Collective Search in Networks*

Niccolò Lomys†

August 30, 2020

Abstract

I study social learning in networks with information acquisition and choice. Bayesian agents act in sequence, observe the choices of their connections, and acquire information via sequential search. Complete learning occurs if search costs are not bounded away from zero and the network is sufficiently connected and has identifiable information paths. If search costs are bounded away from zero, complete learning is possible in many stochastic networks, including almost-complete networks, but even a weaker notion of long-run learning fails in many other networks. When agents observe random numbers of immediate predecessors, the rate of convergence, the probability of wrong herds, and long-run efficiency properties are the same as in the complete network. The density of indirect connections affects convergence rates. Network transparency has short-run implications for welfare and efficiency. Simply letting agents observe the shares of earlier choices reduces inefficiency and welfare losses.

Keywords: Networks; Bayesian Learning; Improvement and Large-Sample Principles; Speed and Efficiency of Social Learning; Search; Bandit Problems; Information Choice.

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

1 Introduction

Individuals increasingly gather information by using search engines or social media's search bars. Google alone receives several billion search queries per day. In such environments, search is never conducted in isolation: others' choices are readily available via online social networks and popularity rankings. For instance, Facebook users observe which movies their friends watch and which restaurants they go to via their check-ins, the photos they share, or the Facebook pages they like. Similarly, Spotify users observe what songs their friends

*I thank Johannes Hörner, Chiara Margaria, and Emanuele Tarantino for excellent feedback. I had valuable conversations with Paolo Conteduca, Daniel Garrett, Amir Habibi, Takakazu Honryo, Nicolas Lambert, Raphaël Levy, Lorenzo Magnolfi, Matthew Mitchell, Pauli Murto, Alexei Parakhonyak, Evan Sadler, Konrad Stahl, Bruno Strulovici, Takuro Yamashita, and Leeat Yariv. I thank various seminar and conference audiences for helpful comments. I acknowledge funding from the European Research Council Grant Nº 714147 and from the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

†Toulouse School of Economics, University of Toulouse Capitole; niccolo.lomys@tse-fr.eu.

listen to, and Flickr users see what cameras were used to take the pictures other users share. In all of these cases, how much and what information individuals gather about a new item (a movie, a restaurant, a song, or a camera) is informed by their connections' choices, and so are their resulting purchase and sharing decisions.[1]

I study the interplay between social information and individual incentives to choose and acquire private information as a mechanism for social learning, influence, and diffusion. If agents can choose what to learn about, how do others' choices and the structure of social ties affect agents' information choice? If agents can choose how much to learn at a cost, do they have the incentive to collect information? Do agents over-exploit the information revealed by others' choices or engage in excessive independent exploration anticipating that a wrong herd may have formed? How do others' choices and the structure of social ties affect this trade-off and the diffusion of new knowledge? When do societies settle on the best course of action or, in contrast, suboptimal behavior persists?

I develop a model of social learning in networks with information acquisition and choice to answer these questions. Countably many Bayes rational agents act in sequence. Each chooses between two actions. Actions' qualities are i.i.d. draws about which agents are initially uninformed. Agents wish to take the action with the highest quality. Each agent observes a subset of earlier agents, the agent's neighborhood. Neighborhoods are drawn from a joint distribution, the *network topology*. After observing his neighbors' choices, an agent engages in *costly sequential search*. Searching perfectly reveals the quality of the sampled action, but comes at a cost (i.i.d. across agents). After sampling an action, the agent decides whether to sample the second action or not. Finally, the agent takes an action from those he has sampled. Individual neighborhoods and sampling decisions remain unobserved.

The network topology shapes agents' ability to learn from others' choices (social information); the *search technology* shapes agents' ability to acquire private information. Social and private information interact: others' choices inform the optimal sampling sequence and timing to stop the search process. I provide two main sets of results. First, I identify sufficient and necessary conditions on network topologies and search technologies for positive learning outcomes; I do so by uncovering which learning principles are, or are not, at play. Second, I characterize how the speed and efficiency of social learning depend on the network structure.

I consider two metrics of social learning. *Complete learning* occurs if the probability that agents take the best action converges to one in the long run. *Maximal learning* occurs if, in the long run, agents take the best action with the same probability as an agent with the lowest possible search cost type and no social information (a "searcher"). Maximal and complete learning coincide if search costs are not bounded away from zero; they may or may not do so otherwise. Thus, maximal learning is a weaker requirement than complete learning.

In equilibrium, agents maximize the value of their search program, rather than the probability of taking the best action, which is what matters for long-run outcomes. With sequential search, the two problems are not equivalent. Nevertheless, I am able to connect an agent's optimal sampling sequence and timing to stop the search process to the probability

[1] For evidence that users learn from their contacts' check-ins, see, e.g., Qiu, Shi and Whinston (2018).

that some of the agents he is directly or indirectly linked to (the agent's subnetwork) has sampled both actions. Since sampling both actions allows agents to assess their relative quality, the latter probability is a lower bound for the agent's probability of taking the best action. This connection makes the analysis of long-run outcomes possible.

Sufficient and necessary conditions for positive learning outcomes are presented in three theorems. Theorem 1 shows that complete learning occurs if search costs are not bounded away from zero and, in the network topology, arbitrarily long information paths occur almost surely and are identifiable—the complete network being the simplest element of this large class. Roughly, complete learning occurs if free experimentation is possible, the network is sufficiently connected, and agents are reasonably informed about the network realization.

I identify sufficient conditions for complete learning by developing an *improvement principle* (IP). According to the IP, a heuristic—improvements upon imitation—is sufficient to take the best action in the long run. Upon observing his neighbors, each agent chooses one of them to rely on and determines his search policy regardless of what others have done. If search costs are not bounded away from zero, there is a strict improvement in the probability of sampling the best action at the first search that an agent has over his chosen neighbor. If agents have reasonably accurate information about the network, they pick the correct neighbor to rely on; if, in addition, information paths are long enough, improvements last until agents sample the best action at the first search. Thus, complete learning occurs.

Theorem 2 characterizes network topologies in which maximal learning occurs. In these networks, there are two sets of agents. In the first set, which forms the core of the network, there are infinitely many isolated agents (i.e. agents with no neighbors) and infinitely many agents that observe all and only their isolated predecessors; isolated agents form a vanishing share of the core. The second set is the rest of the network. Information paths are identifiable and agents are sufficiently connected to recent predecessors in the core. This class of stochastic networks is fairly large and diverse; the core may form the entire network but it may also consist of an arbitrarily small portion of it—such as in almost-complete networks, in which the probability that late-moving agents observe all their predecessors converges to one in the long run. Depending on the primitives of the model, maximal and complete learning may coincide even if arbitrarily low cost draws cannot happen. When they do, an important implication of the theorem is to identify a non-trivial class of networks in which search costs that are not bounded away from zero are not necessary for complete learning.

The intuition for the maximal learning result is the following. The choices of isolated agents are independent of each other and there are infinitely many such agents; thus, the share of observed choices is sufficient for late-moving non-isolated agents in the core to sample at the first search the action a searcher would take. Since the share of isolated agents vanishes, maximal learning occurs within the core. The positive learning result for agents in the core is based on a *large-sample principle* (LSP). According to the LSP, agents learn by observing large samples of individual choices and aggregating the information therein contained. Once it occurs within the core, maximal learning extends to the rest of the network. Late-moving agents have reasonably accurate information about the network and are likely to observe the choices of some recent non-isolated agents in the core. In the long run, beginning search from the action taken by one their most recent neighbors suffices for them to sample at the first search the action a searcher would take.

Theorem 3 characterizes necessary conditions on network topologies and/or search technologies for maximal learning. Independently of search costs, maximal learning fails if agents are (directly or indirectly) connected to only finitely many other agents. When search costs are bounded away from zero, maximal learning fails in networks in which agents have at most one neighbor, in the complete network, and in all networks in which agents observe the choices of possibly correlated random numbers of most immediate predecessors (OIP networks).[2] In this last class of networks, social learning fails discontinuously with respect to its benchmark metric: whereas complete learning occurs if search costs are not bounded away from zero, also maximal learning (the second-best learning outcome) fails as soon as zero is removed from the support of the search costs distribution.

The IP and the LSP have limitations in the setting I study. First, when search costs are bounded away from zero, improvements upon imitation are precluded to late-moving agents; thus, societies that rely on the IP perform worse than a searcher independently of the network structure. Second, the model's information structure leaves large-sample and martingale convergence arguments with little room to operate, as no social belief forming a martingale plays a role in the equilibrium characterization; thus, learning via the LSP remains limited to groups of agents with a very special structure, such as the core in Theorem 2.

The second part of the paper provides insights on the speed and efficiency of social learning as a function of the network structure. First, the speed of learning, the probability of wrong herds, and long-run welfare and efficiency (i.e. as future payoffs are discounted with factor $\delta \to 1$) are the same in all OIP networks. Hence, in OIP networks these equilibrium outcomes are independent of network transparency, the density of connections, and their correlation pattern. The intuition behind this striking result is simple. In OIP networks, agents' subnetworks consist of all their predecessors. As an agent's search policy depends on the probability that some of the agents in his subnetwork has sampled both actions, the probability that the agent takes the best action must be the same in all OIP networks.

Second, I consider a single decision maker who makes all choices, internalizes future gains of today's search, and samples each action only once along the same information path. Equilibrium welfare in OIP networks converges to that implemented by the single decision maker if and only if $\delta \to 1$ and search costs are not bounded away from zero. If $\delta < 1$ or search costs are bounded away from zero, welfare losses remain significant.

Third, reducing network transparency leads to inefficient duplication of costly search: agents who do not observe all prior choices fail to recognize actions that are revealed to be inferior by some of their predecessors' choices and so engage in overeager search. The resulting welfare loss remains sizable for all $\delta < 1$. Simply informing agents about the shares of prior choices restores in all OIP networks the same welfare as that in the complete network.

Finally, the density of indirect connections affects convergence rates. In particular,

[2] To fix ideas, let $1 \leq \ell_n < n$; agents $n - \ell_n, \ldots, n-1$ are the ℓ_n immediate predecessors of agent n. The complete network is the OIP network in which each agent n observes his $n-1$ immediate predecessors.

convergence to the best action is faster than polynomial in OIP networks but only faster than logarithmic under uniform random sampling of one agent from the past. Intuitively, learning is slower under uniform random sampling because in such networks the cardinality of agents' subnetworks grows at a slower rate than in OIP networks, and so does the probability that one agent in the subnetworks samples both actions.

Related Literature. The sequential social learning model (SSLM) originates with the papers of Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992), and Smith and Sørensen (2000). In the SSLM, agents observe a free private signal (informative about the relative quality of all alternatives) and the choices of all prior agents before making their choice. To model the observation structure, I build on Acemoglu, Dahleh, Lobel and Ozdaglar (2011) and Lobel and Sadler (2015), who allow for partial and stochastic observability of prior choices in the SSLM.[3] I discuss in Section 5 how my work relates to the SSLM.

Recent papers study costly acquisition of private information in social learning. In the complete network, my model reduces to that of Mueller-Frank and Pai (2016) (MFP); I discuss in Section 5 how my work relates to theirs. In Burguet and Vives (2000), Chamley (2004), and Ali (2018), agents choose how informative a signal to acquire at a cost which depends on the informativeness. In Hendricks, Sorensen and Wiseman (2012), agents decide whether to purchase a good after observing the aggregate purchase history and can acquire a perfect signal about their value for the good at a given cost. None of these papers focuses on the network structure or information choice. In Board and Meyer-ter-Vehn (2020), agents observe whether their neighbors have adopted an innovation and decide whether to gather information about its quality via costly inspection; they study how the network structure affects learning dynamics at each point in time. Because of substantial differences in the informational environment, results or techniques in these papers do not compare directly to mines.[4]

Recent work (Chamley (2004), Lobel, Acemoglu, Dahleh and Ozdaglar (2009), Monzón and Rapp (2014), Hann-Caruthers, Martynov and Tamuz (2018), Harel, Mossel, Strack and Tamuz (2020), Rosenberg and Vieille (2019), and Dasaratha and He (2020)) studies the speed and efficiency of social learning in the SSLM and related settings. This is a technically challenging topic even in simple networks and with exogenous private information.[5] My paper is the first to study how the speed and efficiency of social learning vary with the network structure in a setting with endogenous private information. Despite the complications introduced by costly information acquisition, a rich and tractable analysis emerges in my setting.

Weitzman (1979) characterizes the optimal sequential search strategy by an agent who faces a bandit problem, each arm representing a distinct alternative with a random prize. Each agent in my model faces the same problem and trade-off between exploration (sampling the second action) and exploitation (taking the best action according to his social

[3] 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.

[4] Garcia and Shelegia (2018) study strategic pricing with consumer search and observational learning.

[5] For instance, for the SSLM little is known about the speed of learning unless all agents observe the most recent choice, a random choice from the past, or all past choices (see Lobel et al. (2009), Rosenberg and Vieille (2019), and Hann-Caruthers et al. (2018)).

information).[6] However, whereas in Weitzman (1979) the prize distributions associated with different actions are exogenous, in my setting they depend on the agent's social information, which is endogenously generated by other agents' equilibrium play. More broadly, my work connects to the literature on the dynamics of information acquisition and choice: Wald (1947), Moscarini and Smith (2001), Fudenberg, Strack and Strzalecki (2018), Che and Mierendorff (2019), Mayskaya (2020), Ke and Villas-Boas (2019), Zhong (2019), Liang, Mu and Syrgkanis (2019, 2020), Liang and Mu (2020), and Bardhi (2020). Studying how social information and the network structure affect information acquisition and choice is new to my paper.

Salish (2017) and Sadler (2020) study learning in networks in which finitely many agents interact repeatedly, acquire private information by experimenting with a two-armed bandit, and observe their neighbors' experimentation. Sadler (2020) allows for complex networks, but agents follow a boundedly rational decision rule. In Salish (2017) agents are rational, but a sharp characterization only obtains for particular network structures. In contrast, I accommodate both Bayes rationality and general network topologies. Perego and Yuksel (2016) study learning with a continuum of Bayesian agents repeatedly choosing between learning from own experimentation or learning from others' experiences. The authors characterize how communication frictions and heterogeneity in connections affect the creation and diffusion of knowledge, but do not focus on network properties other than connectivity.

Kultti and Miettinen (2006, 2007), Celen and Hyndman (2012), Song (2016), and Nei (2019) consider costly observability of past histories in the SSLM. In these papers, private information is free, while which agents' choices to observe is endogenously determined. In contrast, I study costly acquisition of private information in exogenous network structures.

Road Map. Section 2 introduces the model, characterizes equilibrium strategies, and defines the metrics of social learning. Section 3 provides positive and negative learning results with respect to these metrics. Section 4 studies the speed and efficiency of social learning. Section 5 discusses the relation with the SSLM and MFP and presents extensions and further results. Section 6 concludes. Proofs and omitted details are in the Appendices.

2 Collective Search Environment

Agents and Actions. A countably infinite set of agents, indexed by $n \in \mathbb{N} := \{1, 2, \dots\}$, sequentially take a single action each from the set $X := \{0, 1\}$. Agent n acts at time n. Let x denote a typical element of X, $\neg x$ the action in X other than x, and a_n the action agent n takes. Calendar time is common knowledge and the order of moves exogenous.

State Process. Let q_x denote the quality of action x. Qualities q_0 and q_1 are i.i.d. draws from a probability measure $\mathbb{P}_Q$ over $Q \subseteq \mathbb{R}_+ := \{s \in \mathbb{R} : s \geq 0\}$. The state of the world $\omega := (q_0, q_1)$ is drawn once and for all at time 0. The state space is $\Omega := Q \times Q$, with product measure $\mathbb{P}_\Omega := \mathbb{P}_Q \times \mathbb{P}_Q$. This formulation captures finite and infinite state spaces. The probability space $(\Omega, \mathcal{F}_\Omega, \mathbb{P}_\Omega)$ is the *state process*, which is common knowledge.

[6] The trade-off between exploration and exploitation is the distinctive feature of bandit problems. Bergemann and Välimäki (2008) and Hörner and Skrzypacz (2017) survey bandit problems in economics.

Agents wish to take the action with the highest quality. To do so, they have access to *social information*, which is derived from observing a subset of past agents' choices, and to *private information*, which is endogenously acquired by costly sequential search.

Network Topology. Agents observe the choices of a subset of past agents, as in Acemoglu et al. (2011) and Lobel and Sadler (2015). The set of agents whose choices agent n observes, denoted by $B(n)$, is called n's neighborhood. Neighborhoods $B(n) \in 2^{\mathbb{N}_n}$, where $2^{\mathbb{N}_n}$ is the power set of $\mathbb{N}_n := \{m \in \mathbb{N} : m < n\}$, are random variables generated via a probability measure $\mathbb{Q}$ on the product space $\mathbb{B} := \prod_{n \in \mathbb{N}} 2^{\mathbb{N}_n}$. The probability space $(\mathbb{B}, \mathcal{F}_{\mathbb{B}}, \mathbb{Q})$ is the *network topology*, which is common knowledge. Realizations of $B(n)$ are denoted by B_n and are agent n's private information. If $n' \in B_n$, then n not only observes $a_{n'}$, but also knows the identity of this agent. Agent n is isolated if $B_n = \emptyset$. Neighborhoods are independent of the state process and of private search costs (to be soon introduced).

The formulation allows for arbitrary correlations among agents' neighborhoods, as well as for independent neighborhoods and deterministic network topologies. The framework nests most of the networks commonly observed in the data and studied in the literature, such as 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.

Search Technology. Information about the quality of the two actions is acquired via costly sequential search with recall. After observing $B(n)$ and the choices of the agents in $B(n)$, agent n decides which action $s_n^1 \in X$ to sample first. Sampling an action perfectly reveals its quality $q_{s_n^1}$ to the agent. After observing $q_{s_n^1}$, agent n decides whether to sample the remaining action, $s_n^2 = \neg s_n^1$, or to discontinue searching, $s_n^2 = d$. Let S_n denote the set of actions agent n samples. After sampling has stopped, the agent takes an action $a_n \in S_n$. For a single agent, the search problem is a version of Weitzman (1979). When each agent observes all past choices, the model reduces to that of Mueller-Frank and Pai (2016).

The first action is sampled at no cost, while sampling the second action involves a cost $c_n \in C \subseteq \mathbb{R}_+$. Search costs c_n are i.i.d. across agents, are drawn from a probability measure $\mathbb{P}_C$ over C, with CDF F_C, and are independent of the network topology and the state process. The probability space $(C, \mathcal{F}_C, \mathbb{P}_C)$ is the *search technology*, which is common knowledge. An agent's search cost and sampling decisions are his private information.

Search costs are not bounded away from 0 if there is a positive probability of arbitrarily low search costs. The next definition formalizes this idea.

Definition 1. *Let* $\underline{c} := \min \operatorname{supp}(\mathbb{P}_C)$. *Search costs are* bounded away from 0 *if* $\underline{c} > 0$*; search costs are* not bounded away from 0 *if* $\underline{c} = 0$.

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

Information and Strategies. Each agent n has three information sets: $I^1(n) := \{c_n, B(n), a_k \; \forall k \in B(n)\}$ corresponds to n's information prior to sampling any action; $I^2(n) := I^1(n) \cup \{q_{s_n^1}\}$ corresponds to n's information after sampling the first action; finally, $I^a(n) := \{c_n, B(n), a_k \; \forall k \in$

$B(n), \{q_x : x \in S_n\}\}$ corresponds to n's information once his search ends. These sets are random variables whose realizations I denote by I_n^1, I_n^2, and I_n^a. The classes of all possible information sets of agent n are denoted by $\mathcal{I}_n^1$, $\mathcal{I}_n^2$, and $\mathcal{I}_n^a$.

A strategy for agent n is a triple of mappings $\sigma_n := (\sigma_n^1, \sigma_n^2, \sigma_n^a)$, where $\sigma_n^1 \colon \mathcal{I}_n^1 \to \Delta(\{0,1\})$, $\sigma_n^2 \colon \mathcal{I}_n^2 \to \Delta(\{\neg s_n^1, d\})$, and $\sigma_n^a \colon \mathcal{I}_n^a \to \Delta(S_n)$. Given a strategy profile $\sigma := (\sigma_n)_{n\in\mathbb{N}}$, the sequence of decisions $\Big((s_n^1, s_n^2, a_n)\Big)_{n\in\mathbb{N}}$ 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.

Equilibrium Notion. The solution concept is the set of *perfect Bayesian equilibria* of the game of social learning, hereafter referred to as equilibria. A strategy profile $\sigma := (\sigma_n)_{n\in\mathbb{N}}$ is an equilibrium 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 $\sigma_{-n} := (\sigma_1, \dots, \sigma_{n-1}, \sigma_{n+1}, \dots)$.

Agent n' decision problems are discrete choice problems. Thus, they have a well-defined solution that only requires randomization in case of indifference at some stage. Given criteria to break ties, an inductive argument shows that the set of equilibria is nonempty.

I focus on the equilibrium in which agents sample the second action in case of indifference at the second search stage and break other ties uniformly at random. Selecting this equilibrium simplifies the exposition without affecting the results. As no confusion arises, I identify agent n's (equilibrium) strategy $(\sigma_n^1, \sigma_n^2, \sigma_n^a)$ with his decisions (s_n^1, s_n^2, a_n).

2.1 Equilibrium Strategies

I first recall the notion of an agent's subnetwork from Lobel and Sadler (2015) and introduce that of an agent's subnetwork relative to action x. The equilibrium characterization follows.

Definition 2. *Agent m belongs to agent n's* subnetwork*, denoted by $\widehat{B}(n)$, 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. Agent m belongs to agent n's* subnetwork relative to action $x \in X$*, denoted by $\widehat{B}(n,x)$, if $m \in \widehat{B}(n)$ and $a_m = x$.*

In words, $\widehat{B}(n)$ consists of the agents that are, either directly or indirectly (through neighbors, neighbors of neighbors, and so on) observed by/connected to agent n; $\widehat{B}(n,x)$ consists of the agents that are, either directly or indirectly, observed by agent n to take action x. Realizations of the random variables $\widehat{B}(n)$ and $\widehat{B}(n,x)$ are denoted by $\widehat{B}_n$ and $\widehat{B}_{n,x}$.

Choice Stage. If an agent sampled one action, he takes that action; if he sampled both, he takes the best action, randomizing uniformly if the two actions have the same quality.

First Search Stage. Fix n and σ_{-n}. For each action x there are two possibilities:

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 $\mathbb{P}_{\Omega|q_x \geq q_{\neg x}}$. This is so because agents sampling both actions take the one 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 agent n's optimal policy at the first search stage, consider the events

$$E_n^x \coloneqq \left\{ s_k^2 = d \;\forall k \in \widehat{B}(n,x) \right\} \qquad \text{for } x = 0,1. \tag{1}$$

In words, E_n^x occurs when none of the agents in n's subnetwork relative to action x samples both actions. Given σ_{-n} and I_n^1, agent n computes the conditional probabilities

$$P_n(x) \coloneqq \mathbb{P}_{\sigma_{-n}}\left(E_n^x \mid I_n^1\right) \qquad \text{for } x = 0,1. \tag{2}$$

If $P_n(x) < P_n(\neg x)$, agent n's belief about the quality of action x strictly first-order stochastically dominates his belief about the quality of action $\neg x$ (the formal argument is in Appendix A). Thus, according to Weitzman (1979)'s optimal search rule, n samples first action x: $s_n^1 = x$. If $P_n(0) = P_n(1)$, agent n's beliefs about the qualities of the two actions are identical and so n samples the first action uniformly at random.

Second Search Stage. Let I_n^2 be agent n's information set after sampling a first action of quality $q_{s_n^1}$. Agent n will only sample the second action if his search cost c_n is no larger than the expected gain from the second search. If $B_n = \emptyset$, such gain is

$$t^{\emptyset}\left(q_{s_n^1}\right) \coloneqq \mathbb{E}_{\mathbb{P}_Q}\left[\max\left\{q - q_{s_n^1}, 0\right\}\right]. \tag{3}$$

If $B_n \neq \emptyset$, agent n benefits from the second search only if action $\neg s_n^1$ was not sampled by any of the agents in $\widehat{B}(n, s_n^1)$. Thus, he must compute the conditional probability

$$P_n\left(q_{s_n^1}\right) \coloneqq \mathbb{P}_{\sigma_{-n}}\left(E_n^{s_n^1} \mid I_n^2\right). \tag{4}$$

With remaining probability, at least one of those agents sampled action $\neg s_n^1$, but nevertheless took action s_n^1, in which case s_n^1 is (weakly) superior by revealed preferences. Hence, n's expected gain from the second search is

$$t_n\left(q_{s_n^1}\right) \coloneqq P_n\left(q_{s_n^1}\right) t^{\emptyset}\left(q_{s_n^1}\right). \tag{5}$$

Remark 1. The events E_n^x in (1) act as sufficient statistics for the information agent n's subnetwork contains. As a result, the conditional probabilities $P_n(x)$ and $P_n(q_x)$ in (2) and (4) suffice to describe n's equilibrium search policy. In turn, these probabilities link agents' information acquisition to the probability that they take the best action, which is what matters for long-run outcomes (see Section 2.2). The intuition is the following:

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

Here, the first inequality holds as agent n takes the best action between 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 equality holds as the two events at issue are one the complement of another (see definition of $E_n^{s_n^1}$ in (1)).

2.2 Metrics of Social Learning

Complete Learning. Complete learning requires agents to eventually take the best action with probability 1. This outcome would obtain if each agent observed the search decisions of all prior agents and (at least) one of these agents sampled both actions.

Definition 3. Complete learning *occurs if*

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

Maximal Learning. An isolated agent with search cost type $\underline{c}$ has the best search opportunities and the strongest incentives to explore, so call him a "*searcher*".[7] Let $q(\underline{c}) \coloneqq \inf\{q \in \operatorname{supp}(\mathbb{P}_Q) : t^{\emptyset}(q) < \underline{c}\}$ be the quality such that a searcher samples both actions whenever the action he samples first has quality lower than $q(\underline{c})$ and discontinues search otherwise. Thus, a searcher takes the best action with probability 1 if and only if $\omega \notin \Omega(\underline{c}) \coloneqq \{\omega \in \Omega : q_i \geq q(\underline{c}) \text{ for } i = 0, 1 \text{ and } q_0 \neq q_1\}$. Maximal learning requires agents to eventually take the best action with probability 1 whenever a searcher does so.

Definition 4. Maximal learning *occurs if*

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

Complete *vs* Maximal Learning. Maximal and complete learning coincide if and only if $\mathbb{P}_\Omega(\Omega(\underline{c})) = 0$. This is always the case if $\underline{c} = 0$. In contrast, when $\underline{c} > 0$ maximal learning is a weakly weaker requirement than complete learning. To see this, suppose first that q_0, q_1 are uniform draws from $\{0, 1\}$ and $\underline{c} = 1/3$. Here, $q(\underline{c}) = 1$, $\Omega(\underline{c}) = \emptyset$, and $\mathbb{P}_\Omega(\Omega(\underline{c})) = 0$; thus, maximal and complete learning coincide. Next, suppose that q_0, q_1 are uniform draws from $\{0, 1, 2\}$ and that $\underline{c} = 10/29$. Here, $q(\underline{c}) = 1$, $\Omega(\underline{c}) = \{(1, 2), (2, 1)\}$, and $\mathbb{P}_\Omega(\Omega(\underline{c})) = 2/9$; in this case, maximal learning is a weaker requirement than complete learning.

The next assumption rules out uninteresting learning problems.

Assumption 1. *There exist* $\tilde{q}, \tilde{q}' \in \operatorname{supp}(\mathbb{P}_Q)$ *such that:*

1. *(a)* $\mathbb{P}_Q(q(\underline{c}) > q > \tilde{q}) > 0$*; (b)* $1 - F_C(t^{\emptyset}(\tilde{q})) > 0$*, i.e. with positive probability, an isolated agent discontinues search after sampling a first action of quality* $\tilde{q}$ *or higher.*
2. *(a)* $\mathbb{P}_Q(q \leq \tilde{q}') > 0$*; (b)* $F_C(t^{\emptyset}(\tilde{q})) > 0$ *i.e. with positive probability, an isolated agent samples the second action after sampling a first action of quality* $\tilde{q}'$ *or lower.*

If part 1 fails, in equilibrium isolated agents take the best action when $\omega \notin \Omega(\underline{c})$ and never search twice otherwise; in turn, non-isolated agents just follow the behavior of any of their neighbors. Thus, learning trivially obtains for $\omega \notin \Omega(\underline{c})$ and never occurs otherwise. If part 2 fails, no agent ever searches twice and so learning never occurs.

[7] By (3)–(5), $t^{\emptyset}(q_{s_n^1}) \geq t_n(q_{s_n^1})$; that is, given $q_{s_n^1}$, the expected gain from the second search, and so the incentive to explore, is larger for isolated agents than for non-isolated agents (see also Section 5.2).

3 Long-Run Learning

The search technology shapes agents' possibility to acquire private information; the network topology shapes agents' possibility to learn from social information. In this section, I provide sufficient and necessary conditions on these primitives for positive learning results.

3.1 Complete Learning and the Improvement Principle

I begin with 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 the size of $\widehat{B}(n)$ grows without bound as n becomes large.

Definition 5. *A network topology features* expanding subnetworks *if, for all $K \in \mathbb{N}$,*

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

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

The next notions are those of neighbor choice function and chosen neighbor topology. A neighbor choice function represents a particular agent's means of selecting a neighbor; a chosen neighbor topology is a network in which agents discard all unselected neighbors.

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 neighborhoods $B_n \in 2^{\mathbb{N}_n}$, we have $\gamma_n(B_n) \in B_n$ when $B_n \neq \emptyset$, and $\gamma_n(B_n) = 0$ otherwise. Agent $\gamma_n(B_n)$ is called 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 $(\gamma_n)_{n\in\mathbb{N}}$.*

Complete learning occurs if search costs are not bounded away from 0, the network is sufficiently connected, and agents have reasonably accurate information about the network realization. The next theorem formalizes this idea.

Theorem 1. *Complete learning occurs if all the following conditions hold:*

(i) The search technology has search costs that are not bounded away from 0;

(ii) The network topology has a sequence of neighbor choice functions $(\gamma_n)_{n\in\mathbb{N}}$ such that:

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

(b) For all $\varepsilon, \eta > 0$, there exists $N_{\varepsilon\eta} \in \mathbb{N}$ such that, for all agents $n > N_{\varepsilon\eta}$, with probability at least $1 - \eta$,

$$\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n))\right) \geq \mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x\in X} q_x\right) - \varepsilon. \qquad (6)$$

To prove Theorem 1, I develop an *improvement principle* (IP). The IP benchmarks the performance of Bayesian agents against *improvements upon imitation*—a heuristic that is simpler to analyze and can be improved upon by Bayes rationality. It works as follows. Upon observing who his neighbors are, each agent selects one neighbor to rely on. After observing the choice of his chosen neighbor, the agent determines his optimal search policy regardless of what other neighbors have done. The IP holds if: $(*)$ there is an increase in the probability that an agent samples the best action at the first search over his chosen neighbor's probability; $(**)$ the learning mechanism captured by such heuristic suffices for complete learning.

For $(*)$ to hold, search costs must not be bounded away from 0 (condition (i) in Theorem 1). The intuition is the following. Consider an agent, say n, and his chosen neighbor, say $b < n$. Unless b samples the best action with probability 1 at the first search, b's expected gain from the second search is positive. Therefore, if search costs are not bounded away from 0, b samples both actions with positive probability. As b takes the best action between those he samples, with positive probability the action b takes is of better quality than the one he samples first. If n begins searching from the action taken by b, this results in a strict improvement in the probability of sampling the best action at the first search that n has over b, unless b already does so with probability 1.

In turn, $(**)$ requires the network to be sufficiently connected (condition (ii)–(a) in Theorem 1). That is, long information paths must occur almost surely so as for improvements to last until agents sample the best action with probability 1 at the first search.[8] In addition, $(**)$ requires agents to have reasonably accurate information about the network (condition (ii)–(b) in Theorem 1). That is, information paths must be identifiable so as for agents to single out the correct neighbor to rely on. To understand this requirement, note that agent n can rely on agent γ_n only if $\gamma_n \in B(n)$. If neighborhoods are correlated, γ_n's probability of sampling first the best action conditional on n observing γ_n is not the same as γ_n's unconditional probability of sampling first the best action. That is, n earns γ_n's probability of sampling first the best action *conditional* on n choosing to rely on γ_n. Thus, the difference between $\mathbb{P}_\sigma(s^1_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n)$ and $\mathbb{P}_\sigma(s^1_{\gamma_n} \in \arg\max_{x \in X} q_x)$ must be small for large n to ensure that agent n is able to identify the correct neighbor to rely on.

Remark 2. For the IP to work, one needs to show that Bayesian agents who do not ignore all but one of the agents in their neighborhood can obtain the improvements described by the heuristic. In some networks—such as in the complete network—this is an easy task: the probability of sampling the best action at the first search as well as the probability of taking the best action under Bayes rationality are the same as under the IP. In contrast, over general networks, whereas one can show that a Bayesian agent has always a higher probability of *sampling* the best action at the first search than an agent following the heuristic captured by the IP (see Lemma 11 in Appendix C), the same conclusion need not hold true for the probability of *taking* the best action. For this reason, I consider improvements with respect to $\mathbb{P}_\sigma(s^1_n \in \arg\max_{x \in X} q_x)$, and not with respect to $\mathbb{P}_\sigma(a_n \in \arg\max_{x \in X} q_x)$.[9] 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 taking the best

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

[9] The ultimate interest is in the evolution dynamics of $\mathbb{P}_\sigma(a_n \in \arg\max_{x \in X} q_x)$. However, convergence to 1 of $\mathbb{P}_\sigma(s^1_n \in \arg\max_{x \in X} q_x)$ as $n \to \infty$ is sufficient for complete learning.

action, which is what matters for long-run outcomes (see Remark 1). In other words, the IP I develop works as proof technique but does not represent an approximation, even at the heuristic level, of the *true* learning mechanism that Bayesian agents adopt.

Multiple Equilibria. When the social learning game has multiple equilibria, for a given sequence of neighbor choice functions, (6) may be satisfied in one equilibrium but not in a different one. Moreover, there may be a sequence of neighbor choice functions satisfying (6) in one equilibrium, but no sequence of neighbor choice functions satisfying (6) in a different equilibrium. However, a wide variety of intuitive conditions on the network topology ensure the existence of a sequence of neighbor choice functions such that (6) holds in all equilibria of the game. I refer to Appendix B for an overview of such conditions. Note though that Theorem 1 holds more generally than under such conditions.

Limits to the Improvement Principle. When search costs are bounded away from 0, one may conjecture by analogy with Theorem 1 that maximal learning occurs if the network topology is sufficiently connected and has identifiable information paths. However, this conjecture is incorrect. When $\underline{c} > 0$, improvements upon imitation are precluded to late-moving agents and so maximal learning via the IP fails in all network topologies. In other words, the IP is fragile to perturbations in the search technology: it breaks down as soon as 0 is removed from the support of the search costs distribution.

To see why search costs that are bounded away from 0 disrupt the IP, suppose $\underline{c} > 0$, $\omega \notin \Omega(\underline{c})$ and, by contradiction, that the IP holds. Then, there is a chosen neighbor topology in which the probability of none of the agents in $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to 0 as $n \to \infty$. Thus, there must be some sufficiently late-moving agent m for which this probability is so small that his expected gain from the second search falls below $\underline{c} > 0$ and remains below this threshold afterward. As a result, no agent in the chosen neighbor topology moving after agent m will sample the second action. By Assumption 1, the probability that none of the agents in $\widehat{B}(m) \cup \{m\}$ samples both actions in equilibrium is positive for any finite m. This is a contradiction, as then the probability of none of the agents in $\widehat{B}(n) \cup \{n\}$ sampling both actions remains bounded away from 0 in the chosen neighbor topology.[10]

3.2 Maximal Learning and the Large-Sample Principle

The next theorem characterizes sufficient conditions on network topologies for maximal learning. Maximal and complete learning may coincide even if $\underline{c} > 0$. When they do, an immediate implication of the theorem is to identify network topologies in which search costs that are not bounded away from 0 are not necessary for complete learning.

To aid the exposition, I define the following: for all $n \in \mathbb{N}$, $B_n^{\emptyset} := \{k \in \mathbb{N}_n : B(k) = \emptyset\}$ is the set of agent n's isolated predecessors; $S := \{n \in \mathbb{N} : B(n) \in \{\emptyset, B_n^{\emptyset}\}\}$ is the set of agents whose realized neighborhood either is empty or consists of all their isolated predecessors.

[10]One can also construct examples in which asymmetric information about the network disrupts the IP. In such cases, learning via improvements upon imitation fails even if the network is well connected and $\underline{c} = 0$.

Theorem 2. *Maximal learning occurs if the network topology satisfies all the following conditions:*

(i) $\sum_{n\in\mathbb{N}} \mathbb{Q}(B(n)=\emptyset)=\infty$ *and* $\lim_{n\to\infty}\mathbb{Q}(B(n)=\emptyset \mid n\in S)=0$*;*

(ii) $\mathbb{Q}(B(n)=B_n^{\emptyset} \mid 1\in B(n))=1$*;*

(iii) For all $K\in\mathbb{N}$*,*

$$\lim_{n\to\infty}\mathbb{Q}\left(\max_{b\in B(n)\cap S} b<K\right)=0;$$

(iv) For all $\varepsilon,\eta>0$*, there exists* $N_{\varepsilon\eta}\in\mathbb{N}$ *such that, for all agents* $n>N_{\varepsilon\eta}$*, with probability at least* $1-\eta$*,*

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

for all sequences of neighbor choice functions $(\gamma_n)_{n\in\mathbb{N}}$*.*

By Theorem 2, maximal learning occurs in network topologies in which there are two sets of agents. The first set is S and forms the core of the network. Within S there are two groups of agents: an infinite but vanishing group of isolated agents (condition (i) in Theorem 2); agents that observe all and only their isolated predecessors and know that they are observing only isolated agents (condition (ii) in Theorem 2). The first part of the proof shows that maximal learning occurs within S. The intuition is the following. When $\omega\notin\Omega(\underline{c})$, either both actions have the same quality (and so there is nothing to learn), or each isolated agent takes the best action with probability larger than 1/2; moreover, the choices of isolated agents are independent of each other and there are infinitely many such agents. Thus, the share of earlier choices is sufficient for non-isolated agents in S to sample the best action at the first search with probability 1 as $n\to\infty$. As the share non-isolated agents in S converges to 1 as $n\to\infty$, maximal learning occurs within S. In other words, isolated agents form a subsequence of sacrificial lambs who provide just enough information for the other agents in S to sample the best action at the first search. The positive learning result within set S is based on a *large-sample principle* (LSP). According to the LSP, agents learn by observing large samples of individual choices and aggregating the information therein contained.

The second set is $\mathbb{N}\setminus S$ and forms the rest of the network. Maximal learning extends to this second set because of the following argument. Non-isolated agents in S collect and aggregate the information brought in by isolated agents. Condition (iii) in Theorem 2 ensures that late-moving agents are likely to observe recent choices from non-isolated agents in S. Condition (iv) in Theorem 2 ensures that agents have reasonably accurate information about the network realization—similarly to condition (ii)–(b) in Theorem 1, but now (6) has to hold for all sequences of neighbor choice functions. Each late-moving agent in $\mathbb{N}\setminus S$ can now select one of his most recent neighbors to rely on and sample this neighbor's action at the first search. This strategy, together with the properties of the network topology, suffices for agents in $\mathbb{N}\setminus S$ to sample the best action at the first search with probability 1 as $n\to\infty$ whenever $\omega\notin\Omega(\underline{c})$. Thus, maximal learning occurs. Intuitively, once maximal learning occurs via the LSP within set S, the remaining agents can simply learn via imitation (even though improvements upon imitation are precluded because search costs are bounded away from 0).

The class of network topologies in which maximal learning occurs is large and diverse. In particular, the set of agents S may form the entire network but it may also consist of an arbitrarily small portion of the network, as the following example shows.

Example 1. The following network topologies satisfy the conditions of Theorem 2.

1. For all n, let $\mathbb{Q}(B(n) = \emptyset) = 1/n$ and $\mathbb{Q}(B(n) = B_n^{\emptyset}) = 1 - 1/n$. In this case, $S = \mathbb{N}$; that is, all agents in the network are in S.
2. For all $n > 2$, let $\mathbb{Q}(B(n) = \emptyset) = 1/n$, $\mathbb{Q}(B(n) = B_n^{\emptyset}) = 1/\sqrt{n}$, and $\mathbb{Q}(B(n) = \{1, \ldots, n-1\}) = 1 - (1 + \sqrt{n})/n$. In this case, the set S consists of an arbitrarily small portion of the network. This is an almost-complete network, in the sense that the probability that agent n observes all his predecessors converges to 1 as $n \to \infty$.

Limits to the Large-Sample Principle. The positive result in Theorem 2 crucially hinges on the special structure of the set of agents S. In particular, the result relies on the assumption that non-isolated agents in S: (i) observe *only* isolated agents; (ii) *know* that the agents they observe are isolated. Under this premise, the optimal policy at the first search stage for non-isolated agents in S is determined by the relative shares of choices they observe. When non-isolated agents in S observe more than only isolated agents or are unsure whether the agents they observe are isolated, connecting their optimal search policy to the ratio of observed choices is no longer possible. Thus, the positive results in Theorem 2 are hardly extendable to a more general characterization where we allow non-isolated agents in S to only know that *some* agents they observe are isolated or that the agents they observe have *some positive probability* of being isolated. The major impediment arises because no social belief that forms a martingale also plays a role in the characterization of equilibrium strategies. Formally, $\left(\mathbb{P}_\sigma(E_n^x)\right)_{n\in\mathbb{N}}$, $x = 0, 1$, do not form a martingale even when conditioning on public histories $a^{n-1} := (a_1, \ldots, a_{n-1})$. Therefore, large-sample and martingale convergence arguments play a limited role in the setting I study. However, despite learning via the LSP remains restricted to special cases, such as the set of agents S, this is enough for Theorem 2 to show that complete learning occurs in a non-trivial class of network topologies.

3.3 Failure of Maximal Learning

In this section, I focus on negative learning results. To begin, I define a class of networks which will be extensively discussed in the rest of the paper. For all $n \in \mathbb{N}$ and $\ell_n \in \mathbb{N}_n$, let $B_n^{\ell_n} := \{k \in \mathbb{N}_n : k \geq n - \ell_n\}$ be the set consisting of the ℓ_n most immediate predecessors of agent n. Hereafter, the acronym OIP stands for "observation of immediate predecessor".

Definition 7. *A network topology is an* OIP network *if, for all agents n,*

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

OIP networks form a large class of network structures, ranging from deterministic networks to stochastic networks with independent or correlated neighborhoods. For example:

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, each agent observes only his immediate predecessor.
3. Let neighborhoods be independent and, for all n, $\mathbb{Q}(B(n) = B_n^1) = (n-1)/n$ and $\mathbb{Q}(B(n) = B_n^{n-1}) = 1/n$. Here, each agent either observes his immediate predecessor, or all of them, with the latter event becoming less and less likely as n grows large.
4. Let $\mathbb{Q}(B(2) = B_2^1) = 1$, $\mathbb{Q}(B(3) = B_3^1) = \mathbb{Q}(B(3) = B_3^2) = 1/2$, and, for all $n > 3$, $B(n) = B_n^1$ if $B(3) = B_3^1$ and $B(n) = B_n^{n-1}$ if $B(3) = B_3^2$. Here, neighborhoods are correlated and each agent either observes his immediate predecessor, or all of them, depending on agent 3's neighborhood realization.

The next theorem characterizes necessary conditions on network topologies and/or search technologies for maximal (hence, complete) learning.

Theorem 3. *Maximal learning fails if:*

(i) The network topology has non-expanding subnetworks.

(ii) Search costs are bounded away from 0 *and the network topology:*

- *(a) Is an OIP network, or*
- *(b) Satisfies* $\mathbb{Q}(|B(n)| \leq 1) = 1$ *for all agents* n.

Theorem 3–(i) says that maximal learning fails with non-expanding subnetworks, independently of whether search costs are bounded away from 0 or not. The intuition is the following. Suppose $\underline{c} > 0$, $\omega \notin \Omega(\underline{c})$, and that the network topology has non-expanding subnetworks. By Assumption 1 and the characterization of equilibrium search policies, the probability that none of any finite set of agents samples both actions is positive. Since non-expanding subnetworks generate with positive probability an infinite subsequence of agents with finite subnetwork, the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions remains bounded away from 0. As a result, maximal learning fails.

Theorem 3–(ii) says that when search costs are bounded away from 0, maximal learning fails in OIP networks and in networks in which agents have at most one neighbor. In such networks, search costs that are not bounded away from 0 are necessary and sufficient for complete learning (provided the network has expanding subnetworks). The intuition is the following. In networks satisfying the conditions of Theorem 3–(ii), the probability of taking the best action is no larger under Bayes rationality than under imitation as captured by the IP. Therefore, as the IP fails when search costs are bounded away from 0, so does Bayesian learning. The same argument suffices to extend the negative result beyond the network topologies in Theorem 3–(ii). 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 prior choices.

The previous analysis highlights two interesting discontinuity results. First, there is discontinuity in learning outcomes when we compare the complete network with almost-complete networks. By Theorem 3–(ii)–(a), search costs that are not bounded away from 0 are necessary (and sufficient) for complete learning in the complete network. In contrast, part 2 of Example 1 exhibits an almost-complete network in which, whenever maximal and complete learning coincide with $\underline{c} > 0$, search costs that are not bounded away from 0 are not necessary for complete learning. Second, positive learning outcomes are fragile to perturbations in the search technology. In particular, in the networks characterized by Theorem 3–(ii), if 0 is removed from the support of the search costs distribution, not only complete learning fails, but also the second-best learning outcome (maximal learning) does so. That is, social learning fails discontinuously with respect to its benchmark metric.

4 Rate of Convergence, Welfare, and Efficiency

I now study how the speed and the efficiency of social learning and the probability of wrong herds depend on the network structure. Several insights emerge in OIP networks and so I begin by sketching equilibrium strategies in such networks (the formalities are in Appendix E).

4.1 Equilibrium Strategies in OIP Networks

First, I introduce the relevant terminology.

Definition 8. *In OIP networks, we say:*

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

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

(c) Action x is inferior by time n *if there exists an agent $j < n$, who has sampled both actions and such that $a_j = \neg x$.*

Thus, if an action is revealed to be inferior to agent n, then it is also revealed to be inferior by time n

Equilibrium Strategies in OIP Networks. Fix $n \geq 2$. At the first search stage, agent n samples the action taken by his immediate predecessor: $s_n^1 = a_{n-1}$. Hence, if an action is revealed to be inferior by time n, it is also inferior by time n.

At the second search stage, the optimal policy depends on whether action $\neg s_n^1$ is revealed to be inferior to agent n or not.

- If $\neg s_n^1$ is revealed to be inferior to agent n, then n discontinues search and takes action s_n^1. The reason. 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 each agent samples first the action taken by his immediate predecessor, agent $j+1$ must have sampled action $\neg s_n^1$ first, and therefore would only take $a_{j+1} = s_n^1$ 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.
- If $\neg s_n^1$ is not revealed to be inferior to agent n, the expected gain from the second search given $q_{s_n^1}$ 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. The intuition goes as follows. In all OIP networks, agent n's subnetwork is $\{1, \ldots, n-1\}$, which coincides with agent

n's neighborhood in the complete network. Moreover, each agent samples first the action taken by his immediate predecessor. Thus, given $q_{s_n^1}$, the probability that none of the agents in $\widehat{B}(n, s_n^1)$ has sampled both actions must be the same. But then, if $\neg s_n^1$ is not revealed to be inferior to agent n, agent n adopts the same threshold he would use in the complete network to determine whether to search further after sampling an action of the same quality that is not revealed to be inferior by time n.

The previous characterization has important implications: the order of search, the cutoff for sampling a second action that is not revealed to be inferior to an agent, and the probability that each agent takes the best action are the same in all OIP networks. Thus, network transparency, the density of connections, and their correlation pattern do not affect several equilibrium outcomes. The next proposition follows.

Proposition 1. *Fix a state process and a search technology. Then, in all OIP networks:*

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

(ii) If search costs are not bounded away from 0*, so that complete learning occurs, the rate of convergence to the best action is the same as in the complete network.*[11]

4.2 Rate of Convergence

The following property will be useful to establish the results on convergence rates.

Definition 9. *Let* $\underline{q} := \min \operatorname{supp}(\mathbb{P}_Q)$*. The search costs distribution has* polynomial shape *if there exist some real constants* K *and* L*, with* $K \geq 0$ *and* $0 < L < \frac{2^{K+1}}{(K+2)t^{\emptyset}(\underline{q})^K}$*, such that*

$$F_C(c) \geq Lc^K \qquad \forall\ c \in (0, t^{\emptyset}(\underline{q})/2).$$

The density of indirect connections affects the speed of learning: whereas the rate of convergence to the best action is faster than polynomial in OIP networks, it is only faster than logarithmic under uniform random sampling of one past agent. Learning is faster in OIP networks because the cardinality of agents' subnetworks grows at a faster rate, and so does the probability that at least one agent in the subnetworks samples both actions.

Proposition 2. *Suppose that search costs are not bounded away from* 0 *and that the search costs distribution has polynomial shape.*

(a) In OIP networks,

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

(b) If neighborhoods are independent and $\mathbb{Q}(B(n) = \{b\}) = 1/(n-1)$ *for all* $b \in \mathbb{N}_n$*,*

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

[11] In OIP networks and under uniform random sampling of one agent from the past, we can bound the probability of wrong herds as a linear function of the lowest cost in the support of the search costs distribution, as I show in an earlier version of this paper. Details are available upon request.

4.3 Equilibrium Welfare and Efficiency in OIP Networks

In this section, I first characterize how network transparency affects equilibrium welfare. Then, I compare equilibrium welfare against the efficiency benchmark in which a single decision maker makes all choices. Finally, I discuss a simple policy intervention that increases welfare and efficiency. To aid analysis, I assume that $\mathbb{P}_C$ admits density f_C with $f_C(\underline{c}) > 0$.

Equilibrium Welfare across OIP Networks. Equilibrium welfare is not the same across OIP networks. To see this, suppose $a_j = x$ and $a_{j+1} = \neg x$ for some agents $j, j+1$. Hence, 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 $n \geq j+2$; therefore, n fails to realize from agent $j+1$'s choice that $q_x \leq q_{\neg x}$. Thus, agent n inefficiently samples action x with positive probability at the second search stage.[12]

Inefficient duplication of costly search is more severe the shorter in the past agents can observe. Hence, the complete network is the most efficient OIP network and the network in which agents only observe their most immediate predecessor is the least efficient in this class. In all other OIP networks, equilibrium welfare is comprised between these two bounds.

Fix a state process and a search technology and suppose future payoffs are discounted at rate $\delta \in (0,1)$. The next proposition shows that welfare losses due to inefficient duplication of costly search only vanish in arbitrarily patient societies (equivalently, in the long run). These losses, however, remain significant in the short and medium run.

Proposition 3. *For all $\delta \in (0,1)$, the equilibrium social utility is larger in the complete network than in the network in which agents only observe their most immediate predecessor. The difference vanishes as $\delta \to 1$.*

Single Decision Maker Benchmark. Suppose agents are replaced by a single decision maker who draws a new search cost in each period and faces the same connections' structure as the agents in the society. Such decision maker discounts future payoffs at rate $\delta \in (0,1)$, internalizes future gains of current search, and samples each action exactly once along the same information path. In OIP networks, each agent is (directly or indirectly) linked to all his predecessors, and so all agents lie on the same information path. Hence, the single decision maker achieves the same social utility in all OIP networks.[13]

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

(i) Agents do not internalize future gains of today's search. As a result, exploration and convergence to the best action are too slow in equilibrium.

(ii) Equilibrium behavior displays inefficient duplication of costly search:

  (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.

[12] For the descriptive analysis in this section, assume that search costs are not bounded away from 0. The formal details are in Appendix H.

[13] See 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.

(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 $j \in B(n)$. Again, agents sample action x multiple times.

Whereas (a) occurs in all OIP networks, (b) does not in the complete network.

Equilibrium welfare losses disappear if and only if complete learning occurs and the society is arbitrarily patient. If search costs are bounded away from 0, or if the focus is on short- and medium-run outcomes, however, welfare losses can be significant.

Proposition 4. *In OIP networks, the equilibrium social utility converges to the social utility implemented by the single decision maker as $\delta \to 1$ if and only if search costs are not bounded away from* 0.

A Simple Policy to Increase Welfare and Efficiency. Reducing network transparency in OIP networks leads to inefficient duplication of costly search because agents fail to recognize actions that are revealed to be inferior by their predecessors' choices. A simple policy intervention, however, improves efficiency and equilibrium welfare. In particular: for all $\delta \in (0, 1)$, the equilibrium social utility in OIP networks is the same as in the complete network (the most efficient OIP network) if agents observe the aggregate history of past choices in addition to their neighbors' choices. Interestingly, the simple policy of letting agents observe the aggregate history of past choices is a common practice. In particular, online platforms that aggregate past individual decisions by sorting different items according to their popularity or sales rank serve this purpose.

The explanation of this result is simple. First, observing the aggregate history of past choices does not change equilibrium behavior at the first search stage: each agent still samples first the action taken by his immediate predecessor.[14] Second, if an action is revealed to be inferior by time n, that action is never sampled again by any agent $m \geq n$. To see this, suppose $a_j = x$, $a_{j+1} = \neg x$, and consider any agent $n > j + 1$. Agent n samples first action a_{n-1}. Since each agent samples first the action taken by his immediate predecessor and takes the best one between those he samples, it must be that $a_{n-1} = \neg x$. Now, if agent n observes the aggregate history of past choices, he realizes that $q_{\neg x} \geq q_x$ even when $j \notin B(n)$. In fact, from the aggregate history agent n understands 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, inefficient duplication of costly search disappears.

5 Discussion and Extensions

5.1 Relation to the SSLM and MFP

Long-Run Learning in the SSLM. In the SSLM, agents wish to match their actions with an unknown state of nature. They observe a free private signal and the choices of their

[14] This follows by an inductive argument as the one proving part (i) of Lemma 12 in Appendix E.

neighbors in the network before making their choice. The private signal is informative about the relative quality of all alternatives. Information aggregates if, in the long run, agents match their action with the state of nature with probability 1. Information diffuses if, in the long run, agents match their action with the state of nature with at least the same probability as an isolated agent who has access to the most informative private signals. If private signals induce unbounded beliefs, diffusion coincides with aggregation. Each learning metric corresponds to a learning principle: aggregation to a LSP and diffusion to an IP, these learning principles being appropriately defined and developed for the SSLM (see Acemoglu et al. (2011), Lobel and Sadler (2015), and Golub and Sadler (2016)).[15]

There are four fundamental differences between my setting and the SSLM. First, private information is different in kind: here, sampling an action perfectly reveals its quality only, whereas in the SSLM agents receive imperfect signals about the actions' relative quality. Second, private information is generated by equilibrium play rather than being exogenously available. Third, the inferential challenge differs: agents maximize the value of a sequential information acquisition program rather than the probability of matching a state of nature or an ex ante expected utility. Fourth, there is no obvious parallel between the learning metrics in the two settings: whereas aggregation and diffusion are utility-based learning metrics—in the SSLM, each agent's problem of maximizing his expected utility coincides with the problem of maximizing the probability of matching his action with the state of nature—complete and maximal learning are not (see also Remarks 1 and 2). Thus, complete (resp., maximal) learning is not a simple relabeling in my setting of the notions of information aggregation (resp., diffusion) in the SSLM.

These differences have many implications. In particular, whereas an IP and a LSP are also at play in my setting, their notion, mechanics, and applicability are different than in the SSLM. First, the parallel between the IP I establish for my setting and that for the SSLM is fragile. In the SSLM, an IP holds independently of whether private beliefs are bounded or not. Thus, information diffuses—and aggregates with unbounded private beliefs—via improvements upon imitation in all sufficiently connected networks with identifiable information paths. In contrast, in my setting, when search costs are bounded away from 0, maximal learning via an IP fails in all network topologies. Second, large-sample and martingale convergence arguments play a more limited role in my setting than in the SSLM (and in most social learning models). Nevertheless, I leverage on the LSP to characterize a sizeable class of network topologies in which maximal learning occurs. This class neither includes nor is included in the class of networks in which information aggregates with bounded private beliefs via the LSP in the SSLM. Third, an IP (resp., a LSP) corresponds to complete (resp., maximal) learning in my setting, and not vice versa as one may conjecture by (incorrectly) drawing a parallel between complete (resp., maximal) learning and information aggregation (resp., diffusion). Fourth, necessary and sufficient conditions on network topologies for complete (resp., maximal) learning in my setting do not coincide with those for aggregation (resp., diffusion) in the SSLM.

[15] The IP for the SSLM relates to: the welfare improvement principle in Banerjee and Fudenberg (2004) and Smith and Sørensen (2014); the imitation principle in Bala and Goyal (1998) and Gale and Kariv (2003).

Long-Run Learning in MFP. In the complete network, my model reduces to that of MFP. MFP study complete learning but not maximal learning, which is new to my paper. Complete learning occurs in the complete network if and only if search costs are not bounded away from 0. This equivalence no longer holds in general networks. My analysis identifies conditions on both network topologies and search technologies for positive learning outcomes to obtain or fail and uncovers the learning principles that are, or are not, at play.

Theorem 1 characterizes network topologies in which search costs that are not bounded away from 0 are sufficient for complete learning. This is so in all sufficiently connected networks with identifiable information paths—the complete network being the simplest in this class. Establishing an IP is key to make this general characterization possible.

Theorem 2 characterizes a sizeable class of network topologies in which maximal learning occurs. In all such networks, whenever maximal and complete learning coincide with $\underline{c} > 0$, search costs that are not bounded away from 0 are not necessary for complete learning. The analysis also makes clear how and to what extent societies can learn by observing large samples of individual choices and aggregating the information therein contained. These results have no parallel in MFP.

By Theorem 3–(i), maximal learning fails in network topologies with non-expanding subnetworks; in such networks, search costs that are not bounded away from 0 are not sufficient for complete learning. Theorem 3–(ii) characterizes network topologies in which maximal learning fails if search costs are bounded away from 0. In such networks—the complete network being an element of this class—search costs that are not bounded away from 0 are necessary and sufficient (if the network topology has expanding subnetworks) for complete learning. With the notion of maximal learning, I show that, in such networks, social learning fails discontinuously with respect to its benchmark metric by removing 0 from the support of the search costs distribution.

Speed and Efficiency of Social Learning. My paper is the first to study how the speed and efficiency of social learning depend on the network structure in a setting with endogenous private information. Thus, the results in Section 4 have no parallel in MFP or the SSLM. These results are noteworthy for two reasons. First, in OIP networks the rate of convergence, the probability of wrong herds, and long-run welfare are independent of network transparency, the density of connections, and their correlation pattern. Second, the rate of convergence can be analytically characterized in all such networks.

Rosenberg and Vieille (2019) measure efficiency in the SSLM by the expectation of the total welfare loss, which is equal to the total number of incorrect choices under a 0-1 loss function, and define learning to be efficient if the expected welfare loss is finite (see also Hann-Caruthers et al. (2018)). They focus on two polar setups assuming that each agent either observes the entire sequence of earlier choices 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 my setting, the results on the irrelevance of how far in the past agents can observe is much stronger: first, it holds for long-run welfare as well as for the probability of wrong herds and the rate of convergence; second, it neither depends on the number of immediate predecessors that agents observe nor on the density and the dependence structure among connections.

Behavioral Outcomes and Belief Evolution. In my setting, behavioral outcomes and belief evolution are typically very different than in the SSLM. For instance, in my setting, actions and beliefs are essentially identical in all OIP network. In contrast, in the SSLM, equilibrium dynamics dramatically change as the number of immediate predecessors that are observed varies: in the complete network, herding (i.e. all agents conforming in behavior after some time) and informational cascades (i.e. all agents ignoring private signals after some time) always emerge;[16] instead, when each agent only observes his most immediate predecessor's choice, beliefs and, actions cycle indefinitely (see Celen and Kariv (2004)).

In my setting, behavioral outcomes and belief evolution can vary a lot across different network topologies. In the complete network, agents evolve from one action to another until they eventually settle on one action. Herding occurs with probability 1 in finite time and learning occurs by discarding actions that are revealed inferior: once an agent deviates from the action of his predecessor, no later agent will take, or even sample the predecessor's action. Actions are always improving, i.e. each agent takes a weakly better action than his predecessors. Such regularities are typically lost in more complex networks: herding need not occur; agents may generate long patterns of disagreement before settling on one action; inferior actions may be inefficiently sampled multiple times.

5.2 Information Acquisition and Choice with Social Information

The prize distributions associated with the two actions are not exogenous, as in Weitzman (1979), but depend on an agent's social information, which is endogenously generated by other agents' equilibrium play. The equilibrium characterization in Section 2.1 sheds light on how others' actions inform what and how much agents choose to learn in equilibrium.

Social Information and Information Choice. Different network structures result in different optimal sampling sequences. In some networks, agents always find it optimal to sample first the action taken by their most recent neighbor. This is so, for instance, in the complete network, under uniform random sampling of at most two past agents, and when agents observe random numbers of immediate predecessors (see Section 4.1). In contrast, agents who observe only isolated agents always find it optimal to sample first the action with the highest relative share in their neighborhood (see Section 3.2). In more general networks, however, no informational monotonicity property links an agent's optimal sampling sequence to the choices of his most recent neighbors or to the relative shares of choices he observes. In such cases, neither the most recent nor the most popular choices uniquely determine the agent's information choice.

Social Information and Information Acquisition. Now consider how agents trade-off *exploration* (sampling the second action) and *exploitation* (using social information to save on the cost of the second search). First, (3)–(5) imply that $t^{\emptyset}(q_{s_n^1}) \geq t_n(q_{s_n^1})$; that is, given

[16] At least in a limiting sense, see Smith and Sørensen (2000).

$q_{s_n^1}$, the expected gain from the second search, and so the incentive to explore, is larger for isolated agents than for agents who can exploit their social information.

Second, the expected gain from the second search for an isolated agent, and so his incentive to explore, decreases with the quality of the first action sampled: $t^{\emptyset}(q_{s_n^1}) \leq t^{\emptyset}(q'_{s_n^1})$ for $q_{s_n^1} \geq q'_{s_n^1}$. This is standard for single-agent information acquisition problems.

Finally, in contrast to the single-agent problem, exploration incentives for an agent n with $B_n \neq \emptyset$ need not be monotone in the quality of the first action sampled. This is so because n's expected gain from the second search depends on $P_n(q_{s_n^1})$, the probability that no agent in $\widehat{B}(n, s_n^1)$ has sampled action $\neg s_n^1$ given the first action sampled has quality $q_{s_n^1}$. This conditional probability need not be monotone in $q_{s_n^1}$. On the one hand, a high $q_{s_n^1}$ suggests that some agent has explored both alternatives, discarding the one with low quality to adopt the superior one. On the other hand, precisely this effect, combined with the fact that $t^{\emptyset}(\cdot)$ is decreasing, hints that the incentives to explore (exploit social information) decrease (increase) with $q_{s_n^1}$. This is the central trade-off in the environment I study. Depending on the primitives of the model, either force may prevail(see Lomys (2020))..

5.3 Extensions

More than Two Actions. A natural extension of the model is to allow for more than two available actions (i.e. $|X| > 2$). I sketch this extension in Section I. In short: (i) the results on complete and maximal learning (Theorems 1–3) remain unchanged with more than two actions; (ii) the results on the speed and efficiency of social learning need not always hold as stated but most of the insights they highlight remain valid with more than two actions.

Heterogeneous Preferences. Another natural extension of the model is to consider equilibrium learning when preferences are heterogeneous. For instance, the payoff agents get from an action may depend on a common component and a private idiosyncratic shock. Lobel and Sadler (2016) study preference heterogeneity in the SSLM. In their setting: the IP suffers, as imitation no longer guarantees the same payoff that a neighbor obtains when preferences are diverse; the LSP, instead, has more room to operate. In my setting, the IP has more bite than the LSP. Hence, my results suggest that preference heterogeneity may substantially disrupt positive learning outcomes in collective search environments.

Search Technology. I assume that each agent n samples the first action at no cost, whereas sampling the second action involves a cost c_n. All results remain unchanged if both searches cost c_n but the agent cannot abstain, and so must search at least once. That both searches cost the same for an agent captures the idea that the cost of information acquisition is idiosyncratic to (and so heterogeneous across) agents, but not to actions, which are ex ante identical. That all agents take an action is a common assumption in the social learning literature.

That agents can only take an action they have sampled captures the idea that taking an action involves learning about its quality, availability, or functioning. Allowing agents to take an unsampled action is a non-trivial extension of the model. The solution to this problem is not yet well understood even for the simpler case of a single-agent, with Doval

(2018) making recent progress on the topic. Addressing this issue in a social learning setting goes beyond the scope of the paper.

6 Conclusion

The paper studies social learning in networks with information acquisition and choice. Bayesian agents act in sequence, observe the choices of their connections, and acquire information via sequential search. The analysis delivers two main sets of results. First, it provides a systematic characterization of the conditions on network topologies and search costs distributions for positive learning outcomes to obtain or fail. Second, it sheds light on how the speed and efficiency of social learning depend on the observation network.

The paper contributes to both the economic theory of social learning and its applications to the economics of social media and Internet search. The paper's theoretical novelty is to study costly information acquisition and choice in a model of Bayesian learning over general networks. By and large, the literature on social learning in networks neglects the complexity introduced by costly acquisition of private information. Prior work either focuses on particular network structures or posits simple individual decision rules; yet, it acknowledges the importance of a general analysis within the Bayesian benchmark (see, e.g., Sadler (2014) and Golub and Sadler (2016)). From the viewpoint of applications, the model is motivated by the large and growing use that individuals make of search engines to gather information. In online environments, others' choices are readily available via online social networks and popularity rankings. Thus, individuals' search behavior and their resulting purchase, adoption, or sharing decisions are typically informed by their connections' choices. This paper sheds light on the interplay between social information and individual incentives to acquire and choose private information as a central mechanism for social influence, learning, and diffusion.

Several questions remain for future research. First, the exact boundary between possibility and impossibility of maximal learning when search costs are bounded away from zero remains unknown. Second, quantifying the speed and efficiency of social learning in more general network topologies is an important, but complex, task. Third, it remains to study the design of more sophisticated incentives schemes to reduce inefficiencies and foster social exploration.[17] More broadly, one may assume that acquiring private information and observing past histories are both costly activities. If agents are heterogeneous across these two dimensions, in equilibrium some agents will specialize in search, while others in networking, thus enabling information to diffuse. Studying how agents make this trade-off, which network structures endogenously emerge, and the implications for social learning seems a promising direction for future research. Finally, it would be interesting to empirically disentangle the interplay between social information and search behavior from other channels of social influence and to asses its impact on social learning outcomes and diffusion patterns.

[17] A recent and growing literature, such as Smith, Sørensen and Tian (2020), 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 related settings. Addressing these questions in a search setting would be interesting.

A Proofs for Section 2.1

Pick any $x \in X$ and q with $\min \operatorname{supp}(\mathbb{P}_Q) < q < \max \operatorname{supp}(\mathbb{P}_Q)$, and note:

$$\mathbb{P}_Q(q_x \le q) = \mathbb{P}_Q(q_{\neg x} \le q), \tag{7}$$

$$\mathbb{P}_{\Omega|q_{\neg x} \ge q_x}(q_x \le q) = \mathbb{P}_{\Omega|q_x \ge q_{\neg x}}(q_{\neg x} \le q), \tag{8}$$

and

$$\mathbb{P}_{\Omega|q_{\neg x} \ge q_x}(q_x \le q) > \mathbb{P}_Q(q_x \le q). \tag{9}$$

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,

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

Here, E_n^{xC} ($E_n^{\neg x C}$) is the complement of E_n^x ($E_n^{\neg x}$), the third equality holds by (7) and (8), and the inequality follows from (9) and the assumption $P_n(x) < P_n(\neg x)$. ■

B Sufficient Conditions on the Network Topology for Condition (ii)–(b) in Theorem 1

If the network topology satisfies any of the conditions (i)–(vi) below, then it has a sequence of neighbor choice functions satisfying condition (ii)–(b) in Theorem 1 in all equilibria of the model.

(i) The network topology is deterministic.

(ii) The network topology has independent neighborhoods.

(iii) The network topology has deterministic information paths. That is, there exists a sequence of neighbor choice functions such that the corresponding chosen neighbor topology is deterministic (this is so, for instance, in OIP networks).

(iv) The network topology has a sequence of neighbor choice functions $(\gamma_n)_{n\in\mathbb{N}}$ such that neighborhoods $\{B(k)\}_{k=1}^m$ are independent of the event $\gamma_n(B(n)) = m$ for all $n, m \in \mathbb{N}$ with $n > m$.

(v) The network topology has a sequence of neighbor choice functions such that the corresponding chosen neighbor topology has low network distortion.

(vi) The network topology is Markovian. That is, neighborhoods $\{B(n)\}_{n\in\mathbb{N}}$ are conditionally independent given the state of an underlying Markov chain with finitely many states.

Under conditions (i)–(iv), we have

$$\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \underset{x\in X}{\arg\max}\; q_x \mid \gamma_n(B(n))\right) - \mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \underset{x\in X}{\arg\max}\; q_x\right) = 0, \tag{10}$$

and so the claim trivially follows. Regarding conditions (v) and (vi), I refer to Section 5 in Lobel and Sadler (2015) for the formal definitions of low network distortion and Markovian network topology. Under such conditions, the equality in (10) need not hold; however, adapting the arguments in Lobel and Sadler (2015), one can show that, under such conditions, there exists a sequence of neighbor choice functions such that the difference in the left-hand side of (10) is arbitrarily small (with arbitrarily large probability) for large enough n in all equilibria of the model.

C Proof of Theorem 1

Theorem 1 follows by combining the next two propositions which, together, establish an IP for the search setting I study. The first proposition shows that complete learning via improvements upon imitation occur if certain conditions hold.

Proposition 5. *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 $N_{\varepsilon\eta} \in \mathbb{N}$ such that for all $n > N_{\varepsilon\eta}$, with probability at least $1-\eta$,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n))\right) \geq \mathcal{Z}\left(\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x\in X} q_x\right)\right) - \varepsilon.$$

Then, complete learning occurs.

Condition (c) requires the existence of 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 next proposition shows that this is possible if search costs are not bounded away from 0.

Proposition 6. *Suppose search costs are not bounded away from 0, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence of neighbor choice functions. Then, there exists an increasing and continuous function $\mathcal{Z}\colon [1/2,1] \to [1/2,1]$, with $\mathcal{Z}(\beta) > \beta$ for all $\beta \in [1/2, 1)$, $\mathcal{Z}(1) = 1$, and such that*

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

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

The next two sections contain the proofs of Propositions 5 and 6.

C.1 Proof of Proposition 5

Preliminaries. In equilibrium, each agent takes the best action between those he sampled. Since each agent must sample at least one action, the next lemma follows.

Lemma 1. *If $\lim_{n\to\infty} \mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x) = 1$, then complete learning occurs.*

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

Lemma 2. *For all $n \in \mathbb{N}$, we have $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x) \geq \mathbb{P}_\sigma(s_1^1 \in \arg\max_{x\in X} q_x)$.*

Proof. For $n = 1$, the claim trivially holds. Now fix any $n > 1$ 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,

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

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

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = 0\right) = \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = 0\right). \tag{11}$$

Second, suppose $0 < b < n$, so that $B_n \neq \emptyset$. By the characterization of the equilibrium decision s_n^1 in Section 2.1, we have $\mathbb{P}_\sigma(E_n^{s_n^1} \mid c_n, B_n, a_k\ \forall k \in B_n) \leq \mathbb{P}_\sigma(E_n^{s_1^1} \mid c_n, B_n, a_k\ \forall k \in B_n)$ for all c_n, $B_n \neq \emptyset$, and a_k for all $k \in B_n$. By integrating over all possible search costs and choices of the agents in the neighborhood, we obtain $\mathbb{P}_\sigma(E_n^{s_n^1} \mid B_n) \leq \mathbb{P}_\sigma(E_n^{s_1^1} \mid B_n)$ for all $B_n \neq \emptyset$. Integrating further over all B_n such that $\gamma_n(B_n) = b$ we have $\mathbb{P}_\sigma(E_n^{s_n^1} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{s_1^1} \mid \gamma_n(B(n)) = b)$. Thus, 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,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right) \geq \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right). \tag{12}$$

The desired result obtains by observing that

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

where: the equalities hold by the law of total probability; the inequality holds by (11) and (12).■

Proof of Proposition 5. 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

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \geq \phi_k \qquad \forall n \geq \alpha_k. \tag{13}$$

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.

Part 1. By assumptions (a) and (c) of the proposition, for all $\alpha \in \mathbb{N}$ and all $\varepsilon > 0$, there exist $N(\alpha, \varepsilon) \in \mathbb{N}$ and a sequence of neighbor choice functions $(\gamma_k)_{k\in\mathbb{N}}$ such that

$$\mathbb{Q}\big(\gamma_n((B(n)) = b, b < \alpha\big) < \frac{\varepsilon}{2}, \tag{14}$$

$$\mathbb{P}_\sigma\left(\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n))\right) < \mathcal{Z}\left(\mathbb{P}_\sigma\left(s_{\gamma_n(B(n))}^1 \in \arg\max_{x\in X} q_x\right)\right) - \varepsilon\right) < \frac{\varepsilon}{2} \tag{15}$$

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 k to prove (13). Since the qualities of the two actions are i.i.d. draws and agent 1 has no prior information,

$$\mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right) = \frac{1}{2}. \tag{16}$$

From Lemma 2,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \geq \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right) \tag{17}$$

for all n. From (16) and (17) we have

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

which together with $\alpha_1 = 1$ and $\phi_1 = \frac{1}{2}$ establishes (13) for $k = 1$. Assume that (13) holds for an arbitrary k, that is

$$\mathbb{P}_\sigma\left(s_j^1 \in \arg\max_{x\in X} q_x\right) \geq \phi_k \quad \forall j \geq \alpha_k, \tag{18}$$

and consider some agent $n \geq \alpha_{k+1}$. To establish (13) for $n \geq \alpha_{k+1}$ observe that

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

where the inequality follows from (14) and (15), the inductive hypothesis in (18), and the assumption that $\mathcal{Z}$ is increasing.

Part 2. 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 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. ∎

C.2 Proof of Proposition 6

Proposition 6 follows by combining several lemmas, which I next present. Hereafter, let a state of the world $\omega \in \Omega$, 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 $0 \leq b < n$, be agent n's chosen neighbor.

Let $\tilde{s}_n^1$ be agent n's coarse optimal decision at the first search stage when he only uses information from neighbor b. The optimal search policy, as characterized in Section 2.1, requires $\tilde{s}_n^1 \in \arg\min_{x \in X} \mathbb{P}_\sigma(E_n^x \mid \gamma_n(B(n)) = b, a_b)$. Hereafter, I assume n samples first action a_b in case of indifference. This does not affect the results. The next lemma follows.

Lemma 3. *If $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$, then $\tilde{s}_n^1 = a_b$.*

Remark 3. As $\gamma_n(B(n)) = 0$ iff $B(n) = \emptyset$, we have $\tilde{s}_n^1 = s_n^1$ conditional on $\gamma_n(B(n)) = 0$.

Lemma 3 applies to network topologies in which $\mathbb{Q}(|B(n)| \leq 1) = 1$ for all n and so, in particular, to all chosen neighbor topologies. That is, we have the following.

Lemma 4. *Suppose $\mathbb{Q}(|B(n)| \leq 1) = 1$ for all $n \in \mathbb{N}$. Then, $\mathbb{P}_\sigma(E_n^{a_b} \mid \widehat{B}(n) = \widehat{B}_n) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \widehat{B}(n) = \widehat{B}_n)$ for all n and b, with $0 \leq b < n$, and for all $\widehat{B}_n$ that occur with positive probability. It follows that $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$ for all n and b, with $0 \leq b < n$.*

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

Now suppose $\mathbb{P}_\sigma(E_n^{a_b} \mid \widehat{B}(n) = \widehat{B}_n) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \widehat{B}(n) = \widehat{B}_n)$ for all $n \leq k$ and all $\widehat{B}_n$ that occurs with positive probability. 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 $1 < g \leq l$, $B_{\pi_g} = \{\pi_{g-1}\}$. 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.■

Definition 10. *Fix a state of the world $\omega \in \Omega$. The following objects are defined:*

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

Remark 4. For all $b \in \mathbb{N}$, we have $\beta \geq \frac{1}{2}$. This is so because the distribution of the quality of the first action sampled by an agent first-order stochastically dominates 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 search costs distribution, the function $t^\emptyset(\cdot)$ defined in (3), and the thresholds $P_{b,n}^\sigma(q_{\min})$ and $P_{b,n}^\sigma(q_{\max})$.

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,*

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

Proof. By Lemma 3,

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right) = \mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right). \tag{19}$$

Moreover,

$$\begin{aligned}
&\mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right) \\
&= \mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b, s_b^2 = \neg s_b^1\right)\mathbb{P}_\sigma(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b) \\
&+ \mathbb{P}_\sigma\left(a_b \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b, s_b^2 = d\right)\mathbb{P}_\sigma(s_b^2 = d \mid \gamma_n(B(n)) = b) \\
&= \mathbb{P}_\sigma(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b) \qquad (20) \\
&+ \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right)(1 - \mathbb{P}_\sigma(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b)) \\
&= \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right) \\
&+ \mathbb{P}_\sigma(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b)\left(1 - \mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right)\right).
\end{aligned}$$

Here: the first equality holds by the law of total probability; the second equality holds because when b samples both actions, $s_b^2 = \neg s_b^1$, he takes the best one, so that $\mathbb{P}_\sigma(a_b \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b, s_b^2 = \neg s_b^1) = 1$, and when b only samples one action, $s_b^2 = d$, he takes that action, so that $\mathbb{P}_\sigma(a_b \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b, s_b^2 = d) = \mathbb{P}_\sigma(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b)$. The desired result follows from (19) and (20).■

Lemma 6. *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,*

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

Proof. By Lemma 5,

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right) = \beta + \mathbb{P}_\sigma(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b)(1-\beta). \tag{21}$$

Moreover, by the law of total probability,

$$\mathbb{P}_\sigma(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b)$$

$$
\begin{aligned}
&= \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\Big)\mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \\
&+ \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\Big)\mathbb{P}_\sigma\Big(s_b^1 \notin \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \\
&= \beta\, \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\Big) \qquad (22) \\
&+ (1-\beta)\, \mathbb{P}_\sigma\Big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\Big).
\end{aligned}
$$

By the equilibrium characterization in Section 2.1, we have: conditional on $\gamma_n(B(n)) = b$ and $s_b^1 \in \arg\max_{x\in X} q_x$, $s_b^2 = \neg s_b^1 \iff c_b \le P_{b,n}^\sigma(q_{\max})t^\emptyset(q_{\max})$; conditional on $\gamma_n(B(n)) = b$ and $s_b^1 \notin \arg\max_{x\in X} q_x$, $s_b^2 = \neg s_b^1 \iff c_b \le P_{b,n}^\sigma(q_{\min})t^\emptyset(q_{\min})$. Thus,

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

and

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

Thus, equation (22) can be rewritten as

$$
\begin{aligned}
\mathbb{P}_\sigma\big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\big) &= \beta F_C\big(P_{b,n}^\sigma(q_{\max})t^\emptyset(q_{\max})\big) \\
&\quad + (1-\beta)F_C\big(P_{b,n}^\sigma(q_{\min})t^\emptyset(q_{\min})\big).
\end{aligned} \qquad (23)
$$

The desired result follows by combining (21) and (23). ■

By the previous lemma, $(1-\beta)[\beta F_C(P_{b,n}^\sigma(q_{\max})t^\emptyset(q_{\max})) + (1-\beta)F_C(P_{b,n}^\sigma(q_{\min})t^\emptyset(q_{\min}))]$ 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 simpler lower bound on the amount of this improvement.

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,*

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

Proof. If at least one of the agents in $\widehat{B}(b, s_b^1)$ samples both actions, then $s_b^1 \in \arg\max_{x\in X} q_x$. Thus, $\beta \ge 1 - \mathbb{P}_\sigma(E_b^{s_b^1} \mid \gamma_n(B(n)) = b)$, or

$$1-\beta \le \mathbb{P}_\sigma\Big(E_b^{s_b^1} \mid \gamma_n(B(n)) = b\Big). \qquad (24)$$

Moreover, by the law of total probability,

$$
\begin{aligned}
&\mathbb{P}_\sigma\Big(E_b^{s_b^1} \mid \gamma_n(B(n)) = b\Big) \\
&= \mathbb{P}_\sigma\Big(E_b^{s_b^1} \mid \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\Big)\mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \qquad (25) \\
&+ \mathbb{P}_\sigma\Big(E_b^{s_b^1} \mid \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\Big)\mathbb{P}_\sigma\Big(s_b^1 \notin \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \\
&= \beta P_{b,n}^\sigma(q_{\max}) + (1-\beta)P_{b,n}^\sigma(q_{\min}).
\end{aligned}
$$

Combining (24) and (25) yields $1-\beta \leq \beta P^{\sigma}_{b,n}(q_{\min}) + (1-\beta)P^{\sigma}_{b,n}(q_{\min})$, and therefore

$$\max\left\{P^{\sigma}_{b,n}(q_{\min}), P^{\sigma}_{b,n}(q_{\max})\right\} \geq 1-\beta. \tag{26}$$

Finally, observe that

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

Here, the first equality holds by Lemma 6; the first inequality holds as $\beta \geq (1-\beta)$ (by Remark 4, $\beta \geq 1/2$); the second inequality holds as $t^{\emptyset}(q_{\max}) \leq t^{\emptyset}(q_{\min})$ and the CDF F_C is increasing; the third inequality holds because F_C is non-negative; the last inequality follows as $\max\{F_C(P^{\sigma}_{b,n}(q_{\max})t^{\emptyset}(q_{\max})), F_C(P^{\sigma}_{b,n}(q_{\min})t^{\emptyset}(q_{\max}))\} \geq F_C((1-\beta)t^{\emptyset}(q_{\max}))$, which holds because of (26) and the fact that F_C is increasing. The desired result follows.■

The previous lemmas describe the improvement an agent can make over his chosen neighbor by discarding 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 as

$$\overline{\mathcal{Z}}(\beta) := \beta + (1-\beta)^2 F_C((1-\beta)t^{\emptyset}(q_{\max})). \tag{27}$$

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 $\mathbb{P}_{\sigma}(E^{a_b}_n \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_{\sigma}(E^{\neg a_b}_n \mid \gamma_n(B(n)) = b)$, we have

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

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 $\overline{\mathcal{Z}}$, defined by (27), satisfies the following properties:*

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

(b) If search costs are not bounded away from 0, then $\overline{\mathcal{Z}}(\beta) > \beta$ for all $\beta \in [1/2,1)$.

(c) It 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 0, $F_C((1-\beta)t^{\emptyset}(q_{\max})) > 0$.[18] 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.

[18] Note that $t^{\emptyset}(q_{\max}) = 0$ if $q_{s^1_b} = q_{\max} = \max\operatorname{supp}(\mathbb{P}_Q)$ 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.

For part (c), set $\alpha := (1-\beta)t^{\emptyset}(q_{\max})$. 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

$$\mathcal{Z}(\beta) := \frac{1}{2}\left(\beta + \sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)\right). \tag{28}$$

Lemma 9. *The function $\mathcal{Z}$, defined by (28), satisfies the following properties:*

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

(b) If search costs are not bounded away from 0, then $\mathcal{Z}(\beta) > \beta$ for all $\beta \in [1/2,1)$.

(c) It 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. Thus, the average of these two functions, which is $\mathcal{Z}$, is an increasing function, establishing the first part of (c). To establish continuity of $\mathcal{Z}$ in $[1/2,1)$, I argue by contradiction. Suppose $\mathcal{Z}$ is discontinuous at some $\beta' \in [1/2,1)$. If so, $\sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)$ 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 $\delta > 0$, $\sup_{r\in[1/2,\beta'+\delta]} \overline{\mathcal{Z}}(r) > \overline{\mathcal{Z}}(\beta) + \varepsilon$ for all $\beta \in [1/2,\beta')$. This contradicts that $\overline{\mathcal{Z}}$ 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 an *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,*

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

Proof. If $\mathcal{Z}(\beta) = \beta$, the result follows from Lemma 6. Suppose next that $\mathcal{Z}(\beta) > \beta$. By (28), this implies that $\mathcal{Z}(\beta) < \sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)$. Thus, there exists $\overline{\beta} \in [1/2,\beta]$ such that

$$\overline{\mathcal{Z}}(\overline{\beta}) \geq \mathcal{Z}(\beta). \tag{29}$$

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 agent b's choice with some probability. Thus, suppose 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

$$\tilde{a}_b = \begin{cases} a_b & \text{with probability } (2\overline{\beta}-1)/(2\beta-1) \\ 0 & \text{with probability } (\beta-\overline{\beta})/(2\beta-1) \\ 1 & \text{with probability } (\beta-\overline{\beta})/(2\beta-1), \end{cases} \tag{30}$$

with the realization of $\tilde{a}_b$ independent of the rest of n's information set. Under the assumption

$\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \le \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$, we have

$$\mathbb{P}_\sigma\big(E_n^{\tilde{a}_b} \mid \gamma_n(B(n)) = b\big) \le \mathbb{P}_\sigma\big(E_n^{\neg \tilde{a}_b} \mid \gamma_n(B(n)) = b\big). \tag{31}$$

The relation in (31), together with the equilibrium characterization in Section 2.1, 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 $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \le \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$ and (30) also imply that $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \le \mathbb{P}_\sigma(E_n^{\tilde{a}_b} \mid \gamma_n(B(n)) = b)$. 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

$$\mathbb{P}_\sigma\Big(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \ge \mathbb{P}_\sigma\Big(\widetilde{\tilde{s}_n^1} \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big). \tag{32}$$

Now denote with $\tilde{s}_b^1$ the decision of the fictitious agent about which action to sample first. From (30), 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,

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

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

$$\mathbb{P}_\sigma\Big(\widetilde{\tilde{s}_n^1} \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \ge \overline{\mathcal{Z}}(\overline{\beta}). \tag{33}$$

Since $\overline{\mathcal{Z}}(\overline{\beta}) \ge \mathcal{Z}(\beta)$ (see equation (29)), the desired result follows from (32) and (33).■

It remains to show that s_n^1 does at least as well as its coarse version $\tilde{s}_n^1$ given $\gamma_n(B(n)) = b$. This is established with the next lemma and completes the proof of Proposition 6.

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

$$\mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \ge \mathbb{P}_\sigma\Big(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big).$$

Proof. Fix any agent n. If $b = 0$, then $\tilde{s}_n^1 = s_n^1$ by Remark 3, 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 2.1, we have $\mathbb{P}_\sigma(E_n^{s_n^1} \mid c_n, B_n, a_k\ \forall k \in B_n) \leq \mathbb{P}_\sigma(E_n^{\tilde{s}_n^1} \mid c_n, B_n, a_k\ \forall k \in B_n)$ for all c_n, $B_n \neq \emptyset$, and a_k for all $k \in B_n$. By integrating over all possible search costs and choices of the agents in the neighborhood, we obtain $\mathbb{P}_\sigma(E_n^{s_n^1} \mid B_n) \leq \mathbb{P}_\sigma(E_n^{\tilde{s}_n^1} \mid B_n)$ for all $B_n \neq \emptyset$. Integrating further over all B_n such that $\gamma_n(B_n) = b$ we conclude $\mathbb{P}_\sigma(E_n^{s_n^1} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\tilde{s}_n^1} \mid \gamma_n(B(n)) = b)$. Then, conditional on $\gamma_n(B(n)) = b$, the marginal distribution of action s_n^1's quality first-order stochastically dominates the marginal distribution of action $\tilde{s}_n^1$'s quality. Therefore,

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

as desired. ■

D Proof of Theorem 2

Suppose $\omega \notin \Omega(\underline{c})$. If $q_0 = q_1$, there is nothing to learn. Hereafter, suppose $\omega \in \overline{\Omega}(\underline{c}) := \{\omega \in \Omega : \omega \notin \Omega(\underline{c})$ and $q_0 \neq q_1\}$. The proof consists of two parts. In the first part, I show that agents $n \in S$ sample the best action in $\overline{\Omega}(\underline{c})$ at the first search with probability 1 as $n \to \infty$. The second part of the proof extends this result to all agents.

Part 1. Let $\overline{x} := \arg\max_{x \in X} q_x$. Restrict attention to agents $n \in S$, i.e. such that $B(n) \in \{\emptyset, B_n^{\emptyset}\}$. Consider first an agent n with $B(n) = \emptyset$. Agent n takes the best action any time he samples first action $\overline{x}$, which occurs with probability $1/2$, and any time he samples first action $\neg\overline{x}$ and his search cost is smaller that $t^{\emptyset}(q_{\neg\overline{x}})$. Since $\omega \in \overline{\Omega}(\underline{c})$, $q_{\neg\overline{x}} < q(\underline{c})$, and so the latter event occurs with positive probability. Therefore, agent n takes the best action ($a_n = \overline{x}$) with probability $\alpha > 1/2$. Providing an expression for α is irrelevant for the argument.

Next, consider an agent n with $B(n) = B_n^{\emptyset}$. By condition (ii) in Theorem 2, agent n knows he is observing only the choices of all his isolated predecessors. Thus, n's optimal decision at the first search stage depends on the relative shares of choices he observes. In particular, as $B(n) = \widehat{B}(n)$, we have:

$$s_n^1 = \begin{cases} 0 & \text{if } |\widehat{B}(n,0)| > |\widehat{B}(n,1)| \\ 1 & \text{if } |\widehat{B}(n,0)| < |\widehat{B}(n,1)| \end{cases},$$

and agent n samples the first action uniformly at random if $|\widehat{B}(n,0)| = |\widehat{B}(n,1)|$. To see this, note that, for such agent n, $|\widehat{B}(n,x)| > |\widehat{B}(n,\neg x)|$ implies $P_n(x) < P_n(\neg x)$, where $P_n(\cdot)$ is the probability defined by (2).

By condition (i) in Theorem 2, with probability 1 there are infinitely many isolated agents. Moreover, isolated agents' choices form a sequence of independent random variables. Thus, by the weak law of large numbers, the ratio $|\widehat{B}(n,\overline{x})|/n$ converges in probability to α and the ratio $|\widehat{B}(n,\neg\overline{x})|/n$ converges in probability to $1-\alpha$ as $n \to \infty$ with respect to $\mathbb{P}_\sigma$ conditional on $B(n) = B_n^{\emptyset}$ and $\omega \in \overline{\Omega}(\underline{c})$. Hence,

$$\lim_{n \to \infty} \mathbb{P}_\sigma\left(|\widehat{B}(n,\overline{x})| > |\widehat{B}(n,\neg\overline{x})| \mid B(n) = B_n^{\emptyset}, \omega \in \overline{\Omega}(\underline{c})\right) = 1. \tag{34}$$

Finally, for all $n \in S$ we have

$$1 \geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x \in X} q_x \mid n \in S, \omega \in \overline{\Omega}(\underline{c})\right)$$

$$\begin{aligned}
&= \mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) = \emptyset, \omega \in \overline{\Omega}(\underline{c})\Big)\mathbb{Q}\Big(B(n) = \emptyset \mid n \in S\Big) \\
&+ \mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) = B_n^\emptyset, \omega \in \overline{\Omega}(\underline{c})\Big)\mathbb{Q}\Big(B(n) = B_n^\emptyset \mid n \in S\Big) \\
&\geq \frac{1}{2}\mathbb{Q}\Big(B(n) = \emptyset \mid n \in S\Big) \\
&+ \mathbb{P}_\sigma\Big(|\widehat{B}(n,\overline{x})| > |\widehat{B}(n,\neg\overline{x})| \mid B(n) = B_n^\emptyset, \omega \in \overline{\Omega}(\underline{c})\Big)\mathbb{Q}\Big(B(n) = B_n^\emptyset \mid n \in S\Big).
\end{aligned} \tag{35}$$

Here: the equality holds by the law of total probability; the last inequality follows by the optimal policy at the first search stage for agents in S characterized above.

By (34), (35), and since $\lim_{n\to\infty} \mathbb{Q}(B(n) = B_n^\emptyset \mid n \in S) = 1$, we have

$$\lim_{n\to\infty} \mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x \mid n \in S, \omega \in \overline{\Omega}(\underline{c})\Big) = 1,$$

as desired.

Part 2. Consider now any agent $n \in \mathbb{N}$. By the characterization of the equilibrium decision s_n^1 in Section 2.1, we have $\mathbb{P}_\sigma(E_n^{s_n^1} \mid c_n, B_n, a_k \ \forall k \in B_n, \omega \in \overline{\Omega}(\underline{c})) \leq \mathbb{P}_\sigma(E_n^{s_b^1} \mid c_n, B_n, a_k \ \forall k \in B_n, \omega \in \overline{\Omega}(\underline{c}))$ for all c_n, B_n, $b \in B_n$, and a_k for all $k \in B_n$. By integrating over all possible search costs and choices of the agents in the neighborhood, we obtain $\mathbb{P}_\sigma(E_n^{s_n^1} \mid B_n, \omega \in \overline{\Omega}(\underline{c})) \leq \mathbb{P}_\sigma(E_n^{s_b^1} \mid B_n, \omega \in \overline{\Omega}(\underline{c}))$ for all B_n and $b \in B_n$. Thus, conditional on $B(n) = B_n$, the marginal distribution of the quality of action s_n^1 first-order stochastically dominates the marginal distribution of the quality of action s_b^1 for all $b \in B_n$. Hence,

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) = B_n, \omega \in \overline{\Omega}(\underline{c})\Big) \\
&\qquad \geq \max_{b\in B_n} \mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x\in X} q_x \mid B(n) = B_n, \omega \in \overline{\Omega}(\underline{c})\Big).
\end{aligned} \tag{36}$$

Since the state process is independent of the network topology, by condition (iv) in Theorem 2, for all $\varepsilon, \eta > 0$, there exists $N_{\varepsilon\eta} \in \mathbb{N}$ such that, for all agents $n > N_{\varepsilon\eta}$, with probability at least $1 - \eta$,

$$\begin{aligned}
&\max_{b\in B_n} \mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x\in X} q_x \mid B(n) = B_n, \omega \in \overline{\Omega}(\underline{c})\Big) \\
&\qquad \geq \max_{b\in B_n} \mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x\in X} q_x \mid \omega \in \overline{\Omega}(\underline{c})\Big) - \varepsilon.
\end{aligned} \tag{37}$$

Thus, by (36) and (37), for all $\varepsilon, \eta > 0$, there exists $N_{\varepsilon\eta} \in \mathbb{N}$ such that, for all agents $n > N_{\varepsilon\eta}$, with probability at least $1 - \eta$,

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) = B_n, \omega \in \overline{\Omega}(\underline{c})\Big) \\
&\qquad \geq \max_{b\in B_n} \mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x\in X} q_x \mid \omega \in \overline{\Omega}(\underline{c})\Big) - \varepsilon
\end{aligned} \tag{38}$$

for all B_n. By (38), for all $\tilde{\varepsilon} > 0$, there exists $N_{\tilde{\varepsilon}} \in \mathbb{N}$ such that, for all agents $n > N_{\tilde{\varepsilon}}$,

$$\begin{aligned} 1 &\geq \mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x \in X} q_x \mid \omega \in \overline{\Omega}(\underline{c})\Big) \\ &\geq \mathbb{E}_\sigma\Big[\max_{b \in B(n) \cap S} \mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x \in X} q_x\Big) \mid \omega \in \overline{\Omega}(\underline{c})\Big] - \tilde{\varepsilon} \\ &\geq \mathbb{E}_\sigma\Big[\max_{b \in B(n) \cap S} \mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x \in X} q_x\Big) \mid \max_{b \in B(n) \cap S} b \geq K, \omega \in \overline{\Omega}(\underline{c})\Big] \mathbb{Q}\Big(\max_{b \in B(n) \cap S} \geq K\Big) - \tilde{\varepsilon} \end{aligned} \tag{39}$$

for all $K \in \mathbb{N}$. Since agents $n \in S$ sample the best action at the first search with probability 1 as $n \to \infty$ whenever $\omega \in \overline{\Omega}(\underline{c})$,

$$\lim_{n \to \infty} \mathbb{E}_\sigma\Big[\max_{b \in B(n) \cap S} \mathbb{P}_\sigma\Big(s_b^1 \in \arg\max_{x \in X} q_x\Big) \mid \max_{b \in B(n) \cap S} b \geq K, \omega \in \overline{\Omega}(\underline{c})\Big] = 1. \tag{40}$$

Moreover, by condition (iii) in Theorem 2,

$$\lim_{n \to \infty} \mathbb{Q}\Big(\max_{b \in B(n) \cap S} b \geq K\Big) = 1. \tag{41}$$

That

$$\lim_{n \to \infty} \mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x \in X} q_x \mid \omega \in \overline{\Omega}(\underline{c})\Big) = 1$$

follows from (39)–(41), completing the proof. ■

E Equilibrium Strategies in OIP Networks

Let $P_1(q)$ be the posterior probability that agent 1 did not sample both actions given that the action he takes has quality q. The functional form of $P_1(q)$ is irrelevant for the following argument.

Lemma 12. *In OIP networks, equilibrium search policies are as follows:*

(i) At the first search stage, $s_n^1 = a_{n-1}$ for all agents $n \geq 2$.

(ii) At the second search stage, for all agents $n \geq 2$:

(a) $s_n^2 = d$ if $\neg a_{n-1}$ is revealed to be inferior to agent n.

(b) $s_n^2 = \neg a_{n-1}$ if $\neg a_{n-1}$ is not revealed to be inferior to agent n and

$$c_n \leq t_n(q_{s_n^1}) := \begin{cases} P_1(q_{s_n^1}) t^{\emptyset}(q_{s_n^1}) & \text{if } n = 2 \\ P_1(q_{s_n^1}) \big(\prod_{i=2}^{n-1}(1 - F_C(t_i(q_{s_n^1})))\big) t^{\emptyset}(q_{s_n^1}) & \text{if } n > 2. \end{cases} \tag{42}$$

Proof. Part (i) follows by induction. Consider agent 2 and his conditional belief over Ω given that agent 1 has taken action a_1. For action $\neg a_1$, only two cases are possible:

1. Agent 1 sampled $\neg a_1$. If so, $q_{\neg a_1} \leq q_{a_1}$, 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 $\mathbb{P}_Q$.

Under Assumption 1, the first case occurs with positive probability. Thus, agent 2's belief about the quality of action a_1 strictly first-order stochastically dominates his belief about the quality of action $\neg a_1$. That $s_2^1 = a_1$ is optimal follows.

Now consider any agent $n > 2$. Suppose that all agents up to $n-1$ follow this strategy, and that agent $n-1$ takes action a_{n-1}. If action $\neg a_{n-1}$ is revealed to be inferior to agent n, 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 to be inferior to agent n. By the same logic as before, n's belief about the quality of action a_{n-1} strictly first-order stochastically dominates his belief about the quality of action $\neg a_{n-1}$. That $s_n^1 = a_{n-1}$ is optimal follows.

For part (ii)–(a), suppose that $\neg a_{n-1}$ is revealed to be inferior to agent $n \geq 2$. 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), $s_{j+1}^1 = \neg a_{n-1}$. 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 $q_{a_{n-1}} \geq q_{\neg a_{n-1}}$. That $s_n^2 = d$ is optimal follows.

For part (ii)–(b), consider any agent $n \geq 2$ and suppose $\neg a_{n-1}$ is not revealed to be inferior to n. In OIP networks, $\widehat{B}(n) = \{1, \dots, n-1\}$ with probability 1. Moreover, by part (i), each agent samples first the action taken by his immediate predecessor. Thus, none of the agents in $\widehat{B}(n, s_n^1)$ has sampled action $\neg s_n^1$ only if $s_1^1 = s_n^1$, and $s_i^2 = d$ for $1 \leq i \leq n-1$. The thresholds in (42) provide an explicit formula for (5) in OIP networks. 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 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 gain from the second search is $P_1(q_{s_2^1}) t^{\emptyset}(q_{s_2^1})$, which is the the first line on the right-hand side of (42). 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, the probability that no agent in $\{1, \dots, n-1\}$ has sampled action $\neg s_n^1$ is $P_1(q_{s_n^1})(\prod_{i=2}^{n-1}(1 - F_C(t_i(q_{s_n^1}))))$. Hence, the second line on the right-hand side of (42) gives agent n's expected gain from the second search. The optimality of the proposed sequential search policy follows from the equilibrium characterization in Section 2.1.■

By Lemma 12, the probability of none of the first n agents sampling both actions is the same in all OIP networks, and thus so is the probability of agent n taking the best action.

Corollary 1. *Fix a state process, a search technology, and an agent $n \in \mathbb{N}$. Then, $\mathbb{P}_\sigma(a_n \in \arg\max_{x \in X} q_x)$ is the same in all OIP networks.*

F Proof of Theorem 3

Preliminaries. I begin with the notation and a result that will be used in the proof of Theorem 3.

▷ Let $q^D := \min\{\tilde{q} \in \operatorname{supp}(\mathbb{P}_Q) : \text{Assumption 1-part 1 holds}\}$ and $\Omega^D := \{\omega \in \Omega : q_i \geq q^D \text{ for } i = 0, 1\}$. In words, Ω^D consists of all states of the world ω in which, with positive probability, an isolated agent does not sample the second action independently of which action he samples first. Under Assumption 1, there exists some $\delta > 0$ such that

$$\mathbb{P}_\Omega(\Omega^D) \geq \delta. \tag{43}$$

If $\omega \in \Omega^D$, an agent with nonempty neighborhood does not sample the second action either with positive probability, independently of which action he samples first (see Section 2.1). Finally, by Assumption 1, $\mathbb{P}_\Omega(q_0 \neq q_1 \mid \omega \in \Omega^D) > 0$.

▷ Suppose $q_{\min} := \min\{q_0, q_1\} < q(\underline{c})$, so that, by Assumption 1, $\mathbb{P}_\Omega(q_0 \neq q_1 \mid q_{\min} < q(\underline{c})) > 0$. Maximal learning occurs only if the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to 0 as $n \to \infty$. If not, there would be a subsequence of agents who, with probability bounded away from 0, only observe (directly or indirectly) agents who have not compared the quality of the two actions and do not make this comparison either. Thus, maximal learning would 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 13. *Suppose* $\limsup_{n\to\infty} \mathbb{P}_\sigma(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \mid q_{\min} < q(\underline{c})) > 0$. *Then, maximal learning fails.*

Proof of Theorem 3, part (i). Since the network topology has non-expanding subnetworks, there exist some $K \in \mathbb{N}$, some $\varepsilon > 0$, and a subsequence of agents $\mathcal{N}$ such that

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

For all $n \in \mathcal{N}$, we have

$$\begin{aligned}
&\mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \mid q_{\min} < q(\underline{c})\big) \\
&= \mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K\big)\mathbb{Q}\big(|\widehat{B}(n)| < K\big) \\
&+ \mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| \geq K\big)\mathbb{Q}\big(|\widehat{B}(n)| \geq K\big) \\
&\geq \mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K\big)\mathbb{Q}\big(|\widehat{B}(n)| < K\big) \\
&\geq \varepsilon\mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K\big) \\
&= \varepsilon\mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \in \Omega^D\big)\mathbb{P}_\Omega\big(\omega \in \Omega^D\big) \\
&+ \varepsilon\mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \notin \Omega^D\big)\mathbb{P}_\Omega\big(\omega \notin \Omega^D\big) \\
&\geq \varepsilon\mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \in \Omega^D\big)\mathbb{P}_\Omega\big(\omega \in \Omega^D\big) \\
&\geq \varepsilon\delta\mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \in \Omega^D\big)
\end{aligned} \tag{45}$$

where: the two equalities hold by the law of total probability; the second inequality holds by (44); the fourth inequality holds by (43).

Let $\overline{C}(q^D)$ be the set of all search costs for which an isolated agent does not sample the second action when the first action he samples has quality q^D. For all $\omega \in \Omega^D$, any agent k with search cost $c_k \in \overline{C}(q^D)$ does not sample the second action either independently of his neighborhood realization B_k, the choices of his neighbors, and the quality of the first action sampled (see Section 2.1). Then,

$$\begin{aligned}
&\mathbb{P}_\sigma\big(s_k^2 = d\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \in \Omega^D\big) \\
&\geq \mathbb{P}_\sigma\big(c_k \in \overline{C}(q^D)\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \in \Omega^D\big).
\end{aligned} \tag{46}$$

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

$$\begin{aligned}
&\mathbb{P}_\sigma\big(c_k \in \overline{C}(q^D)\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \in \Omega^D\big) \\
&= \mathbb{P}_\sigma\big(c_k \in \overline{C}(q^D)\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K\big).
\end{aligned} \tag{47}$$

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

$$\mathbb{P}_\sigma\big(c_k \in \overline{C}(q^D)\ \forall k \in \widehat{B}(n) \cup \{n\} \,\big|\, q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K\big)$$

$$\geq \mathbb{P}_\sigma\big(c_1 \in \overline{C}(q^D)\big)^K \qquad (48)$$
$$> 0,$$

where the strict inequality holds because $\mathbb{P}_\sigma(c_1 \in \overline{C}(q^D)) > 0$ by the first condition in Assumption 1. Together, (46), (47), and (48) yield that

$$\mathbb{P}_\sigma\big(s_k^2 = d \;\forall k \in \widehat{B}(n) \cup \{n\} \mid q_{\min} < q(\underline{c}), |\widehat{B}(n)| < K, \omega \in \Omega^D\big) > 0. \qquad (49)$$

As $\varepsilon, \delta > 0$, from (45) and (49) we conclude that, for all agents n in the subsequence $\mathcal{N}$,

$$\mathbb{P}_\sigma\big(s_k^2 = d \;\forall k \in \widehat{B}(n) \cup \{n\} \mid q_{\min} < q(\underline{c})\big) > 0.$$

Then, the desired result follows from Lemma 13. ■

Proof of Theorem 3, part (ii)–(a). I prove the result by showing that the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions remains bounded away from 0 as $n \to \infty$ whenever 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 $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to 0 as $n \to \infty$ for all $q < q(\underline{c})$ that the first action sampled by agent 1 can take. That is, $\lim_{n\to\infty} P_1(q)(\prod_{i=2}^n (1 - F_C(t_i(q)))) = 0$ (see Lemma 12 and its proof). Hence, the expected gain from the second search for agent $n+1$, given by $P_1(\hat{q})(\prod_{i=2}^n (1-F_C(t_i(\hat{q}))))t^{\emptyset}(\hat{q})$ (see Lemma 12), where $\hat{q}$ is the quality of the action taken by agent n, also converges to 0 as $n \to \infty$ for all $\hat{q} < q(\underline{c})$. Then, there exists an agent $N_{\hat{q}} + 1$ for which the expected gain from the second search falls below $\underline{c}$.

By Assumption 1, 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 $1 - F_C(t^{\emptyset}(\tilde{q})) > 0$. Hence, with positive probability, agent 1 samples first a suboptimal action with quality, say, $\overline{q}$, and discontinues search. Now suppose the first $N_{\overline{q}}$ agents all have costs larger than $t^{\emptyset}(\overline{q})$, which occurs with positive probability. By Lemma 12, 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 discontinues search either because his expected gain from the second search is smaller than $\underline{c}$. Since the expected 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 0. ■

Proof of Theorem 3, part (ii)–(b). I prove the result by showing that the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions remains bounded away from 0 as $n \to \infty$.

Pick a 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 k. Such a sequence must exist with probability 1; 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.

By way of contradiction, suppose that the probability of no agent in $\widehat{B}(\pi_k) \cup \{\pi_k\}$ sampling both actions converges to 0 as $k \to \infty$ for all 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 probability defined by (4). It follows that the expected 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 0 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 gain from the second search falls below $\underline{c}$, and remains below this threshold for the agents in the sequence moving after $\pi_{K_{\hat{q}}} + 1$.

By Assumption 1, 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 $1 - F_C(t^{\emptyset}(\tilde{q})) > 0$. Therefore, with positive probability, agent π_1 samples first a suboptimal action with quality, say, $\overline{q}$, and discontinues search. 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 discontinues search either because his expected gain from the second search is smaller than $\underline{c}$. Since the expected 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 0. ■

G Proof of Proposition 2

To proof is based on a technique developed by Lobel et al. (2009), which consists in approximating a lower bound on the rate of convergence with an ordinary differential equation. Proposition 2 assumes polynomial shape, but the results extend immediately to search costs distributions with polynomial tail (i.e. such that (9) holds only for $c \in (0, \varepsilon)$ for some $\varepsilon \in (0, t^{\emptyset}(\underline{q})/2)$).

Proof of part (a). It is enough to construct a function $\tilde{\phi} \colon \mathbb{R}_+ \to \mathbb{R}$ such that, for all 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, $\gamma_n = n - 1$. Under the assumptions of the proposition, by Lemmas 7 and 12,

$$\begin{aligned}\mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x \in X} q_x\right) &\geq \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x \in X} q_x\right) \\ &+ \left(1 - \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x \in X} q_x\right)\right)^2 F_C\left(\left(1 - \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x \in X} q_x\right)\right) t^{\emptyset}(q_{\max})\right).\end{aligned} \tag{50}$$

Since the search costs distribution has polynomial shape, from (50) we have

$$\begin{aligned}\mathbb{P}_\sigma\left(s^1_{n+1} \in \arg\max_{x \in X} q_x\right) &\geq \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x \in X} q_x\right) \\ &\qquad + L t^{\emptyset}(q_{\max})^K \left(1 - \mathbb{P}_\sigma\left(s^1_n \in \arg\max_{x \in X} q_x\right)\right)^{K+2}.\end{aligned} \tag{51}$$

Now I build on Lobel et al. (2009) (see their proof of Proposition 2) to construct the function $\tilde{\phi}$. Adapting their procedure to my setup gives that the function $\tilde{\phi}$ is

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

where $\overline{K}$ is some constant of integration (in the construction, $\tilde{\phi}$ is found as the solution to an ordinary differential equation). Note: to apply a construction in the spirit of Lobel et al. (2009), the right-hand side of (51) must be increasing in $\mathbb{P}_\sigma(s^1_n \in \arg\max_{x \in X} q_x)$. This is so under the assumption $0 < L < 2^{K+1}/(K+2) t^{\emptyset}(\underline{q})^K$, which is maintained in the proposition. The same remark applies to the right-hand side of (53) in the proof of part (b).

Proof of part (b). It is enough to construct a function $\tilde{\phi}\colon \mathbb{R}_+ \to \mathbb{R}$ such that, for all n,

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

Under the assumptions of the proposition,

$$\begin{aligned}\mathbb{P}_\sigma\Big(s_{n+1}^1 \in \arg\max_{x\in X} q_x\Big) &= \frac{1}{n}\sum_{b=1}^{n}\mathbb{P}_\sigma\Big(s_{n+1}^1 \in \arg\max_{x\in X} q_x \mid B(n+1)=\{b\}\Big)\\ &= \frac{1}{n}\Big[\mathbb{P}_\sigma\Big(s_{n+1}^1 \in \arg\max_{x\in X} q_x \mid B(n+1)=\{n\}\Big) + (n-1)\mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x\Big)\Big]\end{aligned} \tag{52}$$

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

$$\begin{aligned}\mathbb{P}_\sigma\Big(s_{n+1}^1 \in \arg\max_{x\in X} q_x\Big) &\geq \mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x\Big)\\ &\quad + \frac{Lt^{\emptyset}(q_{\max})^K}{n}\Big(1-\mathbb{P}_\sigma\Big(s_n^1 \in \arg\max_{x\in X} q_x\Big)\Big)^{K+2}.\end{aligned} \tag{53}$$

I now build on Lobel et al. (2009) (see their proof of Proposition 3) to construct the function $\tilde{\phi}$. Adapting their procedure to my setup gives that the function $\tilde{\phi}$ we are looking for is

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

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

H Proofs for Section 4.3

To begin, I set some useful notation. Fix a state process and a search technology. Let $\delta \in (0,1)$ be the discount factor and define $t_1(q) := t^{\emptyset}(q)$ for all $q \in Q$. Suppose agent 1 samples first action x with quality q_x, and let $q_{\neg x}$ be a random variable with probability measure $\mathbb{P}_Q$.

▷ The equilibrium expected discounted social utility normalized by $(1-\delta)$ (hereafter simply referred to as social utility) in the complete network, denoted by $U_\sigma^C(q_x;\delta)$, is

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

Here, the first term is the quality of the first action sampled and the second term is the expected 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 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 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 n. Therefore, all agents $m > n$ with $c_m \leq t_m(q_x)$ will sample action $\neg x$ again.

▷ The social utility when each agent observes only his most immediate predecessor, denoted by $U_\sigma^1(q_x;\delta)$, is

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

$U_\sigma^1(q_x;\delta)$ has the same interpretation as $U_\sigma^C(q_x;\delta)$, except for the cost of search when $q_{\neg x} > q_x$, which now contains an additional term. This is so because agents that observe only their most immediate predecessor fail to recognize actions that are revealed to be inferior by the time of their move. Hence, 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, the expected cost of this additional search is $\mathbb{E}_{\mathbb{P}_Q}[\mathbb{E}_{\mathbb{P}_C}[c \mid c \leq t_n(q_{\neg x})] F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x]$.

▷ Let $U_\sigma^{OIP}(q_x;\delta)$ be the social utility in some arbitrary OIP network. The next lemma is immediate from the discussion in Section 4.3.

Lemma 14. *For all $q_x \in Q$ and $\delta \in (0,1)$, we have $U_\sigma^C(q_x;\delta) \geq U_\sigma^{OIP}(q_x;\delta) \geq U_\sigma^1(q_x;\delta)$.*

▷ Finally, let $U^{DM}(q_x;\delta)$ denote the social utility that is implemented by the single decision maker in any OIP network after sampling action x with quality q_x at the first search in period 1. 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, their derivation applies unchanged to my setting.

H.1 Proof of Proposition 3

The difference in average social utilities, $U_\sigma^C(q_x;\delta) - U_\sigma^1(q_x;\delta)$, is

$$(1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x) \sum_{n=1}^{\infty} \delta^n \mathbb{E}_{\mathbb{P}_Q}\left[\mathbb{E}_{\mathbb{P}_C}\left[c \mid c \leq t_n(q_{\neg x})\right] F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\right] \\ \cdot \left(1 - \prod_{i=1}^{n-1}\left(1 - F_C(t_i(q_x))\right)\right). \tag{54}$$

As (54) is positive for all $\delta \in (0,1)$, that $U_\sigma^C(q_x;\delta) > U_\sigma^1(q_x;\delta)$ for all $\delta \in (0,1)$ follows.

To show that $\lim_{\delta\to 1}[U_\sigma^C(q_x;\delta) - U_\sigma^1] = 0$, we need to show that (54) converges to 0 as $\delta \to 1$. To do so, it is enough to argue that

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

is finite. Notice that

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

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. ■

H.2 Proof of Proposition 4

First, suppose $\underline{c} = 0$. By Proposition 3, Lemma 14, and the sandwich theorem for limits of functions, $\lim_{\delta \to 1} U_\sigma^{OIP}(q_x; \delta) = \lim_{\delta \to 1} U_\sigma^{C}(q_x; \delta)$. By Proposition 3 in MFP, $\lim_{\delta \to 1} U_\sigma^{C}(q_x; \delta) = \lim_{\delta \to 1} U^{DM}(q_x; \delta)$. Thus, that $\lim_{\delta \to 1} U_\sigma^{OIP}(q_x; \delta) = \lim_{\delta \to 1} U^{DM}(q_x; \delta)$ follows by the uniqueness of the limit of a function.

Now suppose that $\lim_{\delta \to 1} U_\sigma^{OIP}(q_x; \delta) = \lim_{\delta \to 1} U^{DM}(q_x; \delta)$. Since the complete network is an OIP network, we have $\lim_{\delta \to 1} U_\sigma^{C}(q_x; \delta) = \lim_{\delta \to 1} U^{DM}(q_x; \delta)$. Thus, that $\underline{c} = 0$ follows from Proposition 3 in MFP. ■

I More than Two Actions

I.1 Collective Search Environment

Suppose the set of available actions is $X \coloneqq \{0, 1, \ldots, N\}$, where $N > 1$; that is, more than two actions are available. Qualities $q_0, q_1, \ldots, q_N$ are i.i.d. draws from a probability measure $\mathbb{P}_Q$ over $Q \subseteq \mathbb{R}_+$ and the state of the world is $\omega \coloneqq (q_0, q_1, \ldots, q_N)$. The model remains otherwise unchanged, with the obvious adjustments to account for more than two available actions.

I.1.1 Equilibrium Strategies

Choice Stage. If an agent sampled one action, he takes that action; if he sampled more than one action, he takes the best sampled action, breaking ties uniformly at random.

First Search Stage. Fix n and σ_{-n}. For all $x, x' \in X$, let $E_n^{x,x'}$ be the event that occurs when none of the agents in n's subnetwork relative to action x has sampled action x'. Given σ_{-n} and I_n^1, agent n computes the conditional probabilities $P_n(x, x') \coloneqq \mathbb{P}_{\sigma_{-n}}(E_n^{x,x'} \mid I_n^1)$ for all $x, x' \in X$. If $P_n(x, x') < P_n(x', x)$, agent n's belief about the quality of action x strictly first-order stochastically dominates his belief about the quality of action x'. Thus, by comparing $P_n(x,x')$ with $P_n(x',x)$ for all $x,x' \in X$, agent n is able to rank his beliefs about the quality of all actions in terms of first order stochastic dominance. In particular, define the linear order $\succsim$ on X as follows: $x \succsim x' \iff P_n(x,x') \leq P_n(x',x)$. According to Weitzman (1979)'s optimal search rule, at the first search stage, agent n: samples the $\succsim$-maximal element of X if there is only one such element; samples uniformly at random one of the $\succsim$-maximal elements of X if there are multiple such elements.

***k*-th Search Stage, $k \geq 2$.** Let I_n^k be agent n's information set after sampling $k-1$ actions. Moreover, let S_n^k be the set of the first $k-1$ actions sampled by agent n. Finally, let $x^* \in \arg\max_{x \in S_n^k} q_x$. Agent n will only sample an additional action if his search cost c_n is no larger than the expected gain from the second search.

- If $B_n = \emptyset$, such gain is $\mathbb{E}_{\mathbb{P}_Q}[\max\{q - q_{x^*}, 0\}]$. If $c_n \leq \mathbb{E}_{\mathbb{P}_Q}[\max\{q - q_{x^*}, 0\}]$, agent n samples uniformly at random one of the actions in $X \setminus S_n^k$.
- If $B_n \neq \emptyset$, for any $x' \in X \setminus S_n^k$, let $E_n^{x',S_n^k} := \cup_{x \in S_n^k} E_n^{x,x'}$. Agent n benefits from sampling action $x' \in X \setminus S_n^k$ only if event E_n^{x',S_n^k} has occurred. Thus, agent n must compute the conditional probabilities $\mathbb{P}_{\sigma_{-n}}(E_n^{x',S_n^k} \mid I_n^k)$ for all $x' \in X \setminus S_n^k$. If $c_n \leq \max_{x' \in X \setminus S_n^k} \mathbb{P}_{\sigma_{-n}}(E_n^{x',S_n^k} \mid I_n^k)\mathbb{E}_{\mathbb{P}_Q}[\max\{q - q_{x^*}, 0\}]$, agent n samples uniformly at random one of the actions in $\arg\max_{x' \in X \setminus S_n^k} \mathbb{P}_{\sigma_{-n}}(E_n^{x',S_n^k} \mid I_n^k)$.

I.1.2 Metrics of Social Learning

The notion of complete learning remains unchanged. Maximal learning occurs if

$$\lim_{n \to \infty} \mathbb{P}_\sigma\Big(q_{a_n} \geq \min\left\{q(\underline{c}), \max\{q_0, q_1, \ldots, q_N\}\right\}\Big) = 1.$$

That is, maximal learning occurs if, as $n \to \infty$, agent n takes with probability 1: (i) any action if all actions have the same quality; (ii) the best action if at most one action has quality larger than or equal to $q(\underline{c})$; (iii) any of the actions with quality larger than or equal to $q(\underline{c})$ if there are multiple such actions.[19]

I.2 Long-Run Learning

I.2.1 Complete Learning and the Improvement Principle

The key intuition behind the IP I develop for the main model remains unchanged with more than two actions. Consider an agent, say n, and his chosen neighbor, say $b < n$. Unless b samples the best action with probability 1 at the first search, b's expected gain from conducting at least one additional search is positive. Therefore, if search costs are not bounded away from 0, b samples at least two actions with positive probability. As b takes the best action among those he samples, with positive probability the action b takes is of better quality than the one he samples first. If n begins searching from the action taken by b, this results in a strict improvement in the probability of sampling the best action at the first search that n has over b, unless b already does so with probability 1. If long information paths occur almost surely in the network topology, such improvements last until agents sample the best action with probability 1 at the first search.

[19] A searcher discontinues search after sampling an action of quality larger than or equal to $q(\underline{c})$. Thus, even a searcher is unable to distinguish among actions with quality larger than or equal to $q(\underline{c})$.

If agents have reasonably accurate information about the network (i.e. information paths are identifiable), agents single out the correct neighbor to rely on. The next result follows.

Result 1. *Theorem 1 remains unchanged with more than two actions.*

I.2.2 Maximal Learning and the Large-Sample Principle

The key intuition behind the LSP also remains unchanged with more than two actions. In particular, consider the set of agents S in Theorem 2. Let α^x be the probability with which an isolated agent takes action x. It is easy to see that: (i) $\alpha^x > \alpha^{x'}$ for all $x, x' \in X$ with $q_{x'} < \min\{q(\underline{c}), q_x\}$; (ii) $\alpha^x = \alpha^{x'}$ for all $x, x' \in X$ with $\min\{q_x, q_{x'}\} \geq q(\underline{c})$ or $q_x = q_{x'}$. Non-isolated agents in S finds it optimal to sample at the first search one of the actions with the largest share in their neighborhood. Thus, by the properties on the network topology, we have

$$\lim_{n\to\infty} \mathbb{P}_\sigma\left(q_{s_n^1} \geq \min\left\{q(\underline{c}), \max\{q_0, q_1, \ldots, q_N\}\right\} \mid B(n) = B_n^\emptyset\right) = 1.$$

Since $q_{a_n} \geq q_{s_n^1}$ and the share non-isolated agents in S converges to 1 as $n \to \infty$, maximal learning occurs within S. An analogous argument as that for Theorem 2 extends the maximal learning result to agents in $\mathbb{N} \setminus S$. The next result follows.

Result 2. *Theorem 2 remains unchanged with more than two actions.*

I.2.3 Failure of Maximal Learning

Identifying the best among more than two actions is no easier for the agents in the society than identifying the best between two actions. Thus, the next result immediately follows.

Result 3. *Theorem 3 remains unchanged with more than two actions.*

I.3 Rate of Convergence, Welfare, and Efficiency

I.3.1 Rate of Convergence

Given a network topology, when more than two actions are available, convergence to the best action occurs more slowly than when only two actions are available. Thus, the upper bounds on the rate of convergence in Proposition 2 (need) no longer apply with more than two actions. However, the insight that the density of indirect connections affects convergence rates remains valid. In particular, also with more than two actions, learning is faster in OIP networks than under uniform random sampling of one past agent. This is so because the cardinality of agents' subnetworks grows at a faster rate in OIP networks, and so does the probability that in the subnetwork all actions have been sampled at least once and the inferior actions have been discarded.

I.3.2 Equilibrium Welfare and Efficiency in OIP Networks

Proposition 3 remains unchanged with more than two actions. When more than two actions are available, reducing network transparency in OIP networks exacerbates the inefficient duplication of costly search. This is so because there are more opportunities to engage in overeager search. The intuition is the following. When more than two actions are available, more than one action can be revealed to be inferior by time n. However, if agent n observes only agent $n-1$ (i.e. in the OIP network in which agents only observe their most immediate predecessor), such actions are not revealed to be inferior to agent n (whereas they would be so in the complete network); thus, agent n may sample at a cost several of such actions again to learn what he would simply learn by observing all his predecessor's choices. An immediate consequence is that simple policy interventions, such as letting agents observe the aggregate history of past choices, play an even more important role in reducing efficiencies when more than two actions are available.

Whether Propositions 1 and 4 holds true with more than two actions is, in contrast, unclear. This is so because establishing these results requires closed-form expressions for equilibrium search policies and for the solution to the single decision maker's problem. With more than two actions, such closed-form expressions are not possible.

References

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

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

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.

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

Bardhi, Arjada (2020), "Attributes: Selective Learning and Influence." *Working Paper*.

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.

Board, Simon and Moritz Meyer-ter-Vehn (2020), "Learning Dynamics in Social Networks." *Working Paper*.

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.

Che, Yeon-Koo and Konrad Mierendorff (2019), "Optimal Dynamic Allocation of Attention." *American Economic Review*, 109, 2993–3029.

Dasaratha, Krishna and Kevin He (2020), "Aggregative Efficiency of Bayesian Learning in Networks." *Working Paper*.

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

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki (2018), "Speed, Accuracy, and the Optimal Timing of Choices." *American Economic Review*, 108, 3651–3684.

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

Garcia, Daniel and Sandro Shelegia (2018), "Consumer Search with Observational Learning." *The RAND Journal of Economics*, 49, 224–253.

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.

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

Harel, Matan, Elchanan Mossel, Philipp Strack, and Omer Tamuz (2020), "Rational Groupthink." *The Quarterly Journal of Economics (forthcoming)*.

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

Hörner, Johannes and Andrzej Skrzypacz (2017), "Learning, Experimentation, and Information Design." Bo Honoré, Ariel Pakes, Monika Piazzesi, and Larry Samuelson eds., *Advances in Economics and Econometrics: Eleventh World Congress (Econometric Society Monographs)*. Cambridge University Press.

Ke, Tony T. and J. Miguel Villas-Boas (2019), "Optimal Learning Before Choice." *Journal of Economic Theory*, 180, 383–437.

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.

Liang, Annie and Xiaosheng Mu (2020), "Complementary Information and Learning Traps." *The Quarterly Journal of Economics*, 135, 389–448.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis (2019), "Optimal and Myopic Information Acquisition." *Working Paper*.

——— (2020), "Dynamically Aggregating Diverse Information." *Working Paper*.

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.

Lomys, Niccolò (2020), "Social Search." *Working Paper*.

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.*

Mayskaya, Tatiana (2020), "Dynamic Choice of Information Sources." *Working Paper.*

Monzón, Ignacio and Michael Rapp (2014), "Observational Learning with Position Uncertainty." *Journal of Economic Theory*, 154, 375–402.

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

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

Nei, Stephen (2019), "Herding with Costs to Observation." *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.*

Qiu, Liangfei, Zhan (Michael) Shi, and Andrew B. Whinston (2018), "Learning from Your Friends' Check-Ins: An Empirical Study of Location-Based Social Networks." *Information Systems Research*, 29, 1044–1061.

Rosenberg, Dinah and Nicolas Vieille (2019), "On the Efficiency of Social Learning." *Econometrica*, 87, 2141–2168.

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

——— (2020), "Innovation Adoption and Collective Experimentation." *Games and Economic Behavior*, 120, 121–131.

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 (2020), "Informational Herding, Optimal Experimentation, and Contrarianism." *Working Paper.*

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

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.

Zhong, Weijie (2019), "Optimal Dynamic Information Acquisition." *Working Paper.*