Essays on Rational Social Learning

Thesis by
Wanying Huang

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
Defended May 7, 2024

Wanying Huang
ORCID: 0009-0000-0413-9171

ACKNOWLEDGEMENTS

The last five years at Caltech have been the most stimulating, challenging, and satisfying period of my life so far. So much of this has been because of the people around me.

First and foremost, I am privileged to have Omer Tamuz as my advisor. He is one of the most genuine and excellent teachers I have ever met, who has taught me much of what I know about economic research. I often wonder how he so effortlessly juggles research, family, hobbies, traveling, and many other responsibilities while still always having time for his students. He is my biggest inspiration. I have also learned much from Federico Echenique, Luciano Pomatto, Philipp Strack, Laura Doval, Yi Xin, Kirby Nielsen, Marina Agranov, and many others who have all collectively enriched and influenced my academic journey. Caltech is a special place to me and these are the people who make it so.

I am lucky to have shared my graduate school years with Wade Hann-Caruthers, Jeffrey Zeidel, Sumit Goel, Joanna Huey, Aldo Lucia, Zhuofang Li, Fan Wu, Po-Hsuan Lin, Alena Buinskaya, Polina Detkova, Shunto Kobayashi, Ke Shi, Peter Doe, and others. I cherish all the memories I have made with them. Wade, thank you for being my academic brother; you are as good as a real brother, if I ever had one. Jeff, I appreciate you for getting through a marathon and a snowstorm with me, which was exactly what I needed at the time. Aldo, whom I have failed to persuade to drink California tap water, I truly enjoy our time together both inside and outside of the office. I have been spoiled to have Brenden Eum, Kimberly Golubeva, and Caltech Alpine Club friends around me, who all share the same passion for adventure and the great outdoors.

I am indebted to my former advisors at the University of Queensland, Priscilla Man, Jeffrey Kline, and Andy McLennan, who formally introduced me to economic theory and encouraged me to pursue graduate studies in the United States.

I thank my family: my mother Wei and father Qingyuan, whose love and support transcend the continents between us. Last but not least, Chen—my partner in life and "partner in crime"—who has moved mountains to join me in the States and stuck by me through thick and thin, I could not have done this without you by my side.

ABSTRACT

This dissertation contains three essays, each contributing to the study of social learning among rational agents in various contexts.

In Chapter 1, I study whether individuals can learn the informativeness of their information technology through social learning. Building on the classic sequential social learning model, I introduce the possibility that a common source is completely uninformative. I then define *asymptotic learning* as the situation in which an outsider, who observes the actions of all agents, eventually distinguishes between uninformative and informative sources. I show that asymptotic learning in this setting is not guaranteed; it depends crucially on the relative tail distributions of private beliefs induced by uninformative and informative signals. Furthermore, I identify the phenomenon of *perpetual disagreement* as the cause of learning and provide a characterization of learning in the canonical Gaussian environment.

In Chapter 2, co-authored with Omer Tamuz and Philipp Strack, we study the asymptotic rate at which the probability of a group of long-lived, rational agents in a social network taking the correct action converges to one. In every period, after observing the past actions of his neighbors, each agent receives a private signal, and chooses an action whose payoff depends only on the state. Since equilibrium actions depend on higher-order beliefs, characterizing agents' behavior becomes difficult. Nevertheless, we show that the rate of learning in any equilibrium is bounded from above by a constant, regardless of the size and shape of the network, the utility function, and the patience of the agents. This bound only depends on the private signal distribution.

In Chapter 3, I study how fads emerge from social learning in a changing environment. I consider a simple sequential learning model in which rational agents arrive in order, each acting only once, and the underlying unknown state is constantly evolving. Each agent receives a private signal, observes all past actions of others, and chooses an action to match the current state. Because the state changes over time, cascades cannot last forever, and actions also fluctuate. I show that despite the rise of temporary information cascades, in the long run, actions change more often than the state. This provides a theoretical foundation for faddish behavior in which people often change their actions more frequently than necessary.

PUBLISHED CONTENT AND CONTRIBUTIONS

Wanying Huang, Philipp Strack, and Omer Tamuz. Learning in repeated interactions on networks. *Econometrica*, 92(1):1–27, 2024. doi: 10.3982/ECTA20806.
All authors contributed equally at all stages of this project.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

INTRODUCTION

It is human nature to observe, imitate, and learn from one another. From everyday decisions like choosing a restaurant to significant financial endeavors such as investing in the stock market, and even in critical domains as in organ transplant procedures, the behavior of social learning permeates and influences many important economic activities. The central theme of this dissertation is to explore the role of social learning in aggregating information in various contexts. It ranges from introducing uncertainty in information sources to examining the dynamics of repeated interactions between agents within networks, as well as the learning processes in evolving environments. These theoretical inquiries yield valuable insights into phenomena such as fads, misinformation, and information cascades.

This dissertation consists of three essays, each contributing to the study of social learning, with the second chapter also contributing to the field of network economics. Chapter 1 examines a sequential social learning model that introduces the possibility of an uninformative source, investigating how and when society can differentiate between uninformative and informative sources. Chapter 2, co-authored with Philipp Strack and Omer Tamuz, explores the efficiency of information aggregation among long-lived agents who repeatedly interact on a general observational network. Finally, Chapter 3 investigates the volatility in behavior relative to the volatility in the state within a dynamic social learning environment.

In Chapter 1, I introduce an additional dimension of uncertainty in a sequential social learning model. Motivated by the recent surge in new information technologies and growing concerns about their accuracy and reliability, I examine whether society can eventually learn the informativeness of a source through social learning. An example of such a potentially uninformative source is a novel AI algorithm that provides a private recommendation to each investor. Rational investors arrive sequentially and each makes a decision based on all past investment decisions of others and their own private recommendation. In contrast to a regime where all private information is observable—thus ensuring the achievement of asymptotic learning about the source's informativeness—my main result demonstrates that when only past decisions are observable, learning is not guaranteed. Specifically, it depends crucially on the relative tail distribution of private beliefs induced by uninformative and informative signals. That is, learning holds when the uninformative signals have fatter tails than the informative ones, but fails when they have thinner tails. For the canonical case of Gaussian signals, I fully characterize the conditions for learning. The key insight behind the main result is that the accumulation of behavioral anomalies, such as an action switch after an extended period of identical actions, helps society distinguish the uninformative source from the informative one.

In Chapter 2, we study how efficiently agents can learn from repeated interactions with each other. Repeated interactions can, for example, describe the exchange of opinions and information among friends on social media, and are especially prominent for agents connected via a social network. Our main result shows that information aggregation is inefficient in the sense that the speed of learning stays bounded below by a constant even in large networks, where efficient aggregation of all agents' private information would lead to arbitrarily fast learning. Methodologically, we introduce new techniques that allow us to relax commonly made assumptions and study general networks with forward-looking, Bayesian agents who interact repeatedly. The mechanism behind our main result is that agents cannot learn too fast because fast individual learning leads to rapid growth in social information. This, in turn, causes agents to ignore their private signals, thereby blocking any inflow of information.

In Chapter 3, I investigate how fads can emerge from social learning in a dynamic environment. A changing underlying state can represent, for example, the development of an economy, technological advancements, or shifts in cultural preferences. My main result shows that individuals who learn from observing others exhibit more volatile behavior than the underlying state itself, resulting in fads. As an example, consider a situation in which the relative quality between two restaurants changes over time. Then, in the long run, customers who wish to dine at the restaurant currently of higher quality would switch between restaurants more frequently than the actual changes in quality warrant. The idea behind these excess action changes is that despite the rise of temporary information cascades, agents tend to overreact to environmental changes, leading to long-run faddish behavior. This provides a theoretical foundation for many commonly observed volatile economic phenomena, such as speculative bubbles, financial fads, and boom-and-bust business cycles.

Chapter 1

LEARNING ABOUT INFORMATIVENESS

1.1 Introduction

Social learning plays a vital role in the dissemination and aggregation of information. The behavior of others reflects their private knowledge about an unknown state of the world, and so by observing others, individuals can acquire additional information, enabling them to make better-informed decisions. A key assumption in most existing social learning models is the presence of an informative source that provides a useful private signal to each individual. In this paper, we explore how the possibility that the source is uninformative interferes with learning, and study the conditions under which individuals can eventually distinguish an uninformative source from an informative one. This question is particularly relevant today due to the proliferation of novel information technologies, raising concerns about the accuracy and credibility of the information they provide.[1]

Formally, we introduce uncertainty regarding the informativeness of the source into the classic sequential social learning model [Banerjee, 1992, Bikhchandani, Hirshleifer, and Welch, 1992]. As in the usual setting, a sequence of short-lived agents arrives in order, each acting once by choosing an action to match an unknown payoff-relevant state that can be either good or bad. Before making their decisions, each agent observes the past actions of her predecessors and receives a private signal from a common source of information. However, unlike in the usual setting, there is uncertainty surrounding this common information source. In particular, we assume that this source can be either informative, generating private signals that are independent and identically distributed (i.i.d.) conditioned on the payoff-relevant state, or uninformative, producing private signals that are i.i.d. but independent of the payoff-relevant state. Both the payoff-relevant state and the informativeness of the source are realized independently at the outset and are assumed to be fixed throughout.

If an outside observer, who aims to evaluate the informativeness of the source, were

[1] For example, the recent surge in the popularity of ChatGPT, a generative AI language model, has led to its widespread usage by a wide range of individuals, including laypeople, artists, and college students. Despite the model's disclaimer stating that "ChatGPT may produce inaccurate information about people, places, or facts," its adoption continues to grow.

to have access to the private signals received by the agents, he would gradually accumulate empirical evidence about the source, and thus eventually learn its informativeness. However, when only the history of past actions is observable, his inference problem becomes more challenging—not only because there is less information available, but also because these past actions are correlated with each other. This correlation stems from social learning behavior, where agents' decisions are influenced by the inferences they draw from observing others' actions. We say that *asymptotic learning* holds if the belief of the outside observer about the source's informativeness converges to the truth, i.e., it converges almost surely to one when the source is informative and to zero when it is uninformative. The questions we aim to address are: Can asymptotic learning be achieved and if so, under what conditions? Furthermore, what are the behavioral implications of asymptotic learning?

Our question of learning about the informativeness of the source is new within the realm of social learning. In the absence of information uncertainty, a common inquiry in this literature is whether or not agents engaged in social learning behavior ultimately choose the *correct* action, i.e., the action that matches the payoff-relevant state. However, in the presence of information uncertainty, we show that agents may still eventually choose the correct action, but can remain uncertain regarding its correctness. Therefore, to establish societal confidence in their decision-making, it is important to understand the conditions for achieving asymptotic learning.

We consider *unbounded signals* [Smith and Sørensen, 2000] under which the agent's private belief induced by an informative signal can be arbitrarily strong. We focus on this setting since otherwise, asymptotic learning can be easily precluded by the agents' lack of response to their private signals.[2] Our main result (Theorem 1) shows that even with unbounded signals, achieving asymptotic learning is far from guaranteed. In fact, the determining factor of asymptotic learning lies in the tail distributions of the private beliefs of the agents. In particular, it depends on whether the belief distribution induced by uninformative signals has *fatter* or *thinner* tails compared to that induced by informative signals. More specifically, we show that asymptotic learning holds when uninformative signals have fatter tails than informative signals, but fails when uninformative signals have thinner tails.

[2] The phenomenon in which agents follow the actions of their predecessors regardless of their private signals is known as *information cascades*. In such cascades, since agents do not respond to their private signals, their actions no longer reveal any private information, and thus information stops accumulating. As shown in Bikhchandani, Hirshleifer, and Welch [1992], information cascades occur almost surely when signals are bounded and the set of possible signal values is finite.

For example, consider an informative source that generates Gaussian signals with unit variance and mean +1 if the payoff-relevant state is good and mean -1 if the state is bad. Meanwhile, the uninformative source generates Gaussian signals with mean 0, independent of the state. If the uninformative source generates signals with a variance strictly greater than one, then the uninformative signals have fatter tails, and thus asymptotic learning holds. In contrast, when the uninformative signals have variance strictly less than one, they exhibit thinner tails, and so asymptotic learning fails.

As another illustration of the main result, consider the case where the informative signals have the same distributions as before, but the uninformative signals are chosen uniformly from the bounded interval $[-\varepsilon, \varepsilon]$ for some $\varepsilon > 0$. In this case, the distribution of private beliefs induced by these uninformative signals also has bounded support. Consequently, it can be viewed as having *extremely* thinner tails compared to those of informative Gaussian signals. Hence, Theorem 1 implies that asymptotic learning fails. Nevertheless, under such an informative source, almost all agents individually learn its informativeness: Once they receive a signal outside the support $[-\varepsilon, \varepsilon]$, which is highly probable, they infer that it can only come from the normal distribution, indicating that the source must be informative. However, an outside observer who only observes the agents' actions is unable to determine the informativeness of the source.

The mechanism behind the main result is as follows. First, in our model, despite information uncertainty, agents always treat signals as informative. So, when the source is indeed informative and generates unbounded signals, agents will eventually reach a consensus on the correct action. Now, suppose that the source is uninformative and generates signals with thinner tails. In this case, it is unlikely that agents will receive signals that are extreme enough to break a consensus, so they usually mimic their predecessors. Consequently, an outside observer who only observes agents' actions does not learn that the source is uninformative, as a consensus will be reached under both uninformative and informative sources. In contrast, suppose that the source is uninformative but generates signals with fatter tails. In this case, extreme signals are more likely, allowing agents to break a consensus; in fact, both actions will be taken infinitely often, so no consensus is ever reached. Hence, an outside observer who observes an infinite number of action switches learns that the source is uninformative.

For some private belief distributions, their relative tail thickness is neither thinner nor fatter. For these, we show that the same holds: Asymptotic learning holds if and only if conditioned on the source being uninformative, agents never reach a consensus (Proposition 1). In terms of behavioral implications, when the source is informative, as mentioned above, agents eventually reach a consensus on the correct action, regardless of the achievement of asymptotic learning. Nevertheless, we show that in this case agents are sure that they are taking the correct action if and only if asymptotic learning holds (Proposition 2). In contrast, when the source is uninformative, agents are clearly not guaranteed to reach a consensus on the correct action; in fact, their actions may or may not converge at all. Proposition 1 demonstrates that an outside observer eventually learns the informativeness of the source if and only if the agents' actions do not converge when the source is uninformative.

A key assumption underlying our results is the assumption of a uniform prior on the payoff-relevant state given an uninformative source. This assumption ensures that despite information uncertainty, rational agents always act as if the signals they receive are informative (Lemma 1). Intuitively, in the absence of any useful information—when the source is uninformative—each agent with a uniform prior is indifferent between the available actions. Therefore, there is no harm in always treating signals as informative. We make this uniform prior assumption for tractability, and it can be viewed as capturing settings in which agents are not very informed a priori, thus making private signals and their informativeness a crucial determining factor of outcomes. Indeed, in many investment settings, the efficient market hypothesis [Fama, 1965, Samuelson, 1965] implies that investors should be close to indifference.[3]

Related Literature

Our paper contributes to a rich literature on sequential social learning. Assuming that the common source of information is always informative, the primary focus of this literature has been on determining whether agents can eventually learn to choose the correct action. Various factors, such as the information structure [Banerjee, 1992, Bikhchandani, Hirshleifer, and Welch, 1992, Smith and Sørensen, 2000] and the observational networks [Acemoglu, Dahleh, Lobel, and Ozdaglar, 2011, Çelen and Kariv, 2004, Lobel and Sadler, 2015], have been extensively studied to analyze their

[3]The efficient market hypothesis states that in the financial market, asset prices should reflect all available information. Thus, if investors are not indifferent, it suggests that they possess private information that is not yet reflected in market prices, thereby challenging the hypothesis.

impact on information aggregation, including its efficiency [Rosenberg and Vieille, 2019] and the speed of learning [e.g., Hann-Caruthers, Martynov, and Tamuz, 2018, Vives, 1993]. However, the question of learning about the informativeness of the source—which is the focus of this paper—remains largely unexplored.[4]

A few papers explore the idea of agents having access to multiple sources of information in the context of social learning. For example, Liang and Mu [2020] consider a model in which agents endogenously choose from a set of correlated information sources, and the acquired information is then made public and learned by other agents. They focus on the externality in agents' information acquisition decisions and show that information complementarity can result in either efficient information aggregation or "learning traps," in which society gets stuck in choosing suboptimal information structures. In a different setting, Chen [2022] examines a sequential social learning model in which ambiguity-averse agents have access to different sources of information. Consequently, information uncertainty arises in his model because agents are unsure about the signal precision of their predecessors. He shows that under sufficient ambiguity aversion, there can be information cascades even with unbounded signals. Our paper differs from these prior works as we focus on rational agents with access to a common source of information of unknown informativeness.

Another way of viewing our model is by considering a social learning model with four states: the source is either informative with the good or bad state, or uninformative with either the good or bad state. In such multi-state settings, recent work by Arieli and Mueller-Frank [2021] demonstrates that pairwise unbounded signals are necessary and sufficient for learning, when the decision problem that agents face includes a distinct action that is uniquely optimal for each state. This is not the case in our model, because the same action is optimal in different states, e.g., when the source is uninformative, and so even when agents observe a very strong signal indicating that the state is uninformative, they do not reveal it in their behavior.

More recently, Kartik, Lee, Liu, and Rappoport [2022] consider a setting with multiple states and actions on general sequential observational networks. They identify a sufficient condition for learning —"excludability" —that jointly depends on the information structure and agents' preferences. Roughly speaking, this condition ensures that agents can always displace the wrong action, which is their driving

[4] For comprehensive surveys on recent developments in the social learning literature, see e.g., Bikhchandani, Hirshleifer, Tamuz, and Welch [2021], Golub and Sadler [2017].

force for learning. In our model, when the source is uninformative, agents cannot displace the wrong action as all signals are pure noise.[5] Conceptually, our approach differs from theirs as we are interested in identifying the uninformative state from the informative one, instead of identifying the payoff-relevant state.

Our paper is also related to the growing literature on social learning with misspecified models. Bohren [2016] investigates a model where agents fail to account for the correlation between actions, demonstrating that different degrees of misperception can lead to distinct learning outcomes. In a broader framework, Bohren and Hauser [2021] show that learning remains robust to minor misspecifications. In contrast, Frick, Iijima, and Ishii [2020] find that an incorrect understanding of other agents' preferences or types can result in a severe breakdown of information aggregation, even with a small amount of misperception. Later, Frick, Iijima, and Ishii [2023] propose a unified approach to establish convergence results in misspecified learning environments where the standard martingale approach fails to hold. On a more positive note, Arieli, Babichenko, Müller, Pourbabaee, and Tamuz [2023] illustrate that by being mildly condescending—misperceiving others as having slightly lower-quality of information—agents may perform better in the sense that on average, only finitely many of them take incorrect actions.

1.2 Leading Example: Testing a New Information Technology

Consider an external evaluator tasked with assessing the informativeness of a novel information technology, such as an AI recommendation system. This system is used by a series of investors who are interested in an investment of unknown quality and have no prior information about it. For concreteness, assume that the quality of the investment is either good or bad, and the corresponding optimal decisions are to invest if it is good and not to invest if it is bad. Although the recommendation system could potentially be uninformative—thus providing no useful information on the quality of the investment—investors consistently treat the private recommendations they receive as informative since they have no other information to follow. Furthermore, before making their own investment decisions, investors also observe the past decisions of their peers who received recommendations from the same system. The evaluator has access only to these investment decisions and not to the private recommendations themselves. Based on these observations, the evaluator is

[5] This observation can also be seen from Theorem 2 in Kartik, Lee, Liu, and Rappoport [2022].

rewarded if he makes the correct assessment of the information technology.[6]

From the perspective of the external evaluator, whether he eventually makes the correct assessment is analogous to whether past investment decisions ultimately reveal the truth about the informativeness of the technology. In this context, our main result indicates that when the uninformative system frequently issues strong but conflicting recommendations, the evaluator eventually learns the truth about the technology, thus reaching the correct assessment. In contrast, when such recommendations occur infrequently under the uninformative system, the evaluator cannot be certain of the truth, and the correct assessment is no longer guaranteed. The frequency at which these recommendations need to occur depends on their tail distributions generated under different systems—they are frequent if the uninformative system has fatter tails and infrequent if it has thinner tails.

Why Relative Tail Thickness Matters for Learning.

Why do uninformative signals with fatter tails induce asymptotic learning, while those with thinner tails do not? To understand why asymptotic learning hinges on a tail condition, recall that agents always treat signals as informative. Thus, when the source is informative and generates unbounded signals, agents eventually reach a consensus on the correct action, and so action switches will eventually cease altogether.

Now suppose the source is uninformative. The condition of having fatter tails means that a very extreme signal—say a signal that is 5 standard deviations away from its mean—is more likely to occur under the uninformative source than the informative one. As a consequence, the presence of such extreme signals suggests that the source is uninformative. Even though an outside observer does not directly observe the private signals received by the agents, he can learn from observing action switches triggered by these extreme signals. For example, consider an agent who deviates from an extended period of consensus on the good action. Upon observing such a deviation, the outside observer infers that this agent must have received an extremely

[6] As another example, we can view this common information source as a scientific paradigm—a set of principles guiding a specific scientific discipline. In this context, the question of learning about the source's (un)informativeness has a similar flavor to the question of making the right *paradigm shift*, as proposed by the scientific philosopher Thomas Kuhn in his influential book "The Structure of Scientific Revolutions" [Kuhn, 1962]. One classic example of a paradigm shift in geology is the acceptance of the theory of plate tectonics, which only occurred in the 20th century, despite the idea of a drifted continent being put forward as early as 1596. See Ortelius [1695] for a printed version of "Thesaurus Geographicus" (in English, "A New Body of Geography") where the cartographer Abraham Ortelius first proposed the hypothesis of continental drift.

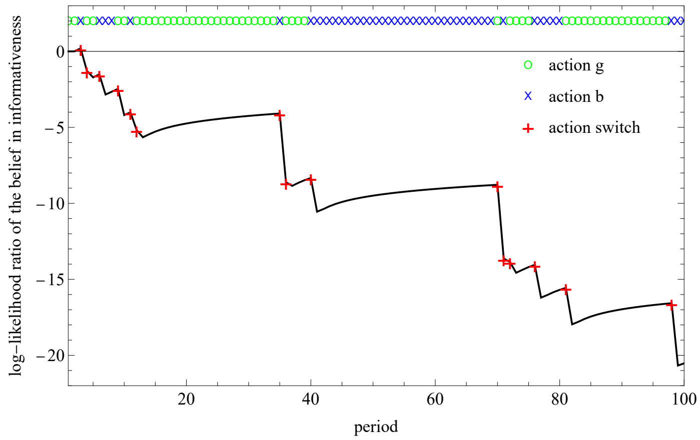

Figure 1.1: This simulation illustrates the log-likelihood ratio of an outside observer's belief in an informative source compared to an uninformative source over 100 periods. The process is generated under uninformative signals with a normal distribution with mean 0 and variance 2. The corresponding informative signals are normally distributed with unit variance and mean -1 and $+1$, respectively. Red crosses mark the periods in which action switches occur.

negative signal, which is more likely to occur under an uninformative source with fatter tails. Consequently, the observer becomes more convinced that the source is uninformative. It turns out that under uninformative signals with fatter tails, these action switches never cease; in fact, they will persist indefinitely. This persistence enables the observer to differentiate between the uninformative and informative sources.

In contrast, the condition of having thinner tails means that the uninformative source tends to produce more moderate signals than the informative source. Consequently, under such an uninformative source, action switches become increasingly rare over time, and eventually they will stop, thus hindering the observer from discerning the uninformative source from the informative one.

We illustrate the aforementioned intuition in Figure 1.1. It depicts a simulation of an outside observer's belief about the source being informative (in its log-likelihood ratio) under uninformative signals with fatter tails. First, observe that every action switch after an extended period of identical actions leads to a significant decrease in the observer's belief that the source is informative. As discussed before, observing these unusual action switches makes the observer less convinced that the source is informative. Second, although his belief gradually increases in the absence of action switches, our main result implies that it will eventually converge to zero, so the corresponding log-likelihood ratio will converge to negative infinity. By contrast, if his belief were generated under uninformative signals with thinner tails, then extreme signals would be less likely, and so would be action switches. Consequently, these action switches would eventually cease, and thus his belief would no longer converge to zero.

1.3 Model

Setup

There is an unknown binary state of the world $\theta \in \{\mathfrak{g}, \mathfrak{b}\}$, chosen at time 0 with equal probability. We refer to $\mathfrak{g}$ as the good state and $\mathfrak{b}$ as the bad state. A countably infinite set of agents indexed by time $t \in \mathbb{N} = \{1, 2, \ldots\}$ arrive in order, each acting once. The action of agent t is $a_t \in A = \{\mathfrak{g}, \mathfrak{b}\}$, with a payoff of one if her action matches the state θ and zero otherwise. Before agent t chooses an action, she observes the history of actions made by her predecessors $H_t = (a_1, \ldots, a_{t-1})$ and receives a private signal s_t, taking values in a measurable space (S, Σ).

A pure strategy of agent t is a measurable function $\sigma_t : A^{t-1} \times S \to A$ that selects an action for each possible pair of observed history and private signal. A pure strategy profile $\sigma = (\sigma_t)_{t \in \mathbb{N}}$ is a collection of pure strategies of all agents. A strategy profile is a Bayesian Nash equilibrium—referred to as equilibrium hereafter—if no agent can unilaterally deviate from this profile and obtain a strictly higher expected payoff conditioned on their information. Given that each agent acts only once, the existence of an equilibrium is guaranteed by a simple inductive argument. In equilibrium, each agent t chooses the action a_t that maximizes her expected payoff given the available information:

$$a_t \in \underset{a \in A}{\arg\max}\, \mathbb{E}[\mathbb{1}(\theta = a) | H_t, s_t].$$

Below, we make a continuity assumption which implies that agents are never indifferent, and so there is a unique equilibrium.

The Informativeness of the Source

So far, the above setup is the canonical setting of the sequential social learning model [Banerjee, 1992, Bikhchandani, Hirshleifer, and Welch, 1992, Smith and Sørensen, 2000]. Our model builds on this setting and introduces another dimension of uncertainty regarding the informativeness of a common source. Specifically, at time 0, independent of the payoff-relevant state θ, nature chooses an additional binary state $\omega \in \{0, 1\}$ with equal probability.[7] When $\omega = 1$, the source is informative and sends i.i.d. signals across agents conditional on θ, with distribution μ_θ. When $\omega = 0$, the source is uninformative and still sends i.i.d. signals, but independently of θ, with distribution μ_0. The realization of ω determines the signal-generating process for all agents. Throughout, we denote by $\mathbb{P}_0[\cdot] := \mathbb{P}[\cdot \mid \omega = 0]$ and $\mathbb{P}_1[\cdot] := \mathbb{P}[\cdot \mid \omega = 1]$ the conditional probability distributions given $\omega = 0$ and $\omega = 1$, respectively. Similarly, we use the notation $\mathbb{P}_{1,\mathfrak{g}}[\cdot] := \mathbb{P}[\cdot \mid \omega = 1, \theta = \mathfrak{g}]$ to denote the conditional probability distribution given $\omega = 1$ and $\theta = \mathfrak{g}$. We use an analogous notation for $\omega = 1$ and $\theta = \mathfrak{b}$.

We first observe that, despite the uncertainty regarding the informativeness of the source, in equilibrium, each agent chooses the action that is most likely to match the state, *conditional* on the source being informative.

Lemma 1. *The equilibrium action for each agent t is*

$$a_t \in \arg\max_{a \in A} \mathbb{P}_1[\theta = a | H_t, s_t]. \tag{1.1}$$

That is, agents always act as if signals are informative, irrespective of the underlying signal-generating process. This lemma holds simply because treating signals as informative—even when they are pure noise—does not adversely affect agents' payoffs, since in the absence of any useful information, each agent is indifferent between the available actions given the uniform prior assumption.

Information Structure

The distributions $\mu_\mathfrak{g}$, $\mu_\mathfrak{b}$, and μ_0 are distinct and mutually absolutely continuous, so no signal fully reveals either state θ or ω. As a consequence, conditioned on $\omega = 1$, the log-likelihood ratio of any signal

$$\ell_t = \log \frac{d\mu_\mathfrak{g}}{d\mu_\mathfrak{b}}(s_t),$$

is well-defined, and we call it the *agent's private log-likelihood ratio*. By Lemma 1, ℓ_t captures how agents update their private beliefs regarding the relative likelihood of the good state over the bad state upon receiving their signals, regardless of the realization of ω. Hence, it is sufficient to consider ℓ_t to capture agents' behavior.[8] We denote by $F_\mathfrak{g}$ and $F_\mathfrak{b}$ the cumulative distribution functions (CDFs) of ℓ_t conditioned

[7] Our results do not depend on the independence assumption between θ and ω. They hold true as long as conditioned on $\omega = 0$, both realizations of θ are equally likely.

[8] Formally, the sequence of actions $a_1, \ldots, a_t$ is determined by $\ell_1, \ldots, \ell_t$.

on the event that $\omega = 1$ and $\theta = \mathfrak{g}$ and the event that $\omega = 1$ and $\theta = \mathfrak{b}$, respectively. We denote by F_0 the CDF of ℓ_t conditioned on $\omega = 0$. All conditional CDFs $F_\mathfrak{g}$, $F_\mathfrak{b}$, and F_0 are mutually absolutely continuous, as $\mu_\mathfrak{g}$, $\mu_\mathfrak{b}$ and μ_0 are. Let $f_\mathfrak{g}$, $f_\mathfrak{b}$ and f_0 denote the corresponding density functions of $F_\mathfrak{g}$, $F_\mathfrak{b}$ and F_0 whenever they are differentiable.

We focus on *unbounded signals* in the sense that the agent's private log-likelihood ratio can take on arbitrarily large or small values, i.e., for any $M \in \mathbb{R}$, there exists a positive probability that $\ell_t > M$ and a positive probability that $\ell_t < -M$. We informally refer to a signal s_t as *extreme* when the corresponding ℓ_t it induces has a large absolute value. A common example of unbounded private signals is the case of Gaussian signals, where s_t follows a normal distribution $\mathcal{N}(m_{(\omega,\theta)}, \sigma^2)$ with variance σ^2 and mean $m_{(\omega,\theta)}$ that depends on the pair of states (ω, θ). An extreme Gaussian signal is a signal that is, for example, $5 - \sigma$ away from the mean $m_{(\omega,\theta)}$.

We make two assumptions for expository simplicity. First, we assume that the pair $(F_\mathfrak{g}, F_\mathfrak{b})$ of informative conditional CDFs is symmetric around zero, i.e., $F_\mathfrak{g}(x) + F_\mathfrak{b}(-x) = 1$. This implies that our model is invariant with respect to a relabeling of the payoff-relevant state. Second, we assume that all conditional CDFs—$F_\mathfrak{g}$, $F_\mathfrak{b}$, and F_0—are continuous, so agents are never indifferent between actions.

In addition, we assume that $F_\mathfrak{b}$ has a differentiable left tail, i.e., is differentiable for all x negative enough and its probability density function $f_\mathfrak{b}$ satisfies the condition that $f_\mathfrak{b}(-x) < 1$ for all x large enough. By the symmetry assumption, this implies that $F_\mathfrak{g}$ also has a differentiable right tail and its density function $f_\mathfrak{g}$ satisfies the condition that $f_\mathfrak{g}(x) < 1$ for all x large enough. This is a mild technical assumption that holds for every non-atomic distribution commonly used in the literature, including the Gaussian distribution. It holds, for instance, whenever the density tends to zero at infinity.

Asymptotic Learning

We denote by $q_t := \mathbb{P}[\omega = 1 | H_t]$ the belief that an outside observer assigns to the source being informative after observing the history of agents' actions from time 1 to $t - 1$. As this observer collects more information over time, his belief q_t converges almost surely since it is a bounded martingale. To ensure that he eventually learns the truth regarding the informativeness of the source, we introduce the following notion of asymptotic learning.

Definition 1. Asymptotic learning *holds if for all* $\omega \in \{0,1\}$,

$$\lim_{t\to\infty} q_t = \omega \quad \mathbb{P}_\omega\text{-almost surely.}$$

That is, conditioned on an informative source, the outside observer's belief that the source is informative converges to one almost surely; meanwhile, conditioned on an uninformative source, his belief that the source is informative converges to zero almost surely. As we explain below in Section 1.5, asymptotic learning is always attainable, regardless of the underlying information structure, when all signals are publicly observable.

1.4 Relative Tail Thickness

To study the conditions for asymptotic learning, it is crucial to understand the concept of relative tail thickness, which compares the tail distributions of agents' private log-likelihood ratios induced by different signals. This comparison is important because it captures the relative likelihood of generating extreme signals from different sources. Formally, for any pair of CDFs (F_0, F_θ) where $\theta \in \{\mathfrak{g}, \mathfrak{b}\}$ and some $x \in \mathbb{R}_+$, we denote their corresponding ratios by

$$L_\theta(x) := \frac{F_0(-x)}{F_\theta(-x)} \quad \text{and} \quad R_\theta(x) := \frac{1-F_0(x)}{1-F_\theta(x)}.$$

For large x, $L_\theta(x)$ and $R_\theta(x)$ represent the left and right tail ratios of F_0 over F_θ, respectively. The following definitions of fatter and thinner tails describe situations in which extreme signals are either more or less likely to occur under an uninformative source compared to an informative one.

Definition 2. *(i) The uninformative signals have* fatter tails *than the informative signals if there exists* $\varepsilon > 0$ *such that*

$$L_\mathfrak{b}(x) \geq \varepsilon \text{ and } R_\mathfrak{g}(x) \geq \varepsilon \quad \text{for all } x \text{ large enough.}$$

(ii) The uninformative signals have thinner tails *than the informative signals if there exists* $\varepsilon > 0$ *such that either*

$$L_\mathfrak{g}(x) \leq 1/\varepsilon \quad \text{for all } x \text{ large enough,}$$

or

$$R_\mathfrak{b}(x) \leq 1/\varepsilon \quad \text{for all } x \text{ large enough.}$$

That is, for the uninformative signals to have fatter tails, both their corresponding left and right tail ratios need to be eventually bounded from below. Conversely, for the uninformative signals to have thinner tails, at least one of their corresponding tail ratios—either left or right—needs to be eventually bounded from above.[9]

Note that the first condition $L_{\mathfrak{b}}(x) \geq \varepsilon$ implies that $L_{\mathfrak{g}}(x) \geq \varepsilon$. This follows from the well-known fact that $F_{\mathfrak{g}}$ exhibits first-order stochastic dominance over $F_{\mathfrak{b}}$, i.e., $F_{\mathfrak{g}}(x) \leq F_{\mathfrak{b}}(x)$ for all $x \in \mathbb{R}$ [see, e.g., Chamley, 2004, Rosenberg and Vieille, 2019, Smith and Sørensen, 2000]. Similar statements apply to the remaining three conditions. Furthermore, note that the uninformative signals cannot have fatter and thinner tails simultaneously, as $F_{\mathfrak{g}}$ and $F_{\mathfrak{b}}$ represent distributions of the agent's private log-likelihood ratio.[10] However, there are distributions under which the uninformative signals have neither fatter nor thinner tails. For these cases, we characterize the conditions for asymptotic learning in the canonical Gaussian environment (see more details in Section 1.5).

Intuitively, compared to informative signals, uninformative signals with fatter tails are more likely to exhibit extreme values. Thus, by Bayes' Theorem, observing an extreme signal suggests that the source is uninformative. In contrast, uninformative signals with thinner tails tend to exhibit moderate values, so observing an extreme signal in this case suggests that the source is informative. Next, we provide three examples of uninformative signals with either fatter or thinner tails.

Example 1 (Gaussian Signals)**.** *Consider the case where $F_{\mathfrak{g}}$ is normal with mean $+1$ and unit standard deviation and $F_{\mathfrak{b}}$ is also normal with mean -1 and unit standard deviation.*

Suppose that F_0 has zero mean. If it has a standard deviation of 17, then the uninformative signals have fatter tails. In this case, if an extreme signal, such as anything above 11, is observed, it is much more likely that the source is uninformative than that an informative $10-\sigma$ signal were generated under $F_{\mathfrak{g}}$. On the other hand, if the standard deviation of F_0 is $1/17$, then the uninformative signals have thinner tails, and thus an extreme signal indicates that the source is informative. A graphical

[9] In statistics, the notion of relative tail thickness has also been explored. Our definition of thinner tails is closest to that of Rojo [1992], where a CDF F is considered not more heavily tailed than G if $\limsup_{x\to\infty}(1-F(x))/(1-G(x)) < \infty$. Other notions of relative tail thickness, represented in terms of density quantile functions, can be found in Parzen [1979] and Lehmann [1988]. See Rojo [1992] for a discussion of the relationship between these existing notions.

[10] In particular, $F_{\mathfrak{g}}$ and $F_{\mathfrak{b}}$ satisfy the following inequality: $\mathrm{e}^x F_{\mathfrak{g}}(-x) \leq F_{\mathfrak{b}}(-x)$, for all $x \in \mathbb{R}$.

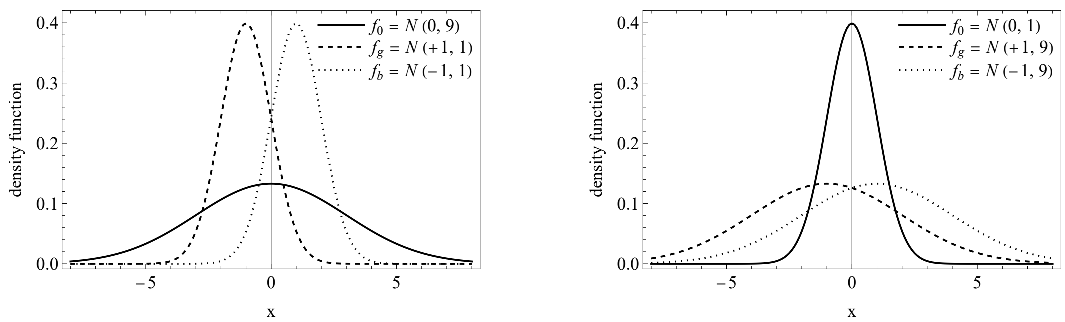

Figure 1.2: Probability density functions of informative Gaussian signals and their corresponding uninformative Gaussian signals with fatter tails (on the left) and with thinner tails (on the right).

example of uninformative Gaussian signals with fatter and thinner tails is shown in Figure 1.2.

Example 2 (First-Order Stochastically Dominated (FOSD) Signals)**.** *Suppose that F_0 first-order stochastically dominates $F_{\mathfrak{g}}$. In this case, since the uninformative signals are more likely to exhibit high values, F_0 has a thinner left tail than $F_{\mathfrak{g}}$, and thus by definition, it has thinner tails than the informative signals.*

Now, suppose that an extreme signal of a positive value is observed. Then, it is highly unlikely that the source is informative and associated with the bad state. Instead, it suggests that the source is more likely to be either uninformative or informative but associated with the good state. In this case, even though the observation of an extremely positive signal provides some evidence of the good state, it remains unclear whether the source is informative, as such a signal is likely to occur under both F_0 and $F_{\mathfrak{g}}$.

Example 3 (Mixture Signals)**.** *For any pair of distributions $(F_{\mathfrak{g}}, F_{\mathfrak{b}})$ and any $\alpha \in (0,1)$, let $F_0 = \alpha F_{\mathfrak{g}} + (1-\alpha)F_{\mathfrak{b}}$. Observe that the uninformative signals represented by F_0 have fatter tails than the informative signals.*[11] *In particular, when $\alpha = 1/2$, the corresponding mixture distribution $F_0 = (F_{\mathfrak{g}} + F_{\mathfrak{b}})/2$ coincides with the unconditional distribution of ℓ_t generated by an informative source. Thus, we can think of this uninformative source as being a priori indistinguishable from the informative one.*

Alternatively, the mixture distribution F_0 can be viewed as generating conditionally i.i.d. signals, but instead of conditioning on the state θ, they are generated condi-

[11] To see this, fix any constant $\alpha \in (0,1)$ and let $F_0 = \alpha F_{\mathfrak{g}} + (1-\alpha)F_{\mathfrak{b}}$. Since CDFs always take nonnegative values, for any $x \in \mathbb{R}$, $F_0(x) \geq (1-\alpha)F_{\mathfrak{b}}(x)$. Similarly, $1 - F_0(x) = \alpha(1 - F_{\mathfrak{g}}(x)) + (1-\alpha)(1 - F_{\mathfrak{b}}(x)) \geq \alpha(1 - F_{\mathfrak{g}}(x))$. Let $\varepsilon = \min\{\alpha, 1-\alpha\}$, and thus by definition F_0 has fatter tails.

tioned on a different state η. Specifically, suppose that in each period, the state η is randomly drawn from the same set $\{\mathfrak{g}, \mathfrak{b}\}$, independent of θ. Say, with probability α, the event $\eta = \mathfrak{g}$ occurs and a signal is drawn from the distribution $F_{\mathfrak{g}}$. With the complementary probability, the event $\eta = \mathfrak{b}$ occurs, and a signal is drawn from the distribution $F_{\mathfrak{b}}$. [12]

1.5 Main Results

A Benchmark

As a benchmark, we briefly discuss the case where all signals are observed by the outside observer.[13] Depending on the realizations of θ and ω, these signals are distributed according to either $\mu_{\mathfrak{g}}$, $\mu_{\mathfrak{b}}$ or μ_0. Since these measures are distinct, as the sample size grows, this observer eventually learns which distribution is being sampled. Formally, at time t, the empirical distribution of the signals $\hat{\mu}_t$ assigns to a measurable set $B \subseteq S$ the probability

$$\hat{\mu}_t(B) := \frac{1}{t} \sum_{\tau=1}^{t} \mathbb{1}(s_\tau \in B).$$

Conditional on both states ω and θ, this is the empirical mean of i.i.d. Bernoulli random variables. Hence, by the strong law of large numbers, for every measurable set $B \subseteq S$,

$$\lim_{t \to \infty} \hat{\mu}_t(B) = \mu_{(\omega,\theta)}(B) \quad \text{almost surely,}$$

where $\mu_{(1,\mathfrak{g})} = \mu_{\mathfrak{g}}$, $\mu_{(1,\mathfrak{b})} = \mu_{\mathfrak{b}}$ and $\mu_{(\cdot,0)} = \mu_0$.

Thus, regardless of the underlying signal-generating process, any uncertainty concerning the informativeness of the source is eventually resolved if all signals are publicly observable. Next, we turn to our main setting where the signals remain private and only the actions are observable. Clearly, in this setting, less information is available—observed actions contain less information than private signals. In addition, one needs to take into account the positive correlation between these observed actions.

Public Actions

We now present our main result (Theorem 1). In contrast to the public signal benchmark, our main result shows that the achievement of asymptotic learning is no

[12] When $\alpha = 1/2$, we can think of these uninformative signals as being generated based on a sequence of fair and independent coin tosses.

[13] Equivalently, one can let the outside observer observe all agents' actions in addition to their signals. Since actions contain no additional payoff relevant information, it suffices to only consider the signals.

longer guaranteed. In fact, the key determinant of asymptotic learning is the relative tail thickness between the uninformative and informative signals, as introduced in Definition 2.

Theorem 1. *When the uninformative signals have fatter tails than the informative signals, asymptotic learning holds. When the uninformative signals have thinner tails than the informative signals, asymptotic learning fails.*

Theorem 1 demonstrates that an outside observer, who learns from observing agents' actions, eventually distinguishes between informative and uninformative sources if the latter source generates signals with greater dispersion than the former. In contrast, when the uninformative signals are relatively concentrated compared to informative ones, such differentiation becomes unattainable for the observer.

For example, consider informative signals that follow a normal distribution with unit variance and mean $+1$ and -1, respectively. When the uninformative signals follow a normal distribution with a variance of 2, they have fatter tails and thus Theorem 1 implies that asymptotic learning holds. In contrast, when the uninformative signals follow a normal distribution with a variance of 1/2, they have thinner tails, and thus Theorem 1 implies that asymptotic learning fails. Indeed, for normal distributions, the relative thickness of the tails is determined solely by their variances: a higher variance corresponds to fatter tails, while a lower variance corresponds to thinner tails (see Lemma 8 in the appendix). An immediate consequence of Theorem 1 is the following result.

Corollary 1. *Suppose all private signals are normal where the informative signals have variance σ^2, and the uninformative signals have variance τ^2. Then, asymptotic learning holds if $\tau > \sigma$ and fails if $\tau < \sigma$.*

The idea behind our proof of Theorem 1 is as follows. First consider the case where the source is informative. In this case, the likelihood of generating extreme signals that overturn a long streak of correct consensus decreases rapidly. Consequently, agents will eventually reach a consensus on the correct action since they always treat signals as informative. Now, suppose that the source is uninformative and instead of reaching a consensus, agents continue to disagree indefinitely, leading to both actions being taken infinitely often. If this were the case, an outside observer would eventually be able to distinguish between informative and uninformative sources, as they induce distinct behavioral patterns among agents. Whether these disagreements persist or not depends on whether the tails of the uninformative signals are thick enough to generate these overturning extreme signals.

In sum, when the tails of uninformative signals are sufficiently thick (i.e., when they have fatter tails), overturning extreme signals occur often enough so that disagreements persist. When the tails are not sufficiently thick (i.e., when they have thinner tails), these signals are less likely, so disagreements eventually cease. Hence, we conclude that the relative tail thickness between uninformative and informative signals plays an important role in determining the achievement of asymptotic learning.

Perpetual Disagreement

As mentioned above, if agents never reach a consensus under an uninformative source, then intuitively the outside observer can infer the uninformativeness of the source. To formally establish the key mechanism underlying our main result, we define the event in which agents never converge to any action as *perpetual disagreement*. Let $S := \sum_{t=1}^{\infty} \mathbb{1}(a_t \neq a_{t+1})$ denote the total number of action switches, so that the event $\{S = \infty\}$ is the perpetual disagreement event. It turns out that perpetual disagreement under an uninformative source is not only a sufficient condition for asymptotic learning but also a necessary condition, as shown in the following proposition.

Proposition 1. *Asymptotic learning holds if and only if conditioned on* $\omega = 0$*, the perpetual disagreement event occurs almost surely.*

The proof of Proposition 1 uses the idea of an outside observer whose goal is to guess the informativeness of the source. Note that the achievement of asymptotic learning means that eventually he can always make the correct guess. Since the outsider observes all agents' actions, he expects to see an eventual consensus when the source is informative. Therefore, if he observes action nonconvergence, he infers that the source must be uninformative and make the correct guess. Conversely, suppose to the contrary that agents could also reach a consensus when the source is uninformative. This implies that action convergence is plausible under both informative and uninformative sources. Consequently, the observer is no longer sure of the source's informativeness, contradicting our hypothesis that he always makes the correct guess.

Gaussian Private Signals

While Theorem 1 provides valuable insight into the role of relative tail thickness in determining the achievement of asymptotic learning, there are situations where uninformative signals have neither fatter nor thinner tails compared to informative signals, rendering Theorem 1 inapplicable. For example, consider a scenario where F_0, $F_{\mathfrak{g}}$, and $F_{\mathfrak{b}}$ are normal distributions with the same variance and mean 0, 1, and -1, respectively. As x approaches infinity, both $F_0(-x)$ and $F_{\mathfrak{b}}(-x)$ approach zero, but the former goes to zero much faster than the latter, leading $L_{\mathfrak{b}}(x)$ converging to zero. As a result, F_0 does not have fatter tails. Similarly, both $L_{\mathfrak{g}}(x)$ and $R_{\mathfrak{b}}(x)$ tend to infinity as x approaches infinity, so F_0 does not have thinner tails either.

To complement the findings of Theorem 1, we focus on the Gaussian environment where all signals are normal and share the same variance σ^2. The informative signals are symmetric with mean $+1$ and -1, respectively, while the uninformative signals have mean $m_0 \in (-1, 1)$.[14] In this setting, a simple calculation shows that the agent's private log-likelihood ratio is directly proportional to the private signal: $\ell_t = 2s_t/\sigma^2$. As a consequence, all distributions $F_{\mathfrak{g}}$, $F_{\mathfrak{b}}$, and F_0 are also Gaussian. In the following result, we show that asymptotic learning is achieved if and only if the uninformative signals are symmetric around zero.

Theorem 2. *Suppose all private signals are Gaussian with the same variance, where informative signals have means* -1 *and* $+1$*, and uninformative signals have mean* $m_0 \in (-1, 1)$*. Then, asymptotic learning holds if and only if* $m_0 = 0$*.*

Together with Corollary 1, Theorem 2 provides a complete characterization of asymptotic learning in the Gaussian environment with symmetric informative signals. In particular, it shows that in such an environment, asymptotic learning holds in a knife-edge case where the uninformative distribution is symmetric around zero. Intuitively, any deviation in the mean of F_0 away from zero would bring F_0 closer to either $F_{\mathfrak{g}}$ or $F_{\mathfrak{b}}$, thus making the uninformative signals more similar to the corresponding informative signals. Consequently, it becomes impossible to fully differentiate between them.

[14] In the case where all Gaussian signals share the same variance and the absolute value of m_0 is strictly greater than one, the uninformative signals clearly have thinner tails. For example, when $m_0 = 2$, it reduces to Example 2. Thus, our main result implies that asymptotic learning fails.

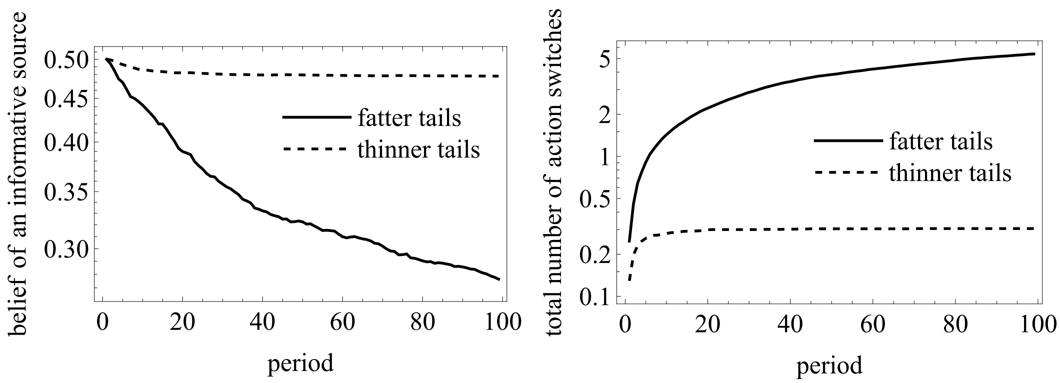

Figure 1.3: The belief that the source is informative (on the left) and the total number of action switches (on the right) under uninformative signals with fatter and thinner tails.

Numerical Simulation

To further illustrate our main result, we use Monte Carlo simulations to numerically simulate the belief process of an outside observer and the corresponding action switches among agents. We fix the pair of informative Gaussian signal distributions to be $(\mu_{\mathfrak{g}}, \mu_{\mathfrak{b}})$ where $\mu_{\mathfrak{g}} = \mathcal{N}(+1, 2)$ and $\mu_{\mathfrak{b}} = \mathcal{N}(-1, 2)$ and simulate these processes under fatter-tailed uninformative Gaussian signals with distribution $\mu_0 = \mathcal{N}(0, 3)$ and thinner-tailed uninformative Gaussian signals with distribution $\mu_0 = \mathcal{N}(0, 1)$, respectively.[15] We conduct these simulations 1,000 times and calculate the averages for each period. This yields approximations for the expected belief and the expected total number of action switches in the presence of uninformative signals. Figure 1.3 displays the results of these simulations.

What immediately stands out is that under fatter-tailed uninformative signals, the belief of the outside observer that the source is informative decreases much faster compared to thinner-tailed uninformative signals. By period 60, this belief is approximately 0.3 under fatter-tailed uninformative signals, which is less than two-thirds of the belief observed under thinner-tailed uninformative signals. These findings align with the predictions of Theorem 1, suggesting that in the former case the observer will eventually learn that the source is uninformative. In contrast, with thinner-tailed uninformative signals, the decline in the belief about the source's informativeness plateaus around 0.48 after period 20, suggesting that the observer will not be able to learn that the source is uninformative.

As shown in Proposition 1, the key mechanism driving asymptotic learning is

[15] Note that in this case, the agent's private log-likelihood ratio $\ell_t = s_t$, so F_0, $F_{\mathfrak{g}}$, and $F_{\mathfrak{b}}$ have the same distribution as μ_0, $\mu_{\mathfrak{g}}$, and $\mu_{\mathfrak{b}}$, respectively.

the persistence of disagreements under an uninformative source. Recall that the intuition is that the uninformative source with fatter tails has a higher probability of generating extreme signals, which in turn, prevents agents from converging to a consensus. This phenomenon is evident in the right plot of Figure 1.3, where the average total number of action switches under fatter-tailed uninformative signals increases over time. In contrast, under the thinner-tailed uninformative signals, the total number of switches plateaus in a short amount of time, suggesting that perpetual disagreement does not occur in this case.

1.6 Asymptotic Learning and Other Notions of Learning

In this section, we discuss the connections between asymptotic learning and other notions of learning that have been studied in the literature. One common notion, concerning the convergence of agents' actions, is known as herding. Another notion, concerning the convergence of agents' beliefs, is known as complete learning. We will show that in the presence of information uncertainty, agents may eventually choose the correct action without ever being certain that it is correct.

Formally, we say that *correct herding* holds if $\lim_{t\to\infty} a_t = \theta$ almost surely. That is, the event that all but finitely many agents take the correct action occurs with probability one. Let $p_t := \mathbb{P}[\theta = \mathfrak{g}|H_t]$ denote the belief assigned by agent t to the good state given the history of observed actions; we refer to it as the *social belief* at time t. We say that *complete learning* holds if the social belief p_t almost surely converges to one when $\theta = \mathfrak{g}$ and to zero when $\theta = \mathfrak{b}$. Essentially, complete learning means that society eventually becomes confident that they are making the correct choice.

In the standard model where there is no information uncertainty and the source is always informative, Smith and Sørensen [2000] show that correct herding holds if and only if signals are unbounded. In our model, since agents always act as if signals are informative, when the source is indeed informative and generates unbounded signals, correct herding also holds:

$$\lim_{t\to\infty} a_t = \theta \quad \mathbb{P}_1\text{-almost surely.}$$

That is, given an informative source, agents eventually converge to the correct action, regardless of the achievement of asymptotic learning. By contrast, given an uninformative source, agents are clearly not guaranteed to converge to the correct action; in fact, their actions may or may not converge at all. Indeed, as suggested by

Proposition 1, in this case, agents' actions do not converge if asymptotic learning holds or they may converge to the wrong action if asymptotic learning fails.

Nevertheless, the achievement of correct herding under an informative source does not imply the attainment of complete learning. In this case, although agents eventually converge to the correct action, they remain uncertain about its correctness unless asymptotic learning holds. Conversely, if the informativeness of the source remains uncertain, so do agents' beliefs about the correctness of their action consensus. We summarize this relationship in the following proposition.

Proposition 2. *Asymptotic learning holds if and only if conditioned on $\omega = 1$, complete learning holds.*

The proof of Proposition 2 uses the idea of an outside observer who observes the same public information as agent t, but does not observe her private signal. So, at time t, the observer's belief is equal to the social belief. As long as the observer remains uncertain about the source's informativeness, he cannot fully trust the public information, which consists of agents' actions. Conversely, once the source's informativeness is confirmed, this public information becomes highly accurate, enabling his belief to converge to the truth. The next result shows that when the source is uninformative, achieving asymptotic learning is equivalent to the social belief converging to the uniform prior.

Proposition 3. *Asymptotic learning holds if and only if conditioned on $\omega = 0$, $\lim_{t\to\infty} p_t = 1/2$ almost surely.*

The proof of Proposition 3 uses a similar approach to Proposition 2 and applies the result of Proposition 1. One direction is straightforward: Once an outside observer learns that the source is uninformative, his belief about the payoff-relevant state remains uniform since agents' actions are based on signals that contain no information. For the other direction, suppose that asymptotic learning fails under an uninformative source. Then, by Proposition 1, it implies that agents may eventually reach a consensus on some action. Since this is also possible under an informative source, the observer cannot completely disregard the possibility that agents' actions are informative. As a consequence, his belief about the payoff-relevant state is not guaranteed to remain uniform.

1.7 Analysis

In this section we first examine how agents update their beliefs. We present two standard yet useful properties of their belief updating process, namely, the *overturning principle* and *stationarity*. Then, based on Proposition 1, we characterize asymptotic learning in terms of *immediate agreement*—the event in which all agents immediately reach a consensus on some action—which simplifies the problem of asymptotic learning. Finally, we provide a proof sketch of our main result (Theorem 1) at the end of this section.

Agents' Beliefs Dynamics

Since agents always act as if signals are informative (Lemma 1), we focus on how agents update their beliefs conditioned on an informative source. We denote by π_t the *public belief* that agent t assigns to the good state after observing the history of actions H_t, i.e., $\pi_t := \mathbb{P}_1[\theta = \mathfrak{g}|H_t]$. The corresponding log-likelihood ratio of π_t is given by

$$r_t := \log \frac{\pi_t}{1-\pi_t} = \log \frac{\mathbb{P}_1[\theta = \mathfrak{g}|H_t]}{\mathbb{P}_1[\theta = \mathfrak{b}|H_t]}.$$

Furthermore, let L_t denote the log-likelihood ratio of the posterior belief that agent t assigns to the good state over the bad state after observing both the history of actions and her private signal:

$$L_t := \log \frac{\mathbb{P}_1[\theta = \mathfrak{g}|H_t, s_t]}{\mathbb{P}_1[\theta = \mathfrak{b}|H_t, s_t]}.$$

Recall that ℓ_t is the log-likelihood ratio of agent t's private belief induced by a signal s_t conditioned on $\omega = 1$. By Bayes' rule, we can write

$$L_t = r_t + \ell_t.$$

It follows from (1.1) that, in equilibrium, agent t chooses action $\mathfrak{g}$ if $\ell_t \geq -r_t$ and action $\mathfrak{b}$ if $\ell_t < -r_t$. Hence, conditioned on the state θ and the event that $\omega = 1$, the probability that agent t chooses action $\mathfrak{g}$ is $1 - F_\theta(-r_t)$ and the probability that she chooses action $\mathfrak{b}$ is $F_\theta(-r_t)$. As a consequence, the agents' public log-likelihood ratios (r_t) evolve as follows:

$$r_{t+1} = r_t + D_{\mathfrak{g}}(r_t) \quad \text{if } a_t = \mathfrak{g}, \tag{1.2}$$

$$r_{t+1} = r_t + D_{\mathfrak{b}}(r_t) \quad \text{if } a_t = \mathfrak{b}, \tag{1.3}$$

where

$$D_{\mathfrak{g}}(r) := \log \frac{1 - F_{\mathfrak{g}}(-r)}{1 - F_{\mathfrak{b}}(-r)} \quad \text{and} \quad D_{\mathfrak{b}}(r) := \log \frac{F_{\mathfrak{g}}(-r)}{F_{\mathfrak{b}}(-r)}.$$

Note that $D_{\mathfrak{g}}$ always takes positive values and $D_{\mathfrak{b}}$ always takes negative values as $F_{\mathfrak{g}}$ first-order stochastically dominates $F_{\mathfrak{b}}$.

Overturning Principle and Stationarity

One important property held by the agent's public belief is known as the *overturning principle* [Smith and Sørensen, 2000, Sørensen, 1996], which asserts that a single action switch is sufficient to change the verdict of π_t.

Lemma 2 (Overturning Principle)**.** *For each agent t, if $a_t = \mathfrak{g}$, then $\pi_{t+1} \geq 1/2$. Similarly, if $a_t = \mathfrak{b}$, then $\pi_{t+1} \leq 1/2$.*

Another important property held by π_t is *stationarity*—the value of π_t captures all past information about the payoff-relevant state, independent of time. This holds in our model because, regardless of the informativeness of the source, agents always update their public log-likelihood ratios according to either (1.2) or (1.3). We further write $\mathbb{P}_{\tilde{\omega},\tilde{\theta},\pi}$ to denote the conditional probability distribution given the pair of state realizations $(\tilde{\omega}, \tilde{\theta})$ while highlighting the different values of the prior π.

Lemma 3 (Stationarity)**.** *For any fixed sequence $(b_\tau)_{\tau=1}^k$ of k actions in $\{\mathfrak{g}, \mathfrak{b}\}$, any prior $\pi \in (0, 1)$ and any pair $(\tilde{\omega}, \tilde{\theta}) \in \{0, 1\} \times \{\mathfrak{g}, \mathfrak{b}\}$*

$$\mathbb{P}_{\tilde{\omega},\tilde{\theta}}[a_{t+1} = b_1, \ldots, a_{t+k} = b_k | \pi_t = \pi] = \mathbb{P}_{\tilde{\omega},\tilde{\theta},\pi}[a_1 = b_1, \ldots, a_k = b_k].$$

This lemma states that regardless of the source's informativeness, if agent t's public belief is equal to π, then the probability of observing a sequence $(b_1, \ldots, b_k)$ of actions of length k is the same as observing this sequence starting from time 1, given that the agents' prior on the payoff-relevant state is π.

Asymptotic Learning and Immediate Agreement

In our model, since the agent's public belief π_t evolves as in the standard model, it remains a martingale when the source is informative. However, when the source is uninformative, π_t ceases to be a martingale under the measure $\mathbb{P}_0$. Therefore, we need to employ a different analytical approach to understand what ensures asymptotic learning.[16] To do so, we focus on the event $\{a_1 = a_2 = \ldots = a\}$ in which all agents immediately reach a consensus on some action $a \in \{\mathfrak{g}, \mathfrak{b}\}$. We denote such an event

[16] A similar approach can be found in Arieli, Babichenko, Müller, Pourbabaee, and Tamuz [2023], where the agent's public belief also ceases to be a martingale under the correct measure because overconfident agents have misspecified beliefs.

by $\{\bar{a} = a\}$ and refer to it as *immediate agreement on action a*. Note that conditioned on this event, the process of public log-likelihood ratios (r_t) is deterministic and evolves according to either (1.2) or (1.3).

The following lemma shows that conditioned on an informative source, immediate agreement on the wrong action is impossible, whereas immediate agreement on the correct action is possible, at least for some prior. For brevity, we state this result only for the case where $\theta = \mathfrak{g}$. By symmetry, analogous statements hold for $\theta = \mathfrak{b}$.

Lemma 4. *Conditioned on $\omega = 1$ and $\theta = \mathfrak{g}$, the following two conditions hold:*

(i) Immediate agreement on action $\mathfrak{b}$ is impossible.
(ii) Immediate agreement on action $\mathfrak{g}$ is possible for some prior $\pi \in (0, 1)$.

The first part of Lemma 4 holds since, as mentioned before, conditioned on an informative source, all but finitely many agents take the correct action. This immediately implies that agents cannot reach a consensus on the wrong action from the outset. Likewise, the second part of Lemma 4 holds because if agents eventually reach a consensus on the correct action, by stationarity, they can also do so immediately at least for some prior.[17]

Next, we focus on the immediate agreement event conditioned on an uninformative source. Recall that in Proposition 1, we have established the relationship between asymptotic learning and perpetual disagreement. Building on this, we now characterize asymptotic learning in terms of immediate agreement, which is crucial in proving our main result.

Proposition 4. *Asymptotic learning holds if and only if conditioned on $\omega = 0$, immediate agreement on any actions is impossible.*

This proposition states that achieving asymptotic learning is equivalent to the absence of immediate agreement starting at a uniform prior given an uninformative source. The proof of Proposition 4 utilizes the idea that the agent's belief updating process is eventually monotonic—a technical property that we establish in Lemma 6 in the appendix. This property ensures that if immediate agreement is impossible for some prior, e.g., the uniform prior, it becomes impossible for *all* priors. By applying stationarity, this implies that agents cannot reach a consensus on any actions for any

[17] In fact, part (ii) of Lemma 4 holds not only for some prior but also for all priors. One can see this by applying a similar argument used in the proof of Lemma 7 in the appendix. We omit the stronger statement here, as it is not required to prove our main result.

priors. Consequently, perpetual disagreement occurs with probability one, and it follows from Proposition 1 that it is equivalent to asymptotic learning.

Therefore, we have reduced the problem of asymptotic learning to determining the possibility of immediate agreement conditioned on an uninformative source, which is much easier to analyze. Specifically, conditioned on the event $\{\bar{a} = \mathfrak{g}\}$, let $r_t^{\mathfrak{g}}$ denote the deterministic process of r_t based on (1.2). Recall that agent t chooses $a_t = \mathfrak{g}$ if $\ell_t \geq -r_t$ and $a_t = \mathfrak{b}$ otherwise. Consequently, the probability of $\{\bar{a} = \mathfrak{g}\}$ is equal to the probability that $\ell_t \geq -r_t^{\mathfrak{g}}$ for all $t \geq 1$. Moreover, conditioned on an uninformative source, since private signals are i.i.d., the corresponding private log-likelihood ratios are also i.i.d. Thus, conditioned on $\omega = 0$, the probability of immediate agreement on action $\mathfrak{g}$ is

$$\mathbb{P}_0[\bar{a} = \mathfrak{g}] = \prod_{t=1}^{\infty}(1 - F_0(-r_t^{\mathfrak{g}})).$$

To determine whether the above probability is positive or zero, by a standard approximation argument, it is equivalent to examining whether the sum of the probabilities of the following events is finite or infinite:

$$\mathbb{P}_0[\bar{a} = \mathfrak{g}] > 0 \ (= 0) \ \Leftrightarrow \ \sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}) < \infty \ (= \infty). \tag{1.4}$$

By the symmetry of the model, we also have

$$\mathbb{P}_0[\bar{a} = \mathfrak{b}] > 0 \ (= 0) \ \Leftrightarrow \ \sum_{t=1}^{\infty} \left(1 - F_0(r_t^{\mathfrak{g}})\right) < \infty \ (= \infty). \tag{1.5}$$

In summary, asymptotic learning holds if and only if conditioned on the source being uninformative, the probability of generating extreme signals decreases slowly enough so that both sums in (1.4) and (1.5) are infinite. As we discuss below, for uninformative signals with fatter tails, these sums diverge, and for signals with thinner tails, at least one of these sums converges.

Proof Sketch of Theorem 1

We conclude this section by providing a sketch of the proof of Theorem 1. On the one hand, by part (i) of Lemma 4, the probability of generating extreme signals that match the payoff-relevant state decreases relatively slowly under informative signals. Hence, if the source is uninformative and generates signals with fatter tails, this probability declines even more slowly. Consequently, both the sums in (1.4) and

(1.5) are infinite, which means that no immediate agreement is possible. Thus, by Proposition 4, asymptotic learning holds. On the other hand, by part (ii) of Lemma 4, the probability of generating extreme signals that mismatch the state decreases relatively fast under informative signals. Hence, if the source is uninformative and generates signals with thinner tails, this probability decreases even more rapidly, at least for some type of extreme signals. Consequently, either the sum in (1.4) or the sum in (1.5) (or both) is finite, which means that immediate agreement on some action is possible. Hence, by Proposition 4, asymptotic learning fails.

1.8 Conclusion

In this paper, we study the sequential social learning problem in the presence of a potentially uninformative source, e.g., an AI recommendation system of unknown quality. We show that achieving asymptotic learning, in which an outside observer eventually discerns the informativeness of the source, is not guaranteed, and it depends on the relationship between the conditional distributions of the private signals. In particular, it hinges on the relative tail distribution between signals: when uninformative signals have fatter tails compared to their informative counterparts, asymptotic learning holds; conversely, when they have thinner tails, asymptotic learning fails. We also characterize the conditions for asymptotic learning in the canonical case of Gaussian private signals, where the relative tail thickness is incomparable.

More generally, our analysis suggests that irregular behavior, such as an action switch (or disagreement) following a prolonged sequence of identical actions, is the driving force behind asymptotic learning. Indeed, contrary to the public-signal benchmark case in which asymptotic learning is always achieved, with private signals, an outside observer can only learn that the source is uninformative from observing these action switches. We show that conditioned on an uninformative source, when action switches accumulate indefinitely (or disagreements occur perpetually), the observer eventually learns the informativeness of the source, and vice versa. We view this characterization of asymptotic learning as the key mechanism behind our main result.[18]

A limitation of our results is that they apply only asymptotically. Our numerical

[18] In the context of scientific paradigms, this mechanism is reminiscent of Kuhn [1962]'s idea that the accumulation of anomalies may trigger scientific revolutions and paradigm shifts. See Ba [2022] for a study on the rationale behind the persistence of a misspecified model, e.g., a wrong scientific paradigm.

simulations suggest that in the case of uninformative Gaussian signals with fatter tails, the asymptotics can already kick in relatively early in the process. It would be interesting to understand the speed at which an outside observer learns about the informativeness of the source. Another promising extension is to explore varying degrees of informativeness, beyond the extreme cases currently considered in this paper. For example, instead of having either informative or completely uninformative signals, one could have a distribution over the variance of conditionally i.i.d. Gaussian signals and ask whether learning about the payoff-relevant state is achievable. We leave these questions for future research.

1.9 Proofs

Proof of Lemma 1. Recall that in equilibrium, each agent t chooses the action that is most likely to match the state θ conditioned on the available information (H_t, s_t). By Bayes' rule, the relative likelihood between the good state and the bad state for agent t is

$$\frac{\mathbb{P}[\theta = \mathfrak{g}|H_t, s_t]}{\mathbb{P}[\theta = \mathfrak{b}|H_t, s_t]} = \frac{\sum_{\tilde{\omega}\in\{0,1\}} \mathbb{P}_{\tilde{\omega}}[\theta = \mathfrak{g}|H_t, s_t] \cdot \mathbb{P}[\omega = \tilde{\omega}|H_t, s_t]}{\sum_{\tilde{\omega}\in\{0,1\}} \mathbb{P}_{\tilde{\omega}}[\theta = \mathfrak{b}|H_t, s_t] \cdot \mathbb{P}[\omega = \tilde{\omega}|H_t, s_t]}.$$

Note that $\mathbb{P}_0[\theta = \mathfrak{g}|H_t, s_t] = \mathbb{P}_0[\theta = \mathfrak{g}]$ and $\mathbb{P}_0[\theta = \mathfrak{b}|H_t, s_t] = \mathbb{P}_0[\theta = \mathfrak{b}]$ as conditioned on $\omega = 0$, neither the history H_t nor the signal s_t contains any information about the payoff-relevant state θ. Since the states ω and θ are independent of each other and the prior on θ is uniform, $\mathbb{P}_0[\theta = \mathfrak{g}|H_t, s_t] = \mathbb{P}_0[\theta = \mathfrak{b}|H_t, s_t] = 1/2$. Thus, it follows from the above equation that

$$\frac{\mathbb{P}[\theta = \mathfrak{g}|H_t, s_t]}{\mathbb{P}[\theta = \mathfrak{b}|H_t, s_t]} \geq 1 \Leftrightarrow \frac{\mathbb{P}_1[\theta = \mathfrak{g}|H_t, s_t]}{\mathbb{P}_1[\theta = \mathfrak{b}|H_t, s_t]} \geq 1.$$

That is, in equilibrium, each agent chooses the most likely action conditioned on the available information and the source being informative. □

Proof of Lemma 2. For any $t \geq 1$, one has that

$$\pi_{t+1} = \mathbb{E}_1[\mathbb{1}(\theta = \mathfrak{g})|H_{t+1}] = \mathbb{E}_1[\mathbb{E}_1[\mathbb{1}(\theta = \mathfrak{g})|H_t, s_t]|H_{t+1}]],$$

where the second equality follows from the law of the iterated expectation. Thus, if $a_t = \mathfrak{g}$, by Lemma 1, $\mathbb{P}_1[\theta = \mathfrak{g}|H_t, s_t] \geq \mathbb{P}_1[\theta = \mathfrak{b}|H_t, s_t]$. It follows from the above equation that $\pi_{t+1} \geq 1 - \pi_{t+1}$, which implies that $\pi_{t+1} \geq 1/2$. The case where $a_t = \mathfrak{b}$ implies that $\pi_{t+1} \leq 1/2$ follows from a symmetric argument. □

The following simple claim will be useful in proving Proposition 1. It employs an idea similar to the *no introspection principle* in Sørensen [1996]. Recall that $q_t = \mathbb{P}[\omega = 1|H_t]$ is the belief that an outside observer assigns to the source being informative based on the history of actions from time 1 to $t-1$.

Claim 1. *For any $a \in (0, 1/2)$ and any $b \in (1/2, 1)$,*

$$\mathbb{P}_0[q_t = \tilde{q}] \leq \frac{1-a}{a} \cdot \mathbb{P}_1[q_t = \tilde{q}], \text{ for all } \tilde{q} \in [a, 1/2];$$
$$\mathbb{P}_0[q_t = \tilde{q}] \geq \frac{1-b}{b} \cdot \mathbb{P}_1[q_t = \tilde{q}], \text{ for all } \tilde{q} \in [1/2, b].$$

Proof of Claim 1. For any $\tilde{q} \in (0, 1)$, let $\tilde{H}_t$ be a history of actions such that the associated belief q_t is equal to $\tilde{q}$. By the law of total expectation,

$$\mathbb{P}[\omega = 1|q_t = \tilde{q}] = \mathbb{E}[\mathbb{E}[\mathbb{1}(\omega = 1)|\tilde{H}_t]|q_t = \tilde{q}] = \mathbb{E}[\tilde{q}|q_t = \tilde{q}] = \tilde{q}.$$

It follows from Bayes' rule that

$$\frac{\mathbb{P}_0[q_t = \tilde{q}]}{\mathbb{P}_1[q_t = \tilde{q}]} = \frac{\mathbb{P}[\omega = 0|q_t = \tilde{q}]}{\mathbb{P}[\omega = 1|q_t = \tilde{q}]} \cdot \frac{\mathbb{P}[\omega = 1]}{\mathbb{P}[\omega = 0]} = \frac{1-\tilde{q}}{\tilde{q}}.$$

Since for any $a \in (0, 1/2)$ and any $\tilde{q} \in [a, 1/2]$, $1 \leq \frac{1-\tilde{q}}{\tilde{q}} \leq \frac{1-a}{a}$, it follows from the above equation that

$$\mathbb{P}_0[q_t = \tilde{q}] = \frac{1-\tilde{q}}{\tilde{q}} \cdot \mathbb{P}_1[q_t = \tilde{q}] \leq \frac{1-a}{a} \cdot \mathbb{P}_1[q_t = \tilde{q}].$$

The second inequality follows from an identical argument. □

In the following proofs for Proposition 1, 2 and 3, we use extensively the idea of an outside observer, say observer x, who observes everyone's actions. The information available to him at time t is thus $H_t = (a_1, \ldots, a_{t-1})$ and at time infinity is $H_\infty = \cup_t H_t$. He gets a utility of one if his guess about ω is correct and zero otherwise. Furthermore, we denote by $q_\infty := \mathbb{P}[\omega = 1|H_\infty]$ the belief that he has at time infinity about the event $\omega = 1$. Similarly, we denote by $p_\infty := \mathbb{P}[\theta = \mathfrak{g}|H_\infty]$ the belief that he has at time infinity about the event $\theta = \mathfrak{g}$.

Proof of Proposition 1. We first show the if direction. Recall that $S = \sum_{t=1}^{\infty} \mathbb{1}(a_t \neq a_{t+1})$ and suppose that $\mathbb{P}_0[S = \infty] = 1$. Denote by a_t^x the guess that the outside observer x would make to maximize his probability of guessing ω correctly at time t. Fix a large positive integer $k \in \mathbb{N}$. Let $A_t(k)$ denote the event that there have been at least k action switches before time t and denote its complementary event by $A_t^c(k)$.

Consider the following strategy $\tilde{a}^x_\infty(k)$ for x at time infinity: $\tilde{a}^x_\infty(k) = 0$ if $A_\infty(k)$ occurs and $\tilde{a}^x_\infty(k) = 1$ otherwise. That is, the observer would guess 0 if there are at least k action switches at time infinity and guess 1 otherwise. The expected payoff of x under this strategy is

$$\mathbb{P}[\tilde{a}^x_\infty(k) = \omega] = \mathbb{P}_0[A_\infty(k)] \cdot \mathbb{P}[\omega = 0] + \mathbb{P}_1[A^c_\infty(k)] \cdot \mathbb{P}[\omega = 1]. \tag{1.9.1}$$

Since conditioned on $\omega = 1$, agents eventually reach a consensus almost surely, it follows that for all k large enough,

$$\mathbb{P}_1[A^c_\infty(k)] = 1. \tag{1.9.2}$$

By assumption, $\mathbb{P}_0[S = \infty] = 1$, which implies that $\mathbb{P}_0[A_\infty(k)] = 1$ for all k. Thus, it follows from (1.9.1) and (1.9.2) that for all k large enough, $\mathbb{P}[\tilde{a}^x_\infty(k) = \omega] = 1$. In other words, for all k large enough, the strategy $\tilde{a}^x_\infty(k)$ achieves the maximal payoff for x.

Meanwhile, note that the optimal strategy for x at time infinity is to make a guess that he believes is most likely, given the information H_∞: $a^x_\infty = 1$ if $q_\infty \geq 1/2$ and $a^x_\infty = 0$ otherwise. Since $\mathbb{P}[\tilde{a}^x_\infty(k) = \omega] = 1$ for all k large enough, the optimal strategy a^x_∞ must also achieve the maximal payoff of one:

$$1 = \mathbb{P}[a^x_\infty = \omega] = \mathbb{P}_1[q_\infty \geq 1/2] \cdot \mathbb{P}[\omega = 1] + \mathbb{P}_0[q_\infty < 1/2] \cdot \mathbb{P}[\omega = 0]. \tag{1.9.3}$$

It follows from (1.9.3) that $\mathbb{P}_0[q_\infty < 1/2] = \mathbb{P}_1[q_\infty \geq 1/2] = 1$. It remains to show that $\mathbb{P}_0[q_\infty = 0] = 1$ and $\mathbb{P}_1[q_\infty = 1] = 1$. To this end, first notice that $\mathbb{P}_0[q_\infty < 1/2] = 1$ implies that $\mathbb{P}_0[q_\infty \geq 1/2] = 0$. Thus, by Claim 1, it further implies that for any $b \in (1/2, 1)$ and all $\tilde{q} \in [1/2, b]$, $\mathbb{P}_1[q_\infty = \tilde{q}] = 0$. Consequently, it follows from $\mathbb{P}_1[q_\infty \geq 1/2] = 1$ that $\mathbb{P}_1[q_\infty = 1] = 1$. The case that $\mathbb{P}_0[q_\infty = 0] = 1$ follows from an identical argument. Together, we conclude that asymptotic learning holds.

Next, we show the only-if direction. Suppose by contraposition that $\mathbb{P}_0[S < \infty] > 0$. Again, since conditioned on $\omega = 1$, agents eventually reach a consensus on the correct action, this implies that $\mathbb{P}_1[S < \infty] = 1$. Thus, there exists a history of actions $\tilde{H}_\infty$ at time infinity that is possible under both probability measures $\mathbb{P}_0$ and $\mathbb{P}_1$: $\mathbb{P}_0[\tilde{H}_\infty] > 0$ and $\mathbb{P}_1[\tilde{H}_\infty] > 0$. It follows from Bayes' rule that $\mathbb{P}[\omega = 1|\tilde{H}_\infty] < 1$ and $\mathbb{P}[\omega = 0|\tilde{H}_\infty] < 1$.

Assume without loss of generality that under this history $\tilde{H}_\infty$, the corresponding belief $\tilde{q}_\infty \geq 1/2$. Therefore, given $\tilde{H}_\infty$, the observer x would guess $a^x_\infty = 1$. As a consequence, the probability of x guessing correctly about ω is strictly less than one:

$$\begin{aligned}\mathbb{P}[a^x_\infty = \omega] &= \mathbb{P}[a^x_\infty = \omega, \tilde{H}_\infty] + \mathbb{P}[a^x_\infty = \omega, \tilde{H}^c_\infty] \\ &= \mathbb{P}[\omega = 1|\tilde{H}_\infty] \cdot \mathbb{P}[\tilde{H}_\infty] + \mathbb{P}[a^x_\infty = \omega, \tilde{H}^c_\infty] \\ &< \mathbb{P}[\tilde{H}_\infty] + \mathbb{P}[\tilde{H}^c_\infty] = 1.\end{aligned}$$

This implies that asymptotic learning fails since otherwise, by definition $\mathbb{P}[a^x_\infty = \omega] = 1$, which is in contradiction with the above strict inequality. □

The following equation will be useful in proving Proposition 2 and 3. Recall that q_∞ and p_∞ are the beliefs of the observer x assigned to the events $\omega = 1$ and $\theta = \mathfrak{g}$ at time infinity, respectively. Denote by $\pi_\infty = \mathbb{P}_1[\theta = \mathfrak{g}|H_\infty]$. By the law of total probability,

$$\begin{aligned}p_\infty &= \mathbb{P}_1[\theta = \mathfrak{g}|H_\infty] \cdot q_\infty + \mathbb{P}_0[\theta = \mathfrak{g}|H_\infty] \cdot (1 - q_\infty) \\ &= \pi_\infty \cdot q_\infty + \frac{1}{2} \cdot (1 - q_\infty), \qquad (1.9.4)\end{aligned}$$

where the second equality holds since conditioned on $\omega = 0$, no action contains any information about θ. Thus $\mathbb{P}_0[\theta = \mathfrak{g}|H_\infty] = \mathbb{P}_0[\theta = \mathfrak{g}] = 1/2$ given the uniform prior.

Proof of Proposition 2. Since agents always act as if signals are informative (Lemma 1), the agent's public belief $\pi_t = \mathbb{P}_1[\theta = \mathfrak{g}|H_t]$ remains a martingale under the measure $\mathbb{P}_1$. Using a standard martingale convergence argument with unbounded signals [Smith and Sørensen, 2000], we have that (i) conditioned on $\omega = 1$ and $\theta = \mathfrak{g}$, $\pi_\infty = 1$ almost surely and (ii) conditioned on $\omega = 1$ and $\theta = \mathfrak{b}$, $\pi_\infty = 0$ almost surely.

Suppose that conditioned on $\omega = 1$, complete learning holds. That is, (i) conditioned on $\omega = 1$ and $\theta = \mathfrak{g}$, $p_\infty = 1$ almost surely and (ii) conditioned on $\omega = 1$ and $\theta = \mathfrak{b}$, $p_\infty = 0$ almost surely. It follows from (1.9.4) that conditioned on $\omega = 1$, $q_\infty = 1$ almost surely so that asymptotic learning holds. Conversely, suppose asymptotic learning holds. By definition, conditioned on $\omega = 1$, $q_\infty = 1$ almost surely. It then follows from (1.9.4) that conditioned on $\omega = 1$, $p_\infty = \pi_\infty$ almost surely, and thus complete learning holds. □

Proof of Proposition 3. The only-if direction is straightforward: suppose asymptotic learning holds. By definition, conditioned on $\omega = 0$, $q_\infty = 0$ almost surely, and thus it follows from (1.9.4) that $p_\infty = 1/2$ almost surely.

Now, we prove the if direction. Suppose asymptotic learning fails. By Proposition 1, conditioned on $\omega = 0$, there exists a history $\tilde{H}_\infty$ that is possible at time infinity in which agents eventually reach a consensus on some action. Since such history is also possible conditioned on $\omega = 1$, the corresponding belief of an outside observer $\tilde{q}_\infty \in (0, 1)$.[19] Furthermore, note that depending on the action to which the history $\tilde{H}_\infty$ converges, the corresponding public belief of the agent $\tilde{\pi}_\infty$ takes values in $\{0, 1\}$. Hence, by (1.9.4), the event $\{p_\infty \neq 1/2\}$ occurs with positive probability. □

Proof of Lemma 4. The proof idea is similar to the proof of Lemma 10 in Arieli, Babichenko, Müller, Pourbabaee, and Tamuz [2023]. Recall that we use $\mathbb{P}_{1,\mathfrak{g}}$ to denote the conditional probability distribution given $\omega = 1$ and $\theta = \mathfrak{g}$ and we use $\mathbb{P}_{1,\mathfrak{g},\pi}$ to denote the same conditional probability distribution while emphasizing the prior value. Since conditioned on $\omega = 1$, correct herding holds, this means that $\mathbb{P}_{1,\mathfrak{g}}[\lim_t a_t = \mathfrak{g}] = 1$. As a consequence, part (i) follows directly from the fact that the events $\{\bar{a} = \mathfrak{b}\}$ and $\{\lim_t a_t = \mathfrak{g}\}$ are disjoint, and thus $\mathbb{P}_{1,\mathfrak{g}}[\bar{a} = \mathfrak{b}] = 0$.

For part (ii), let $\tau < \infty$ denote the last random time at which the agent chooses the wrong action $\mathfrak{b}$. It is well-defined as correct herding holds. Hence, $1 = \mathbb{P}_{1,\mathfrak{g}}[\lim_t a_t = \mathfrak{g}] = \sum_{k=1}^{\infty} \mathbb{P}_{1,\mathfrak{g}}[\tau = k]$. By the overturning principle (Lemma 2), $a_\tau = \mathfrak{b}$ implies that $\pi_{\tau+1} \leq 1/2$. As a consequence,

$$
\begin{aligned}
1 = \sum_{k=1}^{\infty} \mathbb{P}_{1,\mathfrak{g}}[\tau = k] &= \sum_{k=1}^{\infty} \mathbb{P}_{1,\mathfrak{g}}[a_{k+1} = a_{k+2} = \ldots = \mathfrak{g}, \pi_{k+1} \leq 1/2] \\
&= \sum_{k=1}^{\infty} \mathbb{E}_{1,\mathfrak{g}}\big[\mathbb{P}_{1,\mathfrak{g}}[a_{k+1} = a_{k+2} = \ldots = \mathfrak{g}, \pi_{k+1} \leq 1/2 \mid \pi_{k+1}]\big] \\
&= \sum_{k=1}^{\infty} \mathbb{E}_{1,\mathfrak{g}}\big[\mathbb{1}(\pi_{k+1} \leq 1/2) \cdot \mathbb{P}_{1,\mathfrak{g}}[a_{k+1} = a_{k+2} = \ldots = \mathfrak{g} \mid \pi_{k+1}]\big] \\
&= \sum_{k=1}^{\infty} \mathbb{E}_{1,\mathfrak{g}}\big[\mathbb{1}(\pi_{k+1} \leq 1/2) \cdot \mathbb{P}_{1,\mathfrak{g},\pi_{k+1}}[\bar{a} = \mathfrak{g}]\big],
\end{aligned}
$$

where the second equality follows from the law of total expectation, and the last equality follows from the stationarity property (Lemma 3). Suppose that for all prior $\pi \in (0, 1)$, $\mathbb{P}_{1,\mathfrak{g},\pi}[\bar{a} = \mathfrak{g}] = 0$. This implies that the above equation equals zero, a contradiction. □

[19] Note that conditioned on $\omega = 0$, q_∞ has support $\subseteq [0, 1)$ as $\mathbb{P}[\omega = 0 | q_\infty = 1] = 0$.

Proof of Theorem 1

In this section, we prove Proposition 4 and Theorem 1. To prove Proposition 4, we will first prove the following proposition (Proposition 5). Together with Proposition 1, they jointly imply Proposition 4. The proof of Theorem 1 is presented at the end of this section. We write $\mathbb{P}_{0,\pi}$ to denote the conditional probability distribution given $\omega = 0$ while highlighting the value of the prior π on θ.[20]

Proposition 5. *The following are equivalent.*

(i) *For any action* $a \in \{\mathfrak{g}, \mathfrak{b}\}$, $\mathbb{P}_{0,\pi}[\bar{a} = a] = 0$ *for all prior* $\pi \in (0, 1)$.

(ii) $\mathbb{P}_0[S = \infty] = 1$.

(iii) *For any action* $a \in \{\mathfrak{g}, \mathfrak{b}\}$, $\mathbb{P}_{0,\pi}[\bar{a} = a] = 0$ *for some prior* $\pi \in (0, 1)$.

To prove this proposition, we first establish some preliminary results on the process of the agents' public log-likelihood ratios conditioned on the event of immediate agreement. These results lead to Lemma 7, a crucial part in establishing the equivalence between no immediate agreement and perpetual disagreement conditioned on an uninformative source. We present the proof of Proposition 5 towards the end of this section.

Preliminaries

Recall that conditioned on $\{\bar{a} = \mathfrak{g}\}$, the process of the agent's public log-likelihood ratio r_t evolves deterministically according to (1.2), which we denote by $r_t^{\mathfrak{g}}$. Let the corresponding updating function be

$$\phi(x) := x + D_{\mathfrak{g}}(x).$$

That is, $r_{t+1}^{\mathfrak{g}} = \phi(r_t^{\mathfrak{g}})$ for all $t \geq 1$. Since the entire sequence $(r_t^{\mathfrak{g}})$ is determined once its initial value $r_1^{\mathfrak{g}}$ is specified, we denote the value of $r_t^{\mathfrak{g}}$ with an initial value $r_1^{\mathfrak{g}} = r$ by $r_t^{\mathfrak{g}}(r)$. We can thus write $r_t^{\mathfrak{g}}(r) = \phi^{t-1}(r)$ for all $t \geq 1$, where ϕ^t is its t-th composition and $\phi^0(r) = r$.

We remind the readers of two standard properties of the sequence $(r_t^{\mathfrak{g}})$, as summarized in the following lemma. The first part of this lemma states that $(r_t^{\mathfrak{g}})$ tends to infinity as t tends to infinity, and the second part shows that it takes only some bounded time for the sequence $(r_t^{\mathfrak{g}})$ to reach any positive value.

Lemma 5 (The Long-Run and Short-Run Behaviors of $r_t^{\mathfrak{g}}$)**.**

[20] We continue to omit the prior π when it is uniform.

(i) $\lim_{t\to\infty} r_t^{\mathfrak{g}} = \infty$.

(ii) *For any* $\bar{r} \geq 0$, *there exists* t_0 *such that* $r_{t_0}^{\mathfrak{g}}(r) \geq \bar{r}$ *for all* $r \geq 0$.

Proof. See Lemma 6 and Lemma 12 in Rosenberg and Vieille [2019]. □

Note that although the sequence $(r_t^{\mathfrak{g}})$ eventually approaches infinity, it may not do so monotonically without additional assumptions on the distributions $F_{\mathfrak{g}}$ and $F_{\mathfrak{b}}$.[21] The next lemma shows that, under some mild technical assumptions on the left tail of $F_{\mathfrak{b}}$, the function $\phi(x)$ eventually increases monotonically.

Lemma 6 (Eventual Monotonicity)**.** *Suppose that* $F_{\mathfrak{b}}$ *has a differentiable left tail and its probability density function* $f_{\mathfrak{b}}$ *satisfies the condition that, for all* x *large enough,* $f_{\mathfrak{b}}(-x) < 1$. *Then,* $\phi(x) := x + D_{\mathfrak{g}}(x)$ *increases monotonically for all* x *large enough.*

Proof. By assumption, we can find a constant $\rho < 1$ such that for all x large enough, $f_{\mathfrak{b}}(-x) \leq \rho$. By definition, $D_g(x) = \log \frac{1-F_{\mathfrak{g}}(-x)}{1-F_{\mathfrak{b}}(-x)}$. Taking the derivative of $D_{\mathfrak{g}}$,

$$D'_{\mathfrak{g}}(x) = \frac{f_{\mathfrak{g}}(-x)}{1 - F_{\mathfrak{g}}(-x)} - \frac{f_{\mathfrak{b}}(-x)}{1 - F_{\mathfrak{b}}(-x)}.$$

Observe that the log-likelihood ratio of the agent's private log-likelihood ratio ℓ_t is the log-likelihood ratio itself (see, e.g., Chamley [2004]):

$$\log \frac{dF_{\mathfrak{g}}}{dF_{\mathfrak{b}}}(x) = x.$$

It follows that

$$-D'_{\mathfrak{g}}(x) = f_{\mathfrak{b}}(-x)\Big(\frac{1}{1 - F_{\mathfrak{b}}(-x)} - \frac{\mathrm{e}^{-x}}{1 - F_{\mathfrak{g}}(-x)}\Big) \leq \frac{f_{\mathfrak{b}}(-x)}{1 - F_{\mathfrak{b}}(-x)}.$$

Fix some $\varepsilon > 0$ small enough so that $(1 + \varepsilon)\rho \leq 1$. It follows from the above inequality that there exists some x large enough such that $-D'_{\mathfrak{g}}(x) \leq (1+\varepsilon) f_{\mathfrak{b}}(-x)$. Furthermore, for all $x' \geq x$,

$$\begin{aligned} D_{\mathfrak{g}}(x) &= D_{\mathfrak{g}}(x') - \int_x^{x'} D'_{\mathfrak{g}}(y)dy \\ &\leq D_{\mathfrak{g}}(x') + (1+\varepsilon)\int_x^{x'} f_{\mathfrak{b}}(-x)dx \\ &= D_{\mathfrak{g}}(x') - (1+\varepsilon)(F_{\mathfrak{b}}(-x') - F_{\mathfrak{b}}(-x)). \end{aligned}$$

[21] In the case of binary states and actions, Herrera and Hörner [2012] show that the property of increasing hazard ratio is equivalent to the monotonicity of this updating function. See Smith, Sørensen, and Tian [2021] for a general treatment.

Rearranging the above equation,

$$\begin{aligned} D_{\mathfrak{g}}(x) - D_{\mathfrak{g}}(x') &\leq (1+\varepsilon)(F_{\mathfrak{b}}(-x) - F_{\mathfrak{b}}(-x')) \\ &\leq (1+\varepsilon)\rho(x'-x), \end{aligned}$$

where the second last inequality follows from the fact that $f_{\mathfrak{b}}(-x) \leq \rho < 1$. Since $(1+\varepsilon)\rho \leq 1$, the above inequality implies that there exists some x large enough such that $D_{\mathfrak{g}}(x) + x \leq D_{\mathfrak{g}}(x') + x'$ for all $x' \geq x$. That is, $\phi(x)$ eventually increases monotonically. □

Given these lemmas, we are ready to prove the following result. It shows that conditioned on an uninformative source, the possibility of immediate agreement is independent of the prior belief.

Lemma 7. *For any action $a \in \{\mathfrak{g}, \mathfrak{b}\}$, the following statements are equivalent:*

(i) $\mathbb{P}_{0,\pi}[\bar{a} = a] > 0$, *for some prior* $\pi \in (0,1)$*;*
(ii) $\mathbb{P}_{0,\pi}[\bar{a} = a] > 0$, *for all prior* $\pi \in (0,1)$.

Proof. The second implication, namely, $(ii) \Rightarrow (i)$ is immediate. We will show the first implication, $(i) \Rightarrow (ii)$. Fix some prior $\tilde{\pi} \in (0,1)$ such that $\mathbb{P}_{0,\tilde{\pi}}[\bar{a} = \mathfrak{g}] > 0$ and let $\tilde{r} = \log \frac{\tilde{\pi}}{1-\tilde{\pi}}$. Since $r_t^{\mathfrak{g}}(\tilde{r})$ is a deterministic process, the event $\{\bar{a} = \mathfrak{g}\}$ initiated at the prior $\pi_1 = \tilde{\pi}$ is equivalent to the event $\{\ell_t \geq -r_t^{\mathfrak{g}}(\tilde{r}), \forall t \geq 1\}$. Conditioned on $\omega = 0$, since signals are i.i.d., so are the agents' private log-likelihood ratios. Thus, we have

$$\mathbb{P}_{0,\tilde{\pi}}[\bar{a} = \mathfrak{g}] = \prod_{t=1}^{\infty}(1 - F_0(-r_t^{\mathfrak{g}}(\tilde{r})). \tag{1.9.5}$$

As a consequence, $\mathbb{P}_{0,\tilde{\pi}}[\bar{a} = \mathfrak{g}] > 0$ if and only if there exists $M < \infty$ such that

$$-\sum_{t=1}^{\infty} \log\left(1 - F_0(-r_t^{\mathfrak{g}}(\tilde{r}))\right) < M.$$

For two sequences (a_t) and (b_t), we write $a_t \approx b_t$ if $\lim_{t\to\infty}(a_t/b_t) = 1$. Since $r_t^{\mathfrak{g}}(\tilde{r}) \to \infty$ (this follows from part (i) of Lemma 5), $\log\left(1 - F_0(-r_t^{\mathfrak{g}}(\tilde{r}))\right) \approx -F_0(-r_t^{\mathfrak{g}}(\tilde{r}))$. Thus, the above sum is finite if and only if

$$\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}(\tilde{r})) < M. \tag{1.9.6}$$

By the overturning principle (Lemma 2), it suffices to show that (1.9.6) implies that

$$\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}(r)) < M, \quad \text{for any } r \geq 0.$$

By the eventual monotonicity of ϕ (Lemma 6) and the fact that $r_t^{\mathfrak{g}}(\tilde{r}) \to \infty$, we can find a large enough $\bar{t}$ such that $r_{\bar{t}}^{\mathfrak{g}}(\tilde{r}) := \bar{r} \geq 0$ and $\phi(r) \geq \phi(\bar{r})$ for all $r \geq \bar{r}$. By part (ii) of Lemma 5, there exists $t_0 \in \mathbb{N}$ such that $r_{t_0}^{\mathfrak{g}}(r) \geq \bar{r}$ for all $r \geq 0$. Since above $\bar{r}$, ϕ is monotonically increasing, one has $\phi(r_{t_0}^{\mathfrak{g}}(r)) \geq \phi(\bar{r})$ for any $r \geq 0$. Consequently, for all $\tau \geq 1$, $r_{\tau+t_0}^{\mathfrak{g}}(r) = \phi^\tau(r_{t_0}^{\mathfrak{g}}(r)) \geq \phi^\tau(\bar{r}) = r_{\tau+1}^{\mathfrak{g}}(\bar{r})$. Since $r_{\tau+1}^{\mathfrak{g}}(\bar{r}) = r_{\tau+1}^{\mathfrak{g}}(r_{\bar{t}}^{\mathfrak{g}}(\tilde{r})) = r_{\tau+\bar{t}}^{\mathfrak{g}}(\tilde{r})$, it follows that

$$F_0(-r_{\tau+t_0}^{\mathfrak{g}}(r)) \leq F_0(-r_{\tau+\bar{t}}^{\mathfrak{g}}(\tilde{r})).$$

Thus, it follows from (1.9.6) that for any $r \geq 0$, $\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}(r)) < \infty$, as required. The case for action $\mathfrak{b}$ follows from a symmetric argument. □

Now, we are ready to prove Proposition 5.

Proof of Proposition 5. We show that $(i) \Rightarrow (ii)$, $(ii) \Rightarrow (iii)$, and $(iii) \Rightarrow (i)$. To show the first implication, we prove the contrapositive statement. Suppose that $\mathbb{P}_0[S < \infty] > 0$. This implies that there exists a sequence of action realizations $(b_1, b_2, \ldots, b_{k-1}, b_k = \ldots = a)$ for some action $a \in \{\mathfrak{b}, \mathfrak{g}\}$ such that

$$\mathbb{P}_0[a_t = b_t, \forall t \geq 1] > 0.$$

By stationarity, there exists some $\pi' \in (0, 1)$ such that

$$\mathbb{P}_{0,\pi'}[\bar{a} = a] > 0,$$

which contradicts (i).

To show the second implication, suppose towards a contradiction that there exists some action $a \in \{\mathfrak{g}, \mathfrak{b}\}$ such that $\mathbb{P}_{0,\pi}[\bar{a} = a] > 0$ for all prior $\pi \in (0, 1)$. In particular, it holds for the uniform prior. Since the event $\{\bar{a} = a\}$ is contained in the event $\{S < \infty\}$,

$$0 < \mathbb{P}_0[\bar{a} = a] \leq \mathbb{P}_0[S < \infty].$$

This implies that $\mathbb{P}_0[S = \infty] < 1$, a contradiction to (ii).

Finally, we show the last implication by contraposition. Suppose that there exists some $a \in \{\mathfrak{g}, \mathfrak{b}\}$ such that $\mathbb{P}_{0,\pi}[\bar{a} = a] > 0$ for some prior $\pi \in (0, 1)$. By Lemma 7, it also holds for all prior $\pi \in (0, 1)$, which is a contradiction to (iii). This concludes the proof of Proposition 5. □

Proof of Proposition 4. By the equivalence between (ii) and (iii) in Proposition 5 and Proposition 1, we have shown Proposition 4. □

Given Proposition 4, we are now ready to prove our main result.

Proof of Theorem 1. By part (i) of Lemma 4, $\mathbb{P}_{1,\mathfrak{b}}[\bar{a} = \mathfrak{g}] = 0$ and $\mathbb{P}_{1,\mathfrak{g}}[\bar{a} = \mathfrak{b}] = 0$. Following a similar argument that led to (1.9.5), one has

$$0 = \mathbb{P}_{1,\mathfrak{b}}[\bar{a} = \mathfrak{g}] = \prod_{t=1}^{\infty}(1 - F_{\mathfrak{b}}(-r_t^{\mathfrak{g}})).$$

Taking the logarithm on both sides, the above equation is equivalent to $-\sum_{t=1}^{\infty}\log(1-F_{\mathfrak{b}}(-r_t^{\mathfrak{g}})) = \infty$. Since $r_t^{\mathfrak{g}} \to \infty$, $\log(1 - F_{\mathfrak{b}}(-r_t^{\mathfrak{g}})) \approx -F_{\mathfrak{b}}(-r_t^{\mathfrak{g}})$ and the previous sum is infinite if and only if

$$\sum_{t=1}^{\infty} F_{\mathfrak{b}}(-r_t^{\mathfrak{g}}) = \infty. \tag{1.9.7}$$

Similarly, we have that $\mathbb{P}_{1,\mathfrak{g}}[\bar{a} = \mathfrak{b}] = 0$ if and only if $\sum_{t=1}^{\infty}(1 - F_{\mathfrak{g}}(-r_t^{\mathfrak{b}})) = \infty$, where $r_t^{\mathfrak{b}}$ denotes the deterministic process of r_t conditioned on the event $\{\bar{a} = \mathfrak{b}\}$. By symmetry, $r_t^{\mathfrak{b}} = -r_t^{\mathfrak{g}}$ for all $t \geq 1$. Hence, $\mathbb{P}_{1,\mathfrak{g}}[\bar{a} = \mathfrak{b}] = 0$ if and only if

$$\sum_{t=1}^{\infty}(1 - F_{\mathfrak{g}}(r_t^{\mathfrak{g}})) = \infty. \tag{1.9.8}$$

Suppose that the uninformative signals have fatter tails than the informative signals. By definition, there exists $\varepsilon > 0$ such that for all x large enough, $F_0(-x) \geq \varepsilon \cdot F_{\mathfrak{b}}(-x)$ and $1 - F_0(x) \geq \varepsilon \cdot (1 - F_{\mathfrak{g}}(x))$. It then follows from (1.9.7) and (1.9.8) that

$$\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}) = \infty \quad \text{and} \quad \sum_{t=1}^{\infty}(1 - F_0(r_t^{\mathfrak{g}})) = \infty.$$

Using the same logic we used to deduce (1.9.7) and (1.9.8), having these two divergent sums is equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{g}] = 0$ and $\mathbb{P}_0[\bar{a} = \mathfrak{b}] = 0$. Thus, by Proposition 4, asymptotic learning holds.

By part (ii) of Lemma 4, there exist $\pi, \pi' \in (0,1)$ such that $\mathbb{P}_{1,\mathfrak{g},\pi}[\bar{a} = \mathfrak{g}] > 0$ and $\mathbb{P}_{1,\mathfrak{b},\pi'}[\bar{a} = \mathfrak{b}] > 0$. Let $r = \log\frac{\pi}{1-\pi}$ and $r' = \log\frac{\pi'}{1-\pi'}$. Following a similar argument that led to (1.9.6), these are equivalent to

$$\sum_{t=1}^{\infty} F_{\mathfrak{g}}(-r_t^{\mathfrak{g}}(r)) < \infty \quad \text{and} \quad \sum_{t=1}^{\infty}(1 - F_{\mathfrak{b}}(r_t^{\mathfrak{g}}(r'))) < \infty.$$

Now, suppose that the uninformative signals have thinner tails than the informative signals. By definition, there exists $\varepsilon > 0$ such that either (i) $F_0(-x) \leq (1/\varepsilon) \cdot F_{\mathfrak{g}}(-x)$ for all x large enough, or (ii) $1 - F_0(x) \leq (1/\varepsilon) \cdot (1 - F_{\mathfrak{b}}(x))$ for all x large enough. It then follows from the above inequalities that either (i) $\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}(r)) < \infty$ or (ii) $\sum_{t=1}^{\infty}(1 - F_0(r_t^{\mathfrak{g}}(r')))$. Equivalently, we have either (i) $\mathbb{P}_{0,\pi}[\bar{a} = \mathfrak{g}] > 0$ for some $\pi \in (0,1)$ or (ii) $\mathbb{P}_{0,\pi'}[\bar{a} = \mathfrak{b}] > 0$ for some $\pi' \in (0,1)$. By Lemma 7, these also hold for all prior $\pi, \pi' \in (0,1)$, including the uniform prior. So we have either (i) $\mathbb{P}_0[\bar{a} = \mathfrak{g}] > 0$ or (ii) $\mathbb{P}_0[\bar{a} = \mathfrak{b}] > 0$. Thus, by Proposition 4, asymptotic learning fails. □

Gaussian Private Signals

In this section, we prove Corollary 1 and Theorem 2. We consider the canonical environment in which all signal distributions $\mu_{\mathfrak{g}}$, $\mu_{\mathfrak{b}}$ and μ_0 are normal. In particular, the distributions $\mu_{\mathfrak{g}}$ and $\mu_{\mathfrak{b}}$ share the same variance σ^2 and have mean $+1$ and -1, respectively. Meanwhile, μ_0 has mean $m_0 \in (-1, 1)$ and variance τ^2. Note that in this case, the agent's private log-likelihood ratio induced by a signal s_t is

$$\ell_t = \log \frac{f_{\mathfrak{g}}(s_t)}{f_{\mathfrak{b}}(s_t)} = \frac{2}{\sigma^2} s_t. \tag{1.9.9}$$

Since ℓ_t is proportional to s_t, the distributions F_θ and F_0 are also normal, with a variance of $4/\sigma^2$ and $4\tau^2/\sigma^4$, respectively.

Similar to Definition 2, for any pair of CDFs (F, G), we say that F has a *fatter left tail* than G if there exists $\varepsilon > 0$ such that $F(-x)/G(-x) \geq \varepsilon$ for all x large enough. Likewise, we say that F has a *thinner left tail* than G if there exists $\varepsilon > 0$ such that $F(-x)/G(-x) \leq 1/\varepsilon$ for all x large enough. We define the concepts of fatter and thinner right tails analogously. The following lemma demonstrates that for Gaussian signals, the relative thickness of the tails depends solely on their relative variances.

Lemma 8. *Suppose F and G are two Gaussian cumulative distribution functions with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. If $\sigma_1 > \sigma_2$, then F has fatter left and right tails than G. Meanwhile, G has thinner left and right tails than F.*

Proof. Let f and g denote the probability density functions of F and G. Their ratio evaluated at $x \in \mathbb{R}$ is

$$\frac{f(x)}{g(x)} = \frac{\sigma_2}{\sigma_1} \exp\left((\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}) \frac{x^2}{2} + (\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}) x + \frac{1}{2} (\frac{\mu_2^2}{\sigma_2^2} - \frac{\mu_1^2}{\sigma_1^2}) \right).$$

Suppose $\sigma_1 > \sigma_2$. It follows from the above equation that $\lim_{x\to\infty} \frac{f(-x)}{g(-x)} = \infty$ and $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty$. This clearly implies that $F(-x) \geq G(-x)$ and $1-F(x) \geq 1-G(x)$ for all x large enough. So by definition, F has fatter left and right tails than G, and conversely, G has thinner left and right tails than F. □

As a consequence, Corrollary 1 follows directly from Lemma 8 and Theorem 1.

Proof of Corollary 1. If $\tau > \sigma$, then by (1.9.9), F_0 has a strictly higher variance than both $F_\mathfrak{g}$ and $F_\mathfrak{b}$. By Lemma 8, F_0 has fatter left and right tails than both $F_\mathfrak{g}$ and $F_\mathfrak{b}$. Hence, by Definition 2, the uninformative signals have fatter tails and it thus follows from Theorem 1 that asymptotic learning holds. An identical argument applies to the case where $\tau < \sigma$. □

We henceforth focus on the case where $\tau = \sigma$. When all private signals are Gaussian, Hann-Caruthers, Martynov, and Tamuz [2018] show that one can approximate the sequence $r_t^\mathfrak{g}$ by $(2\sqrt{2}/\sigma) \cdot \sqrt{\log t}$ for all t large enough (see their Theorem 4):

$$\lim_{t\to\infty} \frac{r_t^\mathfrak{g}}{(2\sqrt{2}/\sigma) \cdot \sqrt{\log t}} = 1. \tag{1.9.10}$$

Given this approximation and Proposition 4, we are ready to prove Theorem 2.

Proof of Theorem 2. In this proof we use the Landau notation, so that $O(g(t))$ stands for some function $f : \mathbb{N} \to \mathbb{R}$ such that there exists a positive $M \in \mathbb{R}$ and $t_0 \in \mathbb{N}$ such that $|f(t)| \leq M \cdot g(t)$ for all $t \geq t_0$.

Note that by (1.9.9), we can write

$$F_0(-r_t^\mathfrak{g}) = \mathbb{P}_0[\ell_t \leq -r_t^\mathfrak{g}] = \mathbb{P}_0[s_t \leq -(\sigma^2/2) \cdot r_t^\mathfrak{g}].$$

By (1.9.10), we have that for all t large enough,

$$F_0(-r_t^\mathfrak{g}) = \mathbb{P}_0[s_t \leq -\sigma\sqrt{2\log t}] =: \mu_0(-\sigma\sqrt{2\log t}),$$

where μ_0 is the CDF of s_t conditioned on $\omega = 0$. Since μ_0 is the normal distribution with mean $m_0 \in (-1, 1)$ and variance σ^2, observe that $\mu_0(x) = \frac{1}{2}\,\mathrm{erfc}(-\frac{x-m_0}{\sigma\sqrt{2}})$, where $\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \mathrm{e}^{-t^2} dt$ is the complementary error function.

Applying a standard asymptotic expansion of the complementary error function, i.e., $\mathrm{erfc}(x) = \frac{e^{-x^2}}{x\sqrt{\pi}} + O(\mathrm{e}^{-x^2}/x^3)$, we obtain that for all t large enough,

$$\mu_0(-\sigma\sqrt{2\log t}) = \frac{\mathrm{e}^{-\left(\frac{m_0}{\sigma}\sqrt{2\log t} + \frac{m_0^2}{2\sigma^2}\right)}}{t(\sqrt{\pi\log t} + \delta \cdot m_0)} + O\left(\frac{\mathrm{e}^{-m_0\sqrt{2\log t}}}{t(\sigma\sqrt{2\log t} + m_0)^3}\right), \tag{1.9.11}$$

where $\delta > 0$ is a constant.

Case (i): suppose $m_0 = 0$. Then (1.9.11) becomes $\frac{1}{t\cdot\sqrt{\pi \log t}} + O(\frac{1}{t\cdot(\log t)^{3/2}})$. Since the series $\frac{1}{t \log t}$ is divergent and $\frac{1}{t \log t} \leq \frac{1}{t\sqrt{\log t}}$ for all $t \geq 2$, the sum of the first term also diverges. Hence, the sum of (1.9.11) diverges, which implies $\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}) = \infty$. By the same approximation argument used in the proof of Theorem 1, this is equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{g}] = 0$. Using an analogous argument, we obtain $\mathbb{P}_0[\bar{a} = \mathfrak{b}] = 0$. Together, by Proposition 4, we conclude that asymptotic learning holds.

Case (ii): suppose $m_0 \neq 0$ and $m_0 \in (-1, 1)$. Let $c = \frac{m_0\sqrt{2}}{\sigma}$. By the change of variable $x = \sqrt{\log t}$,

$$\int_2^{\infty} \frac{\mathrm{e}^{-c\sqrt{\log t}}}{t\sqrt{\log t}} dt = 2\int_{\sqrt{\log 2}}^{\infty} \mathrm{e}^{-cx} dx.$$

If $m_0 > 0$, then $c = \frac{m_0\sqrt{2}}{\sigma} > 0$. By the integral test, the sum in (1.9.11) converges, and thus $\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}) < \infty$. Again, by the same logic that we use to deduce (1.9.6), this is equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{g}] > 0$. If $m_0 < 0$, it follows from a symmetric argument that $\sum_{t=1}^{\infty}(1 - F_0(r_t^{\mathfrak{g}})) < \infty$, which is equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{b}] > 0$. In either case, it follows from Proposition 4 that asymptotic learning fails. $\square$

Chapter 2

LEARNING IN REPEATED INTERACTIONS ON NETWORKS

2.1 Introduction

We study social learning by long-lived agents who observe each others' actions on a social network. We show that information aggregation fails: the speed of learning stays bounded even in large networks, where efficient aggregation of all private information would lead to arbitrarily fast learning. Methodologically, we introduce new techniques that allow us to relax commonly made assumptions and study general networks, with forward-looking, Bayesian agents who interact repeatedly. Repeated interactions can, for example, describe the exchange of opinions and information among friends on social media or firms that learn from each others actions. Arguably, social learning driven by such interactions can be an important determinant in many choice domains, such as investments, health insurance, schools, technology adoption, or where to live and work.

More formally, we consider a group of agents who repeatedly interact with their neighbors on a network. There is a fixed but unknown state of the world, taking values from a finite set. In every period, each agent receives a private signal about the state and observes all past actions of his neighbors. Based on this information, each agent updates his beliefs, chooses one of finitely many possible actions, and receives a flow payoff that depends on his action and the state (but not on the actions of others). We consider both myopic agents who maximize their instantaneous payoff, as well as strategic agents who are forward-looking, and exponentially discount the future. Since strategic agents care about their future utilities, they may sacrifice their present flow utilities and choose actions that induce others to behave in a way which reveals more information in the future.

The constant influx of private information in our model allows every agent to eventually learn the truth and choose the optimal action. Thus, we focus on how fast agents learn. If the number of agents doubles, the number of private signals available to society also doubles. Hence, if information were aggregated efficiently, the time that it would take for agents to choose correctly (say, with a given high probability) would decrease by a factor of two. In other words, when information aggregation is efficient, the speed of learning increases linearly with the number of agents. The question we ask is: what is the speed of learning in equilibrium, and how does it depend on the number of agents, the structure of the network, the agents' utilities, the signals, and the discount factor?

We focus on *strongly connected* networks, where there is an observational path between every pair of agents, since otherwise efficient aggregation of information is excluded by the lack of informational channels. This is a mild assumption, and follows, for example, from the "six degrees of separation rule", that stipulates that there is a path of length at most six between every two members of a social network.[1] Our main result (Theorem 3) shows that the speed of learning does not increase linearly with the number of agents, and is, in fact, bounded from above by a constant that only depends on the private signal distribution, and is independent of the structure of the network and the remaining parameters of the model.

For example, consider agents who, in each period, observe an independent binary signal that is equal to a binary state with probability 0.9. Then, regardless of the number of agents and network structure, the speed of learning never exceeds ten times the speed at which an agent learns on his own. This is despite the fact that if n agents shared their signals publicly, they would learn n times as fast. Thus, a society of 1,000 agents who observe their neighbors' past actions does not learn faster than a society of ten agents in which information is efficiently aggregated. This means that in a society of 1,000 agents, at least 99% of the information generated by the private signals is lost for any structure of the observational network (and any equilibrium played by the agents with any utility).

As another illustration of this result, consider the finite two dimensional grid graph: The set of agents is $\{1, \ldots, n\}^2$, and i observes j if and only if $|i - j| = 1$. Regardless of the size of the graph, in the early periods each agent is exposed to little information, since (at most) four other agents are observed each period. Theorem 3 shows that even later in the process the size of the graph does not matter substantially, as the speed of learning is bounded.

The mechanism behind this bounded speed of learning is as follows: First, in a strongly connected network all agents learn at the same rate, as each agent could guarantee himself the same learning rate as any of his neighbors. We establish the bound on the learning rate by contradiction. Suppose that agents learn at a rate

[1] The science fiction writer Frigyes Karinthy proposed this rule in his 1929 short story "Láncszemek" (in English, "Chains"). The rule was confirmed empirically on a number of online social network data sets [Leskovec and Horvitz, 2008, Watts and Strogatz, 1998].

that is higher than the rate that their individual private signals can support. This implies that social information, which consists of agents' actions, will become much more precise than each agent's private information. As a result, agents will ignore their private signals and only rely on the social information they observe from their neighbors. This implies that agents' actions no longer reveal any information about the state, so social information cannot grow too precise over time. This contradicts our previous hypothesis, and we conclude that agents cannot learn too fast.

The failure of information aggregation suggested by Theorem 3 is an asymptotic result that describes how fast learning is in late periods. We complement this result with a numerical calculation for the first ten periods on a simple graph: the line graph with bidirectional observations. Assuming myopic agents, and binary state, actions and signals, we find that our asymptotic lower bound on mistake probabilities holds also in the early periods. Quantitatively, in each of the first 10 periods agents choose correctly with a probability that is smaller than that of a group of five agents who share their private signals. This holds independently of the number of agents in the network.

We contribute to the social learning literature in three aspects. First, instead of focusing on short-lived agents who act only once, we consider a more realistic model in which agents are long-lived and repeatedly interact with each other. Second, we extend existing models of social learning from myopic agents to strategic agents who discount their future utilities at a common rate. Analyzing strategic agents is complicated since these agents may have an incentive to choose a sub-optimal action today to learn more information from the actions of others in the future, whereas such an incentive is completely shut down for myopic agents. Third, we generalize previous work on the speed of learning on the complete network where all agents observe each other [Harel et al., 2021] to general social networks where agents only observe their neighbors.

Related Literature

There are few papers addressing repeated interaction between rational agents and its role in information aggregation. This is because it is challenging to analyze the evolution of beliefs of long-lived agents, particularly when these beliefs are influenced by the interactions between them. Most of the literature has focused on either short-lived agents who only act once [Acemoglu et al., 2011, Banerjee, 1992, Bikhchandani et al., 1992, Smith and Sørensen, 2000] or non-fully-rational agents

[Bala and Goyal, 1998, Molavi et al., 2018] and heuristic learning rules [Dasaratha et al., 2020, DeGroot, 1974, Golub and Jackson, 2010]. Nevertheless, as many real-life situations involve repeated interactions, it is natural to study models that allow these interactions. It is likewise interesting to understand rational (and potentially forward-looking) agents as this is an important benchmark case to assess whether or to what extent failures of information aggregation are driven by a lack of patience or lack of rationality of the agents.

Among models which consider repeated interactions, the bandit literature studies endogenous information acquisition and the resulting free rider problem when multiple agents try to learn the state from each other's public signals [Bolton and Harris, 1999, Heidhues et al., 2015, Keller and Rady, 2010, Keller et al., 2005]. Bala and Goyal [1998] extend this to a social network setting for non-fully-rational agents who do aggregate the results of their neighbors' experiments but disregard the information contained in their choices. They examine how the geometry of the social network affects learning outcomes.

Focusing on a repeated interaction setting in which agents and have no experimentation motives, Mossel et al. [2014] consider rational but myopic agents who observe each other's actions. They give conditions for learning to occur on infinite undirected graphs. Mossel et al. [2015] further generalize their setting to allow for forward-looking agents.[2] Unlike our model, agents in these models only receive one signal at the beginning of time. They do not study the speed of learning, and instead focus on identifying the types of social networks in which learning always occurs.

Complementing the previous literature, we take the next natural step: we ask how fast learning occurs and study its relationship with the size of the network. Furthermore, we consider agents with any discount factor with myopic agents as a special case. To our best knowledge, this is the first paper to consider social learning in a network setting with fully-rational agents who interact repeatedly.

A recent paper that considers rational agents in a repeated setting is by Harel, Mossel, Strack, and Tamuz [2021], who study the speed of learning when all agents are myopic and observe each other. They show that similar to our main result, for any number of agents, the speed of learning from actions is bounded above by a constant.

[2] In a similar setting, Migrow [2022] shows that for two forward looking agents who observe each other there does not exist an equilibrium where agents behave myopically, under some assumptions on the signal structure.

Their proofs rely crucially on the symmetry inherent in the complete network in which all agents observe each other and actions are common knowledge. Their analysis relies on a phenomenon called "groupthink": a feedback loop that develops when all players take the wrong action, observe that everyone else also took the same action, become more confident in their wrong beliefs, and then again take the wrong action. In incomplete networks, actions are no longer common knowledge, making the techniques used in the aforementioned paper inapplicable in our setting. The techniques we introduce furthermore allow us to consider non-myopic agents and multiple states of the world.

Our main insight is that learning cannot be too fast since fast individual learning would cause agents to ignore their private signals. This is reminiscent of the information cascades that drive the failure of information aggregation in the classical herding literature [Banerjee, 1992, Bikhchandani et al., 1992, Smith and Sørensen, 2000]. However, the force in our model affects the speed of learning rather than determining whether or not information aggregates, and because of the repeated interactions between agents, it requires a very different analysis. A similar insight also appears in the earlier literature on rational expectations in financial markets, where it implies the breakdown of the efficient market hypothesis [Grossman and Stiglitz, 1980]: prices cannot fully reflect all available information, precisely because if they did, it would eliminate the agents' incentive to respond to their private information, contradicting the assumption that prices contain all information. This is known as the Grossman-Stiglitz paradox.

Following the herding literature, our paper studies the friction that arises for information aggregation when actions are coarse and thus do not fully reveal beliefs. Another strand of the literature shows that information aggregation may still be slow, even with a continuous action space. For example, Vives [1993] considers a Cournot competition model with a common unknown production cost among a continuum of firms. He shows that noisy observations of past actions (through market prices) slow down the speed of learning. Although his environment is different from ours,[3] the force behind his slow speed of learning is related to ours: by examining the asymptotic behavior of public information's informativeness (reflected in prices) and agents' responsiveness to their private information, he found that as public information becomes more informative, agents respond less to their private

[3] More specifically, his setting differs from ours along a number of dimensions: Firms learn from public prices, which are noisy observations of the average actions of others. Furthermore, there is no network and since there is a continuum of agents, any strategic incentive is also abstracted away.

information, leading to a slower increase in how informative the public information can be. Put in his words, "... information revelation through the price system can be slow precisely because it is successful." We contribute to this literature by showing that this intuition in these specific cases generalizes extensively, especially in a general network setting with forward-looking agents.

With a rich-enough action space that fully reveals agents' beliefs, Dasaratha and He [2019] study how different generational networks affect the learning rate with Gaussian signals and continuous actions. They find that learning is slow for large symmetric overlapping generations networks with a uniform bound on the number of signals aggregated per generation. The main mechanism behind the inefficiency is a confounding effect: as earlier generations observe common predecessors, their actions are correlated, reducing the amount of information transmitted to the next generation. This effect is inherent in their overlapping generations network structure, in which information travels unidirectionally. Since information travels bidirectionally in our model, higher-order beliefs pose an obstacle that is not present in Vives [1993] and Dasaratha and He [2019].

2.2 Model

Let $N = \{1, 2, \dots, n\}$ be a finite set of agents. Time is discrete and the horizon is infinite, i.e. $t \in \{1, 2, \dots\}$. In every time period t, each agent i has to choose an action a_t^i from a finite set A. There is an unknown state of the world ω, taking values in a finite set Ω, that is chosen at period 0 and does not change over time. Each state occurs with strictly positive probability. We denote by $\mathfrak{g}, \mathfrak{f}$ generic elements of Ω.

Utility and Optimal Actions

The flow utility for choosing an action $a \in A$ when the state is $\mathfrak{g} \in \Omega$ is $u(a, \mathfrak{g})$ for some $u \colon A \times \Omega \to \mathbb{R}$. We assume that for each state $\mathfrak{g} \in \Omega$ there is a unique action $a^{\mathfrak{g}} \in A$, which maximizes $u(\cdot, \mathfrak{g})$, the utility in that state:

$$\{a^{\mathfrak{g}}\} = \arg\max_{a \in A} u(a, \mathfrak{g}).$$

We also assume that these optimal actions are distinct, i.e. if $\mathfrak{g} \neq \mathfrak{f}$ then $a^{\mathfrak{g}} \neq a^{\mathfrak{f}}$. These assumptions facilitate learning from actions in the sense that observing the optimal action of an agent who knows the state allows other agents to infer the state. Note that these assumptions hold for any generic utility u, as long as the set A has at least as many elements as Ω. When there are two states then this assumption is necessary to make the model non-trivial, since otherwise there is a dominant action which will always be played in every equilibrium.[4]

An important example which the reader may wish to keep in mind is the case of binary states and actions, and where the agent aims to match the state: $\Omega = \{\mathfrak{g}, \mathfrak{f}\}$, $A = \{a^{\mathfrak{g}}, a^{\mathfrak{f}}\}$ with $u(a^{\mathfrak{g}}, \mathfrak{g}) = u(a^{\mathfrak{f}}, \mathfrak{f}) = 1$ and $u(a^{\mathfrak{g}}, \mathfrak{f}) = u(a^{\mathfrak{f}}, \mathfrak{g}) = 0$. This setting already features all the forces and tensions of the general case, and likewise offers the same conclusions.

Agents' Information

In each period t, agent i receives a private signal s_t^i drawn from a finite set S_t^i. Conditional on the state ω, signals s_t^i are independent across agents and time, with distribution $\mu_t^{i,\omega} \in \Delta(S_t^i)$. For distinct $\mathfrak{f}, \mathfrak{g} \in \Omega$, the distributions $\mu_t^{i,\mathfrak{g}}$ and $\mu_t^{i,\mathfrak{f}}$ are distinct and mutually absolutely continuous, so that no signal excludes any state or perfectly reveals the state. Thus, the log-likelihood ratio of any signal s

$$\ell_t^{i,\mathfrak{g},\mathfrak{f}}(s) = \log \frac{\mu_t^{i,\mathfrak{g}}(s)}{\mu_t^{i,\mathfrak{f}}(s)}$$

is well defined.

We focus on *bounded signals* in the sense that the private belief induced by any signal cannot be arbitrarily strong. More specifically, we assume that there exists a constant $M > 0$ that bounds the absolute value of the likelihood induced by any signal:

$$M = 2 \sup_{\mathfrak{f},\mathfrak{g},i,t,s} \left| \ell_t^{i,\mathfrak{g},\mathfrak{f}}(s) \right|. \tag{2.2.1}$$

We make this assumption for tractability and discuss its relaxation in the conclusion. We allow signals to depend on calendar time and the agents' identities to highlight the robustness of our results. However, to understand our main economic insight, it suffices to think of the setting where all signals are i.i.d. across agents and time, as is typically assumed in the literature.

For each agent i there is a subset of agents $N_i \subseteq N$ who are his *social network neighbors*, and whose actions he observes. We include $i \in N_i$ since agent i observes his own actions. The information available to agent i at time t, before taking his

[4]For three states or more this assumption is more restrictive. Indeed, our Lemma 9, which shows that agents eventually choose myopic actions, may not hold without it. This is because even when beliefs concentrate around a state, multiple actions can be optimal for particular likelihood ratios about the other states, and so non-myopic behavior can persist.

action a_t^i, thus consists of a sequence of private signals $(s_1^i, \cdots, s_t^i)$ and the history of actions observed by i,

$$H_t^i = \{a_s^j : s < t, j \in N_i\} .$$

Let $\mathcal{I}_t^i = S_1^i \times \cdots \times S_t^i \times A^{|N_i| \times (t-1)}$ so that the private history $I_t^i = (s_1^i, \cdots, s_t^i, H_t^i)$ is an element of $\mathcal{I}_t^i$.

We assume that the social network is *strongly connected*: there is an observational path between every pair of agents (we relax this assumption in §2.6).[5,6] This assumption avoids situations in which efficient aggregation of information is precluded because there is no channel for information to travel from i to j. Furthermore, we assume that agents observe their neighbors in every period. This assumption is purely to simplify notation; we discuss the case where an agent observes their neighbors only at intermittent and potentially random times in Appendix 2.9.

Strategies and Payoffs

A pure strategy of agent i at time t is a function $\sigma_t^i : \mathcal{I}_t^i \to A$. A pure strategy of agent i is a sequence of functions $\sigma^i = (\sigma_1^i, \sigma_2^i, \cdots)$ and a pure strategy profile is a collection of pure strategies of all agents, $\sigma = (\sigma^i)_{i \in N}$. We write $\sigma = (\sigma^i, \sigma^{-i})$ for any agent $i \in N$, where σ^{-i} denotes the pure strategies of all agents other than i. Given a pure strategy profile σ, the action of agent i at time t is $a_t^i(\sigma) = \sigma_t^i(I_t^i)$. The flow utility of agent i at time t is

$$u(a_t^i(\sigma), \omega).$$

Agents do not observe their flow utilities. Nevertheless, one can incorporate observations of utilities into the private signals that agents receive. A special case of our model is one in which the agent receives an observed payoff $v(a_t^i, s_t^i)$ that depends on the action and the signal realization. In this case, the flow utility corresponds to the expected observed payoff $u(a, \omega) = \sum_s \mu^\omega(s) \cdot v(a, s)$. The agent always observes their payoff as it is only a function of their signal; yet, from an ex-ante perspective, the situation is identical to that of our setting. The assumption of unobserved flow utilities is commonly made in the literature to model learning without an experimentation motive. Indeed, any learning situation without an experimentation

[5] Formally, for each $i, j \in N$ there is a sequence $i = i_1, i_2, \ldots, i_k = j$ such that $i_2 \in N_{i_1}, i_3 \in N_{i_2}, \ldots i_k \in N_{i_{k-1}}$.

[6] In a directed Erdős–Rényi graph with n agents, when the expected number of neighbors is $\ln n + c$ for large c, then with high probability the graph will be strongly connected, and so our assumption will be satisfied with high probability [see Theorem 5 in Graham and Pike, 2008].

motive can be reduced to such a situation [see, e.g., the discussion in Rosenberg et al., 2009, §2.1, p. 981].

We assume that agents discount their future utilities at a common rate $\delta \in [0,1)$. The expected utility of agent i under strategy profile σ is thus

$$u^i(\sigma) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\mathbb{E}[u(a_t^i(\sigma),\omega)].$$

We call agents with $\delta = 0$ *myopic* and agents with $\delta > 0$ *strategic*. Myopic agents fully discount future payoffs and choose their actions in each period to maximize expected flow utilities.

Regardless of whether agents are strategic or myopic, the sole benefit of observing others' past actions is to learn about the state. That is, others' actions reflect the signals they receive, and thus observing others' actions can help agents make better inferences about the state. This pure informational motive is an important feature of the model: each agent's flow utility depends only on his own actions and the state, and it is independent of the actions of the others.

Equilibrium

This is a game of incomplete information, in which agents may have different information regarding the underlying unknown state and the actions of others. We use Nash equilibrium as our equilibrium concept and refer to it as equilibrium thereafter.[7] The existence of a (mixed) equilibrium is guaranteed in this game by standard arguments, since, in the product topology on strategies, the space of strategies is compact and utilities are continuous. We note here that every mixed equilibrium can be mapped to a behaviorally equivalent pure equilibrium by adding to each agent's private signal an additional component that is independent of the state and all other signals, and assuming that the agent uses this signal to randomly choose between actions.[8] As our results will only depend on the information about the state contained in the signal, it thus suffices to establish them for pure strategy equilibria, to show that they hold for all (pure and mixed) equilibria.

[7] Our results, which apply to all Nash equilibria, thus also apply to any refinement of Nash equilibrium such as sequential equilibrium.

[8] Formally, for any signal space S_t^i we can consider $\tilde{S}_t^i = S_t^i \times A^{|\mathcal{I}_t^i|}$ with the signal distributions $\tilde{\mu}_t^{i,\mathfrak{g}}$ equal to the product measure of $\mu_t^{i,\mathfrak{g}}$ and $|\mathcal{I}_t^i|$ independent random variables each taking each value $a \in A$ with probability $\mathbb{P}[a_t^i = a|I_t^i]$. As this transformation does not affect the informational content of the signals it leaves the constant M unchanged. Furthermore, we can replicate any behaviorally mixed strategy by the pure strategy that takes the action a if and only if the entry of the second component corresponding to the private history I_t^i equals a.

As usual, a pure strategy profile σ is an equilibrium if no agent can obtain a strictly higher expected utility by unilaterally deviating from σ. That is, a pure strategy profile σ is an *equilibrium* if for all agents i, and all strategies τ^i

$$u^i(\sigma^i, \sigma^{-i}) \geq u^i(\tau^i, \sigma^{-i}) .$$

Speed of Learning

We say that agent i chooses correctly at time t if $a_t^i = a^\omega$, i.e., if the agent chose the action that is optimal given the state. We measure the speed of learning of agent i by the asymptotic rate at which he converges to the correct action [see, e.g., Hann-Caruthers et al., 2018, Harel et al., 2021, Molavi et al., 2018, Rosenberg and Vieille, 2019, Vives, 1993]. Formally, the *speed of learning* of agent i is

$$\liminf_{t \to \infty} -\frac{1}{t} \log \mathbb{P}[a_t^i \neq a^\omega]. \tag{2.2.2}$$

If this limit exists and is equal to r, then the probability of mistake at large times t is approximately e^{-rt}. As we explain below in §2.3, this is the case for the benchmark case of a single agent who receives conditionally i.i.d. private signals at each period.

2.3 The Public Signals Benchmark

As a benchmark, we briefly discuss the case of *public* signals for a single or multiple agents. We also assume that signals are i.i.d. across time and agents, with $\mu^{\mathfrak{g}} = \mu_t^{i,\mathfrak{g}}$ for all $\mathfrak{g} \in \Omega$. In the single-agent case, a classical large deviations argument shows that the limit $r_a = \lim_{t\to\infty} -\frac{1}{t} \log \mathbb{P}[a_t^i \neq a^\omega]$ exists and is positive.[9] Note that the fact that the limit is positive implies that the agents learn the state; the probability of choosing incorrectly tends to zero. This is a consequence of the assumption that the measures $\mu^{\mathfrak{g}}$ are distinct which ensures that signals are informative.

Next, consider the case where each of n agents observes all n independent public signals in each period, as well as their neighbors' past actions. As actions contain no additional information about the state relative to the signals, this situation is identical to the single-agent case, except that now each agent receives n independent signals at each period. An agent in period t will thus have observed exactly as many signals as a single agent in autarky in period $n \cdot t$. It thus follows from the single-agent case that when signals are public, the speed of learning for n agents is $n \cdot r_a$.

[9]For textbook treatments see, e.g., pp. 380-384 in Cover and Thomas [2006] for the binary state case or, for the general finite state case, Theorem 2.2.30 in Dembo and Zeituni [2009]. See Moscarini and Smith [2002] for an application in economics to single agent decision problems, and Frick et al. [2022] for an application to a multi-agent setting.

These results for the case of public signals immediately bound the speed of learning in the private signals case: In any social network, observed actions contain weakly less information than the private signals. Thus, $n \cdot r_a$ is an upper bound to the speed of learning for any network with n agents and private signals.

2.4 Results

We now state our main result. It turns out that in a strongly connected network, all agents learn at the same speed (by Lemma 10 in §3.4), and we call this common speed of learning *the equilibrium speed of learning*. In contrast to the public signal case, our main result shows that regardless of the size of the network, the equilibrium speed of learning is bounded above by a constant. Recall that in (2.2.1) we defined $M/2$ to be the maximal log-likelihood ratio induced by any signal.

Theorem 3. *The equilibrium speed of learning is at most M, in any equilibrium, on any social network of any size, for any discount factor $\delta \in [0, 1)$, and any utility u.*

Perhaps surprisingly, Theorem 3 shows that adding more agents (thus more information) to the network and expanding the network cannot improve the speed of learning beyond some bound, which is twice the strength of the strongest possible signal, as measured in log-likelihood ratios. Indeed, this upper bound on the learning speed implies that more and more information is lost as the size of the network increases.

For example, for a binary state, binary actions and independent binary signals that are equal to the state with probability 0.9, the speed of learning in a social network of any size is bounded by that of ten agents who observe each other's signals directly.[10] Consequently, in any social network, even if there are 1,000 agents who observe their neighbors' past actions, they cannot learn faster than a group of ten agents who share their private signals. Equivalently, their speed of learning cannot be more than ten times that of a single agent. Thus almost all of the private information in large networks is lost, resulting in inefficient information aggregation.

The idea behind our proof of Theorem 3 is as follows. Intuitively, one might think that larger networks would boost the speed of learning as agents acquire more and more information from their neighbors, as well as indirectly from their

[10] To see this, given the signal distribution, we calculate the speed of learning in the single-agent case, which is approximately equal to 0.51 (see Harel et al. [2021] for exact expressions for the speed of learning in the binary state case). As discussed in §2.3, the learning speed in a network of ten agents with public signals is ten times that of the single-agent case. From Theorem 3, the upper bound M to the equilibrium speed of learning is approximately 4.4, which is less than ten times 0.51.

neighbors' neighbors etc. However, we argue that the social information gathered from observing neighbors' past actions cannot be too precise. Indeed, if that were the case, agents would base their decisions only on the social information. As a result, their actions would no longer reveal any information about their private signals so that information aggregation would cease. Thus, social information cannot grow to be much more precise than private information. But if agents learn quickly, then their actions provide very precise social information. Hence, we conclude that agents cannot learn too quickly.

In sum, regardless of the size of the network, private information must continue to influence agents' decision-making, which can only happen if the social information is not too precise, which in turn can only happen if agents do not learn quickly. Moreover, as we state in the next section, M is an upper bound to how fast the precision of private information increases with time (see Lemma 12 in §3.4), and this bound, too, is independent of the network size. Combining these insights, we conclude that the speed of learning in a social network of any size is bounded by M.

Numerical Calculation on the Line Graph

While Theorem 3 shows that information aggregation fails in the long run, it leaves open the question of what happens in early periods. Clearly, if many agents all observe each other, much information could be aggregated already in the first period, as the first period actions can reveal many independent pieces of information. The answer to this question becomes less clear in a setting where the number of neighbors is bounded, even if there are many agents in total.

To supplement the asymptotic result of Theorem 3 we consider agents who observe both of their adjacent neighbors on a line, i.e., $N_i = \{i-1, i, i+1\} \cap N$. We study a binary state and binary action setting with a uniform prior and assume that in each period, each agent gets a conditionally independent and identically distributed symmetric binary signal that is equal to the state with probability q. We consider myopic agents, i.e., $\delta = 0$, and the tie-breaking rule under which agents follow their first signal when they are indifferent.

Under these assumptions we calculate the exact probabilities of mistakes in the first 10 periods. A naive calculation would require considering some 10^{30} possible signal realizations, which is not feasible.[11] To approach the computational problem,

[11] On a bidirectional line the number of signals that could (potentially indirectly) influence an agent's period t action is the sum of her total number of private signals up to time t and the total number

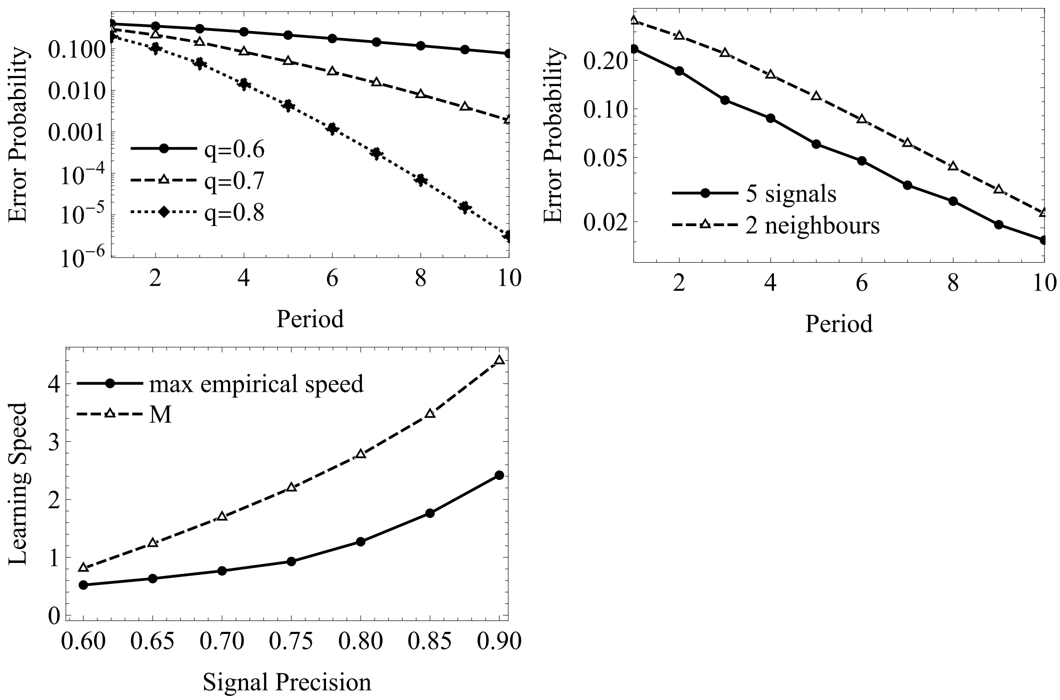

Figure 2.1: On the top left: The log of the inverse of probability of error $-\log \mathbb{P}[a_t^i \neq a^\omega]$ as a function of the period for agents on the bidirectional line graph (away from its ends) for different probabilities q of the signal matching the state. The top right picture compares the error probabilities when the agent observes their 2 neighbors' actions and 4 other agents' signals for $q = 0.65$. On the bottom: The maximal empirical learning speed $\max_{t\in\{2,\ldots,10\}} -\log \frac{1}{t}\mathbb{P}[a_t^i \neq a^\omega]$ for different precisions of the signal on the x-axis.

we use the "dynamic cavity algorithm" proposed by Kanoria and Tamuz [2013] for calculating Bayesian beliefs in social learning environments on tree graphs, which exploits the fact that conditioning on the state and a given agent's actions makes the actions of his left-hand neighbor independent of the actions of his right-hand neighbor. As such a decoupling argument is not available for graphs with cycles, it seems computationally infeasible to perform a similar calculation for, e.g., the two dimensional grid.

We focus on the agents who are not close to the ends of the graph: All agents $i \in \{11, 12, \ldots, n-12, n-11\}$ face the same decision problem in the first ten periods, and we calculate their probabilities of choosing the wrong action. Note that these probabilities are independent of the number of agents n. Equivalently, these are the error probabilities of any agent on a bi-infinite line graph.

of signals observed by his $t-1$ neighbors in each direction, i.e. $t + \sum_{s\leq t-1}(2s) = t + t(t-1) = t^2$. When $t = 10$, this would yield $2^{100} \approx 1.3 \times 10^{30}$ signal realizations.

The results are depicted in Figure 2.1. On the top left, we plot the evolution of error probabilities on a log-scale for different precision of the signals q. Since the ordinate uses a logarithmic scale, an exponential decay of error probabilities would manifest as a downward sloping straight line. Indeed, the graph shows that this decay is approximately exponential from the very early periods. This suggests that at least for myopic agents (the limiting case of very impatient agents), our asymptotic results bounding the speed of learning may begin to apply early on, even before these agents have learned the state very precisely.

On the bottom left we plot, for different signal precisions, the maximum of the empirical learning speed $\max_{t\in\{2,\ldots,10\}} -\log \frac{1}{t}\mathbb{P}[a_t^i \neq a^\omega]$ starting from the second period as well as our asymptotic bound of M. As one can see the asymptotic bound we obtained in Theorem 3 holds for different precision of the signal starting from the second period. Thus, in this example, the conclusion of Theorem 3 do not only hold asymptotically, but already in early periods. For comparison, we also plot the maximum of the rate of mistake for the case where five agents share their private signals on the top right, for $q = 0.65$. The figure shows that in the first ten periods agents do worse on the bidirectional line than they would if they directly observed the private signals of their two neighbors on the right and two neighbors on the left.[12]

2.5 Analysis

In this section we provide a detailed analysis of the agents' beliefs and behavior, leading to a proof sketch for Theorem 3.

Agents' Beliefs

Let $p_t^{i,\mathfrak{g}}$ denote the posterior belief of agent i assigned to event $\omega = \mathfrak{g}$ after observing I_t^i, i.e. $p_t^{i,\mathfrak{g}} = \mathbb{P}[\omega = \mathfrak{g}|I_t^i]$. The log-likelihood ratio of agent i's posterior beliefs of the state being $\mathfrak{g}$ over the state being $\mathfrak{f}$ at time t is

$$L_t^{i,\mathfrak{g},\mathfrak{f}} = \log \frac{p_t^{i,\mathfrak{g}}}{p_t^{i,\mathfrak{f}}} = \log \frac{\mathbb{P}[\omega = \mathfrak{g}|s_1^i, \ldots, s_t^i, H_t^i]}{\mathbb{P}[\omega = \mathfrak{f}|s_1^i, \ldots, s_t^i, H_t^i]}. \tag{2.5.1}$$

Then, it follows from Bayes' rule that this log-likelihood ratio of agent i's posterior beliefs at time t is equal to

$$L_t^{i,\mathfrak{g},\mathfrak{f}} = \log \frac{\mathbb{P}[\omega = \mathfrak{g}]}{\mathbb{P}[\omega = \mathfrak{f}]} + \log \frac{\mathbb{P}[H_t^i|\omega = \mathfrak{g}]}{\mathbb{P}[H_t^i|\omega = \mathfrak{f}]} + \log \frac{\mathbb{P}[s_1^i, \ldots, s_t^i|H_t^i, \omega = \mathfrak{g}]}{\mathbb{P}[s_1^i, \cdots, s_t^i|H_t^i, \omega = \mathfrak{f}]}.$$

[12] While not depicted in the figure, this still holds for any precision of the signal $q \in \{0.6, 0.7, 0.8, 0.9\}$.

We call

$$Q_t^{i,\mathfrak{g},\mathfrak{f}} = \log \frac{\mathbb{P}[\omega = \mathfrak{g}]}{\mathbb{P}[\omega = \mathfrak{f}]} + \log \frac{\mathbb{P}[H_t^i|\omega = \mathfrak{g}]}{\mathbb{P}[H_t^i|\omega = \mathfrak{f}]},$$

the *social likelihood* of agent i at time t. This is the log-likelihood ratio of the social information observed by agent i. Intuitively, Q_t^i measures the inference an outside observer would draw from the observations of the actions of i and his neighbors, without observing i's private signals. Similarly, we call

$$P_t^{i,\mathfrak{g},\mathfrak{f}} = \log \frac{\mathbb{P}[s_1^i, \ldots, s_t^i|H_t^i, \omega = \mathfrak{g}]}{\mathbb{P}[s_1^i, \cdots, s_t^i|H_t^i, \omega = \mathfrak{f}]},$$

the *private likelihood* agent i at time t. Thus, we can write the log-likelihood ratio of agent i's posterior beliefs at time t as

$$L_t^{i,\mathfrak{g},\mathfrak{f}} = S_t^{i,\mathfrak{g},\mathfrak{f}} + P_t^{i,\mathfrak{g},\mathfrak{f}}, \tag{2.5.2}$$

which is the sum of his social likelihood and his private likelihood. We call $L_t^{i,\mathfrak{g},\mathfrak{f}}$ the *posterior likelihood* of agent i at time t.

Agents' Behavior

In the context of a strategy profile σ, the *myopic action* of agent i at time t is

$$m_t^i \in \arg\max_{a \in A} \mathbb{E}[u(a, \omega)|I_t^i].$$

This is the action that maximizes the expected flow utility given the information available at that time, and hence it is the action that a myopic agent would take.[13]

In contrast to a myopic agent, a strategic agent may not always choose the myopic action in equilibrium. Indeed, since a strategic agent is forward-looking, in each period he faces a trade-off between choosing the myopic action and strategically experimenting by choosing a non-myopic action. On the one hand, he needs to bear the immediate cost associated with a non-myopic action. On the other hand, choosing a non-myopic action may allow him to elicit more information from his neighbors' future actions, which he could then use to make better choices in the future. Hence, when the informational gain from experimenting exceeds the current loss caused by a non-myopic action, a strategic agent has an incentive to experiment.

Nevertheless, we show that when an agent's belief of any state is close enough to zero or one, he chooses the unique optimal action at that state which is also the myopic

[13] We assume some deterministic tie-breaking rule when the agent is indifferent. Our results do not depend on this choice and would follow for any tie breaking rule that is common knowledge.

action in equilibrium. This holds despite the fact that the agent is forward-looking. Intuitively, if a strategic agent is very confident about the state, he is expected to pay a high cost if he chooses a non-myopic action and experiments. Consequently, as his expected future gain will not exceed the expected current loss from experimenting, he has no incentive to experiment.

Lemma 9 (Myopic and Strategic Behavior)**.**

(i) *There is a constant* $c > 0$*, independent of* δ *such that, in any equilibrium, if it holds for* $\mathfrak{g} \in \Omega$ *that* $p_t^{i,\mathfrak{g}} > \frac{c}{c+1-\delta}$*, then* $a_t^i = m_t^i = a^{\mathfrak{g}}$*.*

(ii) *There exists a random time* $T < \infty$ *such that in equilibrium, all agents behave myopically after* T*, i.e.* $t \geq T \Rightarrow a_t^i = m_t^i$ *for all* i *almost surely.*

The first part of Lemma 9 applies to a fixed discount factor δ, rather than asymptotically to δ tending to one. So, agents need not have learned the state very precisely at the point in which they become myopic. However, it does imply that as agents become more patient, i.e., δ increases, the posterior belief threshold $\frac{c}{c+1-\delta}$ for choosing the myopic action becomes closer to certainty. Indeed, as δ approaches 1, agents value their future utilities more and the incentive of experimenting becomes stronger. For these agents to forgo their potential expected future informational gains and choose the myopic action, they must be fairly confident about the state. The second part of Lemma 9 states that in finite time the belief of all agents will be sufficiently precise such that they all behave myopically in all future periods.

The next lemma shows that in equilibrium, each agent learns weakly faster than any of his neighbors. Denote the equilibrium speed of learning of agent i by r_i.

Lemma 10 (All agents learn at the same speed)**.**

(i) *If agent* i *can observe agent* j*, i.e.* $j \in N_i$*, then in equilibrium* i *learns weakly faster than* j*, i.e.* $r_i \geq r_j$*.*

(ii) *All agents learn at the same speed, i.e.* $r_i = r_j$ *for all* i, j*, in any strongly connected network.*

We will henceforth call the common speed of learning in a strongly connected network *the equilibrium speed of learning*. The proof of this lemma relies on an extension of the *imitation principle* from myopic agents to strategic agents. For myopic agents, the imitation principle states that if i observes j, then i's actions are not worse than j's:

$$\mathbb{E}[u(a_t^i, \omega)] \geq \mathbb{E}[u(a_{t-1}^j, \omega)],$$

since i can always imitate j [similar arguments are used in Gale and Kariv, 2003, Golub and Sadler, 2017, Smith and Sørensen, 2000, Sørensen, 1996]. We show that for strategic agents, in equilibrium, i's actions are never much worse than j's, even though i may choose myopically sub-optimal actions. To formalize this we denote by $\bar{u} = \mathbb{E}[u(a^\omega, \omega)]$ the expected utility of a decision maker that knows the state. For any i and t it holds that $\mathbb{E}[u(a_t^i, \omega)] \leq \bar{u}$. We can think of $\bar{u} - \mathbb{E}[u(a_t^i, \omega)]$ as the loss in flow utility as compared to the first-best. The imitation principle for strategic agents takes the following form:

$$\bar{u} - \mathbb{E}[u(a_t^i, \omega)] \leq \frac{1}{1-\delta}(\bar{u} - \mathbb{E}[u(a_{t-1}^j, \omega)]).$$

It is obtained by upper-bounding the agent's loss by the loss he would obtain if guessing correctly in every future period and observing that this loss must be less than the loss obtained by taking the action agent j took last period in every future period. By Claim 2 in the Appendix, this upper bound on the loss implies that there is some constant $c > 0$ such that the probability of choosing incorrectly is bounded:

$$\mathbb{P}[a_t^i \neq a^\omega] \leq \frac{c}{1-\delta}\mathbb{P}[a_{t-1}^j \neq a^\omega].$$

One can easily see that when $\delta = 0$, the above equation coincides with the imitation principle for myopic agents.

Social and Private Beliefs

We analyze the agents' beliefs by decomposing their likelihoods into the private and social parts. Recall that by (2.5.2), the posterior likelihood of agent i's at time t, $L_t^{i,\mathfrak{g},\mathfrak{f}}$, is equal to $Q_t^{i,\mathfrak{g},\mathfrak{f}} + P_t^{i,\mathfrak{g},\mathfrak{f}}$, the sum of the social and the private likelihoods. We are interested in the sign of $L_t^{i,\mathfrak{g},\mathfrak{f}}$ as it determines the corresponding myopic action: m_t^i equals $\mathfrak{g}$ if $L_t^{i,\mathfrak{g},\mathfrak{f}} \geq 0$, and $\mathfrak{f}$ otherwise. Let us first focus on the first component of $L_t^{i,\mathfrak{g},\mathfrak{f}}$: agent i's social likelihood $Q_t^{i,\mathfrak{g},\mathfrak{f}}$. In the following lemma, we establish a relationship between the equilibrium speed of learning and the precision of the social likelihood, which is crucial in proving our main theorem.

Lemma 11. *Suppose that the equilibrium speed of learning is at least r. Then, conditioned on $\omega = \mathfrak{g}$, it holds for any $\mathfrak{f} \neq \mathfrak{g}$*

$$\liminf_{t\to\infty} \frac{1}{t} Q_t^{i,\mathfrak{g},\mathfrak{f}} \geq r \quad \textit{almost surely.}$$

This lemma states that a high learning speed implies that the social information inferred from a given agent i and his neighbors' actions must become precise at a high speed. Intuitively, if agents learn quickly, then their actions provide very precise social information.

The proof of Lemma 11 uses the idea of a fictitious outside observer who observes the same social information as agent i and nothing else. Since he observes i's actions, he can achieve the same learning speed as i. This implies that his posterior likelihood increases fast. Hence, as this outside observer's posterior likelihood coincides with agent i's social likelihood, the precision of i's social information increases at a speed of at least r.

Next, we focus on the second component of $L_t^{i,\mathfrak{g},\mathfrak{f}}$: agent i's private likelihood, $P_t^{i,\mathfrak{g},\mathfrak{f}}$. As agents receive more independent private signals over time, their private information about the state becomes more precise. However, the precision of their private information cannot increase without bound, as shown in the following lemma.

Lemma 12. *For any agent i at time t and any distinct $\mathfrak{g}, \mathfrak{f} \in \Omega$, the absolute value of the private likelihood is at most $t \cdot M$, i.e.*

$$\frac{1}{t}|P_t^{i,\mathfrak{g},\mathfrak{f}}| \leq M \quad \textit{almost surely.}$$

This lemma states that that at any given time t, there is an upper bound to the precision of agents' private information, which only depends on the private signals distribution and is independent of the structure of the network and the history of observed actions. Notice that since $P_t^{i,\mathfrak{g},\mathfrak{f}} = L_t^{i,\mathfrak{g},\mathfrak{f}} - S_t^{i,\mathfrak{g},\mathfrak{f}}$ by (2.5.2), it captures the difference between what i knows about the state and what an outside observer who observes i's actions and his neighbors' actions would know about the state. Thus, the bound Mt assigned to i's private signals also applies to the difference between the posterior likelihood L_t^i and the social likelihood S_t^i.

Proof sketch for Theorem 3

We end this section by providing a sketch of the proof of Theorem 3 using our earlier results. Suppose to the contrary that in equilibrium, agents learn at a speed that is strictly higher than M, where M is twice the log-likelihood ratio of the strongest signal. Then, by Lemma 11, the social information would become precise at a speed that is also strictly higher than M. Meanwhile, at any given time t, the precision of the private information is at most Mt, as shown in Lemma 12. Hence, by (2.5.2) the sign of L_t^i would be determined purely by the social information after some (random) time. By Lemma 9, there exists a time T such that from T onward, in equilibrium, all agents act only based on the social information, and furthermore they would choose the myopic action even though they are forward-looking. Consequently, their actions would no longer reveal any information about their private signals and information would cease to be aggregated. This contradicts our hypothesis that the precision of the social information grows at such a high speed. Therefore, we conclude that the equilibrium speed of learning in networks does not increase beyond M.

2.6 Networks which are not Strongly Connected

So far, we have focused on strongly connected networks where there is an observational path between every pair of agents. While on a strongly connected network all agents learn at the same speed, this is not true for general networks. For example, consider a simple star network where there is a single agent at the center who observes everyone, and where the remaining peripheral agents observe no one. Here, the peripheral agents' actions are independent conditional on the state. These actions supply the central agent with $n - 1$ additional independent signals, and he thus learns at a speed that increases linearly with the number of agents.[14] In contrast, all peripheral agents learn at a constant speed r_a, as in the single-agent case. Hence, for general networks, depending on the structure of the network, some agents can learn faster than others.

More importantly, this star network example implies that the bound obtained in Theorem 3 does not hold for all agents in a non-strongly connected network. The intuitive reason is that when the network is not strongly connected, some agents might remain unobservable to others, e.g., the central agent in the star network. These agents can thus learn very fast from observing others since their own past actions do not affect the actions of others, rendering others' past actions more informative. This cannot happen in a strongly connected network where every agent (potentially indirectly) learns about the actions of every other agent.

Nevertheless, even though it is not necessarily true that all agents learn slowly, our next result establishes that it is still true that some agents will learn slowly in any network.

Proposition 6. *Consider an arbitrary (finite) network and let r_i be the speed of learning of agent i. We have that* $\min_i r_i \leq M$.

[14] See Theorem 5 in Harel et al. [2021].

The proof of Proposition 6 relies on the idea that within any general network, there is always a strongly connected sub-network, say E, in which no agent observes any agent outside of this sub-network. Thus, the learning process at E is independent of the agents outside of E. Since E is strongly connected, Theorem 3 implies that the speed of learning on E is bounded by M.

Learning on a Line

In this section we discuss the commonly studied case of an infinite group of agents who learn by observing (a subset) of their predecessors. A prominent feature of the line network is that information is transmitted unidirectionally. This simplified observational structure has thus received particular attention in the herding literature where agents act only once. Here, we extend it to our setting where agents act repeatedly. As we will see below, the arguments we use on line networks are reminiscent of the sequential social learning literature, except that we focus on the speed of learning rather than whether learning occurs or not.

We first consider the case where agents observe a general subset of their predecessors, which in the herding literature was considered in Acemoglu et al. [2011].[15]

Proposition 7. *Suppose that $N = \{1, 2, \ldots\}$ and $N_i \subseteq \{1, \ldots, i\}$. Then there is some constant K such that the speed of learning $r_i \leq K$ for infinitely many agents i.*

That is, under these assumptions, it is impossible that the speed of learning r_i tends to infinity with i. Thus, even though there are infinitely many agents, many agents will have a low speed of learning. This result thus shows that in the line networks, the conclusion of Proposition 6 that the speed of learning of some agent must be bounded generalizes to the conclusion that the speed of an infinite number of agents must be bounded.

The proof of this proposition uses ideas that are similar to those of the proof of Theorem 3. Since observations are unidirectional, agents behave myopically even for $\delta > 0$. To show the result, suppose towards a contradiction, that there is no such K. Then, in particular, there are only finitely many agents whose speed of learning is less than M. Each of the remaining agents eventually stops using their private signals, because the fact that they learn so quickly means that they observe very strong social information. It follows that the only information that is aggregated

[15] Their conclusion is qualitatively different: For some network structures asymptotic learning obtains, and for some it does not. Of course, their notion of learning (namely that agent i takes the correct action with probability that tends to 1 as i tends to infinity) is very different than ours.

asymptotically is that of the finite group of agents who learn slowly. Thus, it is impossible that any agent has a high speed of learning, which is a contradiction.

We now consider the special case of $N_i = \{i-1, i\}$, i.e., each agent observes only their direct predecessor.

Proposition 8. *When $N_i = \{i-1, i\}$, $r_i \leq M$ for all agents i.*

Thus, the conclusion of Theorem 3 applies also to this case of a non-strongly connected network. Again, we prove Proposition 8 by using the ideas behind Theorem 3. Notice first that the imitation principle for myopic agents implies that the speed of learning is weakly increasing in the index of the agent. Now, suppose to the contrary that there exists some agent k who learns at a speed that is strictly greater than M. Then all agents $j > k$ would also learn at a high speed. Eventually, all these high-speed learning agents would stop using their private signals because the fact that they learn so fast means that they observe very precise social information. Consequently, information aggregation stagnates and thus learning cannot be too fast.

2.7 Conclusion

In this paper, we show that information aggregation is highly inefficient for large groups of agents who learn from private signals and by observing their social network neighbors. To overcome the difficulty of constructing equilibria explicitly, we focus on the asymptotic speed of learning, allowing us to prove results that apply to all equilibria. We show that regardless of the size of the network, the speed of learning is bounded above by a constant, which only depends on the private signal distribution (and not on the discount factor, the observational graph, the agents' utilities, or prior belief).

An important limitation of our results is that they only apply asymptotically. As our numerical results show, these asymptotic results can apply already from the early periods, for myopic agents on particular networks. However, for patient agents, it is unclear whether it takes a long time for the asymptotic results to apply. Calculating welfare for patient agents seems beyond what is currently tractable, as it would require a detailed analysis of the equilibria of this game. In fact, even for two myopic agents who observe each other, it seems intractable to calculate welfare, and moreover, it remains unknown whether the asymptotic speed of learning is strictly greater than that of one agent who learns on his own. Nevertheless, even in the most general settings of patient agents on complex networks, it seems reasonable to conjecture that the main economic force driving our result—that learning cannot be too fast because it would lead agents to disregard their own signals—is significant even in the early periods. We leave this for future research.

Another promising direction for future research is the calculation of lower bounds for the speed of learning. Currently, we cannot show that equilibrium speed of learning on any connected network is faster than that of a single agent. Even without imposing equilibrium, this question remains open: What learning speed can be achieved when a social planner is allowed to choose the agents' strategies? For the complete network, it is still unknown how a social planner could achieve any speed of learning that is better than the speed achieved by a single agent. The challenge for the social planner lies in the trade-off between using the actions to communicate between the agents and choosing the correct actions with very high probability at the same time. One conjecture from Harel et al. [2021] is that a better speed can be achieved by having the social planner instruct the agents to behave as if they are myopic and over-weight their own signals, causing their actions to reveal more information. In simpler, sequential settings, this type of mechanism was shown to indeed improve learning outcomes [Arieli et al., 2023].

It is possible to extend our model in a number of directions. In Appendix 2.9 we show that our main result continues to hold if agents do not observe each other every period, but only in some (potentially random) periods. A natural extension, which we leave for future work, is to allow the network to be random and its realization to be only partially observed by the agents. An interesting technical question is that of the robustness of our results to the assumption of bounded private signals. We conjecture that a result similar to our main theorem should hold even if signals are unbounded, with the Kullback-Leibler divergence between the conditional signal distributions playing the bounding role currently played by the maximum log-likelihood ratio. This is indeed the case for myopic agents on the complete network [Harel et al., 2021]. A substantive extension is to allow the underlying state to change over time [see Dasaratha et al., 2020, Frongillo et al., 2011, Moscarini et al., 1998]. For example, the underlying unknown state could capture the quality of a local restaurant or school, which might fluctuate gradually. In such a setting, one could replace the speed of learning metric with the long-run probability of making the correct choice and study whether information gets aggregated in this case and, if so, whether the information aggregation process is efficient.

2.8 Proofs

Proof of Lemma 9. (i) First notice that we can write the agent's expected utility conditioned on his information I_t^i at time t as the sum $U_{<t} + \delta^t U_{\geq t}$, where

$$U_{<t} = (1-\delta)\sum_{k=1}^{t-1} \delta^{k-1}\mathbb{E}[u(a_k^i,\omega)|I_t^i]$$

is the sum of the expected flow utilities until time t, and $U_{\geq t}$ is the expected continuation utility at time t given by

$$\begin{aligned} U_{\geq t} &= (1-\delta)\sum_{k=0}^{\infty} \delta^k \mathbb{E}[u(a_{t+k}^i,\omega)|I_t^i] \\ &= (1-\delta)\mathbb{E}[u(a_t^i,\omega)|I_t^i] + \delta(1-\delta)\sum_{k=0}^{\infty}\delta^k\mathbb{E}[u(a_{t+k+1}^i,\omega)|I_t^i]. \end{aligned}$$

Fix some state $\mathfrak{g} \in \Omega$. Since $a^{\mathfrak{g}}$ is the unique optimal action in state $\mathfrak{g}$, by applying an affine transformation to the flow utility function $u\colon A\times\Omega \to \mathbb{R}$ we can assume that $u(a^{\mathfrak{g}},\mathfrak{g}) = 1$, that $u(a,\mathfrak{g}) \leq 0$ for all $a \neq a^{\mathfrak{g}}$. Let $c^{\mathfrak{g}} = \max_{a,\mathfrak{f}} |u(a,\mathfrak{f})|$. Recall that $p_t^{i,\mathfrak{g}} = \mathbb{P}[\omega = \mathfrak{g}|I_t^i]$ is the agent's posterior at time t. Thus, for any action $a \neq a^{\mathfrak{g}}$, since $u(a,\mathfrak{g}) \leq 0$, the expected flow utility $\mathbb{E}[u(a,\omega)|I_t^i] = \sum_{\mathfrak{f}\in\Omega} u(a,\mathfrak{f})\cdot p_t^{i,\mathfrak{f}}$ is at most $c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}})$. Likewise, $\mathbb{E}[u(a_t^i,\omega)|I_t^i]$ is at most $p_t^{i,\mathfrak{g}} + c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}})$ since $u(a^{\mathfrak{g}},\mathfrak{g}) = 1$.

Now, suppose that $a_t^i = a \neq a^{\mathfrak{g}}$. Then the expected continuation utility is

$$\begin{aligned} U_{\geq t} &= (1-\delta)\mathbb{E}[u(a,\omega)|I_t^i] + \delta(1-\delta)\sum_{k=0}^{\infty}\delta^k\mathbb{E}[u(a_{t+k+1}^i,\omega)|I_t^i] \\ &\leq (1-\delta)c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}}) + \delta(p_t^{i,\mathfrak{g}} + c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}})) \\ &= \delta p_t^{i,\mathfrak{g}} + c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}}). \end{aligned}$$

On the other hand, the strategy that chooses $a^{\mathfrak{g}}$ from period t onward has an expected continuation utility at least $p_t^{i,\mathfrak{g}} - c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}})$. Thus, when

$$p_t^{i,\mathfrak{g}} - c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}}) > \delta p_t^{i,\mathfrak{g}} + c^{\mathfrak{g}}(1-p_t^{i,\mathfrak{g}}) \tag{2.8.1}$$

the agent cannot choose $a_t^i = a \neq a^{\mathfrak{g}}$ in period t. Rearranging, this happens when

$$p_t^{i,\mathfrak{g}} > \frac{2c^{\mathfrak{g}}}{2c^{\mathfrak{g}}+1-\delta}.$$

Thus, under the above condition, agents choose $a_t^i = a^{\mathfrak{g}}$ in equilibrium. Clearly, the myopic action is then also equal to $a^{\mathfrak{g}}$, as this corresponds to the case $\delta = 0$. Part (i) of the lemma now follows by setting $c = \max_{\mathfrak{g}} 2c^{\mathfrak{g}}$.

For part (ii), fix a discount factor $\delta \in [0, 1)$ and let σ be an equilibrium. Let a_t^i be the action taken by i at time t under σ. Since the entire sequence of private signals reveals the state, conditional on $\omega = \mathfrak{g}$, $\lim_t p_t^{i,\mathfrak{g}} = 1$ almost surely. Hence, by part (i) the agent will choose the unique optimal action that is also myopically optimal in equilibrium from some (random) time on. Since there are finitely many agents, this will hold for all i for all t larger than some (random) T. This means that from T onward, in equilibrium all agents will behave myopically. □

The following is a simple consequence of Lemma 9.

Corollary 2. *There is a constant $C > 0$ such that, in any equilibrium, if it holds for $\mathfrak{g} \in \Omega$ and all $\mathfrak{f} \neq \mathfrak{g}$ that $L_t^{i,\mathfrak{g},\mathfrak{f}} > C$, then $a_t^i = m_t^i = a^{\mathfrak{g}}$.*

Proof. Fix a state $\mathfrak{g} \in \Omega$. Suppose that $L_t^{i,\mathfrak{g},\mathfrak{f}} > C$ for some $C > 0$ to be chosen later. Then $p_t^{i,\mathfrak{g}} > \mathrm{e}^C p_t^{i,\mathfrak{f}}$. If this holds for all $\mathfrak{f} \neq \mathfrak{g}$ then

$$p_t^{i,\mathfrak{g}} > \frac{1}{|\Omega| - 1} \mathrm{e}^C \sum_{\mathfrak{f} \neq \mathfrak{g}} p_t^{i,\mathfrak{f}} = \frac{1}{|\Omega| - 1} \mathrm{e}^C (1 - p^{i,\mathfrak{g}}).$$

Thus for any each $c > 0$, u and $\delta \in [0, 1)$, for C large enough it holds that

$$p_t^{i,\mathfrak{g}} > \frac{c}{c + 1 - \delta},$$

and the result follows by part (i) of Lemma 9. □

The following lemma will be useful in the proofs below. Recall that we denote $\bar{u} = \mathbb{E}[u(a^\omega, \omega)]$.

Claim 2. *There exist $\underline{c}, \overline{c} > 0$ such that*

$$\underline{c} \cdot \mathbb{P}[a_t^i \neq a^\omega] \leq \bar{u} - \mathbb{E}[u(a_t^i, \omega)] \leq \overline{c} \cdot \mathbb{P}[a_t^i \neq a^\omega]$$

Recall that we can think of the difference $\bar{u} - \mathbb{E}[u(a_t^i, \omega)]$ as the expected loss in flow utility as compared to the first-best. The lemma above states that this quantity is the same—up to constants—as the probability of choosing the correct action. An immediate consequence of this lemma is that we can express the speed of learning in terms of this loss:

$$r_i = \liminf_t -\frac{1}{t} \log \mathbb{P}[a_t^i \neq a^\omega] = \liminf_t -\frac{1}{t} \log(\bar{u} - \mathbb{E}[w(a_t^i, a^\omega)]). \qquad (2.8.2)$$

Proof of Claim 2. Denote by $\underline{c}$ the minimum loss of utility from choosing incorrectly:

$$\underline{c} = \min_{\mathfrak{g}\in\Omega, a\neq a^{\mathfrak{g}}} u(a^{\mathfrak{g}}, \mathfrak{g}) - u(a, \mathfrak{g}).$$

Since there is a unique optimal action in each state we have that $\underline{c} > 0$. Analogously, denote by $\overline{c} > 0$ the maximum such loss:

$$\overline{c} = \max_{\mathfrak{g}\in\Omega, a\neq a^{\mathfrak{g}}} u(a^{\mathfrak{g}}, \mathfrak{g}) - u(a, \mathfrak{g}).$$

Then

$$\bar{u} - \mathbb{E}[u(a_t^i, \omega)] = \mathbb{E}[u(a^\omega, \omega) - u(a_t^i, \omega)] \leq \mathbb{E}[\overline{c} \cdot 1_{a_t^i \neq a^\omega}] = \overline{c} \cdot \mathbb{P}[a_t^i \neq a^\omega].$$

Likewise,

$$\bar{u} - \mathbb{E}[u(a_t^i, \omega)] = \mathbb{E}[u(a^\omega, \omega) - u(a_t^i, \omega)] \geq \mathbb{E}[\underline{c} \cdot 1_{a_t^i \neq a^\omega}] = \underline{c} \cdot \mathbb{P}[a_t^i \neq a^\omega].$$

□

Proof of Lemma 10. Suppose i observes j. Let σ be an equilibrium and let a_t^i be the action taken by i at time t under σ. We claim that for $t > 1$,

$$(1-\delta)(\bar{u} - \mathbb{E}[u(a_t^i, \omega)]) \leq \bar{u} - \mathbb{E}[u(a_{t-1}^j, \omega)]. \tag{2.8.3}$$

To see that this equation must hold observe that the left-hand side equals the expected continuation loss the agent would have from time t on if he chooses the action a_t^i at time t and suffered no loss in future periods. Hence this is smaller than the expected continuation loss under the strategy profile σ. In equilibrium, this must be smaller than the loss from any deviation, and the right-hand-side equals the loss the agent suffers when imitating j's action a_{t-1}^j from time t onward. Thus the above inequality must hold.

As a consequence,

$$\begin{aligned}\liminf_{t\to\infty} -\frac{1}{t}\log(\bar{u} - \mathbb{E}[u(a_t^i, \omega)]) &\geq \liminf_{t\to\infty} -\frac{1}{t}\log\left(\frac{1}{1-\delta}(\bar{u} - \mathbb{E}[u(a_{t-1}^j, \omega)])\right) \\ &= \liminf_{t\to\infty} -\frac{1}{t}\log(\bar{u} - \mathbb{E}[u(a_{t-1}^j, \omega)]) \\ &= \liminf_{t\to\infty} -\frac{1}{t}\log(\bar{u} - \mathbb{E}[u(a_t^j, \omega)]).\end{aligned}$$

Thus part (i) follows from (2.8.2). Part (ii) follows as in any strongly connected network, there is an observational path from each agent i to each other agent j and the monotonicity of learning speed shown in (i) applied along this path, implies that $r_i \geq r_j$. The opposite inequality holds by the same argument. □

The next simple claim will be helpful in the Proof of Lemma 11.

Claim 3. *Suppose that* X_t *is a sequence of random variables taking values in* $[0,1]$ *such that* $\lim_t -\frac{1}{t}\log\mathbb{E}[X_t|\Sigma_t] \geq r$ *almost surely, for some sequence of sigma-algebras* Σ_t*.Then* $\liminf_t -\frac{1}{t}\log X_t \geq r$*, almost surely.*

Proof of Claim 3. By the claim hypothesis there exists a random $F\colon \mathbb{N}\to\mathbb{R}$, such that $\lim_t F(t)/t = 0$ almost surely and such that

$$\mathbb{E}[X_t|\Sigma_t] \leq \mathrm{e}^{-rt+F(t)}.$$

We can furthermore assume that $F(t) \leq rt$, since $X_t \leq 1$. Taking expectations of both sides yields

$$\mathbb{E}[X_t] \leq \mathrm{e}^{-rt}\cdot\mathbb{E}[\mathrm{e}^{F(t)}].$$

Hence, if we denote $f(t) = \log\mathbb{E}[\mathrm{e}^{F(t)}]$,

$$\mathbb{E}[X_t] \leq \mathrm{e}^{-rt+f(t)}.$$

Furthermore,

$$\lim_t \frac{1}{t}f(t) = \lim_t \frac{1}{t}\log\mathbb{E}[\mathrm{e}^{F(t)}] = \lim_t \log\mathbb{E}[\mathrm{e}^{F(t)/t}] = 0,$$

where the last equality is a consequence of the facts that $F(t)/t \leq r$ and $F(t)/t$ converges almost surely to zero. Hence, by Markov's inequality, for any $c_t > 0$,

$$\mathbb{P}[X_t \geq c_t] \leq \mathbb{E}[X_t]/c_t \leq \mathrm{e}^{-rt+f(t)}/c_t\,.$$

Choosing $c_t = \mathrm{e}^{-rt+f(t)}\cdot t^2$, we get that

$$\mathbb{P}[X_t \geq \mathrm{e}^{-rt+f(t)+2\log t}] \leq \frac{1}{t^2}\,.$$

Hence, by Borel-Cantelli, almost surely $X_t \leq \mathrm{e}^{-rt+f(t)+2\log t}$ for all t large enough, and in particular, $\liminf_t -\frac{1}{t}\log X_t \geq r$. □

Proof of Lemma 11. In this proof we use Landau notation, so that $o(t)$ stands for some function $f\colon\mathbb{N}\to\mathbb{R}$ such that $\lim_t f(t)/t = 0$.

By assumption and the definition of speed of learning in (2.2.2), $\mathbb{P}[a_t^i \neq a^\omega] \leq \mathrm{e}^{-rt+o(t)}$. Let

$$p_t^{x,\mathfrak{f}} = \mathbb{P}[\omega = \mathfrak{f}|a_t^i]$$

be the probability assigned to $\omega = \mathfrak{f}$ by an outside observer x that sees only agent i's action at time t.

By Bayes' Law,

$$\mathbb{P}[\omega = \mathfrak{f} | a_t^i = a^{\mathfrak{g}}] = \frac{\mathbb{P}[a_t^i = a^{\mathfrak{g}}, \omega = \mathfrak{f}]}{\mathbb{P}[a_t^i = a^{\mathfrak{g}}]}.$$

Since $\mathbb{P}[a_t^i \neq a^\omega] \leq \mathrm{e}^{-rt+o(t)}$, we can bound the denominator by

$$\mathbb{P}[a_t^i = a^{\mathfrak{g}}] \geq \mathbb{P}[\omega = \mathfrak{g}, a_t^i = a^\omega] \geq \mathbb{P}[\omega = \mathfrak{g}] - \mathbb{P}[a_t^i \neq a^\omega] \geq \mathbb{P}[\omega = \mathfrak{g}] - \mathrm{e}^{-rt+o(t)}.$$

If $\mathfrak{g} \neq \mathfrak{f}$ then the numerator is at most $\mathbb{P}[a_t^i \neq a^\omega]$ since the event that the agent takes the wrong action contains the event that the agent takes action $a^{\mathfrak{g}}$ in state $\mathfrak{f}$. Hence

$$\mathbb{P}[\omega = \mathfrak{f} | a_t^i = a^{\mathfrak{g}}] \leq \frac{\mathrm{e}^{-rt+o(t)}}{\mathbb{P}[\omega = \mathfrak{g}] - \mathrm{e}^{-rt+o(t)}}.$$

Now, because $\mathbb{P}[a_t^i \neq a^\omega] \leq \mathrm{e}^{-rt+o(t)}$, by Borel-Cantelli, almost surely $a_t^i = a^{\mathfrak{g}}$ for all t large enough, conditioned on $\omega = \mathfrak{g}$. It follows that for all t large enough—again conditioned on $\omega = \mathfrak{g}$—the belief $p_t^{x,\mathfrak{f}}$ will equal $\mathbb{P}[\omega = \mathfrak{f} | a_t^i = a^{\mathfrak{g}}]$. Hence, by the displayed equation above, $\liminf_t -\frac{1}{t} \log p_t^{x,\mathfrak{f}} \geq r$. Since this holds for all $\mathfrak{f} \neq \mathfrak{g}$, we get that $\lim_t p_t^{x,\mathfrak{g}} = 1$.

Now, let

$$p_t^{y,\mathfrak{f}} = \mathbb{P}[\omega = \mathfrak{f} | H_t^i]$$

be the probability assigned to $\omega = \mathfrak{f}$ by an outside observer y that sees only agent i's public history H_t^i.

Then

$$\log \frac{p_t^{y,\mathfrak{g}}}{p_t^{y,\mathfrak{f}}} = \log \frac{\mathbb{P}[H_t^i | \omega = \mathfrak{g}]}{\mathbb{P}[H_t^i | \omega = \mathfrak{f}]} + \log \frac{\mathbb{P}[\omega = \mathfrak{g}]}{\mathbb{P}[\omega = \mathfrak{f}]} = Q_t^{i,\mathfrak{g},\mathfrak{f}}.$$

Since H_t^i includes a_{t-1}^i, the law of total expectations yields that

$$p_{t-1}^{x,\mathfrak{f}} = \mathbb{E}[p_t^{y,\mathfrak{f}} | a_{t-1}^i].$$

It now follows from Claim 3 that identical asymptotics apply to $p_t^{y,\mathfrak{f}}$: $\liminf_t -\frac{1}{t} \log p_t^{y,\mathfrak{f}} \geq r$ and $\lim_t p_t^{y,\mathfrak{g}} = 1$. Thus, conditioned on $\omega = \mathfrak{g}$,

$$\liminf_{t\to\infty} \frac{1}{t} Q_t^{i,\mathfrak{g},\mathfrak{f}} = \liminf_{t\to\infty} \frac{1}{t} \log \frac{p_t^{y,\mathfrak{g}}}{p_t^{y,\mathfrak{f}}} \geq r \quad \text{almost surely.}$$

□

Proof of Lemma 12. Recall that at time t, $H_t^i = \{a_s^j : s < t, j \in N_i\}$ is the history of actions observed by i and $s_1^i, \ldots, s_t^i$ is the sequence of private signals received by i. Given a *pure* strategy profile σ, agent i chooses a unique action $a_t^i \in A$ at time t: $a_t^i = \sigma(s_1^i, \cdots, s_t^i, H_t^i)$. It follows that for each history H_t^i there is a set $S^i(H_t^i) \subseteq S_1^i \times \cdots \times S_{t-1}^i$ of possible private signal realizations $s_1^i, \ldots, s_{t-1}^i$ that are consistent with H_t^i:

$$S^i(H_t^i) = \left\{s_1^i, \ldots, s_{t-1}^i \in S_1^i \times \cdots \times S_{t-1}^i \ : \ \mathbb{P}[s_1^i = s_1^i, \ldots, s_{t-1}^i = s_{t-1}^i | H_t^i] > 0\right\} .$$

In other words, if we imagine an outside observer who sees only H_t^i—i.e. sees i's actions and his neighbors' actions—then $S^i(H_t^i)$ is the set of private signal realizations $(s_1^i, \ldots, s_{t-1}^i)$ to which this observer assigns positive probability.

Consider the numerator $\mathbb{P}[s_1^i, \ldots, s_t^i | H_t^i, \omega = \mathfrak{g}]$ of the private log-likelihood ratio $P_t^{i,\mathfrak{g},\mathfrak{f}}$. Using the definition of $S^i(H_t^i)$, we can write

$$\mathbb{P}[s_1^i, \ldots, s_t^i | H_t^i, \omega = \mathfrak{g}] = \mathbb{P}[s_1^i, \ldots, s_t^i | (s_1^i, \ldots, s_{t-1}^i) \in S^i(H_t^i), \omega = \mathfrak{g}] \quad \text{almost surely.}$$

The above equality holds as conditional on $\omega = \mathfrak{g}$ the signals of different agents are independent and hence the only relevant information about agent i's signals $s_1^i, \ldots, s_t^i$ contained in the history H_t^i is the restriction the history imposes on the realization of these signals.

Let $\mu_{1\ldots t}^{i,\mathfrak{g}}$ be the measure over signal realizations $s_1^i, \ldots, s_t^i$ when $\omega = \mathfrak{g}$. Then

$$\mathbb{P}[s_1^i, \ldots, s_t^i | (s_1^i, \ldots, s_{t-1}^i) \in S^i(H_t^i), \omega = \mathfrak{g}] = \frac{\mu_{1\ldots t}^{i,\mathfrak{g}}(s_1^i, \ldots, s_t^i)}{\mu_{1\ldots t-1}^{i,\mathfrak{g}}(S^i(H_t^i))}.$$

We thus have that

$$P_t^{i,\mathfrak{g},\mathfrak{f}} = \log \frac{\mu_{1\ldots t}^{i,\mathfrak{g}}(s_1^i, \ldots, s_t^i)}{\mu_{1\ldots t}^{i,\mathfrak{f}}(s_1^i, \ldots, s_t^i)} + \log \frac{\mu_{1\ldots t-1}^{i,\mathfrak{f}}(S^i(H_t^i))}{\mu_{1\ldots t-1}^{i,\mathfrak{g}}(S^i(H_t^i))}. \tag{2.8.4}$$

Since the signals are independent over time the first term of (2.8.4) is equal to

$$\sum_{\tau=1}^{t} \log \frac{\mu_\tau^{i,\mathfrak{g}}(s_\tau^i)}{\mu_\tau^{i,\mathfrak{f}}(s_\tau^i)},$$

which is at most $\frac{1}{2}Mt$. The second term of (2.8.4), which is equal to

$$\log \frac{\sum_{(s_1^i, \cdots, s_{t-1}^i) \in S^i(H_t^i)} \prod_{\tau=1}^{t-1} \mu_\tau^{i,\mathfrak{f}}(s_\tau^i)}{\sum_{(s_1^i, \cdots, s_{t-1}^i) \in S^i(H_t^i)} \prod_{\tau=1}^{t-1} \mu_\tau^{i,\mathfrak{g}}(s_\tau^i)}$$

is also at most $\frac{1}{2}M(t-1)$.[16] Thus, it follows that $P_t^{i,\mathfrak{g},\mathfrak{f}}$ is at most Mt. By an analogous argument $P_t^{i,\mathfrak{f},\mathfrak{g}} = -P_t^{i,\mathfrak{g},\mathfrak{f}}$ is at least Mt, and so $|P_t^{i,\mathfrak{g},\mathfrak{f}}|$ is at most Mt. □

Proof of Theorem 3. Fix a discount factor $\delta \in [0,1)$. Let σ be an equilibrium and a_t^i be the action taken by i at time t under σ. Now consider an outside observer x who observes everybody's actions so that the information available to him at time t is $H_t = \{a_s^i, i \in N, s \leq t\}$ and at time infinity is $H_\infty = \cup_t H_t$. Thus at any time t, this observer can calculate the social likelihood Q_t^i for all i.

Suppose that the social likelihood is high, and in particular $Q_t^{i,\mathfrak{g},\mathfrak{f}} > Mt + C$ at some t for some constant C, some state $\mathfrak{g}$ and all $\mathfrak{f} \neq \mathfrak{g}$. Since the private likelihood $P_t^{i,\mathfrak{g},\mathfrak{f}}$ cannot be less than $-Mt$ (Lemma 12), the posterior likelihood $L_t^{i,\mathfrak{g},\mathfrak{f}} = Q_t^{i,\mathfrak{g},\mathfrak{f}} + P_t^{i,\mathfrak{g},\mathfrak{f}} > C$. In this case, supposing C is high enough, by Corollary 2, the agent will choose the myopic action $a_t^i = m_t^i = a^{\mathfrak{g}}$. Thus, under this condition on Q_t^i the outside observer will know which action the agent will choose in equilibrium, and will not learn anything (in particular, about the agent's signals or the state) from observing this action.

Suppose towards a contradiction that the equilibrium speed of learning r is strictly higher than M, i.e. $r = M + \varepsilon$ for some $\varepsilon > 0$. Then it follows from Lemma 11 that for any $C > 0$, $Q_t^{i,\mathfrak{g},\mathfrak{f}} \geq (M+\varepsilon)t > Mt + C$ for all t large enough and all $\mathfrak{f} \neq \mathfrak{g}$. Since there are finitely many agents, this will hold for all i, for all t larger than some (random) T. Hence the outside observer x learns nothing more from the agents' actions after time T.

Let a_t^x be the action that x would choose to maximize the probability of matching the state at time t. Since no new information is gained after time T, the outside observer stops updating their action and so $a_T^x = a_\infty^x$. Hence $\mathbb{P}[a_\infty^x \neq a^\omega] > 0$.

Since x observes everyone's actions, by the imitation principle

$$\mathbb{P}[a_t^i \neq a^\omega] \geq \mathbb{P}[a_\infty^x \neq a^\omega] > 0 \tag{2.8.5}$$

for all agents i and all times t. But since the equilibrium speed of learning is $M + \varepsilon > 0$, by the definition of speed of learning in (2.2.2), $\mathbb{P}[a_t^i \neq a^\omega]$ converges to zero, in contradiction with (2.8.5).

[16]This follows from the fact that for any two sequences of positive numbers $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ it holds that

$$\frac{\sum_k a_k}{\sum_k b_k} = \frac{\sum_k b_k (a_k/b_k)}{\sum_k b_k} \in \left[\min_k(a_k/b_k), \max_k(a_k/b_k)\right].$$

□

Proof of Proposition 6. Recall that $E \subseteq N$ is a strongly connected component if there is an indirect observation path from each agent in E to each other agent in E. By a standard argument[17], there exists a strongly connected component E in which no agent observes agents outside of E. Hence we can analyze the speed of learning of agents in E in isolation and apply Theorem 3 to conclude that $r_i \leq M$ for all $i \in E$. □

Proof of Proposition 7. Suppose towards a contradiction that there does not exist a constant K where $r_i \leq K$ for infinitely many agents, i.e., that $\lim_i r_i = \infty$. This implies that there exists a finite k such that the agents with $r_i \leq M$ constitute a subset of $\{1, \ldots, k\}$. Thus, for all agents $i > k$, $r_i > M$. Fix an agent $n > k$.

Consider an outside observer x who observes agents $\{1, \ldots, n\}$. Since $r_i > M$ for all $i \in \{k+1, \ldots, n\}$, it follows from Lemma 11 that conditioned on $\omega = \mathfrak{g}$, $Q_t^{i,\mathfrak{g},\mathfrak{f}} > M \cdot t$ for all $i \in \{k+1, \ldots, n\}$, for all $\mathfrak{f} \neq \mathfrak{g}$ and for all t large enough. Since the absolute value of the private likelihood $P_t^{i,\mathfrak{g},\mathfrak{f}}$ is always less than or equal to Mt at any given time t (Lemma 12), all agents $k+1, \ldots, n$ will eventually ignore their private information and act based on their social information. I.e., for all $i \in \{k+1, \ldots, n\}$, a_t^i is determined by Q_t^i for all t large enough. Since there are finitely many agents observed by x, there exists a (random) T so that from T onward, x, who knows Q_t^i, learns nothing more from the actions of agents $\{k+1, \ldots, n\}$. Thus, x will learn as fast as he would if he only observed agents $\{1, \ldots, k\}$. It follows that the speed of learning of this outsider observer x is at most kr_a, which is k times the speed of learning of a single agent, or, equivalently, the speed of learning from directly observing k signals every period. Following the argument in the proof of Theorem 3, since x observes i, x learns at least as fast as i, and so $r_i \leq kM$. Since this holds for all i, we have reached a contradiction to the assumption that $\lim_i r_i = \infty$. □

Proof of Proposition 8. By the imitation principle for myopic agents,

$$r_{i+1} \geq r_i$$

[17] Define a preorder on the set of agents N by $i \geq j$ if there is an indirect observation path from i to j. The $\geq$-equivalence classes are the strongly connected components. The set E is any equivalence class of $\geq$-minimal elements.

for all $i = 1, 2, \ldots$ and $r_1 = r_a < M$ as agent 1 sees only their private signals. Suppose towards a contradiction that there exits an agent $k \geq 2$ agent who learns at a speed that is strictly higher than M, i.e., $r_k = M + 2\varepsilon$ for some $\varepsilon > 0$ and let $k > 1$ be the smallest such integer. By Lemma 11 there exists a (random) T such that for all t larger than T, in state $\mathfrak{g}$ the public log-likelihood is greater that the largest likelihood than can be induced by any private signal

$$Q_t^{k,\mathfrak{g},\mathfrak{f}} > M \cdot t. \tag{2.8.6}$$

Consider an outside observer x who observes H_t^k: the actions of agents $k-1$ and k. Hence x knows $Q_t^{k,\mathfrak{g},\mathfrak{f}}$ for all $t > T$, and by the same argument of the proof of Theorem 3, does not learn anything from k's action after time T. Hence x's speed of learning is r_{k-1}. As the outside observer x can observe k, by the imitation principle, x's speed of learning is at least that of agent k. This together implies that $r_{k-1} > M$, which is a contradiction to the original assumption on k. □

2.9 Intermittent and Random Observation Times

In this section we extend our model to allow for some intermittent and random observation times. This extension allows us to consider, for example, a situation in which one pair of agents meet every day, another pair meets every Sunday, and yet another pair meets on a random day of the week. Our main result still applies in this setting, and moreover the same proofs apply, with some additional details that need to be verified, as we explain.

Formally, for each pair of agents i, j such that $j \in N_i$ let $O_{i,j}$ be the set of time periods in which i observes j. These sets can be random, but we assume that they are independent of each other, the state and the signals. We also assume that there is some number $D > 0$ such that, with probability 1, for every i, j such that $j \in N_i$ and every $t \in \{0, 1, 2, \ldots\}$, the intersection $\{t+1, \ldots, t+D\} \cap O_{i,j}$ is not empty. That is, if $j \in N_i$ then i observes j at least once every D periods. Hence, the difference between consecutive $t_1, t_2 \in O_{i,j}$ is at most D.

The history of actions observed by agent i at time t is

$$H_t^i = \{a_s^j \ : \ s < t, s \in O_{i,j}, j \in N_i\}.$$

As before, the private history of agent i at time t is $I_t^i = (s_1^i, \ldots, s_t^i, H_t^i)$, and a pure strategy at time t is a map that assigns an action to each possible realization of I_t^i. The structure of agents' private signals and utilities remain the same. The speed of learning is likewise defined as before.

We now explain why, in this extended model, Theorem 3 still holds as stated. The proof of Lemma 9 applies verbatim, as only one agent is considered, and the observation structure plays no role. Lemma 10 likewise still applies, but an adjustment needs to be made: the imitation principle (2.8.3) again compares the loss the agent suffers in period t, assuming he suffers no loss in the future (on the left-hand-side), to the loss from always taking the action agent j took the last time $t - D$ he was observed by agent i (on the right-hand-side)

$$(1-\delta)(\bar{u} - \mathbb{E}[u(a_t^i, \omega)]) \leq \bar{u} - \mathbb{E}\left[u(a^j_{\tau(t)}, \omega)\right] \leq \max_{t' \in \{t-D-1,\ldots,t-1\}} \bar{u} - \mathbb{E}\left[u(a^j_{t'}, \omega)\right],$$

where $\tau(t) = \max O_{i,j} \cap \{0, \ldots, t-1\}$ is the last time (before t) that i observed j. Now continuing as in the proof of Lemma 10

$$\liminf_{t\to\infty} -\frac{1}{t}\log(\bar{u} - \mathbb{E}[u(a_t^i, \omega)]) \geq \liminf_{t\to\infty} -\frac{1}{t}\log(\bar{u} - \mathbb{E}[u(a_t^j, \omega)])$$

which implies the result of Lemma 10. Here we crucially use the fact that there are at most D periods between observations.

The proof of Lemma 11 remains valid, since the observation structure plays no role. The same holds for Lemma 12: The same proof applies to the modified version of the observed history H_t^i. Finally, the proof of Theorem 3 again applies verbatim.

Chapter 3

THE EMERGENCE OF FADS IN A CHANGING WORLD

3.1 Introduction

The term "fad" describes transient behavior that rises and fades rapidly in popularity, and in particular, these fast changes in behavior cannot be entirely explained by changes in the fundamentals. For example, in macroeconomics, there are boom-and-bust business cycles that cannot be pinned down by changes in the underlying economy.[1] In finance, it has long been documented that price deviation from the asset's intrinsic values can stem from speculative bubbles and fads [Aggarwal and Rivoli, 1990, Camerer, 1989]. While the phenomenon of fads is widely observed in many economic activities, the question of how and why fads emerge has yet to be resolved. In this paper, we show how fads—excessive changes in behavior relative to the fundamentals—can arise from social learning in an ever-changing environment.

The pioneering work in the social learning literature [Banerjee, 1992, Bikhchandani, Hirshleifer, and Welch, 1992, hereafter referred to as BHW] shows that under appropriate conditions, an *information cascade* always occurs. This is the event in which social information swamps agents' private information so that agents would follow the action of their predecessors, even if their private information suggests otherwise. However, since these cascades are typically formed based on limited information, their long-run outcome is fragile to small shocks. For example, as discussed in BHW, the possibility of a one-time change in the underlying state could result in "seemingly whimsical swings in mass behavior without obvious external stimulus," a phenomenon they refer to as fads. Inspired by their original idea, we introduce a formal definition of fads and study their long-term behavior.

While BHW present an early idea of a fad, they mainly focus on learning in a fixed environment where fads cannot recur indefinitely. In contrast, the recurrence of fads is possible in a changing environment, a setting that has recently attracted some attention [see, e.g., Dasaratha, Golub, and Hak, 2020, Lévy, Pęski, and Vieille, 2022]. Indeed, it is important to study this setting, as many applications of social learning, such as investment, employment, cultural norms, and technological advancement,

[1] See a recent study by Schaal and Taschereau-Dumouchel [2021] modeling business cycles through the lens of social learning.

often operate in a dynamic environment.

In this paper, we study social learning in a dynamic environment by adopting a simple two-state Markovian model with a small and symmetric transition probability. In this environment, we consider the canonical model of social learning with binary signals that are symmetric and informative about the current state.[2] We focus on the long-term behavior of agents who arrive sequentially and learn from observing the past actions of others as well as their own private signals. Given their information, each agent updates their belief about the state, acts once, and receives a positive payoff if her action matches the current state. As the underlying state evolves, the optimal action also fluctuates. The questions we aim to address are: how frequently do actions change? And more specifically, how often do they change compared to state changes?

It is important to note that in a dynamic environment, information cascades may occur, but they are only temporary. Hence, unlike in a static environment, agents do not permanently disregard their private signals [Moscarini et al., 1998]. Intuitively, once agents are in a cascade, there is no further influx of new information. However, as the state evolves, the social information supporting this cascade gradually loses relevance to the current agent. Thus, once this social information fades sufficiently, agents begin responding to their private signals again. We focus on the setting where the state evolves slowly, allowing the occurrence of these temporary cascades. This is because if otherwise, it is clear that actions would be more volatile than the state, as volatility in actions is purely caused by the noise in the signals, which are more volatile than the state itself.

Given the rises of temporary information cascades, the question of whether actions change more or less often than the state becomes unclear. On the one hand, due to information cascades, agents sometimes ignore their private signals, and so they do not change their actions even when the state changes. The symmetry of binary states further amplifies this effect: consider a scenario in which the state has changed an even number of times, say twice. However, agents in a cascade would mistakenly perceive the state as unchanged, so they would have no reason to change their actions. On the other hand, because signals are noisy, agents sometimes change actions unnecessarily. We define *fads* as a situation in which there are more action

[2] See other studies of social learning with a Markovian state, e.g., Hirshleifer and Welch [2002], Moscarini, Ottaviani, and Smith [1998] and Lévy, Pęski, and Vieille [2022]. Our model is mostly close to that in Moscarini, Ottaviani, and Smith [1998] except for the tie-breaking rule.

changes than state changes. Our main result (Theorem 4) shows that fads emerge in the long run, so that actions eventually switch more often than the state.

For example, consider a private signal that matches the current state 80 percent of the time. When the state changes once every 100 periods on average, we show that it takes less than sixty-one periods for agents to change their actions. As a consequence, the long-term frequency of action changes must be higher than that of state changes, thus leading to the emergence of fads in the long run. Arguably, this result is in line with the fragility of fads, where small shocks to the system could cause rapid shifts in the behavior of the agents.[3] We stress that in our model, agents act rationally, and hence the emergence of fads is driven by their desire to match with the ever-changing state, rather than any heuristics, irrationalities, or payoff externalities among the agents.

Our proof strategy behind the long-term emergence of fads is as follows. First, for any fixed signal precision and probability of state change, there exists a maximum length of information cascades. As a result, even though the rise of temporary cascades prolongs action inertia, such an effect is limited by its bounded length. Meanwhile, once agents exit a cascade, they only need one opposing signal to change their actions. This allows us to bound the probability of an action change from below, thus establishing an upper bound for the expected time between these changes. We then show that this upper bound is less than the expected time between state changes, implying that action changes occur more frequently than state changes on average. Finally, by translating the expected time between changes for both the state and the action into their long-term relative frequency of changes, we conclude that fads emerge in the long run.

Related Literature

This paper is closely related to a small stream of studies on social learning in a changing state. In BHW, they briefly discuss the case where a one-time shock to the state could break the cascade, even though that shock may never be realized. They provide a numeric example where the probability of an action change is at least 87% higher than the probability of a state change (see their Result 4) which is in line with our main result. Later, Moscarini, Ottaviani, and Smith [1998] show that if the underlying state is evolving in every period and it is sufficiently persistent, an

[3] As discussed in Bikhchandani, Hirshleifer, and Welch [1992, 1998], different kinds of small shocks, such as uncertainty in the underlying state as in our model or the arrival of a better-informed agent, could dislodge the previous trend and cause drastic behavioral changes.

information cascade must arise, but it only last temporarily, i.e., it must end in finite time. Our work builds on their model but with a different focus. Instead of analyzing the short-term patterns of information cascades, e.g., under what conditions do they end or arise, we ask: in the long run, should one expect more volatility in actions or in the state?

In a setting where a single agent repeatedly receives private signals, Hirshleifer and Welch [2002] examine the effect of *memory loss*—a situation in which the agent only recalls past actions but not past signals—on the continuity of the agent's behavior. Using a stylized five-period model, they show that in a relatively stable environment, memory loss induces excessive action inertia compared to a full-recall regime. In contrast, in a more volatile environment, memory loss results in excessive action impulsiveness.[4] Instead of comparing with a single agent with perfect memory, we are interested in understanding whether behavior under social learning (or amnesia in their language) is more or less volatile than the underlying fundamentals.

Among a few more recent studies that consider a dynamic state, the efficiency of learning has been a primary focus of study. For example, Frongillo, Schoenebeck, and Tamuz [2011] consider a specific environment in which the underlying state follows a random walk with non-Bayesian agents who use different linear rules when updating. Their main result is that the equilibrium updating weights may be Pareto suboptimal, causing inefficiency in learning.[5] In a similar but more general environment, Dasaratha, Golub, and Hak [2020] show that having sufficiently diverse network neighbors with different signal distributions improves learning. This is because diverse signals enable agents to extract the most relevant information from the old and confounded data, thereby achieving higher efficiency in information aggregation.

In a setup similar to ours, a recent study by Lévy, Pęski, and Vieille [2022] considers the welfare implication of a dynamic state. In their model, agents observe a random subsample drawn from all past actions and then decide whether to acquire private signals that are potentially costly. These model generalizations allow them to highlight the trade-off between learning efficiency and responsiveness to environmental

[4]Intuitively, as volatility of the environment increases, past actions become less relevant to the current state. At some point, this information weakens enough so that the amnesiac agent would always follow her latest signal, but the full-recall agent may not do so at this point. Hence, there is an increase in the probability of an action change due to amnesia.

[5]See more studies in the computer science literature, e.g., Acemoglu, Nedic, and Ozdaglar [2008], Shahrampour, Rakhlin, and Jadbabaie [2013] that consider a dynamic environment with non-Bayesian agents.

changes in maximizing equilibrium welfare. In contrast, we assume that agents observe the full history of past actions and there is no cost associated with obtaining their private signals. We consider this canonical sequential learning model without further complications as our focus is on comparing the long-term relative frequency of action and state changes—a question that turns out to be nontrivial even in this simple setup.

3.2 Model

Setup

We follow the setup from Moscarini, Ottaviani, and Smith [1998] closely. Time is discrete, and the horizon is infinite, i.e., $t \in \mathbb{N}_+ = \{1, 2, \ldots\}$. There is a binary state $\theta_t \in \{+1, -1\}$ that evolves according to a Markov chain with symmetric transition probability $\varepsilon \in (0, 1)$. That is,

$$\mathbb{P}[\theta_{t+1} \neq i | \theta_t = i] = \varepsilon, \text{ for } i \in \{+1, -1\}.$$

For simplicity, we assume that both states are equally likely at the beginning of time—a uniform distribution that is also the stationary distribution of this Markov chain.

A sequence of short-lived agents indexed by time t arrive in order, each acting once by choosing an action $a_t \in \{+1, -1\}$. For each agent t, she obtains a payoff of one if her action matches the current state, i.e., $a_t = \theta_t$ and zero otherwise. Before choosing an action, she receives a binary private signal $s_t \in \{+1, -1\}$ and observes the history of all past actions made by her predecessors $(a_1, \ldots, a_{t-1}) = h_{t-1} \in \{+1, -1\}^{t-1}$. Conditional on the entire sequence of states, these private signals (s_t) are independent and each s_t follows a Bernoulli distribution $B_{\theta_t}(\alpha)$ where $\alpha \in (1/2, 1)$ is the symmetric probability of matching the current state:

$$\mathbb{P}[s_t = i | \theta_t = i] = \alpha, \text{ for } i \in \{+1, -1\}.$$

We focus on the environment where the state is *sufficiently persistent*: for any signal precision $\alpha \in (1/2, 1)$, the probability of a state change $\varepsilon \in (0, \alpha(1 - \alpha))$. Under this assumption, Moscarini et al. [1998] show that information cascades can occur, but they only last temporarily. Equivalently, we can think of this assumption as follows: in every period, with probability $2\varepsilon \in (0, 2\alpha(1 - \alpha))$ the state will be redrawn from the set $\{+1, -1\}$ with equal probability. Thus, the probability of a state change is equal to ε, which lies in the open interval $(0, \alpha(1 - \alpha))$.

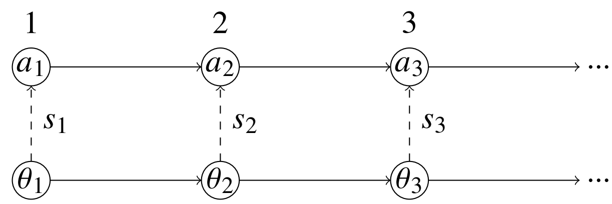

Figure 3.1: An illustration of processes (θ_t) and (a_t).

We summarize the timing of the events as follows (see a graphical illustration in Figure 3.1). At any time t, the agent arrives and observes the history of all past actions $h_{t-1} = (a_1, \ldots, a_{t-1})$. Then the state θ_{t-1} transitions to θ_t with probability ε of switching. After the transition of the state, agent t receives a private signal s_t that matches the current state θ_t with probability α. Finally, she chooses an action a_t that maximizes the probability of matching θ_t conditional on (h_{t-1}, s_t), the information available to her.

Fads

Given that each agent aims to match the current state, as the state evolves, the best action to take also fluctuates. BHW informally discuss the idea of faddish behavior as a situation where action changes occur more frequently than state changes. In other words, fads represent scenarios where there are *excessive* action changes. To formalize this idea, we denote the fraction of time periods $t \leq n$ for which $a_t \neq a_{t+1}$ by

$$Q_a(n) := \frac{1}{n} \sum_{t=1}^{n} \mathbb{1}(a_t \neq a_{t+1}).$$

Similarly, we denote the fraction of time periods $t \leq n$ for which $\theta_t \neq \theta_{t+1}$ by $Q_\theta(n) := \frac{1}{n} \sum_{t=1}^{n} \mathbb{1}(\theta_t \neq \theta_{t+1})$. We say that *fads* emerge at time $n+1$ if

$$Q_a(n) > Q_\theta(n). \tag{3.2.1}$$

Multiplying both sides of (3.2.1) by n, the emergence of fads at time $n+1$ implies that actions have changed more often than the state by time $n+1$.

Agents' Beliefs

Let $q_t := \mathbb{P}[\theta_t = +1 | h_{t-1}]$ denote the *public belief* assigned to the event $\theta_t = +1$ after observing h_{t-1} the history of actions at time t. Let $p_t := \mathbb{P}[\theta_t = +1 | h_{t-1}, s_t]$ denote the *posterior belief* of agent t assigned to the event $\theta_t = +1$ after observing a pair of observed history and private signal (h_{t-1}, s_t). Let L_t denote the log-likelihood ratio

(LLR) of the posterior belief of agent t:

$$L_t := \log \frac{p_t}{1-p_t} = \log \frac{\mathbb{P}[\theta_t = +1 | h_{t-1}, s_t]}{\mathbb{P}[\theta_t = -1 | h_{t-1}, s_t]}.$$

We refer to L_t as the *posterior LLR* at time t. By Bayes' rule,

$$L_t = \log \frac{\mathbb{P}[s_t | \theta_t = +1, h_{t-1}]}{\mathbb{P}[s_t | \theta_t = -1, h_{t-1}]} + \log \frac{\mathbb{P}[\theta_t = +1 | h_{t-1}]}{\mathbb{P}[\theta_t = -1 | h_{t-1}]}. \tag{3.2.2}$$

Since the private signal is independent of the history of actions conditional on the current state, the first term in (3.2.2) reduces to the LLR induced by the signal itself, which is equal to $c_\alpha := \log \frac{\alpha}{1-\alpha}$ if $s_t = +1$ and $-c_\alpha$ if $s_t = -1$. Denote the second term in (3.2.2) by

$$\ell_t := \log \frac{q_t}{1-q_t} = \log \frac{\mathbb{P}[\theta_t = +1 | h_{t-1}]}{\mathbb{P}[\theta_t = -1 | h_{t-1}]},$$

and we refer to it as the *public LLR* at time t. Intuitively, anyone who observes past actions from time 1 to $t-1$ can calculate this log-likelihood ratio.

In summary, depending on the realization of the private signal, the posterior LLR at time t evolves as follows:

$$L_t = \begin{cases} \ell_t - c_\alpha & \text{if } s_t = -1, \\ \ell_t + c_\alpha & \text{if } s_t = +1. \end{cases} \tag{3.2.3}$$

Agents' Behavior

The optimal action for agent t is the action that maximizes her expected payoff conditional on the information available to her:

$$a_t \in \operatorname*{arg\,max}_{a \in \{-1,+1\}} \mathbb{P}[\theta_t = a | h_{t-1}, s_t].$$

Thus $a_t = +1$ if $L_t > 0$ and $a_t = -1$ if $L_t < 0$. When $L_t = 0$, agent t is indifferent between both actions. We assume that she would follow what her immediate predecessor did in the previous period, i.e., $a_t = a_{t-1}$.[6] This tie-breaking rule differs from the one used in Moscarini, Ottaviani, and Smith [1998], where indifferent agents are assumed to follow their own private signals. We make this assumption so that any action changes are driven by agents' strict preference for one action over another, rather than by the specification of the tie-breaking rule.

[6] Our results do not depend on this assumption and are robust to any tie-breaking rule that is common knowledge.

Cascade and Learning Regions

An *information cascade* is the event in which the past actions of others form an overwhelming influence on agents so that they act independently of their private signals. Specifically, from (3.2.3), we see that when $|\ell_t| > c_\alpha$, the sign of L_t is purely determined by the sign of ℓ_t. Since the sign of L_t determines the optimal action of agent t, in this case, a_t is purely determined by the sign of ℓ_t, independent of s_t. That is, $a_t = +1$ if $\ell_t > c_\alpha$ and $a_t = -1$ if $\ell_t < -c_\alpha$. In the case where $|\ell_t| < c_\alpha$, agent t chooses the action according to her private signal: $a_t = s_t$. Finally, in the case where $|\ell_t| = c_\alpha$, the tie-breaking rule at indifference implies that regardless of the realization of s_t, agent t chooses the same action as agent $t-1$, i.e., $a_t = a_{t-1}$. Moreover, notice that $a_{t-1} = \text{sign}(\ell_t)$.[7] Together, we refer to the region in which $|\ell_t| \geq c_\alpha$ as the *cascade region* and the region in which $|\ell_t| < c_\alpha$ the *learning region*.

3.3 Results

A Benchmark

As a benchmark, we briefly discuss the case where each short-lived agent only observes her own private signal but not the actions of her predecessors. In this scenario, agent t simply follows her private signal s_t since it is the only information she has about the state.[8] As a consequence, agents' actions would change as often as their private signals in the long run. Hence, by the strong law of large numbers,

$$\lim_{n\to\infty} Q_a(n) = \mathbb{P}[s_t \neq s_{t+1}] \quad \text{almost surely,}$$

where a simple calculation shows that

$$\mathbb{P}[s_t \neq s_{t+1}] = (1-\alpha^2)(1-\varepsilon) + \alpha^2\varepsilon.$$

First, observe that the above probability is strictly higher than ε, which is equal to the long-term frequency of state changes (see further details in Section 3.4). This is intuitive because in this case agents always follow their signals, so actions would exhibit the same amount of volatility as the signals, which themselves are more volatile than the state. Second, notice that as the private signal becomes more

[7] Without loss of generality, consider the case where $\ell_t = c_\alpha$. On the one hand, if $s_t = +1$, then $L_t = \ell_t + c_\alpha > c_\alpha$, and so $a_t = +1$. On the other hand, if $s_t = -1$, then $L_t = \ell_t - c_\alpha = 0$ and by the tie-breaking rule, $a_t = a_{t-1}$. To see why $a_{t-1} = \text{sign}(\ell_t) = +1$, suppose to the contrary that $a_{t-1} = -1$. Given that $\ell_t = c_\alpha$ and $a_{t-1} = -1$, it must be $\ell_{t-1} > c_\alpha$, which implies $a_{t-1} = +1$. A contradiction.

[8] Note that this case is behaviorally equivalent to the scenario where the state is not sufficiently persistent and there are no temporary information cascades, leading agents to always follow their private signal.

precise (as $\alpha \to 1$), the long-run frequency of action changes converges to that of state changes. As signals become more precise, there are fewer unnecessary changes in actions, and in the limit, actions would change as often as the state in the long run.

Next, we turn to our main setting where each short-lived agent observes both her private signal and the actions of her predecessors. Recall that since the state is sufficiently persistent, i.e., $\varepsilon \in (0, \alpha(1-\alpha))$, there are temporary information cascades, during which agents follow the actions of their predecessors.

Public Actions

We now state our main result. Recall that $Q_a(\cdot)$ and $Q_\theta(\cdot)$ represent the fraction of time periods in which action and state changes occur, respectively, and we define in (3.2.1) the emergence of fads at time $n+1$ as $Q_a(n) > Q_\theta(n)$.

Theorem 4. *Fads emerge in the long run almost surely:*

$$\lim_{n\to\infty} Q_a(n) > \lim_{n\to\infty} Q_\theta(n) \quad \textit{almost surely.}$$

Perhaps surprisingly, Theorem 4 shows that even though there are times in which agents stop responding to their private signals, i.e., when cascades arise temporarily, agents who observe their predecessors' past actions still change their actions more often than the state in the long run. In other words, fads can emerge from social learning even when the underlying state evolves very slowly.

For example, consider a private signal that matches the current state 80 percent of the time. When the probability of state change is 1 percent, on average the state changes once every 100 periods. Meanwhile, the average time for the action to change is strictly less than 61 periods.[9] Thus, in the long run, actions would change strictly more often than the state, resulting in faddish behavior.

The idea behind the proof of Theorem 4 is as follows. Suppose that the agents are in an information cascade. As there is no further influx of new information, the social information contained in the cascade gradually becomes less relevant to the current agent as the state evolves. Consequently, at some point, this social information becomes obsolete enough for the agent to begin responding to her private signal again. In fact, a single opposing private signal to the current belief

[9] This follows from Proposition 9 in §3.4 by substituting $\alpha = 0.8$ and $\varepsilon = 0.01$ into $M(\alpha, \varepsilon)$, and we have $M(0.8, 0.01) \approx 60.7$.

can trigger an action switch. This allows us to bound the probability of initiating a new information cascade from above, establishing an upper bound for the expected time between action switches (Proposition 9). It turns out that the maximum of this upper bound is strictly less than the expected time between state changes. Building on this, we conclude that the long-run relative frequency of action changes is higher than that of state changes.

We note that because action switches are not independent events, the connection between the expected time between action changes and its long-run relative switching frequency does not directly follow from the standard result of the law of large numbers. To address this challenge, we study the process of the random time elapsed between action changes, which has well-defined moments (see Lemma 14 in the appendix). Moreover, unlike the fixed-state model, when the state changes over time, the agent's public belief about the current state ceases to be a martingale—a property that is crucial for analyzing the long-term outcome of learning.[10] Nevertheless, the public belief remains to be a Markov process, and we will rely on its specific transitional patterns to analyze the expected time between its sign switches (see Lemma 13 in Section 3.4).

We make the assumption of a sufficiently persistent state for two reasons. Firstly, in a scenario where the state is not sufficiently persistent—thus temporary information cascades never arise—agents would always follow their signals and change their actions accordingly. As a result, as in our benchmark case, action changes clearly occur more frequently than state changes. Secondly, even with a persistent state, it is a priori unclear whether the state or the action changes more often. This is because as the likelihood of state changes decreases, the rate of action changes also slows down. Intuitively, when state changes are less likely, past actions become more informative about the current state. As a consequence, temporary information cascades tend to persist longer, leading to extended periods of action inertia. Our main result suggests that this prolonged action inertia is eventually surpassed by action impulsiveness induced by noisy signals, leading to excessive changes in actions relative to the state in the long run.

Numerical Simulations

To complement the asymptotic result of Theorem 4, we simulate the empirical frequencies of action and state changes under different parameter values of signal

[10]For example, the martingale property is essential in proving that all agents eventually herd on the correct action in a setting with a fixed state and unbounded signals [Smith and Sørensen, 2000].

$\alpha \setminus \varepsilon$	0.05	0.1	0.2
0.51	16,766 (5,081)	28,564 (10,055)	42,128 (20,024)
0.75	15,240 (5,100)	26,149 (10,034)	–
0.9	14,252 (5,096)	–	–

Table 3.1: The numerical simulations of the number of action and state changes (in parentheses) under different values of α and $\varepsilon \in (0, \alpha(1-\alpha))$ for 100,000 periods.

precision (α) and state volatility (ε) over 100, 000 periods. The results are reported in Table 3.1.

Observe first that these numerical simulations confirm our main result: for any pair of parameter values that we consider, action changes are more frequent than state changes. Next, we examine how the magnitude of action and state changes varies across different parameter values. As illustrated in the first column of Table 3.1, we observe a decrease in the frequency of action changes as the precision of the private signal increases. This aligns with the intuition that more precise signals should mitigate unnecessary action changes. Conversely, as state volatility increases, both action and state changes occur more frequently. However, as shown in the first row of Table 3.1, the ratio between these changes decreases from approximately 3.3 to 2.1. This suggests that the indirect effect of state volatility on action changes is less significant than its direct effect on state changes.

3.4 Analysis

In this section, we begin by analyzing how public belief evolves in different regions, namely the cascade and learning regions. We then compare the expected time between the state and action changes.

The Public Belief Dynamics

Cascade Region

As is well-known, in our setting with a fixed state ($\varepsilon = 0$), the public belief q stays forever at the value at which it first enters the cascade region; thus, an incorrect cascade can persist indefinitely with positive probability. In contrast, with a changing state ($\varepsilon > 0$), the behavior of the public belief varies drastically.

To see this, consider the case where the public belief q is in the cascade region, so its corresponding LLR $\ell = \log \frac{q}{1-q}$ has an absolute value that is greater or equal to c_α. Suppose t is the time at which public belief enters the cascade region from the learning region. Consequently, agent t follows the action of her immediate predecessor, and thus her action a_t contains the same amount of information about θ_t as a_{t-1}. Meanwhile, from time t to $t+1$, the state continues to evolve with probability ε of changing. Given that the process (θ_t) is a Markov chain, conditional on θ_t, the history h_{t-1} contains no additional information about θ_{t+1}. By the law of total probability, the public belief updates deterministically as follows:

$$\begin{aligned} q_{t+1} = \mathbb{P}[\theta_{t+1} = +1 | h_t] &= \sum_{i \in \{-1,+1\}} \mathbb{P}[\theta_{t+1} = +1 | h_t, \theta_t = i] \mathbb{P}[\theta_t = i | h_t] \\ &= \sum_{i \in \{-1,+1\}} \mathbb{P}[\theta_{t+1} = +1 | h_{t-1}, \theta_t = i] \mathbb{P}[\theta_t = i | h_t] \\ &= (1-\varepsilon) q_t + \varepsilon (1 - q_t). \end{aligned}$$

Thus, we can write

$$q_{t+1} = (1-2\varepsilon) q_t + (2\varepsilon)\frac{1}{2}. \tag{3.4.1}$$

Equivalently, it can expressed recursively in terms of its LLR:

$$\ell_{t+1} = \log \frac{q_{t+1}}{1 - q_{t+1}} = \log \frac{(1-\varepsilon)\mathrm{e}^{\ell_t} + \varepsilon}{1 - \varepsilon + \varepsilon \mathrm{e}^{\ell_t}}. \tag{3.4.2}$$

From (3.4.1), we see that q_{t+1} tends to 1/2. Similarly, according to (3.4.2), ℓ_{t+1} moves towards zero over time, so eventually it will exit the cascade region. Intuitively, having a changing state depreciates the value of older information, as actions observed in earlier periods become less relevant to the current agent. Consequently, after some finite number of periods, the public belief will slowly converge towards uniformity, and thus cascades supported by this public belief will eventually cease. This is the main insight from Moscarini, Ottaviani, and Smith [1998], where they show that information cascades—if they arise—must end in finite time with a changing state.

Learning Region

Recall that when the state is fixed ($\varepsilon = 0$) and the public belief is in the learning region, agents will follow their private signals, i.e., $a_t = s_t$. As a result, the public belief at time $t+1$ coincides with the posterior belief of agent t:

$$q_{t+1} := \mathbb{P}[\theta = +1 | h_{t-1}, a_t] = \mathbb{P}[\theta = +1 | h_{t-1}, s_t] = p_t.$$

Hence, the corresponding LLRs also coincide: $\ell_t = L_t$. and they both evolve according to (3.2.3).

In contrast, when the state changes with probability $\varepsilon > 0$ in every period, upon observing the latest history, each agent needs to consider the possibility that the state may have changed after the latest action was taken. However, neither the learning nor the cascade region is affected by a changing state as the state only transitions after the history of past actions is observed. It follows from the Bayes' rule that

$$\ell_{t+1} = \log \frac{\sum_{i\in\{-1,+1\}} \mathbb{P}[\theta_{t+1} = +1, a_t | h_{t-1}, \theta_t = i]\mathbb{P}[\theta_t = i | h_{t-1}]}{\sum_{i\in\{-1,+1\}} \mathbb{P}[\theta_{t+1} = -1, a_t | h_{t-1}, \theta_t = i]\mathbb{P}[\theta_t = i | h_{t-1}]}. \tag{3.4.3}$$

Since the process (θ_t) follows a Markov chain and $a_t = s_t$ in the learning region, conditional on θ_t, both θ_{t+1} and a_t are independent of the history h_{t-1} and independent of each other. Therefore, the process of (ℓ_t) evolves as follows:

$$\ell_{t+1} = \begin{cases} \log \frac{(1-\varepsilon)\alpha e^{\ell_t} + \varepsilon(1-\alpha)}{\varepsilon\alpha e^{\ell_t} + (1-\varepsilon)(1-\alpha)} := f_+(\ell_t) & \text{if } s_t = +1, \\ \log \frac{(1-\varepsilon)(1-\alpha)e^{\ell_t} + \varepsilon\alpha}{\varepsilon(1-\alpha)e^{\ell_t} + (1-\varepsilon)\alpha} := f_-(\ell_t) & \text{if } s_t = -1. \end{cases} \tag{3.4.4}$$

Note that both $f_+(\ell)$ and $f_-(\ell)$ are strictly increasing in ℓ. Moreover, $f_+(\ell) > \ell$ and $f_-(\ell) < \ell$. This means that when an agent starts with a higher prior belief, her posterior belief will always be higher conditioned on receiving a private signal. Similarly, an agent's posterior belief will be higher (or lower) than her prior belief upon receiving a positive signal (or a negative signal).

Observe from (3.4.4) that the magnitude difference between ℓ_t and ℓ_{t+1} depends on both the realization of the private signal s_t and the current value of ℓ_t. The following lemma summarizes the transitional patterns of the public LLR when it is the learning region. At any time t, we say that an action is *opposing* to the current public belief if $a_t \neq \text{sign}(\ell_t)$ and *supporting* otherwise. The following lemma is in spirit close to the *overturning principle* in Smith and Sørensen [2000], but applies to a changing state.

Lemma 13. *Suppose the public belief at time t is in the learning region, i.e., $|\ell_t| < c_\alpha$. Then the following two conditions hold.*

(i) $a_t \neq sign(\ell_t)$ implies that $sign(\ell_{t+1}) = -sign(\ell_t)$.

(ii) $a_t = a_{t+1} = sign(\ell_t)$ implies that $|\ell_{t+2}| \geq c_\alpha$.

The first part of this lemma says that observing one opposing action is sufficient to overturn the sign of public belief. The second part of this lemma states that initiating a cascade requires at most two supporting actions. Intuitively, given that public belief in the learning region tends to be moderate, it is sensitive to opposing evidence; meanwhile, although the belief updating process is impeded by a changing state, consecutive observations of supporting evidence suffice to trigger a cascade as the state evolves slowly.

Another important observation is that regardless of whether the state is fixed or changing, the process (ℓ_t) forms a Markov chain.[11] In the case of a fixed state, the state space of this Markov chain is finite as the magnitude difference between ℓ_t and ℓ_{t+1} is constant given any signal precision. However, in the case of a changing state, its state space is infinite as such magnitude differences also depend on the current value of ℓ_t.[12] This poses a significant challenge in finding its stationary distribution, which is required to calculate the exact expected time between sign switches. We circumvent this problem by providing an upper bound to this expected time instead.

Expected Time Between Switches

We first calculate the expected time between state changes. Since the process (θ_t) follows a simple two-state Markov chain with a symmetric transition probability ε, the expected time between state changes is inversely proportional to the likelihood of a state change. To see this, suppose that the expected time between state changes is equal to some unknown x. Then x satisfies the following equation:

$$x = \varepsilon + (1 - \varepsilon)(1 + x),$$

which implies that $x = 1/\varepsilon$. That is, a higher likelihood of state changes results in a shorter average time between changes.

Given that the process (a_t) is not a Markov chain, determining the average time it takes to change becomes more challenging. Nevertheless, we observe that a_t is a function of ℓ_{t+1}, i.e., $a_t = \text{sign}(\ell_{t+1})$, which is a Markov chain.[13] As discussed before, however, this Markov chain (ℓ_t) is complicated—with infinitely many possible values and different transition probabilities—making it difficult to directly analyze the expected time between its sign switches. We provide an upper bound to this expected time instead, which in turn gives an upper bound to the expected time between action changes.

[11] This is because conditional on the state θ_t, the private signal s_t is independent of ℓ_τ, for any $\tau < t$.

[12] In fact, in almost all cases, two consecutive opposing signals do not exactly offset each other, i.e., $f_+(f_-(\ell)) \neq \ell$ and vice versa.

[13] Note that for a given function $f : \mathcal{X} \to \mathcal{Y}$ with $|\mathcal{X}| < \infty$, and a Markov chain $(X_t)_t$, the process $(Y_t)_t$ where each $Y_t = f(X_t)$, is a Markov chain if and only if f is either constant or injective. Clearly, the sign function is neither constant nor injective.

To do so, consider the maximum length of any cascade. Recall that such a maximum exists since the public belief in the cascade region slowly converges towards uniformity. Note that for any signal precision and probability of a state change, no cascade can last longer than the cascade starting at $f_+(c_\alpha)$, the supremum of the public LLR. Thus, from (3.4.2), we can calculate a tight upper bound to the length of any cascade. Following Moscarini, Ottaviani, and Smith [1998], we denote this upper bound by $K(\alpha, \varepsilon)$ where[14]

$$K(\alpha, \varepsilon) = \frac{\log(1 - 2\alpha(1 - \alpha))}{\log(1 - 2\varepsilon)}.$$

Notice that $K(\alpha, \varepsilon)$ decreases in both α and ε. Intuitively, as private signals become less precise, cascades contain more information relative to private signals. Moreover, as the state becomes less volatile, temporary cascades last longer as social information depreciates at a lower rate. In other words, prolonged cascades may result from less precise private signals or a more stable environment.

Given the maximum length of any cascade and the transitional patterns of the public LLR (Lemma 13), we obtain an upper bound to the expected time between the sign switches of the public LLR, as shown in the next proposition. For $i = 1, 2, \ldots$, denote the random time at which the public LLR switches its sign for the i-th time by $\mathcal{T}_i$ and let $\mathcal{T}_0 = 0$. Furthermore, denote the random time elapsed between the $i-1$-th and i-th sign switch by $\mathcal{D}_i = \mathcal{T}_i - \mathcal{T}_{i-1}$.

Proposition 9. *For any positive integers $i \geq 2$, conditional on the public LLR that just switched its sign for the $i-1$-th time, the expected time to the next sign switch $\mathbb{E}[\mathcal{D}_i | \ell_{\mathcal{T}_{i-1}}]$ is strictly bounded above by $M(\alpha, \varepsilon)$ where*

$$M(\alpha, \varepsilon) = 1 + \frac{K(\alpha, \varepsilon)}{2\alpha(1 - \alpha)}.$$

[14] For completeness, we provide a similar calculation of $K(\alpha, \varepsilon)$ to the one in Section 3.B of Moscarini, Ottaviani, and Smith [1998]. Fix any arbitrary $\alpha \in (1/2, 1)$ and $\varepsilon \in (0, \alpha(1 - \alpha))$. Denote m as the supremum of public belief, where $m = \frac{(1-\varepsilon)\alpha^2 + \varepsilon(1-\alpha)^2}{\alpha^2 + (1-\alpha)^2}$. Since the public belief in a cascade evolves deterministically according to (3.4.1), after h periods, the public belief starting at m equals

$$g(h) := \varepsilon \sum_{i=1}^{h-1} (1 - 2\varepsilon)^i + (1 - 2\varepsilon)^h m.$$

This implies that any public belief after spending h periods in the cascade region would have a value strictly lower than $g(h)$. Thus, whenever $g(h) \leq \alpha$, or equivalently $(1 - 2\varepsilon)^{h+1} \leq 1 - 2\alpha(1 - \alpha)$, the public LLR of value $g(h)$ would have exited the cascade region. Hence, the maximum number of periods in which public LLR can stay in the cascade region is $\frac{\log(1-2\alpha(1-\alpha))}{\log(1-2\varepsilon)}$.

This proposition states that on average, one should expect the public LLR to change its sign at least once every $M(\alpha, \varepsilon)$ periods. For example, with $\alpha = 0.8$ and $\varepsilon = 0.01$, $M(0.8, 0.01)$ is approximately equal to 61, and so once every 61 periods, there is at least one sign switch in the public LLR. Note that $M(\alpha, \varepsilon)$ also decreases in α. As a result, $M(1/2, \varepsilon)$ is the maximum upper bound for any $\varepsilon \in (0, \alpha(1-\alpha))$. This suggests that when private signals are only weakly informative, agents rely more on social information, which potentially leads to longer action inertia.

We illustrate the proof idea of Proposition 9 using a weakly informative signal. Suppose that $\alpha = 1/2 + \delta$ where δ is strictly positive and close to 0. Since $K(\alpha, \varepsilon)$ decreases in α, $K(\frac{1}{2}, \varepsilon)$ is the greatest upper bound to the length of any cascade as δ approaches zero. For sufficiently small δ, upon exiting a cascade, the probability of the public LLR switching its sign is approximately 1/2, given that the agent, who follows her private signal, receives either a positive or a negative signal with almost equal probability. Hence, the expected time between the sign switches is bounded from above by a geometric distribution:

$$
\begin{aligned}
1 + \frac{1}{2}\Big(K(\frac{1}{2}, \varepsilon) + \frac{1}{2}\big(K(\frac{1}{2}, \varepsilon) + \frac{1}{2}(K(\frac{1}{2}, \varepsilon) + \ldots &= 1 + \sum_{i=1}^{\infty} \frac{i}{2^i} K(\frac{1}{2}, \varepsilon) \\
&= 1 + \frac{2\log 2}{-\log(1-2\varepsilon)} = M(1/2, \varepsilon).
\end{aligned}
$$

It is straightforward to verify that $M(1/2, \varepsilon) < 1/\varepsilon$ (see Claim 4 in the appendix). Hence, any upper bound $M(\alpha, \varepsilon)$ provided in Proposition 9 is strictly less than $1/\varepsilon$. Given that the action is a function of the public LLR, an immediate consequence of Proposition 9 is the following result.

Corollary 3. *The expected time between action changes is strictly less than that between state changes.*

That is, on average, actions take less time to change than the state, even for a small probability of state change in which temporary cascades arise. For example, when the probability of state change is equal to 0.05, the state changes every twenty periods on average. In comparison, the maximum average time for the action to change is less than fourteen periods. Our main result (Theorem 4) then builds on this result by connecting the expected time between action changes with its long-term relative frequency of changes.

3.5 Conclusion

We study the long-term behavior of agents who receive a private signal and observe the past actions of their predecessors in a changing environment. As the state evolves, the best action to take also fluctuates. We show that in the long run, the relative frequency of action changes is higher than that of state changes, suggesting fads can emerge from social learning in a changing environment.

Instead of considering the benchmark case where each short-lived agent receives one private signal and acts accordingly, one could study the frequency of action changes for a single long-lived agent who repeatedly receives private signals about a changing state. In this case, we conjecture that action changes would be less frequent than in our main setting—where only past actions are observable—but still more frequent than state changes. Intuitively, shutting down the channel of noisy observations of others' private signals would reduce the frequency of unnecessary action changes but not completely. As a result, we expect that the remaining action changes would still occur more frequently than state changes.

One may wonder if the driving force behind our main result is due to the high frequency of action changes when the posterior belief is around 1/2. Accordingly, we can further restrict the definition of fads to action changes that do not have consecutive changes, i.e., $a_t \neq a_{t-1}$ and $a_{t-1} \neq a_{t-2}$. Simulation results show that actions still change more frequently than the state, even under this more restricted definition of fads. For example, for $\alpha = 0.75$, $\varepsilon = 0.05$ and a total of $100,000$ periods, the action changes about 8,150 times which is more frequent than the number of state changes, which is about 5,100 times.

There are several possible avenues for future research. Recall that Proposition 9 implies that $M(\alpha, \varepsilon)$ is an upper bound to the expected time between action changes. One could ask whether this upper bound $M(\alpha, \varepsilon)$ is tight, and if so, for any finite time N, whether the number of action changes would be close to $N/M(\alpha, \varepsilon)$. Based on the simulation results, we conjecture that it is not a tight bound. E.g., we let $\alpha = 0.9$ and $\varepsilon = 0.05$, and $N = 100,000$. Since $M(0.9, 0.05) \approx 11.5$, it implies that within these hundred thousand periods, the action should change at least about 8700 times. However, our numerical simulation shows that the action changes about $14,200$ times, almost double the number suggested by $M(0.9, 0.05)$. Furthermore, our simulations suggest that as the private signal becomes less informative and the state changes more slowly, i.e., when α approaches $1/2$ and ε approaches 0 at the same rate, the ratio between the frequency of action changes and state changes approaches a constant that is close to 4. This suggests that achieving a very accurate understanding of fads in this regime might be possible.

3.6 Proofs

Proof of Lemma 13. Fix any arbitrary $\alpha \in (1/2, 1)$ and $\varepsilon \in (0, \alpha(1-\alpha))$. Without loss of generality, consider the case where $\ell_t \in (0, c_\alpha)$.

(i) Suppose $a_t = -1$. Given that $f_-(c_\alpha) = 0$, we have that $f_-(\ell_t) < 0$ for all $\ell_t \in (0, c_\alpha)$ since $f_-(\cdot)$ in (3.4.4) is strictly increasing. Since $\ell_t \in (0, c_\alpha)$ is in the learning region, $s_t = a_t = -1$. Thus, it follows from (3.4.4) that $\ell_{t+1} = f_-(\ell_t)$ and $\text{sign}(f_-(\ell_t)) = -1$.

(ii) Since $f_+(\cdot)$ in (3.4.4) is strictly increasing, it suffices to show that $f_+(f_+(0)) \geq c_\alpha$. Denote by c_u the threshold at which exactly one positive signal is required to push the public LLR into a cascade on the positive action:

$$c_u := f_+^{-1}(c_\alpha) = \log \frac{(1-\alpha)(\alpha-\varepsilon)}{\alpha(1-\alpha-\varepsilon)} \in (0, c_\alpha).$$

Notice that $f_+(0) > c_u$ for all $\varepsilon \in (0, \alpha(1-\alpha))$, and thus $f_+(f_+(0)) > f_+(c_u)$. By the definition of c_u, we have $f_+(f_+(0)) > f_+(c_u) = f_+(f_+^{-1}(c_\alpha)) = c_\alpha$, as required.

□

Proof of Proposition 9. Let $\alpha \in (1/2, 1)$, $\varepsilon \in (0, \alpha(1-\alpha))$ and $i \geq 2$ be a positive integer. Recall that $\mathcal{T}_{i-1}$ is the time at which the public LLR changes its sign for the $i-1$-th time and $\mathcal{D}_i = \mathcal{T}_i - \mathcal{T}_{i-1}$. Suppose that $\ell_{\mathcal{T}_{i-1}} > 0$. Then, there are three disjoint intervals in which the value of $\ell_{\mathcal{T}_{i-1}}$ can be: (i) $[c_u, c_\alpha)$ where $c_u = f_+^{-1}(c_\alpha)$; (ii) $(0, c_u)$ and (iii) $[c_\alpha, f_+(c_\alpha))$. We will show that $\mathbb{E}[\mathcal{D}_i | \ell_{\mathcal{T}_{i-1}} > 0] < M(\alpha, \varepsilon)$.

Conditional on $\ell_t = \ell$, let the probability of receiving a positive signal be $\pi(\ell)$.[15] By the law of total probability,

$$\pi(\ell) = \alpha \cdot \frac{e^\ell}{1+e^\ell} + (1-\alpha) \cdot \frac{1}{1+e^\ell} = \frac{1+\alpha(e^\ell - 1)}{1+e^\ell}. \tag{3.6.1}$$

Since $\pi(\ell)$ is strictly increasing in ℓ, the supremum of $\pi(\ell)$ over all $\ell \in (0, c_\alpha)$ is equal to $1 - 2\alpha(1-\alpha)$, and we denote it by $\bar{\pi}$.

Furthermore, let $\kappa(\ell)$ denote the length of a positive cascade triggered by receiving a positive signal conditional on $\ell_t = \ell$. Let $\mathcal{L}(\ell)$ represent the resulting value of the

[15] For ease of notation, we suppress its dependence on α.

public LLR after exiting the cascade region for the first time. We use $\lfloor K(\alpha,\varepsilon)\rfloor$ to denote the greatest integer that is less than or equal to $K(\alpha,\varepsilon)$.

Case (i). Suppose $\ell_{\mathcal{T}_{i-1}} \in [c_u, c_\alpha)$. By part (i) of Lemma 13, since $\ell_{\mathcal{T}_{i-1}}$ is in the learning region, one opposing signal is sufficient to change the sign of $\ell_{\mathcal{T}_{i-1}}$. Thus, the expected time to the next sign switch

$$\begin{aligned}\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] &= 1-\pi(\ell_{\mathcal{T}_{i-1}})+\pi(\ell_{\mathcal{T}_{i-1}})\Big(\kappa(\ell_{\mathcal{T}_{i-1}})+\mathbb{E}[\mathcal{D}_i|\mathcal{L}(\ell_{\mathcal{T}_{i-1}})]\Big)\\ &< 1-\bar{\pi}+\bar{\pi}\Big(\lfloor K(\alpha,\varepsilon)\rfloor+\mathbb{E}[\mathcal{D}_i|\mathcal{L}(\ell_{\mathcal{T}_{i-1}})]\Big), \end{aligned} \tag{3.6.2}$$

where the last inequality follows from the definition of $\bar{\pi}$. There are two possible cases for $\mathcal{L}(\ell_{\mathcal{T}_{i-1}})$: either $\mathcal{L}(\ell_{\mathcal{T}_{i-1}}) \in [c_u, c_\alpha)$ or $\mathcal{L}(\ell_{\mathcal{T}_{i-1}}) \in (0, c_u)$. If the former is the case, then by taking the supremum on both sides of (3.6.2) and rearranging,

$$\sup_{c_u\le \ell_{\mathcal{T}_{i-1}}<c_\alpha} \mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] \le 1+\frac{\bar{\pi}}{1-\bar{\pi}}\lfloor K(\alpha,\varepsilon)\rfloor. \tag{3.6.3}$$

If the latter is the case, by the definition of $\bar{\pi}$,

$$\mathbb{E}[\mathcal{D}_i|\mathcal{L}(\ell_{\mathcal{T}_{i-1}})] < 1-\bar{\pi}+\bar{\pi}\Big(1+\mathbb{E}[\mathcal{D}_i|f_+(\mathcal{L}(\ell_{\mathcal{T}_{i-1}}))]\Big).$$

Substituting the above inequality into (3.6.2),

$$\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] < 1-\bar{\pi}+\bar{\pi}\Big(\lfloor K(\alpha,\varepsilon)\rfloor+1-\bar{\pi}+\bar{\pi}\big(1+\mathbb{E}[\mathcal{D}_i|f_+(\mathcal{L}(\ell_{\mathcal{T}_{i-1}}))]\big)\Big).$$

Since $f_+(\cdot)$ is strictly increasing, by part (ii) of Lemma 13, $f_+(\mathcal{L}(\ell_{\mathcal{T}_{i-1}})) \in [c_u, c_\alpha)$. Thus, taking the supremum on both sides and rearranging,

$$\begin{aligned}\sup_{c_u\le \ell_{\mathcal{T}_{i-1}}<c_\alpha} \mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] &\le \frac{1-\bar{\pi}+(\lfloor K(\alpha,\varepsilon)\rfloor+1)\bar{\pi}}{1-\bar{\pi}^2}\\ &\le 1+\frac{\bar{\pi}}{1-\bar{\pi}}\lfloor K(\alpha,\varepsilon)\rfloor,\end{aligned}$$

where the second inequality holds since $\lfloor K(\alpha,\varepsilon)\rfloor \ge 1$.

Case (ii). Suppose $\ell_{\mathcal{T}_{i-1}} \in (0, c_u)$. By part (i) of Lemma 13 and the definition of $\bar{\pi}$, the expected time to the next sign switch is bounded above:

$$\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] < (1-\bar{\pi})+\bar{\pi}(1+\mathbb{E}[\mathcal{D}_i|f_+(\ell_{\mathcal{T}_{i-1}})]).$$

Since $f_+(\cdot)$ is strictly increasing, part (ii) of Lemma 13 implies that $f_+(\ell_{\mathcal{T}_{i-1}}) \in [c_u, c_\alpha)$. It then follows from (3.6.3) that

$$\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] < \frac{\bar{\pi}^2(\lfloor K(\alpha,\varepsilon)\rfloor-1)+1}{1-\bar{\pi}}. \tag{3.6.4}$$

Case (iii). Suppose $\ell_{\mathcal{T}_{i-1}} \in [c_\alpha, f_+(c_\alpha))$. In this case, after at most $\lfloor K(\alpha,\varepsilon)\rfloor$ periods, the public LLR initiated at $\ell_{\mathcal{T}_{i-1}}$ would have exited the cascade region. Hence, the expected time to the next sign switch is bounded above:

$$\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] \leq \lfloor K(\alpha,\varepsilon)\rfloor + \mathbb{E}[\mathcal{D}_i|\mathcal{L}(\ell_{\mathcal{T}_{i-1}})].$$

Again, there are two possible cases for $\mathcal{L}(\ell_{\mathcal{T}_{i-1}})$: either $\mathcal{L}(\ell_{\mathcal{T}_{i-1}}) \in [c_u, c_\alpha)$ or $\mathcal{L}(\ell_{\mathcal{T}_{i-1}}) \in (0, c_u)$. If the former is the case, it follows from (3.6.3) that

$$\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] < 1 + \frac{1}{1-\bar{\pi}}\lfloor K(\alpha,\varepsilon)\rfloor. \tag{3.6.5}$$

If the latter is the case, then it follows from (3.6.4) that

$$\begin{aligned}\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}] &< \lfloor K(\alpha,\varepsilon)\rfloor + \frac{\bar{\pi}^2(\lfloor K(\alpha,\varepsilon)\rfloor - 1) + 1}{1-\bar{\pi}} \\ &= \lfloor K(\alpha,\varepsilon)\rfloor + 1 + \bar{\pi} + \frac{\bar{\pi}^2}{1-\bar{\pi}}\lfloor K(\alpha,\varepsilon)\rfloor \leq 1 + \frac{1}{1-\bar{\pi}}\lfloor K(\alpha,\varepsilon)\rfloor.\end{aligned}$$

Now, notice that the maximum of these three upper bounds in (3.6.3) to (3.6.5) is $1 + \frac{1}{1-\bar{\pi}}\lfloor K(\alpha,\varepsilon)\rfloor$ and, by definition, $\lfloor K(\alpha,\varepsilon)\rfloor \leq K(\alpha,\varepsilon)$. Hence, we conclude that

$$\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}} > 0] < 1 + \frac{K(\alpha,\varepsilon)}{2\alpha(1-\alpha)}.$$

The case where $\ell_{\mathcal{T}_{i-1}} < 0$ follows from a symmetric argument. Thus, by the law of iterated expectations, the above inequality also holds for $\mathbb{E}[\mathcal{D}_i|\ell_{\mathcal{T}_{i-1}}]$. □

Claim 4. *$M(1/2,\varepsilon) < 1/\varepsilon$ for all $\varepsilon \in (0, 1/4)$.*

Proof. Recall that

$$M(1/2,\varepsilon) = \sup_{\alpha\in(1/2,1)} M(\alpha,\varepsilon) = 1 + \frac{2\log 2}{-\log(1-2\varepsilon)}.$$

By the L'Hôpital's rule

$$\lim_{\varepsilon\to 0} -(\frac{1}{\varepsilon} - 1)\log(1-2\varepsilon) = \lim_{\varepsilon\to 0} 2\frac{(1-\varepsilon)^2}{1-2\varepsilon} = 2 > 2\log 2.$$

Since $-(\frac{1}{\varepsilon} - 1)\log(1-2\varepsilon)$ is strictly increasing in ε, we conclude that $M(1/2,\varepsilon) < 1/\varepsilon$ for all $\varepsilon \in (0, 1/4)$. □

Recall that since $M(\alpha,\varepsilon)$ decreases in α, it follows from Claim 4 and Proposition 9 that for all $\alpha \in (1/2, 1)$ and $\varepsilon \in (0, \alpha(1-\alpha))$,

$$M(\alpha,\varepsilon) < 1/\varepsilon. \tag{3.6.6}$$

The following lemma will be useful in proving Theorem 4. It shows that the process $(\mathcal{D}_i)$ has well-defined moments. In particular, it implies that there is a finite uniform upper bound to $\mathbb{E}[\mathcal{D}_i^2]$, which is required to apply the standard martingale convergence theorem. Intuitively, since any cascade must end after $K(\alpha, \varepsilon)$ periods, the probability that $\mathcal{D}_i$ is larger than some finite periods decreases exponentially fast, and so $\mathcal{D}_i$ must have finite moments.

Lemma 14. *For every $r \in \{1, 2, \ldots\}$ there is a constant c_r that depends on α and ε such that for all i, $\mathbb{E}[|\mathcal{D}_i|^r] < c_r$. I.e., each moment of $\mathcal{D}_i$ is uniformly bounded, independently of i.*

Proof. Fix any arbitrary $\alpha \in (1/2, 1)$, $\varepsilon \in (0, \alpha(1-\alpha))$ and some positive integer $i \geq 2$. Suppose that $\ell_{\mathcal{T}_{i-1}} > 0$ and so $\mathcal{D}_i = \mathcal{T}_i - \mathcal{T}_{i-1}$ is the time elapsed from a positive public LLR to a negative one. For any $n \geq 2$, we denote the minimum number (which may not be an integer) of temporary cascades required for $\mathcal{D}_i > n$ by

$$k(n) := \max\left\{\frac{n-1}{\lfloor K(\alpha, \varepsilon)\rfloor}, 1\right\}.$$

Recall that $\bar{\pi}$ is the supremum of the probability of receiving a high signal conditional on ℓ for all $\ell \in (0, c_\alpha)$. By part (ii) of Lemma 13, for any $n \geq 2$, the probability of the event $\{\mathcal{D}_i > n\}$ is bounded above:

$$\mathbb{P}[\mathcal{D}_i > n] < \bar{\pi}^{2+(\lfloor k(n)\rfloor - 1)}.$$

Since $\mathcal{D}_i$ is a positive random variable, it follows that for any $p > 0$,

$$\begin{aligned} \lim_{n\to\infty} n^p \mathbb{P}[|\mathcal{D}_i| > n] &= \lim_{n\to\infty} \frac{n^p}{1/\mathbb{P}[\mathcal{D}_i > n]} \\ &< \lim_{n\to\infty} \frac{n^p}{(1/\bar{\pi})^{1+\lfloor k(n)\rfloor}} = 0. \end{aligned} \tag{3.6.7}$$

For any $r \geq 1$, the r-th moment of $|\mathcal{D}_i|$ satisfies

$$\begin{aligned} \mathbb{E}[|\mathcal{D}_i|^r] &= \int_0^\infty \mathbb{P}[|\mathcal{D}_i|^r > t]dt \\ &< 1 + \int_1^\infty \mathbb{P}[\mathcal{D}_i > y] r y^{r-1} dy \\ &= 1 + \sum_{n=1}^\infty \int_n^{n+1} \mathbb{P}[\mathcal{D}_i > y] r y^{r-1} dy \\ &< 1 + \sum_{n=1}^\infty \mathbb{P}[\mathcal{D}_i > n] r (n+1)^{r-1}, \end{aligned}$$

where the second inequality follows from a change of variable $y = t^{1/r}$. Since (3.6.7) implies that $\mathbb{P}[\mathcal{D}_i > n] < Cn^{-p}$ for some nonnegative constant C, it follows that for any $p > r$,

$$\mathbb{E}[|\mathcal{D}_i|^r] < 1 + rC \sum_{n=1}^{\infty} \frac{(n+1)^{r-1}}{n^p}$$
$$< 1 + r2^{r-1}C \sum_{n=1}^{\infty} \frac{1}{n^{p-r+1}} < \infty,$$

which holds for all i. Hence, for every $r \in \{1, 2, \ldots\}$, there exists a constant $c_r = 1 + r2^{r-1}C \sum_{n=1}^{\infty} \frac{1}{n^{p-r+1}}$, independently of i, that uniformly bounds $\mathbb{E}[|\mathcal{D}_i|^r]$. □

Given Lemma 14 and Proposition 9, we are ready to prove our main theorem.

Proof of Theorem 4. Fix any arbitrary $\alpha \in (1/2, 1)$ and $\varepsilon \in (0, \alpha(1-\alpha))$. Since the process (θ_t) follows a two-state Markov chain with a symmetric transition probability ε, $(\mathbb{1}(\theta_1 \neq \theta_2), \mathbb{1}(\theta_2 \neq \theta_3), \ldots)$ is a sequence of i.i.d. random variables. By the strong law of large numbers,

$$\lim_{n\to\infty} Q_\theta(n) := \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mathbb{1}(\theta_t \neq \theta_{t+1}) = \mathbb{P}[\theta_t \neq \theta_{t+1}] = \varepsilon \quad \text{almost surely.}$$

Let $\Phi = (\mathcal{F}_1, \mathcal{F}_2, \ldots)$ be the filtration where each $\mathcal{F}_i = \sigma(\mathcal{D}_1, \ldots, \mathcal{D}_i)$ and thus $\mathcal{F}_j \subseteq \mathcal{F}_i$ for any $j \leq i$. Hence, the process $(\mathcal{D}_1, \mathcal{D}_2, \ldots)$ is adapted to Φ since each $\mathcal{D}_i$ is $\mathcal{F}_i$-measurable. By Proposition 9 and (3.6.6), there exists $\delta = 1/\varepsilon - M(\alpha, \varepsilon) > 0$ such that for all $i \geq 2$,

$$\mathbb{E}[\mathcal{D}_i | \ell_{\mathcal{T}_{i-1}}] < 1/\varepsilon - \delta.$$

By the law of iterated expectation and the Markov property of the public LLR,

$$\mathbb{E}[\mathcal{D}_i | \mathcal{F}_{i-1}] = \mathbb{E}[\mathbb{E}[\mathcal{D}_i | \ell_{\mathcal{T}_{i-1}}, \mathcal{F}_{i-1}] | \mathcal{F}_{i-1}] < 1/\varepsilon - \delta. \tag{3.6.8}$$

Let $X_i = \mathcal{D}_i - \mathbb{E}[\mathcal{D}_i | \mathcal{F}_{i-1}]$ for all $i \geq 2$ and since $\mathcal{F}_{i-1} \subseteq \mathcal{F}_i$, each X_i is $\mathcal{F}_i$-measurable. Denote a partial sum of the process $(X_i)_{i\geq 2}$ by

$$Y_n = X_2 + \frac{1}{2} X_3 + \cdots + \frac{1}{n-1} X_n.$$

By definition, $\mathbb{E}[X_i | \mathcal{F}_{i-1}] = 0$ for all $i \geq 2$. Since each Y_{n-1} is $\mathcal{F}_{n-1}$-measurable,

$$\mathbb{E}[Y_n | \mathcal{F}_{n-1}] = \mathbb{E}[\sum_{i=2}^{n} \frac{1}{i-1} X_i | \mathcal{F}_{n-1}] = Y_{n-1} + \frac{1}{n-1} \mathbb{E}[X_n | \mathcal{F}_{n-1}] = Y_{n-1},$$

and so the process $(Y_n)_n$ forms a martingale.

By Lemma 14 and (3.6.8), both $\mathbb{E}[\mathcal{D}_i^2]$ and $\mathbb{E}[\mathcal{D}_i|\mathcal{F}_{i-1}]$ are uniformly bounded. Therefore, $\mathbb{E}[X_i^2]$ is also uniformly bounded for all $i \geq 2$. Furthermore, since $\mathbb{E}[X_i X_j] = 0$ for any $i \neq j$, it then follows that for all $n \geq 2$,

$$\mathbb{E}[Y_n^2] = \sum_{i=2}^{n} \frac{1}{(i-1)^2} \mathbb{E}[X_i^2] < \infty.$$

By the martingale convergence theorem, Y_n converges almost surely. It then follows from Kronecker's lemma that[16]

$$\lim_{n\to\infty} \frac{1}{n-1}(X_2 + \cdots X_n) = 0 \quad \text{almost surely.}$$

Thus, by the definition of X_i, we can write

$$\lim_{n\to\infty} \frac{1}{n-1} \sum_{i=2}^{n} \mathcal{D}_i = \lim_{n\to\infty} \frac{1}{n-1} \sum_{i=2}^{n} \mathbb{E}[\mathcal{D}_i|\mathcal{F}_{i-1}] \quad \text{almost surely.}$$

It follows from (3.6.8) that

$$\lim_{n\to\infty} \frac{1}{n-1} \sum_{i=2}^{n} \mathcal{D}_i \leq 1/\varepsilon - \delta < 1/\varepsilon \quad \text{almost surely.} \tag{3.6.9}$$

Since $a_t = \text{sign}(\ell_{t+1})$ for all $t \geq 2$,

$$\begin{aligned} \lim_{n\to\infty} Q_a(n) &= \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mathbb{1}(a_t \neq a_{t+1}) \\ &= \lim_{n\to\infty} \frac{1}{n-1} \sum_{t=2}^{n} \mathbb{1}(\text{sign}(\ell_{t+1}) \neq \text{sign}(\ell_{t+2})). \end{aligned}$$

By the definitions of $\mathcal{T}_i$ and $\mathcal{D}_i$,

$$\frac{1}{n-1} \sum_{t=2}^{n} \mathbb{1}(\text{sign}(\ell_{t+1}) \neq \text{sign}(\ell_{t+2})) = \frac{n-1}{\mathcal{T}_n - \mathcal{T}_1} = \frac{n-1}{\sum_{i=2}^{n} \mathcal{D}_i}.$$

Hence, we conclude from (3.6.9) that

$$\lim_{n\to\infty} Q_a(n) = \lim_{n\to\infty} \frac{n-1}{\sum_{i=2}^{n} \mathcal{D}_i} > \varepsilon \quad \text{almost surely.}$$

□

[16] This result is also known as the strong law for martingales (See p.238, Feller [1966, Theorem 2]).

Bibliography

Daron Acemoglu, Angelia Nedic, and Asuman Ozdaglar. Convergence of rule-of-thumb learning rules in social networks. In *2008 47th IEEE Conference on Decision and Control*, pages 1714–1720. IEEE, 2008.

Daron Acemoglu, Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar. Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236, 2011.

Reena Aggarwal and Pietra Rivoli. Fads in the initial public offering market? *Financial management*, pages 45–57, 1990.

Itai Arieli and Manuel Mueller-Frank. A general analysis of sequential social learning. *Mathematics of Operations Research*, 46(4):1235–1249, 2021.

Itai Arieli, Yakov Babichenko, Stephan Müller, Farzad Pourbabaee, and Omer Tamuz. The hazards and benefits of condescension in social learning. *arXiv preprint arXiv:2301.11237*, 2023.

Cuimin Ba. Robust misspecified models and paradigm shift. Technical report, working paper, 2022.

Venkatesh Bala and Sanjeev Goyal. Learning from neighbours. *The Review of Economic Studies*, 65(3):595–621, 1998.

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

S. Bikhchandani, D. Hirshleifer, and I. Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, pages 992–1026, 1992.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives*, 12(3):151–170, 1998.

Sushil Bikhchandani, David Hirshleifer, Omer Tamuz, and Ivo Welch. Information cascades and social learning. Technical report, National Bureau of Economic Research, 2021.

J Aislinn Bohren. Informational herding with model misspecification. *Journal of Economic Theory*, 163:222–247, 2016.

J Aislinn Bohren and Daniel N Hauser. Learning with heterogeneous misspecified models: Characterization and robustness. *Econometrica*, 89(6):3025–3077, 2021.

Patrick Bolton and Christopher Harris. Strategic experimentation. *Econometrica*, 67(2):349–374, 1999.

Colin Camerer. Bubbles and fads in asset prices. *Journal of Economic Surveys*, 3 (1):3–41, 1989.

Boğaçhan Çelen and Shachar Kariv. Observational learning under imperfect information. *Games and Economic behavior*, 47(1):72–86, 2004.

Christophe Chamley. *Rational herds: Economic models of social learning*. Cambridge University Press, 2004.

Jaden Yang Chen. Sequential learning under informational ambiguity. *Available at SSRN 3480231*, 2022.

Thomas M Cover and Joy A. Thomas. *Elements of information theory*. John Wiley & Sons, 2006.

Krishna Dasaratha and Kevin He. Aggregative efficiency of bayesian learning in networks. *arXiv preprint arXiv:1911.10116*, 2019.

Krishna Dasaratha, Benjamin Golub, and Nir Hak. Learning from neighbors about a changing state. *Available at SSRN 3097505*, 2020.

Morris H DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

Amir Dembo and Ofer Zeituni. *Large deviations techniques and applications*. Springer, second edition, 2009.

Eugene F Fama. The behavior of stock-market prices. *The journal of Business*, 38 (1):34–105, 1965.

Willliam Feller. *An introduction to probability theory and its applications, vol 2*. John Wiley & Sons, 1966.

Mira Frick, Ryota Iijima, and Yuhta Ishii. Misinterpreting others and the fragility of social learning. *Econometrica*, 88(6):2281–2328, 2020.

Mira Frick, Ryota Iijima, and Yuhta Ishii. Learning efficiency of multi-agent information structures. 2022.

Mira Frick, Ryota Iijima, and Yuhta Ishii. Belief convergence under misspecified learning: A martingale approach. *The Review of Economic Studies*, 90(2):781–814, 2023.

Rafael M Frongillo, Grant Schoenebeck, and Omer Tamuz. Social learning in a changing world. In *International Workshop on Internet and Network Economics*, pages 146–157. Springer, 2011.

Douglas Gale and Shachar Kariv. Bayesian learning in social networks. *Games and economic behavior*, 45(2):329–346, 2003.

Benjamin Golub and Matthew O Jackson. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–49, 2010.

Benjamin Golub and Evan Sadler. Learning in social networks. *Available at SSRN 2919146*, 2017.

Alasdair J Graham and David A Pike. A note on thresholds and connectivity in random directed graphs. *Atl. Electron. J. Math*, 3(1):1–5, 2008.

Sanford J Grossman and Joseph E Stiglitz. On the impossibility of informationally efficient markets. *The American economic review*, 70(3):393–408, 1980.

Wade Hann-Caruthers, Vadim V Martynov, and Omer Tamuz. The speed of sequential asymptotic learning. *Journal of Economic Theory*, 173:383–409, 2018.

Matan Harel, Elchanan Mossel, Philipp Strack, and Omer Tamuz. Rational groupthink. *The Quarterly Journal of Economics*, 136(1):621–668, 2021.

Paul Heidhues, Sven Rady, and Philipp Strack. Strategic experimentation with private payoffs. *Journal of Economic Theory*, 159:531–551, 2015.

Helios Herrera and Johannes Hörner. A necessary and sufficient condition for information cascades. *Columbia and Yale mimeo*, 2012.

David Hirshleifer and Ivo Welch. An economic approach to the psychology of change: Amnesia, inertia, and impulsiveness. *Journal of Economics & Management Strategy*, 11(3):379–421, 2002.

Yashodhan Kanoria and Omer Tamuz. Tractable bayesian social learning on trees. *IEEE Journal on Selected Areas in Communications*, 31(4):756–765, 2013.

Navin Kartik, S Lee, Tianhao Liu, and Daniel Rappoport. Beyond unbounded beliefs: How preferences and information interplay in social learning. Technical report, 2022.

Godfrey Keller and Sven Rady. Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010.

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

Thomas S Kuhn. *The structure of scientific revolutions*. University of Chicago press, 1962.

Erich Leo Lehmann. Comparing location experiments. *The Annals of Statistics*, pages 521–533, 1988.

Jure Leskovec and Eric Horvitz. Planetary-scale views on a large instant-messaging network. In *Proceedings of the 17th international conference on World Wide Web*, pages 915–924, 2008.

Raphaël Lévy, Marcin Pęski, and Nicolas Vieille. Stationary social learning in a changing environment. *arXiv preprint arXiv:2201.02122*, 2022.

Annie Liang and Xiaosheng Mu. Complementary information and learning traps. *The Quarterly Journal of Economics*, 135(1):389–448, 2020.

Ilan Lobel and Evan Sadler. Information diffusion in networks through social learning. *Theoretical Economics*, 10(3):807–851, 2015.

Dimitri Migrow. Strategic observational learning. *arXiv preprint arXiv:2212.09889*, 2022.

Pooya Molavi, Alireza Tahbaz-Salehi, and Ali Jadbabaie. A theory of non-bayesian social learning. *Econometrica*, 86(2):445–490, 2018.

Giuseppe Moscarini and Lones Smith. The law of large demand for information. *Econometrica*, 70(6):2351–2366, 2002.

Giuseppe Moscarini, Marco Ottaviani, and Lones Smith. Social learning in a changing world. *Economic Theory*, 11(3):657–665, 1998.

Elchanan Mossel, Allan Sly, and Omer Tamuz. Asymptotic learning on bayesian social networks. *Probability Theory and Related Fields*, 158(1-2):127–157, 2014.

Elchanan Mossel, Allan Sly, and Omer Tamuz. Strategic learning and the topology of social networks. *Econometrica*, 83(5):1755–1794, 2015.

Abraham Ortelius. *Thesaurus Geographicus: A New Body of Geography: Or, A Compleat Description of the Earth*. London: Printed for Abel Swall and Tim. Child, 1695. URL `https://books.google.com/books?id=RVFdNM38mqYC`.

Emanuel Parzen. Nonparametric statistical data modeling. *Journal of the American statistical association*, 74(365):105–121, 1979.

Javier Rojo. A pure-tail ordering based on the ratio of the quantile functions. *The Annals of Statistics*, pages 570–579, 1992.

Dinah Rosenberg and Nicolas Vieille. On the efficiency of social learning. *Econometrica*, 87(6):2141–2168, 2019.

Dinah Rosenberg, Eilon Solan, and Nicolas Vieille. Informational externalities and emergence of consensus. *Games and Economic Behavior*, 66(2):979–994, 2009.

Paul Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(6):41–49, 1965.

Edouard Schaal and Mathieu Taschereau-Dumouchel. Herding through booms and busts. *CEPR Discussion Paper No. DP16368*, 2021.

Shahin Shahrampour, Sasha Rakhlin, and Ali Jadbabaie. Online learning of dynamic parameters in social networks. *Advances in Neural Information Processing Systems*, 26, 2013.

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

Lones Smith, Peter Norman Sørensen, and Jianrong Tian. Informational herding, optimal experimentation, and contrarianism. *The Review of Economic Studies*, 88(5):2527–2554, 2021.

Peter Norman Sørensen. *Rational social learning*. PhD thesis, Massachusetts Institute of Technology, 1996.

Xavier Vives. How fast do rational agents learn? *The Review of Economic Studies*, 60(2):329–347, 1993.

Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.