Social Learning with Costly Search[*]

Manuel Mueller-Frank[†]   Mallesh M. Pai[‡]

This Version: December 12, 2013

First Version: October 26, 2012.

Abstract

We study a model where rational agents act sequentially in a predetermined order, observing the actions chosen by their predecessors and then selecting an action. Agents engage in costly search to learn about the quality of various actions. Search costs of agents are private to them, and are independently and identically distributed across agents. We show that asymptotic learning, i.e, that late moving agents always select the optimal action, occurs if and only if search costs are not bounded away from zero. We explicitly characterize "common search order equilibria" in which agents choose to search in the same order.

Keywords: social learning, search with recall, herding, informational cascades, asymptotic learning

JEL Classification Numbers: D80, D83

1   INTRODUCTION

Social learning studies settings where agents make a choice while facing payoff relevant uncertainty. In addition to private information about this state of the world, they may gather information from choices made by others. The literature studies whether agents' behavior effectively aggregates the information available to individual agents, or not. In other words, will prior agents' actions be informative enough that agents "learn," or will they "herd" on some suboptimal action. The results in the literature hinge on whether agents may see signals that can overturn a herd, i.e. induce the agent to take a different action than a previous long herd. In an environment with costly information acquisition, agents' incentives to acquire sufficient information for asymptotic learning are unclear. In a putative equilibrium where learning does occur, agents have weaker incentives to acquire information, and may instead free-ride on others, intuitively causing herding. Conversely, if learning may not occur in equilibrium, agents have stronger incentives to acquire information, and thus learning will occur.[1]

This paper introduces an alternative framework for studying social learning. In our model, agents have no private information *a priori*. Agents view the actions taken by their predecessors, and may then conduct costly search among possible actions. We study the possibility of asymptotic learning in our model, i.e., do late movers always take the best action, or may "bad herds" form on suboptimal actions? Our main result is a characterization of asymptotic learning: Asymptotic learning occurs if and only if the distribution of agents' search costs includes zero in its support.

While our main focus in this paper is theoretical, we believe our model is amenable to applications. Several choice situations of economic interest may be well modeled by our choice of information acquisition technology (i.e., costly search), notably several durable consumer goods (e.g. cell phones, cars, etc.). In these cases, a consumer observes choices made by his predecessors—for example, he sees what brands of cars people drive, what kind of phones they carry etc. He may then choose to acquire more information, e.g. take the cars on a test drive, read reviews of specific phones online, and so on. Sampling an alternative (test driving a car, reading a review of a phone) is costly due to the time and effort involved. Further, while it reveals information about the quality of that alternative, it does not directly reveal anything about the quality of other alternatives. After sampling some alternatives (possibly none), the agent stops sampling (i.e., decides he has test-driven enough cars/ read enough reviews), and makes a choice among these alternatives.[2]

---

[1]This is analogous to Grossman and Stiglitz (1976) who study similarly paradoxical incentives to acquire information in a general equilibrium setting when prices may reveal information.

[2]In the taxonomy of search, this is therefore "search with recall."

1.1 Discussion of Model & Result

More precisely, this paper considers a countable set of fully rational agents. In a predetermined sequence, agents take an irreversible action out of a finite set. These actions are a priori identical, and their associated utilities unknown. There are no payoff externalities. Agents observe the history of actions taken by their predecessors and update their beliefs about the respective utilities of each action. Each agent engages in private, costly search before taking his action, taking into account past history, i.e., the actions taken by predecessors. Searching or "sampling" an action reveals the utility of that action perfectly to the agent, but comes at a cost. Each agent has a constant cost per search. Search costs are assumed to be independent and identically distributed across agents. Neither the sampling decisions nor search cost of an agent are observed by subsequent agents.

Our focus lies on the learning properties of equilibria in this setting, that is, whether asymptotic learning obtains in equilibrium. Asymptotic learning requires that the probability of the nth agent selecting the highest utility action converges to one as n goes to infinity, i.e., that late moving agents "learn" the best action or, alternately, do not "herd" on a suboptimal action.[3] We show that asymptotic learning occurs if and only if the distribution of search costs includes zero in its support.[4] That is, in any equilibrium, the optimal action is chosen in the long run with probability one if and only if the support of the distribution of search costs contains arbitrarily small search costs.

We would argue that this result is 'surprising,' in particular that a condition on the support of the distribution of costs is both necessary and sufficient to guarantee learning in equilibrium. In our model, an agent endogenously chooses how much to search given the observed history. In equilibrium, even an agent with low search costs may choose to "free-ride" on his predecessors, not search and therefore be uninformed, thereby propagating a herd. In other words the relationship between an agent's search costs and how well informed he is an endogenous property of the equilibrium. Given this, one may have intuited that sufficient conditions for learning depend more delicately on primitives e.g. the shape of the distribution of search costs, and the actual probability of low search costs. We show that this is not the case. Our result is therefore qualitatively different from the results in the literature on optimal experimentation. There, whether experimentation concludes at the efficient outcome may depend on the shape of the cost function—for example, see Kihlstrom, Mirman, and Postlewaite (1984).[5]

---

[3]Probabilities here are with respect to draws of search costs of each agent, and the utility of each action.

[4]We exclude environments where the first agent always finds an optimal action.

[5]See Aghion, Bolton, Harris, and Jullien (1991) for a characterization of when experimentation converges to the optimal action. Smith and Sorensen (2006) formally connect the social learning and experimentation environments.

It is analytically difficult to explicitly characterize all equilibria of our model. With a view to possible applications, we provide a closed form characterization of a particular subset of these equilibria, which we call *common search order equilibria*. In these equilibria, all agents follow strategies which search actions in a commonly known order, i.e. if an agent sampled action a_3, he must have sampled actions a_1 and a_2. Note that all actions are still ex-ante identical, this common search order is a convention all agents employ. In other words, this is a refinement of the set of all equilibria of our model, not an exogenous constraint on the order in which agents can search. Such equilibria may be of independent interest. There are several settings of applied interests where agents select actions from an ordered list of suggestions, for example listings in the Yellow Pages (alphabetically ordered), competing products down a supermarket aisle, etc. In a recent paper, Athey and Ellison (2011) study a model of Internet advertising where consumers search among a list of advertisements to find a desired product. Additionally, search from lists has attracted attention in single agent decision settings—see, e.g., Rubinstein and Salant (2006) and Meredith and Salant (2007).

To understand the intuition behind our result consider the simplest case where there are only two possible actions, a_1 and a_2. Note that asymptotic learning occurs if and only if at least one agent samples each action. Suppose agent 1 takes action a_1. For agent 2, this is "good news" for action a_1 and "bad news" for action a_2. Agent 2 will now sample action a_1 first, and only sample action a_2 if his cost of search is less than a threshold c_2. This threshold depends on:

i. The observed quality of action a_1: If action 1 has a high quality, then there is little value to searching.

ii. The probability that agent 2 assigns to agent 1 having sampled action 2: If agent 1 sampled action 2 with high probability, then agent 1's choice of action 1 suggests that action 2 is of lower quality.

Note that this threshold of searching further c_2 is less than the corresponding search threshold for agent 1, c_1— agent 2's beliefs are more pessimistic about action 2. Suppose that the first $n-1$ agents have all selected action 1, and consider the problem faced by the nth agent. We show that there is a cutoff cost c_n such that the nth agents samples action 2 if and only if his cost is below this cutoff. Unsurprisingly this cutoff is decreasing in n. Denote the ex-ante probability (with respect to his search costs) of no search by p_n. Note that for asymptotic learning to occur, with probability one, some agent must sample action a_2. Therefore asymptotic learning occurs if and only if the infinite product of p_n's vanishes, i.e. it hinges on the speed of convergence of p_n to one. Surprisingly we show that this occurs as long as zero is in the support of distribution of search costs.

At a high level the idea is this—as long some agents arrive with an arbitrary small

marginal cost for information acquisition, the marginal benefit of information acquisition must converge to zero in the long run. This idea implies long run optimality and is "detail-free."

1.2 COMPARISON TO THE SSLM

In the standard sequential social learning model (henceforth SSLM), rational agents make a choice in a predetermined sequence. Each agent receives a private signal about the unknown state of the world and observes the choices of his predecessors before making his choice. The seminal papers of Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) introduced the SSLM. They showed that agents may "herd" on a suboptimal option, i.e., choose to ignore their signal and follow their predecessors. This is useful in modeling phenomena such as fads, fashions, booms and crashes. The SSLM has thus been widely applied.[6]

In our model, information is endogenously acquired rather than exogenously given to agents. Additionally, information in our model is qualitatively different than that in the SSLM. In our model, sampling an action only reveals the quality of that action, while private information in the SSLM concerns the relative value of all actions.[7] At a technical level, the version of the SSLM that is tractable for applications is a binary-state model, and hence ordinal in nature: either action a is better than action b or vice-versa. However, the better action may be tremendously superior or only marginally so. The SSLM cannot distinguish these two scenarios due to the ordinal nature of the state space. In contrast, our model imposes no restrictions on the cardinality of the state space, and therefore can model situations where both the ordinal ranking of and the cardinal difference between various actions are uncertain. This is potentially useful for understanding the welfare loss from suboptimal herds.

Despite substantial differences in the underlying information structure, our model shares several similarities with the SSLM. First, the characterization of asymptotic learning is conceptually similar to the one in the standard framework. Smith and Sørensen (2000) show that in the SSLM asymptotic learning occurs if and only if private signals induce unbounded beliefs. That is, if the support of the private probability of one of the two states generated by the signals contains 0 and 1. Therefore, the private signals need to be able to lead to posterior beliefs arbitrarily close to zero and one. In our framework the arbitrarily low search costs take the role of unbounded private signals. Second, the most interesting features of

---

[6]For example, to trading in financial markets (Avery and Zemsky, 1998; Park and Sabourian, 2011), pricing (Bose, Orosel, Ottaviani, and Vesterlund, 2006, 2008; Mueller-Frank, 2012), and voting (Dekel and Piccione, 2000; Ali and Kartik, 2011).

[7]To continue the analogy to 'real' world, the information structure of the SSLM can be thought of as agents learning from, e.g., comparison websites.

equilibrium behavior in the SSLM, i.e., local conformity, fragility of mass behavior and the occurrence of incorrect informational cascades, carry forward to our setting. As should be clear from the intuition of our result, fragility of mass behavior takes the following form in our model: Once an agent deviates from the action of his predecessor, no later agent will select, or even sample the predecessor's action. Intuitively, if agent 1 selects action a_1, and agent 2 selects a_2, then action a_2 has been "revealed preferred" to a_1 from the point of view of future agents.

From the point of view of applications, assumptions on search costs might be easier to test from observed data than assumptions on precision of private signals. A large applied literature on search has discussed the estimation of search models, and identifying population search costs from macro-level and micro-level data.[8] In principle therefore, an applied social search model could use search costs estimated from the data. By contrast, an application of the standard model would require an assumption on the distribution of signal qualities that is difficult to verify empirically. We discuss applied work on the subject in further detail in Section 4.

1.3 RELATED LITERATURE

There are few papers we are aware of that consider information acquisition by agents rather than given private signals.

The paper closest to ours in terms of the model is Hendricks, Sorensen, and Wiseman (2012). They consider the case of two actions and two states, where one action has a known payoff (normalized to 0), while the other can have either a high or a low payoff. They study a setting where agents can choose to pay a cost to learn the value of the unknown action. Unlike the present work, agents there have heterogeneous preferences, but identical search costs. These assumptions prevent direct comparisons with our results, or with the SSLM.

Another paper that is similar to ours is the more recent Ali (2013). He studies the SSLM where agents can endogenously choose more precise signals at a cost. For the case of discrete actions, he finds a result similar to ours, i.e., asymptotic learning occurs if and only if agents may be able to procure arbitrarily informative signals for an arbitrarily low cost.

An earlier paper by Burguet and Vives (2000) considers a setting where homogeneous agents take an action to match the (unknown) state of nature, with quadratic loss. Agents observe a noisy signal about the state, and can choose the level of precision of their signal— higher precision is more costly. In this setting, they derive results similar to ours. However, their paper involves a distinct continuum of agents acting at each time period. Further, the specific assumptions they make about information structure imply that the level of precision

---

[8]The literature is too large to comprehensively cite here. For a classic example, see Wolpin (1987).

chosen by agents at any step k depends only on the primitives of the model and k, i.e. it is independent of the choice history they observe. As a result, there is no natural analog to a "herd" in this setting.

Kultti and Miettinen (2006) study a variant of the standard model where agents choose which past agents' actions to observe, and must pay a commonly known cost to observe the action taken by a past agent. They show that if this cost is small, herding may arise deterministically, while a cost of exactly zero recovers the standard model.

Additionally, there have been some attempts to relax the binary state space assumption in the SSLM. Smith and Sørensen (2000) discuss an extension to arbitrary finite state spaces, but at a significant analytical cost. Park and Sabourian (2011) consider a model with 3 states. Arieli and Mueller-Frank (2013) consider the general case of compact metric state and action spaces. They show that if the action set is sufficiently "dense" then information is perfectly aggregated along the equilibrium path for generic continuous utility functions.

2   THE MODEL

In this section we present the formal model analyzed in this paper. A discussion of some of the major assumptions is deferred to Section 4, after we present our main result.

2.1   AGENTS AND ACTIONS

A countably infinite set of agents $\{1, 2, \ldots\}$ sequentially select an action, with agent i acting in period i. The set of actions X available to each agent i is finite and identical across agents. Let k denote the cardinality of the set of actions. A typical element of the action set X will be denoted by x, while the action selected by agent i is denoted by a_i. We denote $h_i = (a_1, ..., a_{i-1})$ as the history of actions of agents preceding agent i. History h_i is common knowledge among all agents $i' \geq i$.

Actions are differentiated in their qualities, but are ex-ante homogeneous.[9] We will denote q_x as the 'quality' associated with action x. These are i.i.d. draws according to probability measure P_Q over $Q \subseteq \mathbb{R}_+$.

The state of the world is the realized quality of each of the k actions $\omega = (q_1, q_2, ..., q_k)$. The state space is then given by $\Omega \equiv Q^k$ with product measure $P_\Omega = (P_Q)^k$. Note that this formulation captures finite, countably infinite and uncountable state spaces.

---

[9]If the qualities of actions are statistically independent but not identically distributed, our asymptotic learning result still obtains.

2.2 SEARCH

Each agent's action choice is based on costly sequential search. After observing the history h_i, agent i decides which action $s_i^1 \in X$ to sample first . Sampling an action perfectly reveals its quality to the agent. We will denote the observed utility of the sampled action as $q_{s_i^1}$. After observing the quality $q_{s_i^1}$ of the sampled action s_i^1, agent i decides whether to continue to search, $s_i^2 \in X$, or discontinue, $s_i^2 = n$,

$$s_i^2 \in X \cup \{n\}.$$

After sampling m actions, agent i selects s_i^{m+1} where

$$s_i^{m+1} \in X \cup \{n\}.$$

Let S_i denote the set of actions agent i samples, $S_i \subseteq X$.

After finishing sampling, the agent chooses an action a_i. We assume that agents can only select an action they sampled, i.e. $a_i \in S_i$. We believe this assumption is a good fit for the settings we model. See Section 4 for further discussion of this assumption.

For simplicity, the first action is sampled at no cost while sampling each other action involves a cost of $c_i \in \mathbb{R}_+$.[10] The search costs c_i are i.i.d. draws from a commonly known probability measure P_C. While search costs are identically distributed across agents, agent i's search cost c_i is privately known to agent i.

The net utility of agent i is therefore the quality of the action he selects minus his search costs,

$$U_i(S_i, a_i, c_i, \omega) = q_{a_i} - c_i \left(|S_i| - 1 \right).$$

For every agent i we distinguish $k + 1$ different information sets. The first stage information set I_i^0 corresponds to i's information set based on the observed history of choices prior to sampling any action. The set I_i^m is the information set agent i has after sampling m actions. The information set I_i^a corresponds to the information set of agent i once his search ends. We have

$$\begin{aligned}
I_i^1 &= (c_i, h_i), \\
I_i^m &= (c_i, h_i, q_{s_i^1}, ..., q_{s_i^m}), \quad m = 2, \dots, k, \\
I_i^a &= (c_i, h_i, (q_x : x \in S_i)).
\end{aligned}$$

---

[10]It is equivalent if all searches cost the same amount c_i, but each agent has to take an action, i.e., he cannot abstain, and therefore must conduct at least one search.

The set of all possible search stage information sets of agent i are denoted as $\mathcal{I}_i^m$, for $m = 0, 1, ..., k$, and the set of action stage information sets as $\mathcal{I}_i^a$. A strategy σ_i of agent i is therefore

$$\sigma_i^0 : \mathcal{I}_i^0 \rightarrow \Delta(X),$$
$$\sigma_i^m : \mathcal{I}_i^m \rightarrow \Delta(X \cup \{n\}),$$
$$\sigma_i^a : \mathcal{I}_i^a \rightarrow \Delta(S_i).$$

Once an agent i stops his search at stage m, then he does not sample further. Strategies $\sigma_i^{m'}$ for $m' > m$ are irrelevant, and he then chooses an action according to σ_i^a. We do not state this constraint formally to avoid burdensome notation.

A strategy profile σ is a sequence of strategies $\{\sigma_i\}_{i \in \mathbb{N}}$. Given a strategy profile σ and probability measures P_Ω and P_C, the sequence of actions $\{a_i\}_{i \in \mathbb{N}}$ is a stochastic process. We denote the resulting probability measure on sequences of actions by P_σ.

3   ASYMPTOTIC LEARNING UNDER SOCIAL SEARCH

The main objective of the paper is to characterize conditions under which agents asymptotically select the correct action. Formally:

DEFINITION 1. *Let σ be a strategy profile with resulting probability measure over actions P_σ. We say asymptotic learning occurs if*

$$\lim_{i \to \infty} P_\sigma(a_i \in \arg\max_{x \in X} q_x) = 1.$$

In words, asymptotic learning occurs if the probability of agent i selecting the best action converges to one as i goes to infinity. Our characterization of learning will hinge on whether the distribution of search costs among agents has 0 in its support.

DEFINITION 2. *We say that search costs are* bounded away from zero *if 0 does not lie in the support of P_C, i.e. there exists $\epsilon > 0$ such that $P_C([0, \epsilon]) = 0$.*

Conversely, search costs are not bounded away from zero if for every $\epsilon > 0$, $P_C([0, \epsilon]) > 0$. In words, search costs are not bounded away from zero if there is a positive probability of arbitrarily low search costs. Finally we assume away "trivial" search environments.

ASSUMPTION 1 (NON-TRIVIAL SEARCH ENVIRONMENT). There exists $\hat{q}$ in the support of P_Q such that:

9

1. $1 - F_Q(\hat{q}) > 0$.

2. The distribution of search costs is such that with positive probability agent 1 does not sample another action when the first action sampled has quality $\hat{q}$ or higher.

In an environment where this fails, the first agent samples actions until he learns the optimal action. The subsequent agents just follow him. This will trivially yield asymptotic learning. Since such environments are uninteresting, Assumption 1 rules them out.

We now characterize asymptotic learning in terms of the support of the distribution of search costs.

THEOREM 1. *In any Perfect Bayesian equilibrium of a non-trivial search environment, asymptotic learning occurs if and only if search costs are not bounded away from zero.*

The theorem states two results. First, the probability of agent i selecting the correct action converges to one if the search costs can be arbitrarily small. Second, if search costs are bounded away from zero, then agents herd on a suboptimal action with positive probability. The theorem closely relates to Smith and Sorensen's (2000) characterization of asymptotic learning in the standard sequential social learning setting. They show that asymptotic learning occurs if private signals are unbounded and fails if private signals are bounded. Our theorem can be interpreted in several ways. First, it exhibits the possibility of inefficient herding with a different underlying informational structure. Second, the theorem implies that the exogenous private signals of the standard framework can be micro-founded. Due to the conceptual similarity of the characterization of asymptotic learning across frameworks, by endogenizing private information, we are providing a micro-foundation of private signals in a sequential social learning environment.

In what follows, we work our way through the proof of the main theorem. Here, we provide a short high level overview to orient the reader. In Observation 1 we describe the search behavior of Agent 1. This is well known from the literature on search with recall. Lemma 1 then shows that every subsequent agent first searches the action taken by his direct predecessor. We then restrict attention to the common search order equilibria described earlier. Lemma 2 fully characterizes these equilibria by describing the search decisions of agents as a function of observed history. We then use this characterization to prove the asymptotic learning result for common search order equilibria. We conclude by providing the intuition for how the asymptotic learning result extends to all equilibria of this model—the formal proof is in the appendix.

3.1 CHARACTERIZING EQUILIBRIA

AGENT 1 Let us begin by considering the first agent. As the utilities of all actions are identically distributed, he is indifferent in regards to which action to sample first. Suppose the action he samples first has utility q. He decides to sample a second action, only if his search cost c_0 is smaller than the expected value of sampling a second good,

$$c_0(q) = \int_{\tilde{q} \geq q} (\tilde{q} - q) dP_Q(\tilde{q}).$$

We can proceed inductively to define his search decisions.

OBSERVATION 1. Suppose the agent has sampled $m < k$ actions, and let $q^\star$ denote the highest utility among the sampled actions. Agent 1 decides to sample an $m + 1^{\text{th}}$ action if and only if his search cost c_1 is smaller or equal to a cutoff search cost $c_0(q^\star)$, where $c_0(\cdot)$ is given by:

$$c_0(q) = \int_{\tilde{q} \geq q} (\tilde{q} - q) dP_Q(\tilde{q}). \tag{Cutoff}$$

It is important to note that the search decision of the agent is "stationary," i.e. it does not depend on the number of unsampled actions remaining.

SUBSEQUENT AGENTS Observation 1 describes the optimal strategy for agent 1. Let us now consider agent 2's problem, given that agent 1 has picked some action x. From his perspective, there are two possibilities:

1. Agent 1 did not sample any other actions than x— in this case agent 1's decision is uninformative as to the utilities of the other actions.

2. Agent 1 did sample other actions— in this case agent 1 selected the best action (x) among the ones he sampled.

Therefore, agent 2's posterior beliefs must be that the action chosen by agent 1, x, is superior to (first order stochastically dominates) all other actions, and therefore agent 2 should sample action x first.

A simple inductive argument establishes that each agent i should sample action x first as long as all his predecessors selected that same action. Suppose now that everyone from 1 through $i - 2$ select action x, while agent $i - 1$ selected $x' \neq x$. By our previous argument, agent $i - 1$ must have sampled action x first, and therefore would only select action x' if it has a higher utility. Agent i onwards infer this, and therefore should sample good x' first, and no longer sample action x at all. Formally,

DEFINITION 3. *We say an action x is* discarded *on history h_i if:*

1. *$a_j = x$ for some $j < i$ and,*
2. *$a_{j'} \neq x$ for some $j < j' < i$.*

Lemma 1 states that each agent i first samples the action chosen by his predecessor $i-1$ as long as it is not a discarded action. This is intuitive since this action is the 'best' action among all those agent $i-1$ searched (see also Weitzman, 1979).

LEMMA 1. *Let σ be a Perfect Bayesian equilibrium of the game. Then for every agent $i > 1$ with cost c_i and history $h_i = (a_1, a_2, \ldots, a_{i-1})$,*

$$\sigma_i^1(c_i, h_i) = a_{i-1}$$

unless a_{i-1} is a discarded action. In that case the agent first samples the most recent non-discarded action a_{i-l}, i.e.,

$$\sigma_i^1(c_i, h_i) = a_{i-l},$$

where all actions a_{i-l+1} to a_{i-1} are discarded.

Before presenting the next lemma, note that for asymptotic learning it is necessary and sufficient that either each action is sampled by at least one agent with probability 1, or an action with the highest possible utility in the support of P_Q, if one exists, is sampled. Lemma 2 characterizes the stopping decision, i.e., whether or not to sample another action, for agents $i > 1$. We will then use this characterization to argue that every action is sampled.

Similar to agent 1's cost cutoff $c_0(q^*)$ being a function of the highest realized utility among the already sampled actions, each agent $i > 1$ faces a cost-cutoff, which determines his stopping condition. Implicitly, this depends on the history h_i that agent i observes.

EQUILIBRIA WITH A COMMON SEARCH ORDER We first focus on a specific subset of equilibria. We assume that whenever any agent's posterior beliefs are such that the expected utilities of a set of actions are the same, then she samples among those actions based on a common order over actions. In other words, numbering the actions from x_1 to x_k, if the expected utility of actions x_g and x_l, $g, l \in \{1, ..., k\}$, are identical then an agent samples x_g before x_l if and only if $g < l$.

There are many other equilibria in this model, which differ in the order in which agents choose to sample actions over which they are indifferent. We restrict attention here to common search order equilibria for analytical tractability in terms of agents' beliefs. In particular, if agent i selects action x_l, then:

1. It is commonly understood that actions x_1 to x_{l-1} are of lower quality than x_l, and therefore are not sampled or selected by any subsequent agent.

2. Subsequent agents $i' > i$ have symmetric posterior beliefs on all goods x_{l+1} to x_k.

This makes the argument more transparent. We will now characterize the stopping decision for agents $i \geq 2$ in equilibria with a common search order, and argue that Theorem 1 holds for equilibria with a common search order.

STOPPING DECISION   The following definition will be helpful. Given $c_0(\cdot)$ as defined in (Cutoff), inductively define $c_n(\cdot)$ for $n > 1$ as:

$$c_n(q) \equiv c_0(q) \prod_{\eta=0}^{n-1} P_C(\tilde{c} > c_\eta(q)), \tag{Cutoff-n}$$

We are now in a position to describe the subsequent search decision of agents $i > 1$.

LEMMA 2. *Let σ be a Perfect Bayesian equilibrium of the game that adheres to a common search order. Consider agent $i > 1$, who observes history $h_i = (a_1, \ldots, a_{i-1})$. Let $a_{i-1} = x_l$ be a non-discarded action, and let $n(h_i)$ be the number of predecessors of i who selected action x_l. The search decision of agent i is given by:*

$$\sigma_i^2(I_i^2) = \begin{cases} x_{l+1} & \text{if } c_i \leq c_{n(h_i)}(q_{x_l}), \\ n & \text{otherwise.} \end{cases} \tag{1}$$

If the agent chooses to sample further, for $m \geq 2$:

$$\sigma_i^{m+1}(I_i^{m+1}) = \begin{cases} x_{l+m} & \text{if } q_{x_l} = q^\star \text{ and } c_i \leq c_{n(h_i)}(q_{x_l}), \\ x_{l+m} & \text{if } q_{x_l} \neq q^\star \text{ and } c_i \leq c_0(q^\star), \\ n & \text{otherwise.} \end{cases} \tag{2}$$

Here c_n is as defined in (Cutoff-n), and $q^\star$ is the highest observed quality by an agent among the products he has searched i.e. $q^\star = \max(x_l, \ldots, x_{l+m-1})$.

Let us discuss this in words. We already know from Lemma 1 that each agent i first samples the action selected by his predecessor if it was non-discarded, which we denote x_l. Intuitively, the fact that agent $i - 1$ selected this action makes it "better" in terms of the agent i's posterior beliefs, and he is subsequently less likely to sample further. Lemma 2 shows that agent i's cutoff cost to sample further is given by $c_{n(h_i)}(q_{x_l})$ (Cutoff-n), and depends only on the number of previous agents $n(h_i)$ who have taken the action x_l, and the

quality of that action q_{x_l}. Intuitively, this is because the common search order implies that these $n(h_i)$ agents are the only agents who could have searched beyond x_l.

If the agent's cost c_i is less than $c_{n(h_i)}(q_{x_l})$ then he continues searching until he either finds a better action or runs out actions. This is expressed formally by (1) and the first case of (2). The latter two cases of (2) describe the continuing search decision after the agent samples an action $l' > l$ with $q_{x_{l'}} > q_{x_l}$. Because of our assumption of a common search order, agent i' infers that no agent before him has sampled any action with index at least l'. Agent i's posterior beliefs about the qualities of actions with index strictly larger than l' are therefore identical to the prior P_Q. Therefore his subsequent search decisions from this point forward are the same as agent 1's, and are as given in Observation 1.

EQUILIBRIUM STRATEGIES   Collecting everything we have argued so far, equilibrium strategies with a common search order are:

1. AGENT 1. He first samples good x_1, and then continues sampling until the marginal value of searching further is exceeded by his search cost

$$c_1 > c_0(\max\{q_{x_1}, q_{x_2}, \ldots, q_{x_l}\}),$$

    where $c_0(\cdot)$ is given in (Cutoff). He then takes an action with the highest utility among those he searched.

2. AGENT i. He observes history $h_i = (a_1, a_2, \ldots, a_{i-1})$. He first samples good a_{i-1}, which is (say) x_l. His cutoff cost to sample further depends on the number of preceding agents who have taken the same action x_l, $n(h_i)$. He will not sample further if

$$c_i > c_{n(h_i)}(q_{x_l}),$$

    where c_n is defined in (Cutoff-n). If c_i is smaller than this cutoff, he will continue to sample $x_{l'} > x_l$ until he finds a better action or exhausts all possible actions. If he does find a better action $x_{l'}$, his subsequent search decisions are identical to agent 1, i.e., he continues searching as long as

$$c_i \leq c_1(q_{x_{l'}}),$$

    where $x_{l'}$ is the best action he has seen so far.

The only observable deviation by a preceding agent is if he selects a discarded action. Off the equilibrium path, all subsequent agents behave as if they only observed the sub-history in which no agent took a discarded action.

It should be clear that given the previous lemmas, these strategies constitute a perfect Bayesian equilibrium of our social search setting, since agents' payoffs depend only on their own sampling decisions and actions taken.

3.2 SKETCH OF PROOF OF THEOREM 1

We now give a short sketch of the proof of the main theorem for the case of common search order equilibria. The formal proof is in the Appendix. Recall that in our setting, asymptotic learning occurs if and only if with probability one, at least one agent samples each action, or an action with the highest possible quality is sampled.

Let us reason through "what it takes" for an action not to be sampled. Consider some action x_l such that $q_{x_l} < \sup Q$,[11] and the first agent i, selecting it. Suppose each agent $i' \geq i$ does not sample further. By Lemma 2 we know that this happens if for each agent $i' = i + g$, his search costs $c_{i'} \geq c_g(q_{x_l})$. Let us denote the probability of this occurring for agent $i' = i + g$ as p_g, i.e. $p_g = P_C(c \geq c_g(q_{x_l}))$.

Therefore, independence of the individual search costs implies that failure of asymptotic learning occurs if and only if

$$\prod_{g=0}^{\infty} p_g > 0,$$

which in turn occurs only if $\lim_{g \uparrow \infty} p_g = 1$.

To show sufficiency, suppose search costs are not bounded away from zero, but asymptotic learning fails. By (Cutoff-n) the lowest cost that leads agent $i' = i + g$ not to search is

$$c_g(q_{x_l}) = c_0(q_{x_l}) \prod_{g=0}^{i'-i} p_g,$$

which has a limit larger than 0 by supposition. Therefore, since search costs are not bounded away from zero, $\lim_{g \uparrow \infty} p_g < 1$.

To see the intuition for necessity, recall equation (Cutoff-n) and note that the cost threshold for person i' to sample action x_{l+1} is strictly decreasing in i'. Suppose that search costs are bounded away from zero, that is there exists a $\underline{c} > 0$ such that $P_C(c \leq \underline{c}) = 0$. The proof shows that for any q_{x_l} within a finite number of agents g, the cutoff cost $c_{g+1}(q_{x_l}) < \underline{c}$. This implies that if agents i to $i + g$ costs are such that they do not sample further, then all agents $i' > i + g$ do not sample further either. Therefore if costs are bounded away from zero, herding is possible with positive probability.

---

[11] If $q_{x_l} = \sup Q$, the agent is done because he has found a best possible action.

GENERAL EQUILIBRIA   Finally, recall that the preceding argument restricted attention to common search order equilibria. We now argue that in a given environment, asymptotic learning occurs in any perfect Bayesian equilibrium σ if and only if the distribution of search costs includes 0 in the support. Necessity and sufficiency of the condition are established along similar lines as in the common search order case but require some more detail.

4   DISCUSSION & CONCLUSION

This paper introduces an alternative framework for social learning in which agents endogenously acquire private information. We assume that agents generate their private information via a costly search process. We show that the probability of agents selecting the optimal action in the long run is equal to one if and only if search costs are not bounded away from zero, i.e., 0 is contained in the support of the distribution of search costs. Our characterization of asymptotic learning is conceptually similar to Smith and Sorensen's (2000) characterization in the standard framework. We provide a micro-foundation for the exogenously given private signals assumed in the SSLM via an information gathering technology motivated by search.

Some assumptions we made to aid our analysis are worth highlighting and discussing. The major one is that agents can only take actions that they have already sampled. A possible justification for this is that taking an action involves learning about its quality. For example if the "actions" in question are websites selling products, then ordering a product involves visiting the website and learning about its quality. Another possible justification is that the effort or cost expended is in actually discovering the existence of the action (for example finding which restaurants are available at a desirable time) rather than learning about its quality. In the absence of this assumption, Lemma 1 would no longer be true. An agent who sees his predecessor take an action x_l knows that this action is "good." It may be a better strategy in expectation to sample other actions, and then take action x_l without sampling if these actions are of low quality. At a more technical level, characterizing a single person's search decisions in the absence of this assumption is not well characterized. We refer the reader to Weitzman (1979) for the classical result, and Doval (2013) for a recent attempt to relax this assumption in the one person setting.

An assumption we make that is often not satisfied in practice is that agents observe the full history of actions made by past agents. For example, the recent papers of Cai, Chen, and Fang (2009) and Tucker and Zhang (2011) empirically separate sequential social learning (observational learning in their terminology) from other channels of learning in a setting where agents only observe aggregate histories, i.e., the number of past agents who chose each action. Similarly the aforementioned work of Hendricks, Sorensen, and Wiseman

(2012) revisits the experimental data of Salganik, Dodds, and Watts (2006) with a structural model. In this setting, again, agents only observe aggregate data. It should be clear that our results continue to hold exactly as written in a setting where agents observe only the aggregate history of past agents and the action taken by their immediate predecessor.

As we demonstrated, social search model behaves similarly to the SSLM: The tendency to local conformity, fragility of mass behavior and the occurrence of incorrect herds all carry forward to our model. There are certain differences, however. For example, in our model the society evolves from one action to another until eventually agents settle on one action. Actions are always "improving," i.e. a later moving agent always takes a weakly better action. Herding occurs with probability one in finite time and learning occurs by discarding actions which are revealed inferior. Those actions, once discarded, are never chosen again. In the SSLM on the other hand agents might go back and forth between different actions, and in case of unbounded signals the society does not settle on one action in finite time.

In terms of future work, a major driving assumption in this paper, standard to the observational learning literature, is that agents observe all predecessors' actions. A more "natural" assumption would be that agents' information about predecessors' actions is determined by a social network— agents observe the actions of the predecessors who are also their neighbors ("friends") in this network. This was considered by Acemoglu, Dahleh, Lobel, and Ozdaglar (2011), extending the SSLM. The counterpart in a search setting would be interesting, and is left to future work.

A PROOFS FROM THE TEXT

PROOF OF LEMMA 1. Consider agent 2 and his conditional distribution over the state space Ω given agent 1's action a_1 and strategy σ_1. Consider an action $x' \neq a_1$. There are two possible cases:

1. Case 1: Agent 1 sampled x'. This implies that $q_{x'} \leq q_{a_1}$, since agent 1 picked action a_1. If agent 2 knew this to be the case, his conditional distribution on Ω is given by $P_\Omega \left(| q_{a_1} \geq q_{x'} \right)$.

2. Case 2: Agent 1 did not sample x'. In this case, there is no new information about action x'. If agent 2 knew this to be case the posterior on action x' is the same as the prior P_Q.

As a result, regardless of the beliefs of agent 2 about agent 1's search decisions, his belief about the quality of any action other than a_1 is first order stochastically dominated by his beliefs about a_1. Further, agent 1 must have sampled all other actions with positive probability— even if he sampled a_1 first, with positive probability his search costs are low

enough that he searched further. Therefore agent 2 selects $\sigma_2^0 = a_1$.

We can now show the Lemma for any agent $i+1 > 2$ by induction. Suppose all agents up to i follow this strategy, and that agent i selects action a_i. It now follows that any discarded action x on this history must have quality $q_x \leq q_{a_i}$, and therefore these are not sampled. Now consider action q_{a_i} versus all other non-discarded actions. By the same logic as before, agent $i+1$'s beliefs about the quality of action a_i strictly first order stochastically dominate his beliefs about the quality of other actions. Therefore $\sigma_{i+1}^0 = a_i$, i.e, he will sample a_i first. □

A.1 COMMON SEARCH ORDER EQUILIBRIA

PROOF OF LEMMA 2. For agent 1, the lemma follows trivially given (Cutoff).

We will now show it for agents $i > 1$ by induction.

BASE: First, let us consider $i = 2$. Suppose agent 1 took action x_l. By Lemma 1 we know that agent 2 samples action x_l first and learns q_{x_l}. Actions x_1 to x_{l-1} are discarded and therefore irrelevant. To determine his subsequent search decisions, we need to reason about his beliefs about actions x_{l+1} to x_k. By stationarity of search decisions, either agent 1 sampled all actions x_{l+1} to x_k or none.

Agent 2 understands therefore that agent 1 sampled all actions only if his cost c_1 was less than agent 1's search cutoff $c_0(q_{x_l})$ (recall Observation 1). Therefore, for any action $l' > l$, agent 2's beliefs about $q_{x_{l'}}$ can be given by

$$
\begin{aligned}
P_Q(q_{x_{l'}}|h_2, q_{x_l}) =& P_Q(q_{x_{l'}}) \times \mathbb{P}[\text{agent 1 did not search}] \\
&+ P_Q(q_{x_{l'}}|q_{x_{l'}} < q_{x_l}) \times \mathbb{P}[\text{agent 1 did search}], \\
=& P_Q(q_{x_{l'}})P_C(c_1 > c_0(q_{x_l})) + P_Q(q_{x_{l'}}|q_{x_{l'}} < q_{x_l})P_C(c_1 \leq c_0(q_{x_l})). \quad (3)
\end{aligned}
$$

Agent 2 will search if and only if his expected gain from searching exceeds his cost, which gives us a cutoff cost of:

$$
c_1(q_{x_l}) = \int_{\tilde{q} \geq q_{x_l}} (\tilde{q} - q_{x_l}) \, dP_Q(\tilde{q}|h_2, q_{x_l}).
$$

Substituting in $P_Q(\tilde{q}|h_2, q_{x_l})$ from (3), we have

$$
\begin{aligned}
c_1(q_{x_l}) =& P_C(c_1 > c_0(q_{x_l})) \int_{\tilde{q} \geq q_{x_l}} (\tilde{q} - q_{x_l}) \, dP_Q(\tilde{q}). \\
=& P_C(c_1 > c_0(q_{x_l})) \, c_0(q_{x_l}).
\end{aligned}
$$

If agent 2's search cost is below this threshold, he will search until he either finds a better action or runs out of actions. If he does find a better action $x_{l'}$ for $l' > l$, he must conclude that agent 1 did not search past l (otherwise he would have found and selected action $x_{l'}$). His posterior beliefs on actions $x_{l''}$ for $l'' > l'$ return to his prior beliefs P_Q, and therefore further search decisions are given by the cutoff function $c_0(\cdot)$.

INDUCTIVE HYPOTHESIS   Let us suppose agents' 2 to $i - 1$ are as described in the Lemma for $i > 2$.

INDUCTIVE STEP   Consider the case of agent $i > 2$. By Lemma 1 we know that agent i samples action x_l first and learns q_{x_l}. Actions x_1 to x_{l-1} are discarded and therefore irrelevant. To determine his subsequent search decisions, we need to reason about his beliefs about actions x_{l+1} to x_k.

Firstly, note that from the point of view of agent i, actions x_{l+1} to x_k are identical— any agent $i' < i$ that sampled one of these actions must have sampled all of them. To see this consider any agent $i' < i$. There are three cases:

1. $a_{i'} \neq x_l$: In this case agent i' clearly didn't search up to x_l, and therefore has not sampled x_{l+1} to x_k.

2. $a_{i'} = x_l$, but agent i' did not sample past x_l.

3. $a_{i'} = x_l$, but agent i' did sample past x_l: If this agent decided to search past x_l, it must have found that these actions were inferior to the x_l, since it chose to take action x_l. But then, by stationarity of search decisions, it would have continued to search until it exhausted all the actions.

Recall that $n(h_i)$ is the number of agents preceding agent i who selected action x_l. Agent i's beliefs about $q_{l'}$ for $l' > l$ can therefore be written as

$$P_Q(q_{x_{l'}}|h_i, q_{x_l}) = P_Q(q_{x_{l'}}) \times \mathbb{P}[\text{agents } i - n(h_i) \text{ to } i - 1 \text{ did not search}]$$
$$+ P_Q(q_{x_{l'}}|q_{x_{l'}} < q_{x_l}) \times \mathbb{P}[\text{some agent } i - n(h_i) \text{ to } i - 1 \text{ did search}].$$

By the inductive hypothesis, this can be written as

$$P_Q(q_{x_{l'}}|h_i, q_{x_l}) = P_Q(q_{x_{l'}}) \prod_{\eta=0}^{n(h_i)-1} P_C(c_{i-n(h_i)+\eta} > c_\eta(q_{x_l}))$$
$$+ P_Q(q_{x_{l'}}|q_{x_{l'}} < q_{x_l}) \left(1 - \prod_{\eta=0}^{n(h_i)-1} P_C(c_{i-n(h_i)+\eta} > c_\eta(q_{x_l}))\right). \tag{4}$$

As before, agent i will only search further if his expected value from search exceeds his costs. This gives us a cutoff cost of:

$$c_{n(h_i)}(q_{x_l}) = \int_{\tilde{q} \geq q_{x_l}} (\tilde{q} - q_{x_l}) \, dP_Q(\tilde{q}|h_i, q_{x_l}).$$

Substituting in $P_Q(\tilde{q}|h_i, q_{x_l})$ from (4), we have

$$c_{n(h_i)}(q_{x_l}) = \left(\prod_{\eta=0}^{n(h_i)-1} P_C(c_{i-n(h_i)+\eta} > c_\eta(q_{x_l})) \right) \int_{\tilde{q} \geq q_{x_l}} (\tilde{q} - q_{x_l}) \, dP_Q(\tilde{q}),$$

$$= \left(\prod_{\eta=0}^{n(h_i)-1} P_C(c_{i-n(h_i)+\eta} > c_\eta(q_{x_l})) \right) c_0(q_{x_l}),$$

as desired.

If agent i's search cost is below this threshold, he will search until he either finds a better action or runs out of actions. If he does find a better action $x_{l'}$ for $l' > l$, he must conclude that previous agents did not search past l (otherwise they would have found and selected action $x_{l'}$). His posterior beliefs on actions $x_{l''}$ for $l'' > l'$ therefore return to his prior beliefs P_Q, and further search decisions are given by the cutoff function $c_0(\cdot)$. $\qquad\square$

PROOF OF THEOREM 1. Asymptotic learning occurs if and only if the probability that any action is sampled by some buyer converges to 1. We need to show that asymptotic learning occurs if and only if search costs are not bounded away from zero.

SUFFICIENCY Suppose our condition holds, i.e. search costs are not bounded away from zero, or to put it alternately, 0 is in the support of the distribution of search costs. We will show that given any history such that there are some un-discarded actions (i.e. the current selected action is x_l for $l < k$), the probability that there is further search by some subsequent agent converges to 1.

Given that action l has quality q_{x_l}, denote

$$p_g = P_C(\tilde{c} > c_g(q_l)),$$

i.e. p_g is the probability that an agent does not search if the last g agents selected action l. Given independence of agents' search costs, search by a subsequent agent occurs with probability 1 if and only if

$$\prod_0^\infty p_g = 0.$$

Suppose that asymptotic learning fails, i.e. suppose that

$$\prod_0^\infty p_g > 0.$$

Since we know that by (Cutoff-n),

$$c_n(q_l) = c_0(v) \prod_0^{n-1} p_g,$$

we have that

$$\lim_{g \to \infty} c_g(q_{x_l}) \equiv L_c > 0.$$

Note that $p_g \leq \mathbb{P}(c \geq L_c)$ for all $g \in \mathbb{N}$. By our support assumption, we therefore have that, $\mathbb{P}(c \geq L_c) \equiv p^\star < 1$. Therefore

$$\prod_{g=0}^n p_g \leq (p^\star)^{n+1},$$

$$\implies \prod_{g=0}^\infty p_g = 0,$$

which is a contradiction.

NECESSITY   We now establish necessity, i.e. that asymptotic learning implies $0 \in \mathrm{supp}(P_C)$. So suppose not, i.e. suppose asymptotic learning occurs but $0 \notin \mathrm{supp}(P_C)$. By our assumption of a non-trivial search environment (Assumption 1) there exists a quality $\hat{q} \in \mathrm{supp}(P_Q)$ such that

1. $P_Q(q > \hat{q}) > 0$, and

2. With positive probability, the first agent does not search further on observing $q_{s_1^1} = \hat{q}$, i.e.,

$$P_C(c_1 > c_0(\hat{q})) > 0.$$

Recall that asymptotic learning occurs if and only if

$$\prod_{g=0}^{\infty} P_C(\tilde{c} > c_g(\hat{q})) = 0. \tag{5}$$

By (Cutoff-n)

$$c_n(q) = c_0(q) \prod_{g=0}^{n-1} P_C(\tilde{c} > c_g(q)).$$

Further, by our assumption that 0 is not in the support of the distribution of search costs, we know there exists $\underline{c} > 0$ such that

$$P_C(c \geq \underline{c}) = 1.$$

Note that by (5), $c_n(q)$ is decreasing in n for a given q, which implies that $P_C(\tilde{c} > c_n(q))$ is increasing in n. Further, since $P_C(\tilde{c} > c_n(q)) \in [0,1]$, $\prod_{g=0}^{n-1} P_C(\tilde{c} > c_g(q))$ is decreasing in n. Therefore, (5) and (Cutoff-n) we can pick n large enough so that

$$\prod_{g=0}^{n-1} P_C(\tilde{c} > c_g(q)) < \frac{\underline{c}}{c_0(q)}$$

$$\implies c_n(q) < \underline{c}$$

$$\implies P_C(c \geq c_n(q)) = 1.$$

$$\implies \prod_{g=0}^{\infty} P_C(\tilde{c} > c_g(\hat{q}))$$

$$= \prod_{g=0}^{n-1} P_C(\tilde{c} > c_g(\hat{q}))$$

$$> 0,$$

which contradicts (5). □

A.2 GENERAL EQUILIBRIA

A.2.1 NECESSITY

Proof of necessity follows on similar lines as the common search order case. Whenever the distribution of search costs is bounded away from 0, with positive probability, early moving agents have a high search cost, and select a suboptimal action $\hat{x}$, possibly not sampling a better action x. Once enough early moving agents have selected $\hat{x}$, later moving agents

have no incentive to search—the largest search cost who would find it profitable to search is smaller than the lower bound of distribution of search costs.

A.2.2 SUFFICIENCY

Consider any equilibrium strategy profile σ. We need to show that asymptotic learning occurs when search costs are not bounded away from zero.

A little bit of additional notation will be useful. Fix equilibrium strategies for each agent. The process of actions is determined by the realization in the probability space $Y = \Omega \times C^\infty \times D^\infty$. Here Ω is the set of possible realizations of qualities of each action, C^∞ the set of realizations of search costs of each agent, and $(D, \mathcal{F}_D, \lambda)$ is a probability space determining the possible mixed strategy realizations of a given agent.

We define $h(\cdot)$ as a function which assigns to each extended state $y \in Y$ the corresponding infinite histories of actions taken by each agent, i.e. $h : Y \to A^\infty$. Consider the following events in Y,

$$E_x^i = \{y \in Y : \text{agents } 1, ..., i-1 \text{ do not sample } x\},$$
$$E_x^\infty = \{y \in Y : \text{no agent samples } x\},$$
$$E_{h_i} = \{y \in Y : h(y) = (h_i, \ldots)\}.$$

In words, the first event is where the first $i-1$ agents do not sample a given action x, the second where no agent samples x, and the third where the ith agents sees history h_i.

COUNTABLE SET OF HISTORIES   First, we establish that the set of infinite equilibrium histories is countable. On the equilibrium path, any action x' that was selected but then discarded is not selected again as it has been revealed inferior to some other action. As the set of actions is finite, each equilibrium history features a finite time period in which herding starts, that is from that time onward all later agents select the same action. Therefore, any infinite equilibrium history h is determined by an initial herding time t^*, a finite sequence h_{t^*} and the corresponding action x which is selected in the herd starting in period t^*. The set of possible equilibrium histories $\mathbf{H}_\sigma$ can therefore be written as

$$\mathbf{H}_\sigma = \bigcup_{t^* \in \mathbb{N}} \bigcup_{h_{t^*} \in A^{t^*}} \bigcup_{x \in X} (h_{t^*}, x^\infty)$$

where x^∞ denotes an infinite sequence of action x. Note that $\mathbf{H}_\sigma^\infty$ is a countable set as it is derived as the countable union of finite sets. As there are only countable equilibrium histories h we restrict attention to those histories which have positive probability, $\mathbb{P}\left[E_h\right] > 0$.

In what follows, we will establish that failure of asymptotic learning implies $\mathbb{P}[E_x^\infty | E_h] > 0$. We then show that this leads to a contradiction with search costs containing 0 in its support, thereby establishing sufficiency. We can trivially discard histories where some agent actually takes action x (because he must have sampled that action first), so consider only histories on which action x is never taken.

DECREASING CONDITIONAL PROBABILITIES   Next, we establish that the conditional probability of no agent up to i having sampled action x, conditional on action x never having been chosen up to i, is strictly decreasing with i, i.e.

$$\mathbb{P}\left[E_x^i | E_{h_i}\right] > \mathbb{P}\left[E_x^{i+1} | E_{h_{i+1}}\right]. \tag{6}$$

To see this, note

$$\mathbb{P}\left[E_x^{i+1} | E_{h_{i+1}}\right] = \mathbb{P}\left[E_x^i | E_{h_{i+1}}\right] \times \mathbb{P}\left[E_x^{i+1} | E_{h_{i+1}}, E_x^i\right]$$
$$= \mathbb{P}\left[E_x^i | E_{h_i}\right] \times \mathbb{P}\left[E_x^{i+1} | E_{h_{i+1}}, E_x^i\right]$$

where the first equality follows from Bayes rule, the second equality from the fact that agent i does not take action x. The inequality now follows since $\mathbb{P}\left[E_x^{i+1} | E_{h_{i+1}}, E_x^i\right] < 1$.

BOUNDING CONDITIONAL PROBABILITIES   We now show that if asymptotic learning fails then there exists a positive probability event in Y such that for every agent i his conditional probability of x not having been sampled, given that no predecessor selected x, is strictly larger than some strictly positive ϵ. Consider the following conditional probability $\mathbb{P}\left[E_x^\infty | E_h\right]$. By (6), this is strictly smaller than $\mathbb{P}\left[E_x^i | E_{h_i}\right]$ for any i.

Suppose now that asymptotic learning fails. This implies that there exists an action x and a positive probability event E_{xx}^∞ such that action x is not sampled along the infinite history while being the uniquely optimal action. Consider any infinite history h such that $\mathbb{P}\left(E_{xx}^\infty \cap E_h\right) > 0$ (note that such an event E_h exists as $\mathbb{P}\left(E_{xx}^\infty\right) > 0$ by assumption and there are only countable histories). We therefore have

$$\mathbb{P}\left[E_{xx}^\infty | E_h\right] = \frac{\mathbb{P}\left(E_{xx}^\infty \cap E_h\right)}{\mathbb{P}\left(E_h\right)} > 0.$$

Therefore,

$$\mathbb{P}\left[E_x^i | E_{h_i}\right] > \mathbb{P}\left[E_x^\infty | E_h\right] \geq \mathbb{P}\left[E_{xx}^\infty | E_h\right].$$

Here the first inequality follows by repeated application of (6), the second since $E_{xx}^\infty \subseteq E_x^\infty$.

BOUNDING BELIEFS AFTER PRIVATE SEARCH   Note that $\mathbb{P}\left[E_x^i | E_{h_i}\right]$ is agent i's conditional probability of action x not having been sampled by his predecessors given history h_i,

prior to agent i sampling any actions himself. His decision whether or not to sample x will be based upon his posterior probability of E_x^i generated according to the quality realizations of the actions he samples.

Let us fix a quality realization $\omega = (q_1, ..., q_k)$ such that $(\omega, c^\infty, d^\infty) \in E_h$ for some c^∞, d^∞. For any such ω and any agent i there exists a subvector $\vec{q} \subset (q_1, ..., q_k)$ which represents the realized qualities of actions agent i knows by the time he decides to stop searching and selects an action. The posterior probability of such an agent i' is given by $\mathbb{P}\left[E_x^{i'} | E_{h^{i'}}, \vec{q}\right]$. Analogous arguments to those that established (6) show

$$\mathbb{P}\left[E_x^i | E_{h_i}, \vec{q}\right] \geq \mathbb{P}\left[E_x^{i+1} | E_{h_{i+1}}, \vec{q}\right] \geq \mathbb{P}\left[E_x^\infty | E_h, \vec{q}\right],$$

for all $i \in \mathbb{N}$.

ESTABLISHING ASYMPTOTIC LEARNING As there are only finitely many actions, there exists a subset X' of X and a positive probability event $E_{h,X'} \subset E_h$ such that for an infinite subsequence of agents searched exactly the actions in $X' \subseteq X$.

Our goal is to show that there exists a positive probability event in $E_{h,X'}$ such that the posterior probability of no agent sampling action x, $\mathbb{P}\left[E_x^\infty | E_h, \vec{q}\right]$ for quality realizations $\vec{q}$ in this event is strictly positive. Suppose not—then

$$\int_{\vec{q} \in Q^{|X'|}} \mathbb{P}\left[E_x^\infty | E_h, \vec{q}\right] d\mathbb{P}\left[\vec{q} | E_h\right] = 0,$$

contradicting $\mathbb{P}\left[E_x^\infty | E_h\right] > 0$.

Therefore, there exists a positive probability event such that for an infinite subsequence of agents the infimum of their posterior probability of action x not having been sampled by their predecessors is bounded away from zero. This in turn implies that the infimum cost cutoff (across all agents) which determines whether or not to search is bounded away from zero. But since search costs are not bounded away from zero and independent across agents, the probability that infinitely many agents have a cost realization above the threshold is equal to zero establishing a contradiction.

REFERENCES

ACEMOGLU, D., M. DAHLEH, I. LOBEL, AND A. OZDAGLAR (2011): "Bayesian learning in social networks," *The Review of Economic Studies*, 78(4), 1201–1236.

AGHION, P., P. BOLTON, C. HARRIS, AND B. JULLIEN (1991): "Optimal learning by experimentation," *The Review of Economic Studies*, 58(4), 621–654.

ALI, S. (2013): "Social Learning with Endogenous Information," Discussion paper, University of California at San Diego.

ALI, S., AND N. KARTIK (2011): "Herding with collective preferences," *Economic Theory*, 51(3), 1–26.

ARIELI, I., AND M. MUELLER-FRANK (2013): "Inferring Beliefs from Actions," Discussion paper, IESE Business School.

ATHEY, S., AND G. ELLISON (2011): "Position auctions with consumer search," *The Quarterly Journal of Economics*, 126(3), 1213–1270.

AVERY, C., AND P. ZEMSKY (1998): "Multidimensional uncertainty and herd behavior in financial markets," *American Economic Review*, 88, 724–748.

BANERJEE, A. (1992): "A simple model of herd behavior," *The Quarterly Journal of Economics*, 107(3), 797–817.

BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, 100, 992–1026.

BOSE, S., G. OROSEL, M. OTTAVIANI, AND L. VESTERLUND (2006): "Dynamic monopoly pricing and herding," *The RAND Journal of Economics*, 37(4), 910–928.

——— (2008): "Monopoly pricing in the binary herding model," *Economic Theory*, 37(2), 203–241.

BURGUET, R., AND X. VIVES (2000): "Social learning and costly information acquisition," *Economic Theory*, 15(1), 185–205.

CAI, H., Y. CHEN, AND H. FANG (2009): "Observational Learning: Evidence from a Randomized Natural Field Experiment," *The American Economic Review*, 99(3), 864–882.

DEKEL, E., AND M. PICCIONE (2000): "Sequential voting procedures in symmetric binary elections," *Journal of Political Economy*, 108(1), 34–55.

DOVAL, L. (2013): "Whether or Not to Open Pandora's Box?," Discussion paper, Northwestern University.

Grossman, S. J., and J. E. Stiglitz (1976): "Information and competitive price systems," *The American Economic Review*, pp. 246–253.

Hendricks, K., A. Sorensen, and T. Wiseman (2012): "Observational learning and demand for search goods," *American Economic Journal: Microeconomics*, 4(1), 1–31.

Kihlstrom, R., L. Mirman, and A. Postlewaite (1984): "Experimental Consumption and the 'Rothschild Effect'," in *Bayesian Models of Economic Theory*. Amsterdam, Elsevier.

Kultti, K., and P. Miettinen (2006): "Herding with costly information," *International Game Theory Review*, 8(01), 21–31.

Meredith, M., and Y. Salant (2007): "On the causes and consequences of ballot order effects," *Political Behavior*, pp. 1–23.

Mueller-Frank, M. (2012): "Market Power, Fully Revealing Prices and Welfare," Discussion paper, IESE Business School.

Park, A., and H. Sabourian (2011): "Herding and contrarian behavior in financial markets," *Econometrica*, 79(4), 973–1026.

Rubinstein, A., and Y. Salant (2006): "A model of choice from lists," *Theoretical Economics*, 1, 3–17.

Salganik, M. J., P. S. Dodds, and D. J. Watts (2006): "Experimental study of inequality and unpredictability in an artificial cultural market," *science*, 311(5762), 854–856.

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

Smith, L., and P. Sorensen (2006): "Informational herding and optimal experimentation," .

Tucker, C., and J. Zhang (2011): "How does popularity information affect choices? A field experiment," *Management Science*, 57(5), 828–842.

Weitzman, M. L. (1979): "Optimal search for the best alternative," *Econometrica*, pp. 641–654.

Wolpin, K. (1987): "Estimating a structural search model: the transition from school to work," *Econometrica*, 55(4), 801–817.