Sequential Collective Search in Networks*

Niccolò Lomys†

August 2, 2018

Abstract

I study social learning in networks where rational agents act in sequence, observe the choices of their connections, and acquire private information via costly sequential search. I characterize perfect Bayesian equilibria of the model by linking individual search policies to the probability that agents select the best action. The information structure of the model precludes information aggregation via martingale convergence arguments. If (and only if) search costs are not bounded away from zero, an improvement principle holds even though the informational environment significantly differs from that of the standard model with exogenous private signals. I leverage the improvement principle to show that asymptotic learning obtains in sufficiently connected networks where information paths are identifiable. When search costs are bounded away from zero, even a weaker notion of long-run learning fails, except in ad hoc network topologies. Networks where agents observe the choices of random numbers of immediate predecessors share many equilibrium properties with the complete network, including the rate of convergence and the probability of wrong herds. Transparency of past histories has short-run, but not long-run, implications for welfare and efficiency. The simple policy intervention of letting agents observe the relative fraction of previous choices reduces inefficiencies and welfare losses.

Keywords: Social Networks; Rational Learning; Herding; Search; Bandit Problems; Sequential Decisions; Improvement and Large-Sample Principles.

JEL Classification: C72; D62; D81; D83; D85.

*I am grateful to Volker Nocke and Emanuele Tarantino for guidance and constant support throughout the project. I thank Alessandra Allocca, Francesco Paolo Conteduca, Daniel Garrett, Bertrand Gobillard, Amir Habibi, Takakazu Honryo, Raphaël Levy, Lorenzo Magnolfi, Chiara Margaria, Pauli Murto, Alexei Parakhonyak, Martin Peitz, Evan Sadler, Konrad Stahl, Thomas Tröger, and Leeat Yariv for valuable conversations. I received helpful comments from seminar audiences at the Toulouse School of Economics, the University of Mannheim, GAMES 2016, EEA–ESEM 2016, the 2016 UECE Lisbon Meetings in Game Theory and Applications, the ENTER Jamboree 2017 at UCL, the Third Annual Conference in Network Science and Economics at WUSTL, the MaCCI IO Day 2017, the 2017 Summer School of Econometric Society in Seoul, EARIE 2017, EWMES 2017, and the 6th European Conference on Networks at Barcelona GSE–UPF. I thank the Department of Economics at Yale University, where part of this research was conducted, for its support and hospitality. The usual disclaimer applies.

†Center for Doctoral Studies in Economics, University of Mannheim; niccolo.lomys@gess.uni-mannheim.de.

1 Introduction

Social learning is the study of how individuals combine their private information with others' experiences to identify the best course of action in the face of payoff-relevant uncertainty. When characterizing conditions under which societies efficiently aggregate dispersed information or, in contrast, herd on suboptimal behavior, it is routine to assume that agents are born with an exogenous information endowment. Contrary to this premise, in most circumstances of social and economic interest information only becomes available at a cost. Agents' incentives to collect the relevant information are ambiguous. On the one hand, the availability of others' experiences weakens individual motivation to acquire independent information and encourages the *exploitation* of others' wisdom, increasing the chances of wrong herds. On the other hand, the possibility of wrong herds fosters independent *exploration*, reducing the odds of suboptimal behavior.

The resulting trade-off is largely neglected in social learning models over general networks because of the technical difficulties that emerge when studying strategic behavior of rational agents in such environments. Prior work deals with these complications by weakening the rationality assumption so as to simplify individual decision rules or by focusing on particular network structures. As the topology of social ties crucially shapes both information flows and individual incentives to acquire independent information, it has been repeatedly acknowledged that progress would be desirable within the Bayesian benchmark (see, e.g., Sadler (2014) and Golub and Sadler (2016)).

In this paper, I address these challenges and develop a tractable model of sequential social learning where agents *(i)* are rational, *(ii)* only observe the choices of their connections over general networks, and *(iii)* endogenously acquire private information by costly sequential search.

Countably many Bayes rational agents act in sequence and must each take one of two feasible actions. The qualities of the two actions are independent draws from the same distribution and agents have no a priori private information about their realization. Agents wish to select the action with the highest quality and payoff externalities are absent. Building on Acemoglu, Dahleh, Lobel and Ozdaglar (2011) and Lobel and Sadler (2015), each agent observes a subset of previous agents, which I call the agent's neighborhood. Neighborhoods are stochastically generated according to a joint distribution, which I refer to as the *network topology*. The framework allows for arbitrary correlations among neighborhoods. After observing his neighbors and their actions, each agent engages in *costly sequential search* with recall before selecting his action. Searching an action perfectly (and only) reveals the quality of that action to the agent, but comes at a cost. After sampling the first action, the agent decides whether to discontinue search or to sample the second alternative. Each agent can only select an action from those he has sampled. For a single agent, the search problem is a version of that proposed by Weitzman (1979), and studied by Mueller-Frank and Pai (2016) (hereafter, MFP) in a social learning model with perfect observation of all previous choices (complete network). Search costs are i.i.d. across agents. Individual neighborhoods, sampling decisions, and search costs are not observed by subsequent agents.

The model results in a dynamic game of incomplete information where the the network topology shapes agents' possibility to learn from others' behavior and the *search technology* shapes agents' possibility to acquire independent information. I characterize conditions on search technologies and network topologies under which positive long-run learning outcomes obtain or fail. The learning model I analyze is non-standard for two reasons. First, while the study of long-run outcomes requires understanding the dynamics of the probability that agents select the best action, agents use their information to maximize the value of their sequential search program. The two problems are not the same; that is, maximizing the probability of selecting the best action is not equivalent to determining the optimal sequential search policy. Second, the information structure of the model precludes information aggregation via martingale convergence arguments, as no social belief that forms a martingale is of some use when characterizing equilibrium behavior.

I describe individual sequential search policies in any perfect Bayesian equilibrium of the model by relating agents' optimization to the probability that they select the best action. This connection makes the analysis of long-run learning outcomes tractable. Upon observing his neighbors and their choices, each agent computes the probability that none of the individuals in his personal subnetwork relative to each action (i.e., the agents he is directly or indirectly linked to who take that action) has sampled both actions. This enables the agent to rank the marginal distributions of the quality of the two actions in terms of first-order stochastic dominance. According to Weitzman (1979)'s search rule, which action to sample first is uniquely determined. Next, the agent combines the information about the quality of the first action with his social information to update the above probability and infer the expected additional gain from the second search. If this gain is larger than his private search cost, the agent samples the second action and then selects the best one. Otherwise, he stops searching and takes the first action sampled.

In equilibrium, agents with no neighbors have the strongest incentives to generate new information. These incentives decrease in the quality of the first action sampled. Remarkably, the incentives to acquire independent information need not be monotonic in the quality of the first action sampled for agents who observe the choices of other individuals. These facts neatly capture how the exploration-exploitation trade-off interacts with individual incentives in my setup.

I establish an *improvement principle* (hereafter, IP) for the present environment. The IP captures the idea that imitation, paired with some individual improvement upon it, is sufficient to learn how to select the best action in the long run. It is based on the following heuristic. Upon observing who his neighbors are, each agent chooses only one neighbor to rely on and determines his optimal search policy regardless of what others have done. If *search costs* are *not bounded away from zero*, there exists a strict lower bound on the increase in the probability that an agent samples first the best action over his chosen neighbor's probability. The improvement occurs unless the chosen neighbor already samples the best action with probability one at the first search.

Acemoglu et al. (2011) and Lobel and Sadler (2015) originally develop an IP for the standard sequential social learning model (henceforth, SSLM) to establish positive learning results in stochastic networks. In the SSLM agents receive a free private signal, which is informative about the relative quality of all alternatives, and wish to match their action with an unknown state of nature.[1] My results extend the reach of the IP and of the learning principle it captures to a new informational environment, which departs from that of the SSLM in three relevant aspects. First, private information is generated by equilibrium play rather than being exogenously given to the agents. Second, while in the SSLM agents have imperfect private information about the relative quality of the two alternatives, in my model sampling an action perfectly reveals the quality of

[1] The SSLM dates back to the seminal work of Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992), and Smith and Sørensen (2000), who propose this class of models, but assume that each agent observes all past actions before making his choice. Smith and Sørensen (2014) introduce neighbor sampling in the SSLM but, differently than in my model, they assume that individuals ignore the identity of the agents they observe.

that action only. Finally, the inferential challenge crucially differs: agents maximize the value of a sequential information acquisition program rather than the probability of matching an underlying state of nature or an ex ante expected utility. The possibility to describe agents' sequential search policies in terms of probabilities, however, bridges the search setting I study to the SSLM. Thus, an IP holds in the two settings in spite of limited comparability of their informational environments.[2]

The first learning metric I consider is *asymptotic learning*, which occurs if the probability that agents take the best action converges to one as the size of the society grows large. I leverage the IP to show that asymptotic learning obtains in network topologies where *arbitrarily long information paths occur almost surely and are identifiable*. That is, if search costs are not bounded away from zero, asymptotic learning obtains in sufficiently connected networks where individual neighborhood realizations do not lead agents astray about the broader network realization. In such networks, agents can identify the correct neighbor to rely on and improvements last long enough for agents to select the best action. The IP is, however, fragile: if zero is not in the support of the search cost distribution, the IP breaks down and learning via improvements upon imitation is precluded.

For *search costs* that are *bounded away from zero*, I introduce a new metric of social learning, which I dub *maximal learning*. Maximal learning occurs if, in the long run, agents take the best action with the same probability as a single agent with the best search opportunities (the lowest search cost type) and the strongest incentives to explore (no social information). This learning requirement is weaker than asymptotic learning and represents the best outcome a society can aim for when zero is not in the support of the search cost distribution.

If search costs are bounded away from zero, maximal learning fails in many common deterministic and stochastic networks. Thus, when search costs are bounded away from zero, asymptotic learning fails discontinuously with respect to the learning metric. Therefore, positive learning results are fragile with respect to perturbations in the support of the search cost distribution.

In a few stochastic networks, maximal (and sometimes also asymptotic) learning obtains despite zero is not in the support of the search cost distribution. Thus, search costs that are not bounded away from zero are not, in general, necessary for asymptotic learning. The positive result, however, is limited to very special network topologies. In fact, the impossibility to develop martingale convergence arguments severely undermines the ability to learn via the aggregation of the information that large samples of other agents' choices contain.

From the viewpoint of selecting the best action, individual search behavior in networks where agents observe the choices of random numbers of immediate predecessors is equivalent to the search behavior in the complete network. These network topologies thus inherit several equilibrium properties from the complete network, including the probability of wrong herds and the speed of learning, which is faster than polynomial.[3] Reducing transparency of past histories, however, leads to inefficient duplication of costly search. I compare equilibrium welfare in the complete network and in the network where each agent only observes his most recent predecessor. The difference only vanishes in the limit of an infinitely patient society but is significant in the short and medium run. Simple policy interventions, such as letting agents observe the relative share of previous choices in addition to their neighbors' choices, reduce inefficiencies and welfare losses.

[2]The informational monotonicity we make use of in the IP is related to the (expected) welfare improvement principle in Banerjee and Fudenberg (2004) and Smith and Sørensen (2014), and to the imitation principle in Bala and Goyal (1998) and Gale and Kariv (2003).

[3]I also show that the rate of convergence is logarithmic under random sampling of one agent from the past.

This paper contributes to both the economic theory of social learning and its applications to the economics of social media and Internet search. The theoretical novelty of the paper is to analyze costly information acquisition in a model of rational learning over general networks. In turn, the information acquisition technology—sequential search, which has received much attention in the applied literature—naturally relates the model to a variety of applications. Many real-world information acquisition and choice problems are well-modeled by sequential search—in particular, situations where taking an action requires learning about its quality, functioning, existence, or availability. Examples are widespread: firms need to be aware of a new technology and assess its merits before adoption; consumers gather information before purchasing an expensive durable good; investors try to understand different financial instruments before making an investment decision; patients inquire into alternative treatments before undergoing an invasive surgery.

A compelling motivation for my model comes from the large evidence that people's online behavior—what they search on web search engines, the order in which they do so, and their resulting purchase decisions—is often inspired by what they observe on social media. For instance, suppose we need to decide which of two recently released comedies to watch. The two movies have a cast and a direction of comparable reputation so that it is ex-ante unclear which one is better. However, we observe on Facebook the movie our friends watched through their check-ins or the Facebook pages they liked, but only have a vague idea of whom they observed in turn. Our friends' decisions give us a first impression of what film is likely to be the best one. We then search on Google for this movie to learn where and when it is played and to read experts' reviews. Looking for movie times and reading reviews takes time and effort, and this idiosyncratic cost depends on factors that are our private information (whether we are in a rush, how much time we can divert from other activities, etc.). Depending on movie times, reviews, and our opportunity cost, we either watch the movie we first learned about, or invest more time searching for information about the other option.[4]

Interestingly, the policy interventions I discuss, such as letting agents observe the relative share of previous choices, are common in online platforms that aggregate individual choices by sorting different items according to their popularity. For instance, when deciding which comedy to watch, agents also have access to box office data and ticket sales rankings.

More broadly, Armstrong (2016) argues that others' choices and aggregate sales rankings may guide the order in which consumers search for new products and influence which items become popular in the long run. For example, people observe on Spotify what songs their connections listen to, and on Flickr the cameras that have been used to take the pictures that other users share. In such cases, the order in which individuals search for a new song or camera is not random, but informed by the previous choices of their connections, and so is their resulting purchase decision. This paper sheds light on the implications of such behavior for social learning, product diffusion and demand, and on the forces that may lead consumers to herd on inferior items.

Road Map. In Section 2, I describe the model. In Section 3, I define asymptotic learning and characterize equilibrium strategies. In Section 4, I establish the improvement principle and the main results on asymptotic learning. In Section 5, I introduce maximal learning and present the

[4] As we need to know where the movie is played and whether it is available at the desired time, we cannot watch a movie we have not searched for. Moreover, reading a movie's review or checking its schedule reveals information about (the quality of) that movie, but does not directly reveal anything about the other movie.

main results with respect to this metric. In Section 6, I present the main results on the rate of convergence, welfare, and efficiency. In Section 7, I discuss the related literature and conclude. Supporting examples are in Appendix A and formal proofs are in Appendix B.

2 Model

2.1 Collective Search Environment

Agents and Actions. A countably infinite set of agents, indexed by $n \in \mathbb{N} \coloneqq \{1, 2, \ldots\}$, sequentially select a single action each, with agent n acting at time n. Each agent has to choose one of two possible alternatives in the set of available actions $X \coloneqq \{0, 1\}$, which is identical across agents. Restricting attention to two actions simplifies the exposition, but does not affect the results. A typical element of X is denoted by x, while the action agent n selects is denoted by a_n. Calendar time is common knowledge and the order of moves exogenous.

State Process. Actions differ in their qualities, but are ex-ante homogeneous. I denote with q_x the quality of action x. Qualities q_0 and q_1 are i.i.d. draws from a probability measure $\mathbb{P}_Q$ over $Q \subseteq \mathbb{R}_+ \coloneqq \{s \in \mathbb{R} : s \geq 0\}$. The state of the world $\omega \coloneqq (q_0, q_1)$ consists of the realized quality of the two actions and is drawn once and for all at time zero. The state space is $\Omega \coloneqq Q \times Q$, with product measure $\mathbb{P}_\Omega \coloneqq \mathbb{P}_Q \times \mathbb{P}_Q$. This formulation captures finite, and countably and uncountably infinite state spaces. The resulting probability space, $(\Omega, \mathcal{F}_\Omega, \mathbb{P}_\Omega)$, is the *state process* of the model and is common knowledge. Whenever convenient, I denote the state process with $(Q, \mathcal{F}_Q, \mathbb{P}_Q)$.

Agents have homogeneous preferences and wish to select the action with the highest quality. To do so, they have access to two sources of information: *social information*, which is derived from observing a subset of other agents' past actions, and *private information*, which is endogenously acquired by costly sequential search. The next two paragraphs describe the two processes in detail.

Network Topology. Agents do not necessarily observe all past actions, but only those of a subset of previous agents according to the structure of the social network, as first modeled in Acemoglu et al. (2011) and generalized by Lobel and Sadler (2015). The set of agents whose actions are observed by agent n, denoted by $B(n)$, is called n's neighborhood. Since agents can only observe actions taken previously, $B(n) \in 2^{\mathbb{N}_n}$, where $2^{\mathbb{N}_n}$ denotes the power set of $\mathbb{N}_n \coloneqq \{m \in \mathbb{N} : m < n\}$. Neighborhoods $B(n)$ are random variables generated via a probability measure $\mathbb{Q}$ on the product space $\mathbb{B} \coloneqq \prod_{n \in \mathbb{N}} 2^{\mathbb{N}_n}$. Given a measure $\mathbb{Q}$ on $\mathbb{B}$, I refer to the resulting probability space $(\mathbb{B}, \mathcal{F}_\mathbb{B}, \mathbb{Q})$ as the *network topology*. Particular realizations of the random variables $B(n)$ are denoted by B_n.

This formulation allows for stochastic network topologies with arbitrary correlations between agents' neighborhoods, as well as for independent neighborhoods (when $B(n)$'s are generated by probability measures $\mathbb{Q}_n$'s on $2^{\mathbb{N}_n}$ and the draws from each $\mathbb{Q}_n$ are independent from each other) and deterministic network topologies (when $\mathbb{Q}$ is a Dirac distribution on a single element of $\mathbb{B}$).

The sequence of neighborhood realizations describes a social network of connections between the agents. The network topology is common knowledge, whereas the realized neighborhood B_n is private information of agent n. If $n' \in B_n$, then n not only observes the choice $a_{n'}$, but also knows the identity of this agent (equivalently, the time at which this agent has acted). Crucially, however, n does not necessarily observe $B_{n'}$ or the actions of the agents in $B_{n'}$.

Neighborhood realizations are independent of the qualities of the two actions and the realizations of private search costs (to be introduced momentarily).

This framework nests most of the network topologies commonly observed in the data and studied in the literature. Among many others, it accommodates for observation of all previous agents (complete network), random sampling from the past, observation of the most recent $M \geq 1$ individuals, networks with influential groups of agents, and the popular preferential attachment and small-world networks (see Acemoglu et al. (2011) and Lobel and Sadler (2015)).

Search Technology. Private information about the quality of the two actions is acquired through costly sequential search with recall. After observing his neighborhood $B(n)$ and the actions of the agents in $B(n)$, agent n decides which action $s_n^1 \in X$ to sample first.[5] Sampling an action perfectly reveals its quality to the agent. I denote the quality of the first action sampled by agent n as $q_{s_n^1}$. After observing $q_{s_n^1}$, agent n decides whether to sample the remaining action, $s_n^2 = \neg s_n^1$, where $\neg s_n^1$ denotes the action in X not sampled initially, or to discontinue searching, $s_n^2 = ns$. That is, $s_n^2 \in \{\neg s_n^1, ns\}$. Let S_n denote the set of actions agent n samples. After sampling has stopped, the agent chooses an action a_n. Agents can only select an action they sampled, that is $a_n \in S_n$. Thus, for a single agent the model of search is that of Weitzman (1979), and proposed by Mueller-Frank and Pai (2016) to study observational learning in the complete network.

For simplicity, the first action is sampled at no cost, while sampling the second action involves a cost $c_n \in C \subseteq \mathbb{R}_+$.[6] Search costs c_n are i.i.d. across agents, are drawn from a commonly known probability measure $\mathbb{P}_C$ over C, with associated CDF F_C, and are independent of the network topology and the quality of the two actions. I refer to the probability space $(C, \mathcal{F}_C, \mathbb{P}_C)$, together with the sequential search rule, denoted by $\mathcal{R}$, as the *search technology* of the model. An agent's search cost and sampling decisions are his private information. That is, for all $n \in \mathbb{N}$, agent n's search cost c_n and sampling decisions are not observed by later moving agents.

Payoffs. The *net utility* of agent n is given by the difference between the quality of the action he takes and the search cost he incurs. That is,

$$U_n(S_n, a_n, c_n, \omega) := q_{a_n} - c_n(|S_n| - 1).$$

Collective Search Environment. A *collective search environment*, denoted by $\mathcal{S}$, consists of the set of agents $\mathbb{N}$, a state process $(\Omega, \mathcal{F}_\Omega, \mathbb{P}_\Omega)$, a network topology $(\mathbb{B}, \mathcal{F}_\mathbb{B}, \mathbb{Q})$, and a search technology $\{(C, \mathcal{F}_C, \mathbb{P}_C), \mathcal{R}\}$. That is,

$$\mathcal{S} := \{\mathbb{N}, (\Omega, \mathcal{F}_\Omega, \mathbb{P}_\Omega), (\mathbb{B}, \mathcal{F}_\mathbb{B}, \mathbb{Q}), \{(C, \mathcal{F}_C, \mathbb{P}_C), \mathcal{R}\}\}.$$

2.2 Information and Strategies

Each collective search environment $\mathcal{S}$ results in a dynamic game of incomplete information (henceforth, game of social learning). For each agent n, I distinguish three different information sets. The first information set $I^1(n)$ corresponds to n's information prior to sampling any action; it

[5] If neighborhoods are correlated, neighborhood realizations convey information about whom an agent's neighbors are likely to have observed.

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

consists of his search cost c_n, his neighborhood $B(n)$, and all actions of agents in $B(n)$:

$$I^1(n) \coloneqq \{c_n, B(n), a_k \text{ for all } k \in B(n)\}.$$

The set $I^2(n)$ is the information set agent n has after sampling the first action, that is

$$I^2(n) \coloneqq \left\{c_n, B(n), a_k \text{ for all } k \in B(n), q_{s_n^1}\right\},$$

which also includes the quality of the first action sampled. Finally, $I^a(n)$ corresponds to the information set of agent n once his search ends:

$$I^a(n) \coloneqq \{c_n, B(n), a_k \text{ for all } k \in B(n), \{q_s : s \in S_n\}\}.$$

$I^1(n), I^2(n)$, and $I^a(n)$ are random variables whose realizations I denote by I^1_n, I^2_n, and I^a_n. I refer to $I^1(n)$ and $I^2(n)$ as agent n's first and second search stage information sets, and to $I^a(n)$ as agent n's choice stage information set. The classes of all possible search stage and choice stage information sets of agent n are denoted by $\mathcal{I}^r_n$, for $r \in \{1, 2\}$, and $\mathcal{I}^a_n$.

A strategy for agent n is an ordered triple of mappings $\sigma_n \coloneqq (\sigma^1_n, \sigma^2_n, \sigma^a)$ with components

$$\sigma^1_n \colon \mathcal{I}^1_n \to \Delta(\{0, 1\}),$$
$$\sigma^2_n \colon \mathcal{I}^2_n \to \left(\left\{\neg s^1_n, ns\right\}\right),$$

and

$$\sigma^a_n \colon \mathcal{I}^a_n \to \Delta(S_n).$$

A strategy profile is a sequence of strategies $\sigma \coloneqq (\sigma_n)_{n\in\mathbb{N}}$. Let $\sigma_{-n} \coloneqq (\sigma_1, \ldots, \sigma_{n-1}, \sigma_{n+1}, \ldots)$ denote the strategies of all agents other than n. Given a collective search environment $\mathcal{S}$ and a strategy profile σ, the sequence of actions $(a_n)_{n\in\mathbb{N}}$ is a stochastic process with probability measure $\mathbb{P}_\sigma$ generated by the state process, the network topology, the search technology, and the mixed strategy of each agent. Formally, for a fixed σ, the sequence $(a_n)_{n\in\mathbb{N}}$ is determined by the realization in the probability space[7] $Y \coloneqq \Omega \times \mathbb{B} \times C^\infty \times D^\infty$. Here, C^∞ is the set of possible realizations of search costs for each agent, $(D, \mathcal{F}_D, \lambda)$ is a probability space determining the possible mixed strategy realizations of a given agent, and Ω and $\mathbb{B}$ have been introduced before.

2.3 Equilibrium Notion

The solution concept is the set of perfect Bayesian equilibria of the game of social learning.

Definition 1. *Fix a collective search environment $\mathcal{S}$. A strategy profile $\sigma \coloneqq (\sigma_n)_{n\in\mathbb{N}}$ is a* perfect Bayesian equilibrium *of the corresponding game of social learning if, for all $n \in \mathbb{N}$, σ_n is an optimal policy for agent n's sequential search and action choice problems given other agents' strategies σ_{-n}.*

Hereafter, I use the term equilibria to mean perfect Bayesian equilibria. I denote with $\Sigma_{\mathcal{S}}$ the set of equilibria of the game of social learning corresponding to $\mathcal{S}$.

In any collective search environment $\mathcal{S}$, given a strategy profile for the agents acting prior to n, and a realization of n's information sets $I^r_n \in \mathcal{I}^r_n$ for $r \in \{1, 2\}$ and $I^a_n \in \mathcal{I}^a_n$, the decision

[7] Formal notation about the corresponding event space and probability measure is standard, and thus omitted.

problems of agent n at the search stages and at the choice stage are discrete choice problems. Therefore, they have a well-defined solution that only requires randomizing according to some mixed strategy in case of indifference at some stage (see Section 3.2.2 for a characterization of individual equilibrium decisions). For given criteria to break ties, an inductive argument shows that the set of equilibria $\Sigma_{\mathcal{S}}$ is nonempty. I note the existence of equilibrium here.

Proposition 1. *For any collective search environment $\mathcal{S}$, the set of equilibria $\Sigma_{\mathcal{S}}$ is nonempty.*

In general, however, the game of social learning admits multiple equilibria since some agents may be indifferent between the available alternatives at the search or choice stage.

Hereafter, whenever a strategy profile or an equilibrium σ is fixed and no confusion arises, I denote agent n's decisions according to his (equilibrium) strategy $\sigma_n := (\sigma_n^1, \sigma_n^2, \sigma_n^a)$ as

$$s_n^1 := \sigma_n^1, \qquad s_n^2 := \sigma_n^2, \qquad a_n := \sigma_n^a.$$

3 Long-Run Learning and Equilibrium Strategies

In this section, I define asymptotic learning, which is the first long-run learning metric considered in the paper. Then, I characterize equilibrium strategies by relating the dynamics of individual sequential search policies to the dynamics of the probability that agents select the correct action. Finally, I discuss how the availability of social information affects agents' search behavior—what to search and the order in which they do so—and their incentives to acquire independent knowledge.

3.1 Asymptotic Learning: Definition

The first aim of the paper is to characterize conditions on collective search environments under which agents asymptotically select the action with the highest quality with probability one. This represents the most natural benchmark for the social learning process—the same limiting outcome that would occur if each agent directly observed the private search decisions of all prior agents and (at least) one of these agents actually sampled both actions.

Definition 2. *Let a collective search environment $\mathcal{S}$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ be given.* Asymptotic learning *occurs in equilibrium σ if*

$$\lim_{n\to\infty} \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X}\, q_x\right) = 1.$$

Studying asymptotic learning requires understanding how the quantity

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X}\, q_x\right) \tag{1}$$

evolves over time. At the same time, agents use their information to optimize the value of their own sequential search program

$$U_n(S_n, a_n, c_n, \omega) := q_{a_n} - c_n(|S_n| - 1),$$

a problem which need not be equivalent to maximizing the quantity in (1) or the ex ante expected utility.[8] This discrepancy raises some conceptual challenges one needs to address before establishing the main results. To this purpose, the next subsection characterizes equilibrium search policies by linking the dynamics of agents' optimization to the dynamics of the quantity $\mathbb{P}_\sigma(a_n \in \arg\max_{x\in X} q_x)$, thus making the analysis of long-run outcomes possible.

3.2 Equilibrium Strategies

Before characterizing equilibrium strategies, I recall the notion of personal subnetwork from Lobel and Sadler (2015) and introduce the concept of personal subnetwork relative to action $x \in X$.

3.2.1 Preliminaries

Definition 3. *Fix a collective search environment $\mathcal{S}$, a strategy profile σ, and an agent $n \in \mathbb{N}$:*

(a) Agent $m < n$ is a member of agent n's personal subnetwork *if there exists a sequence of agents, starting with m and terminating with n, such that each member of the sequence is contained in the neighborhood of the next. The personal subnetwork of agent n is denoted by $\widehat{B}(n)$.*

(b) Agent $m < n$ is a member of agent n's personal subnetwork relative to action $x \in X$ *if $m \in \widehat{B}(n)$ and $a_m = x$. The personal subnetwork of agent n relative to action $x \in X$ is denoted by $\widehat{B}(n, x)$.*

Agent n's personal subnetwork represents the set of all agents in the network that are connected to n, either directly or indirectly, as of the time n must make a decision. Intuitively, the personal subnetwork of agent n consists of those agents that are, either directly or indirectly (through neighbors, neighbors of neighbors, neighbors of neighbors of neighbors, and so on) observed by agent n. Agent n's personal subnetwork relative to action x consists of those agents that are, either directly or indirectly, observed by agent n to choose action x. Clearly, $\widehat{B}(n) = \widehat{B}(n, 0) \cup \widehat{B}(n, 1)$. Particular realizations of the random variables $\widehat{B}(n)$ and $\widehat{B}(n, x)$ are denoted by $\widehat{B}_n$ and $\widehat{B}_{n,x}$.

3.2.2 Characterization of Equilibrium Sequential Search Policies

Fix a collective search environment $\mathcal{S}$. In the corresponding game of social learning, equilibrium behavior is characterized as follows.

Choice stage. To begin, an agent's optimal policy at the choice stage is mechanical: if he only sampled one action, he takes that action; if he sampled both, he takes the action with the highest quality, randomizing according to his mixed strategy whenever the realized quality of the two actions is the same. Therefore, I omit the formal notation.

To characterize equilibrium search policies, I first consider the search problem of an agent with no social connections and then move to the problem of an agent who observes others' choices.

Search policy for an agent with empty neighborhood. Consider an agent n who does not observe any other agent, that is with $B_n = \emptyset$. This is, for instance, the case of the first agent. Fix

[8] An analogous remark applies to maximal learning, introduced in Section 5.

a strategy profile σ_{-n} for agents other than n. Since an agent's neighborhood is independent of the qualities of the two actions and the choices of previous agents, in the absence of any additional information the marginal distributions of the qualities of the two actions are identical (and equal to the prior $\mathbb{P}_Q$). According to Weitzman (1979)'s optimal search rule, either action might be sampled first. Therefore, the strategy of such agent n is described by two non-negative functions, $\pi_n^0(\cdot)$ and $\pi_n^1(\cdot)$, such that $\pi_n^0(I_n^1) + \pi_n^1(I_n^1) = 1$ for all $I_n^1 \in \mathcal{I}_n^1$ with $B_n = \emptyset$. Here, $\pi_n^x(I_n^1)$ denotes the probability that agent n with information set I_n^1 samples action x first.

Suppose the action agent n samples first, s_n^1, has quality $q_{s_n^1}$. Agent n will only sample the second action if his search cost c_n is smaller than the expected additional gain of sampling the second action, denoted by $t^\emptyset(q_{s_n^1})$, where the function $t^\emptyset \colon Q \to \mathbb{R}_+$ is defined pointwise by

$$t^\emptyset\left(q_{s_n^1}\right) := \mathbb{E}_{\mathbb{P}_Q}\left[\max\left\{q - q_{s_n^1}, 0\right\}\right] = \int_{q \geq q_{s_n^1}} \left(q - q_{s_n^1}\right) \mathrm{d}\mathbb{P}_Q(q). \tag{2}$$

If $c_n = t^\emptyset(q_{s_n^1})$, agent n is indifferent between searching further or not. Again, his strategy is described by two non-negative functions, $\pi_n^{\neg s_n^1}(\cdot)$ and $\pi_n^{ns}(\cdot)$, such that $\pi_n^{\neg s_n^1}(I_n^2) + \pi_n^{ns}(I_n^2) = 1$ for all $I_n^2 \in \mathcal{I}_n^2$ with $B_n = \emptyset$. Here, $\pi_n^{\neg s_n^1}(I_n^2)$ ($\pi_n^{ns}(I_n^2)$) is the probability that agent n with information set I_n^2 samples (does not sample) action $\neg s_n^1 \in X$.[9]

Search policy for an agent with nonempty neighborhood. Consider next an agent n who observes the choices of other agents, that is with $B_n \neq \emptyset$. Fix a strategy profile σ_{-n} for agents other than n. The personal subnetwork of agent n contains conclusive information about the relative quality of the two actions if and only if some agents in the subnetwork have sampled both actions. In particular, consider agent n's conditional belief over the state space Ω given his information set I_n^1. For each action $x \in X$ only two mutually exclusive cases are possible:

1. At least one agent in $\widehat{B}(n,x)$ has sampled both actions. If agent n knew this to be the case, his conditional belief on Ω would be $\mathbb{P}_{\Omega|q_x \geq q_{\neg x}}$, where $\neg x$ denotes the action in X other than x. This is so because agents sampling both actions select the alternative with the highest quality at the choice stage.

2. None of the agents in $\widehat{B}(n,x)$ has sampled both actions. If agent n knew this to be the case, the posterior belief on action $\neg x$ would be the same as the prior $\mathbb{P}_Q$.

To understand the optimal search policy of agent n, consider the probability space $Y := \Omega \times \mathbb{B} \times C^\infty \times D^\infty$ and the following events in Y:

$$E_n^x := \left\{y \in Y : s_k^2 = ns \text{ for all } k \in \widehat{B}(n,x)\right\} \qquad \text{for } x = 0, 1. \tag{3}$$

In words, event E_n^x occurs when none of the agents in the personal subnetwork of agent n relative to action x samples both actions. Let $I_n^1 := \{c_n, B_n, a_k \text{ for all } k \in B_n\}$ be agent n's realized information set prior to sampling any action. Given σ_{-n}, agent n can compute the probabilities

$$P_n(x) := \mathbb{P}_{\sigma_{-n}}\left(E_n^x \mid I_n^1\right) \qquad \text{for } x = 0, 1. \tag{4}$$

[9] Henceforth, I omit the formal notation to describe agents' mixed strategies.

These probabilities allow agent n to rank the marginal distributions of the quality of the two actions in terms of first-order stochastic dominance. If $P_n(0) < P_n(1)$, agent n's belief about the quality of action 0 strictly first-order stochastically dominates his belief about the quality of action 1. Therefore, according to Weitzman (1979)'s optimal search rule, agent n samples first action 0: $s_n^1 = 0$. If $P_n(1) < P_n(0)$, by an analogous argument agent n samples first action 1: $s_n^1 = 1$. Finally, if $P_n(0) = P_n(1)$, the marginal distributions of the quality of the two actions are identical in the eyes of agent n, who then selects the action to sample first according to his mixed strategy.

To formalize the previous argument, pick any $x \in X$ and q with $\min\operatorname{supp}(\mathbb{P}_Q) < q < \max\operatorname{supp}(\mathbb{P}_Q)$, and note that:

$$\mathbb{P}_Q(q_x \leq q) = \mathbb{P}_Q(q_{\neg x} \leq q), \tag{5}$$

$$\mathbb{P}_{\Omega|q_{\neg x}\geq q_x}(q_x \leq q) = \mathbb{P}_{\Omega|q_x\geq q_{\neg x}}(q_{\neg x} \leq q), \tag{6}$$

and

$$\mathbb{P}_{\Omega|q_{\neg x}\geq q_x}(q_x \leq q) > \mathbb{P}_Q(q_x \leq q). \tag{7}$$

Suppose $P_n(x) < P_n(\neg x)$. Conditional on I_n^1, agent n's belief about the quality of action x strictly first-order stochastically dominates his belief about action $\neg x$. In fact,

$$\begin{aligned}
\mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid I_n^1\right) &= \mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid E_n^x, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^x \mid I_n^1\right) \\
&\quad + \mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid E_n^{xC}, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^{xC} \mid I_n^1\right) \\
&= \mathbb{P}_Q(q_{\neg x} \leq q)P_n(x) + \mathbb{P}_{\Omega|q_x\geq q_{\neg x}}(q_{\neg x} \leq q)(1 - P_n(x)) \\
&= \mathbb{P}_Q(q_x \leq q)P_n(x) + \mathbb{P}_{\Omega|q_{\neg x}\geq q_x}(q_x \leq q)(1 - P_n(x)) \\
&> \mathbb{P}_Q(q_x \leq q)P_n(\neg x) + \mathbb{P}_{\Omega|q_{\neg x}\geq q_x}(q_x \leq q)(1 - P_n(\neg x)) \\
&= \mathbb{P}_{\sigma_{-n}}\left(q_x \leq q \mid E_n^{\neg x}, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^{\neg x} \mid I_n^1\right) \\
&\quad + \mathbb{P}_{\sigma_{-n}}\left(q_x \leq q \mid E_n^{\neg xC}, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^{\neg xC} \mid I_n^1\right) \\
&= \mathbb{P}_{\sigma_{-n}}\left(q_x \leq q \mid I_n^1\right).
\end{aligned}$$

Here, E_n^{xC} ($E_n^{\neg xC}$) is the complement of E_n^x ($E_n^{\neg x}$), the third equality holds by (5) and (6), and the inequality follows from (7) and the assumption $P_n(x) < P_n(\neg x)$.

Now, let $I_n^2 \coloneqq \{c_n, B_n, a_k \text{ for all } k \in B_n, q_{s_n^1}\}$ be agent n's realized information set after having sampled a first action of quality $q_{s_n^1}$. Given σ_{-n}, agent n needs to infer the posterior probability that action $\neg s_n^1$ was not sampled by any of the agents in $\widehat{B}(n, s_n^1)$, as only in this case he can benefit from the second search. That is, he must compute

$$P_n\left(q_{s_n^1}\right) \coloneqq \mathbb{P}_{\sigma_{-n}}\left(E_n^{s_n^1} \mid I_n^2\right), \tag{8}$$

where also the information about the quality of the first action sampled is used. With remaining probability, at least one of those agents sampled action $\neg s_n^1$, but nevertheless chose action s_n^1, in which case s_n^1 is (weakly) superior by revealed preferences. Agent n's expected benefit from sampling action $\neg s_n^1$ is therefore $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$, where $t^{\emptyset}(\cdot)$ is defined by (2) and describes the gross benefit of the second search (the benefit agent n would have if he did not observe any other agent) when a payoff of $q_{s_n^1}$ has already been secured. It follows that he should only sample further if his search cost c_n is less than $t_n(q_{s_n^1})$, where the function $t_n\colon Q \to \mathbb{R}_+$ is defined pointwise as

$$t_n\Big(q_{s_n^1}\Big) := P_n\Big(q_{s_n^1}\Big)\, t^{\emptyset}\Big(q_{s_n^1}\Big). \tag{9}$$

If $c_n = t_n(q_{s_n^1})$, agent n is indifferent between searching further and discontinuing search; consequently, he resolves the uncertainty according to his mixed strategy.

Unless noted otherwise, hereafter I assume that agents sample the second action in case of indifference at the second search stage, and that they break ties uniformly at random whenever indifferent at the first search stage or at the choice stage. The assumption is consistent with the idea that agents do not prefer an action over the other because of its label, and that labels do not convey any information about agents' behavior. Selecting a particular equilibrium simplifies the exposition, but the results do not depend on the tie-breaking criterion which is adopted.

3.2.3 Discussion of Equilibrium Behavior

Remark 1. For all $n \in \mathbb{N}$, agent n's equilibrium sequential search policy is essentially described by the probabilities $P_n(x)$ and $P_n(q_x)$, for all $x \in X$ and $q_x \in Q$, defined by (4) and (8). This characterization relates the dynamics of agents' optimization to the dynamics of the probability that they select the correct action. Roughly, the intuition is the following.[10] For all $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X}\, q_x\right) &\geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X}\, q_x\right)\\ &\geq \mathbb{P}_\sigma\Big(\Big\{y \in Y : \exists\, k \in \widehat{B}\Big(n, s_n^1\Big) \text{ such that } s_k^2 = \neg s_k^1\Big\}\Big)\\ &= 1 - \mathbb{P}_\sigma\Big(\Big\{y \in Y : s_k^2 = ns \text{ for all } k \in \widehat{B}\Big(n, s_n^1\Big)\Big\}\Big)\\ &= 1 - \mathbb{P}_\sigma\Big(E_n^{s_n^1}\Big).\end{aligned}$$

Here, the first inequality holds as agent n takes the action of better quality among those he has sampled. The second inequality follows because if an agent in $\widehat{B}(n, s_n^1)$ samples both actions and takes action s_n^1, then s_n^1 is superior by revealed preferences. In turn, the first equality holds as the two events at issue are one the complement of another, and the second equality holds by definition of $E_n^{s_n^1}$ (see (3)). This link unravels the complications illustrated at the end of Section 3.1 and will prove a central tool to establish long-run learning results in the analysis to come.

Remark 2. Each agent faces a three-way trade-off between *exploration* (sampling the second action), *exploitation* (using the information revealed by others' choices to save on the cost of the second search), and *individual incentives* (agents are myopically interested in exploiting the wisdom of their neighbors). The characterization of the optimal search policies sheds light on how such trade-off is resolved in equilibrium.

First, (2) and (9) imply $t_n(q) \leq t^{\emptyset}(q)$ for all $q \in Q$, as $P_n(q) \in [0, 1]$. That is, given the quality of the first action sampled, the expected additional gain from the second search is lower for an agent with nonempty neighborhood than for an agent with empty neighborhood. Thus, if an agent with search cost type c and empty neighborhood discontinues search after sampling an action of quality q, so does an agent with the same search cost type and nonempty neighborhood

[10] I refer to Appendix B for the formal details.

after sampling an action of the same quality. In short, agents with no neighbors have stronger incentives to explore than agents who exploit the information revealed by their neighbors' choices.

Second, for agents with empty neighborhood, the expected additional gain from the second search, and so the incentive to explore, decreases with the quality of the first action sampled: $t^{\emptyset}(q) \leq t^{\emptyset}(q')$ for all $q, q' \in Q$ with $q \geq q'$. Thus, if an agent with search cost type c and empty neighborhood discontinues search after sampling an action of quality q, so does an agent with the same search cost type and empty or nonempty neighborhood after sampling an action of quality $q' \geq q$.

Finally, the quality of the first action sampled has ambiguous effects on the incentives to explore of an agent, say n, with nonempty neighborhood. This is so because n's expected additional gain from the second search, $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$, depends on the probability $P_n(q_{s_n^1})$ that none of the agents in his personal subnetwork relative to action s_n^1 has sampled action $\neg s_n^1$ given that the quality of s_n^1 is $q_{s_n^1}$. This probability need not be monotonic in $q_{s_n^1}$ and depends on the network topology as well as on the properties of the state process and the search technology. On the one hand, an action of high quality suggests that some individual has explored both feasible alternatives, discarding the one with low quality to adopt the superior one. On the other hand, precisely this effect, combined with the fact that $t^{\emptyset}(q)$ decreases in q, hints that the incentives to acquire information about the second action (exploit the information revealed by others' choices) decrease (increase) with the quality of the first action sampled. This is the central trade-off in the environment I study. Which force prevails, and so the effect of an increase in the quality of s_n^1 on $P_n(q_{s_n^1})$ and, ultimately, on $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$, is unclear. In Appendix A, I construct two examples to show that $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$ can either increase or decrease as $q_{s_n^1}$ increases depending on the primitives of the model.

Remark 3. In general network topologies, there is no informational monotonicity property linking an agent's equilibrium behavior to the relative fraction of actions he observes or to the actions of his most recent neighbors. This feature is common in models departing from the assumption that agents observe the full history of past actions and motivates the approach I adopt to establish positive learning results in Section 4.

Remark 4. In the collective search environments I study, there is no social belief that is a martingale and, at the same time, is of some use when characterizing equilibrium behavior. Thus, martingale convergence arguments, which are standard tools to study aggregation of dispersed information in social learning settings, have no bite in the present setup. As I will formalize in Section 5.6, this feature undermines the possibility to learn via the direct observation of large samples of other agents and the aggregation of the information that their choices convey.

4 Asymptotic Learning

In a collective search environment, the search technology shapes agents' possibility to acquire independent private information and the network topology shapes agents' possibility to learn by observing others' behavior. In this section and in Section 5, I provide conditions on these primitives under which (different) positive learning results obtain or fail in the long run.

4.1 Preliminaries

Since the characterization of learning outcomes will hinge on the properties of the search technology, I first present the relevant terminology and assumptions.

Definition 4. *Let* $\{(C, \mathcal{F}_C, \mathbb{P}_C), \mathcal{R}\}$ *be a search technology:*

(a) The search cost $\underline{c}$ *is said to be the* lowest cost in the support of $\mathbb{P}_C$ *if, for all* $\varepsilon > 0$, $F_C(\underline{c} + \varepsilon) > 0$ *and* $F_C(\underline{c} - \varepsilon) = 0$.

(b) Search costs are bounded away from zero *if* $\underline{c} > 0$*; conversely, search costs are* not bounded away from zero *if* $\underline{c} = 0$.

In words, search costs are not bounded away from zero if there is a positive probability of arbitrarily low search costs.

The next assumption is a joint restriction on the state process and the search technology which is maintained throughout the paper. It rules out uninteresting learning problems.

Assumption 1 (**Non-Trivial Collective Search Environment**). *There exist* $\tilde{q}, \tilde{q}'$ *in the support of* $\mathbb{P}_Q$*, possibly with* $\tilde{q} = \tilde{q}'$*, such that:*

1. (a) $\mathbb{P}_Q(q > \tilde{q}) > 0$;

   (b) $1 - F_C\Big(t^{\emptyset}(\tilde{q})\Big) > 0$. *That is, the distribution of search costs is such that, with positive probability, an agent* n *with neighborhood realization* $B_n = \emptyset$ *does not sample another action when the first action sampled has quality* $\tilde{q}$ *or higher.*

2. (a) $\mathbb{P}_Q(q \leq \tilde{q}') > 0$;

   (b) $F_C\Big(t^{\emptyset}(\tilde{q})\Big) > 0$. *That is, the distribution of search costs is such that, with positive probability, an agent* n *with neighborhood realization* $B_n = \emptyset$ *samples another action when the first action sampled has quality* $\tilde{q}'$ *or lower.*

When *Part 1.* of the assumption fails, in equilibrium, an agent with empty neighborhood samples both actions and takes the one with the highest quality, while an agent, say n, with $B_n \neq \emptyset$ just follows the behavior of any of his neighbors. This trivially yields asymptotic learning.When *Part 2.* fails, instead, agents never search in equilibrium: each agent samples the first action at no cost and takes that action. As a result, there is no prospect for social learning since both actions must be sampled by at least one agent in order to evaluate their relative quality. Assumption 1 excludes such trivial search environments.

4.2 Sufficient Conditions

For asymptotic learning to occur it is key that costs are not bounded away from zero. Under this premise, I show that an improvement principle holds in the present setup despite the informational environment significantly differ from that of the SSLM. This is the main contribution of Section 4. Then, I leverage the improvement principle to show that asymptotic learning obtains if, in the network topology, arbitrarily long information paths occur almost surely and are identifiable.[11]

[11]Formally, an information path for agent n is a sequence $(\pi_1, \dots, \pi_k)$ of agents such that $\pi_k = n$ and $\pi_i \in B(\pi_{i+1})$ for all $i \in \{1, \dots k-1\}$.

4.2.1 Improvement Principle

The improvement principle benchmarks the equilibrium performance of Bayesian agents against a heuristic that is simpler to analyze and can be improved upon by rational behavior. This heuristic is based on the idea that an agent always has the option to imitate one of his neighbors and improve upon his outcome. It works as follows. Upon observing who his neighbors are, each agent selects only one neighbor to rely on. After observing the action of his chosen neighbor, the agent determines his optimal search policy regardless of what other neighbors have done. An improvement principle holds if: *(i)* there is a lower bound on the increase in the probability that an agent samples first the best action over his chosen neighbor's probability; in particular, this improvement is strict unless the chosen neighbor already samples the best action with probability one at the first search; *(ii)* the learning mechanism captured by such heuristic and the associated improvements lead to asymptotic learning. For condition *(i)* to hold, it is key that search costs are not bounded away from zero. In turn, condition *(ii)* requires that, in the network topology: *(a)* long information paths occur almost surely, so that improvements last until agents sample the best action with probability one at the first search; *(b)* long information paths are identifiable, so that agents can single out the correct neighbor to rely on.

To establish these results, I recall some notions on network topologies introduced by Lobel and Sadler (2015), to which I refer for further discussion. The first notion is a connectivity property requiring that agents are linked, directly or indirectly, to an unbounded subset of other agents.

Definition 5. *A network topology* $(\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q})$ *features* expanding subnetworks *if, for all positive integers* K*,*

$$\lim_{n\to\infty} \mathbb{Q}\Big(\big|\widehat{B}(n)\big| < K\Big) = 0.$$

The network topology has non-expanding subnetworks *if this property fails.*

Under expanding subnetworks, the size of $\widehat{B}(n)$ grows without bound as n becomes large. This condition rules out, for instance, the presence of an excessively influential group of individuals, that is, the existence of infinite subsequences of agents who, with probability uniformly bounded away from zero, only observe the choices of the same finite set of individuals.

Definition 6. *Let* $(\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q})$ *be a network topology:*

(a) A function $\gamma_n \colon 2^{\mathbb{N}_n} \to \mathbb{N}_n \cup \{0\}$ *is a* neighbor choice function *for agent* n *if, for all neighborhood realizations* $B_n \in 2^{\mathbb{N}_n}$*, we have* $\gamma_n(B_n) \in B_n$ *when* $B_n \neq \emptyset$*, and* $\gamma_n(B_n) = 0$ *otherwise. Given a neighbor choice function* γ_n*, we say that* $\gamma_n(B_n)$ *is agent* n*'s* chosen neighbor*.*

(b) A chosen neighbor topology*, denoted by* $(\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q}_\gamma)$*, is derived from the network topology* $(\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q})$ *and a sequence of neighbor choice functions* $\gamma := (\gamma_n)_{n\in\mathbb{N}}$*. It consists only of the links in* $(\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q})$ *selected by the sequence of neighbor choice functions* $(\gamma_n)_{n\in\mathbb{N}}$*.*

In words, a given neighbor choice function represents a particular way in which agents select a neighbor. A chosen neighbor topology then represents a network topology in which agents discard all observations of the neighbors that are not selected by their neighbor choice function.

The next proposition shows that asymptotic learning via the improvement principle occurs if certain conditions (to be soon clarified) hold. For the rest of this subsection, fix a collective search environment $\mathcal{S} := \{\mathbb{N}, (Q, \mathcal{F}_Q, \mathbb{P}_Q), (\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q}), \{(C, \mathcal{F}_C, \mathbb{P}_C), \mathcal{R}\}\}$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$.

Proposition 2. *Suppose there exist a sequence of neighbor choice functions $(\gamma_n)_{n\in\mathbb{N}}$ and a continuous, increasing function $\mathcal{Z}\colon [1/2,1] \to [1/2,1]$ with the following properties:*

(a) The corresponding chosen neighbor topology features expanding subnetworks;

(b) $\mathcal{Z}(\beta) > \beta$ for all $\beta \in [1/2,1)$, and $\mathcal{Z}(1) = 1$;

(c) For all $\varepsilon, \eta > 0$, there exists a positive integer $N_{\varepsilon\eta}$ such that for all $n > N_{\varepsilon\eta}$, with probability at least $1-\eta$,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n))\right) > \mathcal{Z}\left(\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x\in X} q_x\right)\right) - \varepsilon. \tag{10}$$

Then, asymptotic learning occurs in equilibrium σ.[12]

Importantly, one needs to show that Bayesian agents who do not ignore all but one of the individuals in their neighborhood can at least obtain the improvements described by conditions *(b)* and *(c)* in Proposition 2. While a Bayesian agent has always a higher probability of *sampling* first the best action than an agent following the heuristic described above, the same conclusion does not hold true for the probability of *taking* the best action. This is so because agents use their information to optimize the value of their sequential search program, which is not equivalent to maximizing the ex ante probability of selecting the best action. For this reason, I consider improvements with respect to $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x)$, and not with respect to $\mathbb{P}_\sigma(a_n \in \arg\max_{x\in X} q_x)$, although the ultimate interest is in the evolution dynamics of the latter. However, convergence to one of the probability of sampling first the best action is sufficient for asymptotic learning.

Condition *(c)* in Proposition 2 requires the existence of a strict lower bound on the increase in the probability that an individual will sample first the best action over his chosen neighbor's probability except, possibly, for neighbors that γ_n selects with vanishingly small probability. Therefore, for an improvement principle to hold, one must be able to construct a suitable improvement function $\mathcal{Z}$. The next proposition shows that this is possible if search costs are not bounded away from zero. The intuition goes as follows. Consider an agent, say n, and his chosen neighbor, say $b < n$. Unless b samples first the best action with probability one, b's expected additional gain from the second search is positive. Therefore, if search costs are not bounded away from zero, b samples both actions and compares their quality with positive probability. Thus, as b always takes the best action among those he samples, there is a positive probability that the action he takes is of better quality than the one he samples first. Since n finds it optimal to start searching from the action taken by b,[13] this results in a strict improvement in the probability of sampling first the best action that agent n has over his chosen neighbor b, unless b already does so with probability one, in which case the improvement is non-negative.

Proposition 3. *Suppose that the search technology has search costs that are not bounded away from zero, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence of neighbor choice functions. Then, there exists an*

[12]The probabilities in (10), and in (11) below, are random variables.

[13]It is intuitive, and formally proven in Appendix B.1, that, when agent n only relies on agent b disregarding what other agents have done, the marginal distribution of the quality of the action taken by b first-order stochastically dominates the marginal distribution of the quality of the other action in the eyes of n.

increasing and continuous function $\mathcal{Z}\colon [1/2,1] \to [1/2,1]$, satisfying $\mathcal{Z}(\beta) > \beta$ for all $\beta \in [1/2,1)$, $\mathcal{Z}(1) = 1$, and such that

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \geq \mathcal{Z}\left(\mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right)\right)$$

for all agents n and b with $0 \leq b < n$.

The improvement principle is introduced by Acemoglu et al. (2011) in the SSLM as a tool to establish positive learning results in network topologies with independent neighborhoods. Lobel and Sadler (2015) generalize this principle to networks with arbitrarily correlated neighborhoods. In this paper, I extend the scope of the improvement principle to a new environment, where private information is endogenous and fundamentally distinct in nature, which leads to a different inferential problem on the agents' side. Section 3.2.2, however, shows that the equilibrium sequential search policies are essentially described by the probabilities $P_n(x)$ and $P_n(q_x)$ defined in (4) and (8). This characterization relates the dynamics of individual search behavior to the evolution of the quantity $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x)$. This link makes the agents' inference somewhat comparable to the one faced by the agents in the SSLM, despite the very different premises on the information structure. Thus, an improvement principle which is close in spirit holds.

4.2.2 Sufficient Conditions for Asymptotic Learning

To connect Propositions 2 and 3 into a general result, one needs to bound the difference between $\mathbb{P}_\sigma(s_{\gamma_n}^1 \in \arg\max_{x\in X} q_x)$ and $\mathbb{P}_\sigma(s_{\gamma_n}^1 \in \arg\max_{x\in X} q_x \mid \gamma_n)$. Agent n can imitate agent γ_n only if $\gamma_n \in B(n)$. Therefore, if neighborhoods are correlated, agent γ_n's probability of sampling first the best action conditional on agent n observing agent γ_n is not the same as agent γ_n's probability of sampling first the best action. That is, by imitation, agent n earns γ_n's probability of sampling first the best action *conditional* on n choosing to imitate agent γ_n. If $\mathbb{P}_\sigma(s_{\gamma_n}^1 \in \arg\max_{x\in X} q_x)$ and $\mathbb{P}_\sigma(s_{\gamma_n}^1 \in \arg\max_{x\in X} q_x \mid \gamma_n)$ are approximately the same for large n, then Propositions 2 and 3 immediately imply asymptotic learning. In other words, long information paths must be identifiable, in the sense that agents along the path need reasonably accurate information about the network realization. The next theorem formalizes this last step, which is standard from prior work (see, in particular, Golub and Sadler (2016)).

Theorem 1. *Let a collective search environment $\mathcal{S}$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ be given. Suppose that the following two conditions hold:*

(a) The search technology has search costs that are not bounded away from zero;

(b) In the network topology there exists a sequence of neighbor choice functions $(\gamma_n)_{n\in\mathbb{N}}$ such that the corresponding chosen neighbor topology features expanding subnetworks, and for all $\varepsilon, \eta > 0$, there exists a positive integer N_ε such that for all $n > N_\varepsilon$, with probability at least $1-\eta$,

$$\mathbb{P}_\sigma\left(s_{\gamma_n(B(n))}^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n))\right) > \mathbb{P}_\sigma\left(s_{\gamma_n(B(n))}^1 \in \arg\max_{x\in X} q_x\right) - \varepsilon. \tag{11}$$

Then, asymptotic learning occurs in equilibrium σ.

A variety of conditions on the network topology of $\mathcal{S}$ ensure that (11) holds in every equilibrium $\sigma \in \Sigma_{\mathcal{S}}$. In such cases, if search costs are not bounded away from zero and there exists a chosen neighbor topology with expanding subnetworks, we say that asymptotic learning occurs in the collective search environment $\mathcal{S}$. Such conditions have been identified by Acemoglu et al. (2011) and Lobel and Sadler (2015), to which I refer for further details.

The improvement principle not only serves as a proof technique, but also as a learning principle when standard informational monotonicity properties do not hold (cf. Remark 3 in Section 3.2.3). In particular, it captures the idea that a boundedly rational procedure, imitation, combined with some amount of individual improvement upon it, is sufficient to achieve positive learning outcomes in the long run. The improvement principle also lies behind information diffusion in the SSLM.[14] This explains why, in the search setting I study, asymptotic learning occurs in networks where information diffuses in the SSLM, albeit the mechanics behind the two models significantly differ. Namely, in these network topologies the heuristic captured by the improvement principle displays good long-run properties: first, identifiable information paths allow agents to pick the right neighbor to imitate; second, long information paths allow improvements to last as long as they are possible given the information structure of the model.

Theorem 1 also generalizes MFP's insight that arbitrarily low search costs lead to asymptotic learning from the complete network to a much broader class of observation structures. From a technical viewpoint, however, partial observability of past histories considerably changes the characterization of equilibrium behavior and how positive learning results are obtained.

4.3 A Necessary Condition on Network Topologies

Asymptotic learning requires that agents observe, directly or indirectly, the choices of an unbounded subset of other agents. Thus, asymptotic learning fails with non-expanding subnetworks.

Proposition 4. *Let $\mathcal{S}$ be a collective search environment where the network topology has non-expanding subnetworks. Then, there exists no equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ with asymptotic learning.*

The idea behind Proposition 4 is simple. Asymptotic learning requires that the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to zero as n goes to infinity. Otherwise, there would be a subsequence of agents who: *(i)* with probability bounded away from zero, only observe (directly and indirectly) agents who do not compare the quality of the two actions, as none of the agents in their personal subnetworks samples both actions; *(ii)* do not make this comparison either, as agents in the subsequence do not search for the second alternative. Learning would trivially fail because no agent in the subsequence conclusively assesses the relative quality of the two actions. Now suppose that the network topology has non-expanding subnetworks. By Assumption 1 and the characterization of equilibrium search policies, each single agent, with or without neighbors, does not search for the second action with positive probability independently of which action he samples first. Since non-expanding subnetworks generate with positive probability

[14] In the SSLM, information diffuses if a society asymptotically selects the correct action with the same ex ante probability as an agent with no social information who has access to the most informative private signals. Diffusion captures the idea that the strongest signals spread throughout the network. Information aggregates (asymptotic learning occurs) if, in the long run, agents make the correct choice with probability one. Diffusion is thus a weaker learning requirement than aggregation.

an infinite subsequence of agents, say $\mathcal{N}$, with finite personal subnetwork, the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions remains bounded away from zero for the agents in $\mathcal{N}$. As a result, asymptotic learning fails.

The negative result obtains because infinitely many agents remain uninformed about the relative quality of the two actions with positive probability. The society might well have infinitely many perfectly informed agents, but the result of their searches does not spread over the network.

5 Maximal Learning

In this section, I focus on search costs that are bounded away from zero. First, I define the notion of maximal learning, which is the second long-run learning metric considered in the paper. Second, I explain why the improvement principle breaks down when search costs are bounded away from zero. Then, I characterize a large class of network topologies where maximal learning fails when search costs are bounded away from zero. By means of an example, I show that maximal learning obtains in some special network structures despite zero is not in the support of the search cost distribution. Finally, I discuss why large samples and martingale convergence arguments are of little use in the search setting I study.

5.1 A Motivating Example

When search costs are bounded away from zero, the acquisition of relevant information may be precluded even to agents with the best search opportunities (the lowest search cost type) and the strongest incentives to explore (no social information). In such case, asymptotic learning trivially fails. The next example clarifies the point and suggests that asymptotic learning is not the most suitable learning benchmark when zero is not in the support of the search cost distribution.

Example 1. Suppose that the qualities of the two actions are are drawn uniformly at random from $\{0, 1/2, 1\}$ and that the lowest cost in the support of the search cost distribution is $\underline{c} > 1/6$. With probability 2/9, the realized quality of the two actions is $(q_0, q_1) \in \{(1/2, 1), (1, 1/2)\}$. In such cases, in equilibrium an agent with no neighbors and search cost type $\underline{c}$ never samples the second alternative whatever action he samples first, as his expected additional gain from the second search is at most $1/3(1 - 1/2) = 1/6$, which is smaller than his search cost. However, this agent only samples the best action at the first search with probability 1/2. In turn, agents with a higher search cost type and/or nonempty neighborhood do not sample the second action either, independently of which action they sample first (see Section 3.2). Therefore, when $(q_0, q_1) \in \{(1/2, 1), (1, 1/2)\}$, each agent in the social network makes the wrong choice with positive probability. ■

5.2 Maximal Learning: Definition

Fix a collective search environment $\mathcal{S}$ and let $\underline{c} \geq 0$ be the lowest cost in the support of the search cost distribution of $\mathcal{S}$. Define the threshold quality $q(\underline{c}) := \inf\{q \in Q : t^{\emptyset}(q) < \underline{c}\}$, and let

$$\Omega(\underline{c}) := \{\omega := (q_0, q_1) \in \Omega : q_i \geq q(\underline{c}) \text{ for } i = 0, 1 \text{ and } q_0 \neq q_1\}.$$

Consider a hypothetical agent, say an *expert* located outside of the social network, that wishes to select the best alternative in X. Suppose he has access to the lowest search cost $\underline{c}$ in the support of $\mathbb{P}_C$. In the absence of any social information, this agent selects the correct action whenever $\omega \notin \Omega(\underline{c})$. On the contrary, when $\omega \in \Omega(\underline{c})$, the qualities of the two actions are different, but the agent never searches for the second alternative. In such cases, he makes the correct choice only if he samples first the best action, which happens with probability 1/2.

The next definition introduces the notion of maximal learning, which obtains if agents asymptotically select the best action with the same ex ante probability as an expert.

Definition 7. *Let a collective search environment $\mathcal{S}$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ be given.* Maximal learning *occurs in equilibrium σ if*

$$\liminf_{n\to\infty} \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x\right) \geq \alpha(\underline{c}),$$

where $\alpha(\underline{c}) := 1 - \mathbb{P}_\Omega(\Omega(\underline{c}))/2$.

Equivalently, maximal learning obtains in equilibrium σ if, in the long run, agents select the correct action every time $\min\{q_0, q_1\} < q(\underline{c})$. That is, by defining

$$\overline{\Omega}(\underline{c}) := \{\omega := (q_0, q_1) \in \Omega : q_i \geq q(\underline{c}) \text{ for } i = 0, 1\},$$

maximal learning occurs if

$$\lim_{n\to\infty} \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X}\, q_x \;\middle|\; \omega \notin \overline{\Omega}(\underline{c})\right) = 1. \tag{12}$$

When search costs are not bounded away from zero, maximal learning reduces to asymptotic learning. In contrast, when search costs are bounded away from zero, maximal learning may or may not coincide with asymptotic learning. Example 1 suggests that the two notions are distinct. However, this is not always the case. For instance, if the qualities of the two actions are i.i.d. draws from the discrete uniform distribution over $\{0, 1\}$, and the lowest search cost $\underline{c}$ in the support of $\mathbb{P}_C$ is smaller than 1/2, maximal and asymptotic learning coincide. This is so because an expert with search cost $\underline{c}$ samples the second alternative whenever the first action sampled has quality 0. In general, maximal learning is a weaker requirement than asymptotic learning; it represents the best outcome a society can achieve when zero is not in the support of the search cost distribution.

The next assumption, which parallels Assumption 1, is maintained throughout Section 5.

Assumption 2 (Non-Trivial Collective Search Environment Conditional on $\omega \notin \overline{\Omega}(\underline{c})$). *There exists $\tilde{q}$ in the support of $\mathbb{P}_Q$ such that:*

(a) $\mathbb{P}_Q(\tilde{q} < q < q(\underline{c})) > 0$;

(b) $1 - F_C\left(t^{\emptyset}(\tilde{q})\right) > 0$. *That is, the distribution of search costs is such that, with positive probability, an agent n with neighborhood realization $B_n = \emptyset$ does not sample another action when the first action sampled has quality $\tilde{q}$ or higher.*

Assumption 2 rules out uninteresting learning problems where agents with no neighbors always sample both actions when $\omega \notin \overline{\Omega}(\underline{c})$. If this assumption fails, asymptotic learning trivially obtains for $\omega \notin \overline{\Omega}(\underline{c})$, and never obtains otherwise.

By the same argument establishing Proposition 4, also maximal learning fails when the network topology has non-expanding subnetworks.

Proposition 5. *Let $\mathcal{S}$ be a collective search environment where the network topology has non-expanding subnetworks. Then, there exists no equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ with maximal learning.*

5.3 Failure of the Improvement Principle

If search costs are bounded away from zero, improvements upon imitation are precluded to late moving agents. Thus, maximal (hence, asymptotic) learning via the improvement principle fails.

To formalize the argument, consider a collective search environment $\mathcal{S}$ where the lowest cost in the support of the search cost distribution is $\underline{c} > 0$. Assume that $\omega \notin \overline{\Omega}(\underline{c})$. By way of contradiction, suppose that the improvement principle holds. Then, there must be some chosen neighbor topology derived from the network topology of $\mathcal{S}$ where the probability that none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions converges to zero as n grows large. Therefore, in the chosen neighbor topology there is an infinite subsequence of agents $\mathcal{N}$ where, for a sufficiently late moving agent $m \in \mathcal{N}$, this probability is so small that the expected additional gain from the second search falls below $\underline{c} > 0$, and remains below this threshold afterward. As a result, no agent in $\mathcal{N}$ moving after agent m will sample the second action. At the same time, by Assumption 2, the probability that none of the agents in $\widehat{B}(m) \cup \{m\}$ samples both actions is positive for all finite m. But then, this is a contradiction, as the probability that none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions remains bounded away from zero for the infinite subsequence of agents $\mathcal{N}$.

A perturbation of the search technology breaks down the improvement. In contrast, in the SSLM the strongest available signals, whether bounded or not, are transmitted throughout the network via the improvement principle if long information paths occur almost surely and are identifiable. Therefore, information diffuses and the society performs, in the long run, as well as a single agent with no social information who has access to the most informative signals. This is no longer true in collective search environments: when search costs are bounded away from zero, a society that only relies on improvements upon imitation as a learning principle performs strictly worse than a single agent with no social information and the lowest search cost type.

The improvement principle is not the only method agents may use to learn how to select the correct action. Therefore, it is natural to inquire whether there exist network topologies where maximal learning never obtains (i.e., no matter what agents do in order to learn) when search costs are bounded away from zero. Section 5.5 addresses this question.

5.4 OIP Networks

Before stating the next results, I introduce some notation and define a class of network topologies which will be extensively discussed in Section 6 as well. For all $n \in \mathbb{N}$ and $l_n \in \mathbb{N}_n$, let

$$B_n^{l_n} := \{k \in \mathbb{N}_n : k \geq n - l_n\}$$

be the subset of $\mathbb{N}_n$ comprising the l_n most immediate predecessors of n. For instance: if $l_n = 1$, then $B_n^1 = \{n-1\}$; if $l_n = n-1$, then $B_n^{n-1} = \{1, \ldots, n-1\}$.

Definition 8. *A network topology* $(\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q})$ *features* observation of immediate predecessors *if, for all* $n \in \mathbb{N}$,

$$\mathbb{Q}\Big(\cup_{l_n \in \mathbb{N}_n} \left(B(n) = B_n^{l_n} \right) \Big) = 1.$$

I will often refer to network topologies featuring observation of immediate predecessors as *OIP networks*. These represent a fairly large class of network structures, ranging from deterministic network topologies to stochastic networks with rich correlation patterns between neighborhoods.

Example 2. Here are some examples of OIP networks.

1. If $\mathbb{Q}(B(n) = B_n^{n-1}) = 1$ for all n, we have the complete network.

2. If $\mathbb{Q}(B(n) = B_n^1) = 1$ for all n, we have the network topology where each agent only observes his most immediate predecessor.

3. As an example of stochastic network with independent neighborhoods, consider the following: for all $n \in \mathbb{N}$, $\mathbb{Q}_n(B(n) = B_n^1) = (n-1)/n$ and $\mathbb{Q}_n(B(n) = B_n^{n-1}) = 1/n$. In this case, agents either observe their most immediate predecessor, or all of them, with the latter event becoming less and less likely as n grows large.

4. Stochastic networks with correlated neighborhoods are also possible. For instance: $\mathbb{Q}(B(2) = \{1\}) = 1$, $\mathbb{Q}(B(3) = \{2\}) = 1/2 = \mathbb{Q}(B(3) = \{1, 2\})$, and, for all $n > 3$,

$$B(n) = \begin{cases} \{n-1\} & \text{if } \; B_3 = \{2\} \\ \{1, \ldots, n-1\} & \text{if } \; B_3 = \{1, 2\} \end{cases} . \qquad \blacksquare$$

5.5 Failure of Maximal Learning

When search costs are bounded away from zero, maximal learning fails in all OIP networks and in network topologies where each agent has at most one neighbor (for example, under random sampling of one agent from the past).

Theorem 2. *Let $\mathcal{S}$ be a collective search environment where the search technology has search costs that are bounded away from zero and the network topology satisfies one of the following conditions:*

(a) Observation of immediate predecessors;

(b) $\mathbb{Q}(|B(n)| \leq 1) = 1$ *for all* $n \in \mathbb{N}$.

Then, there exists no equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ with maximal learning.

The intuition behind the result is the following. Suppose that the lowest cost in the support of the search cost distribution is $\underline{c} > 0$ and that $\omega \notin \overline{\Omega}(\underline{c})$. By way of contradiction, assume that maximal learning occurs, so that the probability that none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions converges to zero as n grows large. Then, for a sufficiently late moving agent, say m, this probability is so small that the expected additional gain from the second search falls below

$\underline{c} > 0$ and remains below this threshold afterward. As a result, no agent moving after agent m will sample the second action. At the same time, however, by Assumption 2, the probability that none of the agents in $\widehat{B}(m) \cup \{m\}$ samples both actions is positive for all finite m. But then, this is a contradiction with maximal learning, as the probability that none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions remains bounded away from zero.

The negative result on maximal learning extends beyond the observation structures in Theorem 2. For instance, maximal learning fails in OIP networks if, in addition, agents observe the choices of the first K agents or the aggregate history of past actions (see Sections 6.1 and 6.4); it also fails when each agent n samples $M > 1$ agents uniformly and independently from $\{1, \dots, n-1\}$.

Theorem 2 characterizes a class of networks where, when search costs are bounded away from zero, asymptotic learning fails discontinuously with respect to the benchmark learning metric. In fact, even the second best outcome (maximal learning) breaks down. In contrast, when private beliefs are bounded, in the SSLM information diffuses in network topologies satisfying condition *(a)* or *(b)*. Therefore, in these networks, while asymptotic learning is precluded with bounded private beliefs, the second best learning outcome (diffusion) obtains. This is no longer true in collective search environments. The discontinuity emerges both as an inefficiency due to costly information acquisition as well as a consequence of the information structure, which does not allow agents to learn anything about the relative quality of the two actions unless both are sampled.

Theorem 2 also describes a class of network topologies where search costs that are not bounded away from zero are necessary and sufficient for asymptotic learning. The theorem thus generalizes the characterization result of MFP from the complete network to a larger class of network structures. The novel insight that maximal learning fails as well highlights the fragility of positive learning results with respect to perturbations in the support of the search cost distribution.

5.6 Maximal Learning and the Large-Sample Principle

In this section, I investigate whether there exists some network topology where maximal learning obtains when zero is not in the support of the search cost distribution. For the SSLM, Acemoglu et al. (2011) (see their Theorem 4) characterize a class of network topologies where asymptotic learning obtains with bounded private beliefs. Their findings suggest that maximal learning might occur in some networks despite search costs that are bounded away from zero. The next example shows that this intuition is correct in some very special cases.

Example 3. Let $\mathcal{S}$ be a collective search environment where the lowest cost in the support of the search cost distribution is $\underline{c} > 0$. Assume that the network topology satisfies, for all $n \in \mathbb{N}$,

$$\mathbb{Q}(B(n) = \emptyset) = p_n \qquad \text{and} \qquad \mathbb{Q}(B(n) = \{m \in \mathbb{N}_n : B(m) = \emptyset\}) = 1 - p_n,$$

where the sequence $(p_n)_{n\in\mathbb{N}}$ is such that $0 \leq p_n \leq 1$ for all n, $\lim_{n\to\infty} p_n = 0$, and $\sum_{n=1}^{\infty} p_n = \infty$. That is, agent n has empty neighborhood with probability p_n, or observes all and only his predecessors with empty neighborhood with probability $1 - p_n$.

Suppose $(q_0, q_1) \notin \overline{\Omega}(\underline{c})$ and, without loss, $q_0 > q_1$. Consider first an agent, say k, with $B(k) = \emptyset$. By definition of $\overline{\Omega}(\underline{c})$ and $\underline{c}$, k samples the second action with positive probability when he samples action 1 first. Hence, k takes the correct action ($a_k = 0$) with probability $\alpha > 1/2$.[15]

Now consider an agent, say l, with $B(l) \neq \emptyset$. By the assumptions on the network topology, agent l only observes the choices of all his predecessors with empty neighborhood. Thus, l's optimal decision at the first search stage depends on the relative fraction of choices he observes. In particular:

$$s_l^1 = \begin{cases} 0 & \text{if } \left|\widehat{B}(l,0)\right| > \left|\widehat{B}(l,1)\right| \\ 1 & \text{if } \left|\widehat{B}(l,0)\right| < \left|\widehat{B}(l,1)\right| \end{cases},$$

and $s_l^1 \in \Delta(\{0,1\})$ if $|\widehat{B}(l,0)| = |\widehat{B}(l,1)|$. To see this, note that $|\widehat{B}(l,x)| > |\widehat{B}(l,\neg x)|$ immediately implies $P_l(x) < P_l(\neg x)$, where $P_l(\cdot)$ is the probability defined by (4).

The assumptions on $(p_n)_{n\in\mathbb{N}}$ imply that $\lim_{n\to\infty} \mathbb{Q}(|\widehat{B}(n)| < K) = 0$ for all positive integers K. Hence, with probability one, there are infinitely many agents with no social information. Moreover, the actions taken by the agents with empty neighborhood form a sequence of independent random variables. Thus, by the weak law of large numbers, the ratio $|\widehat{B}(l,0)|/|\widehat{B}(l,0)|$ converges in probability to $\alpha > 1/2$ as $l \to \infty$ (with respect to $\mathbb{P}_\sigma$, and conditional on $\widehat{B}(l) \neq \emptyset$). Therefore,

$$\lim_{l\to\infty} \mathbb{P}_\sigma\Big(\left|\widehat{B}(l,0)\right| > \left|\widehat{B}(l,1)\right| \;\Big|\; \widehat{B}(l) \neq \emptyset\Big) = 1. \tag{13}$$

Finally, for all $n \in \mathbb{N}$, note that

$$\begin{aligned} 1 &\geq \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x \;\middle|\; \omega \notin \overline{\Omega}(\underline{c})\right) \\ &\geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \omega \notin \overline{\Omega}(\underline{c})\right) \\ &= \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; B(n) = \emptyset, \omega \notin \overline{\Omega}(\underline{c})\right)\mathbb{Q}(B(n) = \emptyset) \\ &+ \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; B(n) \neq \emptyset, \omega \notin \overline{\Omega}(\underline{c})\right)\mathbb{Q}(B(n) \neq \emptyset) \\ &\geq \frac{1}{2}p_n + \mathbb{P}_\sigma\Big(\left|\widehat{B}(n,0)\right| > \left|\widehat{B}(n,1)\right| \;\Big|\; \widehat{B}(l) \neq \emptyset\Big)(1-p_n). \end{aligned} \tag{14}$$

Here, the second inequality holds as agent n takes the action of better quality among those he has sampled; the first equality holds by the law of total probability; the third inequality follows by the properties of the network topology, the fact that $q_0 > q_1$, the assumption that agents with no neighbors select uniformly at random the action to sample first, and the optimal policy at the first search stage for agents with nonempty neighborhood.

By (13), and since $\lim_{n\to\infty} p_n = 0$, we have

$$\lim_{n\to\infty}\left[\frac{1}{2}p_n + \mathbb{P}_\sigma\Big(\left|\widehat{B}(l,0)\right| > \left|\widehat{B}(l,1)\right|\Big)(1-p_n)\right] = 1. \tag{15}$$

[15] Agent k takes the correct action any time he samples first action 0, which occurs with probability 1/2, and any time he samples first action 1 and his search cost is smaller that $t^{\emptyset}(q_1)$. Since $q_0 > q_1$ and $(q_0,q_1) \notin \overline{\Omega}(\underline{c})$, $q_1 < q(\underline{c})$, and so the latter event occurs with positive probability. Therefore, the overall probability that agent k takes action 0 is larger than 1/2. Providing an expression for α is irrelevant for the following argument.

Together, (14) and (15) imply

$$\lim_{n\to\infty} \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x \;\middle|\; \omega \notin \overline{\Omega}(\underline{c})\right) = 1,$$

showing that maximal learning occurs. ■

The positive result in Example 3 relies on the assumption that agents with nonempty neighborhood *only* observe agents with no social information. Under this premise, the optimal policy at the first search stage for the former group of agents is determined by the relative fraction of choices they observe. When agents with nonempty neighborhood observe more, however, connecting the optimal search policy to the ratio of observed choices is no longer possible. Therefore, it is unclear whether (and to what extent) the insight of Example 3 extends to a more general characterization.

The positive results in Acemoglu et al. (2011) make an extensive use of large samples and martingale convergence arguments, which have no bite in collective search environments (see Remark 4). These arguments are commonly referred to as the *large-sample principle* and capture the idea that agents learn by aggregating the information contained in a large sample of others' choices. The scope of the large-sample principle is severely hampered in the present environment, emphasizing once more the distinction between the inferential challenge in the search setting I study and that in the SSLM. Therefore, if any characterization of networks where maximal learning occurs despite $\underline{c} > 0$ is within reach, it requires a different line of attack.

Recall that maximal and asymptotic learning sometimes coincide despite search costs are bounded away from zero (see Section 5.2). Thus, Example 3 also shows that asymptotic learning may occur when zero is not in the support of the search cost distribution. In other words, search costs that are not bounded away from zero are not, in general, necessary for asymptotic learning.

6 Rate of Convergence, Welfare, and Efficiency

In this section, I present results on the probability of wrong herds forming, the rate of convergence, equilibrium welfare, and efficiency. I also discuss simple policy interventions that enhance welfare in equilibrium. Most of the analysis will focus on OIP networks. Thus, I begin by describing equilibrium behavior in this class of network topologies.

6.1 Equilibrium Strategies in OIP Networks

Fix a state process and a search technology. From the viewpoint of the probability of selecting the best action, equilibrium behavior is equivalent across OIP networks. To illustrate the argument, I first introduce some terminology.

Definition 9. *Let $\mathcal{S}$ be a collective search environment where the network topology features observation of immediate predecessors, and let $\sigma \in \Sigma_{\mathcal{S}}$. We say:*

(a) Action $x \in X$ is revealed to be inferior to agent n *in equilibrium σ if there exist agents $j, j+1 \in B(n)$ such that $a_j = x$ and $a_{j+1} = \neg x$.*

(b) Action $x \in X$ is revealed to be inferior by time *n in equilibrium σ if there exist agents $j, j+1 \in \mathbb{N}$, with $j+1 < n$, such that $a_j = x$ and $a_{j+1} = \neg x$.*

(c) Action $x \in X$ is inferior by time *n in equilibrium σ if there exists an agent $j \in \mathbb{N}$, with $j < n$, who has sampled both actions and such that $a_j = \neg x$.*

If an action is revealed to be inferior to agent n in equilibrium σ, then it is also revealed to be inferior by time n in the same equilibrium. The converse statement is not generally true, but it is so in the complete network, where $B(n) = \{1, \ldots, n-1\}$ with probability one for all n.

In OIP networks, agent $n \geq 2$'s equilibrium behavior is the following.[16] At the first search stage, agent n samples the action taken by his immediate predecessor: $s_n^1 = a_{n-1}$. This is so because n's belief about the quality of action a_{n-1} strictly first-order stochastically dominates his belief about the quality of the other action (the result follows by induction). Hence, if an action is revealed to be inferior by time n in equilibrium σ, it is also inferior by time n in equilibrium σ (the converse statement is not, in general, true).

At the second search stage, the optimal policy depends on whether action $\neg s_n^1$ is revealed to be inferior to agent n in equilibrium or not. If action $\neg s_n^1$ is revealed to be inferior to agent n, then n discontinues search and takes action s_n^1. The reason for not sampling $\neg s_n^1$ is straightforward. Suppose there are agents $j, j+1 \in B(n)$ such that $a_j = \neg s_n^1$ and $a_{j+1} = s_n^1$. Since agents start sampling from the action taken by their immediate predecessor, agent $j+1$ must have sampled action $\neg s_n^1$ first, and therefore would only select $a_j = s_n^1$ at the choice stage if he then sampled action s_n^1 as well, and $q_{s_n^1} \geq q_{\neg s_n^1}$. That is, action $\neg s_n^1$ is revealed to be inferior to action s_n^1 by agent $j+1$'s choice, and so the expected additional gain from the second search is zero. If instead action $\neg s_n^1$ is not revealed to be inferior to agent n, the expected additional gain from the second search given quality $q_{s_n^1}$ is the same as in the complete network for an action of the same quality that is not revealed to be inferior by time n in equilibrium. The intuition goes as follows. In all OIP networks agent n's personal subnetwork is the same, that is $\{1, \ldots, n-1\}$, and coincides with agent n's neighborhood in the complete network. Moreover, all agents start sampling from the action taken by their most immediate predecessor. Thus, given $q_{s_n^1}$, the probability that none of the agents in n's personal subnetwork relative to s_n^1 has sampled both actions must be the same. But then, if s_n^1 is not revealed to be inferior to agent n, this agent adopts the same threshold he would use in the complete network to determine whether to search further or not after having sampled an action of the same quality that is not revealed to be inferior in equilibrium by time n.

Remark 5. Fix a state process and a search technology. By the previous argument, the following equilibrium objects are identical across OIP networks: the order of search; the cutoff for sampling a second action that is not revealed to be inferior to an agent; the probability that each agent n selects the best action. Then:

(a) In OIP networks, the density of connections and their correlation pattern do not affect equilibrium inference and several equilibrium outcomes.

(b) Many equilibrium properties of the game of social learning in the complete network immediately extend to all OIP networks. I will explore this insight in the next subsections.

[16] I refer to Appendix B.3 for the formal characterization.

Remark 6. In all OIP networks actions are always improving; that is, each agent takes a weakly better action than his predecessors.

These properties distinguish the search environment I study from the SSLM, where equilibrium dynamics dramatically change as the number of immediate predecessors that are observed varies. For instance, Celen and Kariv (2004) study the SSLM under the assumption that each agent only observes his most recent predecessor's action and show that beliefs and actions cycle indefinitely.

6.2 Probability of Wrong Herds and Rate of Convergence

OIP Networks. Fix a state process and a search technology. Remark 5 implies the following.

Remark 7. In all OIP networks:

(a) The probability of wrong herds forming is the same as in the complete network;

(b) If search costs are not bounded away from zero, so that asymptotic learning occurs, the rate of convergence is the same as in the complete network.

Consider first the probability of suboptimal herds. The next proposition says that we can bound this probability as a linear function of the lowest cost in the support of the search cost distribution. The result holds by combining Remark 7–*(a)* with Proposition 1 in MFP.

Proposition 6. *Let $\mathcal{S}$ be a collective search environment where the network topology features observation of immediate predecessors, and let $\underline{c}$ be the lowest cost in the support of $\mathbb{P}_C$. Then, in any equilibrium $\sigma \in \Sigma_{\mathcal{S}}$, the quantity*

$$\underline{c}\mathbb{E}_Q\left[\frac{1}{t^{\emptyset}(q_0)}\,\middle|\, q_0 < q_1\right]$$

is an upper bound for the probability of a suboptimal herd forming.

By Proposition 6, the probability that agents asymptotically select the correct action converges to one as $\underline{c}$ approaches zero. Despite this "continuity" result, however, the probability of wrong herds may remain sizable if search costs are bounded away from zero. This is so even when maximal and asymptotic learning coincide, as the next example shows.

Example 4. Suppose the network topology features observation of immediate predecessors. Assume that the qualities of the two actions are drawn uniformly at random from $\{0, 1\}$, and that search costs are drawn from $\{1/2, 2/3\}$, with $\mathbb{P}_C(c = 1/2) = \delta$ and $\mathbb{P}_C(c = 2/3) = 1 - \delta$ for some $\delta \in (0, 1)$. To simplify the exposition, assume that agents sample the other action in case of indifference at the second search stage. For an agent with no neighbors, the expected additional gain from a second search after sampling an action of quality 0 is $1/2 = \underline{c}$. Thus, maximal and asymptotic learning coincide, as an expert would always select the best action.

With probability 1/2, $(q_0, q_1) \in \{(0, 1), (1, 0)\}$. In such cases, agent 1 selects the best action with probability $(1+\delta)/2$. Therefore, the ex ante probability that agent 1 selects a wrong action is $(1-\delta)/4$. Moreover, the expected additional gain from a second search for agent 2 (and for all his successors) after sampling an action of quality 0 is smaller than $1/2 = \underline{c}$, as agent 1 samples both actions with positive probability. Therefore, no agent moving after agent 1 samples both action.

Thus, a suboptimal herd forms whenever agent 1 selects the wrong action. As δ approaches zero, the latter event occurs with probability arbitrarily close to $1/4$. ■

Next, consider the rate of converge. I begin by introducing an important property of search cost distributions that will affect the results on the speed of learning.

Definition 10. *Let $(Q, \mathcal{F}_Q, \mathbb{P}_Q)$ be a state process and $\{(C, \mathcal{F}_C, \mathbb{P}_C), \mathcal{R}\}$ a search technology. Set $\underline{q} := \min \operatorname{supp}\left(\mathbb{P}_Q\right)$. The search cost distribution has* polynomial shape *if there exist some real constants K and L, with $K \geq 0$ and $0 < L < \frac{2^{K+1}}{(K+2)t^{\emptyset}\left(\underline{q}\right)^K}$, such that*

$$F_C(c) \geq Lc^K \qquad \text{for all } \; c \in \left(0, t^{\emptyset}\left(\underline{q}\right)/2\right).$$

Convergence to the correct action is faster than a polynomial rate in OIP networks.

Proposition 7. *Let $\mathcal{S}$ be a collective search environment where the network topology features observation of immediate predecessors. Suppose also that search costs are not bounded away from zero.*

(a) If $\mathbb{P}_C$ admits a density f_C, and $f_C(0) > 0$, then in any equilibrium $\sigma \in \Sigma_{\mathcal{S}}$,

$$\mathbb{P}_\sigma\left(a_n \notin \underset{x \in X}{\arg\max}\ q_x\right) = O\left(\frac{1}{n}\right)$$

for n sufficiently large.

(b) If the search cost distribution has polynomial shape, then in any equilibrium $\sigma \in \Sigma_{\mathcal{S}}$,

$$\mathbb{P}_\sigma\left(s_n^1 \notin \underset{x \in X}{\arg\max}\ q_x\right) = O\left(\frac{1}{n^{\frac{1}{K+1}}}\right).$$

Part *(a)* holds by combining Remark 7–*(b)* with Proposition 1 in MFP. The proof of part *(b)* (and Proposition 8 below) builds on a technique developed by Lobel, Acemoglu, Dahleh and Ozdaglar (2009) to characterize the speed of learning in the SSLM. This technique consists in approximating a lower bound on the rate of convergence with an ordinary differential equation.

Random Sampling from the Past. Convergence occurs at a logarithmic rate under random sampling of one agent from the past. Thus, the speed of learning is slower than in OIP networks. Intuitively, this is so because the cardinality of agents' personal subnetworks grows at a slower rate than in OIP networks, and so does the probability that at least one agent in the personal subnetworks has sampled both actions.

Proposition 8. *Let $\mathcal{S}$ be a collective search environment where the network topology has independent neighborhoods and is such that $\mathbb{Q}_n(|B(n)| = 1) = 1$ for all $n \in \mathbb{N}$. Moreover, assume that the search cost distribution has polynomial shape. Then, in any equilibrium $\sigma \in \Sigma_{\mathcal{S}}$,*

$$\mathbb{P}_\sigma\left(s_n^1 \notin \underset{x \in X}{\arg\max}\ q_x\right) = O\left(\frac{1}{(\log n)^{\frac{1}{K+1}}}\right).$$

6.3 Equilibrium Welfare and Efficiency in OIP Networks

In this section, I first characterize how transparency of past histories affects equilibrium welfare. Then, I compare equilibrium welfare against the efficiency benchmark where agents are replaced by a single decision maker. To aid analysis, I assume throughout this section that the probability measure $\mathbb{P}_C$ admits a density f_C, and that $f_C(\underline{c}) > 0$.

Equilibrium Welfare across OIP Networks. Despite in all OIP networks later moving agents take a weakly better action than their predecessors,[17] equilibrium welfare is not the same across OIP networks. To see this, suppose there exist agents $j, j+1 \in \mathbb{N}$ such that $a_j = x$ and $a_{j+1} = \neg x$. Therefore, action x is revealed to be inferior by time $j+2$ in equilibrium. In the complete network, action x is revealed to be inferior to any agent $n \geq j+2$, and so it is never sampled again. In other OIP networks, instead, agent j is not necessarily in the neighborhood of agent $n \geq j+2$, and therefore n fails to realize from agent $j+1$'s choice that action x is of lower quality than action $\neg x$. Thus, agent n inefficiently samples action x with positive probability at the second search stage.[18]

This kind of inefficient duplication of costly search is more severe the shorter in the past agents can observe. Therefore, the complete network is the most efficient OIP network, and the network where agents only observe their most recent predecessor is the least efficient in this class. In all other OIP networks, equilibrium welfare is comprised between these two bounds.

The next proposition shows that welfare losses arising because agents fail to recognize actions that are revealed to be inferior by the time of their move only vanish in the limit of an arbitrarily patient society (equivalently, in the long run). These losses, however, remain significant in the short and medium run. To ease the statement of the result, let $\mathcal{S}$ and $\mathcal{S}'$ be two collective search environments with identical state process and search technology. Suppose that the network topology of $\mathcal{S}$ is the complete network and that in $\mathcal{S}'$ agents only observe their most immediate predecessor. Let $\sigma \in \Sigma_{\mathcal{S}}$ and $\sigma' \in \Sigma_{\mathcal{S}'}$ and suppose that agents break ties according to the same criterion in σ and σ'. Assume that future payoffs are discounted at rate $\delta \in (0, 1)$.

Proposition 9. *For all $\delta \in (0, 1)$, the average social utility in equilibrium σ is larger than the average social utility in equilibrium σ'. This difference vanishes as δ goes to one.*

The Single Decision Maker Benchmark. Suppose that agents are replaced by a single decision maker (social planner) who has the same search technology available to the agents, draws a new search cost in each time period, and faces the same structure of connections as the agents in the society. The social planner discounts future payoffs at rate $\delta \in (0, 1)$, internalizes future gains of today's search, and needs to sample each of the two actions exactly once along the same information path. Since in OIP networks each agent is (directly or indirectly) linked to all his predecessors, all agents lie on the same (and unique) information path. Therefore, the social planner

[17] This property is lost in general network topologies, where agents may generate long patterns of disagreement before settling on one action. Disagreement, however, does not necessarily impact welfare in a negative way, as it may foster exploration and speed up convergence to the right action.

[18] For the descriptive analysis in this section, assume that search costs are not bounded away from zero. The formal details are in Appendix B.6.

achieves the same average social utility in all collective search environments with the same state process and search technology, but where the network topology is any OIP network.[19]

Equilibrium behavior in OIP networks gives rise to two potential sources of inefficiency:

(i) The single decision maker internalizes future gains of today's search, while agents are myopic. As a result, exploration and convergence to the right action is too slow in equilibrium.

(ii) The single decision maker has more information than the agents in equilibrium and samples each of the two actions exactly once. By contrast, in equilibrium:

(a) Each agent n fails to recognize an action, say x, that is inferior, and not revealed to be so, by time n. Therefore, agents sample action x multiple times.

(b) Each agent n fails to recognize an action, say x, that is revealed to be inferior by time n, i.e. such that $a_j = x$ and $a_{j+1} = \neg x$ for some agents $j, j+1$, with $j+1 < n$, unless $j, j+1 \in B(n)$. Again, agents sample action x multiple times.

As a result, equilibrium behavior displays inefficient duplication of costly search. Note that, while *(a)* occurs in all OIP networks, *(b)* does not in the complete network.

Equilibrium welfare losses disappear in the long run if and only if asymptotic learning occurs. If search costs are bounded away from zero, or if the focus is on short- and medium-run outcomes, the average social utility in equilibrium is lower than under the social planner.

Proposition 10. *Let $\mathcal{S}''$ be collective search environment where the network topology features observation of immediate predecessors. Then, the average social utility in any equilibrium $\sigma'' \in \Sigma_{\mathcal{S}''}$ converges to the average social utility implemented by the single decision maker as δ goes to one if and only if search costs are not bounded away from zero.*

Discussion of Probability of Wrong Herds, Rate of Convergence, and Welfare. The results in Sections 6.2 and 6.3 are surprising for two reasons. First, in OIP networks the probability of wrong herds, the speed of learning, and the long-run (but not short-run) welfare neither depend on transparency of past histories nor on the correlation structure among connections. Second, the rate of convergence can be characterized for a large class of networks. This contrasts with our understanding of the SSLM, for which little is known about convergence rates unless all agents observe the most recent action, a random action from the past, or all past actions (see Lobel et al. (2009), Rosenberg and Vieille (2017), and Hann-Caruthers, Martynov and Tamuz (2018)).

Rosenberg and Vieille (2017) consider two measures of the efficiency of social learning in the SSLM: the expected time until the first correct action and the expected number of incorrect actions (see also Hann-Caruthers et al. (2018)). They focus on two polar setups and assume that each agent either observes the entire sequence of earlier actions or only the previous one. In a similar spirit with my results, they find that whether learning is efficient is independent of the setup: for every signal distribution, learning is efficient in one setup if and only if it is efficient in the other one. In the search setting I study, the results on the irrelevance of how far in the past agents can observe is much stronger: first, it holds for the long-run welfare as well as for the probability of

[19] I refer to Section III.A. in MFP for the solution to the single decision maker's problem in the complete network. As the single decision maker's problem is the same in all OIP networks, their analysis applies unchanged to my setting.

wrong herds and the speed of learning; second, it neither depends on the number of immediate predecessors that agents observe nor on the dependence structure among connections.

6.4 Policy Interventions

Reducing transparency of past histories in OIP networks leads to inefficient duplication of costly search. Straightforward policy interventions, however, can improve efficiency and equilibrium welfare in the short and medium run.

Let $\mathcal{S}$ and $\mathcal{S}'$ be two collective search environments with identical state process and search technology. Assume $\mathcal{S}$ is endowed with the complete network, and let $\sigma \in \Sigma_{\mathcal{S}}$. Suppose that $\mathcal{S}'$ is endowed with any OIP network and that each agent in $\mathcal{S}'$, in addition to the actions of his neighbors, observes the aggregate history of past actions or the action of the first agent (or both). Let σ' be an equilibrium of the game associated to $\mathcal{S}'$ and suppose that agents break ties according to the same criterion in σ and in σ'. Then, we have the following.

Proposition 11. *For all $\delta \in (0,1)$, the average social utility in equilibrium σ' is the same as the average social utility in equilibrium σ.*

Suppose agents observe, in addition to the actions of their neighbors, the relative fraction of past actions or the action of the first agent. Then, according to Proposition 11, in all OIP networks equilibrium welfare is the same as in the complete network (the most efficient network in this class). The intuition behind the result is simple. First, observing the action of the first agent or the aggregate history of past actions (or both) does not change equilibrium behavior at the first search stage: in σ', each agent starts sampling from the action taken by his immediate predecessor. Second, if an action is revealed to be inferior by time n in equilibrium σ', that action is never sampled again by any agent $m \geq n$. To see this, suppose that there exist agents $j, j+1 \in \mathbb{N}$ such that $a_j = x$ and $a_{j+1} = \neg x$, and consider any agent $n > j+1$. Agent n samples first action a_{n-1}. Since each agent starts sampling from the action taken by his immediate predecessor and takes the action of better quality, it must be that $a_{n-1} = \neg x$. Now, if agent n observes the choice of the first agent or the aggregate history of past actions, he realizes that $q_{\neg x} \geq q_x$ even when $j \notin B(n)$. In fact, when n observes $a_1 = x$ and $a_{n-1} = \neg x$, he correctly infers that some agent $j+1$, with $1 \leq j \leq n-2$, has sampled both actions and discarded the inferior action x. Therefore, n stops searching and takes action $\neg x$. The same inference is possible when agent n observes the aggregate history of past choices. In this case, n would observe that j agents have taken action x, while $n-j-1$ agents have taken action $\neg x$. Together with $a_{n-1} = \neg x$, this implies that $a_1 = x$ and that some agent $j+1$, with $1 \leq j \leq n-2$, has sampled both actions and discarded the inferior action x. Therefore, the duplication of costly search that would arise because agents fail to recognize actions that are revealed to be inferior by time n disappears.

Interestingly, the remedies that this section suggests are easy to implement and commonly observed in practice. For instance, online platforms that aggregate past individual choices by sorting different items according to their popularity or sales rank serve the purpose.[20]

[20] Letting agents observe the aggregate history of past actions or the action of the first agent are effective policy interventions in network topologies other than OIP networks (e.g., under random sampling of one agent from the past). The analysis of such cases, however, goes beyond the scope of the paper.

The interventions discussed above, however, do not remove the inefficient duplication of costly search arising when agents fail to recognize actions that are inferior (and not revealed to be so) by some time n. Moreover, they do not incetivize exploration; thus, agents delay search for the second action more than what the single decision maker would do and the rate of convergence remains too slow. A natural step for future research is to understand how and to what extent more complex incentive schemes, which make use of monetary transfers or information management tools, can reduce these other inefficiencies as well.[21]

7 Related Literature and Concluding Remarks

7.1 Related Literature

This paper joins a small but growing literature on costly acquisition of private information in social learning settings. Burguet and Vives (2000) and Chamley (2004) consider a continuum of agents, each choosing an action from a continuous space in every period. Agents wish to match an unknown state of nature in order to minimize a quadratic loss and set the precision of a normally distributed signal at a cost that increases with the signal's precision. In Ali (2018) there is an unknown binary state of nature. Agents select an action from a space, either discrete or continuous, and aim at taking higher actions in the higher state. They act in sequence, observe the choices of all their predecessors, and choose how informative a signal to acquire at a cost which depends on the chosen informativeness about the relative likelihood of the two states. These costs are heterogeneous across agents and are private information. While I focus on some search cost types obtaining perfect signals in a discrete action space, these papers study noisy signals with a continuous (or general, in Ali (2018)) action space. Closer to my setup, Hendricks, Sorensen and Wiseman (2012) study sequential learning when agents choose whether to purchase a product or not. Agents have heterogeneous preferences, which are private information, but identical search costs. At this cost, they can acquire a perfect signal about their value for the product. In their model, however, agents only observe the aggregate purchase history.[22]

My model departs from these papers in two relevant ways. First, I consider a game of social learning which is played over general networks. Thus, I provide conditions on both information acquisition technologies and observation structures that lead to positive or negative learning results. Second, the parametric structure of private information is substantially different. These distinctions require different tools to analyze the social learning process and prevent a direct comparison of the results. A general insight of Ali (2018) and MFP is that, in the complete network, we can trade an assumption of arbitrarily strong exogenous private signals for an assumption of arbitrarily low information acquisition costs. My work shows that this insight generalizes to all network topologies where long information paths occur almost surely and are identifiable.

My model relates to those of sequential information acquisition of Wald (1947), Weitzman

[21] A recent and growing literature in economics and computer science, including Smith, Sørensen and Tian (2017), Kremer, Mansour and Perry (2014), Che and Hörner (2018), Papanastasiou, Bimpikis and Savva (2018), Mansour, Slivkins and Syrgkanis (2015), and Mansour, Slivkins, Syrgkanis and Wu (2016), studies optimal design in the SSLM and other related sequential social learning environments.

[22] Relatedly, Huang (2017) investigates theoretically and empirically the interplay between observational learning and costly information acquisition.

(1979), and Moscarini and Smith (2001), where a single decision maker dynamically chooses how much information to acquire before taking an action. Weitzman (1979) considers a sequential search environment where an agent faces a bandit problem, each arm representing a distinct alternative with a random prize, and characterizes the optimal sampling sequence and the optimal timing to stop the search process. Each agent in my model faces the same problem and trade-off between exploration (sampling the second action) and exploitation (taking the action believed to be the best according to his social information).[23]

Salish (2017) and Sadler (2017) study learning in networks where a finite number of agents acquire private information by strategic experimentation with a two-armed bandit, as in Keller, Rady and Cripps (2005) and Bolton and Harris (1999), and observe the experimentation of their neighbors.[24] In these models, agents interact repeatedly over time, and so the strategic component of their interaction is more involved than in my setting. However, this comes at a cost. Sadler (2017) allows for complex network structures, but agents follow a boundedly rational decision rule. In Salish (2017) agents are rational, but a sharp characterization only obtains for particular network structures. Taking advantage of the sequential nature of the problem, I accommodate both for rational behavior and general network topologies. In a similar spirit, Perego and Yuksel (2016) study a model of learning where a continuum of Bayesian agents repeatedly choose between learning from one's own experimentation or learning from others' experiences. Connections are heterogeneous across agents and peer-to-peer exchange of information is subject to frictions. The authors characterize how frictions and heterogeneity in connections affect the creation and diffusion of knowledge in equilibrium, but do not focus on network properties other than connectivity.

A few papers consider costly observability of past histories in the SSLM (e.g., Kultti and Miettinen (2006, 2007), Song (2016), Nei (2016), and, in an experimental setting, Celen and Hyndman (2012)). In these papers private information is free, while which agents' actions to observe is endogenously determined. In contrast, I study costly acquisition of private information in exogenous network structures.

The literature on social learning in networks is larger than the work surveyed here. It includes: other contributions on Bayesian observational learning, such as Mueller-Frank (2013), Arieli and Mueller-Frank (2018), Mossel, Sly and Tamuz (2015); word-of-mouth learning models, where agents randomly sample others' opinions, as in Banerjee (1993), Ellison and Fudenberg (1995), and Banerjee and Fudenberg (2004); recent work on Bayesian communication learning, such as Acemoglu, Bimpikis and Ozdaglar (2014); models of non-Bayesian learning, including DeGroot (1974), DeMarzo, Vayanos and Zwiebel (2003), Acemoglu, Ozdaglar and ParandehGheibi (2010), Golub and Jackson (2010, 2012), and Molavi, Tahbaz-Salehi and Jadbabaie (2018); models where agents' updating rules combine Bayesian and non-Bayesian features, as in Bala and Goyal (1998) and Jadbabaie, Molavi, Sandroni and Tahbaz-Salehi (2012). Goyal (2007, 2011), Jackson (2008), Vives (2010), Acemoglu and Ozdaglar (2011), Mobius and Rosenblat (2014), and Golub and Sadler (2016) contain excellent (and complementary) accounts of the field.

[23] The fundamental trade-off between exploration and exploitation is the distinctive feature of bandit problems. I refer to Bergemann and Välimäki (2008) for a survey of bandit problems in economics.

[24] Salish (2017) adopts the discrete-time version of Keller et al. (2005), as in Heidhues, Rady and Strack (2015).

7.2 Concluding Remarks

I study observational learning over general networks where rational agents acquire private information via costly sequential search. When search costs are not bounded away from zero, asymptotic learning occurs in sufficiently connected networks where information paths are identifiable. The result relies on two theoretical underpinnings: first, I relate agents' solution to their information acquisition problem to the equilibrium probability that they select the best action; second, I establish an improvement principle for a novel informational environment, which significantly departs from that studied by previous models of social learning. The improvement principle, however, is particularly fragile in collective search environments: it breaks down as soon as zero is removed from the support of the search cost distribution. When search costs are bounded away from zero, even the weaker requirement of maximal learning fails in a large class of networks. Thus, when search costs are bounded away from zero, asymptotic learning fails discontinuously with respect to the benchmark learning metric. In some stochastic networks maximal (and sometimes also asymptotic) learning occurs despite search costs that are bounded away from zero. The impossibility to develop martingale convergence arguments, however, severely prevents the society from learning via the aggregation of dispersed pieces of information. In contrast with previous models of sequential learning, many equilibrium properties of the complete network extend to all networks where agents observe random numbers of immediate predecessors. Reducing transparency of past histories leads to welfare and efficiency losses. Simple policy interventions, such as letting agent observe the relative fraction of previous choices, restore part of the lost welfare.

Several questions remain. First, a general characterization of networks where maximal learning obtains when search costs are bounded away from zero is missing. Finding the demarcation line between possibility and impossibility of maximal learning in terms of network properties would be a valuable addition to this research. Second, quantifying the rate of convergence and efficiency losses in general networks is an important, but complex, task. Third, it remains to study the design of more complex incentives schemes to reduce inefficiencies and foster social exploration.

More broadly, relaxing the assumptions that agents have homogeneous preferences or that they can only take an action they have sampled might generate new insights. Lobel and Sadler (2016) study preference heterogeneity and homophily in the SSLM. They find that the improvement principle suffers, as imitation no longer guarantees the same payoff that a neighbor obtains when preferences are diverse; in contrast, the large-sample principle has more room to operate. In the search setting I study, the improvement principle is the key learning principle, while large-sample arguments have much less bite. Therefore, it is unclear what the analysis of preference heterogeneity would look like in collective search environments. Relaxing the assumption that agents can only take an action they have sampled is also non-trivial; this is a difficult question even for the single-agent sequential search problem (see Doval (2018) for some recent progress).

Alternatively, one might assume that acquiring private information and observing past histories are both costly activities. If individuals are heterogeneous across these two dimensions, in equilibrium some agents will specialize in search, while others in networking, thus enabling information to diffuse throughout the society. Studying how individuals make this trade-off, which network structures endogenously emerge, and the implications for social learning and information diffusion is a promising direction for future investigation.

A Examples for Remark 2 in Section 3.2.3

The first (resp., second) example shows that the incentives to explore for agents with nonempty neighborhood may increase as the quality of the first action sampled increases (resp., decreases).

Example 5. Suppose the qualities of the two actions are drawn uniformly at random from $\left\{0, \frac{49}{100}, \frac{51}{100}, 1\right\}$. Moreover, let $\left\{0, \frac{9}{100}, \frac{1}{8}, \frac{1}{3}\right\}$ be the support of the search cost distribution, with

$$\mathbb{P}_C(c=0) = \frac{1}{200}, \quad \mathbb{P}_C(c=9/100) = \frac{1}{200}, \quad \mathbb{P}_C(c=1/8) = \frac{32}{100}, \quad \text{and} \quad \mathbb{P}_C(c=1/3) = \frac{67}{100}.$$

Assume without loss that $a_1 = 0$ and that agent 2 observes the choice taken by agent 1. By Lemma 13, agent 2 samples first action 0: $s_2^1 = 0$. I will show that agent 2's expected additional gain from the second search is smaller when $q_0 = 49/100$ than when $q_0 = 51/100$. This implies that the incentives to explore of agent 2, who has nonempty neighborhood and can exploit his social information, increase as the quality of the first action sampled increases.

Let q_0 be the quality of action 0. The expected additional gain form the second search for agent 2 is $P_1(q_0)t^{\emptyset}(q_0)$, where $P_1(q_0)$ is the posterior probability that action 1 was not sampled by agent 1 given that action 0 of quality q_0 was taken. Here,

$$P_1(q_0) = \frac{N(q_0)}{D(q_0)},$$

with

$$\begin{aligned} N(q_0) &:= \mathbb{P}_\sigma\Big(s_1^1 = 0, c_1 > t^{\emptyset}(q_0)\Big) \\ &= \frac{1}{2}\mathbb{P}_C\Big(c_1 > t^{\emptyset}(q_0)\Big), \end{aligned} \tag{16}$$

and

$$\begin{aligned} D(q_0) &:= \mathbb{P}_\sigma\Big(s_1^1 = 0, c_1 > t^{\emptyset}(q_0)\Big) + \mathbb{P}_\sigma\Big(s_1^1 = 1, c_1 < t^{\emptyset}(q_1), q_0 > q_1\Big) \\ &\quad + \mathbb{P}_\sigma\Big(s_1^1 = 0, c_1 \le t^{\emptyset}(q_0), q_0 > q_1\Big) + \frac{1}{2}\mathbb{P}_\sigma\Big(s_1^1 = 0, c_1 \le t^{\emptyset}(q_0), q_0 = q_1\Big) \\ &= \frac{1}{2}\Big[\mathbb{P}_C\Big(c_1 > t^{\emptyset}(q_0)\Big) + \mathbb{P}_{C\times Q}\Big(c_1 < t^{\emptyset}(q_1), q_0 > q_1\Big) \\ &\qquad + \mathbb{P}_{C\times Q}\Big(c_1 \le t^{\emptyset}(q_0), q_0 > q_1\Big) + \frac{1}{2}\mathbb{P}_{C\times Q}\Big(c_1 \le t^{\emptyset}(q_0), q_0 = q_1\Big)\Big]. \end{aligned} \tag{17}$$

Above, I denote with $\mathbb{P}_{C\times Q}$ the product measure $\mathbb{P}_C \times \mathbb{P}_Q$ and with c_1 the search cost of agent 1. Consistently with the analysis in the rest of the paper, to derive an expression for $n(q_0)$ and $P_1(q_0)$ I assumed that agent 1 breaks ties uniformly at random at the first search stage and at the choice stage. The chosen tie-breaking rule does not qualitatively affect the results.[25]

Straightforward calculations yield

$$t^{\emptyset}(0) = \frac{1}{2}, \quad t^{\emptyset}(49/100) = \frac{51}{400}, \quad t^{\emptyset}(51/100) = \frac{49}{400}, \quad \text{and} \quad t^{\emptyset}(1) = 0.$$

[25] The same remarks apply to Example 6.

Moreover,

$$\mathbb{P}_C\Big(c_1 > t^{\emptyset}(49/100)\Big) = \frac{67}{100} \quad \text{and} \quad \mathbb{P}_C\Big(c_1 > t^{\emptyset}(51/100)\Big) = \frac{99}{100},$$

$$\mathbb{P}_{C\times Q}\Big(c_1 < t^{\emptyset}(q_1), q_1 < 49/100\Big) = \frac{100}{400} \quad \text{and} \quad \mathbb{P}_{C\times Q}\Big(c_1 < t^{\emptyset}(q_1), q_1 < 51/100\Big) = \frac{133}{400},$$

$$\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(49/100), 49/100 > q_1\Big) + \frac{1}{2}\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(49/100), 49/100 = q_1\Big) = \frac{99}{800},$$

and

$$\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(51/100), 51/100 > q_1\Big) + \frac{1}{2}\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(51/100), 51/100 = q_1\Big) = \frac{5}{800}.$$

Therefore,

$$P_1(49/100) = \frac{536}{800} \quad \text{and} \quad P_1(51/100) = \frac{44}{59}.$$

Note that $t^{\emptyset}(49/100) > t^{\emptyset}(51/100)$, while $P_1(49/100) < P_1(51/100)$. Then,

$$P_1(49/100)t^{\emptyset}(49/100) = \frac{536}{800}\frac{51}{400} \approx 0.086 \quad \text{and} \quad P_1(51/100)t^{\emptyset}(51/100) = \frac{44}{59}\frac{49}{400} \approx 0.091.$$

Since $P_1(49/100)t^{\emptyset}(49/100) < P_1(51/100)t^{\emptyset}(51/100)$, agent 2's incentives to sample the second action increase as the quality of the first action sampled increases, as claimed. In particular, if agent 2's search cost is 9/100, he samples the second action after sampling an action of quality 51/100, but discontinues search after sampling an action of quality 49/100. ■

Example 6. Suppose the qualities of the two actions are drawn uniformly at random from $\left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}$. Moreover, let $\left\{0, \frac{1}{15}, \frac{1}{3}\right\}$ be the support of the search cost distribution, with

$$\mathbb{P}_C(c = 0) = \frac{1}{4}, \quad \mathbb{P}_C(c = 1/15) = \frac{1}{4}, \quad \text{and} \quad \mathbb{P}_C(c = 1/3) = \frac{1}{2}.$$

As in Example 5, assume that agent 1 takes action 0, and that agent 2 observes agent 1. Thus, agent 2 samples first action 0. I will now show that agent 2's expected additional gain from the second search is larger when $q_0 = 1/3$ than when $q_0 = 2/3$. This implies agent 2's incentives to explore increase as the quality of the first action sampled decreases.

Mimicking the analysis in Example 5, we now have

$$t^{\emptyset}(0) = \frac{1}{2}, \quad t^{\emptyset}(1/3) = \frac{1}{4}, \quad t^{\emptyset}(2/3) = \frac{1}{12}, \quad \text{and} \quad t^{\emptyset}(1) = 0.$$

Moreover,

$$\mathbb{P}_C\Big(c_1 > t^{\emptyset}(1/3)\Big) = \frac{1}{2} \quad \text{and} \quad \mathbb{P}_C\Big(c_1 > t^{\emptyset}(2/3)\Big) = \frac{1}{2},$$

$$\mathbb{P}_{C\times Q}\Big(c_1 < t^{\emptyset}(q_1), q_1 < 1/3\Big) = \frac{1}{4} \quad \text{and} \quad \mathbb{P}_{C\times Q}\Big(c_1 < t^{\emptyset}(q_1), q_1 < 2/3\Big) = \frac{3}{8},$$

$$\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(1/3), 1/3 > q_1\Big) + \frac{1}{2}\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(1/3), 1/3 = q_1\Big) = \frac{3}{16},$$

and

$$\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(2/3), 2/3 > q_1\Big) + \frac{1}{2}\mathbb{P}_{C\times Q}\Big(c_1 \leq t^{\emptyset}(2/3), 2/3 = q_1\Big) = \frac{5}{16}.$$

Therefore,

$$P_1(1/3) = \frac{8}{15} \quad \text{and} \quad P_1(2/3) = \frac{8}{19}.$$

Note that now $t^\emptyset(1/3) > t^\emptyset(2/3)$ and $P_1(1/3) > P_1(2/3)$. Then,

$$P_1(1/3)t^\emptyset(1/3) = \frac{8}{15}\frac{1}{4} = \frac{2}{15} \quad \text{and} \quad P_1(2/3)t^\emptyset(2/3) = \frac{8}{19}\frac{1}{12} = \frac{2}{57}.$$

Since $P_1(1/3)t^\emptyset(1/3) > P_1(2/3)t^\emptyset(2/3)$, agent 2's incentives to sample the second action increase as the quality of the first action sampled decreases, as claimed. In particular, if agent 2's search cost is 1/15, he samples the second action after sampling an action of quality 1/3, but discontinues search after sampling an action of quality 2/3. ■

B Proofs

B.1 Proofs for Section 4.2.1

Preliminaries

The first lemma provides an obvious sufficient condition for asymptotic learning.

Lemma 1. *Let a collective search environment $\mathcal{S}$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ be given. If*

$$\lim_{n\to\infty} \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) = 1,$$

then asymptotic learning occurs in equilibrium σ.

Proof. In any equilibrium $\sigma \in \Sigma_{\mathcal{S}}$, each agent takes the action with the highest quality among those he has sampled. Since each agent must sample at least one action, the claim follows. ■

The next lemma shows that each agent does at least as well as the first agent in terms of the probability of sampling first the action with the highest quality.

Lemma 2. *Let a collective search environment $\mathcal{S}$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ be given. Then,*

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \geq \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right)$$

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

Proof. For $n = 1$, the claim trivially holds. Now fix an arbitrary agent $n > 1$ and let b, with $0 \leq b < n$, denote agent n's chosen neighbor. First, suppose $b = 0$. Since $b = 0 \iff B_n = \emptyset$, conditional on $\gamma_n(B(n)) = 0$ agent n faces the same problem as the first agent. Therefore,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = 0\right) = \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right).$$

Since agent 1's decision of which action to sample first is independent of the realization of agent n's neighborhood, the previous equality is equivalent to

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = 0\right) = \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = 0\right). \tag{18}$$

Second, suppose $0 < b < n$, so that $B_n \neq \emptyset$. By the characterization of the equilibrium decision s_n^1 in Section 3.2.2,

$$\mathbb{P}_\sigma\Big(E_n^{s_n^1} \mid c_n, B_n, a_k \text{ for all } k \in B_n\Big) \leq \mathbb{P}_\sigma\Big(E_n^{s_1^1} \mid c_n, B_n, a_k \text{ for all } k \in B_n\Big)$$

holds true for all realizations of $c_n \in C$, $B_n \in 2^{\mathbb{N}_n} \setminus \{\emptyset\}$, and $a_k \in X$ for all $k \in B_n$. By integrating over all possible private search costs and actions of the agents in the neighborhood, we obtain

$$\mathbb{P}_\sigma\Big(E_n^{s_n^1} \mid B_n\Big) \leq \mathbb{P}_\sigma\Big(E_n^{s_1^1} \mid B_n\Big).$$

for all $B_n \in 2^{\mathbb{N}_n} \setminus \{\emptyset\}$. Integrating further over all B_n such that $\gamma_n(B_n) = b$ we have

$$\mathbb{P}_\sigma\Big(E_n^{s_n^1} \mid \gamma_n(B(n)) = b\Big) \leq \mathbb{P}_\sigma\Big(E_n^{s_1^1} \mid \gamma_n(B(n)) = b\Big).$$

Therefore, conditional on $\gamma_n(B(n)) = b$, the marginal distribution of the quality of action s_n^1 first-order stochastically dominates the marginal distribution of the quality of action s_1^1. Hence,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X}\, q_x \;\middle|\; \gamma_n(B(n)) = b\right) \geq \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X}\, q_x \;\middle|\; \gamma_n(B(n)) = b\right). \tag{19}$$

The desired result obtains by observing that

$$\begin{aligned}
\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X}\, q_x\right) &= \sum_{b=0}^{n-1} \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X}\, q_x \;\middle|\; \gamma_n(B(n)) = b\right)\mathbb{Q}\Big(\gamma_n(B(n)) = b\Big) \\
&\geq \sum_{b=0}^{n-1} \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X}\, q_x \;\middle|\; \gamma_n(B(n)) = b\right)\mathbb{Q}\Big(\gamma_n(B(n)) = b\Big) \\
&= \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X}\, q_x\right),
\end{aligned}$$

where the two equalities hold by the law of total probability and the inequality holds by (18) and (19). ■

Proof of Proposition 2

The proof consists of two parts. In the first part, I construct two sequences, $(\alpha_k)_{k\in\mathbb{N}}$ and $(\phi_k)_{k\in\mathbb{N}}$, such that for all $k \in \mathbb{N}$, there holds

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X}\, q_x\right) \geq \phi_k \quad \text{for all } n \geq \alpha_k. \tag{20}$$

In the second part, I show that $\phi_k \to 1$ as $k \to \infty$. The desired result follows by combining these facts with Lemma 1.

By assumptions *(a)* and *(c)* of the proposition, for all positive integer α and all $\varepsilon > 0$, there exist a positive integer $N(\alpha, \varepsilon)$ and a sequence of neighbor choice functions $(\gamma_k)_{k\in\mathbb{N}}$ such that

$$\mathbb{Q}\Big(\gamma_n((B(n)) = b, b < \alpha\Big) < \frac{\varepsilon}{2}, \tag{21}$$

and

$$\mathbb{P}_\sigma\left(\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \,\middle|\, \gamma_n(B(n))\right) < \mathcal{Z}\left(\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x\in X} q_x\right)\right) - \varepsilon\right) < \frac{\varepsilon}{2} \tag{22}$$

for all $n \geq N(\alpha, \varepsilon)$. Now, set $\phi_1 := \frac{1}{2}$ and $\alpha_1 := 1$, and define $(\phi_k)_{k\in\mathbb{N}}$ and $(\alpha_k)_{k\in\mathbb{N}}$ recursively by

$$\phi_{k+1} := \frac{\phi_k + \mathcal{Z}(\phi_k)}{2}, \qquad \text{and} \qquad \alpha_{k+1} := N(\alpha_k, \varepsilon_k),$$

where the sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ is defined by

$$\varepsilon_k := \frac{1}{2}\left(1 + \mathcal{Z}(\phi_k) - \sqrt{1 + 2\phi_k + \mathcal{Z}(\phi_k)^2}\right).$$

Given the assumptions on $\mathcal{Z}$, these sequences are well-defined.

I use induction on the index k to prove relation (20). Since the qualities of the two actions are i.i.d. draws and agent 1 has no a priori information,

$$\mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right) = \frac{1}{2}. \tag{23}$$

From Lemma 2,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \geq \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right) \tag{24}$$

for all $n \in \mathbb{N}$. From (23) and (24) we have

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \geq \frac{1}{2} \qquad \text{for all } n \geq 1,$$

which together with $\alpha_1 = 1$ and $\phi_1 = \frac{1}{2}$ establishes relation (20) for $k = 1$. Assume that relation (20) holds for an arbitrary k, that is

$$\mathbb{P}_\sigma\left(s_j^1 \in \arg\max_{x\in X} q_x\right) \geq \phi_k \qquad \text{for all } j \geq \alpha_k, \tag{25}$$

and consider some agent $n \geq \alpha_{k+1}$. To establish (20) for $n \geq \alpha_{k+1}$ observe that

$$\begin{aligned}\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) &= \sum_{b=0}^{n-1} \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \,\middle|\, \gamma_n((B(n)) = b\right)\mathbb{Q}\big(\gamma_n((B(n)) = b\big) \\ &\geq (1-\varepsilon_k)\,(\mathcal{Z}(\phi_k) - \varepsilon_k) \\ &\geq \phi_{k+1},\end{aligned}$$

where the inequality follows from (21) and (22), the inductive hypothesis in (25), and the assumption that $\mathcal{Z}$ is increasing.

Finally, I prove that $\phi_k \to 1$ as $k \to \infty$. By assumption *(b)* of the proposition, $\mathcal{Z}(\beta) \geq \beta$ for all $\beta \in [1/2, 1]$; it follows from the definition of ϕ_k that $(\phi_k)_{k\in\mathbb{N}}$ is a non-decreasing sequence. Since it is also bounded, it converges to some ϕ^*. Taking the limit in the definition of ϕ_k, we obtain

$$2\phi^* = 2\lim_{k\to\infty} \phi_k = \lim_{k\to\infty} [\phi_k + \mathcal{Z}(\phi_k)] = \phi^* + \mathcal{Z}(\phi^*),$$

where the third equality holds by continuity of $\mathcal{Z}$. This shows that $\phi^* = \mathcal{Z}(\phi^*)$, i.e. ϕ^* is a fixed point of $\mathcal{Z}$. Since the unique fixed point of $\mathcal{Z}$ is 1, we have $\phi_k \to 1$ as $k \to \infty$, as claimed. ■

Proof of Proposition 3

Proposition 3 follows by combining several lemmas, which I next present.

Hereafter, let a collective search environment $\mathcal{S}$, a state of the world $\omega := (q_0, q_1) \in \Omega$, an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$, a sequence of neighbor choice functions $(\gamma_n)_{n\in\mathbb{N}}$, and an agent $n \in \mathbb{N}$ be fixed. Moreover, let b, with $0 \le b < n$, be n's chosen neighbor.

Denote with $\tilde{s}_n^1$ the coarse optimal decision of agent n at the first search stage when he only uses information from neighbor b.[26]. The optimal search policy, as characterized in Section 3.2.2, requires

$$\tilde{s}_n^1 \in \arg\min_{x\in X} \mathbb{P}_\sigma\Big(E_n^x \mid \gamma_n(B(n)) = b, a_b\Big),$$

where indifference is resolved according to agent n's mixed strategy.

Suppose that the probability that none of the agents in $\widehat{B}(n, a_b)$ sampled both actions is smaller than the probability that none of the agents in $\widehat{B}(n, \neg a_b)$ sampled both actions whenever agent n's neighbor choice function selects agent b, with $0 \le b < n$. That is,

$$\mathbb{P}_\sigma\Big(E_n^{a_b} \mid \gamma_n(B(n)) = b\Big) \le \mathbb{P}_\sigma\Big(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b\Big). \tag{26}$$

Then, agent n samples first action a_b: $\tilde{s}_n^1 = a_b$. Henceforth, I assume that agent n samples first action a_b in case of indifference. The assumption does not affect my results. The next lemma summarizes.

Lemma 3. *Suppose $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \le \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$ and $\gamma_n(B(n)) = b$. Then, the coarse version $\tilde{s}_n^1$ of agent n's equilibrium strategy at the first search stage is $\tilde{s}_n^1 = a_b$.*

Remark 8. Since $\gamma_n(B(n)) = 0$ iff $B(n) = \emptyset$, it is without loss of generality to impose $\tilde{s}_n^1 = s_n^1$ conditional on $\gamma_n(B(n)) = 0$. That is, conditional on $\gamma_n(B(n)) = 0$, the coarse version of agent n's equilibrium decision of which action to sample first coincides with his equilibrium decision.

The next lemma shows that network topologies where $\mathbb{Q}(|B(n)| \le 1) = 1$ for all n satisfy condition (26). In particular, this condition is satisfied by all chosen neighbor topologies.

Lemma 4. *Suppose that the network topology $(\mathbb{B}, \mathcal{F}_\mathbb{B}, \mathbb{Q})$ satisfies $\mathbb{Q}(|B(n)| \le 1) = 1$ for all $n \in \mathbb{N}$. Then, $\mathbb{P}_\sigma(E_n^{a_b} \mid \widehat{B}(n) = \widehat{B}_n) \le \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \widehat{B}(n) = \widehat{B}_n)$ for all agents n and b, with $0 \le b < n$, and for all realizations $\widehat{B}_n$ that occurs with positive probability. It follows that $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \le \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$ for all n and b, with $0 \le b < n$.*

Proof. Proceed by induction. The first agent has empty neighborhood. Hence, his personal subnetworks relative to the two actions are empty and the statement is vacuously true.

Now suppose $\mathbb{P}_\sigma(E_n^{a_b} \mid \widehat{B}(n) = \widehat{B}_n) \le \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \widehat{B}(n) = \widehat{B}_n)$ for all $n \le k$ and all $\widehat{B}_n$ that occurs with positive probability. Given a realization $\widehat{B}_{k+1}$ of $\widehat{B}(k+1)$, if $B_{k+1} = \emptyset$, then agent $k+1$ faces the same situation as the first agent, and the desired conclusion follows. If $B_{k+1} = \{b\}$, take $\gamma_{k+1}(\{b\}) = b$ and let $(\pi_1, \ldots, \pi_l)$ be the sequence of agents in $\widehat{B}_{k+1} \cup \{k+1\}$. That is, $\{\pi_1, \ldots, \pi_l\}$ is such that $\pi_1 = \min \widehat{B}_{k+1}$, $\pi_l = k+1$ and, for all g with $1 < g \le l$, $B_{\pi_g} = \{\pi_{g-1}\}$.

[26] By definition of neighbor choice function, the fictitious agent 0 is agent n's chosen neighbor iff $B_n = \emptyset$.

Moreover, for all g with $1 < g \leq l$, say that agent π_{g-1} is the immediate predecessor of agent π_g in $\widehat{B}_{k+1}$. When $\widehat{B}_{k+1} = \{b\}$, the desired result trivially holds. When $\widehat{B}_{k+1}$ contains more than one agent, the desired result follows by observing that, under the inductive hypothesis and the equilibrium decision rule, each agent in $\{\pi_1, \ldots, \pi_{l-1}\}$ samples first the action taken by his immediate predecessor. ■

The next definition introduces the notation that will be used in the following analysis.

Definition 11. *Fix a state of the world $\omega := (q_0, q_1) \in \Omega$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$. The following objects are defined:*

$$\begin{aligned}
q_{\min} &:= \min\{q_0, q_1\}, \\
q_{\max} &:= \max\{q_0, q_1\}, \\
P^{\sigma}_{b,n}(q_{\min}) &:= \mathbb{P}_{\sigma}\left(E_b^{s_b^1} \,\middle|\, \gamma_n(B(n)) = b, q_{s_b^1} = q_{\min}\right) \\
&= \mathbb{P}_{\sigma}\left(E_b^{s_b^1} \,\middle|\, \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x \in X} q_x\right), \\
P^{\sigma}_{b,n}(q_{\max}) &:= \mathbb{P}_{\sigma}\left(E_b^{s_b^1} \,\middle|\, \gamma_n(B(n)) = b, q_{s_b^1} = q_{\max}\right) \\
&= \mathbb{P}_{\sigma}\left(E_b^{s_b^1} \,\middle|\, \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x \in X} q_x\right), \\
\beta &:= \mathbb{P}_{\sigma}\left(s_b^1 \in \arg\max_{x \in X} q_x \,\middle|\, \gamma_n(B(n)) = b\right).
\end{aligned}$$

Remark 9. In any equilibrium $\sigma \in \Sigma_{\mathcal{S}}$, $\beta \geq \frac{1}{2}$ for all $b \in \mathbb{N}$. This is so because the distribution of the quality of the first action sampled by an agent first-order stochastically dominates (although not necessarily strictly so) the distribution of the quality of the other action.

The next two lemmas provide an expression for the probability of agent n sampling first the best action when using $\tilde{s}_n^1$, conditional on agent b being selected by agent n's neighbor choice function, in terms of the probability β of agent b doing so, the private search cost distribution, the function $t^{\emptyset}(\cdot)$ defined in (2), and the thresholds $P^{\sigma}_{b,n}(q_{\min})$ and $P^{\sigma}_{b,n}(q_{\max})$.

Lemma 5. *Suppose $\mathbb{P}_{\sigma}(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_{\sigma}(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$. Then,*

$$\begin{aligned}
&\mathbb{P}_{\sigma}\left(\tilde{s}_n^1 \in \arg\max_{x \in X} q_x \,\middle|\, \gamma_n(B(n)) = b\right) \\
&= \mathbb{P}_{\sigma}\left(s_b^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = b\right) \\
&+ \mathbb{P}_{\sigma}\left(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\right)\left(1 - \mathbb{P}_{\sigma}\left(s_b^1 \in \arg\max_{x \in X} q_x \,\middle|\, \gamma_n(B(n)) = b\right)\right).
\end{aligned} \tag{27}$$

Proof. By Lemma 3,

$$\mathbb{P}_{\sigma}\left(\tilde{s}_n^1 \in \arg\max_{x \in X} q_x \,\middle|\, \gamma_n(B(n)) = b\right) = \mathbb{P}_{\sigma}\left(a_b \in \arg\max_{x \in X} q_x \,\middle|\, \gamma_n(B(n)) = b\right).$$

Moreover,

$$\mathbb{P}_{\sigma}\left(a_b \in \arg\max_{x \in X} q_x \,\middle|\, \gamma_n(B(n)) = b\right)$$

$$
\begin{aligned}
&= \mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b, s_b^2 = \neg s_b^1\right)\mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\Big) \\
&+ \mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b, s_b^2 = ns\right)\mathbb{P}_\sigma\Big(s_b^2 = ns \mid \gamma_n(B(n)) = b\Big) \\
&= \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\Big) \\
&+ \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right)\Big(1 - \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\Big)\Big) \\
&= \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \\
&+ \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\Big)\left(1 - \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right)\right).
\end{aligned}
$$

Here, the first equality holds by the law of total probability; the second equality holds because whenever agent b samples both actions, $s_b^2 = \neg s_b^1$, he takes the one with the highest quality, so that

$$
\mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b, s_b^2 = \neg s_b^1\right) = 1,
$$

and when agent b only samples one action, $s_b^2 = ns$, he takes that action, so that

$$
\mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b, s_b^2 = ns\right) = \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right).
$$

The desired result follows. ■

Lemma 6. *Suppose* $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \le \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$. *Then,*

$$
\begin{aligned}
&\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \\
&= \beta + (1-\beta)\Big[\beta F_C\Big(P_{b,n}^\sigma(q_{\max})t^\emptyset(q_{\max})\Big) + (1-\beta)F_C\Big(P_{b,n}^\sigma(q_{\min})t^\emptyset(q_{\min})\Big)\Big].
\end{aligned}
$$

Proof. By Lemma 5,

$$
\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) = \beta + \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\Big)(1-\beta). \tag{28}
$$

Moreover, by the law of total probability,

$$
\begin{aligned}
&\mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\Big) \\
&= \mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \\
&+ \mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(s_b^1 \notin \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \qquad (29) \\
&= \beta\,\mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\right) \\
&+ (1-\beta)\,\mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\right).
\end{aligned}
$$

By the characterization of equilibrium strategies in Section 3.2.2 we have, conditional on

$\gamma_n(B(n)) = b$ and $s_b^1 \in \arg\max_{x\in X} q_x$,

$$s_b^2 = \neg s_b^1 \iff c_b \le P_{b,n}^{\sigma}(q_{\max})t^{\emptyset}(q_{\max})$$

and, conditional on $\gamma_n(B(n)) = b$ and $s_b^1 \notin \arg\max_{x\in X} q_x$,

$$s_b^2 = \neg s_b^1 \iff c_b \le P_{b,n}^{\sigma}(q_{\min})t^{\emptyset}(q_{\min}),$$

where we assume that agent n samples the second action in case of indifference.[27] It follows that

$$\mathbb{P}_{\sigma}\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\right) = F_C\Big(P_{b,n}^{\sigma}(q_{\max})t^{\emptyset}(q_{\max})\Big),$$

and that

$$\mathbb{P}_{\sigma}\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\right) = F_C\Big(P_{b,n}^{\sigma}(q_{\min})t^{\emptyset}(q_{\min})\Big).$$

Thus, equation (29) can be rewritten as

$$\mathbb{P}_{\sigma}\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\Big) = \beta F_C\Big(P_{b,n}^{\sigma}(q_{\max})t^{\emptyset}(q_{\max})\Big) + (1-\beta)F_C\Big(P_{b,n}^{\sigma}(q_{\min})t^{\emptyset}(q_{\min})\Big). \quad (30)$$

The desired result follows by combining (28) and (30). ■

The previous lemma shows that the quantity

$$(1-\beta)\left[\beta F_C\Big(P_{b,n}^{\sigma}(q_{\max})t^{\emptyset}(q_{\max})\Big) + (1-\beta)F_C\Big(P_{b,n}^{\sigma}(q_{\min})t^{\emptyset}(q_{\min})\Big)\right]$$

acts as an improvement in the probability that agent n samples first the best action over his chosen neighbor's probability. This improvement term is still unsuitable for the analysis to come because it depends on $P_{b,n}^{\sigma}(q_{\min})$ and $P_{b,n}^{\sigma}(q_{\max})$, which are difficult to handle. The next lemma provides a simple lower bound on the amount of this improvement. It also establishes that this lower bound is uniformly bounded away from zero whenever $\beta < 1$, and that it is non-negative when $\beta = 1$.

Lemma 7. *Suppose* $\mathbb{P}_{\sigma}(E_n^{a_b} \mid \gamma_n(B(n)) = b) \le \mathbb{P}_{\sigma}(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$. *Then,*

$$\mathbb{P}_{\sigma}\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \ge \beta + (1-\beta)^2 F_C\Big((1-\beta)t^{\emptyset}(q_{\max})\Big).$$

Proof. Whenever at least one of the agents in the personal subnetwork of agent b relative to action s_b^1 samples both actions, $s_b^1 \in \arg\max_{x\in X} q_x$. Therefore,

$$\beta \ge 1 - \mathbb{P}_{\sigma}\left(E_b^{s_b^1} \;\middle|\; \gamma_n(B(n)) = b\right),$$

or

$$1-\beta \le \mathbb{P}_{\sigma}\left(E_b^{s_b^1} \;\middle|\; \gamma_n(B(n)) = b\right). \quad (31)$$

[27] This assumption does not affect the results.

Moreover, by the law of total probability,

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(E_b^{s_b^1} \;\Big|\; \gamma_n(B(n)) = b\Big) \\
&= \mathbb{P}_\sigma\Big(E_b^{s_b^1} \;\Big|\; \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X}\; q_x\Big)\mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x\in X}\; q_x \;\Big|\; \gamma_n(B(n)) = b\Big) \\
&+ \mathbb{P}_\sigma\Big(E_b^{s_b^1} \;\Big|\; \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X}\; q_x\Big)\mathbb{P}_\sigma\Big(s_b^1 \notin \arg\max_{x\in X}\; q_x \;\Big|\; \gamma_n(B(n)) = b\Big) \\
&= \beta P^\sigma_{b,n}(q_{\max}) + (1-\beta) P^\sigma_{b,n}(q_{\min}).
\end{aligned} \tag{32}$$

Combining (31) and (32) yields

$$1-\beta \le \beta P^\sigma_{b,n}(q_{\max}) + (1-\beta) P^\sigma_{b,n}(q_{\min}), \tag{33}$$

and therefore

$$\max\Big\{P^\sigma_{b,n}(q_{\min}), P^\sigma_{b,n}(q_{\max})\Big\} \ge 1-\beta. \tag{34}$$

Finally, observe that

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(\tilde{s}_n^1 \in \arg\max_{x\in X}\; q_x \;\Big|\; \gamma_n(B(n)) = b\Big) \\
&= \beta + (1-\beta)\Big[\beta F_C\Big(P^\sigma_{b,n}(q_{\max})t^\emptyset(q_{\max})\Big) + (1-\beta)F_C\Big(P^\sigma_{b,n}(q_{\min})t^\emptyset(q_{\min})\Big)\Big] \\
&\ge \beta + (1-\beta)\Big[(1-\beta) F_C\Big(P^\sigma_{b,n}(q_{\max})t^\emptyset(q_{\max})\Big) + (1-\beta)F_C\Big(P^\sigma_{b,n}(q_{\min})t^\emptyset(q_{\min})\Big)\Big] \\
&= \beta + (1-\beta)^2\Big[F_C\Big(P^\sigma_{b,n}(q_{\max})t^\emptyset(q_{\max})\Big) + F_C\Big(P^\sigma_{b,n}(q_{\min})t^\emptyset(q_{\min})\Big)\Big] \\
&\ge \beta + (1-\beta)^2\Big[F_C\Big(P^\sigma_{b,n}(q_{\max})t^\emptyset(q_{\max})\Big) + F_C\Big(P^\sigma_{b,n}(q_{\min})t^\emptyset(q_{\max})\Big)\Big] \\
&\ge \beta + (1-\beta)^2 \max\Big\{F_C\Big(P^\sigma_{b,n}(q_{\max})t^\emptyset(q_{\max})\Big), F_C\Big(P^\sigma_{b,n}(q_{\min})t^\emptyset(q_{\max})\Big)\Big\} \\
&\ge \beta + (1-\beta)^2 F_C\Big((1-\beta)t^\emptyset(q_{\max})\Big).
\end{aligned}$$

Here, the first equality holds by Lemma 6; the first inequality holds because, as $\beta \ge 1/2$ by Remark 9, $\beta \ge (1-\beta)$; the second inequality holds because $t^\emptyset(q_{\max}) \le t^\emptyset(q_{\min})$ and the CDF F_C is increasing; the third inequality holds because the CDF F_C is non-negative; the last inequality follows from

$$\max\Big\{F_C\Big(P^\sigma_{b,n}(q_{\max})t^\emptyset(q_{\max})\Big), F_C\Big(P^\sigma_{b,n}(q_{\min})t^\emptyset(q_{\max})\Big)\Big\} \ge F_C\Big((1-\beta)t^\emptyset(q_{\max})\Big),$$

which holds because of (34) and the fact that F_C is increasing. The desired result follows. ■

The previous lemmas describe the improvement that a single agent can make over her neighbor by employing a heuristic that discards the information from all other neighbors. To study the limiting behavior of these improvements, I introduce the function $\overline{\mathcal{Z}}\colon [1/2, 1] \to [1/2, 1]$ defined by

$$\overline{\mathcal{Z}}(\beta) := \beta + (1-\beta)^2 F_C\Big((1-\beta)t^\emptyset(q_{\max})\Big). \tag{35}$$

Hereafter, I call $(1-\beta)^2 F_C((1-\beta)t^\emptyset(q_{\max}))$ the *improvement term* of function $\overline{\mathcal{Z}}$.

Lemma 7 establishes that, when $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$, we have

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x \in X} q_x \;\Big|\; \gamma_n(B(n)) = b\right) = \overline{\mathcal{Z}}\left(\mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x \in X} q_x \;\Big|\; \gamma_n(B(n)) = b\right)\right).$$

That is, the function $\overline{\mathcal{Z}}$ acts as an *improvement function* for the evolution of the probability of searching first for the best action. The next lemma presents some useful properties of $\overline{\mathcal{Z}}$.

Lemma 8. *The function* $\overline{\mathcal{Z}}\colon [1/2, 1] \to [1/2, 1]$*, defined pointwise by (35), satisfies the following properties:*

(a) For all $\beta \in [1/2, 1]$*,* $\overline{\mathcal{Z}}(\beta) \geq \beta$*.*

(b) If the search technology features search costs that are not bounded away from zero, then $\overline{\mathcal{Z}}(\beta) > \beta$ *for all* $\beta \in [1/2, 1)$*.*

(c) The function $\overline{\mathcal{Z}}$ *is left-continuous and has no upward jumps:*

$$\overline{\mathcal{Z}}(\beta) = \lim_{r \uparrow \beta} \overline{\mathcal{Z}}(r) \geq \lim_{r \downarrow \beta} \overline{\mathcal{Z}}(r).$$

Proof. Since F_C is a CDF and $(1-\beta)^2 \geq 0$, the improvement term of function $\overline{\mathcal{Z}}$ is always non-negative. Part *(a)* follows.

For all $\beta \in [1/2, 1)$, $(1-\beta)t^{\emptyset}(q_{\max}) > 0$ and so, if search costs are not bounded away from zero, $F_C((1-\beta)t^{\emptyset}(q_{\max})) > 0$.[28] Since also $(1-\beta)^2 > 0$ for all $\beta \in [1/2, 1)$, the improvement term of function $\overline{\mathcal{Z}}$ is positive and so part *(b)* holds.

For part *(c)*, set $\alpha \coloneqq (1-\beta)t^{\emptyset}(q_{\max})$. Since F_C is a CDF, it is right-continuous and has no downward jumps in α. Therefore, F_C is left-continuous and has no upward jumps in β. Since β and $(1-\beta)^2$ are continuous functions of β, and so also left-continuous with no upward jumps, the desired result follows because the product and the sum of left-continuous functions with no upward jumps is left-continuous with no upward jumps. ■

Next, I construct a related function $\mathcal{Z}$ that is monotone and continuous while maintaining the same improvement properties of $\overline{\mathcal{Z}}$. In particular, define $\mathcal{Z}\colon [1/2, 1] \to [1/2, 1]$ as

$$\mathcal{Z}(\beta) \coloneqq \frac{1}{2}\left(\beta + \sup_{r \in [1/2, \beta]} \overline{\mathcal{Z}}(r)\right). \tag{36}$$

Lemma 9. *The function* $\mathcal{Z}\colon [1/2, 1] \to [1/2, 1]$ *defined by (36) satisfies the following properties:*

(a) For all $\beta \in [1/2, 1]$*,* $\mathcal{Z}(\beta) \geq \beta$*.*

(b) If the search technology features search costs that are not bounded away from zero, then $\mathcal{Z}(\beta) > \beta$ *for all* $\beta \in [1/2, 1)$*.*

(c) The function $\mathcal{Z}$ *is increasing and continuous.*

Proof. Parts *(a)* and *(b)* immediately result from the corresponding parts of Lemma 8.

The function $\sup_{r \in [1/2, \beta]} \overline{\mathcal{Z}}(r)$ is non-decreasing and the function β is increasing. Therefore, the average of these two functions, which is $\mathcal{Z}$, is an increasing function, establishing the first

[28] Note that $t^{\emptyset}(q_{\max}) = 0$ if $q_{s_b^1} = q_{\max} = \max \operatorname{supp}(\mathbb{P}_Q)$ whenever such sup exists as a real number. However, in such cases we would trivially have $\beta = 1$, which is not the case considered here.

part of *(c)*. Finally, I show that $\mathcal{Z}$ is continuous. To establish continuity in $[1/2, 1)$, I argue by contradiction. Suppose that $\mathcal{Z}$ is discontinuous at some $\beta' \in [1/2, 1)$. This implies that $\sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)$ is discontinuous at β'. Since $\sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)$ is a non-decreasing function, it must be that

$$\lim_{\beta\downarrow\beta'} \sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r) > \sup_{r\in[1/2,\beta']} \overline{\mathcal{Z}}(r),$$

from which it follows that there exists some $\varepsilon > 0$ such that for all $\delta > 0$

$$\sup_{r\in[1/2,\beta'+\delta]} \overline{\mathcal{Z}}(r) > \overline{\mathcal{Z}}(\beta) + \varepsilon \quad \text{for all } \beta \in [1/2, \beta')\,.$$

This contradicts that the function $\overline{\mathcal{Z}}$ has no upward jumps, which was established as property *(c)* in Lemma 8. Continuity of $\mathcal{Z}$ at $\beta = 1$ follows from part *(a)*. ■

The next lemma shows that the function $\mathcal{Z}$ is also a *improvement function* for the evolution of the probability of searching first for the action with highest quality.

Lemma 10. *Suppose that* $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$. *Then,*

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \,\Big|\, \gamma_n(B(n)) = b\right) \geq \mathcal{Z}\left(\mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \,\Big|\, \gamma_n(B(n)) = b\right)\right).$$

Proof. Let again β denote $\mathbb{P}_\sigma(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b)$. If $\mathcal{Z}(\beta) = \beta$, the result follows from Lemma 6. Suppose next that $\mathcal{Z}(\beta) > \beta$. By (36), this implies that $\mathcal{Z}(\beta) < \sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)$. Therefore, there exists $\overline{\beta} \in [1/2, \beta]$ such that

$$\overline{\mathcal{Z}}(\overline{\beta}) \geq \mathcal{Z}(\beta). \tag{37}$$

I next show that $\mathbb{P}_\sigma(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b) \geq \overline{\mathcal{Z}}(\overline{\beta})$. Agent n can always make his decision even coarser by choosing not to observe the choice of agent b with some probability. Suppose that instead of considering b's action directly, agent n bases his decision of which action to sample first on the observation of a fictitious agent whose action, denoted by $\tilde{a}_b$, is generated as

$$\tilde{a}_b = \begin{cases} a_b & \text{with probability } \ (2\overline{\beta} - 1)/(2\beta - 1) \\ 0 & \text{with probability } \ (\beta - \overline{\beta})/(2\beta - 1) \\ 1 & \text{with probability } \ (\beta - \overline{\beta})/(2\beta - 1), \end{cases} \tag{38}$$

with the realization of $\tilde{a}_b$ independent of the rest of n's information set. Under the assumption $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$, we have

$$\mathbb{P}_\sigma\Big(E_n^{\tilde{a}_b} \mid \gamma_n(B(n)) = b\Big) \leq \mathbb{P}_\sigma\Big(E_n^{\neg\tilde{a}_b} \mid \gamma_n(B(n)) = b\Big). \tag{39}$$

The relation in (39), together with the characterization of the equilibrium search policy in Section 3.2.2, implies that agent n samples first action $\tilde{a}_b$ upon observing the choice of the fictitious agent. That is, denoting with $\widetilde{\tilde{s}_n^1}$ the first action sampled by agent n upon observing the choice of the fictitious agent, $\widetilde{\tilde{s}_n^1} = \tilde{a}_b$. Moreover, the assumption $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$ and (38) also imply that $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\tilde{a}_b} \mid \gamma_n(B(n)) = b)$. Therefore, the distribution of the quality of action a_b first-order stochastically dominates the distribution of the quality of action $\tilde{a}_b$. Since $\tilde{s}_n^1 = a_b$ and $\widetilde{\tilde{s}_n^1} = \tilde{a}_b$, it follows that

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \geq \mathbb{P}_\sigma\left(\widetilde{\tilde{s}_n^1} \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right). \tag{40}$$

Now denote with $\tilde{s}_b^1$ the decision of the fictitious agent about which action to sample first. From (38), one can think of $\tilde{s}_b^1$ as generated as

$$\tilde{s}_b^1 = \begin{cases} s_b^1 & \text{with probability } (2\overline{\beta}-1)/(2\beta-1) \\ 0 & \text{with probability } (\beta-\overline{\beta})/(2\beta-1) \\ 1 & \text{with probability } (\beta-\overline{\beta})/(2\beta-1). \end{cases}$$

Therefore,

$$\begin{aligned} \mathbb{P}_\sigma\left(\tilde{s}_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) &= \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right)\frac{2\overline{\beta}-1}{2\beta-1} \\ &+ \mathbb{P}_\sigma\left(0 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right)\frac{\beta-\overline{\beta}}{2\beta-1} \\ &+ \mathbb{P}_\sigma\left(1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right)\frac{\beta-\overline{\beta}}{2\beta-1} \\ &= \beta\frac{2\overline{\beta}-1}{2\beta-1} + (\beta + (1-\beta))\frac{\beta-\overline{\beta}}{2\beta-1} \\ &= \overline{\beta}. \end{aligned}$$

Lemma 7 implies that the first action sampled by agent n based on the observation of this fictitious agent is the one with the highest quality with probability at least $\overline{\mathcal{Z}}(\overline{\beta})$, that is

$$\mathbb{P}_\sigma\left(\widetilde{\tilde{s}_n^1} \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \geq \overline{\mathcal{Z}}(\overline{\beta}). \tag{41}$$

Since $\overline{\mathcal{Z}}(\overline{\beta}) \geq \mathcal{Z}(\beta)$ (see equation (37)), the desired result follows from (40) and (41). ■

It remains to show that the equilibrium search policy s_n^1 does at least as well as its coarse version $\tilde{s}_n^1$ in terms of sampling first the action with the highest quality given $\gamma_n(B(n)) = b$. This is established with the next lemma and completes the proof of Proposition 3.

Lemma 11. *For all agents n and any b, with $0 \leq b < n$, we have*

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \geq \mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right).$$

Proof. Fix any $n \in \mathbb{N}$. If $b = 0$, then $\tilde{s}_n^1 = s_n^1$ by Remark 8, and the claim trivially holds. Now suppose $0 < b < n$, so that $B_n \neq \emptyset$. By the characterization of the equilibrium decision s_n^1 in Section 3.2.2,

$$\mathbb{P}_\sigma\left(E_n^{s_n^1} \mid c_n, B_n, a_k \text{ for all } k \in B_n\right) \leq \mathbb{P}_\sigma\left(E_n^{\tilde{s}_n^1} \mid c_n, B_n, a_k \text{ for all } k \in B_n\right)$$

holds true for all realizations of $c_n \in C$, $B_n \in 2^{\mathbb{N}_n} \setminus \{\emptyset\}$, and $a_k \in X$ for all $k \in B_n$. By integrating over all possible private search costs and actions of the agents in the neighborhood, we obtain

$$\mathbb{P}_\sigma\Big(E_n^{s_n^1} \mid B_n\Big) \leq \mathbb{P}_\sigma\Big(E_n^{\tilde{s}_n^1} \mid B_n\Big) \tag{42}$$

for all $B_n \in 2^{\mathbb{N}_n} \setminus \{\emptyset\}$. Integrating further over all B_n such that $\gamma_n(B_n) = b$ we conclude

$$\mathbb{P}_\sigma\Big(E_n^{s_n^1} \mid \gamma_n(B(n)) = b\Big) \leq \mathbb{P}_\sigma\Big(E_n^{\tilde{s}_n^1} \mid \gamma_n(B(n)) = b\Big).$$

Then, conditional on $\gamma_n(B(n)) = b$, the marginal distribution of the quality of action s_n^1 first-order stochastically dominates the marginal distribution of the quality of action $\tilde{s}_n^1$. Therefore,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \geq \mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right),$$

as desired. ■

B.2 Proofs for Section 4.3

Preliminaries

Definition 12. *Let q^{NS}, Q^{NS}, and Ω^{NS} be defined as follows:*

- $q^{NS} := \inf\{\tilde{q} \in \text{supp}\,(\mathbb{P}_Q) : 1$*–(a) and* 1*–(b) in Assumption 1 hold*$\}$;
- $Q^{NS} := \left\{\tilde{q} \in Q : \tilde{q} \geq q^{NS}\right\}$;
- $\Omega^{NS} := Q^{NS} \times Q^{NS}$.

In words, Ω^{NS} includes all states of the world ω where, with positive probability, an agent with empty neighborhood does not sample the second action independently of which action he samples first. By the first condition in Assumption 1, there exists some $\delta > 0$ such that $\mathbb{P}_Q(Q^{NS}) \geq \sqrt{\delta}$ and so, by definition of product measure,

$$\mathbb{P}_\Omega\Big(\Omega^{NS}\Big) = \mathbb{P}_Q\Big(Q^{NS}\Big) \times \mathbb{P}_Q\Big(Q^{NS}\Big) \geq \delta. \tag{43}$$

When $\omega \in \Omega^{NS}$, an agent with nonempty neighborhood does not sample the second action either with positive probability, independently of which action he samples first (see the characterization and discussion of equilibrium behavior in Section 3.2). Finally, by Assumption 1, conditional on $\omega \in \Omega^{NS}$, the two actions have different quality with positive probability.

Fix a collective search environment $\mathcal{S}$. Asymptotic learning occurs in equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ only if the probability of agent n taking the action with the lowest quality converges to zero with respect to $\mathbb{P}_\sigma$ as n goes to infinity. Because of Assumption 1, a necessary condition for this to happen is that the probability of no agent in $\hat{B}(n) \cup \{n\}$ sampling both actions converges to zero as n goes to infinity with respect to $\mathbb{P}_\sigma$. If this were not the case, there would be a subsequence of agents who, with probability bounded away from zero, only observe (directly and indirectly) agents who have not compared the quality of the two actions (as none of the agents in their personal subnetworks has sampled both actions), and do not make this comparison either (as they do not search for the second alternative). Asymptotic learning would trivially fail as the only way to ascertain the relative quality of the two actions is to sample both of them. The next lemma follows.

Lemma 12. *Let a collective search environment $\mathcal{S}$ and an equilibrium $\sigma \in \Sigma_{\mathcal{S}}$ be given. If asymptotic learning occurs in equilibrium σ, then*

$$\lim_{n\to\infty} \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\Big) = 0.$$

Proof of Proposition 4

Let $\sigma \in \Sigma_{\mathcal{S}}$ be arbitrary. In view of Lemma 12, to prove Theorem 4 it is enough to show that

$$\limsup_{n\to\infty} \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\Big) > 0.$$

Since the network topology has non-expanding subnetworks, there exist some positive integer K, some real number $\varepsilon > 0$,and a subsequence of agents $\mathcal{N}$ such that

$$\mathbb{Q}\Big(\Big|\widehat{B}(n)\Big| < K\Big) \geq \varepsilon \qquad \text{for all } n \in \mathcal{N}. \tag{44}$$

For all $n \in \mathcal{N}$, by the law of total probability we have

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\Big)\\
&= \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K\Big)\mathbb{Q}\Big(\Big|\widehat{B}(n)\Big| < K\Big)\\
&+ \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| \geq K\Big)\mathbb{Q}\Big(\Big|\widehat{B}(n)\Big| \geq K\Big)\\
&\geq \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K\Big)\mathbb{Q}\Big(\Big|\widehat{B}(n)\Big| < K\Big)\\
&\geq \varepsilon\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K\Big),
\end{aligned} \tag{45}$$

where the last inequality follows from (44). By the law of total probability again, we also have

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K\Big)\\
&= \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big)\mathbb{P}_\Omega\Big(\omega \in \Omega^{NS}\Big)\\
&+ \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \notin \Omega^{NS}\Big)\mathbb{P}_\Omega\Big(\omega \notin \Omega^{NS}\Big)\\
&\geq \mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big)\mathbb{P}_\Omega\Big(\omega \in \Omega^{NS}\Big)\\
&\geq \delta\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big),
\end{aligned} \tag{46}$$

where the last inequality holds by (43). Then, by (45) and (46),

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\Big)\\
&\geq \varepsilon\delta\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big)
\end{aligned} \tag{47}$$

holds for all agents $n \in \mathcal{N}$.

Let $\overline{C}_\sigma(q^{NS})$ denote the set of all private search costs for which an agent h with second search stage information set I_h^2 such that $B_h = \emptyset$, $q_{s_h^1} = q^{NS}$, and $c_h \in \overline{C}_\sigma(q^{NS})$, adopts strategy $s_h^2 = ns$ in equilibrium σ. That is, $\overline{C}_\sigma(q^{NS})$ consists of all search costs for which, in equilibrium σ, an agent with empty neighborhood decides not to sample the second action when the first action he samples has quality q^{NS}. For all $\omega \in \Omega^{NS}$, the results of Section 3.2 imply that any agent k with search cost $c_k \in \overline{C}_\sigma(q^{NS})$ adopts strategy $s_k^2 = ns$ at the second search stage independently of his neighborhood realization B_k, the actions of his neighbors, and the quality of the first action sampled (i.e. independently of the realizations of the random variables in his information set other than his search cost). Then,

$$\begin{aligned}&\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big)\\ &\geq \mathbb{P}_\sigma\Big(c_k \in \overline{C}_\sigma\Big(q^{NS}\Big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big).\end{aligned} \tag{48}$$

Moreover, as individual search costs are independent of the network topology and the realized quality of the two actions,

$$\begin{aligned}&\mathbb{P}_\sigma\Big(c_k \in \overline{C}_\sigma\Big(q^{NS}\Big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big)\\ &= \mathbb{P}_\sigma\Big(c_k \in \overline{C}_\sigma\Big(q^{NS}\Big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K\Big).\end{aligned} \tag{49}$$

Finally, as $\Big|\widehat{B}(n)\Big| < K \iff \Big|\widehat{B}(n) \cup \{n\}\Big| \leq K$ and individual search costs are independent of the network topology and i.i.d. across agents, we have

$$\begin{aligned}&\mathbb{P}_\sigma\Big(c_k \in \overline{C}_\sigma\Big(q^{NS}\Big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K\Big)\\ &\geq \mathbb{P}_\sigma\Big(c_1 \in \overline{C}_\sigma\Big(q^{NS}\Big)\Big)^K\\ &> 0,\end{aligned} \tag{50}$$

where the strict inequality holds because $\mathbb{P}_\sigma(c_1 \in \overline{C}_\sigma(q^{NS})) > 0$ by the first condition in Assumption 1. Together, (48), (49), and (50) yield that

$$\mathbb{P}_\sigma\Big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \;\Big|\; \Big|\widehat{B}(n)\Big| < K, \omega \in \Omega^{NS}\Big) > 0. \tag{51}$$

As $\varepsilon, \delta > 0$, from (47) and (51) we conclude

$$\mathbb{P}_\sigma\left(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\right) > 0$$

for all agents n in the subsequence $\mathcal{N}$, which implies

$$\limsup_{n\to\infty} \mathbb{P}_\sigma\left(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\right) > 0,$$

as desired. ■

B.3 Preliminaries for Sections 5 and 6

Characterization of Equilibrium Strategies in OIP Networks

Part *(a)* of Theorem 2 and the results in Section 6 are largely based on the next lemma, which characterizes equilibrium sequential search policies in OIP networks. Let $P_1(q)$ denote the posterior probability that agent 1 did not sample the second action given that the action he takes has quality q. The precise functional form of $P_1(q)$ is irrelevant for the following argument.

Lemma 13. *Let $\mathcal{S}$ be a collective search environment where the network topology features observation of immediate predecessors. Then, in any equilibrium $\sigma \in \Sigma_{\mathcal{S}}$:*

(i) At the first search stage, each agent $n \in \mathbb{N}$, with $n \geq 2$, samples first the action taken by his immediate predecessor. That is, $s_n^1 = a_{n-1}$.

(ii) At the second search stage, each agent n, with $n \geq 2$:

(a) Does not sample action $\neg a_{n-1}$ (i.e. $s_n^2 = ns$) if $\neg a_{n-1}$ is revealed to be inferior to agent n in equilibrium σ.

(b) Samples action $\neg a_{n-1}$ (i.e. $s_n^2 = \neg a_{n-1}$) if $\neg a_{n-1}$ is not revealed inferior to agent n in equilibrium σ, and agent n's search cost c_n is smaller than $t_n(q_{s_n^1})$, where the function $t_n \colon Q \to \mathbb{R}_+$ is defined pointwise by

$$t_n\left(q_{s_n^1}\right) := P_1\left(q_{s_n^1}\right) t^{\emptyset}\left(q_{s_n^1}\right) \tag{52}$$

for $n = 2$, and pointwise recursively as

$$t_n\left(q_{s_n^1}\right) := P_1\left(q_{s_n^1}\right)\left(\prod_{i=2}^{n-1}\left(1 - F_C\left(t_i\left(q_{s_n^1}\right)\right)\right)\right) t^{\emptyset}\left(q_{s_n^1}\right) \tag{53}$$

for $n > 2$.[29]

Proof. To prove part *(i)*, proceed by induction. Consider agent 2 and his conditional belief over Ω given that the first agent has taken action a_1. For action $\neg a_1$, two mutually exclusive cases are possible:

1. Agent 1 sampled $\neg a_1$. In this case, $q_{\neg a_1} \leq q_{a_1}$, as agent 1 picked the best alternative at the choice stage. If agent 2 knew this to be the case, his conditional belief on Ω would be $\mathbb{P}_{\Omega | q_{a_1} \geq q_{\neg a_1}}$.
2. Agent 1 did not sample $\neg a_1$. If agent 2 knew this to be the case, his posterior belief on action $\neg a_1$ would be the same as the prior $\mathbb{P}_Q$.

Then, regardless of the beliefs of agent 2 about agent 1's search decisions, agent 2's belief about the quality of action $\neg a_1$ is strictly first-order stochastically dominated by his beliefs about the quality of action a_1. To see this, note that agent 2 believes that agent 1 has sampled action $\neg a_1$ with positive probability: even if agent 1 sampled a_1 first, by the second condition of Assumption 1, with positive probability, his search costs are low enough that he searched further. Therefore, $s_2^1 = a_1$ is agent 2's optimal policy at the first search stage.

Now consider any agent $n > 2$. Suppose that all agents up to $n-1$ follow this strategy, and that agent $n-1$ selects action a_{n-1}. If action $\neg a_{n-1}$ is revealed inferior to agent n in equilibrium σ, it must be that $q_{\neg a_{n-1}} \leq q_{a_{n-1}}$, and so action $\neg a_{n-1}$ is not sampled at all. Now suppose that action $\neg a_{n-1}$ is not revealed inferior to agent n in equilibrium σ. By the same logic as before, n's beliefs about the quality of action a_{n-1} strictly first-order stochastically dominate his beliefs about the quality of action $\neg a_{n-1}$. Therefore, $s_n^1 = a_{n-1}$, i.e. he will sample action a_{n-1} first.

To establish part *(ii)–(a)*, consider any agent $n \geq 2$, and suppose that $\neg a_{n-1}$ is revealed inferior to agent n in equilibrium σ. Then, there exist $j, j+1 \in B(n)$ such that $a_j = \neg a_{n-1}$ and $a_{j+1} = a_{n-1}$. By part *(i)* we know that $s_{j+1}^1 = \neg a_{n-1}$. Since agents can only take an action they sampled, it follows that $s_{j+1}^2 = a_{n-1}$, that is, agent $j+1$ has sampled both actions. Then, as agents take the best action whenever they sample both of them, we have $q_{a_{n-1}} \geq q_{\neg a_{n-1}}$, and so the expected additional gain of sampling action $\neg a_{n-1}$ is zero. That $s_n^2 = ns$ is optimal follows.

For part *(ii)–(b)*, consider any agent $n \geq 2$ and suppose that $\neg a_{n-1}$ is not revealed inferior to agent n in equilibrium σ. In OIP networks, the personal subnetwork of agent n, $\widehat{B}(n)$, is

[29] Hereafter, I assume that agent n samples the second action in case of indifference. This assumption does not affect the results, but simplifies the derivation of closed form expressions for the $t_n(\cdot)$'s and the ensuing analysis.

$\{1, \ldots, n-1\}$ with probability one. Moreover, by part *(i)*, each agent samples first the action taken by his immediate predecessor. Therefore, none of the agents in the personal subnetwork of agent n relative to action s_n^1 has sampled action $\neg s_n^1$ only if none of the first $n-1$ agents has sampled it; that is, only if $s_1^1 = s_n^1$, and $s_i^2 = ns$ for $1 \leq i \leq n-1$. The thresholds in (52) and (53) provide an explicit formula for (9) when $\widehat{B}(n) = \{1, \ldots, n-1\}$ with probability one for all $n \in \mathbb{N}$. To see this, proceed by induction. Consider first agent 2. By part *(i)*, $s_2^1 = a_1$. Let $P_1(q_{s_2^1})$ be the posterior probability that agent 1 did not sample action $\neg s_2^1$ given that action s_2^1 of quality $q_{s_2^1}$ was taken. Then, agent 2's expected benefit from the second search is $P_1(q_{s_2^1}) t^{\emptyset}(q_{s_2^1})$, which is the right-hand side of (52). Now consider any agent $n > 2$, and let s_n^1 be the action this agent samples first. By part *(i)* and the inductive hypothesis, and since search costs are i.i.d. across agents, it follows that the probability that no agent in $\{1, \ldots, n-1\}$ has sampled action $\neg s_n^1$ is

$$P_1\Big(q_{s_n^1}\Big)\left(\prod_{i=2}^{n-1}\Big(1 - F_C\Big(t_i\Big(q_{s_n^1}\Big)\Big)\Big)\right).$$

Therefore, the right-hand side of (53) gives agent n's expected benefit from the search follows. The optimality of the proposed sequential search policy follows from the characterization of individual equilibrium decisions at the second search stage in Section 3.2.2. ■

Fix a state process and a search technology. Lemma 13 implies that, from the viewpoint of the probability of selecting the best action, the individual search behavior is equivalent across all OIP networks. In particular, we have the following.

Corollary 1. *Let $\mathcal{S}$ and $\mathcal{S}'$ be two collective search environments with identical state process and search technology. Assume that $\mathcal{S}$ is endowed with the complete network, while the network topology of $\mathcal{S}'$ is any OIP network. Finally, let $\sigma \in \Sigma_{\mathcal{S}}$ and $\sigma' \in \Sigma_{\mathcal{S}'}$, and assume that ties are broken according to the same criterion in σ and σ'.*[30] *Then, for all $n \in \mathbb{N}$,*

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x \in X}\ q_x\right) = \mathbb{P}_{\sigma'}\left(a_n \in \arg\max_{x \in X}\ q_x\right).$$

Proof. In OIP networks, each agent starts sampling from the action taken by his immediate predecessor (cf. Lemma 13), and so asymptotic learning trivially occurs when agent 1 takes the best action. Moreover, $\mathbb{P}_\sigma(a_1 \in \arg\max_{x\in X} q_x) = \mathbb{P}_{\sigma'}(a_1 \in \arg\max_{x\in X} q_x)$. Therefore, to establish the result, it suffices to show that $\mathbb{P}_\sigma(a_n \in \arg\max_{x\in X} q_x) = \mathbb{P}_{\sigma'}(a_n \in \arg\max_{x\in X} q_x)$ holds for all $n \in \mathbb{N}$, with $n > 2$, whenever agent 1 does not sample the best action at the first search. In turn, this follows immediately from Lemma 13, which shows that, for all n, the probability that none of the first n agents has sampled both actions is the same across all OIP networks for any fixed quality of the action taken by agent 1. ■

B.4 Proofs for Section 5.5

Proof of Theorem 2

Proof of part *(a)*. Suppose $\omega \notin \overline{\Omega}(\underline{c})$, and that the lowest cost in the support of $\mathbb{P}_C$ is $\underline{c} > 0$. Maximal learning requires that the probability that agent n takes the action with the highest

[30] In particular, assume that agent 1 selects uniformly at random the first action to sample, and that agent n samples the second action in case of indifference.

quality converges to one as $n \to \infty$ (see the characterization of maximal learning in (12)). This, in turn, is equivalent to saying that the probability of the event "none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions" converges to zero as $n \to \infty$ whenever the quality of the first action sampled by agent 1 is lower than $q(\underline{c})$.[31] To establish the failure of maximal learning, I show that the probability of this event remains bounded away from zero when $\underline{c} > 0$.

By way of contradiction, suppose that the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to zero as $n \to \infty$ for any quality q, with $q < q(\underline{c})$, that the first action sampled by agent 1 can take. That is,

$$\lim_{n\to\infty} P_1(q)\left(\prod_{i=2}^{n}\Big(1 - F_C(t_i(q))\Big)\right) = 0$$

(see the proof of Lemma 13 for how to derive this probability). It follows that the expected additional gain from the second search for agent $n+1$, given by

$$P_1(\hat{q})\left(\prod_{i=2}^{n}\Big(1 - F_C(t_i(\hat{q}))\Big)\right)t^{\emptyset}(\hat{q})$$

(see 53 and the proof of Lemma 13), where $\hat{q}$ is the quality of the action taken by agent n, also converges to zero as $n \to \infty$ for all $\hat{q} < q(\underline{c})$. Then, there exists an agent $N_{\hat{q}} + 1$ for which the expected additional gain from the second search falls below $\underline{c}$.

By Assumption 2, there exists $\tilde{q}$ in the support of $\mathbb{P}_Q$ such that:

(i) $\mathbb{P}_Q(\tilde{q} < q < q(\underline{c})) > 0$;

(ii) With positive probability, the first agent does not sample another action if $q_{s_1^1} \geq \tilde{q}$, that is

$$1 - F_C\Big(t^{\emptyset}(\tilde{q})\Big) > 0.$$

Therefore, with positive probability, agent 1 samples first a suboptimal action with quality, say, $\overline{q}$, and does not search further. Now suppose that the first $N_{\overline{q}}$ agents all have costs larger than $t^{\emptyset}(\overline{q})$, and again note that this occurs with positive probability. By Lemma 13, each of these agents will sample the suboptimal action with quality $\overline{q}$ first, and none of these agents will search further. Therefore, all will take this suboptimal action. Agent $N_{\overline{q}}+1$ also samples this action first, and does not search further either because his expected additional gain from the second search is smaller than $\underline{c}$. Since the expected additional gain from the second search in non-increasing in n, there will be no further search by agents $N_{\overline{q}} + 1$ onward, contradicting that the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to zero. The desired result follows. ■

Proof of part *(b)*. Suppose $\omega \notin \overline{\Omega}(\underline{c})$, and that the lowest cost in the support of $\mathbb{P}_C$ is $\underline{c} > 0$. Again, I establish that maximal learning fails because the probability of the event "none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions" remains bounded away from zero as $n \to \infty$.

Pick an infinite sequence of agents $(\pi_1, \pi_2, \ldots, \pi_k, \pi_{k+1}, \ldots)$ such that $B(\pi_1) = \emptyset$ and $\pi_k \in B(\pi_{k+1})$ for all agents $k \in \mathbb{N}$. Such a sequence must exist with probability one; otherwise, the network topology has non-expanding subnetworks and maximal learning fails. Moreover, by Lemma 4, each agent in this sequence samples first the action taken by his neighbor.

[31] By assumption, $\omega \notin \overline{\Omega}(\underline{c})$, and so $\min\{q_0, q_1\} < q(\underline{c})$. Therefore, with positive probability, the quality of the first action sampled by agent 1 is lower than $q(\underline{c})$.

By way of contradiction, suppose that the probability of no agent in $\widehat{B}(\pi_k) \cup \{\pi_k\}$ sampling both actions converges to zero as $k \to \infty$ for any quality q, with $q < q(\underline{c})$, that the first action sampled by agent π_1 can take. That is,

$$\lim_{k\to\infty} P_{\pi_{k+1}}(q) = 0,$$

where $P_{\pi_{k+1}}(\cdot)$ is the function defined by (8). It follows that the expected additional gain from the second search for agent π_{k+1}, given by

$$P_{\pi_{k+1}}(\hat{q})t^{\emptyset}(\hat{q}),$$

where $\hat{q}$ is the quality of the action taken by π_k, also converges to zero as $k \to \infty$ for all $\hat{q} < q(\underline{c})$. Then, there exists an agent $\pi_{K_{\hat{q}}} + 1$ for which the expected additional gain from the second search falls below $\underline{c}$, and remains below this threshold for the other agents in the sequence moving after $\pi_{K_{\hat{q}}} + 1$.

By Assumption 2, there exists $\tilde{q}$ in the support of $\mathbb{P}_Q$ such that:

(i) $\mathbb{P}_Q(\tilde{q} < q < q(\underline{c})) > 0$;

(ii) With positive probability, agent π_1 does not sample another action if $q_{s^1_{\pi_1}} \geq \tilde{q}$, that is

$$1 - F_C\Big(t^{\emptyset}(\tilde{q})\Big) > 0.$$

Therefore, with positive probability, agent π_1 samples first a suboptimal action with quality, say, $\overline{q}$, and does not search further. Now suppose that the first $\pi_{K_{\hat{q}}}$ agents in the sequence all have costs larger than $t^{\emptyset}(\overline{q})$, and again note that this occurs with positive probability. By Lemma 4, each of these agents will sample the suboptimal action with quality $\overline{q}$ first, and none of these agents will search further. Therefore, all will take this suboptimal action. Agent $\pi_{K_{\hat{q}}} + 1$ also samples this action first, and does not search further either because his expected additional gain from the second search is smaller than $\underline{c}$. Since the expected additional gain from the second search remains smaller than $\underline{c}$ afterward, there will be no further search by agents in the sequence moving after agent $\pi_{K_{\hat{q}}} + 1$, contradicting that the probability of no agent in $\widehat{B}(\pi_k) \cup \{\pi_k\}$ sampling both actions converges to zero. The desired result follows. ■

B.5 Proofs for Section 6.2

Proof of Proposition 6

The result follows by combining Corollary 1 with Proposition 1 in MFP. ■

Proof of Proposition 7

Proof of part *(a)*. The result follows by combining Corollary 1 with Proposition 2 in MFP.

Proof of part *(b)*. To establish the result, it is enough to construct a function $\tilde{\phi}\colon \mathbb{R}_+ \to \mathbb{R}$ such that, for all $n \in \mathbb{N}$,

$$\mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X} q_x\right) \geq \tilde{\phi}(n) \qquad \text{and} \qquad 1 - \tilde{\phi}(n) = O\left(\frac{1}{n^{\frac{1}{K+1}}}\right).$$

Consider the sequence of neighbor choice function $(\gamma_n)_{n\in\mathbb{N}}$ where, for all $n \in \mathbb{N}$, $\gamma_n = n-1$. Under the assumptions of the proposition, by Lemmas 7 and 13,

$$\begin{aligned}\mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x\in X}\, q_x\right) &\geq \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X}\, q_x\right)\\ &+ \left(1 - \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X}\, q_x\right)\right)^2 F_C\left(\left(1 - \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X}\, q_x\right)\right) t^{\emptyset}(q_{\max})\right).\end{aligned} \tag{54}$$

If the search cost distribution has polynomial shape, from (54) we have

$$\begin{aligned}\mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x\in X}\, q_x\right) &\geq \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X}\, q_x\right)\\ &+ Lt^{\emptyset}(q_{\max})^K\left(1 - \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X}\, q_x\right)\right)^{K+2}.\end{aligned} \tag{55}$$

From this point forward, I build on Lobel et al. (2009) (see their proof of Proposition 2) to construct the function $\tilde{\phi}$. A simple adaptation of their procedure to my setup gives that the function $\tilde{\phi}$ we are looking for is

$$\tilde{\phi}(n) = 1 - \left(\frac{1}{(K+1)Lt^{\emptyset}(q_{\max})^K\left(n+\overline{K}\right)}\right)^{\frac{1}{K+1}},$$

where $\overline{K}$ is some constant of integration (in the construction, $\tilde{\phi}$ is found as the solution to an ordinary differential equation).[32] ■

Proof of Proposition 8

To establish the result, it is enough to construct a function $\tilde{\phi}\colon \mathbb{R}_+ \to \mathbb{R}$ such that, for all $n \in \mathbb{N}$,

$$\mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X}\, q_x\right) \geq \tilde{\phi}(n) \qquad \text{and} \qquad 1 - \tilde{\phi}(n) = O\left(\frac{1}{(\log n)^{\frac{1}{K+1}}}\right).$$

Under the assumptions of of the proposition,

$$\begin{aligned}\mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x\in X}\, q_x\right) &= \frac{1}{n}\sum_{b=1}^{n} \mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x\in X}\, q_x \mid B(n+1) = \{b\}\right)\\ &= \frac{1}{n}\left[\mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x\in X}\, q_x \mid B(n+1) = \{n\}\right) + (n-1)\mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X}\, q_x\right)\right]\end{aligned} \tag{56}$$

because conditional on observing the same $b < n$, agents n and $n+1$ have identical probabilities of making an optimal decision. By Lemmas 7 and 4, and since the search cost distribution has

[32] To apply a construction in the spirit of Lobel et al. (2009), the right-hand side of (55) must be increasing in $\mathbb{P}_\sigma(s^1_n \in \arg\max_{x\in X} q_x)$. This is so under the assumption $0 < L < \frac{2^{K+1}}{(K+2)t^{\emptyset}\left(\underline{q}\right)^K}$ maintained in the proposition.

polynomial shape, we obtain that

$$\mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x\in X} q_x\right) \geq \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X} q_x\right) + \frac{Lt^\emptyset(q_{\max})^K}{n}\left(1-\mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x\in X} q_x\right)\right)^{K+2}. \tag{57}$$

As for the proof of Proposition 7-part *(b)*, from this point forward, I build on Lobel et al. (2009) (see their proof of Proposition 3) to construct the function $\tilde{\phi}$. A straightforward adaptation of their procedure to my setup gives that the function $\tilde{\phi}$ we are looking for is

$$\tilde{\phi}(n) = 1 - \left(\frac{1}{(K+1)Lt^\emptyset(q_{\max})^K\left(\log n + \overline{K}\right)}\right)^{\frac{1}{K+1}},$$

where $\overline{K}$ is some constant of integration (in the construction, $\tilde{\phi}$ is found as the solution to an ordinary differential equation).[33] ■

B.6 Proofs for Section 6.3

Preliminaries

First, I define the objects and the notation that will be used in the proofs of Propositions 9 and 10.

▷ Let $\mathcal{S}$ and $\mathcal{S}'$ be two collective search environments with identical state process $(Q, \mathcal{F}_Q, \mathbb{P}_Q)$ and search technology $\{(C, \mathcal{F}_C, \mathbb{P}_C), \mathcal{R}\}$. Suppose that the network topology of $\mathcal{S}$ is the complete network and that in $\mathcal{S}'$ agents only observe their most immediate predecessor. Let $\sigma \in \Sigma_{\mathcal{S}}$ and $\sigma' \in \Sigma_{\mathcal{S}'}$. Suppose that agents break ties according to the same criterion in σ and σ'. In particular, assume that agent 1 selects uniformly at random which action to sample first, and that all agents sample the other action whenever indifferent at the second search stage.[34] Suppose also that the first action sampled by the first agent in σ and σ', say x, has the same quality q_x. Let $\delta \in (0,1)$ be the discount rate, and let the function $t_1\colon Q \to \mathbb{R}_+$ be defined pointwise by $t_1(q) := t^\emptyset(q)$.[35] Hereafter, $q_{\neg x}$ is a random variable with probability measure $\mathbb{P}_Q$.

▷ The expected discounted social utility normalized by $(1-\delta)$ in equilibrium σ, denoted by $U_\sigma(q_x;\delta)$, is

$$\begin{aligned} U_\sigma(q_x;\delta) = q_x + t_1(q_x) &- (1-\delta)\sum_{n=1}^{\infty}\delta^n\left(\prod_{i=1}^{n}\Big(1-F_C\big(t_i(q_x)\big)\Big)\right)t_1(q_x) \\ &- (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_x)\Big]F_C(t_n(q_x))\prod_{i=1}^{n-1}\Big(1-F_C(t_i(q_x))\Big) \\ &- (1-\delta)\mathbb{P}_Q(q_{\neg x} \leq q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_x)\Big]F_C(t_n(q_x)). \end{aligned} \tag{58}$$

[33] To apply a construction in the spirit of Lobel et al. (2009), the right-hand side of (57) must be increasing in $\mathbb{P}_\sigma(s^1_n \in \arg\max_{x\in X} q_x)$. This is so under the assumption $0 < L < \frac{2^{K+1}}{(K+2)t^\emptyset\left(\underline{q}\right)^K}$ maintained in the proposition.

[34] This assumption simplifies the notation, but does not qualitatively affect the results.

[35] Redefining function $t^\emptyset$ with t_1 simplifies the notation in the following analysis.

To see this note that the first term is the quality of the first action sampled, and the second term is the additional gain from the second unsampled action. From this, we subtract the sum of the period n discounted gain from the unsampled action times the probability it was not sampled from period 1 to n. Further, we subtract the expected discounted cost of search, which consists of two parts. The first part,

$$(1-\delta)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_x)\Big]F_C(t_n(q_x))\prod_{i=1}^{n-1}\Big(1-F_C(t_i(q_x))\Big),$$

is the expected discounted cost of search when $q_{\neg x} > q_x$. In this case, after agent n samples both actions, action x is revealed to be inferior in equilibrium to all agents moving after agent n. Therefore, no agent $m > n$ will sample action x again. The second part,

$$(1-\delta)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_x)\Big]F_C(t_n(q_x)),$$

is the expected discounted cost of search when $q_{\neg x} \leq q_x$. In this case, after agent n samples both actions, action $\neg x$ is inferior in equilibrium, but not revealed to be so to the agents moving after agent n. Therefore, all agents $m > n$ with $c_m \leq t_m(q_x)$ will sample action $\neg x$ again.

The expected discounted social utility normalized by $(1-\delta)$ in equilibrium σ', denoted by $U_{\sigma'}(q_x;\delta)$, is

$$\begin{aligned}
U_{\sigma'}(q_x;\delta) = {} & q_x + t_1(q_x) - (1-\delta)\sum_{n=1}^{\infty}\delta^n\left(\prod_{i=1}^{n}\Big(1-F_C\Big(t_i\Big(q_x\Big)\Big)\Big)\right)t_1(q_x) \\
& - (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_x)\Big]F_C(t_n(q_x))\prod_{i=1}^{n-1}\Big(1-F_C(t_i(q_x))\Big) \\
& - (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_Q}\Big[\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_{\neg x})\Big]F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\Big] \\
& \cdot\left(1-\prod_{i=1}^{n-1}\Big(1-F_C(t_i(q_x))\Big)\right) \\
& - (1-\delta)\mathbb{P}_Q(q_{\neg x} \leq q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_x)\Big]F_C(t_n(q_x)).
\end{aligned} \tag{59}$$

$U_{\sigma'}(q_x;\delta)$ has the same interpretation as $U_{\sigma}(q_x;\delta)$, except for the expected discounted cost of search when $q_{\neg x} > q_x$, which is now

$$\begin{aligned}
& (1-\delta)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_x)\Big]F_C(t_n(q_x))\prod_{i=1}^{n-1}\Big(1-F_C(t_i(q_x))\Big) \\
& + (1-\delta)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_Q}\Big[\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_{\neg x})\Big]F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\Big]\left(1-\prod_{i=1}^{n-1}\Big(1-F_C(t_i(q_x))\Big)\right).
\end{aligned}$$

When agents only observe their most immediate predecessor, they also fail to recognize actions that are revealed to be inferior in equilibrium by the time of their move. Therefore, in contrast with what happens in the complete network, even if agent n samples both actions and $q_{\neg x} > q_x$, all agents $m > n$ with $c_m \leq t_m(q_{\neg x})$ will now sample action x again. Since the quality of action $\neg x$ is unknown (q_x is fixed, but $q_{\neg x}$ is a random variable), the expected cost of this additional search is

$$\mathbb{E}_{\mathbb{P}_Q}\Big[\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_{\neg x})\Big] F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\Big].$$

▷ Now consider a third collective search environment $\mathcal{S}''$ with the same state process and search technology as in $\mathcal{S}$ and $\mathcal{S}'$, but where the network topology is any OIP network. Let $\sigma'' \in \Sigma_{\mathcal{S}''}$, and suppose that indifferences are resolved in σ'' according to the same tie-breaking criterion as in σ and σ'. Assume also that the first action sampled by agent 1 in σ'', say x, has the same quality q_x as the action sampled at the first search by agent 1 in σ, σ'. Denote with $U_{\sigma''}(q_x;\delta)$ the expected discounted social utility normalized by $(1-\delta)$ in equilibrium σ''. Again, assume that the single decision maker selects the first action to sample uniformly at random, and that he samples the second action in case of indifference. The next lemma is immediate from the discussion in Section 6.3.

Lemma 14. *For all $q_x \in Q$ and $\delta \in (0,1)$, we have*

$$U_\sigma(q_x;\delta) \geq U_{\sigma''}(q_x;\delta) \geq U_{\sigma'}(q_x;\delta).$$

▷ Finally, denote with $U_{DM}(q_x;\delta)$ the expected discounted social utility normalized by $(1-\delta)$ that is implemented by the single decision maker in any OIP network after sampling an action, say x, of quality q_x at the first search at time period 1. Again, assume that the single decision maker selects the action to sample first uniformly at random at time period 1, and that he samples the second action whenever indifferent. I refer to Section III.A. in MFP for the derivation of $U_{DM}(q_x;\delta)$. Since the single decision maker's problem is the same in all OIP networks, of which the complete network is an example, the same analysis applies unchanged in my setting.

Proof of Proposition 9

The difference in average social utilities is

$$\begin{aligned} &U_\sigma(q_x;\delta) - U_{\sigma'}(q_x;\delta) \\ &\quad= (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n \mathbb{E}_{\mathbb{P}_Q}\Big[\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_{\neg x})\Big] F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\Big] \\ &\quad\cdot \left(1 - \prod_{i=1}^{n-1}\Big(1 - F_C(t_i(q_x))\Big)\right). \end{aligned} \tag{60}$$

The right-hand side of (60) is positive for all $\delta \in (0,1)$. That $U_\sigma(q_x;\delta) > U_{\sigma'}(q_x;\delta)$ for all $\delta \in (0,1)$ follows.

To show that

$$\lim_{\delta \to 1}\Big[U_\sigma(q_x;\delta) - U_{\sigma'}(q_x;\delta)\Big] = 0,$$

we need to show that the right-hand side of (60) converges to zero as $\delta \to 1$. To do so, it is enough to argue that

$$\sum_{n=1}^{\infty}\delta^n \mathbb{E}_{\mathbb{P}_Q}\Big[\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_{\neg x})\Big] F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\Big]$$

is finite. Notice that

$$0 \leq \sum_{n=1}^{\infty}\delta^n \mathbb{E}_{\mathbb{P}_Q}\Big[\mathbb{E}_{\mathbb{P}_C}\Big[c \mid c \leq t_n(q_{\neg x})\Big] F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\Big]$$

$$
\begin{aligned}
&\leq \sum_{n=1}^{\infty} \delta^n \mathbb{E}_{\mathbb{P}_Q}\Big[t_n(q_{\neg x})F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\Big] \\
&\leq \sum_{n=1}^{\infty} \delta^n \sup_{q>q_x} t_n(q)F_C(t_n(q)) \\
&\leq \sum_{n=\bar{n}+1}^{\infty} \delta^n \sup_{q>q_x} t_n(q)F_C(t_n(q)) + \bar{n} \sup_{q>q_x} t^{\emptyset}(q) \\
&\approx \sum_{n=\bar{n}+1}^{\infty} \delta^n \sup_{q>q_x} \Big(t_n(q)\Big)^2 f_C(0) + \bar{n} \sup_{q>q_x} t^{\emptyset}(q) \\
&\approx \sum_{n=\bar{n}+1}^{\infty} \delta^n \sup_{q>q_x} \Big(t^{\emptyset}(q)\Big)^2 \frac{1}{f_C(0)t^2} + \bar{n} \sup_{q>q_x} t^{\emptyset}(q),
\end{aligned}
$$

where $\bar{n}$ is large enough for $t_n(q)$ to be close to 0. Since $\sum_{n=\bar{n}+1}^{\infty} \frac{1}{n^2}$ and $\bar{n} \sup_{q>q_x} t^{\emptyset}(q)$ are finite, the desired result follows. ■

Proof of Proposition 10

First, suppose $\underline{c} = 0$. Wee need to show that $\lim_{\delta \to 1} U_{\sigma''}(q_x; \delta) = \lim_{\delta \to 1} U_{DM}(q_x; \delta)$. By Proposition 9, $\lim_{\delta \to 1} U_{\sigma}(q_x; \delta) = \lim_{\delta \to 1} U_{\sigma'}(q_x; \delta)$. Moreover, by Lemma 14, $U_{\sigma}(q_x; \delta) \geq U_{\sigma''}(q_x; \delta) \geq U_{\sigma'}(q_x; \delta)$ Therefore, by the sandwich theorem for limits of functions,

$$
\lim_{\delta \to 1} U_{\sigma''}(q_x; \delta) = \lim_{\delta \to 1} U_{\sigma}(q_x; \delta). \tag{61}
$$

By Proposition 3 in MFP,

$$
\lim_{\delta \to 1} U_{\sigma}(q_x; \delta) = \lim_{\delta \to 1} U_{DM}(q_x; \delta). \tag{62}
$$

Then, by (61) and (62), and the uniqueness of the limit of a function, we have

$$
\lim_{\delta \to 1} U_{\sigma''}(q_x; \delta) = \lim_{\delta \to 1} U_{DM}(q_x; \delta),
$$

which gives the desired result.

Now suppose that $\lim_{\delta \to 1} U_{\sigma''}(q_x; \delta) = \lim_{\delta \to 1} U_{DM}(q_x; \delta)$. We need to show that $\underline{c} = 0$. Since the complete network is an OIP network, it follows that $\lim_{\delta \to 1} U_{\sigma}(q_x; \delta) = \lim_{\delta \to 1} U_{DM}(q_x; \delta)$. That $\underline{c} = 0$ immediately follows by Proposition 3 in MFP. ■

B.7 Proofs for Section 6.4

Proof of Proposition 11

An inductive argument analogous to the one establishing part *(i)* of Lemma 13 shows that each agent starts sampling from the action taken by his immediate predecessor. Then, the result follows directly from the discussion in Section 6.4. ■

References

Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar (2014), “Dynamics of Information Exchange in Endogenous Social Networks.” *Theoretical Economics*, 9, 41–97.

Acemoglu, Daron, Munther A. Dahleh, Ilan Lobel, and Asuman Ozdaglar (2011), “Bayesian Learning in Social Networks.” *The Review of Economic Studies*, 78, 1201–1236.

Acemoglu, Daron and Asuman Ozdaglar (2011), “Opinion Dynamics and Learning in Social Networks.” *Dynamic Games and Applications*, 1, 1–43.

Acemoglu, Daron, Asuman Ozdaglar, and Ali ParandehGheibi (2010), “Spread of (Mis)information in Social Networks.” *Games and Economic Behavior*, 70, 194–227.

Ali, S. Nageeb (2018), “Herding with Costly Information.” *Journal of Economic Theory*, 175, 713–729.

Arieli, Itai and Manuel Mueller-Frank (2018), “Multi-Dimensional Social Learning.” *Review of Economic Studies (forthcoming).*

Armstrong, Mark (2016), “Ordered Consumer Search.” *Working Paper.*

Bala, Venkatesh and Sanjeev Goyal (1998), “Learning from Neighbours.” *The Review of Economic Studies*, 65, 595–621.

Banerjee, Abhijit V. (1992), “A Simple Model of Herd Behavior.” *The Quarterly Journal of Economics*, 107, 797–817.

——— (1993), “The Economics of Rumors.” *The Review of Economic Studies*, 60, 309–327.

Banerjee, Abhijit V. and Drew Fudenberg (2004), “Word-of-Mouth Learning.” *Games and Economic Behavior*, 46, 1–22.

Bergemann, Dirk and Juuso Välimäki (2008), “Bandit Problems.” Steven N. Durlauf and Lawrence E. Blume eds., *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), “A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades.” *Journal of Political Economy*, 100, 992–1026.

Bolton, Patrick and Christopher Harris (1999), “Strategic Experimentation.” *Econometrica*, 67, 349–374.

Burguet, Robert and Xavier Vives (2000), “Social Learning and Costly Information Acquisition.” *Economic Theory*, 15, 185–205.

Celen, Bogachan and Kyle Hyndman (2012), “Social Learning Through Endogenous Information Acquisition: An Experiment.” *Management Science*, 158, 1525–1548.

Celen, Bogachan and Shachar Kariv (2004), “Observational Learning under Imperfect Information.” *Games and Economic Behavior*, 47, 72–86.

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

Che, Yeon-Koo and Johannes Hörner (2018), "Recommender Systems as Mechanisms for Social Learning." *The Quarterly Journal of Economics*, 133, 871–925.

DeGroot, Morris H. (1974), "Reaching a Consensus." *Journal of the American Statistical Association*, 69, 118–121.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel (2003), "Persuasion Bias, Social Influence, and Unidimensional Opinions." *The Quarterly Journal of Economics*, 118, 909–968.

Doval, Laura (2018), "Whether or Not to Open Pandora's Box." *Journal of Economic Theory*, 175, 127–158.

Ellison, Glenn and Drew Fudenberg (1995), "Word-of-Mouth Communication and Social Learning." *The Quarterly Journal of Economics*, 110, 93–125.

Gale, Douglas and Shachar Kariv (2003), "Bayesian Learning in Social Networks." *Games and Economic Behavior*, 45, 329–346.

Golub, Benjamin and Matthew Jackson (2010), "Naïve Learning in Social Networks and the Wisdom of Crowds." *American Economic Journal: Microeconomics*, 2, 112–149.

——— (2012), "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *The Quarterly Journal of Economics*, 127, 1287–1338.

Golub, Benjamin and Evan Sadler (2016), "Learning in Social Networks." Yann Bramoullé, Andrea Galeotti, and Brian Rogers eds., *The Oxford Handbook of the Economics of Network*. Oxford University Press.

Goyal, Sanjeev (2007), *Connections: An Introduction to the Economics of Networks*. Princeton University Press.

——— (2011), "Learning in Networks." Jess Benhabib, Alberto Bisin, and Matthew O. Jackson eds., *Handbook of Social Economics*. North-Holland.

Hann-Caruthers, Wade, Vadim V. Martynov, and Omer Tamuz (2018), "The Speed of Sequential Asymptotic Learning." *Journal of Economic Theory*, 173, 383–409.

Heidhues, Paul, Sven Rady, and Philipp Strack (2015), "Strategic Experimentation with Private Payoffs." *Journal of Economic Theory*, 159, 531–551.

Hendricks, Kenneth, Alan Sorensen, and Thomas Wiseman (2012), "Observational Learning and Demand for Search Goods." *American Economic Journal: Microeconomics*, 4, 1–31.

Huang, Jin (2017), "To Glance or to Peruse: Observational and Active Learning from Peer Consumers." *Working Paper*.

Jackson, Matthew O. (2008), *Social and Economic Networks*. Princeton University Press.

Jadbabaie, Ali, Pooya Molavi, Alvaro Sandroni, and Alireza Tahbaz-Salehi (2012), "Non-Bayesian Social Learning." *Games and Economic Behavior*, 76, 210–225.

Keller, Godfrey, Sven Rady, and Martin Cripps (2005), "Strategic Experimentation with Exponential Bandits." *Econometrica*, 73, 39–68.

Kremer, Ilan, Yishay Mansour, and Motty Perry (2014), "Implementing the "Wisdom of the Crowd"." *Journal of Political Economy*, 122, 988–1012.

Kultti, Klaus and Paavo Miettinen (2006), "Herding with Costly Information." *International Game Theory Review*, 8, 21–31.

——— (2007), "Herding with Costly Observation." *The B.E. Journal of Theoretical Economics*, 7, 1 (Topics), Article 28.

Lobel, Ilan, Daron Acemoglu, Munther A. Dahleh, and Asuman Ozdaglar (2009), "Rate of Convergence of Learning in Social Networks." *Proceedings of the American Control Conference, ACC 2009.*

Lobel, Ilan and Evan Sadler (2015), "Information Diffusion in Networks through Social Learning." *Theoretical Economics*, 10, 807–851.

——— (2016), "Preferences, Homophily, and Social Learning." *Operations Research*, 64, 564–584.

Mansour, Yishay, Aleksandrs Slivkins, and Vasilis Syrgkanis (2015), "Bayesian Incentive-Compatible Bandit Exploration." *16th ACM Conference on Economics and Computation, EC 2015.*

Mansour, Yishay, Aleksandrs Slivkins, Vasilis Syrgkanis, and Zhiwei Steven Wu (2016), "Bayesian Exploration: Incentivizing Exploration in Bayesian Games." *17th ACM Conference on Economics and Computation, EC 2016.*

Mobius, Markus and Tanya Rosenblat (2014), "Social Learning in Economics." *Annual Review of Economics*, 6, 827–847.

Molavi, Pooya, Alireza Tahbaz-Salehi, and Ali Jadbabaie (2018), "A Theory of Non-Bayesian Social Learning." *Econometrica*, 86, 445–490.

Moscarini, Giuseppe and Lones Smith (2001), "The Optimal Level of Experimentation." *Econometrica*, 69, 1629–1644.

Mossel, Elchanan, Allan Sly, and Omer Tamuz (2015), "Strategic Learning and the Topology of Social Networks." *Econometrica*, 83, 1755–1794.

Mueller-Frank, Manuel (2013), "A General Framework for Rational Learning in Social Networks." *Theoretical Economics*, 8, 1–40.

Mueller-Frank, Manuel and Mallesh M. Pai (2016), "Social Learning with Costly Search." *American Economic Journal: Microeconomics*, 8, 83–109.

Nei, Stephen (2016), "Frictions to Information Aggregation in Social Learning Environments." *Working Paper.*

Papanastasiou, Yiangos, Kostas Bimpikis, and Nicos Savva (2018), "Crowdsourcing Exploration." *Management Science*, 64, 1727–1746.

Perego, Jacopo and Sevgi Yuksel (2016), "Searching for Information and the Diffusion of Knowledge." *Working Paper*.

Rosenberg, Dinah and Nicolas Vieille (2017), "On the Efficiency of Social Learning." *Working Paper*.

Sadler, Evan (2014), "Bounded Rationality and the Network Economist's Dilemma." *Working Paper*.

——— (2017), "Innovation Adoption and Collective Experimentation." *Working Paper*.

Salish, Mirjam (2017), "Strategic Experimentation in Networks." *Working Paper*.

Smith, Lones and Peter Norman Sørensen (2000), "Pathological Outcomes of Observational Learning." *Econometrica*, 68, 371–398.

——— (2014), "Rational Social Learning with Random Sampling." *Working Paper*.

Smith, Lones, Peter Norman Sørensen, and Jianrong Tian (2017), "Informational Herding, Optimal Experimentation, and Contrarianism." *Working Paper*.

Song, Yangbo (2016), "Social Learning with Endogenous Observation." *Journal of Economic Theory*, 166, 324–333.

Vives, Xavier (2010), *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton University Press.

Wald, Ábrahám (1947), "Foundations of a General Theory of Sequential Decision Functions." *Econometrica*, 15, 279–313.

Weitzman, Martin L. (1979), "Optimal Search for the Best Alternative." *Econometrica*, 43, 641–654.