CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 688

Collective Search in Networks

Niccolò Lomys

October 2023

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

Centre for Studies in Economics and Finance

WORKING PAPER NO. 688

Collective Search in Networks

Niccolò Lomys*

Abstract

I study the dynamics of collective search in networks. Bayesian agents act in sequence, observe the choices of their connections, and privately acquire information about the qualities of different actions via sequential search. If search costs are not bounded away from zero, maximal learning occurs in sufficiently connected networks where individual neighborhood realizations weakly distort agents' beliefs about the realized network. If search costs are bounded away from zero, maximal learning is possible in several stochastic networks, including almost-complete networks, but generally fails otherwise. When agents observe random numbers of immediate predecessors, the learning rate, 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.

JEL Classification: C7; D6; D8.

Keywords: Networks; Bayesian Learning; Search; Speed and Efficiency of Social Learning.

Acknowledgements: I thank various seminars, workshops, and conference audiences, and, in particular, Ali Aouad, Paolo Conteduca, Daniel Garrett, Amir Habibi, Takakazu Honryo, Johannes Hörner, Nicolas Lambert, Raphaël Levy, Lorenzo Magnolfi, Suraj Malladi, Chiara Margaria, Matthew Mitchell, Pauli Murto, Volker Nocke, Alexei Parakhonyak, Martin Peitz, Evan Sadler, Konrad Stahl, Bruno Strulovici, Emanuele Tarantino, Thomas Tröger, Takuro Yamashita, and Leeat Yariv for their helpful comments. This work was done in part while visiting the Simons Institute for the Theory of Computing at UC Berkeley (program on Data-Driven Decision Processes). I acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreements No°714147 and No°714693).

* CSEF and Università degli Studi di Napoli Federico II. E-mail: niccolomys@gmail.com.

1 Introduction

Search models provide a classical framework to study information acquisition and choice problems. When searching for the best option, agents seldom act in isolation. Social information, i.e., the choices and experiences of other agents, is readily available via direct observation and online social networks. Despite the abundant theoretical and empirical evidence in the social learning literature demonstrating how heavily social information shapes individual behavior (for comprehensive reviews, see Mobius and Rosenblat, 2014; Golub and Sadler, 2016; Bikhchandani, Hirshleifer, Tamuz, and Welch, 2022), search models typically posit that agents search in isolation, ignoring the information contained in the choices of their connections.

However, the interplay between learning from others and the incentives for individual search raises new questions. Does social information mitigate or even eliminate search frictions? That is, can societies learn to perform as well as an agent with access to the best search technology in the environment, thanks to the exploitation of social information? Conversely, could excessive exploitation of social information reduce individual incentives for search, effectively exacerbating search frictions? In this paper, I address these questions by studying the dynamics of collective search in networks and show that the answers to them depend significantly on the properties of the search technology *and* on those of the network structure.

Formally, I embed the standard search model of Weitzman (1979) into an observational learning setting in general networks à la Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) and Lobel and Sadler (2015). A countably infinite number of Bayesian agents act in sequence, each choosing between two actions whose qualities are i.i.d. draws about which agents are initially uninformed. First, each agent observes a subset of earlier choices, the agent's neighborhood. Neighborhoods are drawn from a joint distribution, the *network topology*. Next, the agent engages in *costly sequential search*. Searching for an action reveals its quality. After the first free search, the agent decides to sample the second action at a cost (i.i.d. across agents) or not. Finally, the agent takes the best between the sampled actions. Individual neighborhoods, search costs, and sampling decisions are private information, i.e., remain unobserved to other agents. The network topology shapes how effectively agents can learn from social information (exploitation), while the *search technology* shapes their ability to acquire private information (exploration). *Maximal learning* occurs if, in the long run, the probability that agents take the best action converges to that with which a *searcher*—an agent with the lowest possible search cost and no social information—does so.

I provide two main sets of results. First, I identify sufficient and necessary conditions on network topologies and search technologies for maximal learning. Second, by allowing for general network topologies, I obtain new insights into how the speed and efficiency of collective search depend on the network structure. In the complete network, my model reduces to that of Mueller-Frank and Pai (2016) (MFP). However, we study different long-run learning metrics. In particular, MFP focus on the more demanding benchmark of *complete learning*, which occurs if, in the long run, the probability that agents take the best action converges to 1. Given a search technology, I argue that maximal learning is the best achievable long-run outcome, from which my focus on this less stringent requirement follows. Moreover, as an immediate corollary of my results, I show that no side of MFP's equivalence that complete learning occurs if and only if search costs are not bounded away from 0 needs to hold beyond the complete network.

I characterize equilibrium search behavior by connecting an agent's optimal search policy to the probability that some agents in his subnetwork (i.e., the agents directly or indirectly linked to him) sampled both actions. Since by sampling both actions agents can 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 study of maximal learning possible.

Theorem 1 shows that if search costs are not bounded away from zero (i.e., arbitrarily low search costs have positive probability), maximal learning occurs if the network topology is sufficiently connected and neighborhood realizations weakly distort agents' beliefs about the realized network—the complete network being the simplest in this class.

I identify sufficient conditions for maximal learning by developing an *improvement principle* (IP). According to the IP, improvements upon imitation—a heuristic—suffice for maximal learning. If agents' beliefs about the realized network conditional on their neighborhood are not too distorted—in a specific sense—compared to the actual network topology, each agent can pick a reliable neighbor to observe and determine the search policy regardless of what others have done. If search costs are not bounded away from zero, the agent samples both actions and takes the best one with a higher probability than his chosen neighbor unless the latter already does so. If, in addition, information paths are long enough, such improvements last until maximal learning occurs.

Theorem 2 characterizes network topologies where maximal learning occurs independently of whether search costs are bounded away from zero. In these networks, there are two sets of agents. The first set—the core—consists of infinitely many isolated agents (i.e., agents with no neighbors) and infinitely many agents that observe all and only their isolated predecessors. The share of the former vanishes in the long run. The second set is the rest of the network. Neighborhood realizations weakly distort agents' beliefs, and agents are sufficiently connected to recent predecessors in the core. The core may form the entire network but also consist of an arbitrarily small portion of it—such as almost-complete networks, where the probability that late-moving agents observe all predecessors converges to one.

The intuition is the following. As the choices of the infinitely many isolated agents are independent, the share of observed choices suffices to late-moving non-isolated agents in the core to sample first the action a searcher would take. Since isolated agents vanish, maximal learning occurs within the core. Agents in the core learn via a *large-sample principle* (LSP), i.e., by observing and aggregating the information in large samples of individual choices. Once it occurs within the core, maximal learning extends to the rest of the network via imitation, as late-moving agents are likely to observe some recent non-isolated agents in the core and can pick the correct neighbor to rely on.

Theorem 3 characterizes necessary conditions 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 the complete network, if agents have at most one neighbor, and if agents observe (possibly correlated) random numbers of immediate predecessors (OIP networks).[1]

The IP and the LSP have limitations. 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 in all network topologies. 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 agents with a specific structure, such as the core in Theorem 2.

[1]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 where each agent n observes his $n-1$ immediate predecessors.

The second part of the paper studies the speed and efficiency of learning as a function of the network structure. First, the convergence rate to the best action, the probability of wrong herds, and long-run welfare and efficiency (i.e., when discounting future payoffs with factor $\delta \to 1$) are the same in all OIP networks. Hence, in such networks, these equilibrium outcomes are independent of network transparency, the density of connections, and their correlation. This striking result holds because, in all OIP networks, agents' subnetworks consist of all their predecessors. As agents' search policy depends on the probability that some agent in their subnetwork sampled both actions, agents' performance in all OIP networks must be the same.

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

Third, reducing network transparency leads to inefficient duplication of costly searches as agents who do not observe all prior choices fail to recognize actions revealed as inferior by some of their predecessors' choices. The resulting welfare loss remains sizable for all $\delta < 1$.

Finally, the density of indirect connections affects convergence rates. Convergence to the best action is faster than polynomial in OIP networks but only faster than logarithmic under uniform random sampling of one past agent. Learning is slower under uniform random sampling because the cardinality of agents' subnetworks grows more slowly than in OIP networks, and so does the probability that some agent in the subnetworks samples both actions.

Related Literature. In the sequential social learning model (SSLM), agents wish to match their actions with an unknown state of nature, observe a free private signal—informative about all options' relative quality—and the choices of all (Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992; Smith and Sørensen, 2000), or subsets of (Acemoglu and Ozdaglar, 2011; Lobel and Sadler, 2015), predecessors before making their choice. My setting differs from the SSLM in three key aspects. First, private information is generated in equilibrium and not exogenously available. Second, sampling an action reveals its quality only and perfectly, whereas exogenous signals are imperfectly informative about all actions' relative quality. Third, the inferential challenge differs: agents maximize the value of a sequential information acquisition program and not an ex-ante expected utility.

Like most earlier work on observational learning implicitly or explicitly does, I organize positive and negative long-run learning results around an IP and an LSP.[2] Because of the substantial differences in informational environments and learning metrics outlined above, the working and applicability of these principles differ from those in the SSLM. First, the IP only holds if agents have access to arbitrarily low search costs in my mode; in contrast, it does independent of whether private signals are arbitrarily informative in the SSLM. Second, the scope of large-sample and martingale convergence arguments in my model is more limited than in the SSLM (and most other social learning models).

Burguet and Vives (2000), Hendricks, Sorensen, and Wiseman (2012), Ali (2018), and

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

Bobkova and Mass (2022) study information acquisition in a social learning game over the complete network. None of these papers focuses on the network structure. Because of technical challenges, previous work on social learning in networks with costly information acquisition (see, e.g., Perego and Yuksel, 2016; Salish, 2017; Sadler, 2020; Board and Meyer-ter Vehn, 2021, 2023) mostly focuses on particular network structures or properties, or posits boundedly rational decision rules. Yet, it acknowledges the importance of a general analysis within the Bayesian benchmark (see, e.g., Sadler, 2014; Golub and Sadler, 2016). Assuming agents move sequentially, I accommodate Bayes rationality and general network topologies.

Recent papers (Lobel, Acemoglu, Dahleh, and Ozdaglar, 2009; Monzón and Rapp, 2014; Hann-Caruthers, Martynov, and Tamuz, 2018; Harel, Mossel, Strack, and Tamuz, 2021; Dasaratha and He, 2023) study the speed and efficiency of social learning with exogenous private information. Because of technical challenges, existing results focus on simple network structures. My paper is the first to study how the speed and efficiency of learning vary with the network with endogenous private information. Despite the complications introduced by costly information acquisition, a rich and tractable analysis emerges in my setting.

Rosenberg and Vieille (2019) define learning to be efficient in the SSLM if the expected number of incorrect choices under a 0-1 loss function is finite. Learning is efficient in the complete network if and only if it is so when agents only observe the immediate predecessor. In my setting, the irrelevance of how far in the past agents observe holds more generally. First, not only does it hold for long-run welfare, but also for the probability of wrong herds and the convergence rate. Second, it neither depends on the number of immediate predecessors that agents observe nor on the density and correlation among connections.

Road Map. Section 2 introduces the model. Section 3 provides positive and negative learning results. Section 4 studies the speed and efficiency of learning. Section 5 concludes and discusses extensions. 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 an action in X, $\neg x$ the action in X other than x, and a_n the action agent n takes.

State Process. Actions' qualities, denoted by 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 $\omega := (q_0, q_1) \in \Omega := Q \times Q$ has probability measure $\mathbb{P}_\Omega := \mathbb{P}_Q \times \mathbb{P}_Q$ and is drawn once and for all at time 0.

To take the action with the highest quality, each agent n first observes a subset of past choices and then privately engages in costly sequential search.

Network Topology. Agent n observes the choices of a subset $B(n)$ of past agents, referred to as agent 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 with probability measure $\mathbb{Q}$ on $\mathbb{B} := \prod_{n \in \mathbb{N}} 2^{\mathbb{N}_n}$. Realizations of $B(n)$ are agent n's private information. If $n' \in B(n)$, then n observes $a_{n'}$ and knows the identity of n'. Agent n is isolated if $B(n) = \emptyset$. Neighborhoods are independent of the state process and search costs (introduced below).

The network model follows Acemoglu et al. (2011) and Lobel and Sadler (2015); it allows for correlated neighborhoods, independent neighborhoods, and deterministic networks. The framework nests most networks observed in the data and studied in the literature: complete network, observation of the most recent predecessors, random sampling from the past, networks with influential agents, preferential attachment, small-world networks, etc.

Search Technology. 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 to perfectly learn its quality $q_{s_n^1}$. After observing $q_{s_n^1}$, agent n decides whether to sample the remaining action, $s_n^2 = \neg s_n^1$, to perfectly learn $q_{\neg s_n^1}$, or to discontinue searching, $s_n^2 = d$. Let S_n denote the set of actions agent n samples. After the sampling has stopped, the agent takes an action $a_n \in S_n$. When each agent observes all past choices, the model reduces to that of Mueller-Frank and Pai (2016).

The first search is free. The second search costs $c_n \in C \subseteq \mathbb{R}_+$. Search costs c_n are i.i.d. draws across agents from a probability measure $\mathbb{P}_C$ over C, with CDF F_C, and are independent of the network topology and the state process. 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.

Definition 1. *Let* $\underline{c} \coloneqq \min \operatorname{supp}(\mathbb{P}_C)$. *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: $q_{a_n} - c_n(|S_n| - 1)$.

Information. Agent n has three information sets: $I_n^1 \coloneqq \{c_n, B(n), a_k \; \forall \, k \in B(n)\}$ corresponds to n's information prior to sampling any action; $I_n^2 \coloneqq I^1(n) \cup \{(s_n^1, q_{s_n^1})\}$ corresponds to n's information after sampling the first action; $I_n^a \coloneqq \{c_n, B(n), a_k \; \forall \, k \in B(n), \{(x, q_x) : x \in S_n\}\}$ corresponds to n's information once search ends. The classes of all possible information sets of agent n are denoted by $\mathcal{I}_n^1$, $\mathcal{I}_n^2$, and $\mathcal{I}_n^a$. The game is common knowledge.

Strategies. A strategy for agent n is a triple of mappings $\sigma_n \coloneqq (\sigma_n^1, \sigma_n^2, \sigma_n^a)$, 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 $\sigma_n^a \colon \mathcal{I}_n^a \to \Delta(S_n)$. The sequence of decisions $((s_n^1, s_n^2, a_n))_{n \in \mathbb{N}}$ is a stochastic process whose probability measure—generated by the state process, the network topology, the search technology, and agents' strategy profiles $\sigma \coloneqq (\sigma_n)_{n \in \mathbb{N}}$—I denote by $\mathbb{P}_\sigma$.

Equilibrium Notion. The solution concept is *perfect Bayesian equilibrium*, hereafter referred to as equilibrium. A strategy profile σ 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 $\sigma_{-n} \coloneqq (\sigma_1, \ldots, \sigma_{n-1}, \sigma_{n+1}, \ldots)$.

Agents' decisions are discrete choice problems with a well-defined solution that only requires randomization in case of indifference. Given tie-breaking criteria, an inductive argument shows that the set of equilibria is nonempty. To ease exposition, I focus on the equilibrium where agents sample the second action in case of indifference and break other ties uniformly at random. This equilibrium selection does not affect 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 is in agent n's* subnetwork, *denoted by $\widehat{B}(n)$, if there is a sequence of agents, starting with m and terminating with n, with each element of the sequence in the neighborhood of the next. Agent m is in agent n's* subnetwork relative to action x, *denoted by $\widehat{B}(n,x)$, if $m \in \widehat{B}(n)$ and $a_m = x$.*

$\widehat{B}(n)$ consists of the agents that are, directly or indirectly, connected to agent n. $\widehat{B}(n,x)$ consists of the agents that are, directly or indirectly, connected to agent n and take action x.

Choice. Agent n takes the best between sampled actions, randomizing uniformly if indifferent.

First Search. Consider agent n and the events

$$E_n^x := \left\{ s_k^2 = d \; \forall \, k \in \widehat{B}(n,x) \right\} \qquad \text{for all } x \in X.$$

E_n^x occurs when none of the agents in agent n's subnetwork relative to action x samples both actions. For $x = 0, 1$, the complement of E_n^x, denoted by E_n^{xC}, occurs when at least one agent in n's subnetwork relative to action x samples both actions. For each action x:

1. If agent n knew that event E_n^x occurred, his posterior belief on the quality of action $\neg x$ is the same as the prior $\mathbb{P}_Q$.
2. If agent n knew that event E_n^{xC} occurred, his posterior belief on Ω is $\mathbb{P}_{\Omega | q_x \geq q_{\neg x}}$, as agents sampling both actions take the one with the highest quality.

Thus, agent n computes the conditional probabilities

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

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 $\neg x$ (see Appendix A for the formal argument). By Weitzman (1979)'s optimal search rule, as extended by Gergatsouli and Tzamos (2023) to correlated actions' qualities, agent n samples action x first (actions qualities need no longer be independent in the eyes of agent n, i.e., conditional on I_n^1). If $P_n(x) = P_n(\neg x)$, agent n's beliefs about the qualities of the two actions are identical, and n samples the first action uniformly at random.

Second Search. After sampling a first action of quality $q_{s_n^1}$, agent n samples the second action if and only if the expected gain from doing is no less than his search cost. If $B(n) = \emptyset$, the expected gain from the second search is

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

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, agent n computes the conditional probability

$$P_n\Big(q_{s_n^1}\Big) := \mathbb{P}_{\sigma_{-n}}\Big(E_n^{s_n^1} \mid I_n^2\Big). \tag{3}$$

Otherwise, at least one of those agents sampled action $\neg s_n^1$ but took action s_n^1, in which case s_n^1 must be (weakly) superior. Hence, n's expected gain from the second search is

$$t_n\Big(q_{s_n^1}\Big) := P_n\Big(q_{s_n^1}\Big)\, t^{\emptyset}\Big(q_{s_n^1}\Big). \tag{4}$$

2.2 Maximal Learning

An isolated agent having access to the smallest search costs the search technology allows for has (i) the best search opportunities and (ii) the strongest motivation to explore. To understand (ii), note that, by (2)–(4), $t^{\emptyset}(q_{s_n^1}) \geq t_n(q_{s_n^1})$; that is, given $q_{s_n^1}$, the expected gain from the second search when agent n is isolated is larger than when he is not. Such an agent, hereafter called a *searcher*, takes the best action with probability 1 if and only if

$$\omega \in \Omega(\underline{c}) := \Big\{\omega \in \Omega : F_C\Big(\max\Big\{t^{\emptyset}(q_0), t^{\emptyset}(q_1)\Big\}\Big) > 0 \text{ or } q_0 = q_1\Big\}.$$

The next assumption rules out uninteresting environments. If part (i) fails, no agent searches twice. If part (ii) fails, isolated agents always take the best action when $\omega \in \Omega(\underline{c})$, and non-isolated agents copy any neighbor. The assumption implies $\mathbb{P}_\Omega(\omega \in \Omega(\underline{c})) > 0$.

Assumption 1. *Let $Q(\underline{c}) := \{q \in Q : F_C(t^{\emptyset}(q)) > 0\}$. Assume that: (i) $\mathbb{P}_Q(Q(\underline{c})) > 0$; (ii) there is $\tilde{q} \in Q(\underline{c})$ such that $\mathbb{P}_Q(q > \tilde{q}) > 0$ and $F_C(t^{\emptyset}(\tilde{q})) < 1$, i.e., with positive probability, isolated agents discontinue searching after sampling an action of quality $\tilde{q}$ or higher.*

Maximal learning occurs if the probability that agents take the best action converges to 1 whenever a searcher does so. It is the best achievable long-run outcome in this environment.

Definition 3. Maximal learning *occurs if*

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

Remark 1. The probabilities $P_n(x)$ and $P_n(q_x)$ in definitions (1) and (3) suffice to describe agent n's equilibrium search policy and link agent n's search policy to the probability that he takes the best action, which is what matters for maximal learning, as

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x\right) \geq \mathbb{P}_{\sigma_{-n}}\Big(\Big\{s_k^2 = \neg s_k^1 \text{ for some } k \in \widehat{B}\Big(n, s_n^1\Big)\Big\}\Big) = 1 - \mathbb{P}_{\sigma_{-n}}\Big(E_n^{s_n^1}\Big).$$

The inequality holds as if some agent in $\widehat{B}(n, s_n^1)$ samples both actions and takes action s_n^1, then s_n^1 (and so a_n) must be superior. The equality holds as the events are complements.

3 Long-Run Learning

The following definitions recall some notions on network topologies introduced by Lobel and Sadler (2015), to which I refer for further discussion. Expanding subnetworks is a connectivity property requiring that the size of $\widehat{B}(n)$ grows without bound as n becomes large.

Definition 4. *A network topology has* expanding subnetworks *if, for all $K \in \mathbb{N}$,*

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

If this property fails, the network topology has non-expanding subnetworks*.*

A neighbor choice function is a particular agent's means of selecting a neighbor from any realization of his neighborhood. Given a network topology and a sequence of neighbor choice functions, one implicitly defines a new network topology, called the chosen neighbor topology, that includes only those links in the original network selected by the neighbor choice functions. In other words, the chosen neighbor topology is the network where agents discard all unselected neighbors.

Definition 5. *Let a network topology be given:*

(a) A function $\gamma_n : 2^{\mathbb{N}_n} \to \mathbb{N}_n \cup \{0\}$ is a neighbor choice function *for agent n if, for all $B(n) \in 2^{\mathbb{N}_n}$, we have $\gamma_n \in B(n)$ when $B(n) \neq \emptyset$, and $\gamma_n = 0$ otherwise. Agent γ_n is called agent n's* chosen neighbor*.*

(b) A chosen neighbor topology *consists of the links in the given network topology selected by a sequence of neighbor choice functions $\gamma := (\gamma_n)_{n\in\mathbb{N}}$.*

3.1 Maximal Learning via the Improvement Principle

If search costs are not bounded away from 0, maximal learning occurs if the network is sufficiently connected, and agents' beliefs about the network conditional on their neighborhood are not too distorted—in a specific sense—compared to the actual network topology.

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

(i) Search costs 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 has expanding subnetworks;

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

$$\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X}\, q_x \;\middle|\; \gamma_n\right) \geq \mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X}\, q_x\right) - \varepsilon. \quad (5)$$

To prove Theorem 1, I establish an auxiliary result: the *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. The IP works as follows. Each agent selects one neighbor to rely on and determines his optimal policies regardless of the other neighbors' choices. The IP holds if: $(*)$ the probability with which an agent takes the best action increases compared to that of his chosen neighbor; $(**)$ such increases last until maximal learning occurs.

For $(*)$ to hold, search costs must not be bounded away from 0 (condition (i) in Theorem 1). Consider agent n and his chosen neighbor b. Agent n finds it optimal to begin searching from the action taken by b. Moreover, unless b takes the best action with probability 1, n's expected gain from the second search is positive. Therefore, if search costs are not bounded away from 0, agent n samples both actions with positive probability and takes a better action than the one he samples first. Hence, there is a strict improvement in the probability of taking the best action that n has over b.

In turn, $(**)$ requires the network to be sufficiently connected (condition (ii)–(a) in Theorem 1). An information path for agent n is a sequence $(\pi_1, \ldots, \pi_k)$ with $\pi_k = n$ and $\pi_i \in B(\pi_{i+1})$ for all $i \in \mathbb{N}_k$. If arbitrarily long information paths almost surely occur, improvements last until agents take the best action with probability 1. In addition, $(**)$ requires that agents can single out the correct neighbor to rely on (condition (ii)–(b) in Theorem 1). Agent n can rely on agent γ_n only if $\gamma_n \in B(n)$. With correlated neighborhoods, the probability that γ_n takes the best action conditional on n observing γ_n is not the same as the unconditional probability. That is, n earns γ_n's probability of taking the best action *conditional* on choosing to rely on γ_n. Thus, $\mathbb{P}_\sigma(a_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n)$ and $\mathbb{P}_\sigma(a_{\gamma_n} \in \arg\max_{x \in X} q_x)$ must be approximately equal for large n to ensure that agent n can single out the correct neighbor to rely on.

A variety of explicit and intuitive conditions on the network topology only suffice for condition (ii)–(b) in Theorem 1 to hold (independent of equilibrium selection). I refer to Appendix B for such conditions. However, Theorem 1 is more general: condition (ii)–(b) holds whenever the network topology is such that an agent's belief about the realized network given his neighborhood realizations—a conditional probability measure—is not too different from the unconditional probability measure—i.e., the network topology.

Limits to the Improvement Principle. If search costs are bounded away from 0, late-moving agents cannot improve upon imitation, and maximal learning via the IP fails in all network topologies. To understand why the IP is fragile to perturbations of the search technology, suppose $\underline{c} > 0$, $\omega \in \Omega(\underline{c})$, and the IP holds. Then, in some chosen neighbor topology, the probability of none of the agents in $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to 0 as $n \to \infty$. Thus, the expected gain from the second search for all agents moving after some sufficiently late time m is below $\underline{c} > 0$. As a result, no such agent in the chosen neighbor topology will sample the second action. By Assumption 1 and the equilibrium characterization, the probability that none of the agents in $\widehat{B}(m) \cup \{m\}$ samples both actions is positive for any finite m. Thus, a contradiction arises, as the probability of none of the agents in $\widehat{B}(n) \cup \{n\}$ sampling both actions remains bounded away from 0.

3.2 Maximal Learning via the Large-Sample Principle

The next theorem provides sufficient conditions on network topologies under which maximal learning occurs independently of whether search costs are bounded away from 0.

First, I introduce some notation. For all $n \in \mathbb{N}$, let $B_n^{\emptyset} := \{k \in \mathbb{N}_n : B(k) = \emptyset\}$ be the set of agent n's isolated predecessors. Moreover, let $S := \{n \in \mathbb{N} : B(n) \in \{\emptyset, B_n^{\emptyset}\}\}$ the set of agents whose neighborhood is either empty or consists of all their isolated predecessors.

Theorem 2. *For all search technologies, maximal learning occurs if the network topology satisfies the following conditions:*

(i) *For all* $n \in \mathbb{N}$, $\mathbb{Q}(n \in S) > 0$*;*

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

(iii) *For all* $n \in \mathbb{N}$ *such that* $\mathbb{Q}(1 \in B(n)) > 0$, $\mathbb{Q}(B(n) = B_n^{\emptyset} \mid 1 \in B(n)) = 1$*;*

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

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

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

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

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

By Theorem 2, maximal learning occurs in networks where there are two sets of agents. The first set is S and forms the core of the network. Within set S, there are two groups of agents: an infinite but vanishing group of isolated agents (conditions (i) and (ii) in Theorem 2); agents that observe all and only their isolated predecessors and know they are doing so (condition (iii) in Theorem 2). Maximal learning occurs within S because of the following argument. When $\omega \in \Omega(\underline{c})$ and actions do not have the same quality, each isolated agent takes the best action with a probability larger than 1/2. Moreover, the choices of the infinitely many isolated agents are independent. 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. Agents in S learn via a *large-sample principle*, i.e., by observing and aggregating the information in large samples of individual choices.

The second set is $\mathbb{N} \setminus S$ and forms the rest of the network. By condition (iv) in Theorem 2, late-moving The second set is $\mathbb{N} \setminus S$ and forms the rest of the network. By condition (iv) in Theorem 2, late-moving agents observe recent choices from non-isolated agents in S with a high probability. By condition (v) in Theorem 2, agents' beliefs about the network conditional on their neighborhood are not too distorted compared to the actual network topology—similarly to condition (ii)–(b) in Theorem 1, but now (5) has to hold for all sequences of neighbor choice functions. Thus, each late-moving agent in $\mathbb{N} \setminus S$ can carefully select one of his most recent neighbors and sample this neighbor's action at the first search.

This strategy 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 $\omega \in \Omega(\underline{c})$. Intuitively, once maximal learning occurs via the LSP within the core, the remaining agents learn via imitation (even though search costs that are bounded away from 0 may preclude *improvements* upon imitation).

Example 1. The class of networks where maximal learning occurs is large and diverse, and the core S may form the entire network or consist of an arbitrarily small portion of it. For instance, 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, $S = \mathbb{N}$.
2. For all $n > 2$, let $\mathbb{Q}(B(n) = \emptyset) = 1/n$, $\mathbb{Q}(B(n) = B_n^{\emptyset}) = 1/\sqrt{n}$, and $\mathbb{Q}(B(n) = \{2, \ldots, n-1\}) = 1 - (1 + \sqrt{n})/n$. In this case, 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 relies on the assumption that non-isolated agents in S: (i) observe *only* isolated agents; (ii) *know* that they are observing isolated agents. Under this premise, the first search for non-isolated agents in S depends only on the relative shares of choices they observe. When non-isolated agents in S observe more 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 general characterization where we allow non-isolated agents in S to know only 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 equilibrium characterization. Formally, $(\mathbb{P}_\sigma(E_n^x))_{n\in\mathbb{N}}$, $x = 0, 1$, 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.

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 6. *A network topology is an* OIP network *if, for all* $n \in \mathbb{N}$,

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

The class of OIP networks is large, ranging from deterministic networks to stochastic networks with independent or correlated neighborhoods, as the following examples show.

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 $n > 3$, $\mathbb{Q}(B(n) = B_n^1 \mid B(3) = B_3^1) = 1$ and $\mathbb{Q}(B(n) = B_n^{n-1} \mid B(3) = B_3^2) = 1$. Here, neighborhoods are correlated, and each agent observes either the immediate predecessor or all of them, depending on agent 3's neighborhood realization.

The next theorem characterizes necessary conditions for maximal 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* $n \in \mathbb{N}$.

By Theorem 3–(i), maximal learning always fails with non-expanding subnetworks. The intuition is the following. Suppose $\underline{c} > 0$ and $\omega \in \Omega(\underline{c})$. By Assumption 1 and the equilibrium characterization, the probability that none of finitely many agents samples both actions is positive. Since non-expanding subnetworks have 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.

By Theorem 3–(ii), when search costs are bounded away from 0, maximal learning fails in OIP networks and if agents have at most one neighbor. In such networks, the probability of taking the best action under Bayes rationality is the same as under the IP. As the IP fails when search costs are bounded away from 0, so does Bayesian learning.

The analysis shows a discontinuity in learning outcomes between complete and almost-complete networks. By Theorem 3–(ii)–(a), search costs that are not bounded away from 0 are necessary for maximal learning in the complete network. In contrast, part 2 of Example 1 exhibits an almost-complete network where search costs that are not bounded away from 0 are not necessary for maximal learning.

4 Convergence Rate, Welfare, and Efficiency

Since several insights on the speed and the efficiency of social learning emerge in OIP networks, I begin by sketching equilibrium strategies in such networks (the formalities are in Appendix E). First, I introduce the relevant terminology.

Definition 7. *In OIP networks, action x is:*

(a) Revealed inferior to agent n *if $a_j = x$ and $a_{j+1} = \neg x$ for some agents $j, j+1 \in B(n)$.*

(b) Revealed inferior by time n *if $a_j = x$ and $a_{j+1} = \neg x$ for some agents $j, j+1 < n$.*

(c) Inferior by time n *if an agent $j < n$ sampled both actions and $a_j = \neg x$.*

Equilibrium Strategies in OIP Networks. Fix $n \geq 2$. At the first search, agent n samples the action taken by his immediate predecessor: $s_n^1 = a_{n-1}$.

At the second search, the optimal policy is as follows.

- If $\neg s_n^1$ is revealed inferior to agent n, then n discontinues searching. To see why, suppose $a_j = \neg s_n^1$ and $a_{j+1} = s_n^1$ for some $j, j+1 \in B(n)$. Since each agent samples the action taken by his immediate predecessor first, agent $j+1$ must have sampled action $\neg s_n^1$ first and takes $a_{j+1} = s_n^1$ only if he sampled action s_n^1 as well, and $q_{s_n^1} \geq q_{\neg s_n^1}$.
- If $\neg s_n^1$ is not revealed inferior to agent n, the expected gain from the second search is the same as in the complete network if the action is not revealed inferior by time n. The intuition goes as follows. In all OIP networks, agent n's subnetwork, $\{1, \ldots, n-1\}$, coincides with n's neighborhood in the complete network. Moreover, each agent samples first the action taken by the immediate predecessor. Thus, given $q_{s_n^1}$, the probability that none of the agents in $\widehat{B}(n, s_n^1)$ sampled both actions must be the same. But then, if $\neg s_n^1$ is not revealed inferior to agent n, the expected gain from the second search is the same as in the complete network if $\neg s_n^1$ is not revealed inferior by time n.

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

Proposition 1. *In all OIP networks, the probability of wrong herds—i.e., that all sufficiently late-moving agents take the same wrong action—is the same as in the complete network. Moreover, if search costs are not bounded away from* 0*, the convergence rate to the best action is the same as in the complete network.*

4.1 Convergence Rate

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

Definition 8. *Let* $\underline{q} := \min \operatorname{supp}(\mathbb{P}_Q)$*. Search costs have a* polynomial shape *if, for some* $K, L \in \mathbb{R}$ *with* $K \geq 0$ *and* $0 < L < \frac{2^{K+1}}{(K+2)t^{\emptyset}(\underline{q})^K}$*,* $F_C(c) \geq Lc^K$ *for all* $c \in \left(0, t^{\emptyset}\left(\underline{q}\right)/2\right)$*.*

The density of indirect connections affects convergence rates. Whereas the convergence rate 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 occurs faster in OIP networks because the cardinality of agents' subnetworks grows faster, and so does the probability that at least one agent in the subnetworks samples both actions.

Theorem 4. *Suppose search costs are not bounded away from* 0 *and have a polynomial shape.*

(a) In OIP networks,

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

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

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

4.2 Equilibrium Welfare and Efficiency in OIP Networks

In this section, I characterize how network transparency affects equilibrium welfare and compare equilibrium welfare against the efficiency benchmark in which a social planner makes all choices. To aid analysis, I assume $\mathbb{P}_C$ 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 why, suppose $a_j = x$ and $a_{j+1} = \neg x$ for some agents $j, j+1$. Hence, action x is revealed inferior by time $j+2$. In the complete network, action x is revealed inferior to any agent $n \geq j+2$ and so is never sampled again. In other OIP networks, instead, agent j need not be in the neighborhood of agent $n \geq j+2$. Therefore, n fails to realize from agent $j+1$'s choice that $q_x \leq q_{\neg x}$ and may inefficiently samples action x again.

Inefficient duplication of costly searches is more severe the shorter in the past agents can observe. Hence, the complete network is the most efficient OIP network, whereas the network where agents observe only the most immediate predecessor is the least efficient OIP network. In all other OIP networks, equilibrium welfare is between these two bounds. The next proposition shows that welfare losses due to duplication of costly searches vanish in arbitrarily patient societies (equivalently, in the long run) but remain significant otherwise.

Proposition 2. *Suppose future payoffs are discounted at rate* $\delta \in (0,1)$*. For all* $\delta \in (0,1)$*, the equilibrium social utility is larger in the complete network than in the network where agents observe only their most immediate predecessor. The difference vanishes as* $\delta \to 1$*.*

Social Planner Benchmark. Consider a social planner who draws a new search cost in each period and faces the same connections' structure as the agents. The social planner discounts future payoffs at rate $\delta \in (0,1)$, internalizes future gains of the current search, and samples each action 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 social planner achieves the same social utility in all OIP networks.

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

(i) Agents do not internalize future gains of the current 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 for two reasons:

   (a) Agent n fails to recognize an action that is inferior and not revealed so by time n.

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

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

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

Proposition 3. *In OIP networks, the equilibrium social utility converges to the social planner's as $\delta \to 1$ if and only if search costs are not bounded away from* 0*.*

5 Concluding Remarks

I conclude by discussing some modeling choices and extensions.

Search Technology. No result changes if both searches are costly, but agents cannot abstain and must search at least once. That both searches cost the same captures that information acquisition costs are idiosyncratic to agents but not to actions, which are ex-ante identical. Assuming that agents cannot abstain is standard in the social learning literature.

More than Two Actions. In Section I, I sketch the analysis for more than two available actions. In short, with more than two actions: (i) the results on maximal learning (Theorems 1–3) remain unchanged; (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.

Heterogeneous Preferences. Preference heterogeneity arises if agents' payoffs depend on a common and a private component. In the SSLM with heterogeneous preferences, the IP suffers, as imitation no longer guarantees the neighbor's payoff, whereas the LSP has more room to operate (see Lobel and Sadler, 2016). In my setting, the IP has more bite than the LSP, suggesting that preference heterogeneity may disrupt positive learning results.

Design of Collective Search. In Section J, I show that letting agents observe the shares of earlier choices reduces inefficiency in several network topologies. Interestingly, this simple policy is common in practice: many online platforms aggregate past decisions by sorting items according to their popularity or sales rank. Characterizing more sophisticated incentive schemes, which combine monetary transfers with information management tools to further reduce inefficiencies and foster social exploration, may be interesting for future work.[3]

Related Work. In Lomys and Tarantino (2023), we investigate how social learning affects the identification and estimation of search models. In Bigoni, Boldrini, Lomys, and Tarantino (2023), we experimentally study how social information influences individual search behavior and how behavioral biases and the perceived reliability of the information source affect this process.

[3] A growing literature study incentivized social learning (see, e.g., Slivkins, 2019, 2023, for surveys).

A Proofs for Section 2.1

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

$$\mathbb{P}_Q(q_x \leq q) = \mathbb{P}_Q(q_{\neg x} \leq q), \tag{6}$$

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

and

$$\mathbb{P}_{\Omega|q_{\neg x} \geq q_x}(q_x \leq q) > \mathbb{P}_Q(q_x \leq q). \tag{8}$$

Pick any $n \in \mathbb{N}$. Under Assumption 1, E_n^x and E_n^{xC} occur with positive probability for all x and n. From the discussion of the optimal first search in Section 2.1, we have:

1. Given E_n^x, agent n's posterior belief on the quality of action $\neg x$ is the same as the prior $\mathbb{P}_Q$:

$$\mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid E_n^x, I_n^1\right) = \mathbb{P}_Q(q_{\neg x} \leq q). \tag{9}$$

2. Given E_n^{xC}, agent n's posterior belief on Ω is $\mathbb{P}_{\Omega|q_x \geq q_{\neg x}}$:

$$\mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid E_n^{xC}, I_n^1\right) = \mathbb{P}_{\Omega|q_x \geq q_{\neg x}}(q_{\neg x} \leq q). \tag{10}$$

Suppose $P_n(x) < P_n(\neg x)$. Given I_n^1, agent n's belief about the quality of action x strictly first-order stochastically dominates that about the quality of action $\neg x$. To see why, note that

$$\begin{aligned}
\mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid I_n^1\right) &= \mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid E_n^x, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^x \mid I_n^1\right) \\
&\quad + \mathbb{P}_{\sigma_{-n}}\left(q_{\neg x} \leq q \mid E_n^{xC}, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^{xC} \mid I_n^1\right) \\
&= \mathbb{P}_Q(q_{\neg x} \leq q)P_n(x) + \mathbb{P}_{\Omega|q_x \geq q_{\neg x}}(q_{\neg x} \leq q)(1 - P_n(x)) \\
&= \mathbb{P}_Q(q_x \leq q)P_n(x) + \mathbb{P}_{\Omega|q_{\neg x} \geq q_x}(q_x \leq q)(1 - P_n(x)) \\
&> \mathbb{P}_Q(q_x \leq q)P_n(\neg x) + \mathbb{P}_{\Omega|q_{\neg x} \geq q_x}(q_x \leq q)(1 - P_n(\neg x)) \\
&= \mathbb{P}_{\sigma_{-n}}\left(q_x \leq q \mid E_n^{\neg x}, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^{\neg x} \mid I_n^1\right) \\
&\quad + \mathbb{P}_{\sigma_{-n}}\left(q_x \leq q \mid E_n^{\neg xC}, I_n^1\right)\mathbb{P}_{\sigma_{-n}}\left(E_n^{\neg xC} \mid I_n^1\right) \\
&= \mathbb{P}_{\sigma_{-n}}\left(q_x \leq q \mid I_n^1\right),
\end{aligned} \tag{11}$$

where: the first and the last equalities hold by the law of total probability; the second equality holds by (1), (9), and (10); the third equality holds by (6) and (7); the inequality holds by (8) and the assumption $P_n(x) < P_n(\neg x)$; the fourth equality holds by (1), (9), and (10). If, instead, $P_n(x) = P_n(\neg x)$, the inequality in (11) becomes equality, and so $\mathbb{P}_{\sigma_{-n}}(q_{\neg x} \leq q \mid I_n^1) = \mathbb{P}_{\sigma_{-n}}(q_x \leq q \mid I_n^1)$, i.e., agent n's beliefs about the qualities of the two actions conditional on I_n^1 are identical.■

B Sufficient Conditions for (ii)–(b) in Theorem 1

If the network topology satisfies any of the conditions 1–6 below, then it has a sequence of neighbor choice functions satisfying condition (ii)–(b) in Theorem 1.

1. The network topology is deterministic.
2. The network topology has independent neighborhoods.

3. The network topology has deterministic information paths. That is, there is a sequence of neighbor choice functions such that the corresponding chosen neighbor topology is deterministic (this is so, for instance, in OIP networks).
4. The network topology has a sequence of neighbor choice functions such that neighborhoods $\{B(k)\}_{k=1}^{m}$ are independent of the event $\gamma_n = m$ for all $n, m \in \mathbb{N}$ with $n > m$.
5. The network topology has a sequence of neighbor choice functions such that the corresponding chosen neighbor topology has low network distortion.
6. The network topology is Markovian: neighborhoods $\{B(n)\}_{n\in\mathbb{N}}$ are conditionally independent given the state of an underlying Markov chain with finitely many states.

Under conditions 1–4, we have

$$\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X}\, q_x \mid \gamma_n\right) - \mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X}\, q_x\right) = 0, \tag{12}$$

and so condition (ii)–(b) in Theorem 1 trivially follows. Regarding conditions 5 and 6, 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 (12) need not hold. However, Lobel and Sadler (2015) show that, under such conditions, there is a sequence of neighbor choice functions such that the difference in the left-hand side of (12) is arbitrarily small (with arbitrarily large probability) for large enough n.

C Proof of Theorem 1

Theorem 1 follows by combining two propositions which, combined, form the IP. The first proposition provides sufficient conditions for maximal learning via improvements upon imitation to occur.

Proposition 4. *Maximal learning occurs if there is 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 has 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 is $N_{\varepsilon\eta} \in \mathbb{N}$ such that, for all $n > N_{\varepsilon\eta}$, with probability at least $1-\eta$,

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X}\, q_x \mid \gamma_n\right) \geq \mathcal{Z}\left(\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X}\, q_x\right)\right) - \varepsilon.$$

Condition (c) requires the existence of a strict lower bound on the increase in the probability that an agent takes the best action over his chosen neighbor's probability. The second proposition shows that this is possible if search costs are not bounded away from 0.

Proposition 5. *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 is an increasing and continuous function $\mathcal{Z}\colon [1/2,1] \to [1/2,1]$, with $\mathcal{Z}(\beta) > \beta$ for all $\beta \in [1/2,1)$, and $\mathcal{Z}(1) = 1$, such that, for all $n \in \mathbb{N}$,*

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X}\, q_x \mid \gamma_n\right) \geq \mathcal{Z}\left(\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X}\, q_x \mid \gamma_n\right)\right).$$

The next two sections contain the proofs of Propositions 4 and 5.

C.1 Proof of Proposition 4

Preliminaries. 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 1. *For all $n \in \mathbb{N}$, we have*

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

Proof. If $n = 1$, the claim trivially holds. Now fix any $n > 1$. If $B(n) = \emptyset$, agent n's problem is identical to agent 1's, and the desired result follows. If $B(n) \neq \emptyset$, by the characterization of the equilibrium decision s_n^1 in Section 2.1, we have $\mathbb{P}_\sigma\left(E_n^{s_n^1} \mid I_n^1\right) \leq \mathbb{P}_\sigma\left(E_n^{s_1^1} \mid I_n^1\right)$. By integrating over all possible search costs, choices of the agents in the neighborhood, and neighborhoods, we have $\mathbb{P}_\sigma\left(E_n^{s_n^1}\right) \leq \mathbb{P}_\sigma\left(E_n^{s_1^1}\right)$. Then, the distribution of the quality of action s_n^1 first-order stochastically dominates (in the case of a strict inequality), or is the same as (in the case of equality), that of action s_1^1. That $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x \in X} q_x) \geq \mathbb{P}_\sigma(s_1^1 \in \arg\max_{x \in X} q_x)$ follows. ■

C.1.1 Proof of Proposition 4

First, 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

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x \in X} q_x\right) \geq \phi_k \quad \text{for all } n \geq \alpha_k. \tag{13}$$

Second, I show that $\phi_k \to 1$ as $k \to \infty$. The desired result follows.

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

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

$$\mathbb{P}_\sigma\left(\mathbb{P}_\sigma\left(a_n \in \arg\max_{x \in X} q_x \;\middle|\; \gamma_n\right) < \mathcal{Z}\left(\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x \in X} q_x\right)\right) - \varepsilon < \frac{\varepsilon}{2}\right. \tag{15}$$

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

$$\phi_{k+1} := \frac{\phi_k + \mathcal{Z}(\phi_k)}{2}, \qquad \alpha_{k+1} := N(\alpha_k, \varepsilon_k), \qquad \varepsilon_k := \frac{1}{2}\left(1 + \mathcal{Z}(\phi_k) - \sqrt{1 + 2\phi_k + \mathcal{Z}(\phi_k)^2}\right).$$

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. and agent 1 has no prior information,

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

Moreover, note that

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x \in X} q_x\right) \geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x \in X} q_x\right) \geq \frac{1}{2} \quad \forall n \geq 1, \tag{17}$$

where: the first inequality holds because agent n takes the best action between those he sampled;

the second inequality holds by Lemma 1 and (16). Then, (17), together with $\alpha_1 = 1$ and $\phi_1 = \frac{1}{2}$, establishes (13) for $k = 1$. Assume that (13) holds for an arbitrary k, and consider some agent $n \geq \alpha_{k+1}$. Let $\mathcal{B}_{\gamma_n}$ be the set of agents $0 \leq b < n$ that γ_n selects with positive probability. To establish (13) for $n \geq \alpha_{k+1}$, observe that

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

where the inequalities follows from (14) and (15), the inductive hypothesis, the assumption that $\mathcal{Z}$ is increasing, and the definitions of ϕ_k, ϕ_{k+1}, and ε_k.

Part 2. By assumption (b), $\mathcal{Z}(\beta) \geq \beta$ 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 $2\phi^* = 2\lim_{k\to\infty}\phi_k = \lim_{k\to\infty}\left[\phi_k + \mathcal{Z}(\phi_k)\right] = \phi^* + \mathcal{Z}(\phi^*)$, 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 5

Proposition 5 follows by combining several lemmas, which I next present.

Let $(\tilde{s}_n^1, \tilde{s}_n^2, \tilde{a}_n)$ denote agent n's coarse optimal policies when he only uses information from his chosen neighbor. The following remark, which is obvious, states that an agent's coarse optimal policies coincide with the optimal one if the agent has at most one neighbor.

Remark 2. If $|B(n)| \leq 1$, then $(\tilde{s}_n^1, \tilde{s}_n^2, \tilde{a}_n) = (s_n^1, s_n^2, a_n)$.

Hereafter, I assume agent n samples action a_{γ_n} if indifferent. This does not affect the results. Moreover, I use the convention $a_{\gamma_n} = \frac{1}{2}\circ 0 + \frac{1}{2}\circ 1$, where $\sum_x \xi(x)\circ x$ denotes the mixture assigning probability $\xi(x)$ to action x, whenever $\gamma_n = 0$ (or, equivalently, $B(n) = \emptyset$).

Lemma 2. *If* $\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}) \leq \mathbb{P}_\sigma(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n})$*, then* $\tilde{s}_n^1 = a_{\gamma_n}$.

Proof. By the optimal search policy, $\tilde{s}_n^1 \in \arg\min_{x\in X} \mathbb{P}_\sigma(E_n^x \mid c_n, \gamma_n, a_{\gamma_n})$. The desired result follows by observing that E_n^x is independent of c_n for all x. ■

The next lemma shows that Lemma 2 applies to network topologies where $\mathbb{Q}(|B(n)| \leq 1) = 1$ for all n, and so to all chosen neighbor topologies.

Lemma 3. *If* $\mathbb{Q}(|B(n)| \leq 1) = 1$ *for all* $n \in \mathbb{N}$*, then* $\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}) \leq \mathbb{P}_\sigma(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n})$.

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

Now suppose $\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}) \leq \mathbb{P}_\sigma(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n})$ for all $n \leq k$. If $B(k+1) = \emptyset$, agent k's problem is identical to that of agent 1, and the desired result 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, $\pi_1 = \min \widehat{B}(k+1)$, $\pi_l = k+1$ 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, each agent in $(\pi_1, \ldots, \pi_{l-1})$ finds it optimal to sample the action taken by his immediate predecessor first. ■

Definition 9. *The following objects are defined:*

$$q_{\min} \coloneqq \min\{q_0, q_1\} \quad \textit{and} \quad q_{\max} \coloneqq \max\{q_0, q_1\},$$
$$P_n^b(q_{\min}) \coloneqq \mathbb{P}_\sigma\left(E_n^{\tilde{s}_n^1} \mid \gamma_n, q_{\tilde{s}_n^1} = q_{\min}\right),$$
$$P_n^b(q_{\max}) \coloneqq \mathbb{P}_\sigma\left(E_n^{\tilde{s}_n^1} \mid \gamma_n, q_{\tilde{s}_n^1} = q_{\max}\right),$$
$$\beta \coloneqq \mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n\right).$$

Remark 3. For all n and γ_n, we have $\beta \geq \frac{1}{2}$. That is, any agent takes the best action at least with the same probability with which he would do so by sampling an action uniformly at random.

The next two lemmas provide an expression for the probability of agent n taking the best action when using the coarse policies $(\tilde{s}_n^1, \tilde{s}_n^1, \tilde{a}_n)$.

Lemma 4. *Suppose* $\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}) \leq \mathbb{P}_\sigma(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n})$. *Then,*

$$\mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x \in X} q_x \mid \gamma_n\right) = \beta + \mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n\right)(1 - \beta). \tag{18}$$

Proof. Note that

$$\begin{aligned}
\mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x \in X} q_x \mid \gamma_n\right) &= \mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x \in X} q_x \mid \gamma_n, \tilde{s}_n^2 = \neg\tilde{s}_n^1\right)\mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n\right) \\
&\quad + \mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x \in X} q_x \mid \gamma_n, \tilde{s}_n^2 = d\right)\mathbb{P}_\sigma\left(\tilde{s}_n^2 = d \mid \gamma_n\right) \\
&= \mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n\right) \qquad (19) \\
&\quad + \mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n\right)\left(1 - \mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n\right)\right) \\
&= \mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n\right) \\
&\quad + \mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n\right)\left(1 - \mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n\right)\right).
\end{aligned}$$

The first equality holds by the law of total probability. The second equality holds because: (i) when agent n samples both actions, he takes the best one, and so $\mathbb{P}_\sigma(\tilde{a}_n \in \arg\max_{x \in X} q_x \mid \gamma_n, \tilde{s}_n^2 = \neg\tilde{s}_n^1) = 1$; (ii) when agent n only samples one action, he takes that action, and so, since $\tilde{s}_n^1 = a_{\gamma_n}$ by Lemma 2, $\mathbb{P}_\sigma(\tilde{a}_n \in \arg\max_{x \in X} q_x \mid \gamma_n, \tilde{s}_n^2 = d) = \mathbb{P}_\sigma(a_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n)$. The desired result follows from (19) and the definition of β. ■

Lemma 5. *Suppose* $\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}) \leq \mathbb{P}_\sigma(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n})$. *Then,*

$$\mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x \in X} q_x \mid \gamma_n\right) = \beta + (1 - \beta)\left[\beta F_C\left(P_n^b(q_{\max})t^{\emptyset}(q_{\max})\right) + (1 - \beta)F_C\left(P_n^b(q_{\min})t^{\emptyset}(q_{\min})\right)\right].$$

Proof. Note that

$$\begin{aligned}\mathbb{P}_\sigma\Big(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n\Big) &= \mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n, \tilde{s}_n^1 \in \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n\right)\\ &+ \mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n, \tilde{s}_n^1 \notin \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(\tilde{s}_n^1 \notin \arg\max_{x\in X} q_x \mid \gamma_n\right)\\ &= \beta\mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n, \tilde{s}_n^1 \in \arg\max_{x\in X} q_x\right) \qquad (20)\\ &+ (1-\beta)\mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n, \tilde{s}_n^1 \notin \arg\max_{x\in X} q_x\right),\end{aligned}$$

where: the first equality holds by the law of total probability; the second equality holds because, by Lemma 2, $\tilde{s}_n^1 = a_b$, and the definition of β. By the equilibrium characterization in Section 2.1, we have: conditional on γ_n and $\tilde{s}_n^1 \in \arg\max_{x\in X} q_x$, $\tilde{s}_n^2 = \neg\tilde{s}_n^1 \iff c_n \le P_n^b(q_{\max})t^\emptyset(q_{\max})$; conditional on γ_n and $\tilde{s}_n^1 \notin \arg\max_{x\in X} q_x$, $\tilde{s}_n^2 = \neg\tilde{s}_n^1 \iff c_n \le P_n^b(q_{\min})t^\emptyset(q_{\min})$. Thus,

$$\mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n, \tilde{s}_n^1 \in \arg\max_{x\in X} q_x\right) = F_C\Big(P_n^b(q_{\max})t^\emptyset(q_{\max})\Big),$$

and

$$\mathbb{P}_\sigma\left(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n, \tilde{s}_n^1 \notin \arg\max_{x\in X} q_x\right) = F_C\Big(P_n^b(q_{\min})t^\emptyset(q_{\min})\Big).$$

Hence, (20) can be rewritten as

$$\mathbb{P}_\sigma\Big(\tilde{s}_n^2 = \neg\tilde{s}_n^1 \mid \gamma_n\Big) = \beta F_C\Big(P_n^b(q_{\max})t^\emptyset(q_{\max})\Big) + (1-\beta)F_C\Big(P_n^b(q_{\min})t^\emptyset(q_{\min})\Big). \qquad (21)$$

The desired result follows by combining (18) and (21). ■

The improvement $(1-\beta)[\beta F_C(P_n^b(q_{\max})t^\emptyset(q_{\max})) + (1-\beta)F_C(P_n^b(q_{\min})t^\emptyset(q_{\min}))]$ in the probability that agent n takes the best action over his chosen neighbor's probability is still unsuitable for the analysis to come as it depends on $P_n^b(q_{\min})$ and $P_n^b(q_{\max})$, which are difficult to handle. The next lemma provides a simpler lower bound on the amount of this improvement.

Lemma 6. *Suppose* $\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}) \le \mathbb{P}_\sigma(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n})$. *Then,*

$$\mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x\in X} q_x \mid \gamma_n\right) \ge \beta + (1-\beta)^2 F_C\Big((1-\beta)t^\emptyset(q_{\max})\Big).$$

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

$$1-\beta \le \mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n). \qquad (22)$$

Moreover, by the law of total probability,

$$\begin{aligned}\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n) &= \mathbb{P}_\sigma\left(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n} \in \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X} q_x \mid \gamma_n\right)\\ &+ \mathbb{P}_\sigma\left(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n} \notin \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(a_{\gamma_n} \notin \arg\max_{x\in X} q_x \mid \gamma_n\right) \qquad (23)\\ &= \beta P_n^b(q_{\max}) + (1-\beta)P_n^b(q_{\min}).\end{aligned}$$

Combining (22) and (23) yields $1-\beta \leq \beta P_n^b(q_{\max}) + (1-\beta)P_n^b(q_{\min})$, and therefore

$$\max\left\{P_n^b(q_{\min}), P_n^b(q_{\max})\right\} \geq 1-\beta. \tag{24}$$

Finally, observe that

$$\begin{aligned}
&(1-\beta)\Big[\beta F_C\Big(P_n^b(q_{\max})t^{\emptyset}(q_{\max})\Big) + (1-\beta)F_C\Big(P_n^b(q_{\min})t^{\emptyset}(q_{\min})\Big)\Big] \\
&\geq (1-\beta)\Big[(1-\beta)F_C\Big(P_n^b(q_{\max})t^{\emptyset}(q_{\max})\Big) + (1-\beta)F_C\Big(P_n^b(q_{\min})t^{\emptyset}(q_{\min})\Big)\Big] \\
&= (1-\beta)^2\Big[F_C\Big(P_n^b(q_{\max})t^{\emptyset}(q_{\max})\Big) + F_C\Big(P_n^b(q_{\min})t^{\emptyset}(q_{\min})\Big)\Big] \\
&\geq (1-\beta)^2\Big[F_C\Big(P_n^b(q_{\max})t^{\emptyset}(q_{\max})\Big) + F_C\Big(P_n^b(q_{\min})t^{\emptyset}(q_{\max})\Big)\Big] \\
&\geq (1-\beta)^2 \max\Big\{F_C\Big(P_n^b(q_{\max})t^{\emptyset}(q_{\max})\Big), F_C\Big(P_n^b(q_{\min})t^{\emptyset}(q_{\max})\Big)\Big\} \\
&\geq (1-\beta)^2 F_C\Big((1-\beta)t^{\emptyset}(q_{\max})\Big).
\end{aligned} \tag{25}$$

Here, the first inequality holds as $\beta \geq (1-\beta)$ (by Remark 3, $\beta \geq 1/2$); 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 $\max\{F_C(P_n^b(q_{\max})t^{\emptyset}(q_{\max})), F_C(P_n^b(q_{\min})t^{\emptyset}(q_{\max}))\} \geq F_C((1-\beta)t^{\emptyset}(q_{\max}))$, which holds because of (24) and the fact that F_C is increasing. The desired result follows from Lemma 5 and (25). ■

To study the limiting behavior of improvements, I define the function $\overline{\mathcal{Z}}\colon [1/2,1] \to [1/2,1]$ as

$$\overline{\mathcal{Z}}(\beta) := \beta + (1-\beta)^2 F_C\Big((1-\beta)t^{\emptyset}(q_{\max})\Big). \tag{26}$$

Hereafter, I call $(1-\beta)^2 F_C((1-\beta)t^{\emptyset}(q_{\max}))$ the *improvement term* of function $\overline{\mathcal{Z}}$. Lemma 6 establishes that, when $\mathbb{P}_\sigma(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}) \leq \mathbb{P}_\sigma(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n})$, we have

$$\mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x\in X} q_x \mid \gamma_n\right) = \overline{\mathcal{Z}}\left(\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X} q_x \mid \gamma_n\right)\right).$$

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

Lemma 7. *The function $\overline{\mathcal{Z}}$, defined by (26), 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, $F_C((1-\beta)t^{\emptyset}(q_{\max})) > 0$.[4] Since also $(1-\beta)^2 > 0$ for all $\beta \in [1/2,1)$, the improvement term of function $\overline{\mathcal{Z}}$ is positive. Part (b) follows.

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. Hence, 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,

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

the desired result follows because products and sums of left-continuous functions with no upward jumps are left-continuous with no upward jumps. ■

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

$$\mathcal{Z}(\beta) \coloneqq \frac{1}{2}\left(\beta + \sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)\right). \tag{27}$$

Lemma 8. *The function $\mathcal{Z}$, defined by (27), 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 7. 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 increasing, establishing the first part of (c). I establish continuity of $\mathcal{Z}$ in $[1/2,1)$ 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 non-decreasing, 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 is 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 by Lemma 7–(c). 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 taking the best action.

Lemma 9. *Suppose $\mathbb{P}_\sigma\big(E_n^{a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}\big) \leq \mathbb{P}_\sigma\big(E_n^{\neg a_{\gamma_n}} \mid \gamma_n, a_{\gamma_n}\big)$. Then,*

$$\mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x\in X} q_x \mid \gamma_n\right) \geq \mathcal{Z}\left(\mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X} q_x \mid \gamma_n\right)\right).$$

Proof. If $\mathcal{Z}(\beta) = \beta$, the result follows from Lemma 5. Suppose next that $\mathcal{Z}(\beta) > \beta$. By (27), this implies that $\mathcal{Z}(\beta) < \sup_{r\in[1/2,\beta]} \overline{\mathcal{Z}}(r)$. Thus, there is $\overline{\beta} \in [1/2,\beta]$ such that

$$\overline{\mathcal{Z}}(\overline{\beta}) \geq \mathcal{Z}(\beta). \tag{28}$$

I next show that $\mathbb{P}_\sigma\big(\tilde{a}_n \in \arg\max_{x\in X} q_x \mid \gamma_n\big) \geq \overline{\mathcal{Z}}(\overline{\beta})$, from which the desired result follows. Agent n can always make his decision even coarser by observing a fictitious agent whose action, denoted by $\tilde{a}_{\gamma_n}$, is generated as

$$\tilde{a}_{\gamma_n} = \begin{cases} a_{\gamma_n} & \text{with probability } \ (2\overline{\beta}-1)/(2\beta-1) \\ 0 & \text{with probability } \ (\beta-\overline{\beta})/(2\beta-1) \\ 1 & \text{with probability } \ (\beta-\overline{\beta})/(2\beta-1) \end{cases}, \tag{29}$$

with the realization of $\tilde{a}_{\gamma_n}$ independent of the rest of n's information set. Let $\widetilde{\tilde{a}_n}$ denote the choice of agent n upon observing the choice of the fictitious agent. Under the assumption $\mathbb{P}_\sigma\big(E_n^{a_{\gamma_n}} \mid \gamma_n\big) \leq \mathbb{P}_\sigma\big(E_n^{\neg a_{\gamma_n}} \mid \gamma_n\big)$, we have

$$\mathbb{P}_\sigma\left(E_n^{\tilde{a}_{\gamma_n}} \mid \gamma_n\right) \leq \mathbb{P}_\sigma\left(E_n^{\neg\tilde{a}_{\gamma_n}} \mid \gamma_n\right). \tag{30}$$

Moreover, note that

$$\begin{aligned}\mathbb{P}_\sigma\left(\tilde{a}_{\gamma_n} \in \arg\max_{x\in X} q_x \mid \gamma_n\right) &= \mathbb{P}_\sigma\left(a_{\gamma_n} \in \arg\max_{x\in X} q_x \mid \gamma_n\right)\frac{2\overline{\beta}-1}{2\beta-1} \\ &+ \mathbb{P}_\sigma\left(0 \in \arg\max_{x\in X} q_x \mid \gamma_n\right)\frac{\beta-\overline{\beta}}{2\beta-1} \\ &+ \mathbb{P}_\sigma\left(1 \in \arg\max_{x\in X} q_x \mid \gamma_n\right)\frac{\beta-\overline{\beta}}{2\beta-1} \\ &= \beta\frac{2\overline{\beta}-1}{2\beta-1} + (\beta+(1-\beta))\frac{\beta-\overline{\beta}}{2\beta-1} \\ &= \overline{\beta}.\end{aligned} \tag{31}$$

From Lemma 6, (30), and (31), it follows that

$$\mathbb{P}_\sigma\left(\widetilde{\tilde{a}_n} \in \arg\max_{x\in X} q_x \mid \gamma_n\right) \geq \overline{\mathcal{Z}}(\overline{\beta}). \tag{32}$$

By the characterization of the equilibrium decision s_n^1 in Section 2.1, we have $\mathbb{P}_\sigma\left(E_n^{\tilde{s}_n^1} \mid I_n^1\right) \leq \mathbb{P}_\sigma\left(E_n^{\widetilde{\tilde{a}_n}} \mid I_n^1\right)$. Let b be agent n's chosen neighbor. By integrating over all possible search costs, choices of the agents in the neighborhood, and neighborhoods such that $\gamma_n = b$, we have $\mathbb{P}_\sigma\left(E_n^{\tilde{s}_n^1} \mid \gamma_n = b\right) \leq \mathbb{P}_\sigma(E_n^{\widetilde{\tilde{a}_n}} \mid \gamma_n = b)$. Then, conditional on $\gamma_n = b$, the distribution of the quality of action $\tilde{s}_n^1$ first-order stochastically dominates or is the same as that of action $\widetilde{\tilde{a}_n}$. It follows that

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n\right) \geq \mathbb{P}_\sigma\left(\widetilde{\tilde{a}_n} \in \arg\max_{x\in X} q_x \mid \gamma_n\right). \tag{33}$$

Finally, since agent n takes the best action between those he sampled,

$$\mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x\in X} q_x \mid \gamma_n\right) \geq \mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n\right). \tag{34}$$

The desired result follows from (32), (33), and (34). ■

It remains to show that a_n does at least as well as its coarse version $\tilde{a}_n$ given γ_n. This is established with the next lemma and completes the proof of Proposition 5.

Lemma 10. *For all $n \in \mathbb{N}$, we have*

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x \mid \gamma_n\right) \geq \mathbb{P}_\sigma\left(\tilde{a}_n \in \arg\max_{x\in X} q_x \mid \gamma_n\right).$$

Proof. Fix any agent n. If $|B(n)| \leq 1$, then $\tilde{a}_n = a_n$ by Remark 2, and the desired result follows. Now suppose $|B(n)| > 1$, and let b be agent n's chosen neighbor. By the characterization of the equilibrium decision s_n^1 in Section 2.1, we have $\mathbb{P}_\sigma\left(E_n^{s_n^1} \mid I_n^1\right) \leq \mathbb{P}_\sigma\left(E_n^{\tilde{a}_n} \mid I_n^1\right)$. By integrating over all possible search costs, choices of the agents in the neighborhood, and neighborhoods such that $\gamma_n = b$, we have $\mathbb{P}_\sigma\left(E_n^{s_n^1} \mid \gamma_n = b\right) \leq \mathbb{P}_\sigma(E_n^{\tilde{a}_n} \mid \gamma_n = b)$. Then, conditional on $\gamma_n = b$, the distribution of the quality of action s_n^1 first-order stochastically dominates or is the same as that of action $\tilde{a}_n$. Hence, $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n = b) \geq \mathbb{P}_\sigma(\tilde{a}_n \in \arg\max_{x\in X} q_x \mid \gamma_n = b)$. Since agent n takes the best action between those he sampled, we have $\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x \mid \gamma_n = b\right) \geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n = b\right)$. The desired result follows. ■

D Proof of Theorem 2

Fix $\omega \in \Omega(\underline{c})$. Let $\mathbb{P}_\sigma^\omega$ denote the probability measure of $\left((s_n^1, s_n^2, a_n)\right)_{n\in\mathbb{N}}$ when the state is ω. I show that

$$\lim_{n\to\infty} \mathbb{P}_\sigma^\omega\left(a_n \in \arg\max_{x\in X} q_x\right) = 1. \tag{35}$$

If $q_0 = q_1$, there is nothing to prove. Hereafter, suppose $q_0 \neq q_1$. To establish (35), I first show that (35) holds once we restrict attention to agents in S and then extend the result to all agents.

Part 1. Let $x^* := \arg\max_{x\in X} q_x$. If $B(n) = \emptyset$, agent n takes the best action when he samples action x^* first, which occurs with probability 1/2, and when he samples action $\neg x^*$ first and $c_n \leq t^\emptyset(q_{\neg x^*})$. Since $\omega \in \Omega(\underline{c})$ and $q_0 \neq q_1$, the latter event occurs with positive probability. Therefore, $a_n = x^*$ with probability $\alpha > 1/2$.

If $B(n) = B_n^\emptyset$, by condition (iii) in Theorem 2, agent n knows he is observing only the choices of all his isolated predecessors. Thus, n's optimal first search decision depends on the relative shares of choices he observes. In particular, as $B(n) = \widehat{B}(n)$,

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

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

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

$$\lim_{n\to\infty} \mathbb{P}_\sigma^\omega\left(|\widehat{B}(n,x^*)| > |\widehat{B}(n,\neg x^*)| \mid B(n) = B_n^\emptyset\right) = 1. \tag{36}$$

Finally, we have

$$\begin{aligned}
1 &\geq \mathbb{P}_\sigma^\omega\left(s_n^1 \in \arg\max_{x\in X} q_x \mid n \in S\right) \\
&= \mathbb{P}_\sigma^\omega\left(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) = \emptyset\right)\mathbb{Q}\left(B(n) = \emptyset \mid n \in S\right) \\
&+ \mathbb{P}_\sigma^\omega\left(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) = B_n^\emptyset\right)\mathbb{Q}\left(B(n) = B_n^\emptyset \mid n \in S\right) \\
&= \frac{1}{2}\mathbb{Q}\left(B(n) = \emptyset \mid n \in S\right) \\
&+ \mathbb{P}_\sigma^\omega\left(|\widehat{B}(n,x^*)| > |\widehat{B}(n,\neg x^*)| \mid B(n) = B_n^\emptyset\right)\mathbb{Q}\left(B(n) = B_n^\emptyset \mid n \in S\right).
\end{aligned} \tag{37}$$

Here: the first equality holds by the law of total probability; the second equality holds by the optimal first search policy for agents in S characterized above.

By (36), (37), and since $\lim_{n\to\infty} \mathbb{Q}(B(n) = B_n^{\emptyset} \mid n \in S) = 1$ by condition (ii) in Theorem 2, we have

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

Since each agent takes the best action between those he sampled, the desired result follows.

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 $\mathbb{P}_\sigma\left(E_n^{s_n^1} \mid I_n^1\right) \le \mathbb{P}_\sigma(E_n^{a_b} \mid I_n^1)$ for all $b \in B(n)$. By integrating over all possible search costs and choices of agents in the neighborhood, we obtain $\mathbb{P}_\sigma\left(E_n^{s_n^1} \mid B(n)\right) \le \mathbb{P}_\sigma(E_n^{a_b} \mid B(n))$ for all $b \in B(n)$. Thus, conditional on $B(n)$, the distribution of the quality of action s_n^1 first-order stochastically dominates or is the same as that of action a_b for all $b \in B(n)$. Hence,

$$\mathbb{P}_\sigma^\omega \left(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) \right) \ge \max_{b\in B(n)} \mathbb{P}_\sigma^\omega \left(a_b \in \arg\max_{x\in X} q_x \mid B(n) \right). \tag{38}$$

By condition (v) in Theorem 2, for all $\varepsilon, \eta > 0$, there is $N_{\varepsilon\eta} \in \mathbb{N}$ such that, for all $n > N_{\varepsilon\eta}$, with probability at least $1 - \eta$,

$$\max_{b\in B(n)} \mathbb{P}_\sigma^\omega \left(a_b \in \arg\max_{x\in X} q_x \mid B(n) \right) \ge \max_{b\in B(n)} \mathbb{P}_\sigma^\omega \left(a_b \in \arg\max_{x\in X} q_x \right) - \varepsilon. \tag{39}$$

Thus, by (38) and (39), for all $\varepsilon, \eta > 0$, there is $N_{\varepsilon\eta} \in \mathbb{N}$ such that, for all $n > N_{\varepsilon\eta}$, with probability at least $1 - \eta$,

$$\mathbb{P}_\sigma^\omega \left(s_n^1 \in \arg\max_{x\in X} q_x \mid B(n) \right) \ge \max_{b\in B(n)} \mathbb{P}_\sigma^\omega \left(a_b \in \arg\max_{x\in X} q_x \right) - \varepsilon \tag{40}$$

By (40), for all $\tilde{\varepsilon} > 0$, there is $N_{\tilde{\varepsilon}} \in \mathbb{N}$ such that, for all $n > N_{\tilde{\varepsilon}}$,

$$\begin{aligned} 1 &\ge \mathbb{P}_\sigma^\omega \left(s_n^1 \in \arg\max_{x\in X} q_x \right) \\ &\ge \mathbb{E}_\sigma^\omega \left[\max_{b\in B(n)\cap S} \mathbb{P}_\sigma^\omega \left(a_b \in \arg\max_{x\in X} q_x \right) \right] - \tilde{\varepsilon} \\ &\ge \mathbb{E}_\sigma^\omega \left[\max_{b\in B(n)\cap S} \mathbb{P}_\sigma^\omega \left(a_b \in \arg\max_{x\in X} q_x \right) \mid \max_{b\in B(n)\cap S} b \ge K \right] \mathbb{Q}\left(\max_{b\in B(n)\cap S} \ge K \right) - \tilde{\varepsilon} \end{aligned} \tag{41}$$

for all $K \in \mathbb{N}$. Since agents $n \in S$ take the best action with probability 1 as $n \to \infty$,

$$\lim_{n\to\infty} \mathbb{E}_\sigma^\omega \left[\max_{b\in B(n)\cap S} \mathbb{P}_\sigma^\omega \left(a_b \in \arg\max_{x\in X} q_x \right) \mid \max_{b\in B(n)\cap S} b \ge K \right] = 1. \tag{42}$$

Moreover, by condition (iv) in Theorem 2,

$$\lim_{n\to\infty} \mathbb{Q}\left(\max_{b\in B(n)\cap S} b \ge K \right) = 1. \tag{43}$$

From (41)–(43), we have $\lim_{n\to\infty} \mathbb{P}_\sigma^\omega(s_n^1 \in \arg\max_{x\in X} q_x) = 1$. Since each agent takes the best action between those he sampled, the desired result follows. ■

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.

Lemma 11. *In OIP networks, equilibrium search policies are as follows:*

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

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

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

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

$$c_n \leq t_n\Big(q_{s_n^1}\Big) := \begin{cases} P_1\Big(q_{s_n^1}\Big) t^{\emptyset}\Big(q_{s_n^1}\Big) & \text{if } n = 2 \\ P_1\Big(q_{s_n^1}\Big)\Big[\prod_{i=2}^{n-1}\Big(1 - F_C\Big(t_i\Big(q_{s_n^1}\Big)\Big)\Big)\Big] t^{\emptyset}\Big(q_{s_n^1}\Big) & \text{if } n > 2. \end{cases} \tag{44}$$

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 took the best action. If agent 2 knew this to be the case, his conditional belief on Ω is $\mathbb{P}_{\Omega | q_{a_1} \geq q_{\neg a_1}}$.
2. Agent 1 did not sample $\neg a_1$. If agent 2 knew this to be the case, his posterior belief on the quality of action $\neg a_1$ is 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 that about the quality of action $\neg a_1$. That $s_2^1 = a_1$ follows.

Consider any agent $n > 2$. Suppose 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 inferior to agent n, it must be that $q_{\neg a_{n-1}} \leq q_{a_{n-1}}$, and so $\neg a_{n-1}$ is not sampled at all. If action $\neg a_{n-1}$ is not revealed 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 $s_n^1 = a_{n-1}$ follows.

For part (ii)–(a), suppose $\neg a_{n-1}$ is revealed inferior to agent $n \geq 2$. Then, $a_j = \neg a_{n-1}$ and $a_{j+1} = a_{n-1}$ for some $j, j+1 \in B(n)$. By part (i), $s_{j+1}^1 = \neg a_{n-1}$. Since agents can only take an action they sampled, it must be that $s_{j+1}^2 = a_{n-1}$. Then, as agents take the best action whenever they sample both of them, $q_{a_{n-1}} \geq q_{\neg a_{n-1}}$. That $s_n^2 = d$ follows.

For part (ii)–(b), consider any agent $n \geq 2$ and suppose $\neg a_{n-1}$ is not revealed inferior to n. In OIP networks, $\widehat{B}(n) = \{1, \ldots, n-1\}$. Moreover, by part (i), each agent samples first the action taken by his immediate predecessor. Thus, none of the agents in $\widehat{B}(n, s_n^1)$ sampled action $\neg s_n^1$ if and only if $s_i^1 = s_n^1$, and $s_i^2 = d$ for $1 \leq i \leq n-1$. The thresholds in (44) provide an explicit formula for (4) in OIP networks. To see why, proceed by induction. First, consider agent 2. By part (i), $s_2^1 = a_1$. Let $P_1\Big(q_{s_2^1}\Big)$ 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. Agent 2's expected gain from the second search is $P_1\Big(q_{s_2^1}\Big) t^{\emptyset}\Big(q_{s_2^1}\Big)$, which is the first line on the right-hand side of (44). Next, consider any agent $n > 2$. Let s_n^1 be the action agent n 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 $\{1, \ldots, n-1\}$ sampled action $\neg s_n^1$ is $P_1\Big(q_{s_n^1}\Big)\Big[\prod_{i=2}^{n-1}\Big(1 - F_C\Big(t_i\Big(q_{s_n^1}\Big)\Big)\Big)\Big]$. Hence, the second line on the right-hand side of (44) gives agent n's expected gain from the second search. The optimality of the proposed search policy follows from the equilibrium characterization in Section 2.1. ■

Remark 4. By Lemma 11, 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.

F Proof of Theorem 3

Fix any $\omega \in \Omega(\underline{c})$ such that $q_0 \neq q_1$ and $F_C(t^{\emptyset}(\min\{q_0, q_1\})) < 1$. Let $\mathbb{P}^{\omega}_{\sigma}$ denote the probability measure of $\big((s^1_n, s^2_n, a_n)\big)_{n\in\mathbb{N}}$ when the state is ω. I show that $\limsup_{n\to\infty} \mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\}\Big) > 0$. That is, there is a subsequence of agents who, with probability bounded away from 0, do not compare the quality of the two actions. But then, since the only way to ascertain the relative quality of the two actions is to sample both of them, we have $\liminf_{n\to\infty} \mathbb{P}^{\omega}_{\sigma}(a_n \in \arg\max_{x\in X} q_x) < 1$. By Assumption 1, the probability that the state ω satisfies $q_0 \neq q_1$ and $F_C(t^{\emptyset}(\min\{q_0, q_1\})) < 1$ is positive, and so that maximal learning fails follows.

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

$$\mathbb{Q}\Big(|\widehat{B}(n)| < K\Big) \geq \varepsilon \quad \text{for all } n \in \mathcal{N}. \tag{45}$$

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

$$\begin{aligned}
&\mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\}\Big) \\
&= \mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\} \mid |\widehat{B}(n)| < K\Big)\mathbb{Q}\Big(|\widehat{B}(n)| < K\Big) \\
&+ \mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\} \mid |\widehat{B}(n)| \geq K\Big)\mathbb{Q}\Big(|\widehat{B}(n)| \geq K\Big) \\
&\geq \mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\} \mid |\widehat{B}(n)| < K\Big)\mathbb{Q}\Big(|\widehat{B}(n)| < K\Big) \\
&\geq \varepsilon\mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\} \mid |\widehat{B}(n)| < K\Big),
\end{aligned} \tag{46}$$

where: the equality holds by the law of total probability; the last inequality holds by (45).

Let $\overline{C} \coloneqq \{c \in C : c > t^{\emptyset}(\min\{q_0, q_1\})\}$ be the set of all search costs for which an isolated agent does not sample the second action independently of which action he samples first. Any other agent k with search cost $c_k \in \overline{C}$ does not sample the second action either independently of his neighborhood realization, the choices of his neighbors, and the quality of the first action sampled (see Section 2.1). Then, as $|\widehat{B}(n)| < K \iff |\widehat{B}(n) \cup \{n\}| \leq K$ and search costs are i.i.d. across agents,

$$\mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\} \mid |\widehat{B}(n)| < K\Big) \geq \mathbb{P}^{\omega}_{\sigma}\Big(c_1 \in \overline{C}\Big)^K > 0, \tag{47}$$

where the strict inequality holds because $F_C(t^{\emptyset}(\min\{q_0, q_1\})) < 1$ by assumption. As $\varepsilon > 0$, from (46) and (47) we conclude that $\mathbb{P}^{\omega}_{\sigma}\Big(s^2_k = d \ \forall k \in \widehat{B}(n) \cup \{n\}\Big) > 0$ for all agents n in the subsequence $\mathcal{N}$. The desired result follows. ■

Proof of Theorem 3, part (ii)–(a). Let q be the quality of the action agent 1 takes. By way of contradiction, suppose the probability of no agent in $\widehat{B}(n) \cup \{n\}$ sampling both actions converges to 0 as $n \to \infty$. That is, $\lim_{n\to\infty} P_1(q)[\prod_{i=2}^{n}(1 - F_C(t_i(q)))] = 0$ (see Lemma 11 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 11), where $\hat{q}$ is the quality of the action taken by agent n, also converges to 0 as $n \to \infty$. Then, there is an agent $N_{\hat{q}} + 1$ for which the expected gain from the second search falls below $\underline{c}$.

Without loss, assume $q_0 < q_1$. With positive probability, agent 1 samples action 0 first and does not sample action 1 (by assumption, $F_C(t^\emptyset(q_0)) < 1$). Now suppose the first N_{q_0} agents all have costs larger than $t^\emptyset(q_0)$, which occurs with positive probability. By Lemma 11, each of these agents will sample action 0 first, and none of these agents will search further. Hence, all will take action 0. Agent $N_{q_0} + 1$ also samples action 0 first and discontinues searching because his expected gain from the second search is smaller than $\underline{c}$. Since the expected gain from the second search is non-increasing in n, there will be no further search by agents $N_{q_0} + 1$ 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). Pick a sequence of agents $(\pi_1, \pi_2, \ldots, \pi_k, \pi_{k+1}, \ldots)$ with $B(\pi_1) = \emptyset$ 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 3, each agent in this sequence samples first the action taken by his neighbor.

By way of contradiction, suppose 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 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 (3). Thus, 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 $k \to \infty$. Then, there is 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, $q_0 \neq q_1$. Without loss, assume $q_0 < q_1$. With positive probability, agent π_1 samples action 0 at the first search and does not sample action 1 (by assumption, $F_C(t^\emptyset(q_0)) < 1$). Now suppose the first N_{q_0} agents all have costs larger than $t^\emptyset(q_0)$, which occurs with positive probability. By Lemma 3, each of these agents will sample action 0 first, and none of these agents will search further. Therefore, all will take action 0. Agent $\pi_{K_{q_0}} + 1$ also samples action 0 first and discontinues searching 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 $\pi_{K_{q_0}} + 1$, contradicting that the probability of no agent in $\widehat{B}(\pi_k) \cup \{\pi_k\}$ sampling both actions converges to 0. ■

G Proof of Theorem 4

The proof builds on a technique developed by Lobel et al. (2009), which approximates a lower bound on the convergence rate with an ordinary differential equation.

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

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

By Proposition 1, the convergence rate to the best action in all OIP networks is the same as in the OIP network where each agent observes only his immediate predecessor. Thus, consider the OIP network where $\mathbb{Q}(B(n) = B_n^1) = 1$ for all $n \in \mathbb{N}$and the sequence of neighbor choice function

$(\gamma_n)_{n\in\mathbb{N}}$ where, for all n, $\gamma_n = n-1$. By Remark 2 and Lemma 6,

$$\begin{aligned}\mathbb{P}_\sigma\left(a_{n+1} \in \arg\max_{x\in X} q_x\right) &\geq \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x\right) \\ &+ \left(1-\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x\right)\right)^2 F_C\left(\left(1-\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x\right)\right)t^{\emptyset}(q_{\max})\right).\end{aligned} \tag{48}$$

Since search costs have a polynomial shape, from (48) we have

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

Now I build on the proof of Proposition 2 in Lobel et al. (2009) to construct the function $\tilde{\phi}$. To apply their construction, the right-hand side of the previous inequality must be increasing in $\mathbb{P}_\sigma(a_n \in \arg\max_{x\in X} q_x)$. This is so under the assumption $0 < L < 2^{K+1}/(K+2)t^{\emptyset}(\underline{q})^K$. Adapting Lobel et al. (2009)'s procedure to my setup gives that the function $\tilde{\phi}$ we are looking for is

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

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

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

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

Under the assumptions of the proposition,

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

because, conditional on observing the same agent $b < n$, agents n and $n+1$ have identical probabilities taking the best action. By Remark 2 and Lemma 6, and since search costs have a polynomial shape, we obtain that

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

Now I build on the proof of Proposition 3 in Lobel et al. (2009) to construct the function $\tilde{\phi}$. To apply their construction, the right-hand side of the previous inequality must be increasing in $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x)$. This is so under the assumption $0 < L < 2^{K+1}/(K+2)t^{\emptyset}(\underline{q})^K$. Adapting

Lobel et al. (2009)'s procedure to my setup gives that the function $\tilde{\phi}$ we are looking for is

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

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

H Proofs for Section 4.2

Let $\delta \in (0,1)$ be the discount factor. Define $t_1(q) := t^{\emptyset}(q)$ for all $q \in Q$. Suppose agent 1 samples first action x with quality q_x. 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 is

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

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

The social utility when each agent observes only his most immediate predecessor is

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

$U_\sigma^1(q_x;\delta)$ 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 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 $c_m \le t_m(q_{\neg x})$ will now sample action x again. Since the quality of action $\neg x$ is unknown, the expected cost of this additional search is $\mathbb{E}_{\mathbb{P}_Q}[\mathbb{E}_{\mathbb{P}_C}[c \mid c \le t_n(q_{\neg x})]F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x]$.

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 4.2.

Lemma 12. *For all $q_x \in Q$ and $\delta \in (0,1)$, we have $U^C_\sigma(q_x;\delta) \geq U^{OIP}_\sigma(q_x;\delta) \geq U^1_\sigma(q_x;\delta)$.*

Finally, let $U^{DM}(q_x;\delta)$ denote the social utility that is implemented by the social planner in any OIP network after sampling action x with quality q_x at the first search in period 1. I Refer to Section III.A. in MFP for the derivation of $U^{DM}(q_x;\delta)$. Since the social planner's problem is the same in all OIP networks, their derivation applies unchanged to my setting.

H.1 Proof of Proposition 2

The difference in average social utilities, $U^C_\sigma(q_x;\delta) - U^1_\sigma(q_x;\delta)$, is

$$U^C_\sigma(q_x;\delta) - U^1_\sigma(q_x;\delta) = (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x) \cdot \sum_{n=1}^{\infty} \delta^n \mathbb{E}_{\mathbb{P}_Q}\left[\mathbb{E}_{\mathbb{P}_C}\left[c \mid c \leq t_n(q_{\neg x})\right] F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\right]\left(1 - \prod_{i=1}^{n-1}\left(1 - F_C(t_i(q_x))\right)\right). \tag{49}$$

As (49) is positive for all $\delta \in (0,1)$, that $U^C_\sigma(q_x;\delta) > U^1_\sigma(q_x;\delta)$ for all $\delta \in (0,1)$ follows. To show that $\lim_{\delta\to 1}[U^C_\sigma(q_x;\delta) - U^1_\sigma q_x;\delta)] = 0$, we need to show that the right-hand side of (49) 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

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

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 3

Suppose $\underline{c} = 0$. By Proposition 2, Lemma 12, and the sandwich theorem for limits, $\lim_{\delta\to 1} U^{OIP}_\sigma(q_x;\delta) = \lim_{\delta\to 1} U^C_\sigma(q_x;\delta)$. By Proposition 3 in MFP, $\lim_{\delta\to 1} U^C_\sigma(q_x;\delta) = \lim_{\delta\to 1} U^{DM}(q_x;\delta)$. That $\lim_{\delta\to 1} U^{OIP}_\sigma(q_x;\delta) = \lim_{\delta\to 1} U^{DM}(q_x;\delta)$ follows by the uniqueness of the limit.

Suppose $\lim_{\delta\to 1} U^{OIP}_\sigma(q_x;\delta) = \lim_{\delta\to 1} U^{DM}(q_x;\delta)$. As the complete network is an OIP network, $\lim_{\delta\to 1} U^C_\sigma(q_x;\delta) = \lim_{\delta\to 1} U^{DM}(q_x;\delta)$. That $\underline{c} = 0$ follows from Proposition 3 in MFP.■

I More than Two Actions

I.1 Collective Search Environment

Suppose the set of actions is $X := \{0, 1, \ldots, N\}$, where $N > 1$. 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 $\omega := (q_0, q_1, \ldots, q_N)$. The model remains otherwise the same, with obvious adjustments.

I.1.1 Equilibrium Strategies

Choice. Agent n takes the best between sampled actions, randomizing uniformly if indifferent.

First Search. 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 $\widehat{B}(n, x)$ sampled action x'. Given I_n^1, agent n computes the conditional probabilities $P_n(x, x') := \mathbb{P}_{\sigma_{-n}}\left(E_n^{x,x'} \mid I_n^1\right)$ 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 can 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: $x \succsim x' \iff P_n(x, x') \leq P_n(x', x)$. By Weitzman (1979)'s and Gergatsouli and Tzamos (2023)'s optimal search rule, at the first search, 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, $k \geq 2$.** Let I_n^k be agent n's information set after sampling $k-1$ actions, S_n^k the set of the first $k-1$ actions sampled by agent n, and $x^* \in \arg\max_{x \in S_n^k} q_x$. Agent n samples an additional action if and only if the expected gain from doing is no less than his search cost.

- If $B(n) = \emptyset$, and $c_n \leq \mathbb{E}_{\mathbb{P}_Q}[\max\{q - q_{x^*}, 0\}]$, agent n samples uniformly at random an action in $X \setminus S_n^k$.
- If $B(n) \neq \emptyset$, agent n benefits from sampling action $x' \in X \setminus S_n^k$ if and only if $E_n^{S_n^k, x'} := \bigcap_{x \in S_n^k} E_n^{x,x'}$ has occurred. Thus, he samples an action in $\arg\max_{x' \in X \setminus S_n^k} \mathbb{P}_{\sigma_{-n}}\left(E_n^{S_n^k, x'} \mid I_n^k\right)$ uniformly at random if $c_n \leq \max_{x' \in X \setminus S_n^k} \mathbb{P}_{\sigma_{-n}}\left(E_n^{S_n^k, x'} \mid I_n^k\right) \mathbb{E}_{\mathbb{P}_Q}[\max\{q - q_{x^*}, 0\}]$.

I.1.2 Maximal Learning

For all $\omega \in \Omega$, let x_ω^* be a fixed element of $\arg\max_{x \in X} q_x$. A searcher takes the best action with probability 1 if and only if $\omega \in \Omega(\underline{c}) := \left\{\omega \in \Omega : F_C\left(t^{\emptyset}(q_x)\right) > 0 \text{ for all } x \neq x_\omega^* \text{ or } q_0 = q_1 = \cdots = q_N\right\}$. Maximal learning occurs if $\lim_{n \to \infty} \mathbb{P}_\sigma(a_n \in \arg\max_{x \in X} q_x \mid \omega \in \Omega(\underline{c})) = 1$.

I.2 Long-Run Learning

I.2.1 Maximal Learning via the Improvement Principle

The IP remains valid with more than two actions. Consider an agent, say n, and his chosen neighbor, say $b < n$. Unless b takes the best action with probability 1, n's expected gain from an additional search is positive. Therefore, if search costs are not bounded away from 0, with positive probability, agent n samples at least two actions and takes an action of better quality than the one he samples first. Since agent n finds it optimal to begin searching from the action taken by agent b, there is a strict improvement in the probability of taking the best action that n has over b. If arbitrarily long information paths almost surely occur, such improvements last until agents take the best action with probability 1. If agents' beliefs about the network conditional on their neighborhood are not too distorted compared to the actual network topology, 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 via the Large-Sample Principle

The key intuition behind the LSP also remains the same with more than two actions. In particular, fix any $\omega \in \Omega(\underline{c})$ and 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' with $q_x > q_{x'}$; (ii) $\alpha^x = \alpha^{x'}$ for all x, x' with $q_x = q_{x'}$. Non-isolated agents in S find it optimal to sample one of the actions with the largest share in their neighborhood at the first search. Thus, by the properties of the network topology, $\lim_{n\to\infty} \mathbb{P}^{\omega}_{\sigma}\left(s^1_n \in \arg\max_{x\in X} q_x \mid n \in S\right) = 1$. Since 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 between more than two actions is not easier than identifying the best between two actions. The next result follows.

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

I.3 Convergence Rate, Welfare, and Efficiency

I.3.1 Convergence Rate

Given a network topology, when more than two actions are available, convergence to the best action occurs more slowly than with two actions. Thus, the upper bounds on the convergence rate in Theorem 4 need no longer apply. However, the insight that the density of indirect connections affects convergence rates remains valid. In particular, with more than two actions, learning remains faster in OIP networks than under uniform random sampling of one past agent. The reason is that the cardinality of agents' subnetworks grows faster in OIP networks, and the same happens to the probability that agents in the subnetwork sample all actions and the inferior ones.

I.3.2 Equilibrium Welfare and Efficiency in OIP Networks

Proposition 2 remains unchanged with more than two actions. In this case, reducing network transparency in OIP networks exacerbates the inefficient duplication of costly search because there are more opportunities to engage in overeager search. The intuition is the following. With more than two actions, multiple actions can be revealed inferior by time n. However, if agent n observes only agent $n-1$, such actions are not revealed 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 infer for free by observing all his predecessor's choices. An immediate consequence is that letting agents observe the aggregate history of past choices play a more central role in reducing inefficiencies when more than two actions are available.

Whether Propositions 1 and 3 hold with more than two actions is unclear. The reason is that establishing these results requires closed-form expressions for equilibrium search policies and the solution to the social planner's problem. With more than two actions, such closed-form expressions are not possible.

J Simple Policies 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 inferior by their predecessors' choices. A simple policy intervention, however, improves equilibrium welfare.

Proposition 6. *For all $\delta \in (0,1)$, the equilibrium social utility in OIP networks is the same as in the complete network if, in addition, agents observe the aggregate history of past choices.*

Proof. First, upon observing the aggregate history of past choices, each agent still samples the action taken by the immediate predecessor first. This follows by an inductive argument as the one proving part (i) of Lemma 11 in Appendix E. Second, if an action is revealed inferior by time n, that action is never sampled again by any agent $m \geq n$. To see why, 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 the immediate predecessor and takes the best one between those he samples, it must be that $a_{n-1} = \neg x$. Upon observing the aggregate history, eveb if $j \notin B(n)$, agent n infers 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 $1 \leq j \leq n-2$, sampled both actions and discarded the inferior action x. Therefore, inefficient duplication of costly searches disappears. ■

The observability of the aggregate history of past choices does not remove the inefficient duplication of costly search arising when agents fail to recognize actions that are inferior—but not revealed so—by some time n. Moreover, it does not incentivize exploration. Thus, agents delay searching for the second action more than the single decision maker, and the convergence rate to the best action remains too slow.

References

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011): "Bayesian Learning in Social Networks," *The Review of Economic Studies*, 78, 1201–1236.

Acemoglu, D. and A. Ozdaglar (2011): "Opinion Dynamics and Learning in Social Networks," *Dynamic Games and Applications*, 1, 1–43.

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

Bala, V. and S. Goyal (1998): "Learning from Neighbours," *The Review of Economic Studies*, 65, 595–621.

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

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

Bigoni, M., M. Boldrini, N. Lomys, and E. Tarantino (2023): "Why Don't You Take a Leaf Out of Her Book? An Experiment on Social Search," *Working Paper.*

Bikhchandani, S., D. Hirshleifer, O. Tamuz, and I. Welch (2022): "Information Cascades and Social Learning," *Working Paper.*

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

Board, S. and M. Meyer-ter Vehn (2021): "Learning Dynamics in Social Networks," *Econometrica*, 89, 2601–2635.

——— (2023): "Experimentation in networks," *Working Paper.*

Bobkova, N. and H. Mass (2022): "Two-Dimensional Information Acquisition in Social Learning," *Journal of Economic Theory*, 202, 105451.

Burguet, R. and X. Vives (2000): "Social Learning and Costly Information Acquisition," *Economic Theory*, 15, 185–205.

Dasaratha, K. and K. He (2023): "Aggregative Efficiency of Bayesian Learning in Networks," *Working Paper.*

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

Gergatsouli, E. and C. Tzamos (2023): "Weitzman's Rule for Pandora's Box with Correlations," *Working Paper.*

Golub, B. and E. Sadler (2016): "Learning in Social Networks," in *The Oxford Handbook of the Economics of Network*, ed. by Y. Bramoullé, A. Galeotti, and B. Rogers, Oxford University Press.

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

Harel, M., E. Mossel, P. Strack, and O. Tamuz (2021): "Rational Groupthink," *The Quarterly Journal of Economics*, 136, 621–668.

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

Lobel, I., D. Acemoglu, M. A. Dahleh, and A. Ozdaglar (2009): “Rate of Convergence of Learning in Social Networks,” *Proceedings of the American Control Conference, ACC 2009.*

Lobel, I. and E. 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, N. and E. Tarantino (2023): “Identification and Estimation in Search Models with Social Information,” *Working Paper.*

Mobius, M. and T. Rosenblat (2014): “Social Learning in Economics,” *Annual Review of Economics*, 6, 827–847.

Monzón, I. and M. Rapp (2014): “Observational Learning with Position Uncertainty,” *Journal of Economic Theory*, 154, 375–402.

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

Perego, J. and S. Yuksel (2016): “Searching for Information and the Diffusion of Knowledge,” *Working Paper.*

Rosenberg, D. and N. Vieille (2019): “On the Efficiency of Social Learning,” *Econometrica*, 87, 2141–2168.

Sadler, E. (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, M. (2017): “Strategic Experimentation in Networks,” *Working Paper.*

Slivkins, A. (2019): “Introduction to Multi-Armed Bandits,” *Foundations and Trends® in Machine Learning*, 12, 1–286.

——— (2023): “Exploration and Persuasion,” in *Online and Matching-Based Markets*, ed. by F. Echenique, N. Immorlica, and V. V. Vazirani, Cambridge University Press.

Smith, L. and P. N. Sørensen (2000): “Pathological Outcomes of Observational Learning,” *Econometrica*, 68, 371–398.

——— (2014): “Rational Social Learning with Random Sampling,” *Working Paper.*

Weitzman, M. L. (1979): “Optimal Search for the Best Alternative,” *Econometrica*, 43, 641–654.