with sample size two or more.

They assume a continuum of privately informed individuals of mass 1, who each act. In an inessential difference, they endow everyone with a private signal by allowing payoff observations — a function of the state of the world alone, and not the action history. That learning is "word-of-mouth" is an inessential difference. Each period, a fraction $\alpha > 0$ is replaced by privately informed newcomers, who first sample from their predecessors.

In a discrete agent setting, social learning is doomed if it is based only on early private signals — for any finite number may well mislead. But with a continuum of agents, enough information is theoretically available at the outset, with the first continuum realization. Sure enough, BF's complete learning obtains even absent any new private information.[14]

We simplify BF's model to render an easier comparison with ours. Assume no deaths but a constant entry of mass 1 of agents each period. Assume a symmetric binary action ($u = 1$), binary signal (private beliefs $p > 1/2$ and $1 - p < 1/2$) and sample size two. Assume that individuals after period 2 can observe history. In this case, one mimics pure $\{a, a\}$ or $\{b, b\}$ samples, while one's private signal is decisive for the mixed sample $\{a, b\}$. With this simple decision rule, we may solve the model. First assume discrete agents. The initially pure histories will forever remain pure. The chances in state H of the unordered action histories $(n, 0), (n - 1, 1), \ldots (1, n - 1), (0, n)$ are then

- p^2, $2p^2(1 - p)$, $2p(1 - p)^2$, and $(1 - p)^2$ for $n = 3$,
- p^2, $2(1 + 2p)(1 - p)p^2/3$, $8p^2(1 - p)^2/3$, $2(3 - 2p)(1 - p)^2p/3$, and $(1 - p)^2$ for $n = 4$.

Taking expectations, the chances $\bar{q}_n$ in state H of a random sampled individual of the first n choosing action b are $\bar{q}_2 = p$, $\bar{q}_3 = p(2 - p)(1 + 2p)/3$, $\bar{q}_4 = p(1 + 3p - 2p^2)/2, \ldots$.

In the continuum agent model, we cannot speak of separate history realizations. Rather, only a single number $\hat{q}_n$ is relevant: namely, the fraction of individuals choosing action b

[14] One might profitably view this also as a model of information transmission rather than social learning, like the continuum agent model of Vives (1993). The finite agent analogue is a model where an initial incredibly large cohort decides at time 0 on the basis of private information alone, and everyone else tries to discern what these individuals knew.

in the population in period n. By assumption, $\hat{q}_2 = p$, while it is easy to see that

$$\hat{q}_n = \hat{q}_{n-1}(1 - 1/n) + (\hat{q}_{n-1}^2 + 2\hat{q}_{n-1}(1 - \hat{q}_{n-1})p)/n. \tag{10}$$

For a share $1 - 1/n$ of the population is old, while all new agents either chose b having seen $\{b, b\}$, or because they saw a mixed sample but their private signal was pivotal. For instance, $\hat{q}_3 = \bar{p}_3 = p(2 - p)(1 + 2p)/3$ by (10), but the models then diverge with $\hat{q}_4 \neq \bar{q}_4$.

This divergence occurs because individuals average over separate mutually exclusive realizations of the discrete setting. In period three, this does not matter, as no learning has yet taken place. But in period 4, those in the continuum setting *update as if all four mutually exclusive histories* $(3, 0), (2, 1), (1, 2), (0, 3)$ *have been realized, weighted by their chances.* So "path-simultaneity" replaces path-dependence in the continuum model. Averaging across disjoint stochastic outcomes of a discrete model as time passes is the essence of the continuum model, and why the complete learning conditions are different. Indeed, the solution to (10) satisfies $\hat{q}_n \to 1$ as $n \to \infty$; so there is complete learning in the continuum model. But obviously, with chance at least p^2 the fraction of individuals choosing action b is 1; so there is expected incomplete learning in the discrete agent model.

C OMITTED PROOFS

C.1 Complete Learning: Proof of Proposition 1 (*a*)

Proof of (*a*). Since anyone may ignore his sample and rely on his private signal, V_n has a lower bound strictly above -1. If the result fails, then some subsequence V_{n_k} tends to $v = \liminf V_n < u$. Let a sampled predecessor of individual n choose action b with chance R_n^θ in state $\theta = L, H$. Since eventually $uR_{n_k}^H - R_{n_k}^L = W_{n_k} \leq V_{n_k} < (u + v)/2 < u$, the chance of all correct sampled actions $R_{n_k}^H(1 - R_{n_k}^L)$ is bounded away from one. By Lemma 2, there exists $\eta > 0$ such that eventually $V_{n_k} > W_{n_k} + \eta$. Choose $\delta > 0$ smaller than $\eta/3$. Now there exists N with $V_n > v - \delta$ when $n > N$ (by definition of v), and $|V_{n_k} - v| < \delta$ when $n_k > N$, as v is the limit of (V_{n_k}). Let $3\varepsilon = \eta/(1+u)$. Since Σ does not over-sample the past, there exists an $M > N$ so large that any $n > M$ samples $m \leq N$ with chance less than ε. Thus, the expected welfare of any $n_k > M$ obeys $V_{n_k} > W_{n_k} + \eta \geq \varepsilon(-1) + (1 - \varepsilon)(v - \delta) + \eta = v - \varepsilon(1 + v - \delta) - \delta + \eta > v - \eta/3 - \eta/3 + \eta > v + \delta$. Since the past is not over-sampled, if V_n converges to u, then so does the running average W_n. □

C.2 Almost Sure Convergence: Proof of Proposition 2

We prove that a.s. convergence of the cumulative forecast error process (μ_n) implies the same of the deviations process $(R_n^H - X_n)$. So if the beliefs are unbounded, then $R_n^H \to 1$ by Proposition 1. Since R_n^H is the state H mean of X_n, and $X_n \leq 1$, we have $X_n \to 1$ a.s.

Fix a realization with convergent accumulated forecast errors, and study the sequences of real numbers (X_n) and (ϵ_n^H). Subtracting (5) and (6):

$$X_{n+1} - R_{n+1}^H = X_n - R_n^H + \pi_n[\chi_n^H(X_n) - \chi_n^H(R_n^H) + R_n^H - X_n + \epsilon_n^H] \qquad (11)$$

Since χ_n^H is an increasing contraction, we can merge factors in $X_n - R_n^H$ into the first term, and get $(X_{n+1} - R_{n+1}^H)^2 \leq (X_n - R_n^H)^2 + 4\pi_n|\epsilon_n^H||X_n - R_n^H| + \pi_n^2|\epsilon_n^H|^2$. Simplify this using $|X_n - R_n^H| \leq 1$ and $|\epsilon_n^H| = |i_n - \chi_n^H| \leq 2$, and thereby deduce the inequality for squares:

$$(X_{n+1} - R_{n+1}^H)^2 \leq (X_n - R_n^H)^2 + 4\pi_n\epsilon_n^H + 4\pi_n^2 \qquad (12)$$

Now, equations (5) and (6) yields $|X_{n+1} - X_n| \leq \pi_n$ and also $|R_{n+1}^H - R_n^H| \leq \pi_n$, by the contraction character of χ_n^H. Since $\sum_{n=1}^{\infty} \pi_n^2 < \infty$, we have $\pi_n \to 0$ as $n \to \infty$, and so the step sizes of X_n and R_n^H vanish. Hence, one of the following two alternatives holds:

(*i*) there exists N such that $X_n > R_n^H$ for all $n > N$, or $X_n < R_n^H$ for all $n > N$, *or*

(*ii*) there exists an infinite subsequence (n_k) such that $R_{n_k}^H - X_{n_k} \to 0$ as $n_k \to \infty$.

Consider first alternative (*i*). Note that $|\chi_n^H(X_n) - \chi_n^H(R_n^H)| \leq |X_n - R_n^H|$, as χ_n^H is an contraction. Then $\chi(X_n) - \chi(R_n^H) \leq X_n - R_n^H$ for $n > N$, since χ is increasing. Thus, (11) implies

$$X_{n+1} - R_{n+1}^H \leq X_n - R_n^H + \pi_n\epsilon_n \qquad (13)$$

Let $\bar{\omega} = \liminf_n(X_n - R_n^H)$. We will prove that $X_n - R_n^H \to \bar{\omega}$. Fix $\xi > 0$, and choose $N' > N$ such that our cumulative forecast error process (μ_n) obeys $|\mu_n - \mu_m| < \xi/2$, for all $n, m > N'$. Choose $N'' > N'$ such that $X_{N''} - R_{N''}^H < \bar{\omega} + \xi/2$. Iterate (13) to give:

$$X_{N''+k} - R_{N''+k}^H \leq X_{N''} - R_{N''}^H + \sum_{n=N''}^{N''+k-1} \pi_n\epsilon_n = X_{N''} - R_{N''}^H + \mu_{N''} - \mu_{N''+k} \leq \bar{\omega} + \xi.$$

for all $k > 0$. Thus, $\limsup_n X_n - R_n^H \leq \liminf X_n - R_n^H$, and hence the limit exists.

Next, consider alternative (*ii*). We prove that $X_n - R_n^H$ converges to zero. Fix $\xi > 0$. Choose $N' > N$ such that $\sum_{k=N'}^{\infty} \pi_k^2 < \xi/6$ and $|\mu_n - \mu_m| < \xi/12$ for all $n, m > N'$. Pick

$N'' > N'$ with $(X_{N''} - R^H_{N''})^2 < \xi/3$. We iterate inequality (12) to deduce for all $k > 0$,

$$(X_{N''+k} - R^H_{N''+k})^2 \leq (X_{N''} - R^H_{N''})^2 + 2\sum_{n=N''}^{N''+k-1} \pi_n^2 + 4\sum_{n=N''}^{N''+k-1} \pi_n \epsilon_n \leq \xi. \qquad \square$$

References

Acemoglu, D., M. Dahleh, I. Lobel, and A. Ozdaglar (2008): "Bayesian Learning in Social Networks," MIT mimeo.

Arthur, W. B., Y. M. Ermoliev, and Y. M. Kaniovski (1986): "Strong Laws For a Class of Path-Dependent Stochastic Processes with Applications," in *Stochastic Optimization: Proceedings of the International Conference, Kiev, 1984*, ed. by I. Arkin, A. Shiraev, and R. Wets. Springer-Verlag, New York.

Arthur, W. B., and D. Lane (1994): "Information Contagion," in *Increasing Returns and Path Dependence in the Economy*, ed. by W. B. Arthur, pp. 69–97. The University of Michigan Press, Ann Arbor.

Banerjee, A., and D. Fudenberg (2004): "Word-of-Mouth Learning," *Games and Economic Behavior*, 46, 1–22.

Banerjee, A. V. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797–817.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy*, 100, 992–1026.

Breiman, L. (1968): *Probability.* Addison-Wesley, Reading, Mass.

Celen, B., and S. Kariv (2004): "Observational learning under imperfect information," *Games and Economic Behavior*, 47, 72–86.

Chamley, C. P. (2004): *Rational Herds: Economic Models of Social Learning*. Cambridge University Press, New York.

Eggenberger, F., and G. Polya (1923): "Uber die Statistik verketter Vorginge," *Zeitschrift fur Angewandte Mathematik und Mechanik*, 3, 279–89.

Freedman, D. A. (1965): "Bernard Friedman's Urn," *Annals of Mathematical Statistics*, 36, 956–970.

Gale, D., and S. Kariv (2003): "Bayesian Learning in Social Networks," *Games and Economic Behavior,*, 45, 329–46.

Mahmoud, H. M. (2009): *Polya Urn Models*. CRC Press, Boca Raton, FL.

Monzon, I., and M. Rapp (2011): "Observational Learning with Position Uncertainty," Collegio Carlo Alberto Working Paper.

Sgroi, D. (2002): "Optimizing Information in the Herd: Guinea Pigs, Profits, and Welfare," *Games and Economic Behavior*, 39, 137–166.

Shiller, R. J. (1995): "Conversation, Information, and Herd Behavior," *American Economic Review*, 85, 181–185.

Smith, L. (1991): "Error Persistence, and Experiential versus Observational Learning," Foerder Series working paper.

Smith, L., and P. Sørensen (1996): "Rational social learning by random sampling," MIT mimeo.

——— (2000): "Pathological Outcomes of Observational Learning," *Econometrica*, 68, 371–398.

Sørensen, P. (1996): "Essays on Informational Herding," Unpublished MIT PhD Thesis.

Surowiecki, J. (2004): *The Wisdom of Crowds*. Random House, New York.

Vives, X. (1993): "How Fast do Rational Agents Learn?," *Review of Economic Studies*, 60, 329–347.