Essays on Dynamic Games

Inauguraldissertation zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften der Universität Mannheim

Niccolò Lomys

Herbstsemester 2018

Abteilungssprecher: Prof. Dr. Jochen Streb
Referent: Prof. Volker Nocke, Ph.D.
Korreferent: Prof. Martin Peitz, Ph.D.

Tag der mündlichen Prufung: 4. September 2018

Contents

To my parents, Simonetta and Luciano.

Acknowledgements

First of all, I wish to express my sincere gratitude to my supervisors Volker Nocke, Martin Peitz, Thomas Tröger, and Emanuele Tarantino for their encouragement and guidance throughout my doctoral studies at the University of Mannheim. I am grateful for the time they invested in advising me, the numerous discussions we had about my work and research in general, and the detailed feedback I received from them. On a less formal note, I would like to thank my supervisors also because of the following reasons: Volker, by masterfully playing the role of the bad cop, has always kept incentives high powered; Martin, with his incredible patience, immensely supported me to approach the job market season; Thomas gave me a stark example of what it means not to ever make a cheap comment; Emanuele did a great job in bridging the gap between me, as a doctoral student, and the more senior faculty members.

I would like to thank the Department of Economics at Yale University, which I visited during my doctoral studies. Taking part to the many activities of the Economic Theory group at Yale University has been a unique chance to be exposed to the research frontier of my field of study. Special thanks go to Johannes Hörner, with whom I had many valuable conversations; I learned a lot from his dedication to research and his commitment to support the efforts of the new generations of scholars.

More generally, I would like to thank the Graduate School of Economic and Social Sciences (GESS) at the University of Mannheim for the many opportunities it provided me with. Moreover, I would like to acknowledge the friendly support I received in uncountably many circumstances from the administrative staff both at the GESS and at the Department of Economics (especially the staff on the 3rd floor).

My life would be way more tedious without my friends. Among them, Lorenzo Magnolfi deserves a special mention. Since we met half a lifetime ago, we never stopped sharing our experiences. Conversing with Lorenzo is always a pleasure and a privilege. Knowing that I could count on his help, support, and advice at anytime has been a great source of comfort during my doctoral studies.

In Mannheim, I have had the fortune to get to know Francesco Paolo Conteduca, Alessandra Allocca, and Giuseppe Corbisiero. Paolo and Alessandra not only are outstanding human beings and extremely caring friends, but they are also unique: Paolo is the only person I have ever met to be as much entertaining as I am; Alessandra is the only person I have ever met to be as much histrionic as I am. Any time they appeared at the door of my office, I knew I was about to have the most joyful half-an-hour of the day. Giuseppe, being both entertaining and histrionic, was simply the cherry on the cake in all of this.

My time in New Haven has been memorable. It would not have been the same without Vittorio Bassi, Federico Esposito, Stefan Lamp, Chiara Margaria, Monica Morlacco, Tommaso Porzio, Camilla

Roncoroni, Gabriella Santangelo, and Benjamin Schickner. Any time we meet, it feels as though we had never been apart.

I would also like to thank Enrico Camarda, Franco Esteban Cattaneo, Ekaterina Kazakova, Marcel Preuss, Xiaodi Wang, and Christoph Wolf; they proved to be not only good office mates (or office neighbors), but also good friends.

Additionally, I owe thanks to Antonio Villanacci and Domenico Menicucci, my supervisors at the Università di Firenze, and to Lorenzo Cocchi, Marco Doretti, Katia Fiorucci, Filippo Macaluso, Niccolò Pagliarini, Alessandro Parrini, and Antonella Sorini, my good old friends outside of the Economics community.[1]

Last but foremost, my deepest gratitude goes to my family. My parents Luciano (il babbone d'oro) and Simonetta (la mammina d'oro) are by far and large the brightest part of my days. Besides that, words cannot describe what they represent for me. Unfortunately, none of my grandparents got to see the completion of my doctoral studies, but they are always present in my thoughts.

[1]Francesco Paolo Conteduca, Lorenzo Magnolfi, Martin Peitz, Camilla Roncoroni, and Emanuele Tarantino also became my coauthors. Collaborating with them is a very enriching experience for me.

Chapter 1

Introduction

This thesis consists of three research papers on dynamic games. All the three papers focus on learning in strategic environments, but each of them is a one-off exploration of a different topic.

Chapter 2. Sequential Collective Search in Networks. In this paper, I investigate the impact of costly information acquisition on the process of social learning. Social learning is the study of how individuals combine their private information with others' experiences to identify the best course of action in the face of payoff-relevant uncertainty. In most circumstances of social and economic interest—from product markets to technology adoption—private information only becomes available at a cost. When others' experiences are available, agents' incentives to collect the relevant information are ambiguous. On the one hand, the availability of others' experiences weakens individual motivation to acquire independent information and encourages the exploitation of others' wisdom, increasing the chances of wrong herds. On the other hand, the possibility of wrong herds fosters independent exploration, reducing the odds of suboptimal behavior.

I propose a model of social learning in networks to understand this trade-off. In particular, I explore how the structure of social ties (the network topology) shapes individual incentives to acquire independent information and, subsequently, the diffusion of newly created knowledge. To do so, I study a setup where countably many rational agents act in sequence, observe the choices of their connections, and acquire private information via costly sequential search. The sequential nature of the problem allows for a rich characterization of the properties of search technologies and network topologies under which different positive learning results obtain or fail. The framework also allows for welfare analysis, the study of convergence rates, and the discussion of policy interventions to alleviate the inefficiency of equilibrium outcomes.

I characterize perfect Bayesian equilibria of the model by linking individual search policies to the probability that agents select the best action. The information structure of the model precludes information aggregation via martingale convergence arguments. Besides the technical challenges it raises, the (negative) result prevents agents from learning via the aggregation of the information contained in large samples of other agents' choices. If search costs are not bounded away from zero, however, imitation, paired with some amount of individual improvement upon it, is sufficient for agents to learn how to select the best action in the long run. In fact, when search costs are not bounded away from zero, asymptotic learning occurs in sufficiently connected networks where the structure of individual connections does not lead agents astray about the broader network realization. In such networks, agents can identify the correct social connections to rely on and improvements last long enough for agents to select the best action at the first search. Learning via improvements upon imitation, however, is fragile: it breaks down as soon as zero is removed from the support of the search cost distribution. When search costs are bounded away from zero, even a weaker notion of long-run learning fails, except in ad hoc network topologies. Networks where agents observe the choices of random numbers of immediate predecessors share many equilibrium properties with the complete network, including the rate of convergence and the probability of wrong herds. Transparency of past histories has short-run, but not long-run, implications for welfare and efficiency. The simple policy intervention of letting agents observe the relative fraction of previous choices reduces inefficiencies and welfare losses.

The paper contributes to both the economic theory of social learning and its applications to the economics of social media and Internet search. The theoretical novelty of the paper is to analyze costly information acquisition in a model of rational learning over general networks. In turn, the information acquisition technology—sequential search, which has received much attention in the applied literature—naturally relates the model to a variety of applications. In particular, the model speaks to the large evidence that people's online behavior—what they search on web search engines, the order in which they do so, and their resulting purchase decisions—is often inspired by what they observe on social media. Thus, the paper sheds light on the implications of such behavior for social learning, product diffusion and demand, and on the forces that may lead consumers to herd on inferior items.

Chapter 3. Learning While Bargaining: Experimentation and Coasean Dynamics. Bargaining is ubiquitous. Many economic interactions involve negotiations on a variety of issues. For example, prices of commodities are often the outcome of negotiations between the concerned parties, wages are set as an arrangement between firms and workers, and takeovers require an agreement over the price of the transaction. As such, bargaining relationships are the cornerstone of many theory of markets, from industrial organization to labor economics. Classical models of bargaining with incomplete information are typically presented as bilateral monopolies. Such models posit common knowledge of gains from trade and assume the relevant information to reach an agreement to be available to parties—perhaps asymmetrically—since the outset of their negotiations. Yet, in many real-world bargaining situations, superior outside opportunities may become available to either or both parties during their negotiations—parties routinely investigating what their best opportunities are, as a large literature on search documents.

Motivated by these considerations, I develop a framework to understand bargaining relationships in such an environment in which there is uncertainty about whether and when superior outside opportunities are available and parties may want to wait to reach an agreement in order to learn about their best opportunities during negotiations. In particular, I study dynamic bilateral bargaining with one-sided incomplete information when superior outside opportunities may arrive during negotiations. Gains from trade are uncertain: in a good-match market environment, outside opportunities are not available; in a bad-match market environment, superior outside opportunities stochastically arrive for either or both parties. The two parties begin their negotiations with the same belief on the type of the market environment. Arrivals are public and learning about the market environment is common. One party, the seller (he), makes price offers at every instant to the other party, the buyer (she). The seller has no commitment power and the buyer is privately informed about her own valuation.

I show that the option value of waiting to learn about the existence of better opportunities is of first-order importance in shaping the bargaining relationship. It affects the timing of agreements, the dynamics of prices, surplus division, and the seller's ability to exercise market power. In equilibrium, there is either an initial period with no trade or trade starts with a burst. Afterward, the seller screens out buyer types one by one as uncertainty about the market environment unravels. Delay is always present, but it is inefficient only if valuations are interdependent. Whether prices increase or decrease over time depends on which party has a higher option value of learning. The seller may exercise market power. In particular, when the seller can clear the market in finite time at a positive price, prices are higher than the competitive price. Market power, however, need not be at odds with efficiency.

The model has a number of applications, including durable-good monopoly without commitment, wage bargaining in markets for skilled workers, and takeover negotiations. On the methodological side, posing the model in continuous time not only simplifies the analysis, but also allows for additional economic insights. Continuous time captures the idea that there are no institutional frictions in the bargaining protocol (besides incomplete information). Thus, my analysis clearly disentangles the effect of learning about the market environment on equilibrium outcomes from that of other frictions in the protocol. In addition, closed-form expressions for all the relevant equilibrium outcomes of the game open the doors to comparative statics as well as to empirical studies and more applied research.

Chapter 4. Dynamic Foundations for Empirical Static Games. This paper is joint work with Lorenzo Magnolfi and Camilla Roncoroni. We propose a simple estimation strategy when data on strategic interaction are interpreted as the long-run result of a history of game plays. Players interact repeatedly in an incomplete information game, possibly while learning how to play in such a game. We remain agnostic on the details of the learning process and only impose a minimal behavioral assumption describing an optimality condition for the long-term outcome of players' interaction. In particular, we assume that play satisfies a property of "asymptotic no regret" (ANR). This property requires that the time average of the counterfactual increase in past payoffs, had different actions been played, becomes approximately zero in the long run. The ANR property is satisfied by a large class of well-known algorithms for the repeated play of the underlying one-shot game, once they are appropriately extended to games of incomplete information.

We show that, under the ANR assumption, it is possible to partially identify the structural parameters of players' payoff functions. We establish our result in two steps. First, we prove that the time average of play that satisfies ANR converges to the set of Bayes correlated equilibria of the underlying static game. To do so, we extend to incomplete information environments prior results on dynamic foundations for equilibrium play in static games of complete information. Second, we show how to use the limiting model to obtain consistent estimates of the parameters of interest. Our approach gives rise to non-standard econometric issues, as it is not possible to fully characterize a single limit distribution of the observables, but only the set it belongs to. Yet, we show that we can use the limiting model to obtain a consistent estimator for the parameters of interest.

The ANR property is weaker than the one-shot no-ex post-regret property of pure-strategy Nash equilibrium that is sometimes invoked to motivate the choice of modeling cross-sectional data as equilibrium outcomes of a static game. Indeed, this descriptive interpretation of static models is often paired with the assumptions of complete information and pure-strategy Nash equilibrium. The rationale for these assumptions is that the no-ex post-regret property of pure-strategy Nash equilibria reflects the stable nature of long-run outcomes. Although appropriate for some environments, the static notion of no-ex post regret is a strong requirement: our work is thus complementary to standard equilibrium models of strategic interaction and provides an alternative whenever Nash equilibrium does not represent an appropriate restriction on behavior. In fact, Nash equilibrium of the static game is neither a natural long-run outcome of many simple game dynamics, nor easy to compute in large games.

Chapter 2

Sequential Collective Search in Networks[1]

2.1 Introduction

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

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

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

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

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

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

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

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

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

Acemoglu et al. (2011) and Lobel and Sadler (2015) originally develop an IP for the standard sequential social learning model (henceforth, SSLM) to establish positive learning results in stochastic networks. In the SSLM agents receive a free private signal, which is informative about the relative quality of all alternatives, and wish to match their action with an unknown state of nature.[2] My results extend the reach of the IP and of the learning principle it captures to a new informational environment, which departs from that of the SSLM in three relevant aspects. First, private information is generated by equilibrium play rather than being exogenously given to the agents. Second, while in the SSLM agents have imperfect private information about the relative quality of the two alternatives, in my model sampling an action perfectly reveals the quality of that action only. Finally, the inferential challenge crucially differs: agents maximize the value of a sequential information acquisition program rather than the probability of matching an underlying state of nature or an ex ante expected utility. The possibility to describe agents' sequential search policies in terms of probabilities, however, bridges the search setting I study to the SSLM. Thus, an IP holds in the two settings in spite of limited comparability of their informational environments.[3]

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

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

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

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

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

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

From the viewpoint of selecting the best action, individual search behavior in networks where agents observe the choices of random numbers of immediate predecessors is equivalent to the search behavior in the complete network. These network topologies thus inherit several equilibrium properties from the complete network, including the probability of wrong herds and the speed of learning, which is faster than polynomial.[4] Reducing transparency of past histories, however, leads to inefficient duplication of costly search. I compare equilibrium welfare in the complete network and in the network where each agent only observes his most recent predecessor. The difference only vanishes in the limit of an infinitely patient society but is significant in the short and medium run. Simple policy interventions, such as letting agents observe the relative share of previous choices in addition to their neighbors' choices, reduce inefficiencies and welfare losses.

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

A compelling motivation for my model comes from the large evidence that people's online behavior—what they search on web search engines, the order in which they do so, and their resulting purchase

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

decisions—is often inspired by what they observe on social media. For instance, suppose we need to decide which of two recently released comedies to watch. The two movies have a cast and a direction of comparable reputation so that it is ex-ante unclear which one is better. However, we observe on Facebook the movie our friends watched through their check-ins or the Facebook pages they liked, but only have a vague idea of whom they observed in turn. Our friends' decisions give us a first impression of what film is likely to be the best one. We then search on Google for this movie to learn where and when it is played and to read experts' reviews. Looking for movie times and reading reviews takes time and effort, and this idiosyncratic cost depends on factors that are our private information (whether we are in a rush, how much time we can divert from other activities, etc.). Depending on movie times, reviews, and our opportunity cost, we either watch the movie we first learned about, or invest more time searching for information about the other option.[5]

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

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

Road Map. In Section 2.2, I describe the model. In Section 2.3, I define asymptotic learning and characterize equilibrium strategies. In Section 2.4, I establish the improvement principle and the main results on asymptotic learning. In Section 2.5, I introduce maximal learning and present the main results with respect to this metric. In Section 2.6, I present the main results on the rate of convergence, welfare, and efficiency. In Section 2.7, I discuss the related literature and conclude. Supporting examples are in Appendix 2.8 and formal proofs are in Appendix 2.9.

2.2 Model

2.2.1 Collective Search Environment

Agents and Actions. A countably infinite set of agents, indexed by $n \in \mathbb{N} := \{1, 2, \dots\}$, sequentially select a single action each, with agent n acting at time n. Each agent has to choose one of two possible alternatives in the set of available actions $X := \{0, 1\}$, which is identical across agents. Restricting

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

attention to two actions simplifies the exposition, but does not affect the results. A typical element of X is denoted by x, while the action agent n selects is denoted by a_n. Calendar time is common knowledge and the order of moves exogenous.

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

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

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

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

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

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

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

Search Technology. Private information about the quality of the two actions is acquired through costly sequential search with recall. After observing his neighborhood $B(n)$ and the actions of the agents in $B(n)$, agent n decides which action $s_n^1 \in X$ to sample first.[6] Sampling an action perfectly reveals its quality to the agent. I denote the quality of the first action sampled by agent n as $q_{s_n^1}$. After observing $q_{s_n^1}$, agent n decides whether to sample the remaining action, $s_n^2 = \neg s_n^1$, where $\neg s_n^1$ denotes the action in X not sampled initially, or to discontinue searching, $s_n^2 = ns$. That is, $s_n^2 \in \{\neg s_n^1, ns\}$. Let S_n denote the set of actions agent n samples. After sampling has stopped, the agent chooses an action a_n. Agents can only select an action they sampled, that is $a_n \in S_n$. Thus, for a single agent the model of search is that of Weitzman (1979), and proposed by Mueller-Frank and Pai (2016) to study observational learning in the complete network.

For simplicity, the first action is sampled at no cost, while sampling the second action involves a cost $c_n \in C \subseteq \mathbb{R}_+$.[7] Search costs c_n are i.i.d. across agents, are drawn from a commonly known probability measure $\mathbb{P}_C$ over C, with associated CDF F_C, and are independent of the network topology and the quality of the two actions. I refer to the probability space $(C, \mathcal{F}_C, \mathbb{P}_C)$, together with the sequential search rule, denoted by $\mathcal{R}$, as the *search technology* of the model. An agent's search cost and sampling decisions are his private information. That is, for all $n \in \mathbb{N}$, agent n's search cost c_n and sampling decisions are not observed by later moving agents.

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

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

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

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

2.2.2 Information and Strategies

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned} &\sigma_n^1 \colon \mathcal{I}_n^1 \to \Delta(\{0,1\}),\\ &\sigma_n^2 \colon \mathcal{I}_n^2 \to \left(\left\{\neg s_n^1, ns\right\}\right),\\ \text{and} \quad &\sigma_n^a \colon \mathcal{I}_n^a \to \Delta(S_n). \end{aligned}$$

A strategy profile is a sequence of strategies $\sigma := (\sigma_n)_{n\in\mathbb{N}}$. Let $\sigma_{-n} := (\sigma_1, \ldots, \sigma_{n-1}, \sigma_{n+1}, \ldots)$ denote the strategies of all agents other than n. Given a collective search environment $\mathcal{S}$ and a strategy profile σ, the sequence of actions $(a_n)_{n\in\mathbb{N}}$ is a stochastic process with probability measure $\mathbb{P}_\sigma$ generated by the state process, the network topology, the search technology, and the mixed strategy of each agent. Formally, for a fixed σ, the sequence $(a_n)_{n\in\mathbb{N}}$ is determined by the realization in the probability space[8] $Y := \Omega \times \mathbb{B} \times C^\infty \times D^\infty$. Here, C^∞ is the set of possible realizations of search costs for each agent, $(D, \mathcal{F}_D, \lambda)$ is a probability space determining the possible mixed strategy realizations of a given agent, and Ω and $\mathbb{B}$ have been introduced before.

2.2.3 Equilibrium Notion

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

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

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

In any collective search environment $\mathcal{S}$, given a strategy profile for the agents acting prior to n, and a realization of n's information sets $I_n^r \in \mathcal{I}_n^r$ for $r \in \{1, 2\}$ and $I_n^a \in \mathcal{I}_n^a$, the decision problems of agent n at the search stages and at the choice stage are discrete choice problems. Therefore, they have

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

a well-defined solution that only requires randomizing according to some mixed strategy in case of indifference at some stage (see Section 2.3.2 for a characterization of individual equilibrium decisions). For given criteria to break ties, an inductive argument shows that the set of equilibria $\Sigma_{\mathcal{S}}$ is nonempty. I note the existence of equilibrium here.

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

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

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

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

2.3 Long-Run Learning and Equilibrium Strategies

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

2.3.1 Asymptotic Learning: Definition

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

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

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

Studying asymptotic learning requires understanding how the quantity

$$\mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X}\, q_x\right) \tag{2.1}$$

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

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

a problem which need not be equivalent to maximizing the quantity in (2.1) or the ex ante expected utility.[9] This discrepancy raises some conceptual challenges one needs to address before establishing the main results. To this purpose, the next subsection characterizes equilibrium search policies by linking the dynamics of agents' optimization to the dynamics of the quantity $\mathbb{P}_\sigma(a_n \in \arg\max_{x\in X} q_x)$, thus making the analysis of long-run outcomes possible.

2.3.2 Equilibrium Strategies

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

Preliminaries

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

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

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

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

Characterization of Equilibrium Sequential Search Policies

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

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

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

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

[9] An analogous remark applies to maximal learning, introduced in Section 2.5.

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

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

$$t^\emptyset\big(q_{s_n^1}\big) := \mathbb{E}_{\mathbb{P}_Q}\big[\max\big\{q - q_{s_n^1}, 0\big\}\big] = \int_{q \geq q_{s_n^1}} \big(q - q_{s_n^1}\big) \mathrm{d}\mathbb{P}_Q(q). \tag{2.2}$$

If $c_n = t^\emptyset(q_{s_n^1})$, agent n is indifferent between searching further or not. Again, his strategy is described by two non-negative functions, $\pi_n^{\neg s_n^1}(\cdot)$ and $\pi_n^{ns}(\cdot)$, such that $\pi_n^{\neg s_n^1}(I_n^2) + \pi_n^{ns}(I_n^2) = 1$ for all $I_n^2 \in \mathcal{I}_n^2$ with $B_n = \emptyset$. Here, $\pi_n^{\neg s_n^1}(I_n^2)$ ($\pi_n^{ns}(I_n^2)$) is the probability that agent n with information set I_n^2 samples (does not sample) action $\neg s_n^1 \in X$.[10]

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

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

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

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

$$E_n^x := \Big\{ y \in Y : s_k^2 = ns \text{ for all } k \in \widehat{B}(n, x) \Big\} \qquad \text{for } x = 0, 1. \tag{2.3}$$

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

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

$$P_n(x) := \mathbb{P}_{\sigma_{-n}}\big(E_n^x \mid I_n^1\big) \qquad \text{for } x = 0, 1. \tag{2.4}$$

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

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

$$\mathbb{P}_Q(q_x \le q) = \mathbb{P}_Q(q_{\neg x} \le q), \tag{2.5}$$

$$\mathbb{P}_{\Omega|q_{\neg x} \ge q_x}(q_x \le q) = \mathbb{P}_{\Omega|q_x \ge q_{\neg x}}(q_{\neg x} \le q), \tag{2.6}$$

and

$$\mathbb{P}_{\Omega|q_{\neg x} \ge q_x}(q_x \le q) > \mathbb{P}_Q(q_x \le q). \tag{2.7}$$

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

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

Here, E_n^{xC} ($E_n^{\neg xC}$) is the complement of E_n^x ($E_n^{\neg x}$), the third equality holds by (2.5) and (2.6), and the inequality follows from (2.7) and the assumption $P_n(x) < P_n(\neg x)$.

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

$$P_n\big(q_{s_n^1}\big) := \mathbb{P}_{\sigma_{-n}}\big(E_n^{s_n^1} \mid I_n^2\big), \tag{2.8}$$

where also the information about the quality of the first action sampled is used. With remaining probability, at least one of those agents sampled action $\neg s_n^1$, but nevertheless chose action s_n^1, in which case s_n^1 is (weakly) superior by revealed preferences. Agent n's expected benefit from sampling action $\neg s_n^1$ is therefore $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$, where $t^{\emptyset}(\cdot)$ is defined by (2.2) and describes the gross benefit of the second search (the benefit agent n would have if he did not observe any other agent) when a payoff of $q_{s_n^1}$ has already been secured. It follows that he should only sample further if his search cost c_n is less than $t_n(q_{s_n^1})$, where the function $t_n \colon Q \to \mathbb{R}_+$ is defined pointwise as

$$t_n\big(q_{s_n^1}\big) \coloneqq P_n\big(q_{s_n^1}\big)\, t^{\emptyset}\big(q_{s_n^1}\big). \tag{2.9}$$

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

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

Discussion of Equilibrium Behavior

Remark 1. For all $n \in \mathbb{N}$, agent n's equilibrium sequential search policy is essentially described by the probabilities $P_n(x)$ and $P_n(q_x)$, for all $x \in X$ and $q_x \in Q$, defined by (2.4) and (2.8). This characterization relates the dynamics of agents' optimization to the dynamics of the probability that they select the correct action. Roughly, the intuition is the following.[11] For all $n \in \mathbb{N}$,

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

Here, the first inequality holds as agent n takes the action of better quality among those he has sampled. The second inequality follows because if an agent in $\widehat{B}(n, s_n^1)$ samples both actions and takes action s_n^1, then s_n^1 is superior by revealed preferences. In turn, the first equality holds as the two events at issue are one the complement of another, and the second equality holds by definition of $E_n^{s_n^1}$ (see (2.3)). This link unravels the complications illustrated at the end of Section 2.3.1 and will prove a central tool to establish long-run learning results in the analysis to come.

Remark 2. Each agent faces a three-way trade-off between *exploration* (sampling the second action), *exploitation* (using the information revealed by others' choices to save on the cost of the second search),

[11] I refer to Appendix 2.9 for the formal details.

and *individual incentives* (agents are myopically interested in exploiting the wisdom of their neighbors). The characterization of the optimal search policies sheds light on how such trade-off is resolved in equilibrium.

First, (2.2) and (2.9) imply $t_n(q) \leq t^{\emptyset}(q)$ for all $q \in Q$, as $P_n(q) \in [0,1]$. That is, given the quality of the first action sampled, the expected additional gain from the second search is lower for an agent with nonempty neighborhood than for an agent with empty neighborhood. Thus, if an agent with search cost type c and empty neighborhood discontinues search after sampling an action of quality q, so does an agent with the same search cost type and nonempty neighborhood after sampling an action of the same quality. In short, agents with no neighbors have stronger incentives to explore than agents who exploit the information revealed by their neighbors' choices.

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

Finally, the quality of the first action sampled has ambiguous effects on the incentives to explore of an agent, say n, with nonempty neighborhood. This is so because n's expected additional gain from the second search, $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$, depends on the probability $P_n(q_{s_n^1})$ that none of the agents in his personal subnetwork relative to action s_n^1 has sampled action $\neg s_n^1$ given that the quality of s_n^1 is $q_{s_n^1}$. This probability need not be monotonic in $q_{s_n^1}$ and depends on the network topology as well as on the properties of the state process and the search technology. On the one hand, an action of high quality suggests that some individual has explored both feasible alternatives, discarding the one with low quality to adopt the superior one. On the other hand, precisely this effect, combined with the fact that $t^{\emptyset}(q)$ decreases in q, hints that the incentives to acquire information about the second action (exploit the information revealed by others' choices) decrease (increase) with the quality of the first action sampled. This is the central trade-off in the environment I study. Which force prevails, and so the effect of an increase in the quality of s_n^1 on $P_n(q_{s_n^1})$ and, ultimately, on $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$, is unclear. In Appendix 2.8, I construct two examples to show that $P_n(q_{s_n^1})t^{\emptyset}(q_{s_n^1})$ can either increase or decrease as $q_{s_n^1}$ increases depending on the primitives of the model.

Remark 3. In general network topologies, there is no informational monotonicity property linking an agent's equilibrium behavior to the relative fraction of actions he observes or to the actions of his most recent neighbors. This feature is common in models departing from the assumption that agents observe the full history of past actions and motivates the approach I adopt to establish positive learning results in Section 2.4.

Remark 4. In the collective search environments I study, there is no social belief that is a martingale and, at the same time, is of some use when characterizing equilibrium behavior. Thus, martingale convergence arguments, which are standard tools to study aggregation of dispersed information in social learning settings, have no bite in the present setup. As I will formalize in Section 2.5.6, this feature undermines the possibility to learn via the direct observation of large samples of other agents and the aggregation of the information that their choices convey.

2.4 Asymptotic Learning

In a collective search environment, the search technology shapes agents' possibility to acquire independent private information and the network topology shapes agents' possibility to learn by observing others' behavior. In this section and in Section 2.5, I provide conditions on these primitives under which (different) positive learning results obtain or fail in the long run.

2.4.1 Preliminaries

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

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

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

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

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

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

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

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

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

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

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

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

2.4.2 Sufficient Conditions

For asymptotic learning to occur it is key that costs are not bounded away from zero. Under this premise, I show that an improvement principle holds in the present setup despite the informational environment significantly differ from that of the SSLM. This is the main contribution of Section 2.4. Then, I leverage the improvement principle to show that asymptotic learning obtains if, in the network topology, arbitrarily long information paths occur almost surely and are identifiable.[12]

Improvement Principle

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n))\right) > \mathcal{Z}\left(\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x\in X} q_x\right)\right) - \varepsilon. \tag{2.10}$$

Then, asymptotic learning occurs in equilibrium σ*.*[13]

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

Condition *(c)* in Proposition 2 requires the existence of a strict lower bound on the increase in the probability that an individual will sample first the best action over his chosen neighbor's probability except, possibly, for neighbors that γ_n selects with vanishingly small probability. Therefore, for an

[13]The probabilities in (2.10), and in (2.11) below, are random variables.

improvement principle to hold, one must be able to construct a suitable improvement function $\mathcal{Z}$. The next proposition shows that this is possible if search costs are not bounded away from zero. The intuition goes as follows. Consider an agent, say n, and his chosen neighbor, say $b < n$. Unless b samples first the best action with probability one, b's expected additional gain from the second search is positive. Therefore, if search costs are not bounded away from zero, b samples both actions and compares their quality with positive probability. Thus, as b always takes the best action among those he samples, there is a positive probability that the action he takes is of better quality than the one he samples first. Since n finds it optimal to start searching from the action taken by b,[14] this results in a strict improvement in the probability of sampling first the best action that agent n has over his chosen neighbor b, unless b already does so with probability one, in which case the improvement is non-negative.

Proposition 3. *Suppose that the search technology has search costs that are not bounded away from zero, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence of neighbor choice functions. Then, there exists an increasing and continuous function $\mathcal{Z}\colon [1/2,1] \to [1/2,1]$, satisfying $\mathcal{Z}(\beta) > \beta$ for all $\beta \in [1/2,1)$, $\mathcal{Z}(1) = 1$, and such that*

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

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

The improvement principle is introduced by Acemoglu et al. (2011) in the SSLM as a tool to establish positive learning results in network topologies with independent neighborhoods. Lobel and Sadler (2015) generalize this principle to networks with arbitrarily correlated neighborhoods. In this paper, I extend the scope of the improvement principle to a new environment, where private information is endogenous and fundamentally distinct in nature, which leads to a different inferential problem on the agents' side. Section 2.3.2, however, shows that the equilibrium sequential search policies are essentially described by the probabilities $P_n(x)$ and $P_n(q_x)$ defined in (2.4) and (2.8). This characterization relates the dynamics of individual search behavior to the evolution of the quantity $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x)$. This link makes the agents' inference somewhat comparable to the one faced by the agents in the SSLM, despite the very different premises on the information structure. Thus, an improvement principle which is close in spirit holds.

Sufficient Conditions for Asymptotic Learning

To connect Propositions 2 and 3 into a general result, one needs to bound the difference between $\mathbb{P}_\sigma(s_{\gamma_n}^1 \in \arg\max_{x\in X} q_x)$ and $\mathbb{P}_\sigma(s_{\gamma_n}^1 \in \arg\max_{x\in X} q_x \mid \gamma_n)$. Agent n can imitate agent γ_n only if $\gamma_n \in B(n)$. Therefore, if neighborhoods are correlated, agent γ_n's probability of sampling first the best action conditional on agent n observing agent γ_n is not the same as agent γ_n's probability of sampling first the best action. That is, by imitation, agent n earns γ_n's probability of sampling

[14] It is intuitive, and formally proven in Appendix 2.9.1, that, when agent n only relies on agent b disregarding what other agents have done, the marginal distribution of the quality of the action taken by b first-order stochastically dominates the marginal distribution of the quality of the other action in the eyes of n.

first the best action *conditional* on n choosing to imitate agent γ_n. If $\mathbb{P}_\sigma(s^1_{\gamma_n} \in \arg\max_{x \in X} q_x)$ and $\mathbb{P}_\sigma(s^1_{\gamma_n} \in \arg\max_{x \in X} q_x \mid \gamma_n)$ are approximately the same for large n, then Propositions 2 and 3 immediately imply asymptotic learning. In other words, long information paths must be identifiable, in the sense that agents along the path need reasonably accurate information about the network realization. The next theorem formalizes this last step, which is standard from prior work (see, in particular, Golub and Sadler (2016)).

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

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

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

$$\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x \in X} q_x \;\Big|\; \gamma_n(B(n))\right) > \mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x \in X} q_x\right) - \varepsilon. \tag{2.11}$$

Then, asymptotic learning occurs in equilibrium σ.

A variety of conditions on the network topology of $\mathcal{S}$ ensure that (2.11) holds in every equilibrium $\sigma \in \Sigma_{\mathcal{S}}$. In such cases, if search costs are not bounded away from zero and there exists a chosen neighbor topology with expanding subnetworks, we say that asymptotic learning occurs in the collective search environment $\mathcal{S}$. Such conditions have been identified by Acemoglu et al. (2011) and Lobel and Sadler (2015), to which I refer for further details.

The improvement principle not only serves as a proof technique, but also as a learning principle when standard informational monotonicity properties do not hold (cf. Remark 3 in Section 2.3.2). In particular, it captures the idea that a boundedly rational procedure, imitation, combined with some amount of individual improvement upon it, is sufficient to achieve positive learning outcomes in the long run. The improvement principle also lies behind information diffusion in the SSLM.[15] This explains why, in the search setting I study, asymptotic learning occurs in networks where information diffuses in the SSLM, albeit the mechanics behind the two models significantly differ. Namely, in these network topologies the heuristic captured by the improvement principle displays good long-run properties: first, identifiable information paths allow agents to pick the right neighbor to imitate; second, long information paths allow improvements to last as long as they are possible given the information structure of the model.

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

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

2.4.3 A Necessary Condition on Network Topologies

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

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

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

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

2.5 Maximal Learning

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

2.5.1 A Motivating Example

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

Example 1. Suppose that the qualities of the two actions are are drawn uniformly at random from $\{0, 1/2, 1\}$ and that the lowest cost in the support of the search cost distribution is $\underline{c} > 1/6$. With probability $2/9$, the realized quality of the two actions is $(q_0, q_1) \in \{(1/2, 1), (1, 1/2)\}$. In such cases, in equilibrium an agent with no neighbors and search cost type $\underline{c}$ never samples the second alternative whatever action he samples first, as his expected additional gain from the second search is at most $1/3(1 - 1/2) = 1/6$, which is smaller than his search cost. However, this agent only samples the best action at the first search with probability $1/2$. In turn, agents with a higher search cost type and/or nonempty neighborhood do not sample the second action either, independently of which action they sample first (see Section 2.3.2). Therefore, when $(q_0, q_1) \in \{(1/2, 1), (1, 1/2)\}$, each agent in the social network makes the wrong choice with positive probability. ■

2.5.2 Maximal Learning: Definition

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

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

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

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

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

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

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

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

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

maximal learning occurs if

$$\lim_{n\to\infty} \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x \;\middle|\; \omega \notin \overline{\Omega}(\underline{c})\right) = 1. \tag{2.12}$$

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

The next assumption, which parallels Assumption 1, is maintained throughout Section 2.5.

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

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

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

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

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

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

2.5.3 Failure of the Improvement Principle

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

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

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

The improvement principle is not the only method agents may use to learn how to select the correct action. Therefore, it is natural to inquire whether there exist network topologies where maximal learning never obtains (i.e., no matter what agents do in order to learn) when search costs are bounded away from zero. Section 2.5.5 addresses this question.

2.5.4 OIP Networks

Before stating the next results, I introduce some notation and define a class of network topologies which will be extensively discussed in Section 2.6 as well. For all $n \in \mathbb{N}$ and $l_n \in \mathbb{N}_n$, let

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

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

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

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

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

Example 2. Here are some examples of OIP networks.

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

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

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

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

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

2.5.5 Failure of Maximal Learning

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

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

(a) Observation of immediate predecessors;

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

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

The intuition behind the result is the following. Suppose that the lowest cost in the support of the search cost distribution is $\underline{c} > 0$ and that $\omega \notin \overline{\Omega}(\underline{c})$. By way of contradiction, assume that maximal learning occurs, so that the probability that none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions converges to zero as n grows large. Then, for a sufficiently late moving agent, say m, this probability is so small that the expected additional gain from the second search falls below $\underline{c} > 0$ and remains below this threshold afterward. As a result, no agent moving after agent m will sample the second action. At the same time, however, by Assumption 2, the probability that none of the agents in $\widehat{B}(m) \cup \{m\}$ samples both actions is positive for all finite m. But then, this is a contradiction with maximal learning, as the probability that none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions remains bounded away from zero.

The negative result on maximal learning extends beyond the observation structures in Theorem 2. For instance, maximal learning fails in OIP networks if, in addition, agents observe the choices of the first K agents or the aggregate history of past actions (see Sections 2.6.1 and 2.6.4); it also fails when each agent n samples $M > 1$ agents uniformly and independently from $\{1, \dots, n-1\}$.

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

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

2.5.6 Maximal Learning and the Large-Sample Principle

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

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

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

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

Suppose $(q_0, q_1) \notin \overline{\Omega}(\underline{c})$ and, without loss, $q_0 > q_1$. Consider first an agent, say k, with $B(k) = \emptyset$. By definition of $\overline{\Omega}(\underline{c})$ and $\underline{c}$, k samples the second action with positive probability when he samples action 1 first. Hence, k takes the correct action ($a_k = 0$) with probability $\alpha > 1/2$.[16]

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

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

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

and $s_l^1 \in \Delta(\{0,1\})$ if $|\widehat{B}(l,0)| = |\widehat{B}(l,1)|$. To see this, note that $|\widehat{B}(l,x)| > |\widehat{B}(l,\neg x)|$ immediately implies $P_l(x) < P_l(\neg x)$, where $P_l(\cdot)$ is the probability defined by (2.4).

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

$$\lim_{l\to\infty} \mathbb{P}_\sigma\left(\left|\widehat{B}(l,0)\right| > \left|\widehat{B}(l,1)\right| \;\middle|\; \widehat{B}(l) \neq \emptyset\right) = 1. \tag{2.13}$$

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

$$\begin{aligned}
1 &\geq \mathbb{P}_\sigma\left(a_n \in \arg\max_{x\in X} q_x \;\middle|\; \omega \notin \overline{\Omega}(\underline{c})\right) \\
&\geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \omega \notin \overline{\Omega}(\underline{c})\right) \\
&= \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; B(n) = \emptyset, \omega \notin \overline{\Omega}(\underline{c})\right)\mathbb{Q}(B(n) = \emptyset) \\
&+ \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; B(n) \neq \emptyset, \omega \notin \overline{\Omega}(\underline{c})\right)\mathbb{Q}(B(n) \neq \emptyset) \\
&\geq \frac{1}{2}p_n + \mathbb{P}_\sigma\left(\left|\widehat{B}(n,0)\right| > \left|\widehat{B}(n,1)\right| \;\middle|\; \widehat{B}(l) \neq \emptyset\right)(1-p_n).
\end{aligned} \tag{2.14}$$

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

By (2.13), and since $\lim_{n\to\infty} p_n = 0$, we have

$$\lim_{n\to\infty}\left[\frac{1}{2}p_n + \mathbb{P}_\sigma\left(\left|\widehat{B}(l,0)\right| > \left|\widehat{B}(l,1)\right|\right)(1-p_n)\right] = 1. \tag{2.15}$$

Together, (2.14) and (2.15) imply

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

showing that maximal learning occurs. ■

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

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

Recall that maximal and asymptotic learning sometimes coincide despite search costs are bounded away from zero (see Section 2.5.2). Thus, Example 3 also shows that asymptotic learning may occur when zero is not in the support of the search cost distribution. In other words, search costs that are not bounded away from zero are not, in general, necessary for asymptotic learning.

2.6 Rate of Convergence, Welfare, and Efficiency

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

2.6.1 Equilibrium Strategies in OIP Networks

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

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

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

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

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

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

In OIP networks, agent $n \geq 2$'s equilibrium behavior is the following.[17] At the first search stage, agent n samples the action taken by his immediate predecessor: $s_n^1 = a_{n-1}$. This is so because n's belief about the quality of action a_{n-1} strictly first-order stochastically dominates his belief about the quality of the other action (the result follows by induction). Hence, if an action is revealed to be inferior by time n in equilibrium σ, it is also inferior by time n in equilibrium σ (the converse statement is not, in general, true).

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

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

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

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

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

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

[17] I refer to Appendix 2.9.3 for the formal characterization.

2.6.2 Probability of Wrong Herds and Rate of Convergence

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

Remark 7. In all OIP networks:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

for n *sufficiently large.*

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

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

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

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

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

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

2.6.3 Equilibrium Welfare and Efficiency in OIP Networks

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

Equilibrium Welfare across OIP Networks. Despite in all OIP networks later moving agents take a weakly better action than their predecessors,[18] equilibrium welfare is not the same across OIP networks. To see this, suppose there exist agents $j, j+1 \in \mathbb{N}$ such that $a_j = x$ and $a_{j+1} = \neg x$. Therefore, action x is revealed to be inferior by time $j+2$ in equilibrium. In the complete network, action x is revealed to be inferior to any agent $n \geq j+2$, and so it is never sampled again. In other OIP networks, instead, agent j is not necessarily in the neighborhood of agent $n \geq j+2$, and therefore n fails to realize from agent $j+1$'s choice that action x is of lower quality than action $\neg x$. Thus, agent n inefficiently samples action x with positive probability at the second search stage.[19]

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

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

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

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

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

[19] For the descriptive analysis in this section, assume that search costs are not bounded away from zero. The formal details are in Appendix 2.9.6.

utility in all collective search environments with the same state process and search technology, but where the network topology is any OIP network.[20]

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

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

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

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

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

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

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

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

Discussion of Probability of Wrong Herds, Rate of Convergence, and Welfare. The results on OIP networks presented in Sections 2.6.2 and 2.6.3 are surprising for two reasons. First, in OIP networks the probability of wrong herds, the speed of learning, and the long-run (but not short-run) welfare neither depend on transparency of past histories nor on the correlation structure among connections. Second, the rate of convergence can be characterized (and in a simple way) for a large class of networks. This contrasts with our understanding of the SSLM, for which little is known about learning rates unless all agents observe the most recent action, a random action from the past, or all past actions (see Lobel et al. (2009), Rosenberg and Vieille (2017), and Hann-Caruthers, Martynov, and Tamuz (2018)).

Rosenberg and Vieille (2017) consider two measures of the efficiency of social learning in the SSLM: the expected time until the first correct action and the expected number of incorrect actions (see also Hann-Caruthers et al. (2018)). They focus on two polar setups and assume that each agent either observes the entire sequence of earlier actions or only the previous one. In a similar spirit with my results,

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

they find that whether learning is efficient is independent of the setup: for every signal distribution, learning is efficient in one setup if and only if it is efficient in the other one. In the search setting I study, the results on the irrelevance of how far in the past agents can observe is much stronger: first, it holds for the long-run welfare as well as for the probability of wrong herds and the speed of learning; second, it neither depends on the number of immediate predecessors that agents observe nor on the dependence structure among connections.

2.6.4 Policy Interventions

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

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

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

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

Interestingly, the remedies that this section suggests are easy to implement and commonly observed in practice. For instance, online platforms that aggregate past individual choices by sorting different items according to their popularity or sales rank serve the purpose.[21]

The Limits of Simple Policy Interventions. The interventions discussed above do not remove the inefficient duplication of costly search arising when agents fail to recognize actions that are inferior (and not revealed to be so) by some time n. Moreover, they do not incetivize exploration; thus, agents delay search for the second action more than what the single decision maker would do and the rate of convergence remains too slow. A natural step for future research is to understand how and to what extent more complex incentive schemes, which make use of monetary transfers or information management tools, can reduce these other inefficiencies as well.[22]

2.7 Related Literature and Concluding Remarks

2.7.1 Related Literature

This paper joins a small but growing literature on costly acquisition of private information in social learning settings. Burguet and Vives (2000) and Chamley (2004) consider a continuum of agents, each choosing an action from a continuous space in every period. Agents wish to match an unknown state of nature in order to minimize a quadratic loss and set the precision of a normally distributed signal at a cost that increases with the signal's precision. In Ali (2018) there is an unknown binary state of nature. Agents select an action from a space, either discrete or continuous, and aim at taking higher actions in the higher state. They act in sequence, observe the choices of all their predecessors, and choose how informative a signal to acquire at a cost which depends on the chosen informativeness about the relative likelihood of the two states. These costs are heterogeneous across agents and are private information. While I focus on some search cost types obtaining perfect signals in a discrete action space, these papers study noisy signals with a continuous (or general, in Ali (2018)) action space. Closer to my setup, Hendricks, Sorensen, and Wiseman (2012) study sequential learning when agents choose whether to purchase a product or not. Agents have heterogeneous preferences, which are private information, but identical search costs. At this cost, they can acquire a perfect signal about their value for the product. In their model, however, agents only observe the aggregate purchase history.[23]

My model departs from these papers in two relevant ways. First, I consider a game of social learning which is played over general networks. Thus, I provide conditions on both information acquisition

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

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

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

technologies and observation structures that lead to positive or negative learning results. Second, the parametric structure of private information is substantially different. These distinctions require different tools to analyze the social learning process and prevent a direct comparison of the results. A general insight of Ali (2018) and MFP is that, in the complete network, we can trade an assumption of arbitrarily strong exogenous private signals for an assumption of arbitrarily low information acquisition costs. My work shows that this insight generalizes to all network topologies where long information paths occur almost surely and are identifiable.

My model relates to those of sequential information acquisition of Wald (1947), Weitzman (1979), and Moscarini and Smith (2001), where a single decision maker dynamically chooses how much information to acquire before taking an action. Weitzman (1979) considers a sequential search environment where an agent faces a bandit problem, each arm representing a distinct alternative with a random prize, and characterizes the optimal sampling sequence and the optimal timing to stop the search process. Each agent in my model faces the same problem and trade-off between exploration (sampling the second action) and exploitation (taking the action believed to be the best according to his social information).[24]

Salish (2017) and Sadler (2017) study learning in networks where a finite number of agents acquire private information by strategic experimentation with a two-armed bandit, as in Keller, Rady, and Cripps (2005b) and Bolton and Harris (1999b), and observe the experimentation of their neighbors.[25] In these models, agents interact repeatedly over time, and so the strategic component of their interaction is more involved than in my setting. However, this comes at a cost. Sadler (2017) allows for complex network structures, but agents follow a boundedly rational decision rule. In Salish (2017) agents are rational, but a sharp characterization only obtains for particular network structures. Taking advantage of the sequential nature of the problem, I accommodate both for rational behavior and general network topologies. In a similar spirit, Perego and Yuksel (2016) study a model of learning where a continuum of Bayesian agents repeatedly choose between learning from one's own experimentation or learning from others' experiences. Connections are heterogeneous across agents and peer-to-peer exchange of information is subject to frictions. The authors characterize how frictions and heterogeneity in connections affect the creation and diffusion of knowledge in equilibrium, but do not focus on network properties other than connectivity.

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

The literature on social learning in networks is larger than the work surveyed here. It includes: other contributions on Bayesian observational learning, such as Mueller-Frank (2013), Arieli and Mueller-Frank (2018), Mossel, Sly, and Tamuz (2015); word-of-mouth learning models, where agents randomly sample others' opinions, as in Banerjee (1993), Ellison and Fudenberg (1995), and Banerjee and Fudenberg (2004); recent work on Bayesian communication learning, such as Acemoglu, Bimpikis, and

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

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

Ozdaglar (2014); models of non-Bayesian learning, including DeGroot (1974), DeMarzo, Vayanos, and Zwiebel (2003), Acemoglu, Ozdaglar, and ParandehGheibi (2010), Golub and Jackson (2010, 2012), and Molavi, Tahbaz-Salehi, and Jadbabaie (2018); models where agents' updating rules combine Bayesian and non-Bayesian features, as in Bala and Goyal (1998) and Jadbabaie, Molavi, Sandroni, and Tahbaz-Salehi (2012). art; Goyal (2011), Jackson (2008), Vives (2010), Acemoglu and Ozdaglar (2011), Mobius and Rosenblat (2014), and Golub and Sadler (2016) contain excellent (and complementary) accounts of the field.

2.7.2 Concluding Remarks

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

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

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

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

2.8 Examples for Remark 2 in Section 2.3.2

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

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

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

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

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

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

with

$$\begin{aligned} N(q_0) &:= \mathbb{P}_\sigma\big(s_1^1=0, c_1 > t^{\emptyset}(q_0)\big) \\ &= \frac{1}{2}\mathbb{P}_C\big(c_1 > t^{\emptyset}(q_0)\big), \end{aligned} \tag{2.16}$$

and

$$\begin{aligned} D(q_0) &:= \mathbb{P}_\sigma\big(s_1^1=0, c_1 > t^{\emptyset}(q_0)\big) + \mathbb{P}_\sigma\big(s_1^1=1, c_1 < t^{\emptyset}(q_1), q_0 > q_1\big) \\ &\quad + \mathbb{P}_\sigma\big(s_1^1=0, c_1 \le t^{\emptyset}(q_0), q_0 > q_1\big) + \frac{1}{2}\mathbb{P}_\sigma\big(s_1^1=0, c_1 \le t^{\emptyset}(q_0), q_0 = q_1\big) \\ &= \frac{1}{2}\Big[\mathbb{P}_C\big(c_1 > t^{\emptyset}(q_0)\big) + \mathbb{P}_{C\times Q}\big(c_1 < t^{\emptyset}(q_1), q_0 > q_1\big) \\ &\quad + \mathbb{P}_{C\times Q}\big(c_1 \le t^{\emptyset}(q_0), q_0 > q_1\big) + \frac{1}{2}\mathbb{P}_{C\times Q}\big(c_1 \le t^{\emptyset}(q_0), q_0 = q_1\big)\Big]. \end{aligned} \tag{2.17}$$

Above, I denote with $\mathbb{P}_{C\times Q}$ the product measure $\mathbb{P}_C \times \mathbb{P}_Q$ and with c_1 the search cost of agent 1. Consistently with the analysis in the rest of the paper, to derive an expression for $n(q_0)$ and $P_1(q_0)$ I assumed that agent 1 breaks ties uniformly at random at the first search stage and at the choice stage. The chosen tie-breaking rule does not qualitatively affect the results.[26]

[26] The same remarks apply to Example 6.

Straightforward calculations yield

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

Moreover,

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

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

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

and

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

Therefore,

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

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

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

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

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

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

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

Mimicking the analysis in Example 5, we now have

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

Moreover,

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

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

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

and

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

Therefore,

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

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

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

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

2.9 Proofs

2.9.1 Proofs for Section 2.4.2

Preliminaries

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

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

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

then asymptotic learning occurs in equilibrium σ.

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

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

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

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

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

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

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

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

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = 0\right) = \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = 0\right). \tag{2.18}$$

Second, suppose $0 < b < n$, so that $B_n \neq \emptyset$. By the characterization of the equilibrium decision s_n^1 in Section 2.3.2,

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

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

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

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

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

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

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = b\right) \geq \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = b\right). \tag{2.19}$$

The desired result obtains by observing that

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

where the two equalities hold by the law of total probability and the inequality holds by (2.18) and (2.19). ■

Proof of Proposition 2

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

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \geq \phi_k \quad \text{for all } n \geq \alpha_k. \tag{2.20}$$

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

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

$$\mathbb{Q}\big(\gamma_n((B(n)) = b, b < \alpha\big) < \frac{\varepsilon}{2}, \tag{2.21}$$

and

$$\mathbb{P}_\sigma\left(\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n))\right) < \mathcal{Z}\left(\mathbb{P}_\sigma\left(s^1_{\gamma_n(B(n))} \in \arg\max_{x\in X} q_x\right)\right) - \varepsilon\right) < \frac{\varepsilon}{2} \tag{2.22}$$

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

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

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

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

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

I use induction on the index k to prove relation (2.20). Since the qualities of the two actions are i.i.d. draws and agent 1 has no a priori information,

$$\mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right) = \frac{1}{2}. \tag{2.23}$$

From Lemma 2,

$$\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \geq \mathbb{P}_\sigma\left(s_1^1 \in \arg\max_{x\in X} q_x\right) \tag{2.24}$$

for all $n \in \mathbb{N}$. From (2.23) and (2.24) we have

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

which together with $\alpha_1 = 1$ and $\phi_1 = \frac{1}{2}$ establishes relation (2.20) for $k = 1$. Assume that relation (2.20) holds for an arbitrary k, that is

$$\mathbb{P}_\sigma\left(s_j^1 \in \arg\max_{x \in X} q_x\right) \geq \phi_k \quad \text{for all } j \geq \alpha_k, \tag{2.25}$$

and consider some agent $n \geq \alpha_{k+1}$. To establish (2.20) for $n \geq \alpha_{k+1}$ observe that

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

where the inequality follows from (2.21) and (2.22), the inductive hypothesis in (2.25), and the assumption that $\mathcal{Z}$ is increasing.

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

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

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

Proof of Proposition 3

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

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

Denote with $\tilde{s}_n^1$ the coarse optimal decision of agent n at the first search stage when he only uses information from neighbor b.[27]. The optimal search policy, as characterized in Section 2.3.2, requires

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

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

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

$$\mathbb{P}_\sigma\big(E_n^{a_b} \mid \gamma_n(B(n)) = b\big) \leq \mathbb{P}_\sigma\big(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b\big). \tag{2.26}$$

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

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

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

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

The next lemma shows that network topologies where $\mathbb{Q}(|B(n)| \leq 1) = 1$ for all n satisfy condition (2.26). In particular, this condition is satisfied by all chosen neighbor topologies.

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

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

Now suppose $\mathbb{P}_\sigma(E_n^{a_b} \mid \widehat{B}(n) = \widehat{B}_n) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \widehat{B}(n) = \widehat{B}_n)$ for all $n \leq k$ and all $\widehat{B}_n$ that occurs with positive probability. Given a realization $\widehat{B}_{k+1}$ of $\widehat{B}(k+1)$, if $B_{k+1} = \emptyset$, then agent $k+1$ faces the same situation as the first agent, and the desired conclusion follows. If $B_{k+1} = \{b\}$, take $\gamma_{k+1}(\{b\}) = b$ and let $(\pi_1, \ldots, \pi_l)$ be the sequence of agents in $\widehat{B}_{k+1} \cup \{k+1\}$. That is, $\{\pi_1, \ldots, \pi_l\}$ is such that $\pi_1 = \min \widehat{B}_{k+1}$, $\pi_l = k+1$ and, for all g with $1 < g \leq l$, $B_{\pi_g} = \{\pi_{g-1}\}$. Moreover, for all g with $1 < g \leq l$, say that agent π_{g-1} is the immediate predecessor of agent π_g in $\widehat{B}_{k+1}$. When $\widehat{B}_{k+1} = \{b\}$, the desired result trivially holds. When $\widehat{B}_{k+1}$ contains more than one agent, the desired result follows by observing that, under the inductive hypothesis and the equilibrium decision rule, each agent in $\{\pi_1, \ldots, \pi_{l-1}\}$ samples first the action taken by his immediate predecessor. ■

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

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

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

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

The next two lemmas provide an expression for the probability of agent n sampling first the best action when using $\tilde{s}_n^1$, conditional on agent b being selected by agent n's neighbor choice function, in terms of the probability β of agent b doing so, the private search cost distribution, the function $t^{\emptyset}(\cdot)$ defined in (2.2), and the thresholds $P_{b,n}^{\sigma}(q_{\min})$ and $P_{b,n}^{\sigma}(q_{\max})$.

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

$$\begin{aligned}
&\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \underset{x\in X}{\arg\max}\ q_x \mid \gamma_n(B(n)) = b\right)\\
&= \mathbb{P}_\sigma\left(s_b^1 \in \underset{x\in X}{\arg\max}\ q_x \mid \gamma_n(B(n)) = b\right)\\
&+ \mathbb{P}_\sigma\big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\big)\left(1 - \mathbb{P}_\sigma\left(s_b^1 \in \underset{x\in X}{\arg\max}\ q_x \mid \gamma_n(B(n)) = b\right)\right).
\end{aligned} \tag{2.27}$$

Proof. By Lemma 3,

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

Moreover,

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

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

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

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

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

The desired result follows. ■

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

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

Proof. By Lemma 5,

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) = \beta + \mathbb{P}_\sigma\big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\big)(1-\beta). \tag{2.28}$$

Moreover, by the law of total probability,

$$\begin{aligned}
&\mathbb{P}_\sigma\big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\big) \\
&= \mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \\
&+ \mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(s_b^1 \notin \arg\max_{x\in X} q_x \;\middle|\; \gamma_n(B(n)) = b\right) \qquad (2.29)\\
&= \beta\,\mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\right) \\
&+ (1-\beta)\,\mathbb{P}_\sigma\left(s_b^2 = \neg s_b^1 \;\middle|\; \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\right).
\end{aligned}$$

By the characterization of equilibrium strategies in Section 2.3.2 we have, conditional on $\gamma_n(B(n)) = b$ and $s_b^1 \in \arg\max_{x\in X} q_x$,

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

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

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

where we assume that agent n samples the second action in case of indifference.[28] It follows that

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

[28] This assumption does not affect the results.

and that

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

Thus, equation (2.29) can be rewritten as

$$\mathbb{P}_\sigma\big(s_b^2 = \neg s_b^1 \mid \gamma_n(B(n)) = b\big) = \beta F_C\big(P_{b,n}^\sigma(q_{\max})t^\emptyset(q_{\max})\big) + (1-\beta)F_C\big(P_{b,n}^\sigma(q_{\min})t^\emptyset(q_{\min})\big). \quad (2.30)$$

The desired result follows by combining (2.28) and (2.30). ■

The previous lemma shows that the quantity

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

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

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

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

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

$$\beta \geq 1 - \mathbb{P}_\sigma\Big(E_b^{s_b^1} \mid \gamma_n(B(n)) = b\Big),$$

or

$$1-\beta \leq \mathbb{P}_\sigma\Big(E_b^{s_b^1} \mid \gamma_n(B(n)) = b\Big). \quad (2.31)$$

Moreover, by the law of total probability,

$$\begin{aligned}
&\mathbb{P}_\sigma\Big(E_b^{s_b^1} \mid \gamma_n(B(n)) = b\Big) \\
&= \mathbb{P}_\sigma\left(E_b^{s_b^1} \mid \gamma_n(B(n)) = b, s_b^1 \in \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(s_b^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right) \\
&+ \mathbb{P}_\sigma\left(E_b^{s_b^1} \mid \gamma_n(B(n)) = b, s_b^1 \notin \arg\max_{x\in X} q_x\right)\mathbb{P}_\sigma\left(s_b^1 \notin \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\right) \quad (2.32)\\
&= \beta P_{b,n}^\sigma(q_{\max}) + (1-\beta)P_{b,n}^\sigma(q_{\min}).
\end{aligned}$$

Combining (2.31) and (2.32) yields

$$1-\beta \leq \beta P_{b,n}^\sigma(q_{\max}) + (1-\beta)P_{b,n}^\sigma(q_{\min}), \quad (2.33)$$

and therefore

$$\max\left\{P^{\sigma}_{b,n}(q_{\min}), P^{\sigma}_{b,n}(q_{\max})\right\} \geq 1-\beta. \tag{2.34}$$

Finally, observe that

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

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

$$\max\left\{F_C\big(P^{\sigma}_{b,n}(q_{\max})t^{\emptyset}(q_{\max})\big), F_C\big(P^{\sigma}_{b,n}(q_{\min})t^{\emptyset}(q_{\max})\big)\right\} \geq F_C\big((1-\beta)t^{\emptyset}(q_{\max})\big),$$

which holds because of (2.34) and the fact that F_C is increasing. The desired result follows. ■

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

$$\overline{\mathcal{Z}}(\beta) \coloneqq \beta + (1-\beta)^2 F_C\big((1-\beta)t^{\emptyset}(q_{\max})\big). \tag{2.35}$$

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

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

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

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

Lemma 8. *The function $\overline{\mathcal{Z}}\colon [1/2,1] \to [1/2,1]$, defined pointwise by (2.35), satisfies the following properties:*

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

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

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

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

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

For all $\beta \in [1/2, 1)$, $(1-\beta)t^{\emptyset}(q_{\max}) > 0$ and so, if search costs are not bounded away from zero, $F_C((1-\beta)t^{\emptyset}(q_{\max})) > 0$.[29] Since also $(1-\beta)^2 > 0$ for all $\beta \in [1/2, 1)$, the improvement term of function $\overline{\mathcal{Z}}$ is positive and so part *(b)* holds.

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

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

$$\mathcal{Z}(\beta) := \frac{1}{2}\left(\beta + \sup_{r \in [1/2, \beta]} \overline{\mathcal{Z}}(r)\right). \tag{2.36}$$

Lemma 9. *The function $\mathcal{Z}\colon [1/2, 1] \to [1/2, 1]$ defined by (2.36) satisfies the following properties:*

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

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

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

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

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

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

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

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

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

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

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

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

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

Proof. Let again β denote $\mathbb{P}_\sigma(s_b^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = b)$. If $\mathcal{Z}(\beta) = \beta$, the result follows from Lemma 6. Suppose next that $\mathcal{Z}(\beta) > \beta$. By (2.36), this implies that $\mathcal{Z}(\beta) < \sup_{r \in [1/2, \beta]} \overline{\mathcal{Z}}(r)$. Therefore, there exists $\overline{\beta} \in [1/2, \beta]$ such that

$$\overline{\mathcal{Z}}(\overline{\beta}) \geq \mathcal{Z}(\beta). \tag{2.37}$$

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

$$\tilde{a}_b = \begin{cases} a_b & \text{with probability } (2\overline{\beta} - 1)/(2\beta - 1) \\ 0 & \text{with probability } (\beta - \overline{\beta})/(2\beta - 1) \\ 1 & \text{with probability } (\beta - \overline{\beta})/(2\beta - 1), \end{cases} \tag{2.38}$$

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

$$\mathbb{P}_\sigma\big(E_n^{\tilde{a}_b} \mid \gamma_n(B(n)) = b\big) \leq \mathbb{P}_\sigma\big(E_n^{\neg \tilde{a}_b} \mid \gamma_n(B(n)) = b\big). \tag{2.39}$$

The relation in (2.39), together with the characterization of the equilibrium search policy in Section 2.3.2, implies that agent n samples first action $\tilde{a}_b$ upon observing the choice of the fictitious agent. That is, denoting with $\widetilde{\tilde{s}}_n^1$ the first action sampled by agent n upon observing the choice of the fictitious agent, $\widetilde{\tilde{s}}_n^1 = \tilde{a}_b$. Moreover, the assumption $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\neg a_b} \mid \gamma_n(B(n)) = b)$ and (2.38) also imply that $\mathbb{P}_\sigma(E_n^{a_b} \mid \gamma_n(B(n)) = b) \leq \mathbb{P}_\sigma(E_n^{\tilde{a}_b} \mid \gamma_n(B(n)) = b)$. Therefore, the distribution of the quality of action a_b first-order stochastically dominates the distribution of the quality of action $\tilde{a}_b$. Since $\tilde{s}_n^1 = a_b$ and $\widetilde{\tilde{s}}_n^1 = \tilde{a}_b$, it follows that

$$\mathbb{P}_\sigma\left(\tilde{s}_n^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = b\right) \geq \mathbb{P}_\sigma\left(\widetilde{\tilde{s}}_n^1 \in \arg\max_{x \in X} q_x \mid \gamma_n(B(n)) = b\right). \tag{2.40}$$

Now denote with $\tilde{s}_b^1$ the decision of the fictitious agent about which action to sample first. From (2.38), one can think of $\tilde{s}_b^1$ as generated as

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

Therefore,

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

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

$$\mathbb{P}_\sigma\Big(\widetilde{s}_n^1 \in \arg\max_{x\in X} q_x \mid \gamma_n(B(n)) = b\Big) \geq \overline{\mathcal{Z}}(\overline{\beta}). \tag{2.41}$$

Since $\overline{\mathcal{Z}}(\overline{\beta}) \geq \mathcal{Z}(\beta)$ (see equation (2.37)), the desired result follows from (2.40) and (2.41). ■

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

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

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

Proof. Fix any $n \in \mathbb{N}$. If $b = 0$, then $\tilde{s}_n^1 = s_n^1$ by Remark 8, and the claim trivially holds. Now suppose $0 < b < n$, so that $B_n \neq \emptyset$. By the characterization of the equilibrium decision s_n^1 in Section 2.3.2,

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

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

$$\mathbb{P}_\sigma\big(E_n^{s_n^1} \mid B_n\big) \leq \mathbb{P}_\sigma\big(E_n^{\tilde{s}_n^1} \mid B_n\big) \tag{2.42}$$

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

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

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

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

as desired. ■

2.9.2 Proofs for Section 2.4.3

Preliminaries

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

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

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

$$\mathbb{P}_\Omega\big(\Omega^{NS}\big) = \mathbb{P}_Q\big(Q^{NS}\big) \times \mathbb{P}_Q\big(Q^{NS}\big) \geq \delta. \tag{2.43}$$

When $\omega \in \Omega^{NS}$, an agent with nonempty neighborhood does not sample the second action either with positive probability, independently of which action he samples first (see the characterization and discussion of equilibrium behavior in Section 2.3.2). Finally, by Assumption 1, conditional on $\omega \in \Omega^{NS}$, the two actions have different quality with positive probability.

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

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

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

Proof of Proposition 4

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

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

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

$$\mathbb{Q}\big(\big|\widehat{B}(n)\big| < K\big) \geq \varepsilon \qquad \text{for all } n \in \mathcal{N}. \tag{2.44}$$

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

$$\begin{aligned}
&\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\big) \\
&= \mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K\big)\mathbb{Q}\big(\big|\widehat{B}(n)\big| < K\big) \\
&+ \mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| \geq K\big)\mathbb{Q}\big(\big|\widehat{B}(n)\big| \geq K\big) \\
&\geq \mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K\big)\mathbb{Q}\big(\big|\widehat{B}(n)\big| < K\big) \\
&\geq \varepsilon\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K\big),
\end{aligned} \tag{2.45}$$

where the last inequality follows from (2.44). By the law of total probability again, we also have

$$\begin{aligned}
&\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K\big) \\
&= \mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big)\mathbb{P}_\Omega\big(\omega \in \Omega^{NS}\big) \\
&+ \mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \notin \Omega^{NS}\big)\mathbb{P}_\Omega\big(\omega \notin \Omega^{NS}\big) \\
&\geq \mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big)\mathbb{P}_\Omega\big(\omega \in \Omega^{NS}\big) \\
&\geq \delta\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big),
\end{aligned} \tag{2.46}$$

where the last inequality holds by (2.43). Then, by (2.45) and (2.46),

$$\begin{aligned}
&\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\}\big) \\
&\geq \varepsilon\delta\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big)
\end{aligned} \tag{2.47}$$

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

Let $\overline{C}_\sigma(q^{NS})$ denote the set of all private search costs for which an agent h with second search stage information set I_h^2 such that $B_h = \emptyset$, $q_{s_h^1} = q^{NS}$, and $c_h \in \overline{C}_\sigma(q^{NS})$, adopts strategy $s_h^2 = ns$ in equilibrium σ. That is, $\overline{C}_\sigma(q^{NS})$ consists of all search costs for which, in equilibrium σ, an agent with empty neighborhood decides not to sample the second action when the first action he samples has quality q^{NS}. For all $\omega \in \Omega^{NS}$, the results of Section 2.3.2 imply that any agent k with search cost $c_k \in \overline{C}_\sigma(q^{NS})$ adopts strategy $s_k^2 = ns$ at the second search stage independently of his neighborhood realization B_k, the actions of his neighbors, and the quality of the first action sampled (i.e. independently of the realizations of the random variables in his information set other than his search cost). Then,

$$\begin{aligned}&\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big) \\ &\geq \mathbb{P}_\sigma\big(c_k \in \overline{C}_\sigma\big(q^{NS}\big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big).\end{aligned} \tag{2.48}$$

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

$$\begin{aligned}&\mathbb{P}_\sigma\big(c_k \in \overline{C}_\sigma\big(q^{NS}\big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big) \\ &= \mathbb{P}_\sigma\big(c_k \in \overline{C}_\sigma\big(q^{NS}\big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K\big).\end{aligned} \tag{2.49}$$

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

$$\begin{aligned}&\mathbb{P}_\sigma\big(c_k \in \overline{C}_\sigma\big(q^{NS}\big) \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K\big) \\ &\geq \mathbb{P}_\sigma\big(c_1 \in \overline{C}_\sigma\big(q^{NS}\big)\big)^K \\ &> 0,\end{aligned} \tag{2.50}$$

where the strict inequality holds because $\mathbb{P}_\sigma(c_1 \in \overline{C}_\sigma(q^{NS})) > 0$ by the first condition in Assumption 1. Together, (2.48), (2.49), and (2.50) yield that

$$\mathbb{P}_\sigma\big(s_k^2 = ns \text{ for all } k \in \widehat{B}(n) \cup \{n\} \mid \big|\widehat{B}(n)\big| < K, \omega \in \Omega^{NS}\big) > 0. \tag{2.51}$$

As $\varepsilon, \delta > 0$, from (2.47) and (2.51) we conclude

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

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

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

as desired. ■

2.9.3 Preliminaries for Sections 2.5 and 2.6

Characterization of Equilibrium Strategies in OIP Networks

Part *(a)* of Theorem 2 and the results in Section 2.6 are largely based on the next lemma, which characterizes equilibrium sequential search policies in OIP networks. Let $P_1(q)$ denote the posterior probability that agent 1 did not sample the second action given that the action he takes has quality q. The precise functional form of $P_1(q)$ is irrelevant for the following argument.

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

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

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

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

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

$$t_n\left(q_{s_n^1}\right) := P_1\left(q_{s_n^1}\right) t^{\emptyset}\left(q_{s_n^1}\right) \tag{2.52}$$

for $n = 2$, and pointwise recursively as

$$t_n\left(q_{s_n^1}\right) := P_1\left(q_{s_n^1}\right)\left(\prod_{i=2}^{n-1}\left(1 - F_C\left(t_i\left(q_{s_n^1}\right)\right)\right)\right) t^{\emptyset}\left(q_{s_n^1}\right) \tag{2.53}$$

for $n > 2$.[30]

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

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

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

Now consider any agent $n > 2$. Suppose that all agents up to $n-1$ follow this strategy, and that agent $n-1$ selects action a_{n-1}. If action $\neg a_{n-1}$ is revealed inferior to agent n in equilibrium σ, it must be that $q_{\neg a_{n-1}} \leq q_{a_{n-1}}$, and so action $\neg a_{n-1}$ is not sampled at all. Now suppose that action $\neg a_{n-1}$ is not revealed inferior to agent n in equilibrium σ. By the same logic as before, n's beliefs about the quality

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

of action a_{n-1} strictly first-order stochastically dominate his beliefs about the quality of action $\neg a_{n-1}$. Therefore, $s_n^1 = a_{n-1}$, i.e. he will sample action a_{n-1} first.

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

For part *(ii)–(b)*, consider any agent $n \geq 2$ and suppose that $\neg a_{n-1}$ is not revealed inferior to agent n in equilibrium σ. In OIP networks, the personal subnetwork of agent n, $\widehat{B}(n)$, is $\{1, \ldots, n-1\}$ with probability one. Moreover, by part *(i)*, each agent samples first the action taken by his immediate predecessor. Therefore, none of the agents in the personal subnetwork of agent n relative to action s_n^1 has sampled action $\neg s_n^1$ only if none of the first $n-1$ agents has sampled it; that is, only if $s_1^1 = s_n^1$, and $s_i^2 = ns$ for $1 \leq i \leq n-1$. The thresholds in (2.52) and (2.53) provide an explicit formula for (2.9) when $\widehat{B}(n) = \{1, \ldots, n-1\}$ with probability one for all $n \in \mathbb{N}$. To see this, proceed by induction. Consider first agent 2. By part *(i)*, $s_2^1 = a_1$. Let $P_1(q_{s_2^1})$ be the posterior probability that agent 1 did not sample action $\neg s_2^1$ given that action s_2^1 of quality $q_{s_2^1}$ was taken. Then, agent 2's expected benefit from the second search is $P_1(q_{s_2^1})t^{\emptyset}(q_{s_2^1})$, which is the right-hand side of (2.52). Now consider any agent $n > 2$, and let s_n^1 be the action this agent samples first. By part *(i)* and the inductive hypothesis, and since search costs are i.i.d. across agents, it follows that the probability that no agent in $\{1, \ldots, n-1\}$ has sampled action $\neg s_n^1$ is

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

Therefore, the right-hand side of (2.53) gives agent n's expected benefit from the search follows. The optimality of the proposed sequential search policy follows from the characterization of individual equilibrium decisions at the second search stage in Section 2.3.2. ■

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

Corollary 1. *Let $\mathcal{S}$ and $\mathcal{S}'$ be two collective search environments with identical state process and search technology. Assume that $\mathcal{S}$ is endowed with the complete network, while the network topology of $\mathcal{S}'$ is any OIP network. Finally, let $\sigma \in \Sigma_{\mathcal{S}}$ and $\sigma' \in \Sigma_{\mathcal{S}'}$, and assume that ties are broken according to the same criterion in σ and σ'.*[31] *Then, for all $n \in \mathbb{N}$,*

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

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

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

2.9.4 Proofs for Section 2.5.5

Proof of Theorem 2

Proof of part *(a)*. Suppose $\omega \notin \overline{\Omega}(\underline{c})$, and that the lowest cost in the support of $\mathbb{P}_C$ is $\underline{c} > 0$. Maximal learning requires that the probability that agent n takes the action with the highest quality converges to one as $n \to \infty$ (see the characterization of maximal learning in (2.12)). This, in turn, is equivalent to saying that the probability of the event "none of the agents in $\widehat{B}(n) \cup \{n\}$ samples both actions" converges to zero as $n \to \infty$ whenever the quality of the first action sampled by agent 1 is lower than $q(\underline{c})$.[32] To establish the failure of maximal learning, I show that the probability of this event remains bounded away from zero when $\underline{c} > 0$.

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

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

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

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

(see 2.53 and the proof of Lemma 13), where $\hat{q}$ is the quality of the action taken by agent n, also converges to zero as $n \to \infty$ for all $\hat{q} < q(\underline{c})$. Then, there exists an agent $N_{\hat{q}} + 1$ for which the expected additional gain from the second search falls below $\underline{c}$.

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

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

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

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

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

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

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

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

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

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

where $P_{\pi_{k+1}}(\cdot)$ is the function defined by (2.8). It follows that the expected additional gain from the second search for agent π_{k+1}, given by

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

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

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

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

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

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

Therefore, with positive probability, agent π_1 samples first a suboptimal action with quality, say, $\overline{q}$, and does not search further. Now suppose that the first $\pi_{K_{\hat{q}}}$ agents in the sequence all have costs larger than $t^{\emptyset}(\overline{q})$, and again note that this occurs with positive probability. By Lemma 4, each of these agents will sample the suboptimal action with quality $\overline{q}$ first, and none of these agents will search further. Therefore, all will take this suboptimal action. Agent $\pi_{K_{\hat{q}}} + 1$ also samples this action first, and does not search further either because his expected additional gain from the second search is smaller than

$\underline{c}$. Since the expected additional gain from the second search remains smaller than $\underline{c}$ afterward, there will be no further search by agents in the sequence moving after agent $\pi_{K_{\hat{q}}}+1$, contradicting that the probability of no agent in $\widehat{B}(\pi_k) \cup \{\pi_k\}$ sampling both actions converges to zero. The desired result follows. ■

2.9.5 Proofs for Section 2.6.2

Proof of Proposition 6

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

Proof of Proposition 7

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

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

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

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

$$\begin{aligned}\mathbb{P}_\sigma\left(s_{n+1}^1 \in \arg\max_{x\in X} q_x\right) &\geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \\ &+ \left(1-\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right)\right)^2 F_C\left(\left(1-\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right)\right)t^{\emptyset}(q_{\max})\right).\end{aligned} \tag{2.54}$$

If the search cost distribution has polynomial shape, from (2.54) we have

$$\begin{aligned}\mathbb{P}_\sigma\left(s_{n+1}^1 \in \arg\max_{x\in X} q_x\right) &\geq \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \\ &+ Lt^{\emptyset}(q_{\max})^K\left(1-\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right)\right)^{K+2}.\end{aligned} \tag{2.55}$$

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

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

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

Proof of Proposition 8

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

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

Under the assumptions of of the proposition,

$$\begin{aligned}\mathbb{P}_\sigma\left(s_{n+1}^1 \in \arg\max_{x\in X} q_x\right) &= \frac{1}{n}\sum_{b=1}^{n}\mathbb{P}_\sigma\left(s_{n+1}^1 \in \arg\max_{x\in X} q_x \mid B(n+1)=\{b\}\right) \\ &= \frac{1}{n}\left[\mathbb{P}_\sigma\left(s_{n+1}^1 \in \arg\max_{x\in X} q_x \mid B(n+1)=\{n\}\right) + (n-1)\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right)\right]\end{aligned} \tag{2.56}$$

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

$$\begin{aligned}\mathbb{P}_\sigma\left(s_{n+1}^1 \in \arg\max_{x\in X} q_x\right) \geq{}& \mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right) \\ &+ \frac{Lt^{\emptyset}(q_{\max})^K}{n}\left(1-\mathbb{P}_\sigma\left(s_n^1 \in \arg\max_{x\in X} q_x\right)\right)^{K+2}.\end{aligned} \tag{2.57}$$

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

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

where $\overline{K}$ is some constant of integration (in the construction, $\tilde{\phi}$ is found as the solution to an ordinary differential equation).[34] ■

[33] To apply a construction in the spirit of Lobel et al. (2009), the right-hand side of (2.55) must be increasing in $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x)$. This is so under the assumption $0 < L < \frac{2^{K+1}}{(K+2)t^{\emptyset}\left(\underline{q}\right)^K}$ maintained in the proposition.

[34] To apply a construction in the spirit of Lobel et al. (2009), the right-hand side of (2.57) must be increasing in $\mathbb{P}_\sigma(s_n^1 \in \arg\max_{x\in X} q_x)$. This is so under the assumption $0 < L < \frac{2^{K+1}}{(K+2)t^{\emptyset}\left(\underline{q}\right)^K}$ maintained in the proposition.

2.9.6 Proofs for Section 2.6.3

Preliminaries

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

▷ Let $\mathcal{S}$ and $\mathcal{S}'$ be two collective search environments with identical state process $(Q, \mathcal{F}_Q, \mathbb{P}_Q)$ and search technology $\{(C, \mathcal{F}_C, \mathbb{P}_C), \mathcal{R}\}$. Suppose that the network topology of $\mathcal{S}$ is the complete network and that in $\mathcal{S}'$ agents only observe their most immediate predecessor. Let $\sigma \in \Sigma_{\mathcal{S}}$ and $\sigma' \in \Sigma_{\mathcal{S}'}$. Suppose that agents break ties according to the same criterion in σ and σ'. In particular, assume that agent 1 selects uniformly at random which action to sample first, and that all agents sample the other action whenever indifferent at the second search stage.[35] Suppose also that the first action sampled by the first agent in σ and σ', say x, has the same quality q_x. Let $\delta \in (0,1)$ be the discount rate, and let the function $t_1 \colon Q \to \mathbb{R}_+$ be defined pointwise by $t_1(q) \coloneqq t^{\emptyset}(q)$.[36] Hereafter, $q_{\neg x}$ is a random variable with probability measure $\mathbb{P}_Q$.

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

$$\begin{aligned}
U_\sigma(q_x;\delta) = q_x + t_1(q_x) &- (1-\delta)\sum_{n=1}^{\infty}\delta^n\Bigg(\prod_{i=1}^{n}\big(1-F_C\big(t_i(q_x)\big)\big)\Bigg)t_1(q_x) \\
&- (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\big[c \mid c \le t_n(q_x)\big]F_C(t_n(q_x))\prod_{i=1}^{n-1}\big(1-F_C(t_i(q_x))\big) \qquad (2.58)\\
&- (1-\delta)\mathbb{P}_Q(q_{\neg x} \le q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\big[c \mid c \le t_n(q_x)\big]F_C(t_n(q_x)).
\end{aligned}$$

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

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

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

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

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

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

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

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

$$
\begin{aligned}
U_{\sigma'}(q_x;\delta) = q_x + t_1(q_x) &- (1-\delta)\sum_{n=1}^{\infty}\delta^n\Big(\prod_{i=1}^{n}\big(1-F_C\big(t_i(q_x)\big)\big)\Big)t_1(q_x) \\
&- (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\big[c \mid c \leq t_n(q_x)\big]F_C(t_n(q_x))\prod_{i=1}^{n-1}\big(1-F_C(t_i(q_x))\big) \\
&- (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_Q}\big[\mathbb{E}_{\mathbb{P}_C}\big[c \mid c \leq t_n(q_{\neg x})\big]F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\big] \\
&\cdot\Big(1-\prod_{i=1}^{n-1}\big(1-F_C(t_i(q_x))\big)\Big) \\
&- (1-\delta)\mathbb{P}_Q(q_{\neg x} \leq q_x)\sum_{n=1}^{\infty}\delta^n\mathbb{E}_{\mathbb{P}_C}\big[c \mid c \leq t_n(q_x)\big]F_C(t_n(q_x)).
\end{aligned}
\tag{2.59}
$$

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

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

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

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

▷ Now consider a third collective search environment $\mathcal{S}''$ with the same state process and search technology as in $\mathcal{S}$ and $\mathcal{S}'$, but where the network topology is any OIP network. Let $\sigma'' \in \Sigma_{\mathcal{S}''}$, and suppose that indifferences are resolved in σ'' according to the same tie-breaking criterion as in σ and σ'. Assume also that the first action sampled by agent 1 in σ'', say x, has the same quality q_x as the action sampled at the first search by agent 1 in σ, σ'. Denote with $U_{\sigma''}(q_x;\delta)$ the expected discounted social utility normalized by $(1-\delta)$ in equilibrium σ''. Again, assume that the single decision maker selects the first action to sample uniformly at random, and that he samples the second action in case of indifference. The next lemma is immediate from the discussion in Section 2.6.3.

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

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

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

Proof of Proposition 9

The difference in average social utilities is

$$\begin{aligned} &U_\sigma(q_x;\delta) - U_{\sigma'}(q_x;\delta) \\ &\quad = (1-\delta)\mathbb{P}_Q(q_{\neg x} > q_x)\sum_{n=1}^{\infty}\delta^n \mathbb{E}_{\mathbb{P}_Q}\left[\mathbb{E}_{\mathbb{P}_C}\left[c \mid c \leq t_n(q_{\neg x})\right]F_C(t_n(q_{\neg x})) \mid q_{\neg x} > q_x\right] \\ &\quad \cdot \left(1 - \prod_{i=1}^{n-1}\left(1 - F_C(t_i(q_x))\right)\right). \end{aligned} \tag{2.60}$$

The right-hand side of (2.60) is positive for all $\delta \in (0,1)$. That $U_\sigma(q_x;\delta) > U_{\sigma'}(q_x;\delta)$ for all $\delta \in (0,1)$ follows.

To show that

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

we need to show that the right-hand side of (2.60) converges to zero as $\delta \to 1$. To do so, it is enough to argue that

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

is finite. Notice that

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

$$
\begin{aligned}
&\leq \sum_{n=\bar{n}+1}^{\infty} \delta^n \sup_{q>q_x} t_n(q) F_C(t_n(q)) + \bar{n} \sup_{q>q_x} t^{\emptyset}(q) \\
&\approx \sum_{n=\bar{n}+1}^{\infty} \delta^n \sup_{q>q_x} \left(t_n(q)\right)^2 f_C(0) + \bar{n} \sup_{q>q_x} t^{\emptyset}(q) \\
&\approx \sum_{n=\bar{n}+1}^{\infty} \delta^n \sup_{q>q_x} \left(t^{\emptyset}(q)\right)^2 \frac{1}{f_C(0)t^2} + \bar{n} \sup_{q>q_x} t^{\emptyset}(q),
\end{aligned}
$$

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

Proof of Proposition 10

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

$$
\lim_{\delta\to 1} U_{\sigma''}(q_x;\delta) = \lim_{\delta\to 1} U_{\sigma}(q_x;\delta). \tag{2.61}
$$

By Proposition 3 in MFP,

$$
\lim_{\delta\to 1} U_{\sigma}(q_x;\delta) = \lim_{\delta\to 1} U_{DM}(q_x;\delta). \tag{2.62}
$$

Then, by (2.61) and (2.62), and the uniqueness of the limit of a function, we have

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

which gives the desired result.

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

2.9.7 Proofs for Section 2.6.4

Proof of Proposition 11

An inductive argument analogous to the one establishing part *(i)* of Lemma 13 shows that each agent starts sampling from the action taken by his immediate predecessor. Then, the result follows directly from the discussion in Section 2.6.4. ■

Chapter 3

Learning While Bargaining: Experimentation and Coasean Dynamics[1]

3.1 Introduction

Bargaining is ubiquitous. Many economic interactions involve negotiations on a variety of issues. For example, prices of commodities are often the outcome of negotiations between the concerned parties, wages are set as an arrangement between firms and workers, and takeovers require an agreement over the price of the transaction. As such, bargaining relationships are the cornerstone of many theories of markets, from industrial organization to labor economics. Classical models of bargaining with incomplete information are typically presented as bilateral monopolies. These models posit common knowledge of gains from trade and assume the relevant information to reach an agreement to be available to parties—perhaps asymmetrically—since the outset of their negotiations. Yet, in many real-world bargaining situations, superior outside opportunities may become available to either or both parties during their negotiations—parties routinely investigating what their best options are, as a large literature on search documents. This paper characterizes bargaining dynamics in such an environment in which there is uncertainty about whether and when superior outside opportunities are available and new information about these opportunities may arrive during negotiations. In such an environment, gains from trade are ex ante uncertain and parties may want to wait to reach an agreement in order to learn about their best opportunities during negotiations.

I address this question in a one-sided incomplete information bargaining model between a seller who is unable to commit to future prices and a privately informed buyer. I show that the option value of waiting to learn about the existence of superior outside opportunities is of first-order importance in shaping the bargaining relationship. It affects the timing of agreements, the dynamics of prices, surplus

[1]I am grateful to Volker Nocke, Martin Peitz, Emanuele Tarantino, and Thomas Tröger for guidance and constant support throughout this project. I thank Vincenzo Denicolò, Francesc Dilmé, Georg Dürnecker, Daniel Garrett, Andreas Gulyas, Johannes Hörner, Bruno Jullien, Stephan Lauermann, Konrad Mierendorff, Marco Pagnozzi, Sven Rady, Patrick Rey, Nicolas Schutz, Ernst-Ludwig von Thadden, Veikko Thiele, and Jean Tirole for excellent suggestions that improved the paper. I received helpful comments from seminar audiences at the Center for Studies in Economics and Finance in Naples, University College London, the University of Bologna, the Toulouse School of Economics, the University of Bonn, the University of Mannheim, IIOC 2018, and the 2017 Bonn–Mannheim Economics Ph.D. Workshop. The usual disclaimer applies.

division, and the seller's ability to exercise market power. In equilibrium, there is either an initial period with no trade or trade starts with a burst. Afterward, the seller screens out buyer types one by one as uncertainty about the existence of superior outside opportunities unravels. Delay is always present, but not necessarily inefficient; inefficiently timed transactions only occur if valuations are interdependent. Whether prices increase or decrease over time depends on which party has a higher option value of waiting to learn. When the seller can clear the market in finite time at a positive price, prices are higher than the competitive price. Market power, however, need not be at odds with efficiency.

For concreteness, consider a capacity-constrained supplier of a leading technology making price offers to a downstream buyer which is privately informed about its own valuation for the technology. If the two firms were to negotiate in isolation, they could do no better than reaching a immediate agreement, at least in terms of joint surplus to share. In innovative industries, however, the supplier's competitors may develop a new disruptive technology in the future. The downstream buyer has then the option to wait for a new technology to arrive and so may only be willing to accept favorable trading conditions. Uncertainty about outside opportunities may also be present for the supplier. At some point, a new buyer with a higher valuation for the technology or without the time to engage in lengthy negotiations may approach the supplier. In this case, the supplier has the chance to conclude a favorable deal, whereas the downstream buyer loses the opportunity to trade. Parties are unlikely to know for sure whether such opportunities will arrive. The two parties, however, may learn by waiting. For instance, as time elapses with no breakthrough from the supplier's competitor, the original parties revise downward their beliefs about the chances of such an R&D success. They reach a corresponding conclusion about the existence of buyers interested in the technology if no new buyer shows up for a while. This snapshot of economic activity raises a number of questions. How do parties choose their bargaining posture in the face of market uncertainty? How do strategies depend on which party hopes for superior outside opportunities to arrive? Does the option to wait for the uncertainty to unravel lead to inefficiently late agreements? Or, rather, does the threat to leave the negotiation empty-handed leads the supplier (resp., buyer) to propose (resp., accept) a particularly favorable (resp., unfavorable) deal to (resp., from) the counterparty, to the effect that negotiations conclude inefficiently early?[2] Will the supplier exercise market power as uncertainty unravels in its favor? How will the supplier price its technology over time? How do learning about market opportunities and learning about the downstream buyer's private valuation interact in equilibrium?

In this paper, I develop a framework to answer these questions. Formally, I study a dynamic bargaining game between two risk-neutral players: a long-lived seller (he) and a long-lived buyer (she). Time is continuous and the time horizon is infinite. The seller has an indivisible durable good (or asset) to sell. His valuation for the good is normalized to zero. The buyer has a positive private valuation for the good (her type). There is uncertainty about whether and when superior outside opportunities become available. Negotiations take place in a market that can be of two types. In a market of type 0, no outside opportunity is available (bad-match market environment). In a market of type 1, superior outside opportunities arrive stochastically on either or both sides according to a Poisson process with

[2]When the option value of waiting to learn is high enough, the efficient benchmark is not immediate agreement but rather calls for an optimal degree of delay.

commonly known intensity (good-match market environment). From the viewpoint of time zero, gains from trade are uncertain. Bargaining begins with a common prior about the type of the market. The arrival of a superior outside opportunity is public and concludes the game. Thus, learning about the type of the market environment is common. As time elapses with no event, the two parties' belief that the market is of type 1 drifts downward and an agreement with the current trading partner becomes more attractive. The seller has no commitment power and makes price offers to the buyer at every instant. The buyer accepts or rejects.

Without arrivals (i.e., when it is common knowledge that the market is of type 0) the model reduces to the standard bargaining game with one-sided incomplete information. By the classic Coase Conjecture argument (Coase (1972)) the seller prices at the lowest buyer valuation as soon as negotiations begin; thus, trade occurs "in the twinkling of an eye", with the seller being unable to extract any rent from the transaction, and the market outcome is efficient. The intuition for the result is simple. For any given price, high valuation buyers are more likely to purchase than low valuation buyers, leading to negative selection in the demand pool. Accordingly, the seller cuts his price over time. A forward-looking buyer expects prices to fall, so she is unwilling to pay a high price in the first place. The seller's inability to commit thus leads its later selves to exert a negative externality on its former selves, reducing its overall profit to the lowest buyer valuation. The formalization of this argument is due to the seminal contributions by Stokey (1981), Bulow (1982), Fudenberg, Levine, and Tirole (1985), Gul, Sonnenschein, and Wilson (1986), and Ausubel and Deneckere (1989), among others.

Equilibrium bargaining dynamics drastically change when learning about the availability of superior outside opportunities is taken into account. Such a natural extension of the baseline model allows me to gain new insights on the original problem and to establish a set of novel results. To begin, I show that trade occurs over time in equilibrium, with the seller serving different (groups of) buyer types at different points in time. In particular, trade begins with a burst or following a silent period with no agreement. Afterward, the seller slowly screens out buyer types one by one as uncertainty about outside opportunities unravels. Absent outside opportunities or learning about their existence, the two parties would either trade immediately, upon meeting, or never reach an agreement.

When parties have the option of waiting to learn whether superior outside opportunities are available, immediate agreement is, in general, not efficient. Under efficiency, trade occurs when the joint benefit of market experimentation, as measured by the sum of the two players' option value of waiting to learn, equals its joint cost, as measured by the foregone gains from trade in terms of discounting. The optimal delay is different for different (groups of) buyer types. Thus, periods with no trade, as well as bursts of trade, followed by periods where different types trade one by one, are possible as efficient outcomes.

Whether players' incentives point toward inefficient hurry or inefficient delay is unclear. I show that delay is always present in equilibrium. Learning alone, however, only accounts for delay, but not for inefficiently timed agreements. Inefficiently late agreements only arise when the seller's payoff from the outside opportunity is correlated with the buyer type. This dependence endogenously creates a bargaining environment with interdependent values, as in Evans (1898), Vincent (1989), and Deneckere and Liang (2006). The main economic intuition behind the delay is similar to that in those papers. There is, however, an important difference between my findings and those of Deneckere and Liang (2006). In the setting without arrivals of Deneckere and Liang (2006), delay occurs because the equilibrium is characterized by burst of trade followed by periods of delay. During a period of delay, the sellers' belief must be exactly such that the Coasean desire to speed up trade is absent. With the addition of learning, the seller's belief cannot remain constant at such a belief over any time interval after trade begins. As a result, dynamics where bursts of trade alternate with periods of delay countably many times do not arise in my model, independently of whether there is a gap between the seller's valuation and the support of buyer valuations or not. Rather, the seller smoothly screens out buyer types one by one after trade begins.[3]

These results yield three main takeaways. First, they provide a novel and particularly natural rationale for both equilibrium delay and non-trivial trade dynamics in one-sided incomplete information bargaining environments. As striking as it is, the Coase Conjecture is at odds with how negotiations often occur in practice.[4] Second, as inefficient delay only arises when additional frictions are present in the trading environment (namely, interdependent values), the Coasean force toward efficiency remains overwhelming when parties are learning about the bargaining environment. Third, the result questions the view that long disputes result in inefficient outcomes: in markets with search and learning, examples of which are countless, this need not be true.

There is price discrimination in equilibrium. Prices smoothly decrease or increase over time depending on which party has the higher option value of waiting to learn. In particular, if superior outside opportunities may only arrive for the buyer, the price schedule is increasing. If, instead, superior outside opportunities may only arrive for the seller, the price schedule decreases over time.

The seller exercises market power if he has the option to clear the market in finite time at positive prices. In this case, prices are higher than the competitive price and the seller's payoff is higher than what he would get if he were: *(i)* awaiting for the possible arrival of a superior outside opportunity; *(ii)* unable to screen using prices; *(iii)* selling to a market in which all buyers had the lowest valuation.

My framework has a number of applications. A prominent application sees the bargaining game as the problem of a monopolist who is selling a perfectly divisible and infinitely durable good to a demand curve of atomless buyers and is unable to commit to future prices. The connection obtains because to every actual buyer type in the durable goods model, there corresponds an equivalent potential buyer type in the bargaining model. With this interpretation, my results shed new light on the dynamics of sales and the determinants of market power in monopolistic industries. Other applications include takeover negotiations, wage bargaining in markets for skilled workers, and negotiations in the housing market. In addition, we can interpret the model as a job search problem where an unemployed worker sets his reservation wage while learning about employment opportunities.

On the methodological side, posing the model in continuous time not only simplifies the analysis, but also allows for additional economic insights. Continuous time captures the idea that there are

[3]See also Daley and Green (2018) for a result in a similar spirit to mine.

[4]For instance, Ausubel, Cramton, and Deneckere (2002) argue that the Coase Conjecture "has the unfortunate implication that real bargaining delays can only be explained by either exogenous limitations on the frequency with which bargaining partners can make offers, or by significant differences in the relative degree of impatience between the bargaining parties."

no institutional frictions in the bargaining protocol (in addition to incomplete information). Thus, my analysis clearly disentangles the effect of learning about the market environment on equilibrium outcomes from that of other frictions in the protocol. In addition, optimality conditions, equilibrium strategies, and equilibrium outcomes have a clean characterization in continuous time. These conditions are described by means of Hamilton-Jacobi-Bellman equations and (solutions to) partial and ordinary differential equations with a clear economic interpretation. Closed-form expressions for all the relevant equilibrium outcomes of the game open the doors to comparative statics as well as to empirical studies and more applied research.[5]

Road Map. In Section 3.2, I introduce the general bargaining game and formalize the equilibrium notion. In Section 3.3, I describe two benchmark cases: efficient trade and the bargaining game without arrivals. I present the main results in Section 3.4. To gain insight, I develop the analysis in two steps: In Section 3.4.1, I characterize equilibrium bargaining dynamics with independent private valuations; in Section 3.4.2, I analyze equilibrium bargaining dynamics when valuations are interdependent. In Section 3.5, I discuss extensions of the general model and robustness checks. There, I also present the two other relevant benchmarks: the complete information outcome and the bargaining game with arrivals but no learning about the market environment. In Section 3.6, I discuss the related literature and conclude. In the main text, I provide a detailed account of the equilibrium characterization and develop the intuition behind the main results. The more technical proofs and additional details of the analysis are in Appendix 3.7.

3.2 Model and Equilibrium Notion

In this section, I first present the general bargaining game and discuss the main assumptions of the model. Then, I formalize players' strategies and the equilibrium notion.

3.2.1 The General Bargaining Game

Players and Values. There are two players, a seller (he) and a buyer (she). The seller has an indivisible durable good (or asset) to sell. His valuation for the good is normalized to zero. The buyer has a privately known type $v \in [\underline{v}, \overline{v}]$ that represents her valuation for the good. I assume $\overline{v} > \underline{v} \geq 0$. Type v is distributed according to a c.d.f. F, which is an atomless distribution with full support and density f. Following the standard terminology in the literature, if $\underline{v} > 0$ (resp., $\underline{v} = 0$), I refer to the model as the "gap" case (resp., "no gap" case) bargaining game. Hereafter, I use the words type and valuation interchangeably.

Time and Payoffs. Time, denoted by t, is continuous. The game starts at time zero and has a potentially infinite horizon: $t \in \mathbb{R}_+ \cup \{+\infty\}$. Players are long-lived, risk-neutral expected utility maximizers with common discount rate $r > 0$. The seller has no commitment power and makes a price offer p_t to the

[5] Fuchs and Skrzypacz (2010, 2013b), Ortner (2017), Daley and Green (2018), and Chaves Villamizar (2018) achieve a similar simplification. I refer to Section 3.6.1 for further discussion.

buyer at every instant t. If the buyer accepts the price offer p_t at time t, trade is executed and the game ends. If so, the seller's payoff is $e^{-rt}p_t$ and the buyer's payoff is $e^{-rt}(v - p_t)$.

Market Environment and News. Negotiations take place in a market of type $m \in \{0, 1\}$. If $m = 1$, an event stochastically occurs according to a Poisson process with intensity $\lambda > 0$. If $m = 0$, such event never occurs. The type of the market realizes once and for all before negotiations begin and is unknown to both players, who share a common prior $\mu^0 \in (0, 1)$ on $m = 1$ at the beginning of the game. The type of the market and the Poisson process governing the arrival of the event when $m = 1$ are independent of all other stochastic elements in the model. The arrival of the event at time t is public and concludes the game with (possibly) type-dependent payoffs $e^{-rt}O^S(v)$ for the seller and $e^{-rt}O^B(v)$ for the buyer. For now, think of the event as a reduced-form of some continuation play, and of $O^S(v)$ and $O^B(v)$ as the reduced-form payoffs associated to it. More structure will be imposed momentarily. The joint surplus conditional on the arrival of the event is $O(v) := O^S(v) + O^B(v)$.

Remark 10. The model has *independent private values (IPV)* if the function $O^S(v)$ is constant. If $O^S(v)$ is not constant in v, instead, the correlation of the seller's payoff from the outside opportunity with the buyer type endogenously gives rise to a bargaining environment with *interdependent values (IV).*

Either a transaction or the occurrence of an event conclude the game. Hereafter, whenever I refer to time t, I do so with the understanding that the game is still in place by then.

For future reference, let

$$\underline{O}^S(k) := \mathbb{E}\left[O^S(v) \mid v \leq k\right] = \int_{\underline{v}}^{k} O^S(v)\frac{f(v)}{F(k)}\mathrm{d}v$$

be the seller's expected payoff conditional on the arrival of the event and the buyer type being distributed according to the right-truncation of F over $[\underline{v}, k]$.

Learning. Since the arrival of the event is public, the seller and the buyer always share the same belief about the type of the market. To derive the law of motion of the common belief, suppose the two players start with the belief μ_t at time t and no event occurs in the interval $[t, t + \mathrm{d}t)$. By Bayes' rule, the updated belief at the end of the time interval is

$$\mu_t + \mathrm{d}\mu_t = \frac{\mu_t(1 - \lambda \mathrm{d}t)}{1 - \mu_t + \mu_t(1 - \lambda \mathrm{d}t)}.$$

Simplifying, we obtain that, as long as no event occurs, the common belief changes by $\mathrm{d}\mu_t = -\lambda\mu_t(1 - \mu_t)\mathrm{d}t$; its law of motion is described by the ordinary differential equation (henceforth, ODE)

$$\dot{\mu}_t = -\lambda\mu_t(1 - \mu_t), \qquad \mu_0 = \mu^0, \tag{3.1}$$

with solution

$$\mu_t = \frac{\mu^0 e^{-\lambda t}}{\mu^0 e^{-\lambda t} + (1 - \mu^0)}.^{6} \tag{3.2}$$

[6]Here, $e^{-\lambda t}$ is the probability that the event has not occurred by time t if the market is of type $m = 1$.

Thus, if the event does not occur, the common belief on $m = 1$ drifts downward over time. Once the event occurs, instead, the belief jumps to 1 and the game ends. Denote with t_μ the time at which the common belief μ_t equals $\mu \in [0, \mu^0]$, with the convention that $t_0 = +\infty$. Since μ_t strictly decreases over time, t_μ is well-defined.

Heuristic Timeline. Let $t \in \mathbb{R}_+$. The heuristic timeline within "period" $[t, t + \mathrm{d}t)$ is the following:

(i) The period begins with a common belief μ_t on $m = 1$.

(ii) If the market is of type $m = 1$, the event occurs with instantaneous probability $\lambda \mathrm{d}t$, terminating the game, and players collect payoffs; with complementary probability, no event occurs. If the market is of type $m = 0$, no event occurs. From the agents' viewpoint, the event occurs with subjective probability $\mu_t \lambda \mathrm{d}t$.

(iii) If no event occurs, the seller makes a price offer p_t, which the buyer accepts or rejects:

- *(a)* If the buyer accepts, the game ends and players collect payoffs.
- *(b)* If the buyer rejects, players update their belief about the market environment to $\mu_t + \mathrm{d}\mu_t$, and the game moves to the next period.[7]

Interpreting the Event. The event corresponds to an *outside opportunity* arriving for either or both players. I will assume that the outside opportunity is *superior* to an agreement with the current counterparty in terms of joint surplus to share, at least for some buyer types. In practice, the event may correspond to one of the two parties being rematched to an alternative trading partner or finding a more satisfactory use of his/her resources, thus disappearing from the original negotiation. It may also capture the arrival of a new agent offering better terms of trade to one of the two players. The event may represent a breakthrough in some underlying (on-the-market) search activity that parties are engaged in in parallel to their negotiations. Alternatively, it may correspond to a major technological step forward rendering obsolete the object that is originally for sale. Additionally, the event may correspond to favorable information arriving for either or both players. Finally, to capture the complexity of market interactions and bargaining relationships, the event may even correspond to the realization from some probability distribution over the previous cases. In short, many natural interpretations of the event—and of the associated learning process—are possible.

[7] The results of the paper hold unchanged under the following alternative timeline within "period" $[t, t + \mathrm{d}t)$:

(i) The period begins with a common belief μ_t on $m = 1$.

(ii) The seller makes a price offer p_t, which the buyer accepts or rejects.

(iii) If the buyer accepts, the game ends and players collect payoffs.

(iv) If the buyer rejects:

- *(a)* If the market is of type $m = 1$, the event occurs with probability $\lambda \mathrm{d}t$, terminating the game, and players collect payoffs; with complementary probability, no event occurs.
- *(b)* If the market is of type $m = 0$, no event occurs.

From the agents' viewpoint, the event occurs with subjective probability $\mu_t \lambda \mathrm{d}t$. If no event occurs, players update their belief about the market environment to $\mu_t + \mathrm{d}\mu_t$ and the game moves to the next period.

Under the timing convention I adopt, however, the notation and the analysis are cleaner.

The focus of this paper, however, is not to model what the event stands for or the strategic interaction it gives rise to upon its arrival. Rather, it is to investigate how *uncertainty* about *whether and when* superior outside opportunities are available and *learning* about their availability during negotiations affect the bargaining relationship. In this spirit, I replace the event with the continuation payoffs $O^S(v)$ and $O^B(v)$ that would arise upon its arrival. Assumption 3 below introduces the relevant restrictions on the payoffs $O^S(v)$ and $O^B(v)$. These restrictions only reflect the motivation of the paper and are minimal. Therefore, the framework is flexible enough to capture a variety of applications.[8]

Assumption 3. *Throughout the paper, I assume the following.*

A1 $O^S(v)$ *and* $O^B(v)$ *are non-negative differentiable functions.*

A2 $v - O^B(v)$ *is non-decreasing on* $[\underline{v}, \overline{v}]$.

A3 $O^S(v)$ *is either a constant or increasing on* $[\underline{v}, \overline{v}]$.

A4 $v/O(v)$ *is increasing on* $[\underline{v}, \overline{v}]$.

A5 *Learning is non-trivial. That is, there exists* $v^* > \underline{v}$ *such that*

$$\frac{\mu^0 \lambda}{\mu^0 \lambda + r} O(v) > v \qquad \textit{for all } v < v^*.$$

Part A1 is a technical requirement that simplifies the exposition. It does not impose restrictions affecting the main insights of the model. Part A4 ensures that there is non-trivial heterogeneity in buyer types.

To understand A2, note that $v - O^B(v)$ is a measure of how eager to trade the buyer of type v is. When $v - O^B(v)$ is non-decreasing, higher types are more eager to trade and the *skimming property* holds (see Section 3.2.2 for the details). For part A3, note that $O^S(v)$ is a measure of how attractive to the seller the outside opportunity is. Assumption A3 states that the seller's payoff upon the arrival of the event either does not depend on the buyer's type or, if it does, the seller prefers higher v's upon the arrival of the event.

Part A5 is central to the paper. It says that at time zero the value of waiting for outside opportunities is larger than the value of trading, at least for a positive measure of buyer types. Thus, gains from trade are ex ante uncertain. The uncertainty unravels over time if players postpone reaching an agreement and engage in market experimentation. Therefore, some delay in the transaction may be bilaterally efficient. Importantly, while players may individually learn over time that there are gains from trade, this fact does not necessarily become common knowledge. Whether it does so in equilibrium depends on the specific assumptions on the model and affects trading dynamics and other equilibrium outcomes. I will discuss this point extensively in the next sections. Note, however, that Part A5 is silent with

[8] The same argument justifies the assumption that the arrival of the event is public and concludes the game. In particular, this paper neither studies the role of transparency of outside options on bargaining dynamics (see, e.g., Hwang and Li (2017)), nor bargaining dynamics in the shadow of preexisting outside options that players may decide to exercise (see, e.g., Lee and Liu (2013) and Board and Pycia (2014)) during their negotiations.

respect to the two parties' individual incentives to actually postpone or advance the transaction in time. Finally, part A5 implies that the joint surplus associated to the arrival of the event is larger than the joint surplus from the transaction for a positive measure of buyer types. That is, in a market of type $m = 1$ there are superior opportunities available to the two players, at least in terms of joint surplus to share.

The analysis will make clear the role of the restrictions in Assumption 3. Meanwhile, note that they are natural in the settings this paper models. In Section 3.5, I extend the analysis to situations where A1–A5 fail and discuss the robustness of the main insights of the model with respect to these assumptions.

Finally, I assume that the game is common knowledge among the players, which is standard.

Example 7. There are three natural benchmark specifications of the general model, each of them corresponding to a different market configuration.

Sellers' Market. When $O^B(v) = 0$ for all $v \in [\underline{v}, \overline{v}]$, the buyer does not reap any benefit from the potential arrival of the event. This is an extreme form of a sellers' market, where superior opportunities, if existing, only benefit the seller.

Buyers' Market. When $O^S(v) = 0$ for all $v \in [\underline{v}, \overline{v}]$, the seller does not reap any benefit from the potential arrival of the event. This is an extreme form of a buyers' market, where superior opportunities, if existing, only benefit the buyer.

General Market. In a general market, post-arrival payoffs are non-trivial for both parties. One way to think of a general market is to assume that the arrival of the event alters both parties' payoffs at the same time. Alternatively, one may assume that, upon arrival, the event is favorable to the seller with probability α and to the buyer with probability $1 - \alpha$. Since players are risk neutral, this is a parsimonious way to model the possibility that either side of the transaction may benefit from the existence of superior opportunities independently of the other side.

Example 8. Here are a few examples of what the event may represent.

1. The event may correspond to the arrival of a new buyer coming to offer the seller price c. In this case, $O^S(v) = c$ and $O^B(v) = 0$ for all v. As long as $\mu^0\lambda c/(\mu^0\lambda + r) > \underline{v}$, Assumption 3 is satisfied. Similarly, the event may correspond to the arrival of a short-lived buyer with high valuation, say V. This valuation is known to the seller who can then charge the buyer her willingness to pay V upon arrival. In this case, $O^S(v) = V$ and $O^B(v) = 0$ for all v. As long as $\mu^0\lambda V/(\mu^0\lambda + r) > \underline{v}$, Assumption 3 is satisfied.

2. The event may correspond to the arrival of a new seller leaving all buyer types with a surplus of $V > v$ (e.g, by offering an upgraded version of the good). In this case, $O^S(v) = 0$ and $O^B(v) = V$ for all v. As long as $\mu^0\lambda V/(\mu^0\lambda + r) > \underline{v}$, Assumption 3 is satisfied. Alternatively, the new seller may leave the buyer of type v with surplus $v + V$ for some $V > 0$. In this case, $O^S(v) = 0$ and $O^B(v) = v + V$. As long as $\mu^0\lambda(\underline{v} + V)/(\mu^0\lambda + r) > \underline{v}$, Assumption 3 is satisfied.

3. The event may correspond to the arrival of a new short-lived buyer replacing the original byers. The new buyer's valuation $\tilde{v}$ is uniformly distributed over $[v, v + 1]$, whereas the original buyer's

valuation v is uniformly distributed over $[0,1]$. Let $\tilde{F}$ be the c.d.f. of $\tilde{v}$. In this case, the seller offers price $(1+v)/2 = \arg\max_p p(1-\tilde{F}(p))$ to the new buyer and trade occurs with probability $1-\tilde{F}((1+v)/2) = (1+v)/2$. In this case, $O^B(v) = 0$ and $O^S(v) = (1+v)^2/4$. As long as $\left[\mu^0\lambda(1+\underline{v})^2/4\right]/(\mu^0\lambda + r) > \underline{v}$, Assumption 3 is satisfied.

4. Finally, the event may correspond to the realization from some probability distribution over the previous cases.

The Benefits of Continuous Time. I formulate the model directly in continuous time for two reasons. First, continuous time captures the idea that the seller loses all his commitment power and/or that there are no institutional frictions in the bargaining protocol besides private information.[9] As a consequence, my analysis clearly disentangles the implications of learning about the existence of (possibly superior) outside opportunities for bargaining dynamics from that of other frictions in the trading environment. Second, equilibrium strategies in discrete-time bargaining games are in general analytically intractable. In contrast, they are easier to characterize in continuous time. Moreover, continuous-time methods are particularly suitable to perform the option value calculations that arise when studying learning problems of this kind. As a result, I will be able to describe optimality conditions, as well as equilibrium strategies and outcomes, by means of Hamilton-Jacobi-Bellman equations and (solutions to) partial and ordinary differential equations which carry a clear economic intuition. Closed-form solutions and relatively simple expressions for all equilibrium outcomes open the doors to comparative statics. In addition, when the model is specialized to particular applications, closed-form solutions yield sharp predictions for empirical studies and more applied research.

3.2.2 Strategies and Equilibrium Notion

Preliminaries

There are well-known technical issues that arise when modeling games in continuous time (see, in particular, Simon and Stinchcombe (1989) and Bergin and MacLeod (1993)). To address these issues, I introduce an ad hoc equilibrium concept for the bargaining game I study. The equilibrium notion, which builds on Daley and Green (2018) and Ortner (2017), captures a set of basic properties that would hold in any perfect Bayesian analysis of a discrete-time counterpart of the model.[10] These properties are the following.

Property 1. The buyer solves an optimal stopping problem. Given her type, the evolution of the common belief on the type of the market, the seller's pricing rule, and conditional on the event not having occurred, the buyer decides when to accept the offer and conclude the bargaining process.

Property 2. The buyer types remaining at the end of each time "period" are a truncated sample of the original distribution. This is the so called *skimming property*. By Assumption 3–A2, $v - O^B(v)$ is non-decreasing on $[\underline{v}, \overline{v}]$, which implies that it is more costly for the high types to delay trade than

[9] The two interpretations are mathematically equivalent

[10] See also Chaves Villamizar (2018).

it is for the low types. Thus, at the end of each time "period", the pool of remaining buyer types is a right-truncation of the original type distribution, implying that negative selection in the poll of buyer types occurs in equilibrium.[11]

Property 3. The current truncation of the original type distribution describes the seller's current belief on the buyer type. Therefore, given Property 2, the type defining the current truncation (hereafter, the *cutoff type*), together with the current belief on the type of the market, describe the payoff-relevant state of the game on which players can condition their strategies. In particular, this is so for stationary strategies, where the seller (resp., the buyer) conditions his price offers (resp, her acceptances) at each point in time only on the current cutoff type and the current belief.

Property 4. To any given equilibrium price history, there corresponds a history of realized cutoff types. Thus, along the equilibrium path, the seller can be thought of as choosing his own future beliefs on buyer types (as described by the future path of cutoff types) as a function of his current belief (as described by the current cutoff type). That is, the seller can be thought of as choosing how quickly to screen through buyer types instead of choosing prices.

Property 5. At each point in time, the willingness to pay of the buyer of type v is the difference between her valuation and the current present discounted value of waiting for the outside opportunity. At time t this difference is

$$v - \frac{\mu_t \lambda O^B(v)}{\lambda + r}. \tag{3.3}$$

In any perfect Bayesian analysis of a discrete-time counterpart of the model, it is straightforward to show that if the seller proposes at time t a price that is smaller than (3.3) for *all* buyer types that have not traded by time t, then all remaining types accept the price offer and the game concludes.[12] However, such a time and non-negative price combination need not exist. When $\underline{v} = 0$, (3.3) is negative for a positive-measure subset of buyer types at any time $t \in \mathbb{R}_+$. Therefore, there is no finite time at which the seller can make a non-negative price offer that all remaining buyer types accept.[13]

In the next two subsections I build on the previous insights to introduce the equilibrium notion. Formalizing consistently players' strategies and equilibrium conditions in continuous time requires the introduction of some technical concepts and notation. Before moving to the next subsections, I note the following.

- Let $(\Omega_A, \mathcal{A}, \mathbb{P}_A)$ be the (sufficiently rich) probability space where the Poisson process governing the arrival of the event when $m = 1$ is defined. Moreover, let $\mathbb{A} := (\mathcal{A}_t)_{t \geq 0}$ be the natural filtration of $(\Omega_A, \mathcal{A}, \mathbb{P}_A)$ associated to the process.

  For all $t \in \mathbb{R}_+$, if neither the event nor trade has occurred by time t, the two players share a common belief μ_t on $m = 1$. This common belief is derived by Bayes rule given the information

[11] Under Assumption 3-A2, the skimming property follows by standard arguments, which I thus omit. See, for instance, Fudenberg et al. (1985).

[12] I refer to Fudenberg and Tirole (1991) for an argument along these lines. A similar reasoning applies here, once extended to the setting I consider.

[13] A negative price offer that is accepted by a positive-measure subset of buyer types can never be part of an equilibrium.

available at the time. Formally, the process $(\mu_t)_{t\geq 0}$ is a time-homogeneous $\mathcal{A}_t$-Markov process adapted to $\mathbb{A}$. It is càdlàg and piecewise differentiable, with a random jump time T defined by the arrival of the event, and satisfies the ODE in (3.1) for all $t < T$, with initial condition $\mu_0 = \mu^0$; the common belief μ_t on $m = 1$ is given by (3.2) for all $t < T$, with $\mu_t = 1$ for $t \geq T$.[14]

From the viewpoint of the two players, the arrival of the event follows a Poisson process with subjective intensity function $t \mapsto \lambda_t := \mu_t \lambda$. Denote with T_μ the first jump time associated to this Poisson process with initial condition $\lambda_0 = \mu\lambda$. The distribution of T_μ is uniquely determined by the law of the process $(\mu_t)_{t\geq 0}$ and T_μ is adapted to the filtration $\mathbb{A}$.

- For technical convenience, consider the auxiliary filtered probability space $(\Omega_F, \mathcal{F}, \mathbb{F}, \mathbb{P}_F)$ where: *(i)* the underlying probability space $(\Omega_F, \mathcal{F}, \mathbb{P}_F)$ is complete; *(ii)* the filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 0}$ of the probability space $(\Omega_F, \mathcal{F}, \mathbb{P}_F)$ is both complete and right-continuous. Assume that the filtration $\mathbb{F}$ is independent of all other stochastic elements in the model (namely, the buyer type and the arrival of the event).

- Let $\left\{(K_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\right\}$ be a class of non-increasing càdlàg stochastic processes adapted to the filtration $\mathbb{F}$, describing the possible paths of the cutoff type, one path for each initial cutoff type $K_0 = k \in [\underline{v}, \overline{v}]$.

- I refer to any realization $(k, \mu) \in [\underline{v}, \overline{v}] \times [0, \mu^0]$ of the process $\big((K_t, \mu_t)\big)_{t\geq 0}$ as the *state* of the game.

Equilibrium Conditions

To begin, I lay out the components of and requirements for equilibrium.

Stationarity. In keeping with the literature, I focus on behavior that is stationary, using the current cutoff type and the current belief about the type of the market as state variables. Stationarity requires that, as long as the game is still in place, both the current price offer and the evolution of the cutoff type depend only on the current state of the game. Formally, we have the following.

Equilibrium Condition 1 (Stationarity). The seller's price offer in state (k, μ) is given by $P(k, \mu)$, where $P\colon [\underline{v}, \overline{v}] \times [0, \mu^0] \to \mathbb{R}$ is a Borel-measurable function and $(K_t)_{t\geq 0}$, $K_t = k$, is a time-homogeneous $\mathcal{F}_t$-Markov process.

Define the filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P}) := \big(\Omega_F \times \Omega_A, \mathcal{F} \otimes \mathcal{A}, (\mathcal{F}_t \otimes \mathcal{A}_t)_{t\geq 0}, \mathbb{P}_F \times \mathbb{P}_A\big)$, and let $\mathcal{G}_t := \mathcal{F}_t \otimes \mathcal{A}_t$ for all $t \geq 0$. Note that $\left\{(K_t, \mu_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\right\}$ is a class of processes defined over the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and adapted to the filtration $\mathbb{G}$. By stationarity, $\left\{\big((K_t, \mu_t)\big)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\right\}$ is a class of time-homogeneous $\mathcal{G}_t$-Markov processes.

[14]As the arrival of the event concludes the game, any time $t > T$ does not play any role in the analysis, neither does so any belief $\mu_t = 1$.

Buyer's Problem. The buyer takes the price offer function P and the law of motion for $\big((K_t, \mu_t)\big)_{t\geq 0}$ as given. A pure strategy for the buyer of type v is a $\mathcal{G}_t$-adapted stopping time (when to accept) $\tau^v : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$. Let $\mathcal{T}$ be the set of all $\mathcal{G}_t$-adapted stopping times. For all states $(k, \mu) \in [\underline{v}, \overline{v}] \times [0, \mu^0]$, the buyer of type v solves the following optimal stopping problem:

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}^K_{(k,\mu)} \Big[\mathbb{1}_{\{\tau < T_\mu\}} e^{-r\tau} \big(v - P(K_\tau, \mu_\tau) \big) + \mathbb{1}_{\{\tau \geq T_\mu\}} e^{-rT_\mu} O^B(v) \Big], \tag{BP^v}$$

where $\mathbb{E}^K_{(k,\mu)}$ is the expectation with respect to the law of the process $\big((K_t, \mu_t)\big)_{t\geq 0}$ induced by $\big\{(K_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]$ conditional on $(K_{0^-}, \mu_{0^-}) = (k, \mu)$. Henceforth, when the random jump time T_μ appears as the argument of the expectation $\mathbb{E}^K_{(k,\mu)}$, it has to be interpreted as the first jump time of a Poisson process with intensity function $t \mapsto \lambda_t := \mu_t \lambda$ and initial condition $\lambda_{0^-} = \mu\lambda$. Here, the first term in the expectation reflects the surplus from trading before the arrival of the event, and the second part stands for the possibility that the event occurs before time τ.

Given any price offer function P and process $\big((K_t, \mu_t)\big)_{t\geq 0}$, when $v - O^B(v)$ is non-decreasing (Assumption 3-A2), the buyer's objective satisfies increasing differences in $(-\tau, v)$, and any selection of maximizers of the above problem will be decreasing in v. Therefore, higher types will accept sooner, and at any time t, for any history of prices offers and $T_{\mu^0} > t$, there will exist some cutoff type k_t such that all $v \geq k_t$ would have accepted weakly before t. In particular, if the seller were to observe the buyer respond to any P according to (BP^v), then his beliefs about v at time t conditional on no acceptance would be right-truncation of F at k_t. In short, the skimming property always holds in equilibrium. Therefore, it is without loss to restrict attention to pure strategies.

Equilibrium Condition 2 (Buyer's Optimality)**.** Let τ^v be the $\mathcal{G}_t$-adapted stopping time chosen by the buyer of type v. Given the price offer function P and the law of the process $\big((K_t, \mu_t)\big)_{t\geq 0}$, τ^v solves the optimal stopping problem (BP^v).[15]

Given stationarity, the buyer's value function depends only on the current state. In particular, the expected payoff of the buyer of type v when the game is in state (k, μ), denoted by $B^v(k, \mu)$, is

$$B^v(k, \mu) := \mathbb{E}^K_{(k,\mu)} \Big[\mathbb{1}_{\{\tau^v < T_\mu\}} e^{-r\tau^v} \big(v - P(K_{\tau^v}, \mu_{\tau^v}) \big) + \mathbb{1}_{\{\tau^v \geq T_\mu\}} e^{-rT_\mu} O^B(v) \Big]. \tag{3.4}$$

Consistency. If neither the event nor trade has occurred by time t, the seller's belief about the buyer type is conditioned on the fact that the buyer has rejected all past offers. This belief is summarized by the current cutoff type, where "cutoff type at time t" should be interpreted to mean before observing the buyer's decision at time t. Consistency simply requires that the cutoff type is derived from the buyer's optimal strategy.

Equilibrium Condition 3 (Consistency)**.** For all $t < T$,

$$K_t = \underline{v} + \int_{\underline{v}}^{\overline{v}} \mathbb{1}_{\{\tau^v \geq t\}} \mathrm{d}v. \tag{3.5}$$

[15]Note that τ^v does not specify how to handle off-path price offers. This will be addressed by Equilibrium Condition 5.

Option for Immediate Trade. The next condition says that if the seller offers a price that is smaller than the willingness to pay of *all* buyer types that have not yet traded, then all remaining types accept the price offer and the game concludes.

Equilibrium Condition 4 (Option for Immediate Trade)**.** Let $(K_{t^-}, \mu_{t^-}) = (k, \mu)$. If

$$P(k, \mu) \leq v - \frac{\mu\lambda O^B(v)}{\lambda + r},$$

for all $v \in [\underline{v}, k]$, then $\tau^v = t$ for all $v \in [\underline{v}, k]$.

Seller's Problem. Instead of writing the seller's problem in terms of price offers, I will write it as an "optimal stopping + impulse control problem" over the seller's beliefs about the buyer type.[16] In this case, the seller's problem is to choose a stopping time $\overline{T}_S : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ at which he exercises the option for immediate trade, and a class of processes for cutoff types, $\{(Q_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\}$, one path for each initial cutoff type $Q_{0^-} = k \in [\underline{v}, \overline{v}]$, for the intensity of trade at any time $t \leq \min\{T_{\mu^0}, \overline{T}_S\}$.[17] One way to interpret this formulation is to think of the seller as setting quantities instead of prices: rather than choosing prices that induce rejections leading to belief cutoffs, the seller can choose cutoffs, evaluating them according to the prices that would be consistent with those cutoffs.[18]

Let the game be in state (k, μ). I refer to the pair $\gamma_k := \left(\overline{T}_S, \{(Q_t)_{t\geq 0}, Q_{0^-} = k\}\right)$ as a *policy*. A policy γ_k is *feasible* if $\overline{T}_S$ is a $\mathcal{G}_t$-adapted stopping time and $(Q_t)_{t\geq 0}$, $Q_{0^-} = k$, is a non-increasing, càdlàg, and $\mathcal{G}_t$-adapted process on $[\underline{v}, k]$. Let Γ_k denote the set of such feasible policies and define $\overline{T}_\mu^{\min} := \min\{T_\mu, \overline{T}_S\}$.

Equilibrium Condition 5 (Seller's Optimality)**.** For all states $(k, \mu) \in [\underline{v}, \overline{v}] \times [0, \mu^0]$, given the price offer function P and the law of the process $(\mu_t)_{t\geq 0}$, $\left(T_S, \{(K_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\}\right)$ solves

$$\sup_{\gamma_k \in \Gamma_k} \mathbb{E}^Q_{(k,\mu)} \left[\int_0^{\overline{T}_\mu^{\min}} e^{-rt} P(Q_t, \mu_t) \mathrm{d}F(Q_t) + e^{-r\overline{T}_\mu^{\min}} \underline{O}^S\left(Q_{\overline{T}_\mu^{\min}}\right) \right], \tag{3.6}$$

where $\mathbb{E}^Q_{(k,\mu)}$ is the expectation with respect to the law of the process $\left((Q_t, \mu_t)\right)_{t\geq 0}$ induced by $\{(Q_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\}$, conditional on $(Q_{0^-}, \mu_{0^-}) = (k, \mu)$.

Given stationarity, the seller's value function depends only on the current state. In particular, the seller's expected payoff when the game is in state (k, μ), denoted by $S(k, \mu)$, is

$$S(k, \mu) := \mathbb{E}^Q_{(k,\mu)} \left[\int_0^{T_\mu^{\min}} e^{-rt} P(K_t, \mu_t) \mathrm{d}F(K_t) + e^{-rT_\mu^{\min}} \underline{O}^S\left(K_{T_\mu^{\min}}\right) \right],$$

where $T_\mu^{\min} := \min\{T_\mu, T_S\}$.

[16] For a standard reference on impulse control problems, see Harrison (2013).

[17] The seller's problem can also be stated as an impulse control problem only. However, the formulation as an "optimal stopping + impulse control problem" simplifies the exposition.

[18] Formally dealing with continuation play following deviations from P poses well-known existence problems in a continuous-time setting (again, see Simon and Stinchcombe (1989) and Bergin and MacLeod (1993)). Thus, it would require a substantially more complicated set of available strategies for the seller.

Reservation Price Strategies. Suppose that the game is in state (k, μ) at time t. The intensity of trade at time t, $\mathrm{d}K_t$, determines the belief about the buyer type conditional on rejection according to condition (3.5). Therefore, the price at time t must be the expected payoff of the cutoff type at time t conditional on accepting the offer; that is,

$$P(k, \mu) = k - B^{v=k}(k, \mu), \tag{3.7}$$

where $B^k(k, \mu)$ is defined by (3.4). Implicitly, (3.7) assumes that the seller can resolve buyers' indifference in his favor. In other words, $P(k, \mu)$ is not only the price offer in state $(K_t, \mu_t) = (k, \mu)$, but also the reservation price strategy for the buyer of type $v = k$. Given the interpretation of K_t as a "quantity", P has a corresponding interpretation as an (endogenous) inverse demand curve faced by the seller. Formally, we have the following.

Equilibrium Condition 6 (Reservation Price Strategies)**.** For all $(k, \mu) \in [\underline{v}, \overline{v}] \times [0, \mu^0]$, $P(k, \mu)$ is an optimal reservation price strategy for the buyer of type $v = k$, taking as given the law of the process $\big((K_t, \mu_t)\big)_{t \geq 0}$ and future prices given by $\big(P(K_t, \mu_t)\big)_{t \geq 0}$ induced by $\big\{(K_t)_{t \geq 0}, k \in [\underline{v}, \overline{v}]\big\}$, conditional on $(K_{0^-}, \mu_{0^-}) = (k, \mu)$; that is, $P(k, \mu) = k - B^k(k, \mu)$.

Note that the equilibrium condition on P only really species that the marginal type (the one whose type equals the state, $v = k$) is stopping optimally at $K_t = k$. However, since the buyer's problem satisfies a single-crossing property in type and stopping time, whenever $v = k$ wants to stop at $K_t = k$, so will any type $v' > k$, while all $v'' < k$ will want to continue. Therefore, all buyer types will be stopping optimally at all histories if $P(k, \mu)$ is an optimal reservation price for every $v = k$, and one can derive optimal stopping times at all states and for all types from P.

Remark 11. The filtration $\mathbb{F}$ and, ultimately, $\mathbb{G}$, serve as a public correlation device. In my model, however, the scope for randomization is limited. In particular, the only scope for randomization is in the seller's choice of paths for the cutoff types. As I will show below, the equilibria of the bargaining game mostly evolve deterministically along the path of play. Therefore, the main purpose of the filtration $\mathbb{G}$ is to allow me to define strategies and the equilibrium notion without incurring in the usual non-existence problems that arise when modeling games in continuous time.

Remark 12. Even though the buyer and the seller can in principle condition their strategies on the public correlation device $\mathcal{G}_t$, their equilibrium payoffs will depend on $\mathcal{G}_t$ only through (K_t, μ_t). Conditional on no event, the seller's future payoff depend on $\mathcal{G}_t$ only through $\big((K_{t+s}, \mu_{t+s})\big)_{s \geq 0}$. The seller chooses $(K_{t+s})_{s \geq 0}$ taking $K_t = k$ as given and $(\mu_{t+s})_{s \geq 0}$ evolves exogenously given $\mu_t = \mu$. Therefore, the seller's value function depends on $\mathcal{G}_t$ only through the current state $(K_t, \mu_t) = (k, \mu)$. Likewise for the buyer, at any time t v's future payoff depend on $\mathcal{G}_t$ only through $\big((K_{t+s}, \mu_{t+s})\big)_{s \geq 0}$, and the future law of motion for $\big((K_{t+s}, \mu_{t+s})\big)_{s \geq 0}$ depends only on the current state $(K_t, \mu_t) = (k, \mu)$. Therefore, v's value function depends on $\mathcal{G}_t$ only through $(K_t, \mu_t) = (k, \mu)$.

Remark 13. The "optimal-stopping + impulse-control" formulation of the seller's problem allows me to identify three qualitatively different dynamics: time intervals during which the probability mass of trade is infinitesimal, so that the seller screens buyer types one by one; bursts of trade, so that the seller screens through a positive mass of buyer types in an instant; and periods of silent trade, where no buyer type trades.

Regular Stationary Equilibria

Definition 13 (Stationary Equilibrium)**.** *A stationary equilibrium is a pair*

$$\Big(\big(T_S, \big\{(K_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\big\}\big), P\Big)$$

that satisfies Equilibrium Conditions 1-6.

I focus on a subset of stationary equilibria with the property that the seller alternates between periods of sufficiently gradual trade and a few instants with bursts of trade.

Definition 14 (Regular Stationary Equilibrium)**.** *A stationary equilibrium is regular if, for any initial condition $k \in [\underline{v}, \overline{v}]$,*

$$K_t = K_t^{abs} + K_t^{jump},$$

where K_t^{abs} is absolutely continuous in t and K_t^{jump} is a step function with finitely many jumps. The acronym RSE denotes a regular stationary equilibrium.

Hereafter, I use the term equilibrium to mean regular stationary equilibrium, thus omitting the "regular stationary" qualifier. Before discussing the equilibrium restriction, I introduce the following terminology.

Definition 15. *Let $s, \underline{s}, \overline{s} \in \mathbb{R}_+$, with $\underline{s} < \overline{s}$. In the RSE $\big(\big(T_S, \big\{(K_t)_{t\geq 0}, k \in [\underline{v}, \overline{v}]\big\}\big), P\big)$:*

(i) Trade is smooth over the time interval $[\underline{s}, \overline{s})$ if K_t is absolutely continuous over $[\underline{s}, \overline{s})$;

(ii) Smooth trade is silent over the time interval $[\underline{s}, \overline{s})$ if, in addition, $K_{\overline{s}} - K_{\underline{s}} = 0$;

(iii) There is a burst of trade at time s if $K_{s^-}^{jump} \neq K_s^{jump}$.[19]

If trade is smooth over the time interval $(\underline{s}, \overline{s})$ and $s \in (\underline{s}, \overline{s})$, I refer to $\dot{K}_s \in (-\infty, 0]$ as the *speed* of trade at time s. If there is a burst of trade at time t, I write $\dot{K}_t \in \{-\infty, +\infty\}$.

By the skimming property, the function K_t is monotone. Thus, it has a Lebesgue decomposition of the form $K_t = K_t^{abs} + K_t^{Jump} + K_t^{sing}$, where K_t^{abs} is an absolutely continuous function (in t), K_t^{Jump} is a piecewise constant jump function, and K_t^{sing} is a singular continuous function (i.e., a non-constant continuous function with first derivative equal to zero almost everywhere). Imposing regularity on K_t introduces two additional restrictions. First, it says that there are only finitely many jumps, implying that there are only finitely many bursts of trade in equilibrium. Second, it says that the continuous part of K_t is sufficiently smooth, implying that over a smooth trade region the buyer sees the price changing gradually over time, rather than K_t only moving in twitches.

[19] I use the convention that $K_{0^-}^{jump} = \overline{v}$.

Remark 14. In principle, the restriction to regularity rules out potential dynamics. However, it is worth noting the following.

- When values are interdependent and $\underline{v} > 0$ ("gap" case), one may worry that the continuous-time limit of the discrete-time analog of the model may exhibit dynamics that do not satisfy my equilibrium restriction. In particular, these dynamics may neither be regular nor generated by stationary strategies (see Deneckere and Liang (2006)).[20] In the spirit of Daley and Green (2018), however, one can show that singular dynamics do not arise in my setup even when $\underline{v} > 0$. In fact, for equilibrium to alternate between bursts of trade and silent periods, during a silent period the sellers' belief must be exactly such that the Coasian desire to speed up trade is absent. With the addition of learning, the seller's belief cannot remain constant at such a belief over any time interval. As a result, in contrast with Deneckere and Liang (2006)'s findings, singular dynamics do not arise in my model.

- The restriction to regularity is weaker than taking the continuous-time *atomless* limit of a selection of perfect Bayesian stationary equilibria of the corresponding discrete-time model. Such an exercise, in fact, would exclude bursts of trade (see, e.g., Fuchs and Skrzypacz (2010)).

Consequently, it is unclear whether the restriction to regularity rules out interesting dynamics, if any at all.

3.3 Benchmarks

In this section, I first characterize the bilaterally efficient trading dynamics, where the transaction occurs so as to maximize the sum of the two players' payoffs. Then, I present the stationary perfect Bayesian equilibria of the bargaining game where the event never occurs (i.e., where it is common knowledge that the market is of type $m = 0$). The two cases serve as natural benchmarks for the analysis to come.

3.3.1 Efficient Trading Dynamics

Consider a social planner who knows the buyer type and wishes to maximize the sum of the two players' payoffs from the transaction.[21] The planner has to choose the surplus-maximizing time at which parties stop waiting for outside opportunities and the seller serves the buyer. The planner's optimal stopping problem can be written as a dynamic programming problem where the planner's current belief on $m = 1$ serves as state variables. Formally, we have the following.

Fix a buyer type $v \in [\underline{v}, \overline{v}]$. First, suppose that $O(v) \leq v$. As the planner becomes increasingly pessimistic over time that the market is of type $m = 1$, under efficiency the transaction with such a buyer type occurs at belief μ_0 (equivalently, at time $t = 0$).

[20] Fuchs and Skrzypacz (2013a) show that equilibria become stationary and regular as $\underline{v} \to 0$.

[21] Here, the benchmark I consider is *first-best* efficiency.

Next, suppose that $O(v) > v$ and let μ be the planner's current belief on $m = 1$. The expected joint surplus from delaying trade so as to learn about the market environment, which I denote by $L(\mu; v)$, satisfies the continuous-time recursion

$$L(\mu; v) = \mu\lambda O(v)\mathrm{d}t + e^{-r\mathrm{d}t}\mathbb{E}\big[L(\mu + \mathrm{d}\mu; v) \mid \mu\big]. \tag{3.8}$$

Here, the first term on the right-hand side is the expected instantaneous surplus from waiting for news, where the event with associated joint payoff $O(v)$ occurs with subjective instantaneous probability $\mu\lambda \mathrm{d}t$; the second term is the discounted expected continuation surplus from waiting for news.[22] As to the latter, with subjective probability $\mu\lambda \mathrm{d}t$ the event occurs and the game ends, so that the expected surplus from waiting jumps to $L(1; v) = 0$; with subjective probability $\mu(1-\lambda \mathrm{d}t)+(1-\mu) = 1-\mu\lambda \mathrm{d}t$, no event occurs and, assuming that the function $L(\mu; v)$ is differentiable, the expected surplus from waiting changes to $L(\mu; v) + L'(\mu; v)\mathrm{d}\mu = L(\mu; v) - \lambda\mu(1 - \mu)L'(\mu; v)\mathrm{d}t$. Using these expectations, together with $1 - r\mathrm{d}t$ as an approximation to $e^{-r\mathrm{d}t}$ as $\mathrm{d}t \to 0$, I replace the second term in equation (3.8), simplify, and rearrange, to obtain that $L(\mu; v)$ satisfies the first-order ODE

$$(\mu\lambda + r)L(\mu; v) + \lambda\mu(1 - \mu)L'(\mu; v) = \mu\lambda O(v).$$

This has solution[23]

$$L(\mu; v) = \frac{\mu\lambda O(v)}{\lambda + r} + C_e(1 - \mu)\left(\frac{1 - \mu}{\mu}\right)^{r/\lambda}, \tag{3.9}$$

where C_e is the constant of integration.[24]

At any belief μ on $m = 1$, the joint surplus from the transaction with the buyer is v. Thus, by imposing the optimality conditions $L(\mu; v) = v$ (value matching) and $L'(\mu; v) = 0$ (smooth pasting) we obtain

$$\mu\lambda\big[O(v) - v\big] = rv. \tag{3.10}$$

Define $\widehat{\mu}(v) := rv/\lambda\big(O(v)-v\big)$. If $\widehat{\mu}(v) < \mu^0$, then $\widehat{\mu}(v)$ is the belief at which the planner stops waiting for the arrival of the event and the transaction takes place. If instead $\widehat{\mu}(v) \geq \mu^0$, it is never efficient to engage in market experimentation, and the transaction takes place at belief μ^0. That is, if we denote with $\mu(v)$ the efficient trading belief with the buyer, who has type v, we have

$$\mu(v) = \min\left\{\mu^0, \widehat{\mu}(v)\right\}.$$ [25]

The left-hand side of (3.10) is the expected value of a jump in the joint surplus from v to $O(v)$ should the event occur. This is the flow benefit of market experimentation at belief μ. The right-hand side is

[22] Henceforth, I refer to subjective instantaneous probabilities as subjective probabilities, omitting the instantaneous qualifier.

[23] I refer to Polyanin and Zaitsev (2003) for the closed-form solutions to the ODEs that appear in the paper.

[24] The closed-form solution to the ODE shows that $L(\mu; v)$ is differentiable in μ, and so it was legitimate to assume differentiability.

[25] Optimality of the planner's strategy follows by standard verification arguments. To justify smooth pasting, it is enough to determine the constant of integration C from value matching and to check that $\mu(v)$ maximizes the planner's objective with respect to μ.

the flow cost, due to discounting, of postponing a transaction whose joint value is v. This is the cost of market experimentation. Condition (3.10) is thus intuitive: as long as there is sufficient optimism on the market being of type $m = 1$, the parties engage in costly experimentation. As soon as the benefit of waiting for the event equates the cost of delaying trade, however, the transaction occurs.

The previous characterization determines the buyer types that trade at belief $\mu < \mu_0$ (equivalently, at time $t_\mu > 0$). These are all types $v \in [\underline{v}, \overline{v}]$ such that $\mu\lambda(O(v) - v) = rv$ or, equivalently,

$$\frac{\mu\lambda}{\mu\lambda + r} = \frac{v}{O(v)}. \tag{3.11}$$

By Assumption 3–A4, the ratio $v/O(v)$ is increasing, implying that at most one $v \in [\underline{v}, \overline{v}]$ solves (3.11) for $t_\mu > 0$. The next proposition summarizes.

Proposition 12 (Efficient Trading Dynamics)**.** *Suppose trade is efficient. Then:*

(i) At time $t = 0$, trade occurs with all buyer types $v \in [\underline{v}, \overline{v}]$ for which $\mu^0\lambda O(v) \leq (\mu^0\lambda + r)v$. At time $t_\mu > 0$, trade occurs with the buyer type $v \in [\underline{v}, \overline{v}]$ satisfying

$$\frac{\mu\lambda}{\mu\lambda + r} = \frac{v}{O(v)} \qquad \text{or, equivalently,} \qquad k = \frac{\mu\lambda O(k)}{\mu\lambda + r}. \tag{3.12}$$

(ii) The efficient trading time with the buyer of type $v \in [\underline{v}, \overline{v}]$, denoted by $t(v)$, is

$$t(v) = \begin{cases} 0 & \text{if } v \geq \frac{\mu^0\lambda O(v)}{\mu^0\lambda + r} \\ T_v & \text{if } v < \frac{\mu^0\lambda O(v)}{\mu^0\lambda + r} \end{cases}, \tag{3.13}$$

where T_v solves $\mu_t\lambda O(v) = (\mu_t\lambda + r)v$ for t.

The next remarks follow immediately from Proposition 12 and describe the main properties of the efficient trading dynamics.

(a) The left-hand side of (3.11) strictly decreases over time. Thus, as $v/O(v)$ is strictly increasing by assumption A4, learning endogenously gives rise to negative selection in the buyer type distribution.

(b) Trade may begin with a burst or after a silent period with no trade. There is a burst of trade at $t = 0$ if $v \geq \mu^0\lambda O(v)/(\mu^0\lambda + r)$ holds for buyers types in a positive-measure subset of $[\underline{v}, \overline{v}]$. Trade begins silently if, instead, $v < \mu^0\lambda O(v)/(\mu^0\lambda + r)$ for all buyer types. After trade begins, it proceeds smoothly until the end, as $v/O(v)$ is strictly monotone. Finally, trade may also begin smoothly (and proceed so afterward), This happens when $\overline{v} = \mu^0\lambda O(\overline{v})/(\mu^0\lambda + r)$. The latter case, however, is non-generic in the space of parameters.

(c) The market may or may not clear in finite time. In particular, whether the market clears in finite time depends on whether the lowest buyers valuation is positive ("gap" case) or not ("no gap" case). If $\underline{v} > 0$, the last instant of trade is at the finite time $T^e_{\underline{v}}$ at which $\mu_{T^e_{\underline{v}}}/(\mu_{T^e_{\underline{v}}}\lambda + r) = \underline{v}/O(\underline{v})$;

if, instead, $\underline{v} = 0$, the market does not clear in finite time, as the right-hand side of (3.11) is equal to zero when evaluated at $\underline{v} = 0$, whereas its left-hand side is positive at any finite time.

The rich trading dynamics under efficiency are entirely driven by learning. Without arrivals (i.e., $\mu^0 = 0$) or without learning (i.e., $\mu^0 = 1$) the efficient benchmark is trivial and interesting dynamics are absent. In particular, if there are no arrivals, under efficiency all buyer types trade at time $t = 0$. If, instead, there is no learning, efficiency prescribes that trade occurs at time $t = 0$ with all buyer types v for which $v \geq \lambda O(v)/(\lambda + r)$, and never occurs with all types v for which $v < \lambda O(v)/(\lambda + r)$ (when $v < \lambda O(v)/(\lambda + r)$, it is efficient to wait for the joint surplus to grow).

3.3.2 The Bargaining Game without Arrivals

Assume that $\mu^0 = 0$ (or, equivalently, $\lambda = 0$), but that the bargaining game remains otherwise the same. In this case, the event never arrives and learning plays no role. Suppose that time is continuous but that the seller makes price offers at times $t = 0, \Delta, 2\Delta, \ldots$, so that the model reduces to the canonical bargaining game with one-sided incomplete information studied in the seminal contributions of Stokey (1981), Bulow (1982), Fudenberg et al. (1985), Gul et al. (1986), and Ausubel and Deneckere (1989).[26] In any stationary perfect Bayesian equilibrium of the game, as $\Delta \to 0$ (i.e., in the frictionless, continuous-time limit of arbitrarily frequent offers, where the seller loses all his commitment power):

(i) The initial price offer converges to $\underline{v}$, the lowest buyer type;

(ii) The expected time to trade converges to zero (no delay);

(iii) All buyer types are served at the same time and at the same price (there is neither intertemporal nor price discrimination);

(iv) The outcome of "freestyle" bargaining is efficient;

(v) The seller's profit converges to $\underline{v}$ and so the seller is unable to extract rents from the buyer with higher valuations.

These are the classic Coase Conjecture dynamics, named so after Coase (1972).[27] The result holds because a monopolist seller lacking the ability to commit to future prices faces the competition of his own future selves, thereby dissipating all of his own monopoly power. After the buyer rejects the initial price offer, the seller would necessarily benefit by lowering prices so as to sell to the buyer with lower valuations who did not yet purchase. Thus, prices would decline after each offer. A forward-looking buyer expecting prices to fall would then be unwilling to pay the initial high price. Consequently, if

[26] The assumption that the seller makes offers only at times $t = 0, \Delta, 2\Delta, \ldots$ is made for a direct comparison with the results in the existing literature, which considers the bargaining game in discrete time and takes the limit of the period length to zero.

[27] In the "gap" case (i.e., $\underline{v} > 0$), for any $\Delta > 0$, there exists a unique perfect Bayesian equilibrium (generically) and this equilibrium is stationary. In the "no gap" case (i.e., $\underline{v} = 0$), a stationary equilibrium exists, but there may be perfect Bayesian equilibria which are not stationary. The Coase Conjecture fails when consumers use non-stationary strategies (see Ausubel and Deneckere (1989)). As I focus on the stationary equilibria of my model, I compare my results to those that arise in the stationary equilibria of prior models.

the time between offers were to vanish, the opening price would converge to the lowest buyer type and the transaction would occur at the opening of the negotiations.

As I will argue in the next sections, bargaining dynamics significantly depart from the Coasean benchmark if parties are uncertain about whether and when superior outside opportunities become available. Some Coasean forces, however, will still be present; I will explain how they generalize to or need to be reinterpreted in the present environment.

3.4 Equilibrium Characterization

In this section, I characterize the regular stationary equilibria of the bargaining game. To gain insights, I proceed in two steps: in Section 3.4.1, I consider the case of independent private values; in Section 3.4.2 I consider the case of interdependent values.

I proceed by construction. To begin, I characterize bargaining dynamics over smooth trade regions. By the definition of RSE, there is at least one such region in equilibrium, unless all buyer types trade in a single instant. A simple argument, however, excludes that equilibrium consists of a single burst of trade. I then show that there exists a unique candidate RSE. In the candidate RSE, bargaining dynamics are mostly determined by smooth trade. I characterize when trade is smooth, when it is smooth and silent, and when a burst of trade occurs. To identify the unique candidate RSE I only rely on the necessary optimality conditions for the seller's and the buyer's problems. A standard verification argument shows that the candidate is indeed an equilibrium. While I carefully explain the steps of equilibrium construction in the main text, I refer to Appendix 3.7 for the more technical details of the analysis.

3.4.1 Bargaining with Independent Private Values

Seller's Problem. Consider any state (k, μ) in the interior of a smooth-trade region. Since the probability of leaving the interior of such a region in the next $\mathrm{d}t$ is negligible, the seller's expected payoff in state (k, μ), $S(k, \mu)$, satisfies the Hamilton-Jacobi-Bellman (hereafter, HJB) equation

$$rS(k,\mu) = \sup_{\dot{K}\in(-\infty,0]} \Bigg\{ \mu\lambda\big[O^S - S(k,\mu)\big] + \big[P(k,\mu) - S(k,\mu)\big]\frac{f(k)}{F(k)}\big(-\dot{K}\big) + S_1(k,\mu)\dot{K} + S_2(k,\mu)\dot{\mu} \Bigg\}, \tag{3.14}$$

where $S_1(k, \mu)$ (resp., $S_2(k, \mu)$) denotes the partial derivative of $S(k, \mu)$ with respect to its first (resp., second) argument.[28] Condition (3.14) has a direct interpretation. The left-hand side is the seller's expected equilibrium payoff expressed in flow terms. The right-hand side represents the possible sources of the flow: upon arrival of the event, which happens with a subjective probability flow $\mu\lambda$, the

[28] I provide below the closed-form expression for $S(k, \mu)$, which is differentiable. Thus, it is legitimate to assume differentiability. Hereafter, I omit to state this argument when taking (partial) derivatives of a (payoff) function whose closed-form solution is derived throughout the analysis.

game ends with the seller earning O^S and forgoing $S(k,\mu)$. With a flow probability $[f(k)/F(k)](-\dot{K})$ the buyer accepts the current offer, $P(k,\mu)$, which also ends the game with the seller earning $P(k,\mu)$ and forgoing $S(k,\mu)$. Finally, if the game does not end immediately, the continuation payoff changes by $S_1(k,\mu)\dot{K} + S_2(k,\mu)\dot{\mu}$.

The right-hand side of (3.14) is linear in $\dot{K}$. This linearity is the source of Coasean dynamics when outside opportunities or learning about their existence are absent. In that case, for any non-decreasing price offer function, the seller wants to run down the demand function as fast as possible. In particular, without outside opportunities ($\mu^0 = 0$ or, equivalently, $\lambda = 0$), the equilibrium $P(k)$ becomes flat at $\underline{v}$ and trade happens immediately (see Section 3.3.2); similarly, without learning ($\mu^0 = 1$), the equilibrium $P(k)$ becomes flat at $p := \lambda O^S/(\lambda + r)$ and either trade happens immediately (if $v \geq p$) or parties wait for the arrival of the event (if $v < p$) (see Section 3.5.2). In contrast, in the setting I study, the two parties' option value of waiting to learn about the existence of superior outside opportunities act as a counterbalance to the seller's temptation to run down the demand curve instantaneously. Interestingly, the force against immediate trade is present in all market configurations (buyers', sellers', and general market). One may conjecture that, in a buyers' market, the seller's incentives to reach an immediate agreement are even stronger than in the standard bargaining model. This is so because the arrival of the event prevents the seller from concluding any transaction and from reaping any benefit from the relationship. In contrast with this intuition, however, the equilibrium analysis will show that the buyer's option value of waiting to learn prevents agreements with all types at time zero.

Since the right-hand side of (3.14) is linear in $\dot{K}$, the sum of the coefficients on $\dot{K}$ must be non-negative on the interior of a smooth-trade region. In fact, if the coefficients on $\dot{K}$ in (3.14) added up to something negative, the seller would maximize his payoff by trading as fast as possible, that is by setting $\dot{K} = -\infty$, which is incompatible with smooth trade. Thus, the coefficients on $\dot{K}$ either add up to zero or to something positive. The seller finds it optimal to set $\dot{K} = 0$ (i.e., silent trade) if the coefficients on $\dot{K}$ add up to something positive. The next lemma, whose proof also uses the necessary conditions for the buyers' problem (see below), shows that $\dot{K} = 0$ cannot occur after a positive measure of buyer types has traded.

Lemma 15. *In any RSE of the bargaining game with independent private values, if $K_t = k < \overline{v}$, then $\dot{K}_s < 0$ for all $s \geq t$.*

Suppose that $\dot{K} < 0$. In this case, the coefficients on $\dot{K}$ must add up to zero, which means that the seller must be indifferent between speeds of trade. Setting the coefficients on $\dot{K}$ to zero in (3.14) yields the partial differential equations (hereafter, PDEs)

$$S_1(k,\mu) = \big[P(k,\mu) - S(k,\mu)\big]\frac{f(k)}{F(k)}, \tag{3.15}$$

$$S_2(k,\mu)\dot{\mu} = (\mu\lambda + r)S(k,\mu) - \mu\lambda O^S, \tag{3.16}$$

which describe the seller's best response problem on the interior of a smooth-trade region where $\dot{K} < 0$ of any candidate equilibrium. The PDE in (3.16) has general solution

$$S(k,\mu) = \frac{\mu\lambda O^S}{\lambda + r} + C_{ipv}\frac{(1-\mu)^{(\lambda+r)/\lambda}}{\mu^{r/\lambda}}, \tag{3.17}$$

where C_{ipv} is the constant of integration. Note that, by (3.17), the seller's expected payoff in state (k, μ) does not depend on the current cutoff type k; therefore,

$$S_1(k,\mu) = 0. \tag{3.18}$$

As $f(k)/F(k) > 0$ by assumption, (3.15) and (3.18) yield

$$P(k,\mu) = S(k,\mu), \tag{3.19}$$

that is, on the interior of a smooth trade region, equilibrium prices coincide with the seller's expected payoff. Therefore, by (3.16) and (3.19), the price offer function must satisfy the PDE

$$P_2(k,\mu)\dot{\mu} = (\mu\lambda + r)P(k,\mu) - \mu\lambda O^S, \tag{3.20}$$

where $P_2(k, \mu)$ denotes the partial derivative of $P(k, \mu)$ with respect to its second argument. The next lemma summarizes the previous discussion.

Lemma 16. *In any RSE of the bargaining game with independent private values, smooth-trade prices are determined by the seller's indifference between speeds of trade. In particular,*

$$P(k,\mu) = \frac{\mu\lambda O^S}{\lambda + r} + C_{ipv}\frac{(1-\mu)^{(\lambda+r)/\lambda}}{\mu^{r/\lambda}}, \tag{3.21}$$

for some constant of integration C_{ipv}. Moreover, the seller's expected payoff is equal to the price; that is,

$$S(k,\mu) = P(k,\mu).$$

Remark 15. An immediate consequence of Lemma 16 is that the path of prices and the seller's expected payoff are independent of the distribution of buyer types and of the buyer type that trades at any given instant. Notably, equilibrium prices only depend on the two players' common belief on the type of the market environment, which evolves exogenously. By Equilibrium Condition 6, $P(k, \mu)$ is not only the price offer in state $(K_t, \mu_t) = (k, \mu)$, but also the reservation price strategy for the buyer of type $v = k$; it follows that reservation prices of different buyer types are also independent of the distribution of buyer types and of the buyer type that trades at any given instant. Hence, in the bargaining game with independent private values, there is a strong sense in which the equilibrium structure is robust to the details of the distribution of values.

Buyer's Problem. The buyer's indifference between accepting and rejecting price offers helps to pin down the speed of smooth trade. At any state (k, μ) in the interior of a smooth-trade region, the expected payoff of the buyer of type $v = k$ (i.e., the current cutoff type) satisfies the HJB equation

$$rB^k(k, \mu) = \mu\lambda\big[O^B(k) - B^k(k, \mu)\big] - \Big[B_1^k(k, \mu)\dot{K} + B_2^k(k, \mu)\dot{\mu}\Big], \tag{3.22}$$

where $B_1^k(k, \mu)$ (resp., $B_2^k(k, \mu)$) is the partial derivative of $B^k(k, \mu)$ with respect to its first (resp., second) argument. Again, condition (3.22) has a direct interpretation. The left-hand side is the buyer's expected equilibrium payoff expressed in flow terms. The right-hand side represents the possible sources of the flow: upon arrival of the event, which happens with a subjective probability flow $\mu\lambda$, the game ends with the buyer earning $O^B(k)$ and forgoing $B^k(k, \mu)$. If the game does not end immediately, the continuation payoff changes by $-\big[B_1^k(k, \mu)\dot{K} + B_2^k(k, \mu)\dot{\mu}\big]$. In the HJB equation in (3.22) there is no term corresponding to the flow payoff from accepting the offer, $k - P(k, \mu)$, and foregoing $B^k(k, \mu)$. This is so because, by Equilibrium Condition 6, the equilibrium price at any time is equal to the payoff of the cutoff type at that time conditional on accepting the price offer: $P(k, \mu) = k - B^k(k, \mu)$. That $k - P(k, \mu) - B^k(k, \mu) = 0$ follows.

Taking the total derivative of the equilibrium condition

$$B^k(k, \mu) = k - P(k, \mu) \tag{3.23}$$

with respect to time yields

$$B_1^k(k, \mu)\dot{K} + B_2^k(k, \mu)\dot{\mu} = P_1(k, \mu)\dot{K} + P_2(k, \mu)\dot{\mu} = P_2(k, \mu)\dot{\mu}, \tag{3.24}$$

where the second equality follows from Lemma 16, which shows that smooth trade prices do not depend on the buyer type that trades at any given instant (see Remark 15), so that $P_1(k, \mu) = 0$. Replacing (3.23) and (3.24) into the HJB equation in (3.22) gives

$$P_2(k, \mu)\dot{\mu} = (\mu\lambda + r)P(k, \mu) - (\mu\lambda + r)k + \mu\lambda O^B(k). \tag{3.25}$$

Trade Dynamics. Together, the necessary equilibrium conditions (3.20) and (3.25) determine the speed of smooth trade. In particular, (3.20) and (3.25) imply that the buyer type that trades at time t_μ on the interior of a smooth trade region satisfies

$$\frac{\mu\lambda}{\mu\lambda + r} = \frac{k}{O(k)} \qquad \text{or, equivalently,} \qquad k = \frac{\mu\lambda O(k)}{\mu\lambda + r}, \tag{3.26}$$

which corresponds to the efficiency condition (3.11). That is, smooth trade is efficient.

Since trade proceeds smoothly after it begins (see Lemma 15), there are only two candidate RSE. One in which all types trade at the same instant and one in which trade is efficient. The next lemma says that the seller can profitably deviate from any price schedule that sustains trade with all buyer types in a single instant. Thus, the latter case can never be part of a RSE.

Lemma 17. *There is no RSE of the bargaining game with independent private values where all buyer types trade in a single instant.*

Moreover, a standard verification argument shows that the unique candidate RSE is indeed an equilibrium. The next proposition follows.

Proposition 13 (Equilibrium Uniqueness and Efficient Trade – IPV)**.** *The bargaining game with independent private values has a unique RSE. In the unique RSE, trade is efficient.*

With IPV, equilibrium trade dynamics inherit the properties of the efficiency benchmark. Trade is not immediate, but rather occurs over time, with the seller serving different (groups of) buyer types at different times. In particular, trade may begin with a burst at time $t = 0$ or after a silent period with no trade. There is a burst of trade at $t = 0$ if $v \geq \mu^0 \lambda O(v)/(\mu^0 \lambda + r)$ for a positive-measure subset of buyer types; trade begins silently if, instead, $v < \mu^0 \lambda O(v)/(\mu^0 \lambda + r)$ for all buyer types. After trade begins, it proceeds smoothly until the end, with the seller screening out buyer types one by one as the uncertainty about the market environment unravels.

The result in Proposition 13 yields three main takeaways. First, the result provides a novel and natural rationale for equilibrium delay, intertemporal discrimination of buyer types, and non-trivial trading dynamics in bargaining games with one-sided incomplete information. Second, the result suggests that the Coasean force toward efficient agreements remains overwhelming when parties are learning about what their best market opportunities are during their negotiations. Third, the result questions the view that long disputes result in inefficient outcomes: in markets with search and learning—examples of which are countless—this need not be true.

In equilibrium, the market may or may not clear in finite time. In particular, whether the market clears in finite time or not depends on whether the lowest buyer valuation is positive or not. If $\underline{v} > 0$, the last instant of trade is at the finite time $T_{ipv}^{\underline{v}}$ at which

$$\frac{\mu_{T_{ipv}^{\underline{v}}} \lambda}{\mu_{T_{ipv}^{\underline{v}}} \lambda + r} = \frac{\underline{v}}{O(\underline{v})}. \tag{3.27}$$

If, instead, $\underline{v} = 0$, the market does not clear in finite time, as the right-hand side of (3.26) is equal to zero when evaluated at $k = \underline{v} = 0$, whereas its left-hand side is positive at any finite time.

Remark 16 (Market Clearing and Common Knowledge of Gains from Trade – IPV)**.** In the unique RSE of the bargaining game with independent private values, the market clears in finite time if and only if gains from trade become common knowledge in finite time. When the market clears in finite time, it does so precisely at the instant at which gains from trade become common knowledge.

Price Dynamics, Market Clearing, and Market Power. For future reference, let $\Omega_{ipv}(\mu)$ be the function defined pointwise as

$$\Omega_{ipv}(\mu) := \left(\frac{\mu_{T_{ipv}^{\underline{v}}}}{\mu} \right)^{r/\lambda} \left(\frac{1 - \mu}{1 - \mu_{T_{ipv}^{\underline{v}}}} \right)^{(\lambda + r)/\lambda}. \tag{3.28}$$

When $\underline{v} > 0$, along the equilibrium path the seller knows that only the buyer of type $\underline{v}$ has still to trade at time $T_{ipv}^{\underline{v}}$. Therefore, the asymmetric information vanishes at time $T_{ipv}^{\underline{v}}$ and the seller can charge the last remaining buyer type $\underline{v}$ her willingness to pay at that time. That is, the seller exercises the option for immediate trade at time $T_S = T_{ipv}^{\underline{v}}$ (Equilibrium Condition 4) with

$$P\left(\underline{v}, \mu_{T_{ipv}^{\underline{v}}}\right) = \underline{v} - \frac{\mu_{T_{ipv}^{\underline{v}}} \lambda O^B(\underline{v})}{\lambda + r} > 0.^{29} \tag{3.29}$$

Now, (3.29) can be used as the terminal condition to determine the constant of integration in (3.21) and provide an exact expression for the equilibrium price schedule. Together, (3.21), (3.28), and (3.29) imply that equilibrium prices must satisfy

$$P(k, \mu) = \frac{\mu \lambda O^S}{\lambda + r} + \left(\underline{v} - \frac{\mu_{T_{ipv}^{\underline{v}}} \lambda O(\underline{v})}{\lambda + r}\right) \Omega_{ipv}(\mu).$$

If, instead, $\underline{v} = 0$, the seller cannot exercise the option for immediate trade at any finite time and non-negative price. Therefore, equilibrium prices must satisfy

$$P(k, \mu) = \frac{\mu \lambda O^S}{\lambda + r}.$$

The next proposition follows.

Proposition 14 (Price Dynamics and Seller's Payoff – IPV)**.** *In the unique RSE of the bargaining game with independent private values, the price offer function and the seller's payoff are:*

(i) If $\underline{v} = 0$,

$$P(k, \mu) = S(k, \mu) = \frac{\mu \lambda O^S}{\lambda + r};$$

(ii) If $\underline{v} > 0$,

$$P(k, \mu) = S(k, \mu) = \frac{\mu \lambda O^S}{\lambda + r} + \left(\underline{v} - \frac{\mu_{T_{ipv}^{\underline{v}}} \lambda O(\underline{v})}{\lambda + r}\right) \Omega_{ipv}(\mu).$$

The previous results provide several interesting insights. To begin, note that $\mu \lambda O^S/(\lambda + r)$ is both the seller's option value of waiting to learn about the existence of superior outside opportunities and the competitive price at belief μ on the type of the market environment. When $\underline{v} = 0$, prices and the seller's expected payoff equal $\mu \lambda O^S/(\lambda + r)$; in this case, the seller cannot do anything better than trading at his "marginal cost". When, instead, $\underline{v} > 0$, prices are higher than the competitive price and

[29] To see that the right-hand side of the equality in (3.29) is strictly positive, note that

$$0 = \underline{v} - \frac{\mu_{T_{ipv}^{\underline{v}}} \lambda O(\underline{v})}{\mu_{T_{ipv}^{\underline{v}}} \lambda + r} \leq \frac{\mu_{T_{ipv}^{\underline{v}}} \lambda O^B(\underline{v})}{\mu_{T_{ipv}^{\underline{v}}} \lambda + r} < \underline{v} - \frac{\mu_{T_{ipv}^{\underline{v}}} \lambda O^B(\underline{v})}{\lambda + r},$$

where: the equality holds by (3.27), the weak inequality holds because $O(\underline{v}) := O^B(\underline{v}) + O^S$ and $O^S \geq 0$, and the strict inequality holds because $\mu_{T_{ipv}^{\underline{v}}} < 1$.

the seller's expected payoff is larger than what he would get if he were awaiting for the possible arrival of the outside opportunity. The markup over the competitive price and expected payoffs is

$$\left(\underline{v} - \frac{\mu_{T^{\underline{v}}_{ipv}} \lambda O(\underline{v})}{\lambda + r}\right)\Omega_{ipv}(\mu),$$

which increases over time.

In short, the seller may or may not exercise market power. Prices are competitive when the seller cannot credibly commit to clear the market in finite time at a positive price. When, instead, the seller can clear the market in finite time at a positive price, prices are higher than under competition. When $\underline{v} > 0$, what provides the seller with a credible commitment not to lower further the price offer is the fact that, at some finite future date, it becomes common knowledge that there exist gains from trade and the information asymmetry vanishes. The next corollary follows.

Corollary 2 (Market Clearing and Market Power – IPV)**.** *In the unique RSE of the bargaining game with independent private values, the market clears in finite time if and only if market clearing prices are positive. Moreover:*

(i) If the market does not clear in finite time, prices are competitive and the seller's expected payoff equals what he would get if he were awaiting for the possible arrival of the outside opportunity.

(ii) If market clearing prices are positive, prices are higher than the competitive price and the seller's payoff is larger than what he would get if he were awaiting for the possible arrival of the outside opportunity.

Price discrimination and market power, however, do not prevent the bargaining outcome from being efficient. In fact, although trade occurs over time and the seller may gain from the ability to screen using prices, the Coasean force toward efficient agreements remains overwhelming. The efficiency result obtains because the seller screens out buyer types by conditioning his price offers only on the two parties' common belief about the type of the market environment, which evolves exogenously over time, and not directly on the current cutoff type.

Although in equilibrium different (groups of) buyer types are served at different points in time, price discrimination may or may not be present. In particular, in a buyers' market with $\underline{v} = 0$, the price offer function is identically equal to zero. In all other cases, different (groups of) buyer types are served at different prices.

Finally, whether equilibrium prices increase or decrease over time depends on which party has a higher option value of waiting to learn. In particular, prices decrease over time in a sellers' market and increase over time in a buyers' market (except when $\underline{v} = 0$, in which case prices are identically equal to zero).

3.4.2 Bargaining with Interdependent Values

Seller's Problem. Consider any state (k, μ) in the interior of a smooth-trade region. Since the probability of leaving the interior of such a region in the next dt is negligible, the seller's expected payoff in state (k, μ), $S(k, \mu)$, satisfies the HJB equation

$$rS(k,\mu) = \sup_{\dot{K}\in(-\infty,0]} \Bigg\{ \mu\lambda\big[\underline{O}^S(k) - S(k,\mu)\big] + \big[P(k,\mu) - S(k,\mu)\big]\frac{f(k)}{F(k)}\big(-\dot{K}\big) + S_1(k,\mu)\dot{K} + S_2(k,\mu)\dot{\mu}\Bigg\}. \tag{3.30}$$

The interpretation of (3.30) is analogous to that of equation (3.14), but with an important difference. Now, the sellers' expected payoff upon arrival of the event, $\underline{O}^S(k)$, is not constant but depends on the current cutoff type. This is the source of interdependency between the two players' payoffs.

Again, as the right-hand side of (3.30) is linear in $\dot{K}$, the sum of the coefficients on $\dot{K}$ must be non-negative on the interior of a smooth-trade region. The seller finds it optimal to set $\dot{K} = 0$ (i.e., silent trade) if the coefficients on $\dot{K}$ add up to something positive. The next lemma parallels Lemma 15 and shows that $\dot{K} = 0$ cannot occur after a positive measure of buyer types has traded.

Lemma 18. *In any RSE of the bargaining game with interdependent values, if $K_t = k < \overline{v}$, then $\dot{K}_s < 0$ for all $s \geq t$.*

Suppose that $\dot{K} < 0$. In this case, the coefficients on $\dot{K}$ must add up to zero, which means that the seller must be indifferent between speeds of trade. Setting the coefficients on $\dot{K}$ to zero in (3.30) yields the PDEs

$$S_1(k,\mu) = \big[P(k,\mu) - S(k,\mu)\big]\frac{f(k)}{F(k)}, \tag{3.31}$$

$$S_2(k,\mu)\dot{\mu} = (\mu\lambda + r)S(k,\mu) - \mu\lambda\underline{O}^S(k), \tag{3.32}$$

which describe the seller's best response problem on the interior of a smooth-trade region where $\dot{K} < 0$ of any candidate equilibrium. The PDE in (3.32) has general solution

$$S(k,\mu) = \frac{\mu\lambda\underline{O}^S(k)}{\lambda + r} + C^S_{iv}\frac{(1-\mu)^{(\lambda+r)/\lambda}}{\mu^{r/\lambda}}, \tag{3.33}$$

where C^S_{iv} is the constant of integration. Moreover, note that (3.31) is equivalent to

$$P(k,\mu) = \frac{\frac{\partial}{\partial k}[S(k,\mu)F(k)]}{f(k)}. \tag{3.34}$$

Together, (3.33) and (3.34) imply that equilibrium prices must satisfy

$$P(k,\mu) = \frac{\mu\lambda O^S(k)}{\lambda + r} + C^P_{iv}\frac{(1-\mu)^{(\lambda+r)/\lambda}}{\mu^{r/\lambda}},$$

for some constant of integration C_{iv}^{P}. The next lemma summarizes the previous discussion.

Lemma 19. *In any RSE of the bargaining game with independent private values, smooth-trade prices are determined by the seller's indifference between speeds of trade. In particular,*

$$P(k,\mu) = \frac{\mu\lambda O^S(k)}{\lambda + r} + C_{iv}^{P}\frac{(1-\mu)^{(\lambda+r)/\lambda}}{\mu^{r/\lambda}}, \tag{3.35}$$

for some constant of integration C_{iv}^{P}. Moreover, the seller's expected payoff is

$$S(k,\mu) = \frac{\mu\lambda \underline{O}^S(k)}{\lambda + r} + C_{iv}^{S}\frac{(1-\mu)^{(\lambda+r)/\lambda}}{\mu^{r/\lambda}}, \tag{3.36}$$

for some constant of integration C_{iv}^{S}.

As an immediate consequence of Lemma 19, the price offer function must satisfy the PDE

$$P_2(k,\mu)\dot{\mu} = (\mu\lambda + r)P(k,\mu) - \mu\lambda O^S(k). \tag{3.37}$$

Remark 17. By Lemma 19, the seller's expected payoff always depends on the distribution of buyer types; so does the path of prices, unless $O^S(v)$ is independent of the distribution of values. The result contrasts with the findings for the bargaining game with independent private values, where the path of prices and the seller's expected payoff neither depend on the distribution of buyer types nor on the buyer type that trades at any given instant. (cf. Remark 15)

Buyer's Problem. Again, the buyer's indifference between accepting and rejecting price offers helps to pin down the speed of smooth trade. Note that the buyer's problem is the same as that for the case of independent private values. Thus, at any state (k,μ) in the interior of a smooth-trade region, the expected payoff of the buyer of type $v = k$ satisfies the HJB equation in (3.22).

Taking the total derivative of the equilibrium condition (3.23) with respect to time yields

$$B_1^k(k,\mu)\dot{K} + B_2^k(k,\mu)\dot{\mu} = P_1(k,\mu)\dot{K} + P_2(k,\mu)\dot{\mu}. \tag{3.38}$$

Now, as prices depend on the buyer type that trades at any given instant, the term $P_1(k,\mu)\dot{K}$ does not disappear from the right-hand side of (3.38) (cf. 3.24). Replacing (3.23) and (3.38) into the HJB equation in (3.22) gives

$$P_2(k,\mu)\dot{\mu} = (\mu\lambda + r)P(k,\mu) - (\mu\lambda + r)k + \mu\lambda O^B(k). \tag{3.39}$$

Trade Dynamics. Together, the necessary equilibrium conditions (3.37) and (3.39) determine the speed of smooth trade. In particular, (3.37) and (3.39) imply that the buyer type that trades at time t_μ on the interior of a smooth trade region satisfies

$$k = \frac{\mu\lambda O(k)}{\mu\lambda + r} - \frac{P_1(k,\mu)\dot{K}}{\mu\lambda + r}. \tag{3.40}$$

Note that $\dot{K} < 0$ and $P_1(k,\mu) = \mu\lambda O^{S'}(k)/(\lambda+r) > 0$ (by Assumption 3–A3). Thus, by comparing (3.40) to the efficiency condition (3.11), it follows immediately that smooth trade is inefficiently slow when valuations are interdependent. That is, in the bargaining game with interdependent values, there is inefficient delay.

Since trade proceeds smoothly after it begins (see Lemma 18), there are only two candidate RSE. One in which all types trade at the same instant and one in which trade is efficient. The next lemma, paralleling Lemma 17, says that the seller can profitably deviate from any price schedule that sustains trade with all buyer types in a single instant. Thus, the latter case can never be part of a RSE.

Lemma 20. *There is no RSE of the bargaining game with interdependent values where all buyer types trade in a single instant.*

Moreover, a standard verification argument shows that the unique candidate RSE is indeed an equilibrium. The next proposition follows.

Proposition 15 (Equilibrium Uniqueness and Inefficient Trade – IV)**.** *The bargaining game with interdependent values has a unique RSE. In the unique RSE, trade is inefficiently slow.*

Similarly to the case with independent private values, trade may begin with a burst at time $t = 0$ or after a silent period with no trade; after trade begins, it proceeds smoothly until the end, with the seller screening out buyer types one by one as the uncertainty about the market environment unravels. The market clears in finite time if and only if the lowest buyer valuation is positive, that is, if and only if gains from trade become common knowledge in finite time. However, when the market clears in finite time, it does so *after* the instant at which gains from trade become common knowledge.

Price Dynamics, Market Clearing, and Market Power. The remaining part of the analysis closely mimics that for the bargaining game with independent private values. I refer to Section 3.4.1 for the details that I omit.

For future reference, let $\Omega_{iv}(\mu)$ be the function defined pointwise as

$$\Omega_{iv}(\mu) := \left(\frac{\mu_{T_{iv}^{\underline{v}}}}{\mu}\right)^{r/\lambda}\left(\frac{1-\mu}{1-\mu_{T_{iv}^{\underline{v}}}}\right)^{(\lambda+r)/\lambda}. \tag{3.41}$$

When $\underline{v} > 0$, along the equilibrium path the seller knows that only the buyer of type $\underline{v}$ has still to trade at time $T_{iv}^{\underline{v}}$. Therefore, the asymmetric information vanishes at time $T_{iv}^{\underline{v}}$ and the seller can charge the last remaining buyer type $\underline{v}$ her willingness to pay at that time. That is, the seller exercises the option for immediate trade at time $T_S = T_{iv}^{\underline{v}}$ (Equilibrium Condition 4) with

$$S\left(\underline{v}, \mu_{T_{iv}^{\underline{v}}}\right) = P\left(\underline{v}, \mu_{T_{iv}^{\underline{v}}}\right) = \underline{v} - \frac{\mu_{T_{iv}^{\underline{v}}}\lambda O^B(\underline{v})}{\lambda + r}. \tag{3.42}$$

Now, (3.42) can be used as the terminal condition to determine the constants of integration in (3.35) and (3.36) and provide an exact expression for the equilibrium price schedule and the seller's expected payoffs. If, instead, $\underline{v} = 0$, the seller cannot exercise the option for immediate trade at any finite time and non-negative price. The next proposition follows.

Proposition 16 (Price Dynamics and Seller's Payoff – IV)**.** *In the unique RSE of the bargaining game with interdependent values, the price offer function and the seller's payoff are:*

(i) If $\underline{v} = 0$,

$$P(k,\mu) = \frac{\mu\lambda O^S(k)}{\lambda + r} \qquad \text{and} \qquad S(k,\mu) = \frac{\mu\lambda \underline{O}^S(k)}{\lambda + r};$$

(ii) If $\underline{v} > 0$,

$$P(k,\mu) = \frac{\mu\lambda O^S(k)}{\lambda + r} + \left(\underline{v} - \frac{\mu_{T_{iv}^{\underline{v}}}\lambda O(\underline{v})}{\lambda + r}\right)\Omega_{iv}(\mu)$$

and

$$S(k,\mu) = \frac{\mu\lambda \underline{O}^S(k)}{\lambda + r} + \left(\underline{v} - \frac{\mu_{T_{iv}^{\underline{v}}}\lambda O(\underline{v})}{\lambda + r}\right)\Omega_{iv}(\mu).$$

When $\underline{v} = 0$, $S(k,\mu)$ has the property that at any point in the game the expected payoff of the seller is equal to his payoff from waiting for the possible arrival of the outside opportunity. With interdependent values, although the Coase conjecture does not hold any longer in terms of efficient trade, the Coasean dynamics force down the seller's profit to his outside option; that is, the seller cannot do anything better than trading at his "marginal cost". Moreover, for each state (k,μ), $P(k,\mu)$ is exactly the expected present value the seller would have earned from type k if he waited for the possible arrival of the outside opportunity—a kind of no-ex post regret property—and upon the price being accepted the seller does not regret not slowing down the trade.[30] When, instead, $\underline{v} > 0$, prices are higher than the competitive price and the seller's expected payoff is larger than what he would get if he were awaiting for the possible arrival of the outside opportunity. The markup over the competitive price and expected payoffs is

$$\left(\underline{v} - \frac{\mu_{T_{iv}^{\underline{v}}}\lambda O(\underline{v})}{\lambda + r}\right)\Omega_{iv}(\mu),$$

which increases over time.

Again, the seller may or may not exercise market power. Prices are competitive when the seller cannot credibly commit to clear the market in finite time at a positive price. When, instead, the seller can clear the market in finite time at a positive price, prices are higher than under competition. The next corollary follows.

Corollary 3 (Market Clearing and Market Power – IV)**.** *In the unique RSE of the bargaining game with interdependent values, the market clears in finite time if and only if market clearing prices are positive. Moreover:*

(i) If the market does not clear in finite time, prices are competitive and the seller's expected payoff equals what he would get if he were awaiting for the possible arrival of the outside opportunity.

(ii) If market clearing prices are positive, prices are higher than the competitive price and the seller's payoff is larger than what he would get if he were awaiting for the possible arrival of the outside opportunity.

[30] A similar property arises in Fuchs and Skrzypacz (2010).

3.5 Extensions and Discussion

3.5.1 The Case of Complete Information

Suppose that the buyer's valuation $v \in [\underline{v}, \overline{v}]$ for the good is common knowledge between the seller and the buyer. If the event does not occur before trade takes place, then:

(a) Price discrimination is perfect and the seller, who has all the bargaining power, extracts all the surplus from the transaction. As a result, the transaction occurs at time $t(v)$, the efficient trading time with the buyer of type v (see Proposition 12–*(ii)*). The transaction takes place at price p equal to the buyer's willingness to pay at time $t(v)$. This is given by the difference between the buyer's valuation v and the present discounted value of her outside opportunity at time $t(v)$, which is $\mu_{t(v)}\lambda O^B(v)/(\lambda + r)$. That is,

$$p = v - \frac{\mu_{t(v)}\lambda O^B(v)}{\lambda + r}.$$

(b) The seller implements the complete information outcome by setting any price schedule which is greater than the buyer's willingness to pay, $v - \mu_t\lambda O^B(v)/(\lambda + r)$, at all times $t \neq t(v)$, and equal to $v - \mu_{t(v)}\lambda O^B(v)/(\lambda + r)$ at time $t = t(v)$. In such case, the buyer of type v is willing to pay price $p_{t(v)}$ at time $t(v)$, as this leaves him indifferent between trading and waiting for the arrival of the outside opportunity, but refuses to trade at any other time.

The next proposition characterizes the complete information outcome. Parts *(i)* and *(ii)* follow from the previous discussion. Part *(iii)* holds by straightforward calculations, which are in Appendix 3.7.3.

Proposition 17 (Complete Information Outcome)**.** *Suppose the buyer type $v \in [\underline{v}, \overline{v}]$ is common knowledge, and that the event does not occur before trade takes place. Then:*

(i) Trade is efficient.

(ii) The seller extracts all the surplus from the transaction by implementing any price schedule $p\colon \mathbb{R}_+ \to \mathbb{R}_+$, $t \mapsto p_t$, such that

$$p_t \begin{cases} = v - \frac{\mu_t\lambda O^B(v)}{\lambda+r} & \textit{if } t = t(v) \\ > v - \frac{\mu_t\lambda O^B(v)}{\lambda+r} & \textit{if } t \neq t(v) \end{cases},$$

where $t(v)$ is the efficient trading time for the buyer of type v.

(iii) The seller's and the buyer's payoffs, denoted by $S_C(v)$ and $B_C(v)$, are:

$$S_C(v) = \begin{cases} v - \frac{\mu^0\lambda O^B(v)}{\lambda+r} & \textit{if } v \geq \frac{\mu^0\lambda O(v)}{\mu^0\lambda+r} \\ v - \frac{r\lambda v O^B(v)}{\lambda(\lambda+r)(O(v)-v)} & \textit{if } v < \frac{\mu^0\lambda O(v)}{\mu^0\lambda+r} \end{cases},$$

and

$$B_C(v) = \begin{cases} \frac{\mu^0 \lambda O^B(v)}{\lambda + r} & \text{if } v \geq \frac{\mu^0 \lambda O(v)}{\mu^0 \lambda + r} \\ \frac{r \lambda v O^B(v)}{\lambda(\lambda+r)(O(v)-v)} & \text{if } v < \frac{\mu^0 \lambda O(v)}{\mu^0 \lambda + r} \end{cases}.$$

The next corollary, which follows from straightforward calculations collected in Appendix 3.7.3, contains comparative statics for the complete information outcome. The results have a natural interpretation.

Corollary 4 (Comparative Statics)**.** *Fix a buyer type $v \in [\underline{v}, \overline{v}]$. In the complete information outcome:*

(i) If $v \geq \frac{\mu^0 \lambda O(v)}{\mu^0 \lambda + r}$, then

(a) $S_C(v)$ is decreasing in $O^B(v)$ and independent of $O^S(v)$;

(b) $B_C(v)$ is increasing in $O^B(v)$ and independent of $O^S(v)$.

(ii) If $v < \frac{\mu^0 \lambda O(v)}{\mu^0 \lambda + r}$, then

(a) $S_C(v)$ is increasing in $O^S(v)$, decreasing in $O^B(v)$ when $O^S(v) > v$, increasing in $O^B(v)$ when $O^S(v) < v$, and independent of $O^B(v)$ when $O^S(v) = v$;

(b) $B_C(v)$ is decreasing in $O^S(v)$, increasing in $O^B(v)$ when $O^S(v) > v$, decreasing in $O^B(v)$ when $O^S(v) < v$, and independent of $O^B(v)$ when $O^S(v) = v$.

3.5.2 The Bargaining Game without Learning

Fuchs and Skrzypacz (2010) study a one-sided incomplete information bargaining game where a new trader arrives according to a Poisson process. There are two main differences between their setup and mine. First, in my setting there is uncertainty about the type of the market; in contrast, in their setup outside opportunities arrive stochastically, but are known to exist (i.e., it is common knowledge that the market is of type 1). Thus, learning about the market environment and learning about the buyer's private information do not interact in their model. Second, in their setup arrivals do not correspond to superior, but rather to alternative, trading opportunities. Therefore, the efficient benchmark calls for immediate agreement, and not for an optimal level of market experimentation.[31]

These contrasts lead to distinct insights and equilibrium dynamics. Fuchs and Skrzypacz (2010) show that trade occurs over time only if valuations are interdependent. With independent private values, instead, trade occurs either immediately or never (see also Inderst (2008)). In particular, Fuchs and Skrzypacz (2010) argue as follows: "Arrival of new traders or outside options is necessary for delay, but another important ingredient for slow equilibrium screening is that the seller's outside value depends on the buyer's type." In contrast, when learning about the market environment interacts with learning about parties' private information, I show that the seller serves different (groups) of buyer types at different points in time and charges them different prices even when private valuations are

[31] In addition, they write the model in discrete time and then study the atomless continuous-time limit of stationary perfect Bayesian equilibria. In contrast, I pose the game directly in continuous time and develop a suitable framework for equilibrium analysis.

independent. A common insight of our models, however, is that interdependent values are necessary for inefficiently timed transactions. Fuchs and Skrzypacz (2010) also propose the following generalization of the Coase Conjecture: although there is inefficient delay and the price does not drop immediately to zero, the Coasean dynamics force down the seller's profit to his outside opportunity. In my setting, this insight only holds if the seller cannot clear the market in finite time at a positive price. In contrast, when the seller has the option to do so, I show that he prices above his marginal cost, may exercise substantial market power, and his payoff is larger from that he would obtain by simply awaiting for the possible arrival of an outside opportunity. Moreover, the result holds independently of whether private valuations are interdependent or not.

3.5.3 Optimal Sales Mechanisms under Commitment

In section 3.3.1, I characterize the first-best efficient benchmark without discussing its implementation. When valuations are independent, my analysis shows that "freestyle" bargaining implements the efficient outcome. This is no longer true with interdependent values. In this case, it is natural to inquire whether an efficient mechanism exists and, if so, whether it can be implemented in prices. If the answer to either question is negative, then the bargaining outcome is necessarily inefficient.

In some work in progress, I adopt a mechanism design approach to study the same trading environment as the one in this paper. In particular, I consider the design of profit maximizing mechanisms when the seller has full commitment power and the design of efficient trading mechanisms. A preliminary investigation suggests that the bargaining outcome is not second-best efficient when valuations are interdependent.[32]

3.5.4 Burst of Trade after Trade Begins

In my model, trade may begin with a burst, but proceeds smoothly afterward. What drives slow screening of buyer types in equilibrium is the assumption that $v/O(v)$ is increasing. If $v/O(v)$ were to be constant over some interval subset of $[\underline{v}, \overline{v}]$, bursts of trade might occur after trade begins. As long as $v/O(v)$ is constant over finitely many interval subset of $[\underline{v}, \overline{v}]$, regular stationary equilibria allow to capture bargaining dynamics where bursts of trade alternate with periods of smooth trade. This alternative specification of the benchmark model may be appropriate in some applications.

3.5.5 Different Learning Processes

I assume that learning occurs via conclusive news. Whereas this is a natural modeling choice in my setting, one may consider more gradual learning processes. For instance, news may arrive via a non-conclusive Poisson process (see Keller and Rady (2010)). I can nest this case in my model by assuming that the event corresponds to the first jump of the Poisson process and using $O^S(v)$ and $O^B(v)$ to capture the two parties' payoffs in the continuation game that follows the jump. Alternatively, one may assume that news about the market environment are revealed via a Brownian diffusion process

[32] See Deneckere and Liang (2006) for a result in a similar spirit. However, while theirs is a static mechanism design problem, mine is dynamic.

(see Bolton and Harris (1999a)). This is an interesting theoretical extension of the model, which I plan to investigate in future work. I expect the main insights of the analysis to remain valid in such setup.

3.6 Related Literature and Concluding Remarks

3.6.1 Related Literature

This paper joins a recent literature that explicitly models changing (stochastic) features of the bargaining environment. Huang and Li (2013), Ortner (2017), Daley and Green (2018), and Ishii, Öry, and Vigier (2018) are the closest contributions to mine.[33] Huang and Li (2013) analyze a discrete-time bargaining game where a seller makes all price offers to a privately informed buyer. A new buyer with a higher valuation for the seller's good may arrive in the future. As time elapses with no arrival, the two players revise downward their common belief about the existence of such a buyer. Their main result shows that prices fluctuate in equilibrium. This is so because the seller posts a price at the very beginning of each time period and, in discrete-time, he has to commit to that price for the whole period. I can specialize my model to capture their setup by assuming independent private values, that there is a gap between the seller and the lowest buyer's valuation, and that outside opportunities, if existing, are only available to the seller. In this case, however, I show that price fluctuations disappear in continuous time. The price, instead, gently declines over time as the seller becomes more pessimistic about his outside opportunities. An interpretation of the difference in our findings is that price fluctuations are not driven exclusively by the option value of waiting to learn; rather, they result from the combined effect of learning with that of other frictions in the protocol.

Ortner (2017) studies the problem of a durable-good monopolist who lacks commitment power and whose marginal cost of production varies stochastically over time. He suggests a generalization of the Coase Conjecture according to which the monopolist seller earns the same profit as he would earn if he were selling to a market in which all consumers had the lowest valuation. I show that the seller's profit is larger than this lower bound when he has the option to clear the market at a positive price in finite time. Ortner (2017) also shows that the seller exercises market power if the distribution of buyer valuations is discrete but is unable to do so when there is a continuum of types. In contrasts, my findings on market power do not rely on the distribution of buyer valuations being discrete, but rather on a market clearing condition.

Daley and Green (2018) propose a one-sided incomplete information bargaining model with news. Their setup differs from mine in two relevant ways. First, in their setting news are about the informed party's private information, and not about the existence of outside opportunities. Second, the social value of waiting for news is nil, and so efficiency calls for an immediate agreement.[34] They show that the uninformed party's ability to leverage public information to extract more surplus from the transaction

[33] More broadly, this paper adds to the work studying the role of outside options or the arrival of new traders in bargaining games with asymmetric information (e.g., Fudenberg, Levine, and Tirole (1987), Samuelson (1992), Inderst (2008), and, more recently, Chang (2015) and Hwang (2018)).

[34] From a more technical viewpoint, in Daley and Green (2018) the uninformed party learns about the informed party's private information by observing a public Brownian news process, whereas in my setting the two players learn about the existence of superior outside opportunities by publicly observing a conclusive Poisson process.

is remarkably limited. They suggest a novel interpretation of the Coase Conjecture: because of his perfect lack of commitment, the uninformed party derives no benefit from the ability to screen using prices. The equilibrium may involve delay of trade and positive profits, depending on the environment, but the uninformed party's payoffs must be exactly what it would receive if it were unable to make offers at all. In contrast, I show the seller's payoff exceeds what he would get if he were unable to screen using prices when he has the option to clear the market in finite time at a positive price.

Ishii et al. (2018) study wage bargaining between a worker and two firms, with public learning about worker-specific productivity. Firms make take-it-or-leave-it offers over time, and hiring is irreversible. Search frictions delay the arrival of one firm, the entrant, while informational frictions prevent the incumbent from always observing the entrant's arrival. They show that the combined effect of search and informational frictions induces unraveling in all equilibria: parties reach inefficiently early agreements and the average talent of hired workers is lower than socially optimal. They also show that without market frictions, or when the search friction is present whereas the informational friction is not, there is no unraveling in equilibrium. In my model, one can interpret the arrival of the outside opportunity as a breakthrough in some underlying (on-the-market) search activity that parties engage in in parallel to their negotiations. With this interpretation, my results say that search and learning do not give rise to inefficient bargaining outcomes, unless they are paired with an additional friction in the environment, which in my model takes the form of interdependent values. At a high level, this insight is evocative of the one of Ishii et al. (2018), who show that learning and the search friction do not give rise to inefficiencies, unless they are paired with the informational friction.

My work relates to the recent contribution by Nava and Schiraldi (2018). They analyze the problem of a durable-good monopolist who sells multiple varieties of a good without the ability to commit to future prices. The authors show that, in such setting, the seller regains some ability to command positive profits. They propose a robust interpretation of the Coase Conjecture by arguing that the force driving any Coasean equilibrium is market-clearing, and not efficiency or minimal pricing (that is, pricing equal to the maximum between marginal cost and the minimal value). This is so because any market-clearing price (that is, any price at which all consumers are willing to buy) provides a credible commitment to the monopolist (as it is no longer compelled to lower prices). Although our settings are very different in nature and Nava and Schiraldi (2018) do not model any dynamic feature of the environment, their result is reminiscent of my finding that a monopolist seller exercises some form of market power any time he has the option to clear the market at a positive price in finite time.

From a methodological viewpoint, I add to the recent and growing work on bargaining games in continuous time. In particular, I build on Ortner (2017) and Daley and Green (2018) to develop an ad hoc equilibrium notion for the game I study by introducing strategy restrictions directly into the equilibrium definition.[35] Although not being fully Nash, the equilibrium concept captures the key features of any perfect Bayes (stationary) analysis of the discrete-time analog of the model.[36]

[35] Chaves Villamizar (2018) adopts a similar approach.

[36] Other notable contributions on bargaining games in continuous time, although less immediately connected to my work, are Perry and Reny (1993), Sákovics (1993), Ambrus and Lu (2015), and Ortner (2016). More broadly, my paper relates to literature that uses continuous-time techniques to study strategic interactions.

More broadly, my work relates to the literature on equilibrium delay in bargaining. For example, delay occurs in a model with two-sided private information about fundamentals and overlap in values (e.g., Cramton (1984), Chatterjee and Samuelson (1987), and Cho (1990)), with reputational concerns (e.g., Abreu and Gul (2000), Compte and Jehiel (2002), and Atakan and Ekmekci (2014)), with higher order uncertainty (Feinberg and Skrzypacz (2005)), with disagreement about continuation play (Yildiz (2004)), with externalities (Jehiel and Moldovanu (1995)), with the possibility that players can commit to not responding to offers (Admati and Perry (1987) and Freshtman and Seidmann (1993)), or when outside options are history-dependent (Compte and Jehiel (2004)). Efficient delay may emerge in bargaining games with complete information where the size of the cake and the identity of the proposer evolves stochastically over time (Merlo and Wilson (1995) and Merlo and Wilson (1998)). (Inefficient) delay when players may receive new information while bargaining also arises in the complete information setting of (Avery and Zemsky (1994)). In my environment, there is incomplete information and the identity of the proposer is fixed; in a similar spirit, however, parties do not trade as long as the option value of waiting for news is, in expectation, sufficiently larger that the surplus from the transaction.

My paper also relates to the literature that checks the robustness of Coase's insight or identifies different ways in which a dynamic monopolist can exercise market power. For instance, a monopolist could relax its commitment problem and increase its profit by renting the good rather than selling it (Bulow (1982)), by introducing best-price provisions (Butz (1990)), or by introducing new updated versions of the durable good over time (Levinthal and Purohit (1989), Waldman (1993, 1999), Choi (1994), Fudenberg and Tirole (1998), and Lee and Lee (1998)). Other studies have analyzed environments which preclude the market from fully deteriorating. These include environments with capacity constraints (Kahn (1986) and McAfee and Wiseman (2008)), with entry of new buyers (Sobel (1991)), where buyers' valuations are subject to idiosyncratic stochastic shocks (Biehl (2001), Deb (2014), and Garrett (2016)), where buyers can exercise an outside option (Board and Pycia (2014)), where goods depreciate over time (Bond and Samuelson (1987)), and with demand is discrete (Bagnoli, Salant, and Swierzbinski (1989), von der Fehr and Kuhn (1995), and Montez (2013)).

Finally, I model learning and market experimentation building on the exponential bandit framework pioneered by Keller, Rady, and Cripps (2005a).

3.6.2 Concluding Remarks

Parties to a negotiation often have reasons to inquire whether the current counterparty offers the best available trading opportunity or, in contrast, a more satisfactory use of their resources exists in the market. Superior outside opportunities not only may take time to arrive; they are often of uncertain existence. Uncertainty, however, unravels over time if parties engage in some form of market experimentation or search. These features are common in many markets—durable goods, labor, housing, and financial markets, just to name a few—and the trade-offs they give rise to are arguably a defining feature of many bargaining relationships. In this paper, I develop a framework to understand bilateral bargaining relationships with one-sided incomplete information when gains from trade are ex ante uncertain and parties may learn whether superior outside opportunities are available during their negotiations. I show that the resulting tension between an immediate agreement and the option value of waiting to learn is of first-order importance in shaping the bargaining relationship. It affects the timing of agreements, the dynamics of prices, realized surplus and its division, and the seller's ability to exercise market power.

Trade no longer takes place immediately with the informed party capturing all the rents. In contrast, the seller screens out buyers over time by charging different prices to different types. While learning accounts for delay, inefficiently timed agreements only occur if valuations are interdependent. Although other explanations have been proposed for the observed delay in bargaining, mine is a very natural one. It shows that delay is to be expected in markets with search and learning. My results, however, question the view that long disputes result in inefficient bargaining outcomes: absent additional frictions in the protocol (e.g., interdependent values), the Coasean force toward efficiency remains overwhelming when parties engage in market experimentation during their negotiations. I also show that the seller may exercise market power. In particular, market power is present when the seller is able to clear the market in finite time at positive prices. In this cases, prices are higher than the competitive price and the seller's payoff is larger than what he would get if he were awaiting for the possible arrival of a superior outside opportunity or unable to screen using prices (or both).

My model is flexible enough to serve as a stepping stone for future research. I discuss several extensions in Section 3.5. I plan to build on this setup to explore further bargaining and trade relationships when the economic environment is non-stationary because of learning.

3.7 Remaining Proofs

3.7.1 Proofs for Section 3.4.1

Proof of Lemma 17

Suppose $\underline{v} = 0$. At time $t = t_\mu$ the willingness to pay of the buyer of type v is

$$v - \frac{\mu\lambda O^B(v)}{\lambda + r}. \tag{3.43}$$

As (3.43) is negative for a positive-measure subset of buyer types at any $t_\mu \in \mathbb{R}_+$, trade with all buyer types at a single instant can only occur at a negative price. Since the seller can always secure himself a payoff of zero (or larger, if $O^S > 0$) by not trading, this can never be an equilibrium.

Now suppose $\underline{v} > 0$. Let $\underline{T}$ be the time satisfying

$$\underline{v} - \frac{\mu_{\underline{T}}\lambda O^B\left(\underline{v}\right)}{\lambda + r} = 0.$$

Trade with all buyer types at any single instant $t < T_{\underline{v}}$ requires a negative price. Again, as the seller can always secure himself a payoff of zero (or larger, if $O^S > 0$) by not trading, this can never be an equilibrium. It remains to rule out instantaneous trade with all buyer types at any $t \geq T_{\underline{v}}$. ■

3.7.2 Proofs for Section 3.4.2

Proof of Lemma 17

Suppose $\underline{v} = 0$. At time $t = t_\mu$ the willingness to pay of the buyer of type v is

$$v - \frac{\mu\lambda O^B(v)}{\lambda + r}. \tag{3.44}$$

As (3.44) is negative for a positive-measure subset of buyer types at any $t_\mu \in \mathbb{R}_+$, trade with all buyer types at a single instant can only occur at a negative price. Since the seller can always secure himself a payoff of zero (or larger, if $O^S > 0$) by not trading, this can never be an equilibrium.

Now suppose $\underline{v} > 0$. Let $\underline{T}$ be the time satisfying

$$\underline{v} - \frac{\mu_{\underline{T}}\lambda O^B\left(\underline{v}\right)}{\lambda + r} = 0.$$

Trade with all buyer types at any single instant $t < T_{\underline{v}}$ requires a negative price. Again, as the seller can always secure himself a non-negative payoff by not trading, this can never be an equilibrium. It remains to rule out instantaneous trade with all buyer types at any $t \geq T_{\underline{v}}$. ■

3.7.3 Proofs for Section 3.5.1

Proof of Proposition 17

If $v \geq \mu^0 \lambda O(v)/(\mu^0\lambda + r)$, trade occurs at time $t = 0$ (equivalently, belief μ^0). The seller's payoff is

$$S_C(v) = p_0 = v - \frac{\mu_0 \lambda O^B(v)}{\lambda + r},$$

and the buyer's payoff is

$$B_C(v) = v - p_0 = \frac{\mu_0 \lambda O^B(v)}{\lambda + r}.$$

If $v < \mu^0 \lambda O(v)/(\mu^0\lambda + r)$, trade occurs at belief

$$\mu(v) = \frac{rv}{\lambda(O(v) - v)}.$$

Let $t_{\mu(v)}$ be the time at which $\mu_t = \mu(v)$. Thus, the seller's payoff is

$$S_C(v) = p_{t_{\mu(v)}} = v - \frac{\mu(v)\lambda O^B(v)}{\lambda + r} = v - \frac{rv}{\lambda(O(v) - v)}\frac{\lambda O^B(v)}{\lambda + r} = v - \frac{r\lambda v O^B(v)}{\lambda(\lambda + r)(O(v) - v)},$$

and the buyer's payoff is

$$B_C(v) = v - p_{t_{\mu(v)}} = \frac{r\lambda v O^B(v)}{\lambda(\lambda + r)(O(v) - v)}. \qquad \blacksquare$$

Proof of Corollary 4

Part *(i)* is immediate. The only non-obvious statement in part *(ii)* is the dependence of $S_C(v)$ and $B_C(v)$ on $O^B(v)$. This is so because $O^B(v)$ appears in both the nominator and the denominator of $S_C(v)$ and $B_C(v)$. Suppose that $v < \mu^0 \lambda O(v)/(\mu^0\lambda + r)$, so that also $v < O(v)$. Then:

$$\frac{\partial S_C(v)}{\partial O^B(v)} = \frac{\partial}{\partial O^B(v)}\left[v - \frac{r\lambda v O^B(v)}{\lambda(\lambda + r)(O(v) - v)}\right] = -\frac{r\lambda v\big(O^S(v) - v\big)}{O(v) - v},$$

and

$$\frac{\partial B_C(v)}{\partial O^B(v)} = \frac{\partial}{\partial O^B(v)}\left[\frac{r\lambda v O^B(v)}{\lambda(\lambda + r)(O(v) - v)}\right] = \frac{r\lambda v\big(O^S(v) - v\big)}{O(v) - v}.$$

By observing that

$$\frac{\partial S_C(v)}{\partial O^B(v)} > 0 \Longleftrightarrow O^S(v) < v \qquad \text{and} \qquad \frac{\partial S_C(v)}{\partial O^B(v)} = 0 \Longleftrightarrow O^S(v) = v,$$

and that

$$\frac{\partial B_C(v)}{\partial O^B(v)} > 0 \Longleftrightarrow O^S(v) < v \qquad \text{and} \qquad \frac{\partial B_C(v)}{\partial O^B(v)} = 0 \Longleftrightarrow O^S(v) = v,$$

the desired result follows. ■

Chapter 4

Dynamic Foundations for Empirical Static Games[1]

4.1 Introduction

When data are generated by strategic interaction, the analyst needs to specify a solution concept to interpret players' outcomes and leverage the observables to perform identification of the game's primitives. It is common practice to assume that equilibrium play is observed. This assumption is justified whenever players are able to form correct expectations[2] on the strategic environment, and to (optimally) behave accordingly. However, for many real-world strategic environments, it is not obvious that behavior satisfies these requirements: players may need to learn how to play. An external analyst needs then to determine when the learning phase terminates, and whether it is possible to interpret subsequent behavior as the result of equilibrium play.

In this paper, we explicitly allow for the possibility that players are learning how to interact. However, we remain agnostic on the details of the learning process. Moreover, instead of assuming that players reach the ability to choose correct strategies at some given point in time, we only impose a minimal behavioral assumption describing an optimality condition for the long-term result of players' interaction. More specifically, we model players as interacting repeatedly, playing an incomplete information game, and assume that long-run outcomes satisfy a property of "asymptotic no regret" (hereafter, ANR). The ANR property requires that the time average of the counterfactual increase in past payoffs, had different actions been played, becomes approximately equal to zero in the long run. Intuitively, no matter which specific learning rule players are adopting, we assume that they eventually eliminate the regret of not having played differently in the past. After having imposed that players' behavior satisfies the ANR property, we derive the implications for identification of the game's primitives.

We show that, under the ANR assumption, it is possible to partially identify the structural parameters of players' payoff functions. We do so in two steps. First, we show that the time average of play that

[1]The content of this chapter is joint work with Lorenzo Magnolfi and Camilla Roncoroni. We thank Dirk Bergemann for valuable conversations. All errors and omissions are our own.

[2]Given their information sets, which are assumed to be correctly defined by the analyst.

satisfies ANR converges to the set of Bayes correlated equilibria of the underlying static game (although, in general, it does not converge to a particular point in this set).[3] To establish this property we extend to games of incomplete information prior results on dynamic foundations for equilibrium play in static games of complete information. Second, we show how to use the limiting model to obtain consistent estimates of the parameters of interest. Our empirical approach is based on a behavioral assumption that has only implications on the predictions of a limiting model, as opposed to the full data generating process. Our approach gives rise to non-standard econometric issues, as it is not possible to fully characterize a single limit distribution of the observables, similarly to Epstein, Kaido, and Seo (2016), but only the set it belongs to. Yet, we show that we can use the limiting model to obtain a consistent estimator for the parameters of interest. Since behavior is not specified, our model is incomplete in the sense of Tamer (2003).

The ANR property is weaker than the one-shot no-ex post-regret property of pure-strategy Nash equilibrium that is sometimes invoked to motivate the choice of modeling cross-sectional data as equilibrium outcomes of a static game. Indeed, this descriptive[4] interpretation of static models is often paired with the assumptions of complete information and pure-strategy Nash equilibrium. The rationale for these assumptions is that the no-ex post-regret property of pure-strategy Nash equilibria reflects the stable nature of long-run outcomes.[5] Although appropriate for some environments, the static notion of no-ex post regret is a strong requirement: our work is thus complementary to standard equilibrium models of strategic interaction and provides an alternative whenever Nash equilibrium does not represent an appropriate restriction on behavior. In fact, Nash equilibrium of the static game is neither a natural long-run outcome of many simple game dynamics (for a review, see Hart and Mas-Colell (2013)), nor easy to compute in large games (BR2 (2017)).

In contrast, the ANR property is satisfied by a large class of well-known algorithms for the repeated play of the underlying one-shot game, once they are appropriately extended to games of incomplete information. This class includes simple adaptive heuristics, fictitious-play-like dynamics, more sophisticated learning rules involving active experimentation, calibrated learning, and several equilibrium dynamics. Since we do not fully specify what the behavior of players is or what they do to play according to this minimal long-run requirement, we depart from the current literature on empirical dynamic games that typically imposes the Markov perfect (or related) solution concept (for a review, see Ackerberg, Benkard, Berry, and Pakes (2007)).

Related Literature. Our work is related to the literature on learning in games, especially to Hart and Mas-Colell (2000, 2013), whose convergence results we extend to incomplete information environments. In contrast to these authors, the emphasis of our work is on connecting learning dynamics to the inference problem of an external observer. Recent contributions in computer science offer both related theoretical results (Hartline, Syrgkanis, and Tardos (2015)), and connection to empirical work

[3]Bayes Correlated Equilibrium is a generalization of Correlated Equilibrium to incomplete information environments developed by Bergemann and Morris (2016).

[4]In the sense of Pakes (2016)

[5]For instance, Ciliberto and Tamer (2009) argue as follows: "The idea behind cross-section studies is that in each market, firms are in a long-run equilibrium."

(Nekipelov, Syrgkanis, and Tardos (2015)), although the latter is specialized to online auction environments. Instead, we consider a general model where the primitive to be recovered is not the payoff type of players, but rather the structural features of payoffs, in line with the econometric literature on empirical games.

We are not the first to leverage on results in the literature on learning in games to perform empirical analysis. Lee and Pakes (2009) develop a learning-based procedure to compute counterfactuals in dynamic games. Several recent advances in the estimation of dynamic games investigate tractable and less restrictive empirical models (e.g., Doraszelski, Lewis, and Pakes (2018)). Our paper proposes a valid descriptive approach that is complementary to these structural methods.

Magnolfi and Roncoroni (2017) also consider estimation of discrete games under the assumption of Bayes correlated equilibrium behavior. Although this paper proposes the use of a similar estimation technique, the motivation is very different. In fact, Magnolfi and Roncoroni (2017) exploit the link between equilibrium behavior and information to establish that Bayes correlated equilibrium allows to estimate static discrete games under minimal assumptions on information. In this paper, instead, we motivate the use of Bayes correlated equilibrium as a behavioral restriction when the data are generated by repeated interaction.

Road Map. In Section 4.2, we present the basic theoretical setup. First, we formalize the notions of regret and asymptotic no-conditional regret for the repeated play of a one-shot incomplete information game. Then, we study the convergence properties of no-conditional regret dynamics. In Section, 4.3 we specialize the theoretical model to study what features of the underlying economic environment we can empirically recover, and how, under minimal assumptions on behavior when the one-shot game is played repeatedly over time. In Section 4.4, we present several extensions of the main results and outline the next steps of this research. In Section 4.5, we conclude.

4.2 Theoretical Model

4.2.1 Basic Setup

Incomplete Information Game. Following Bergemann and Morris (2011, 2016) (and a standard practice in the literature), we decompose an incomplete information game into a basic game and an information structure. Formally, there is a finite set of I players, $\mathcal{I} := \{1, \ldots, I\}$, and we write i for a typical player. There is a finite set of payoff states, Θ, and we write θ for a typical state. A *basic game* G consists of: *(i)* for each player i, a finite set of actions A_i, where we write $A := A_1 \times \cdots \times A_I$, and a utility function $u_i\colon A \times \Theta \to \mathbb{R}$; and *(ii)* a full support common prior $\psi \in \Delta_{++}(\Theta)$. Thus, $G := \big((A_i, u_i)_{i=1}^{I}, \psi\big)$. An *information structure* S consists of: *(i)* for each player i, a finite set of signals (or types) T_i, where we write $T := T_1 \times \cdots \times T_I$; and *(ii)* a signal distribution $\pi\colon \Theta \to \Delta(T)$. Thus, $S := \big((T_i)_{i=1}^{I}, \pi\big)$. Together, the basic game G and the information structure S define a standard *incomplete information game*, which we identify with the pair (G, S). We use the standard notation a_{-i} to denote a profile of actions for players other than i, i.e., $a_{-i} := (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_I)$. Analogously, t_{-i} denotes a profile of signals for players other than i.

When Θ is a singleton, the game is one of *complete information* A possible (and natural) information structure is the *null* information structure, in which each player's set of signals is a singleton, i.e. $T_i := \{t_i\}$. This corresponds to the situation in which each player has no information over and above the common prior ψ.

Bayes Correlated Equilibrium. The relevant space of uncertainty in the incomplete information game (G, S) is $A \times T \times \Theta$. We write ν for a typical element of $\Delta(A \times T \times \Theta)$. The equilibrium notion we are interested in, *Bayes correlated equilibrium*, is defined through the restrictions that we impose on ν. The first restriction, *consistency*, is a feasibility constraint which simply says that the marginal of distribution ν on the exogenous variables T and Θ is consistent with the description of the game (G, S).

Definition 16 (Consistency). *A probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is consistent for (G, S) if, for all $t \in T$ and $\theta \in \Theta$, we have*

$$\sum_{a \in A} \nu(a, t, \theta) = \pi(t \,|\, \theta)\psi(\theta).$$

The second condition, *interim obedience*, is an incentive constraint. Intuitively, a probability distribution ν is interim obedient if any player i who knows ν and is told his action–signal pair (a_i, t_i) from a realization of ν weakly prefers to play a_i, given that the other players, who know their realized action–signal pair, are going to play their part of the realized action profile.

Definition 17 (Interim ε-Obedience). *Let $\varepsilon \geq 0$. A probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is interim ε-obedient for (G, S) if for each $i \in \mathcal{I}$, $t_i \in T_i$, and $a_i \in A_i$, we have*

$$\sum_{a_{-i}, t_{-i}, \theta} \left[u_i\big((a_i', a_{-i}), \theta\big) - u_i\big((a_i, a_{-i}), \theta\big)\right] \nu\big((a_i, a_{-i}), (t_i, t_{-i}), \theta\big) \leq \varepsilon \tag{4.1}$$

for all $a_i' \in A_i$.

We now define the notion of Bayes correlated equilibrium of (G, S).

Definition 18 (Bayes Correlated ε-Equilibrium). *Let $\varepsilon \geq 0$. The probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is a Bayes correlated ε-equilibrium (hereafter, ε-BCE) of (G, S) if it is consistent and interim ε-obedient for (G, S). When $\varepsilon = 0$, we say that ν is a Bayes correlated equilibrium (hereafter, BCE) of (G, S). Denote with $E(\varepsilon)$ the set of ε-BCE of (G, S). We denote with $E(\varepsilon)$ the set of ε-BCE of (G, S).*

The notion of BCE we adopt is due to BM2 (2013); Bergemann and Morris (2011, 2016); it can be seen as an incomplete information version of correlated equilibrium (Aum; Aumann (1987)). When Θ is a singleton, the definition of ε-BCE reduces to the Aumann (1987) definition of correlated ε-equilibrium for a complete information game.

We say that the distribution $\nu \in \Delta(A \times T \times \Theta)$ is a Bayes Nash equilibrium action-signal-state distribution of (G, S) if there exists a Bayes Nash equilibrium $\beta := (\beta_1, \dots, \beta_I)$ of (G, S), where $\beta_i \colon T_i \to \Delta(A_i)$ for all $i \in \mathcal{I}$, such that

$$\nu(a, t, \theta) = \psi(\theta)\pi(t \mid \theta) \prod_{i=1}^{I} \beta_i(a_i \mid t_i)$$

for all $a \in A$, $t \in T$, and $\theta \in \Theta$. It is straightforward to show that every Bayes Nash equilibrium action-signal-state distribution of (G, S) is a BCE of (G, S) (see Lemma 1 in Bergemann and Morris (2011)). Moreover, $E(\varepsilon)$ is convex and contains the convex hull of all Bayes Nash equilibrium action-signal-state distributions of (G, S).

Conditional Regrets and No-Conditional Regret Dynamics. Suppose that the game (G, S) is played repeatedly over time. Time is discrete, and periods are indexed by $n \in \mathbb{N} := \{1, 2, \dots\}$. We maintain the following assumptions on the timing of events within periods and on what players observe about the game being played.

Assumption 4. *The timing within each period $n \in \mathbb{N}$ is the following:*

(i) A new state $\theta^n \in \Theta$ is drawn form ψ;

(ii) Given a realized state θ^n, a profile of signals $(t_1^n, \dots, t_I^n)$ is drawn from $\pi(\cdot \mid \theta^n)$;

(iii) After observing his signal $t_i^n \in T_i$, each player $i \in \mathcal{I}$ selects an action $a_i^n \in A_i$ and payoffs realize;

(iv) At the end of the period, each player $i \in \mathcal{I}$ observes the realized state θ^n, the profile of actions a^n that has been played, and his own realized payoff $u_i(a^n, \theta^n)$.

Definition 19 (Sequence of Actions, Signals, and States)**.** *We say that $\big((a^n, t^n, \theta^n)\big)_{n \in \mathbb{N}}$ is a sequence of actions, signals, and states from (G, S) if the game (G, S) is played repeatedly over time under Assumption 4 and $(a^n, t^n, \theta^n) \in A \times T \times \Theta$ for all $n \in \mathbb{N}$.*

Under Assumption 4, $(\theta^n)_{n \in \mathbb{N}}$ is a sequence of i.i.d. realizations of payoff states. Moreover, the following objects are well-defined.

- For all $i \in \mathcal{I}$ and $t_i \in T_i$, denote by $U_i(t_i; N)$ the average payoff that player i with signal t_i has obtained up to time N; that is,

$$U_i(t_i; N) := \frac{1}{N} \sum_{n=1}^{N} u_i\big((a_i^n, a_{-i}^n), \theta^n\big) \mathbb{1}_{\{t_i\}}(t_i^n).$$

- Let j be the last action played by player i with signal t_i up to time N. For each action $k \in A_i$, $k \neq j$, let $V_i(t_i, j, k; N)$ be the average payoff player i with signal t_i would have obtained had he played k instead of j every time in the past that he actually played j; that is,

$$V_i(t_i, j, k; N) := \frac{1}{N} \sum_{n=1}^{N} v_i^n(t_i, j, k),$$

where, for each $n \in \mathbb{N}$,

$$v_i^n(t_i, j, k) := \begin{cases} u_i\big((k, a_{-i}^n), \theta^n\big)\mathbb{1}_{\{t_i\}}(t_i^n) & \text{if } a_i^n = j \\ u_i\big((a_i^n, a_{-i}^n), \theta^n\big)\mathbb{1}_{\{t_i\}}(t_i^n) & \text{if } a_i^n \neq j \end{cases} .$$

Definition 20 (Conditional Regret)**.** *Suppose that the game (G, S) is played repeatedly over time under Assumption 4. For all $i \in \mathcal{I}$, $t_i \in T_i$ and $j, k \in A_i$, the conditional regret $R_i(j, k; t_i, N)$ for action k with respect to action j before play at time $N + 1$ is defined by*

$$R_i(j, k; t_i, N) := \max \{V_i(t_i, j, k; N) - U_i(t_i; N), 0\} .$$

The expression $R_i(j, k; t_i, N)$ has a clear interpretation as a measure of the (average) "regret" experienced by player i with signal t_i at period N for not having played, every time that j was played in the past, the different action k. The notion of conditional regrets for the repeated play of a complete information game is due to Hart and Mas-Colell (2000). We extend their notion to incomplete information games in a natural way—with each player computing his own conditional regrets signal-by-signal.[6]

Definition 21 (Asymptotic ε-No-Conditional Regret)**.** *Let $\varepsilon \geq 0$. A sequence of actions, signals, and states $\big((a^n, t^n, \theta^n)\big)_{n \in \mathbb{N}}$ from (G, S) has asymptotic ε-no-conditional regret (hereafter, ε-ANCR) if*

$$\limsup_{N \to \infty} R_i(j, k; t_i, N) \leq \varepsilon$$

for all $i \in \mathcal{I}$, $t_i \in T_i$, and $j, k \in A_i$ with $j \neq k$. When $\varepsilon = 0$, we say that sequence of actions, signals, and states has asymptotic no-conditional regret (hereafter, ANCR).

Asymptotic ε-no-conditional regret can be interpreted as a minimal long-run optimality condition for the repeated play of the one-shot game (G, S). When the sequence of actions, signals, and states has ε-ANCR, the average "regret" experienced by each type of each player for not having played different actions vanishes (or is ε-close to vanish) in the long-run. There are many strategies that players can follow in the repetition of the one-shot game (G, S) and that generate a sequence of actions, signals, and states satisfying ε-ANCR (see Section 4.2.3 for a few examples). For our empirical exercise (see Section 4.3), however, we do not take any stand on how players play the one-shot game (G, S) in each period. We only assume that players (learn to) play game (G, S) sufficiently well for the sequence of actions, signals, and states to satisfy, in the long-run, the mild optimality condition captured by ε-ANCR.

[6]This is a natural choice under the assumptions that each player's set of signals is finite and that payoff states are i.i.d. across time periods.

4.2.2 Convergence of No-Conditional Regret Dynamics

Definition 22 (Empirical Distribution). *Let $\big((a^n, t^n, \theta^n)\big)_{n\in\mathbb{N}}$ be a sequence of actions, signals, and states from (G, S). For every $N \in \mathbb{N}$, the empirical distribution $Z^N \in \Delta(A \times T \times \Theta)$ is defined pointwise as*

$$Z^N(a, t, \theta) := \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\{a\}}(a^n)\mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n) \tag{4.2}$$

for all $(a, t, \theta) \in A \times T \times \Theta$.

That is, $Z^N(a, t, \theta)$ is the relative frequency of the action-signal-state profile (a, t, θ) in the first N periods.

The next theorem shows that a necessary and sufficient condition for the sequence of empirical distributions to converge almost surely to the set of ε-BCE of (G, S) is that the sequence of actions, signals, and states has ε-ANCR.

Theorem 3 (ε-ANCR and ε-BCE). *The sequence of actions, signals, and states $\big((a^n, t^n, \theta^n)\big)_{n\in\mathbb{N}}$ from (G, S) has asymptotic ε-no-conditional regret almost surely for some $\varepsilon \geq 0$ if and only if, as $N \to \infty$, the sequence of empirical distributions $\big(Z^N\big)_{N\in\mathbb{N}}$ converges almost surely to the set of Bayes correlated ε-equilibria of (G, S).*

For our empirical exercise, we are mostly interested in the sufficiency part of the previous result. Suppose that players play the one-shot game (G, S) repeatedly over time under Assumption 4. If their play satisfies the minimal long-run optimality condition captured by ε-ANCR, then, by Theorem 3, the empirical distribution converges almost surely to $E(\varepsilon)$ of the underlying incomplete information game (G, S); that is, from some time on, the empirical distribution is almost surely close to an ε-BCE of (G, S). The convergence here is to the set of ε-BCE, not necessarily to a point in that set. Moreover, observe that it is the empirical distribution that becomes essentially an ε-BCE, not (necessarily) the actual play.

Proof of Theorem 3. [$\Longrightarrow$] Suppose the sequence of actions, signals, and states $\big((a^n, t^n, \theta^n)\big)_{n\in\mathbb{N}}$ from (G, S) has ε-ANCR almost surely for some $\varepsilon \geq 0$. Consider any subsequence $\big(Z^{N_l}\big)_{l\in\mathbb{N}}$ of $\big(Z^N\big)_{N\in\mathbb{N}}$ that converges almost surely to some $\nu \in \Delta(A \times T \times \Theta)$. We need to show that ν is almost surely an ε-BCE of (G, S), i.e., that ν is almost surely consistent and interim ε-obedient for (G, S).

Consistency. Pick any $(t, \theta) \in T \times \Theta$. Note the following:

$$\begin{aligned}
\sum_a \nu(a, t, \theta) &= \sum_a \lim_{l\to\infty} Z^{N_l}(a, t, \theta) \\
&= \lim_{l\to\infty} \sum_a Z^{N_l}(a, t, \theta) \\
&= \lim_{l\to\infty} \left[\frac{\sum_a Z^{N_l}(a, t, \theta)}{\sum_{a,t} Z^{N_l}(a, t, \theta)} \sum_{a,t} Z^{N_l}(a, t, \theta) \right]
\end{aligned} \tag{4.3}$$

$$= \lim_{l\to\infty} \frac{\sum_a Z^{N_l}(a,t,\theta)}{\sum_{a,t} Z^{N_l}(a,t,\theta)} \lim_{l\to\infty} \sum_{a,t} Z^{N_l}(a,t,\theta)$$

$$= \lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} \lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)}{N_l}.$$

The ratio

$$\frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} \tag{4.4}$$

is the empirical frequency of the signal profile t when filtered at time steps where the state is θ. As $(\theta^n)_{n\in\mathbb{N}}$ is an i.i.d. sequence, and the t's are drawn from $\pi(\cdot\,|\,\theta)$, (4.4) is the empirical frequency of $\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)$ independent observations from $\pi(\cdot\,|\,\theta)$. Moreover, as ψ has full support, $\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n) \to \infty$ as $l\to\infty$. Thus, by the strong law of large numbers,

$$\lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} = \pi(t\,|\,\theta) \qquad \text{almost surely.} \tag{4.5}$$

Again, as $(\theta^n)_{n\in\mathbb{N}}$ is an i.i.d. sequence, by the strong law of large numbers,

$$\lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)}{N_l} = \psi(\theta) \qquad \text{almost surely.} \tag{4.6}$$

Together, (4.3), (4.5), and (4.6) give

$$\sum_a \nu(a,t,\theta) = \pi(t\,|\,\theta)\psi(\theta) \qquad \text{almost surely.} \tag{4.7}$$

As $(t,\theta)\in T\times\Theta$ was arbitrarily chosen, we conclude from (4.7) that ν is almost surely consistent for (G,S).

Interim ε-obedience. To begin, note the following:

$$\begin{aligned} V_i(t_i,j,k;N) - U_i(t_i;N) &= \frac{1}{N}\sum_{n=1}^{N} \big[u_i\big((k,a_{-i}^n),\theta^n\big) - u_i\big((a_i^n,a_{-i}^n),\theta^n\big)\big]\mathbb{1}_{\{j\}}(a_i^n)\mathbb{1}_{\{t_i\}}(t_i^n) \\ &= \frac{1}{N}\sum_{\theta\in\Theta}\sum_{n=1}^{N} \big[u_i\big((k,a_{-i}^n),\theta^n\big) - u_i\big((a_i^n,a_{-i}^n),\theta^n\big)\big]\mathbb{1}_{\{j\}}(a_i^n)\mathbb{1}_{\{t_i\}}(t_i^n)\mathbb{1}_{\{\theta\}}(\theta^n) \\ &= \sum_{a_{-i},t_{-i},\theta} \big[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\big] Z^N\big((j,a_{-i}),(t_i,t_{-i}),\theta\big). \end{aligned} \tag{4.8}$$

Now pick any $i\in\mathcal{I}$, $t_i\in T_i$, and $j,k\in A_i$ with $j\neq k$. As $\limsup_{N\to\infty} R_i(j,k;t_i,N)\le\varepsilon$ almost surely, by definition of $R_i(j,k;t_i,N)$, we also have $\limsup_{N\to\infty}\big[V_i(t_i,j,k;N)-U_i(t_i;N)\big]\le\varepsilon$ almost surely. But then, by (4.8),

$$\limsup_{N\to\infty} \sum_{a_{-i},t_{-i},\theta} \big[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\big] Z^N\big((j,a_{-i}),(t_i,t_{-i}),\theta\big) \le \varepsilon \qquad \text{almost surely.} \tag{4.9}$$

Moreover, on the subsequence $\left(Z^{N_l}\right)_{l\in\mathbb{N}}$ we get

$$\begin{aligned}
&\lim_{l\to\infty}\sum_{a_{-i},t_{-i},\theta}\left[u_i\big((k,a_{-i}),\theta\big)-u_i\big((j,a_{-i}),\theta\big)\right]Z^{N_l}\big((j,a_{-i}),(t_i,t_{-i}),\theta\big)\\
&=\sum_{a_{-i},t_{-i},\theta}\lim_{l\to\infty}\left[u_i\big((k,a_{-i}),\theta\big)-u_i\big((j,a_{-i}),\theta\big)\right]Z^{N_l}\big((j,a_{-i}),(t_i,t_{-i}),\theta\big)\\
&=\sum_{a_{-i},t_{-i},\theta}\left[u_i\big((k,a_{-i}),\theta\big)-u_i\big((j,a_{-i}),\theta\big)\right]\nu\big((j,a_{-i}),(t_i,t_{-i}),\theta\big).
\end{aligned}\tag{4.10}$$

Together, (4.9) and (4.10) give

$$\sum_{a_{-i},t_{-i},\theta}\left[u_i\big((k,a_{-i}),\theta\big)-u_i\big((j,a_{-i}),\theta\big)\right]\nu\big((j,a_{-i}),(t_i,t_{-i}),\theta\big)\le\varepsilon\qquad\text{almost surely.}\tag{4.11}$$

As $i\in\mathcal{I}$, $t_i\in T_i$, and $j,k\in A_i$ with $j\neq k$ were arbitrarily chosen, we conclude from (4.11) that ν is almost surely interim ε-obedient for (G,S).

[$\Longleftarrow$] Now suppose the sequence of empirical distributions $\left(Z^N\right)_{N\in\mathbb{N}}$ converges almost surely to the set of Bayes correlated ε-equilibria of (G,S) for some $\varepsilon\geq 0$. Pick any $i\in\mathcal{I}$, $t_i\in T_i$, and $j,k\in A_i$ with $j\neq k$. By interim ε-obedience,

$$\limsup_{N\to\infty}\sum_{a_{-i},t_{-i},\theta}\left[u_i\big((k,a_{-i}),\theta\big)-u_i\big((j,a_{-i}),\theta\big)\right]Z^{N}\big((j,a_{-i}),(t_i,t_{-i}),\theta\big)\le\varepsilon\qquad\text{almost surely.}\tag{4.12}$$

By (4.8) and (4.12),

$$\limsup_{N\to\infty}V_i(t_i,j,k;N)-U_i(t_i;N)\le\varepsilon\qquad\text{almost surely.}$$

This, by definition of conditional regret,

$$\limsup_{N\to\infty}R_i(j,k;t_i,N)\le\varepsilon\qquad\text{almost surely.}$$

As $i\in\mathcal{I}$, $t_i\in T_i$, and $j,k\in A_i$ with $j\neq k$ were arbitrarily chosen, the desired result follows. ■

Remark 18. The almost sure convergence of $\left(Z^N\right)_{N\in\mathbb{N}}$ to $E(\varepsilon)$ means that the sequence $\left(Z^N\right)_{N\in\mathbb{N}}$ eventually enters any neighborhood of the set $E(\varepsilon)$ and stays there forever. An equivalent way of stating this is as follows: given any $\varepsilon'>\varepsilon$, there is a time $N(\varepsilon')$ after which the empirical distribution is always a ε'-BCE of (G,S) almost surely; that is, $Z^N\in E(\varepsilon')$ for all $N>N(\varepsilon')$ almost surely.

4.2.3 Asymptotic-No-Conditional-Regret Strategies

Suppose game (G,S) is played repeatedly over time under Assumption 4. In this section, we provide some examples of algorithms that players can follow in the repetition of the one-shot game (G,S) and that generate a sequence of actions, signals, and states satisfying ε-ANCR (hereafter, ε-ANCR algorithms). By Theorem 3, under ε-ANCR algorithms the sequence of empirical distributions $\left(Z^N\right)_{N\in\mathbb{N}}$

converges almost surely to the set of ε-BCE of (G, S). Although we only sketch here a few examples, it is worthwhile noting that the class of ε-ANCR algorithms is very large, ranging from simple adaptive heuristics to sophisticated learning dynamics, and even to repeated equilibrium play of the one-shot game (G, S).

Conditional Regret Matching and Generalizations. It is natural to extend the conditional regret matching algorithm—introduced by Hart and Mas-Colell (2000) for complete information games—to the incomplete information game (G, S). Let t_i be player i's signal in period $N+1$ and let j be the action played by player i the last time in the past he observed signal t_i. *Conditional regret matching* stipulates that each action $k \neq j$ is played in period $N+1$ with a probability that is proportional to its regret $R_i(j, k; t_i, N)$, and, with the remaining probability, the same action j is played in period $N+1$. Formally, denote with $p_i^{N+1}(k; t_i, j, N)$ the probability of playing action k in period $N+1$ by player i with signal t_i, given that i has played action j the last time in the past he observed signal t_i. Then, conditional regret matching prescribes that

$$p_i^{N+1}(k; t_i, j, N) = \begin{cases} cR_i(j, k; t_i, N) & \text{if } k \neq j \\ 1 - \sum_{k \neq j} cR_i(j, k; t_i, N) & \text{if } k = j \end{cases} \tag{4.13}$$

for some sufficiently small constant $c > 0$.[7] The play in the first period can be arbitrary.

Following the logic of Hart and Mas-Colell (2000), we can show that, if each player plays a conditional regret matching strategy in each period N, then, the sequence of actions, signals, and states has ANCR almost surely. The proof that all regrets vanish in the limit uses arguments suggested by Blackwell (1956)'s approachability.

Instead of the switching probability being proportional to the conditional regret $R_i(j, k; t_i, N)$, i.e., equal to $cR_i(j, k; t_i, N)$, we may want to allow this switching probability to be given by a general function $f(R_i(j, k; t_i, N))$ of $R_i(j, k; t_i, N)$. If f is sign-preserving (i.e., $f(x) > 0$ for $x > 0$ and $f(0) = 0$) and Lipschitz continuous, we call the resulting strategies *generalized conditional regret matching strategies*, in the spirit of Hart and Mas-Colell (2001a) and Cahn (2004). Building on these authors' results, we can show that, if each player plays a generalized conditional regret matching strategy in each period N, then, the sequence of actions, signals, and states has ANCR almost surely. In fact, the full class of conditional-regret-based strategies—for which, if played by all players, the sequence of actions, signals, and states has ANCR almost surely—is even larger. We refer to Hart (2005) for an extensive survey of regret-based strategies for the repeated play of complete information games. Their extension to incomplete information games mimics that for conditional regret matching and its generalization outlined above.

Even procedures that fall in the class of "reinforcement learning" algorithms (see, for example, Roth and Erev (1995), Börgers and Sarin (1997, 2000), and Erev and Roth (1998)), broadly defined as including those procedures whereby individuals react to past payoffs without full knowledge of the game, may lead to convergence results such as the ones we obtain in this paper. A relevant case is presented

[7] The constant c must guarantee that (4.13) yields a probability distribution over A_i and, moreover, that the probability of j is strictly positive.

in Hart and Mas-Colell (2001b), who develop a modified regret matching algorithm, dubbed proxy-regret matching, for environments where each player initially knows only his own set of actions and is informed, after each period of play, only of his realized payoff.

Calibrated Learning, Variants of Fictitious Play, and Other Dynamics. Previous work has identified several dynamics for the repeated play of a one-shot complete information game that converge to the set of correlated equilibria or correlated ε-equilibria of the underlying game. A notable example of such dynamics is the *calibrated learning* of (Foster and Vohra (1997)), Here, each player computes calibrated forecasts on the behavior of the other players, and then plays a best reply to these forecasts (Foster and Vohra (1997)).[8] Other notable examples are *conditional smooth fictitious play eigenvector strategies* (Fudenberg and Levine (1998, 1999a)) and *smooth conditional fictitious play* (Cahn (2004)), where each player i plays at each period a smoothed-out best reply to the distribution of the play of the opponents in those periods where i played the same action j as in the previous period.

With some work, it is possible to extend such dynamics to the repeated play of the one-shot incomplete information game (G, S) signal-by-signal, in the spirit of what we outline above for conditional regret matching and its generalizations. Building on previous work we can then show that such dynamics converge to the set of BCE or ε-BCE of (G, S). It follows from Theorem 3 that the sequence of actions, signals, and states has ANCR or ε-ANCR.

Equilibrium Play. Suppose that, in each period n, players play a Bayes Nash equilibrium of (G, S) (not necessarily the same). If so, the empirical distribution lies in the convex hull of all Bayes Nash equilibrium action-signal-state distributions of (G, S), which is contained in the set of BCE of (G, S). It follows from Theorem 3 that the sequence of actions, signals, and states has ANCR. Similarly, if players play a BCE of (G, S) in each period n, then the sequence of actions, signals, and states has ANCR.

Remark 19. Several ε-ANCR algorithms (e.g., regret matching, its generalizations, and smooth variants of fictitious play) only require regrets as an input. Therefore, to achieve ε-ANCR players do not need to know the utility functions of the other players or their signals. That is, there exist algorithms that achieve ε-ANCR and only require each player to know what other players do, not what their objectives or private information are.

4.3 From the Model to the Data

We now specialize the model developed in Section 4.2 to consider the following question: when players interact repeatedly in an incomplete information game, what features of the underlying one-shot game can we recover under minimal assumptions on behavior?

[8] Forecasts are calibrated if, roughly speaking, probabilistic forecasts and long-run frequencies are close: for example, an event must occur approximately $\pi\%$ of the times for which the forecast was a $\pi\%$ chance of the event. There are various ways to generate calibrated forecasts (see, among others, Foster and Vohra (1997, 1998, 1999), Foster (1997), Fudenberg and Levine (1999b), and Kakade and Foster (2008).

We leverage on the convergence results established in Section 4.2 to recover features of the underlying one-shot game under the assumption that the data we observe are the outcomes of an ongoing dynamic interaction. The literature on empirical games typically maintains that the observable actions results from equilibrium play in a cross-section of simultaneous games (e.g., Berry (1992), Tamer (2003), and Ciliberto and Tamer (2009)) or from fully rational dynamic equilibrium play in a panel of dynamic games (e.g., Ericson and Pakes (1995), Benkard (2004), Jofre-Bonet and Pesendorfer (2003), and Ryan (2012)). In contrast, we do not make any strong assumption on how players play the one-shot game in each period. We only assume that players (learn to) play the one-shot game sufficiently well for the sequence of actions, signals, and states to satisfy, in the long-run, the minimal optimality condition captured by ε-ANCR.

Empirical Model. To begin, we lay out the main assumptions on the empirical model and the observables. In particular, we assume the following.

(i) The incomplete information game $\left(G\left(\lambda^G\right), S\left(\lambda^S\right)\right)$ belongs to a parametrized class with structural parameters $\lambda := \left(\lambda^G, \lambda^S\right) \in \Gamma \subseteq \mathbb{R}^{d_g} \times \mathbb{R}^{d_s}$.

(ii) $\lambda_0 := \left(\lambda_0^G, \lambda_0^S\right)$ are the true structural parameters.

(iii) The sequence of actions, signals, and states $\left(\left(a^n, t^n, \theta^n\right)\right)_{n\in\mathbb{N}}$

  (a) Is generated by play in $\left(G\left(\lambda_0^G\right), S\left(\lambda_0^S\right)\right)$ that satisfies Assumption 4;

  (b) Has asymptotic ε-no-conditional regret for some $\varepsilon \geq 0$.

(iv) The econometrician only observes the realized sequence of actions $\left(a^n\right)_{n\in\mathbb{N}}$.

Assumption 5. *The empirical model and the observables are summarized by (i)–(iv) above.*

Although players observe the sequence of actions, signals, and states $\left(\left(a^n, t^n, \theta^n\right)\right)_{n\in\mathbb{N}}$, in applied contexts outside analysts have typically less information than players. Therefore, we assume that the econometrician only observes the realized sequence of actions $\left(a^n\right)_{n\in\mathbb{N}}$ (i.e., *(iv)* above). This assumption is the most common in the empirical literature, but it is not the only one. For instance, BM2 (2013) consider identification under BCE behavior in a model where both actions and payoff states, i.e., $\left(a^n, \theta^n\right)_{n\in\mathbb{N}}$, are observable to the econometrician.

In the remaining part of this section, we investigate what we can recover of λ_0, and how, under Assumption 5.

Bayes Correlated Equilibrium and Restrictions on Parameters. To recover structural parameters λ_0 under Assumption 5, we leverage on the convergence results in Section 4.2, which motivates the adoption of our equilibrium restrictions. In particular, we will show the ε-BCE assumption allows to recover valid bounds on the structural parameter. We have already defined the notion of ε-BCE; for future reference, denote with $E(\lambda; \varepsilon)$ the set of ε-BCE of the incomplete information game $\left(G\left(\lambda^G\right), S\left(\lambda^S\right)\right)$ with structural parameters $\lambda := \left(\lambda^G, \lambda^S\right) \in \Gamma$. We now expand on the restrictions that the ε-BCE equilibrium assumption implies for the structural parameters.

Definition 23 (ε-BCE Prediction)**.** *Let $\varepsilon \geq 0$. A probability distribution $q \in \Delta(A)$ is an ε-BCE prediction if there exists $\nu \in E(\lambda; \varepsilon)$ such that*

$$q(a) = \sum_{t \in T, \theta \in \Theta} \nu(a, t, \theta)$$

for all $a \in A$. The set of ε-BCE predictions for a game with structural parameters λ is denoted by $Q(\lambda; \varepsilon)$.

Definition 24 (Identified Set)**.** *Let $q \in \Delta(A)$ be a distribution of actions. The set of parameters identified by q under the ε-BCE assumption, denoted by $\Lambda_I(q; \varepsilon)$, is*

$$\Lambda_I(q; \varepsilon) \coloneqq \{\lambda \in \Lambda : q \in Q(\lambda; \varepsilon)\}.$$

In our model, we do not observe a limiting "population" distribution of the observables, i.e., a fixed limiting $q \in \Delta(A)$. To see this, let the *empirical distribution of actions* $q^N \in \Delta(A)$ be defined pointwise as

$$q^N(a) \coloneqq \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\{a\}}(a^n)$$

for all $a \in A$. Under Assumption 5, Theorem 3 only ensures that the sequence of empirical distributions $\left(Z^N\right)_{N \in \mathbb{N}}$ converges almost surely to $E(\lambda; \varepsilon)$ as $N \to \infty$. Therefore, as q^N is the marginal on A of the empirical distribution Z^N for all $N \in \mathbb{N}$, i.e.,

$$q^N(a) = \sum_{t \in T, \theta \in \Theta} Z^N(a, t, \theta)$$

for all $a \in A$, Theorem 3 only ensures that, as the sample size gets large, time averages of actions converge almost surely to the set $Q(\lambda_0; \varepsilon)$, not necessarily to a point in that set. To overcome this complication, instead of focusing on the identified set $\Lambda_I(q; \varepsilon)$, we consider the set of parameters that can be recovered when *any* $q \in Q(\lambda_0; \varepsilon)$ may describe the data, which leads to the next definition.

Definition 25 (Recoverable Set)**.** *The set of recoverable parameters under Assumption 5, denoted by $\Lambda_R(\varepsilon)$, is*

$$\Lambda_R(\varepsilon) \coloneqq \bigcup_{q \in Q(\lambda_0; \varepsilon)} \Lambda_I(q; \varepsilon).$$

The bounds imposed by $\Lambda_R(\varepsilon)$ are valid, in the sense that $\lambda_0 \in \Lambda_R(\varepsilon)$.[9]

Recovering Bounds on Parameters. Consider the "plug-in" estimator

$$\hat{\Lambda}_N(\varepsilon) \coloneqq \left\{\lambda \in \Lambda : \ q^N \in Q(\lambda; \varepsilon)\right\} = \Lambda_I\left(q^N; \varepsilon\right),$$

where q^N is the observed empirical distribution of N actions.

Theorem 4 (Properties of $\hat{\Lambda}_N$)**.** *Under Assumption 5, for any $\varepsilon' > \varepsilon$, the following statements hold almost surely as $N \to \infty$:*

[9] That $\lambda_0 \in \Lambda_R(\varepsilon)$ holds by construction of $\Lambda_R(\varepsilon)$.

(i) $\lambda_0 \in \hat{\Lambda}_N(\varepsilon')$;

(ii) $\hat{\Lambda}_N(\varepsilon) \subseteq \Lambda_R(\varepsilon')$.

Theorem 4 says that a static equilibrium notion, ε-BCE, provides an adequate behavioral restriction for the estimation of dynamic interactions that satisfy the minimal long-run optimality condition captured by ε-ANCR assumption. Part *(i)* of the theorem establishes that the restriction of ε-ANCR leads to estimating a set of parameters which contains the true structure of the data generating process. Part *(ii)* describes bounds on this estimated set, which is contained within the (theoretical) recoverable set. The width of the bounds, in practice, will depend on the specific model and on the informativeness of the data. Despite data not being generated by the repetition of identical experiments, we bound structural parameters without statistical assumptions on the sampling process on top of the economic assumption of ε-no-conditional regret in the limit.

Proof of Theorem 4. To establish part *(i)*, fix any $\varepsilon' > \varepsilon$ and note that, by definition of $\hat{\Lambda}_N(\varepsilon')$,

$$\lambda_0 \in \hat{\Lambda}_N(\varepsilon') \Longleftrightarrow q^N \in Q(\lambda_0; \varepsilon'). \tag{4.14}$$

Under Assumption 5, by Theorem 3 we have that the sequence of empirical distributions $\left(Z^N\right)_{N\in\mathbb{N}}$ converges almost surely to $E(\varepsilon)$ as $N \to \infty$. Then, by Remark 18, there exists $N(\varepsilon')$ such that $Z^N \in E(\varepsilon')$ for all $N > N(\varepsilon')$ almost surely. As q^N is the marginal on A of the empirical distribution Z^N for all $N \in \mathbb{N}$, it follows that $q^N \in Q(\lambda_0; \varepsilon')$ for all $N > N(\varepsilon')$ almost surely. Combining this fact with (4.14) gives the desired result.

To establish part *(ii)*, fix any $\varepsilon' > \varepsilon$ and note that, by definition of $\Lambda(\varepsilon')$ and of $\hat{\Lambda}_N(\varepsilon')$,

$$q^N \in Q(\lambda_0; \varepsilon') \Longleftrightarrow \Lambda_I\left(q^N; \varepsilon'\right) \subseteq \Lambda_R(\varepsilon') \Longleftrightarrow \hat{\Lambda}_N(\varepsilon') \subseteq \Lambda_R(\varepsilon'). \tag{4.15}$$

Moreover, as $\varepsilon' > \varepsilon$,

$$\hat{\Lambda}_N(\varepsilon) \subseteq \hat{\Lambda}_N(\varepsilon'). \tag{4.16}$$

Under Assumption 5, by Theorem 3 we have that $q^N \in Q(\lambda_0; \varepsilon')$ for all $N > N(\varepsilon')$ almost surely (see the proof of part *(i)*). Combining this fact with (4.15) and (4.16) gives the desired result. ■

Remark 20. Assumption 5 maintains that the econometrician observes the entire realized sequence of actions $(a^n)_{n\in\mathbb{N}}$. Often, however, data do not capture the complete path of play and may come without precise time identifiers. Our analysis easily extends to the case where the econometrician only observes a subsequence $(a^{n_l})_{l\in\mathbb{N}}$ (i.e., a sample) of the realized sequence of actions.

To see this, suppose that the econometrician only observes the subsequence $(a^{n_l})_{l\in\mathbb{N}}$ of the realized sequence of actions $(a_n)_{n\in\mathbb{N}}$. Under Assumption 5, the realized sequence of actions, signals, and states $\left((a^n, t^n, \theta^n)\right)_{n\in\mathbb{N}}$ has ε-ANCR. Therefore,

$$\limsup_{N\to\infty} R_i(j, k; t_i, N) \leq \varepsilon \tag{4.17}$$

for all $i \in \mathcal{I}$, $t_i \in T_i$, and $j, k \in A_i$ with $j \neq k$. By standard properties of the $\limsup$ operator, on the subsequence of actions, signals, and states $\big((a^{n_l}, t^{n_l}, \theta^{n_l})\big)_{l\in\mathbb{N}}$ corresponding to $(a^{n_l})_{n\in\mathbb{N}}$ we have that

$$\limsup_{L\to\infty} R_i(j,k;t_i,L) \leq \varepsilon \tag{4.18}$$

for all $i \in \mathcal{I}$, $t_i \in T_i$, and $j, k \in A_i$ with $j \neq k$. That is, also the sequence $\big((a^{n_l}, t^{n_l}, \theta^{n_l})\big)_{l\in\mathbb{N}}$ has ε-ANCR. Thus, by Theorem 3, as $L \to \infty$, the sequence of empirical distributions $\big(Z^L\big)_{L\in\mathbb{N}}$ converges almost surely to the set of ε-BCE of (G, S). It follows that Theorem 4 applies unchanged to the "plug-in" estimator $\hat{\Lambda}_L(\varepsilon) := \Lambda_I\big(q^L;\varepsilon\big)$, where $q^L \in \Delta(A)$ is the *observed* empirical distribution of actions defined pointwise as

$$q^L(a) := \frac{1}{L}\sum_{l=1}^{L} \mathbb{1}_{\{a\}}(a^{n_l})$$

for all $a \in A$.

4.4 Extensions and Discussion

4.4.1 Bayes Coarse Correlated Equilibrium and Unconditional Regrets

Bayes Coarse Correlated Equilibrium. The form of obedience we impose on ν distinguishes the notion of Bayes correlated equilibrium from that of *Bayes coarse correlated equilibrium*. Intuitively, a probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is *ex ante obedient* if any player i who knows ν, is told his signal t_i (but not his action a_i) from a realization of ν, and is given a choice between *(i)* committing to whatever joint action profile (a_i, a_{-i}) has realized from ν, and *(ii)* committing to a fixed action a_i', weakly prefers *(i)* to *(ii)*, given that the other players, who know their realized signal (but not their realized action), are committed to playing their part of whatever joint action has realized.

Definition 26 (Ex Ante ε-Obedience)**.** *Let $\varepsilon \geq 0$. A probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is ex ante ε-obedient for (G, S) if, for each $i \in \mathcal{I}$ and $t_i \in T_i$, we have*

$$\sum_{a,t_{-i},\theta} \big[u_i\big((a_i', a_{-i}), \theta\big) - u_i(a,\theta)\big]\nu\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon \tag{4.19}$$

for all $a_i' \in A_i$.

We now define the notion of Bayes coarse correlated equilibrium of (G, S).

Definition 27 (Bayes Coarse Correlated ε-Equilibrium)**.** *Let $\varepsilon \geq 0$. The probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is a Bayes coarse correlated ε-equilibrium (hereafter, ε-BCCE) of (G, S) if it is consistent and ex ante ε-obedient for (G, S). When $\varepsilon = 0$, we say that ν is a Bayes coarse correlated equilibrium (hereafter, BCCE) of (G, S).*

The notion of BCCE can be seen as an incomplete information version of coarse correlated equilibrium (Hannan (1957), Moulin and Vial (1978), and Young (2004)). We extend the notion of BCE due to BM2 (2013); Bergemann and Morris (2011, 2016) to its coarse analogue in the natural way. When Θ is a singleton, the definition of ε-BCCE reduces to the definition of coarse correlated ε-equilibrium for a complete information game. For fixed incomplete information game (G, S) and $\varepsilon \geq 0$, it is straightforward to show that the set of ε-BCCE of (G, S) is convex and contains the set of ε-BCE of (G, S).

Unconditional Regrets and No-Unconditional Regret Dynamics. Suppose that the game (G, S) is played repeatedly over time under Assumption 4. For each action $k \in A_i$, let $\widehat{V}_i(t_i, k; N)$ be the average payoff player i with signal t_i would have obtained had he played k in all periods up to time N; that is,

$$\widehat{V}_i(t_i, k; N) := \frac{1}{N} \sum_{n=1}^{N} u_i\big((k, a_{-i}^n), \theta^n\big) \mathbb{1}_{\{t_i\}}(t_i^n).$$

Definition 28 (Unconditional Regret)**.** *For all $i \in \mathcal{I}$, $t_i \in T_i$ and $k \in A_i$, the unconditional regret $\widehat{R}_i(k; t_i, N)$ for action k before play at time $N+1$ is defined by*

$$\widehat{R}_i(k; t_i, N) := \max\left\{\widehat{V}_i(t_i, k; N) - U_i(t_i; N), 0\right\}.$$

The expression $\widehat{R}_i(k; t_i, N)$ has a clear interpretation as a measure of the (average) "regret" experienced by player i with signal t_i at period N for not having played action k in all past periods up to N. Unconditional regrets are a rougher measure of regret than conditional regrets; namely, they are based on the increase in the average payoff, if any, were player i with signal t_i to replace all past plays, and not just the j-plays, by k.

Definition 29 (Asymptotic No Unconditional Regret)**.** *Let $\varepsilon \geq 0$. A sequence of actions, signals, and states $\big((a^n, t^n, \theta^n)\big)_{n\in\mathbb{N}}$ from (G, S) has asymptotic ε-no unconditional regret (hereafter, ε-ANUR) if*

$$\limsup_{N\to\infty} \widehat{R}_i(k; t_i, N) \leq \varepsilon$$

for all $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$.

Convergence of No-Unconditional Regret Dynamics. The next theorem parallels Theorem 3 and shows that a necessary and sufficient condition for the sequence of empirical distributions to converge (almost surely) to the set of ε-BCCE is that the sequence of actions, signals, and states has ε-ANUR.

Theorem 5 (ε-ANUR and ε-BCCE)**.** *The sequence of actions, signals, and states $\big((a^n, t^n, \theta^n)\big)_{n\in\mathbb{N}}$ from (G, S) has asymptotic ε-no unconditional regret almost surely for some $\varepsilon \geq 0$ if and only if, as $N \to \infty$, the sequence of empirical distributions $\big(Z^N\big)_{N\in\mathbb{N}}$ converges almost surely to the set of Bayes coarse correlated ε-equilibria of (G, S).*

Proof. See Appendix 4.6 ■

From the Model to the Data with Unconditional Regrets. The empirical exercise we perform in Section 4.3 extends in the obvious way when we modify condition *(iii)–(b)* of Assumption 5 by requiring that the sequence $\left((a^n, t^n, \theta^n)\right)_{n\in\mathbb{N}}$ satisfies ε-ANUR instead of ε-ANCR.

A motivation for our work is to provide valid bounds on structural parameters under minimal assumptions on behavior. Bounds on parameters are less sharp under the ε-ANUR assumption than under the ε-ANCR assumption, as the set of ε-BCE of (G, S) is contained in the set of ε-BCCE of (G, S). However, the ε-ANUR assumption is more robust, at least in the following sense. Call an algorithm a player can follow in the repetition of the one-shot game (G, S) *Hannan-consistent* if it guarantees, for any algorithms other players may follow, that all the regrets of this player vanish in the limit with probability one. That is, an algorithm is Hannan-consistent if it unilaterally assures vanishing regrets independently of what other players do. Hannan-consistent algorithms exist for both conditional and unconditional regrets (see Blackwell (1954), Hannan (1957), Fudenberg and Levine (1995, 1998), Foster and Vohra (1993, 1998, 1999), Freund and Schapire (1999), Hart and Mas-Colell (2000), and Young (2004)—their extension to incomplete information settings is straightforward). Importantly, however, unconditional regret matching and its generalizations[10] are Hannan-consistent, whereas conditional regret matching and its generalizations are not.[11] As Remark 19 points out, regret matching algorithms only require regrets as an input. It follows that to achieve the long-run optimality condition captured by ε-ANUR players need to know neither the utility functions of the other players and their signals nor the algorithms that other players adopt. In contrast, to achieve the long-run optimality condition captured by ε-ANCR: *(a)* If players do not know the utility functions or the signals of other players, they need to know what regret-minimizing algorithm the other players adopt and to coordinate on this algorithm; *(b)* If, instead, they do not know what other players do, they need to design more sophisticated algorithms to assure that their own conditional regrets vanish—these algorithms requiring common knowledge of the underlying one-shot game.

4.4.2 Robustness of No-Conditional Regret Dynamics

There are at least two natural orderings on information structures: an "incentive ordering" and a "statistical ordering". Roughly speaking, we have the following.[12]

- Incentive ordering: an information structure is more incentive constrained than another if it gives rise to a smaller set of BCE.
- Statistical ordering: an information structure is individually sufficient for another if there exists a combined information structure where each player's signal from the former information structure is a sufficient statistic for the state and other players' signals in the latter information

[10]The definition of unconditional regret matching mimics that of conditional regret matching, except for using unconditional regrets instead of conditional regrets.

[11]That is, unconditional regret matching assures non-positive unconditional regrets for any player who uses it irrespective of the behavior of the other players. In contrast, if a single player uses conditional regret matching, there is no assurance that his conditional regrets will become non-positive over time unless we assume that the other players use the same algorithm.

[12]We refer to Bergemann and Morris (2016) for the formal definitions and the discussion of other orderings on information structures.

structure; individual sufficiency captures intuitively when one information structure contains more information than another.

Bergemann and Morris (2016) show that one information structure is more incentive constrained than another if and only if the former is individually sufficient for the latter. That is, the statistical ordering is equivalent to the incentive ordering.

Building on the latter equivalence, we can provide a robustness result for our empirical exercise. Fix a basic game G. Suppose the econometrician knows that players observe at least information structure S, but may observe more—in the sense that they may observe information structure S', for some S' that is individually sufficient for S, the exact S' being unknown to the econometrician. Can the econometrician recover valid bounds on the payoff structure of (G, S') under the ε-ANCR assumption for the repeated play of the misspecified model (G, S)? The answer to this question is positive. Suppose that the sequence of actions, signals and states is generated by play in (G, S')—the true game—and has ε-ANCR. Thus, by Theorem 3, the sequence of empirical distributions converges almost surely to the set of ε-BCE of (G, S'). As S' is individually sufficient for S, by Bergemann and Morris (2016)'s equivalence result, S' is also more incentive constrained than S, and so the set of ε-BCE of (G, S') is contained in the set of ε-BCE of (G, S). But then, the sequence of empirical distributions converges almost surely also to the set of ε-BCE of (G, S). As a result, the bounds on the payoff structure under the ε-ANCR assumption for the misspecified model remain valid for the true model, although they might not be as sharp as those one would obtain under the correct specification of the information structure.

4.4.3 How Long to Equilibrium?

In empirical applications, the question often arises of how many observations one needs to consistently estimate the parameters of interest. In our setting, this concern needs to be paired with an assessment of how long it takes for the empirical distribution to converge to the set of ε-BCE (or ε-BCCE) of (G, S).

The answer to the latter question depends on the particular ε-ANCR (or ε-ANUR) algorithm that players follow in the repetition of the one-shot game (G, S). For instance, if players play a Bayes Nash equilibrium or a BCE of (G, S) in each period n, then the empirical distribution is in the set of BCE of (G, S) since period 1. When players follow regret-based algorithms, Hart and Mansour (2010) show that the rate of convergence to the set of correlated ε-equilibria of the underline complete information game is polynomial in the number of players; given our signal-by-signal extension of regret-based algorithms to incomplete information environments, one can show that the rate of convergence is longer—as each player now needs to accumulate experiences for each of their signal—but remains of the same order when players follow regret-based algorithms for the repeated play of (G, S).

In short, we cannot provide sharp rate-of-convergence results under minimal assumptions on behavior (i.e, without selecting a specific ε-ANCR (or ε-ANUR) algorithm). However, it is worthwhile noting that converge to the set ε-BCE (or ε-BCCE) of (G, S) is, in general, faster that convergence to Nash or Bayes Nash equilibria (or related solution concepts).

4.4.4 An Alternative Empirical Model

An alternative setup is one where the same players take part in a panel of games. Formally, we allow for a set $\mathcal{M} \coloneqq \{1, \ldots, M\}$ of different games, where we write m for a typical game, to be played in every period n. Each game m has a (possibly overlapping) set of players $\mathcal{I}_m$. Game m in period n is characterized by a vector of payoff shifters $x_{m,n}$ in a subset of $\mathbb{R}^{d_x}$, so that

$$u_i\left(x_{m,n}\right) : A \times \Theta \rightarrow \mathbb{R}$$

for all $x_{m,n} \in \mathbb{R}^{d_x}$, $m \in \mathcal{M}$, and $n \in \mathbb{N}$; payoff shifters are assumed to be common knowledge to all players. We maintain the same structure of payoff states as in Section 4.2, so that a basic game m at time n can be denoted as $G_{m,n} \coloneqq \left((A_i, u_i(x_{m,n}))_{i \in \mathcal{I}_m}, \psi\right)$. Timing is as in Assumption 4, so that play now generates a sequence a sequence of actions, signals and states

$$\left((a^{m,n}, t^{m,n}, \theta^{m,n})\right)_{m \in \mathcal{M}, n \in \mathbb{N}}$$

from the repeated play of $(G_{m,n}, S_m)$.

The interesting aspect of this setup is that now players may learn *across* different markets by pooling their experiences. We conjecture that this type of learning, formally defined with an appropriate no-regret condition, could result in restrictions on the empirical distribution of actions in a *cross-section* of markets, i.e., when the empirical distribution of actions is computed, for a given period n, across all values of m. In turn, this may allow the econometrician to identify the underlying structure from just cross-sectional data. Further work is needed to establish results in this direction.

4.5 Concluding Remarks

We propose an estimation strategy that is valid when data on strategic interaction are interpreted as the long-run result of a history of game plays. We model players as interacting repeatedly, playing an incomplete information game, and learning how to play. We remain agnostic on the details of the learning process and we do not require the analyst to determine whether or when the learning phase terminates and equilibrium behavior is observed. Instead, we only impose a minimal behavioral assumption describing an optimality condition for the long-term outcome of players' interaction. In particular, we assume that play satisfies a property of "asymptotic no regret" (ANR). This condition requires that the time average of the counterfactual increase in past payoffs, had different actions been played, becomes approximately zero in the long run. A large class of well-known dynamics satisfies the ANR property: for example, this is the case for regret matching algorithms, calibrated learning, and variants of fictitious play. Moreover, the condition is trivially satisfied if observed outcome are the result of equilibrium play.

We show that, under the ANR assumption, it is possible to partially identify the structural parameters of the underlying static game of incomplete information game. Identification relies on the result that the time average of play that satisfies ANR converges to the set of Bayes correlated ε-equilibria of the underlying static game. Consequently, we can use the limiting model to obtain consistent estimates of the parameters of interest.

In future work, we plan to explore the extensions of our result for general data generating processes, allowing for persistence in the process that determines the evolution of payoff states over time.

4.6 Remaining Proofs

Proof of Theorem 5. [$\Longrightarrow$] Suppose the sequence of actions, signals, and states $\big((a^n, t^n, \theta^n)\big)_{n\in\mathbb{N}}$ from (G, S) has ε-ANUR almost surely for some $\varepsilon \geq 0$. Consider any subsequence $\big(Z^{N_l}\big)_{l\in\mathbb{N}}$ of $\big(Z^N\big)_{N\in\mathbb{N}}$ that converges almost surely to some $\nu \in \Delta(A \times T \times \Theta)$. We need to show that ν is a Bayes coarse correlated ε-equilibrium of (G, S), i.e., that ν is almost surely consistent and ex ante ε-obedient for (G, S).

Consistency. The proof of consistency is the same as for Theorem 3.

Ex ante ε-obedience. To begin, note the following:

$$
\begin{aligned}
\widehat{V}_i(t_i, k; N) - U_i(t_i; N) &= \frac{1}{N}\sum_{n=1}^{N}\big[u_i\big((k, a^n_{-i}), \theta^n\big) - u_i\big((a^n_i, a^n_{-i}), \theta^n\big)\big]\mathbb{1}_{\{t_i\}}(t^n_i) \\
&= \frac{1}{N}\sum_{\theta\in\Theta}\sum_{n=1}^{N}\big[u_i\big((k, a^n_{-i}), \theta^n\big) - u_i(a^n, \theta^n)\big]\mathbb{1}_{\{t_i\}}(t^n_i)\mathbb{1}_{\{\theta\}}(\theta^n) \\
&= \sum_{a, t_{-i}, \theta}\big[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\big]Z^N\big(a, (t_i, t_{-i}), \theta\big).
\end{aligned}
\tag{4.20}
$$

Now pick any $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$. As $\limsup_{N\to\infty}\widehat{R}_i(k; t_i, N) \leq \varepsilon$ almost surely, by definition of $\widehat{R}_i(k; t_i, N)$, we also have $\limsup_{N\to\infty}\big[\widehat{V}_i(t_i, k; N) - U_i(t_i; N)\big] \leq \varepsilon$ almost surely. But then, by (4.20),

$$
\limsup_{N\to\infty}\sum_{a, t_{-i}, \theta}\big[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\big]Z^N\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon \qquad \text{almost surely.} \tag{4.21}
$$

Moreover, on the subsequence $\big(Z^{N_l}\big)_{l\in\mathbb{N}}$ we get

$$
\begin{aligned}
&\lim_{l\to\infty}\sum_{a, t_{-i}, \theta}\big[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\big]Z^N_l\big(a, (t_i, t_{-i}), \theta\big) \\
&= \sum_{a, t_{-i}, \theta}\lim_{l\to\infty}\big[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\big]Z^N_l\big(a, (t_i, t_{-i}), \theta\big) \\
&= \sum_{a, t_{-i}, \theta}\big[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\big]\nu\big(a, (t_i, t_{-i}), \theta\big).
\end{aligned}
\tag{4.22}
$$

Together, (4.21) and (4.22) give

$$
\sum_{a, t_{-i}, \theta}\big[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\big]\nu\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon \qquad \text{almost surely.} \tag{4.23}
$$

As $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$ were arbitrarily chosen, we conclude from (4.23) that ν is almost surely ex ante ε-obedient for (G, S).

[$\Longleftarrow$] Now suppose the sequence of empirical distributions $\left(Z^N\right)_{N\in\mathbb{N}}$ converges almost surely to the set of Bayes coarse correlated ε-equilibria of (G,S) for some $\varepsilon \geq 0$. Pick any $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$. By ex ante ε-obedience,

$$\limsup_{N\to\infty} \sum_{a,t_{-i},\theta} \left[u_i\big((k,a_{-i}),\theta\big) - u_i(a,\theta)\right] Z^N\big(a,(t_i,t_{-i}),\theta\big) \leq \varepsilon \qquad \text{almost surely.} \tag{4.24}$$

By (4.20) and (4.24),

$$\limsup_{N\to\infty} \widehat{V}_i(t_i,k;N) - U_i(t_i;N) \leq \varepsilon \qquad \text{almost surely.}$$

This, by definition of unconditional regret,

$$\limsup_{N\to\infty} \widehat{R}_i(k;t_i,N) \leq \varepsilon \qquad \text{almost surely.}$$

As $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$ were arbitrarily chosen, the desired result follows. ■

Bibliography

Subjectivity and Correlation in Randomized Strategies, author=Robert J. Aumann, journal=Journal of Mathematical Economics, year=1974, volume = 1, pages = 67–96.

Connections: An Introduction to the Economics of Networks, author=Sanjeev Goyal, isbn=9780691126500, lccn=2007927826, year=2007, publisher=Princeton University Press.

Robust Predictions in Games with Incomplete Information, author=Dirk Bergemann and Stephen Morris. *Econometrica*, 11(2):487—-522, 2013.

Communication Complexity of Approximate Nash Equilibria, author=Yakov Babichenko and Aviad Rubinstein. *STOC 2017*, pages 878–889, 2017.

D. Abreu and F. Gul. Bargaining and Reputation. *Econometrica*, 68(1):85–117, 2000.

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

D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi. Spread of (Mis)information in Social Networks. *Games and Economic Behavior*, 70:194–227, 2010.

D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar. Bayesian Learning in Social Networks. *The Review of Economic Studies*, 78(4):1201–1236, 2011.

D. Acemoglu, K. Bimpikis, and A. Ozdaglar. Dynamics of Information Exchange in Endogenous Social Networks. *Theoretical Economics*, 9:41–97, 2014.

D. Ackerberg, C. L. Benkard, S. Berry, and A. Pakes. Econometric Tools for Analyzing Market Outcomes. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics*. Elsevier, 2007.

A. R. Admati and M. Perry. Strategic Delay in Bargaining. *The Review of Economic Studies*, 54(3): 345–364, 1987.

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

A. Ambrus and S. E. Lu. A Continuous-Time Model of Multilateral Bargaining. *American Economic Journal: Microeconomics*, 7(1):208–49, 2015.

I. Arieli and M. Mueller-Frank. Multi-Dimensional Social Learning. *Review of Economic Studies (forthcoming)*, 2018.

M. Armstrong. Ordered Consumer Search. *Working Paper*, 2016.

A. E. Atakan and M. Ekmekci. Bargaining and Reputation in Search Markets. *The Review of Economic Studies*, 81(1):1–29, 2014.

R. J. Aumann. Correlated Equilibrium as an Expansion of Bayesian Rationality. *Econometrica*, 55:1–18, 1987.

L. M. Ausubel and R. J. Deneckere. Reputation in Bargaining and Durable Goods Monopoly. *Econometrica*, 57(3):511–531, 1989.

L. M. Ausubel, P. Cramton, and R. J. Deneckere. Bargaining with Incomplete Information. In R. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, pages 1897–1945. 2002.

C. Avery and P. B. Zemsky. Option Values and Bargaining Delays. *Games and Economic Behavior*, 7: 139–153, 1994.

M. Bagnoli, S. W. Salant, and J. E. Swierzbinski. Durable-Goods Monopoly with Discrete Demand. *Journal of Political Economy*, 97(6):1459–1478, 1989.

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

A. V. Banerjee. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107(3):797–817, Aug. 1992.

A. V. Banerjee. The Economics of Rumors. *The Review of Economic Studies*, 60(2):309–327, 1993.

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

L. Benkard. A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft. *The Review of Economic Studies*, 71(3):581–611, 2004.

D. Bergemann and S. Morris. Correlated Equilibrium in Games with Incomplete Information. *Cowles Foundation Discussion Paper No. 1822*, 2011.

D. Bergemann and S. Morris. Bayes Correlated Equilibrium and the Comparison of Information Structures in Games. *Theoretical Economics*, 11(2):487—-522, 2016.

D. Bergemann and J. Välimäki. Bandit Problems. In S. N. Durlauf and L. E. Blume, editors, *The New Palgrave Dictionary of Economics*, The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke, 2008.

J. Bergin and W. B. MacLeod. Continuous Time Repeated Games. *International Economic Review*, 34 (1):21–37, 1993.

S. Berry. Estimation of a Model of Entry in the Airline Industry. *Econometrica*, 60:889–917, 1992.

A. R. Biehl. Durable-Good Monopoly with Stochastic Values. *The RAND Journal of Economics*, 32: 565–577, 2001.

S. Bikhchandani, D. Hirshleifer, and I. Welch. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5):992–1026, 1992.

D. Blackwell. Controlled Random Walks. *Proceedings of the International Congress of Mathematicians*, 3:336–338, 1954.

D. Blackwell. An Analog of the Minmax Theorem for Vector Payoffs. *Pacific Journal of Mathematics*, 6: 1–8, 1956.

S. Board and M. Pycia. Outside Options and the Failure of the Coase Conjecture. *The American Economic Review*, 104(2):656–671, 2014.

P. Bolton and C. Harris. Strategic Experimentation. *Econometrica*, 67:349–374, 1999a.

P. Bolton and C. Harris. Strategic Experimentation. *Econometrica*, 67:349–374, 1999b.

E. W. Bond and L. Samuelson. Durable Goods, Market Structure and the Incentives to Innovate. *Economica*, 54(213):57–67, 1987.

T. Börgers and R. Sarin. Learning Through Reinforcement and Replicator Dynamics. *Journal of Economic Theory*, 77:1—-14, 1997.

T. Börgers and R. Sarin. Naive Reinforcement Learning With Endogenous Aspirations. *Internation Economic Review*, 41:921—-950, 2000.

J. I. Bulow. Durable Goods Monopolists. *Journal of Political Economy*, 90:1062–1076, 1982.

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

D. A. Butz. Durable-Good Monopoly and Best-Price Provision. *The American Economic Review*, 80(2): 314–32, 1990.

A. Cahn. General Procedures Leading to Correlated Equilibria. *International Journal of Game Theory*, 33:21—-40, 2004.

B. Celen and K. Hyndman. Social Learning Through Endogenous Information Acquisition: An Experiment. *Management Science*, 158:1525–1548, 2012.

B. Celen and S. Kariv. Observational Learning under Imperfect Information. *Games and Economic Behavior*, 47:72–86, 2004.

C. P. Chamley. *Rational Herds: Economic Models of Social Learning*. Cambridge University Press, 2004. ISBN 9780521530927.

D. Chang. Bargaining with Arriving Sellers. *Working Paper*, 2015.

K. Chatterjee and L. Samuelson. Bargaining with Two-Sided Incomplete Information: An Infinite Horizon Model with Alternating Offers. *The Review of Economic Studies*, 54(2):175–192, 1987.

I. Chaves Villamizar. Privacy in Bargaining: The Case of Endogenous Entry. *Working Paper*, 2018.

Y.-K. Che and J. Hörner. Recommender Systems as Mechanisms for Social Learning. *The Quarterly Journal of Economics*, 133:871–925, 2018.

I.-K. Cho. Uncertainty and Delay in Bargaining. *The Review of Economic Studies*, 57(4):575–595, 1990.

J. P. Choi. Network Externality, Compatibility Choice, and Planned Obsolescence. *The Journal of Industrial Economics*, 42(2):167–182, 1994.

F. Ciliberto and E. Tamer. Market Structure and Multiple Equilibria in Airline Markets. *Econometrica*, 77:1791–1828, 2009.

R. H. Coase. Durability and Monopoly. *Journal of Law and Economics*, 15(1):143–149, 1972.

O. Compte and P. Jehiel. On the Role of Outside Options in Bargaining with Obstinate Parties. *Econometrica*, 70(4):1447–1517, 2002.

O. Compte and P. Jehiel. Gradualism in Bargaining and Contribution Games. *The Review of Economic Studies*, 71(4):975–1000, 2004.

P. C. Cramton. Bargaining with Incomplete Information: An Infinite-Horizon with Two-Sided Uncertainty. *The Review of Economic Studies*, 51(4):579–593, 1984.

B. Daley and B. Green. Bargaining and News. *Working Paper*, 2018.

R. Deb. Intertemporal Price Discrimination with Stochastic Values. *Working Paper*, 2014.

M. H. DeGroot. Reaching a Consensus. *Journal of the American Statistical Association*, 69:118–121, 1974.

P. M. DeMarzo, D. Vayanos, and J. Zwiebel. Persuasion Bias, Social Influence, and Unidimensional Opinions. *The Quarterly Journal of Economics*, 118:909–968, 2003.

R. Deneckere and M.-Y. Liang. Bargaining with Interdependent Values. *Econometrica*, 74(5):1309–1364, 2006.

U. Doraszelski, G. Lewis, and A. Pakes. Just Starting Out: Learning and Equilibrium in a New Market. *The American Economic Review*, 108(3):565–615, 2018.

L. Doval. Whether or Not to Open Pandora's Box. *Journal of Economic Theorys*, 175:127–158, 2018.

G. Ellison and D. Fudenberg. Word-of-Mouth Communication and Social Learning. *The Quarterly Journal of Economics*, 110:93–125, 1995.

L. Epstein, H. Kaido, and K. Seo. Robust Confidence Regions for Incomplete Models. *Econometrica*, 84 (5):1799–1838, 2016.

I. Erev and A. E. Roth. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategies. *American Economic Review*, 88:848—-881, 1998.

R. Ericson and A. Pakes. Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies*, 62(1):53–82, 1995.

R. Evans. Sequential Bargaining with Correlated Values. *The Review of Economic Studies*, 56(4):499—-510, 1898.

Y. Feinberg and A. Skrzypacz. Uncertainty About Uncertainty and Delay in Bargaining. *Econometrica*, 73(1):69–91, 2005.

D. P. Foster. Proof of Calibration via Blackwell's Approachability Theorem. *Games and Economic Behavior*, 29:74–78, 1997.

D. P. Foster and R. V. Vohra. A Randomization Rule for Selecting Forecasts. *Operations Research*, 41: 704–709, 1993.

D. P. Foster and R. V. Vohra. Calibrated Learning and Correlated Equilibrium. *Games and Economic Behavior*, 21:40–55, 1997.

D. P. Foster and R. V. Vohra. Asymptotic Calibration. *Biometrika*, 85:379–390, 1998.

D. P. Foster and R. V. Vohra. Regret in the On-Line Decision Problem. *Games and Economic Behavior*, 29:7–35, 1999.

C. Freshtman and D. J. Seidmann. Deadline Effects and Inefficient Delay in Bargaining with Endogenous Commitment. *Journal of Economic Theory*, 60:306–321, 1993.

Y. Freund and R. E. Schapire. Adaptive Game Playing Using Multiplicative Weights. *Games and Economic Behavior*, 29:79–103, 1999.

W. Fuchs and A. Skrzypacz. Bargaining with Arrival of New Traders. *The American Economic Review*, 100(3):802–836, 2010.

W. Fuchs and A. Skrzypacz. Bridging the Gap: Bargaining with Interdependent Values. *Journal of Economic Theory*, 148:1226–1236, 2013a.

W. Fuchs and A. Skrzypacz. Bargaining with Deadlines and Private Information. *American Economic Journal: Microeconomics*, 5:219–243, 2013b.

D. Fudenberg and D. K. Levine. Consistency and Cautious Fictitious Play. *Journal of Economic Dynamics and Control*, 19:1065–1090, 1995.

D. Fudenberg and D. K. Levine. The Theory of Learning in Games. *MIT Press*, 1998.

D. Fudenberg and D. K. Levine. Conditional Universal Consistency. *Games and Economic Behavior*, 29:104–130, 1999a.

D. Fudenberg and D. K. Levine. An Easier Way to Calibrate. *Games and Economic Behavior*, 29:131–137, 1999b.

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.

D. Fudenberg and J. Tirole. Upgrades, Tradeins, and Buybacks. *The RAND Journal of Economics*, 29(2):235–258, 1998.

D. Fudenberg, D. K. Levine, and J. Tirole. Infinite-Horizon Models of Bargaining with One-Sided Incomplete Information. In A. E. Roth, editor, *Game-Theoretic Models of Bargaining*, pages 73–98. Cambridge University Press, 1985.

D. Fudenberg, D. K. Levine, and J. Tirole. Incomplete Information Bargaining with Outside Opportunities. *The Quarterly Journal of Economics*, 102(1):37–50, 1987.

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

D. F. Garrett. Intertemporal Price Discrimination: Dynamic Arrivals and Changing Values. *The American Economic Review*, 106:3275–3299, 2016.

B. Golub and M. Jackson. Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*, 2:112–149, 2010.

B. Golub and M. Jackson. How Homophily Affects the Speed of Learning and Best-Response Dynamics. *The Quarterly Journal of Economics*, 127:1287–1338, 2012.

B. Golub and E. Sadler. Learning in Social Networks. In Y. Bramoullé, A. Galeotti, and B. Rogers, editors, *The Oxford Handbook of the Economics of Network*. Oxford University Press, 2016.

S. Goyal. Learning in Networks. In J. Benhabib, A. Bisin, and M. O. Jackson, editors, *Handbook of Social Economics*. North-Holland, 2011.

F. Gul, H. Sonnenschein, and R. Wilson. Foundations of Dynamic Monopoly and the Coase Conjecture. *Journal of Economic Theory*, 39(1):155–190, 1986.

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

J. Hannan. Approximation to Bayes Risk in Repeated Plays. In M. Dresher, A. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games, Vol. 3*. Princeton University Press, 1957.

J. M. Harrison. *Brownian Models of Performance and Control*. Cambridge University Press, 2013.

S. Hart. Adaptive Heuristics. *Econometrica*, 73:1401–1430, 2005.

S. Hart and Y. Mansour. How Long to Equilibrium? The Communication Complexity of Uncoupled Equilibrium Procedures. *Games and Economic Behavior*, 69:107–126, 2010.

S. Hart and A. Mas-Colell. A Simple Adaptive Procedure Leading to Correlated Equilibrium. *Econometrica*, 68:1127–1150, 2000.

S. Hart and A. Mas-Colell. A General Class of Adaptive Strategies. *Journal of Economic Theory*, 98: 26—-54, 2001a.

S. Hart and A. Mas-Colell. A Reinforcement Procedure Leading to Correlated Equilibrium. In G. Debreu, W. Neuefeind, and W. Trockel, editors, *Economic Essays: A Festschrift for Werner Hildenbrand*. Springer, 2001b.

S. Hart and A. Mas-Colell. Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics. *World Scientific Series in Economic Theory*, 2013.

J. Hartline, V. Syrgkanis, and E. Tardos. No-Regret Learning in Bayesian Games. *NIPS 2015*, 2015.

P. Heidhues, S. Rady, and P. Strack. Strategic Experimentation with Private Payoffs. *Journal of Economic Theory*, 159:531–551, 2015.

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

C. Huang and F. Li. Bargaining while Learning About New Arrivals. *Working Paper*, 2013.

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

I. Hwang. A Theory of Bargaining Deadlock. *Games and Economic Behavior*, 109:501–522, 2018.

I. Hwang and F. Li. Transparency of Outside Options in Bargaining. *Journal of Economic Theory*, 167: 116–147, 2017.

R. Inderst. Dynamic Bilateral Bargaining Under Private Information with a Sequence of New Buyers. *Review of Economic Dynamics*, 11:220–236, 2008.

Y. Ishii, A. Öry, and A. H. Vigier. Competing for Talent: Wage Bargaining, Frictions, and Unraveling. *Working Paper*, 2018.

M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008. ISBN 0691134405, 9780691134406.

A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi. Non-Bayesian Social Learning. *Games and Economic Behavior*, 76:210–225, 2012.

P. Jehiel and B. Moldovanu. Negative Externalities May Cause Delay in Negotiation. *Econometrica*, 63 (6):1321–1335, 1995.

M. Jofre-Bonet and M. Pesendorfer. Estimation of a Dynamic Auction Game. *Econometrica*, 71(5): 1443–1489, 2003.

C. Kahn. The Durable Goods Monopolist and Consistency with Increasing Costs. *Econometrica*, 54(2): 275–294, 1986.

S. M. Kakade and D. P. Foster. Deterministic Calibration and Nash Equilibrium. *Journal of Computer and System Science*, 74:115–130, 2008.

G. Keller and S. Rady. Strategic Experimentation with Poisson Bandits. *Theoretical Economics*, pages 275–311, 2010.

G. Keller, S. Rady, and M. Cripps. Strategic Experimentation with Exponential Bandits. *Econometrica*, 73:39–68, 2005a.

G. Keller, S. Rady, and M. Cripps. Strategic Experimentation with Exponential Bandits. *Econometrica*, 73:39–68, 2005b.

I. Kremer, Y. Mansour, and M. Perry. Implementing the "Wisdom of the Crowd". *Journal of Political Economy*, 122:988–1012, 2014.

K. Kultti and P. Miettinen. Herding with Costly Information. *International Game Theory Review*, 8: 21–31, 2006.

K. Kultti and P. Miettinen. Herding with Costly Observation. *The B.E. Journal of Theoretical Economics*, 7, 1 (Topics), Article 28, 2007.

I. H. Lee and J. Lee. A Theory of Economic Obsolescence. *Econometrica*, 46(3):383—-401, 1998.

J. Lee and Q. Liu. Gambling Reputation: Repeated Bargaining with Outside Options. *Econometrica*, 81 (4):1601–1672, 2013.

R. S. Lee and A. Pakes. Multiple Equilibria and Selection by Learning in an Applied Setting. *Economics Letters*, 104(1):13–16, 2009.

D. A. Levinthal and D. Purohit. Durable Goods and Product Obsolescence. *Marketing Science*, 8:35–56, 1989.

I. Lobel and E. Sadler. Information Diffusion in Networks through Social Learning. *Theoretical Economics*, 10(3):807–851, 2015.

I. Lobel and E. Sadler. Preferences, Homophily, and Social Learning. *Operations Research*, 64(3):564–584, 2016.

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

L. Magnolfi and C. Roncoroni. Estimation of Discrete Games with Weak Assumptions on Information. *Working Paper*, 2017.

Y. Mansour, A. Slivkins, and V. Syrgkanis. Bayesian Incentive-Compatible Bandit Exploration. *16th ACM Conference on Economics and Computation, EC 2015*, 2015.

Y. Mansour, A. Slivkins, V. Syrgkanis, and Z. S. Wu. Bayesian Exploration: Incentivizing Exploration in Bayesian Games. *17th ACM Conference on Economics and Computation, EC 2016*, 2016.

P. R. McAfee and T. Wiseman. Capacity Choice Counters the Coase Conjecture. *The Review of Economic Studies*, 75:317–331, 2008.

A. Merlo and C. Wilson. A Stochastic Model of Sequential Bargaining with Complete Information. *Econometrica*, 62:371–399, 1995.

A. Merlo and C. Wilson. Efficient Delays in a Stochastic Model of Bargaining. *Economic Theory*, 11: 39–55, 1998.

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

P. Molavi, A. Tahbaz-Salehi, and A. Jadbabaie. A Theory of Non-Bayesian Social Learning. *Econometrica*, 86:445–490, 2018.

J. Montez. Inefficient Sales Delay by a Durable-Good Monopoly Facing a Finite Number of Buyers. *The RAND Journal of Economics*, 44:425–437, 2013.

G. Moscarini and L. Smith. The Optimal Level of Experimentation. *Econometrica*, 69:1629–1644, 2001.

E. Mossel, A. Sly, and O. Tamuz. Strategic Learning and the Topology of Social Networks. *Econometrica*, 83:1755–1794, 2015.

H. Moulin and J.-P. Vial. Strategically Zero-Sum games: The Class of Games whose Completely Mixed Equilibria Cannot Be Improved Upon. *Internation Journal of Game Theory*, 7:201–221, 1978.

M. Mueller-Frank. A General Framework for Rational Learning in Social Networks. *Theoretical Economics*, 8:1–40, 2013.

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

F. Nava and P. Schiraldi. Differentiated Durable Goods Monopoly: A Robust Coase Conjecture. *Working Paper*, 2018.

S. Nei. Frictions to Information Aggregation in Social Learning Environments. *Working Paper*, 2016.

D. Nekipelov, V. Syrgkanis, and E. Tardos. Econometrics for Learning Agents. *EC 2015*, 2015.

J. Ortner. A Continuous Time Model of Bilateral Bargaining. *Working Paper*, 2016.

J. Ortner. Durable Goods Monopoly with Stochastic Costs. *Theoretical Economics*, 12(2):817—-861, 2017.

A. Pakes. Methodological Issues in Analyzing Market Dynamics. In F. Thuijsman and F. Wagener, editors, *Advances in Dynamic and Evolutionary Games: Theory, Applications, and Numerical Methods*. Springer International Publishing, 2016.

Y. Papanastasiou, K. Bimpikis, and N. Savva. Crowdsourcing Exploration. *Management Science*, 64: 1727–1746, 2018.

J. Perego and S. Yuksel. Searching for Information and the Diffusion of Knowledge. *Working Paper*, 2016.

M. Perry and P. J. Reny. A Non-cooperative Bargaining Model with Strategically Timed Offers. *Journal of Economic Theory*, 59(1):50–77, 1993.

A. D. Polyanin and V. F. Zaitsev. *Handbook of Exact Solutions for Ordinary Differential Equations*. Chapman and Hall, 2003.

D. Rosenberg and N. Vieille. On the Efficiency of Social Learning. *Working Paper*, 2017.

A. E. Roth and I. Erev. Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior*, 8:164—-212, 1995.

S. P. Ryan. The Costs of Environmental Regulation in a Concentrated Industry. *Econometrica*, 80(3): 1019–1061, 2012.

E. Sadler. Bounded Rationality and the Network Economist's Dilemma. *Working Paper*, 2014.

E. Sadler. Innovation Adoption and Collective Experimentation. *Working Paper*, 2017.

M. Salish. Strategic Experimentation in Networks. *Working Paper*, 2017.

L. Samuelson. Disagreement in Markets with Matching and Bargaining. *The Review of Economic Studies*, 59:177–185, 1992.

L. K. Simon and M. B. Stinchcombe. Extensive Form Games in Continuous Time: Pure Strategies. *Econometrica*, 57(5):1171–1214, 1989.

L. Smith and P. N. Sørensen. Pathological Outcomes of Observational Learning. *Econometrica*, 68(2): 371–398, 2000.

L. Smith and P. N. Sørensen. Rational Social Learning with Random Sampling. *Working Paper*, 2014.

L. Smith, P. N. Sørensen, and J. Tian. Informational Herding, Optimal Experimentation, and Contrarianism. *Working Paper*, 2017.

J. Sobel. Durable Good Monopoly with Entry of New Consumers. *Econometrica*, 59:1455–1485, 1991.

Y. Song. Social Learning with Endogenous Observation. *Journal of Economic Theory*, 166:324–333, 2016.

N. Stokey. Rational Expectations and Durable Goods Pricing. *Bell Journal of Economics*, 12(1):112–128, 1981.

J. Sákovics. Delay in Bargaining Games with Complete Information. *Journal of Economic Theory*, 59 (1):78–95, 1993.

E. Tamer. Incomplete Simultaneous Discrete Response Model with Multiple Equilibria. *The Review of Economic Studies*, 70(1):147–165, 2003.

D. R. Vincent. Bargaining with Common Values. *Journal of Economic Theory*, 48(1):47—-62, 1989.

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

N.-H. von der Fehr and K.-U. Kuhn. Coase versus Pacman: Who Eats Whom in the Durable-Goods Monopoly? *Journal of Political Economy*, 103(4):785–812, 1995.

A. Wald. Foundations of a General Theory of Sequential Decision Functions. *Econometrica*, 15(4): 279–313, 1947.

M. Waldman. A New Perspective on Planned Obsolescence. *The Quarterly Journal of Economics*, 108 (1):273–283, 1993.

M. Waldman. Planned Obsolescence and the R&D Decision. *The RAND Journal of Economics*, 108(27): 583–595, 1999.

M. L. Weitzman. Optimal Search for the Best Alternative. *Econometrica*, 43(3):641–654, 1979.

M. Yildiz. Waiting to Persuade. *The Quarterly Journal of Economics*, 119(1):223–248, 2004.

H. P. Young. *Strategic Learning and Its Limits.* Oxford University Press, 2004.

Curriculum Vitae

Niccolò Lomys

— 2012–2018	*University of Mannheim* Ph.D. Student in Economics
— 2015–2016	*Yale University* Visiting Assistant in Research
— 2011–2012	*Barcelona GSE–Universitat Pompeu Fabra* M.Sc. in Economics
— 2009–2011	*Università degli Studi di Firenze* *Laurea Magistrale* in Statistics
— 2003–2009	*Università degli Studi di Firenze* *Laurea* in Business Economics

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich die Dissertation selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Mannheim, den 01.08.2018 Niccolò Lomys