LEARNING ABOUT INFORMATIVENESS

WANYING HUANG[†]

—Click here to view the most recent draft—

ABSTRACT. We study whether individuals can learn the informativeness of their information technology through social learning. As in the classic sequential social learning model, rational agents arrive in order and make decisions based on the past actions of others and their private signals. There is uncertainty regarding the informativeness of the common signal-generating process. We show that learning in this setting is not guaranteed, and depends crucially on the relative tail distributions of private beliefs induced by uninformative and informative signals. We identify the phenomenon of perpetual disagreement as the cause of learning and provide a characterization of learning in the canonical Gaussian environment.

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 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]

As a leading example, consider an external evaluator tasked with assessing the informativeness of a novel information source, such as an AI recommendation system. The evaluator is rewarded if his assessment is correct. Meanwhile, this AI recommendation system has been employed by a sequence of investors who are interested in buying into a new start-up company of unknown quality. Each investor not only observes the decisions

Date: November 30, 2023.

[†]California Institute of Technology. Email: whhuang@caltech.edu.

I am indebted to my advisor Omer Tamuz for his continued support and encouragement. I thank (in alphabetical order) Krishna Dasaratha, Laura Doval, Federico Echenique, Wade Hann-Caruthers, Kirby Nielson, Fan Wu, and seminar audiences for their helpful comments and suggestions.

[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.

made by prior investors but also receives a private recommendation from this information source. While this source could be informative, offering valuable insights into the quality of the start-up, there is some positive probability that it is completely uninformative, providing no useful information at all. However, the external evaluator only has access to the investors' buying or selling decisions, but not to the private recommendations they received from the system. Can the evaluator eventually learn the informativeness of the recommendation system and thus make the correct assessment? What information structures would ensure learning in this sense and thus enable society to eventually adopt informative sources while discarding uninformative ones?[2]

Formally, we incorporate the uncertainty regarding the informativeness of the source into the classic sequential social learning model (Banerjee, 1992; Bikhchandani et al., 1992). A sequence of short-lived agents arrive in order, each acting once by choosing an action to match with a fixed, unknown, and payoff-relevant state that can be either good or bad. Before choosing an action, each agent observes the past actions of her predecessors and receives a private signal from a common source of information. 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. The informativeness of the source is realized at the beginning of time.

If an outside observer, who aims to evaluate the informativeness of the source, were to have access to the private signals received by the agents, he would gradually accumulate empirical evidence from this source, thus eventually learning its informativeness. However, when only the history of past actions is observable, the 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 arises from the fact that agents also base their decisions on the inferences they draw from others' actions. In this case, we say *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?

We focus on *unbounded signals* (Smith and Sørensen, 2000), where the belief induced by a private signal conditioned on an informative source can be arbitrarily strong, since otherwise learning can be easily precluded by agents' lack of responses to their private

[2]More broadly, the source of information can be viewed as a scientific paradigm—a widely accepted theoretical framework within a specific scientific discipline over an extended period. In such a context, the question of whether society can learn about the informativeness of the source has a similar flavor to Kuhn's question of whether society can make the right scientific paradigm shifts (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 proposed earlier by Abraham Ortelius in his 1596 book "Thesaurus Geographicus" (in English, "A New Body of Geography").

signals.[3] 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 private beliefs, particularly in whether the distribution of private beliefs induced by uninformative signals has *fatter* or *thinner* tails compared to that of informative signals. More specifically, we show that asymptotic learning holds when uninformative signals have fatter tails than informative signals, but it fails when uninformative signals have thinner tails.

For example, consider an informative source that generates Gaussian signals with unit variance and mean $+1$ if the payoff-relevant state is good, and -1 if the state is bad. If the uninformative source also generates Gaussian signals, but with a variance greater than one, then the uninformative signals have fatter tails, and so asymptotic learning holds by Theorem 1. In contrast, uninformative Gaussian signals with a variance less than one exhibit thinner tails, and thus, Theorem 1 implies that asymptotic learning fails. When all Gaussian signals share the same variance and the uninformative signals have a mean that lies between -1 and $+1$, they exhibit neither fatter nor thinner tails than the informative signals. To complement our main result, we demonstrate that in this canonical case, asymptotic learning occurs if and only if uninformative signals are symmetric around zero, which is equal to the ex-ante expected value of the payoff-relevant state (Theorem 2).

As another illustration of the main result, consider the case where the informative signals have the same distributions as before but the uninformative source generates signals chosen uniformly from the 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 that of the Gaussian informative signals, and thus, Theorem 1 implies that asymptotic learning fails. However, note that under such an informative source, almost all agents individually learn that it is informative. This is because once they receive a signal outside the support $[-\varepsilon, \varepsilon]$, which happens with high probability, they can infer that this signal can only be generated from the normal distribution, and thus, learn that the source must be informative. Nevertheless, the outside observer who only observes agents' actions would fail to determine the informativeness of the source.

Intuitively, the condition of having fatter tails means that it is more likely to observe a very extreme signal—say a $5-\sigma$ signal—under an uninformative source than an informative one. As a consequence, the presence of extreme signals suggests that the source is uninformative. However, since the outside observer does not directly observe the private signals received by the agents, he can only learn from observing the "irregular" behavior triggered by these extreme signals. For example, when an agent deviates from a long

[3]The phenomenon that agents follow the action of their predecessors irrespective of their private signals is known as information cascades, which occur almost surely with bounded and finite signals (Banerjee, 1992; Bikhchandani et al., 1992).

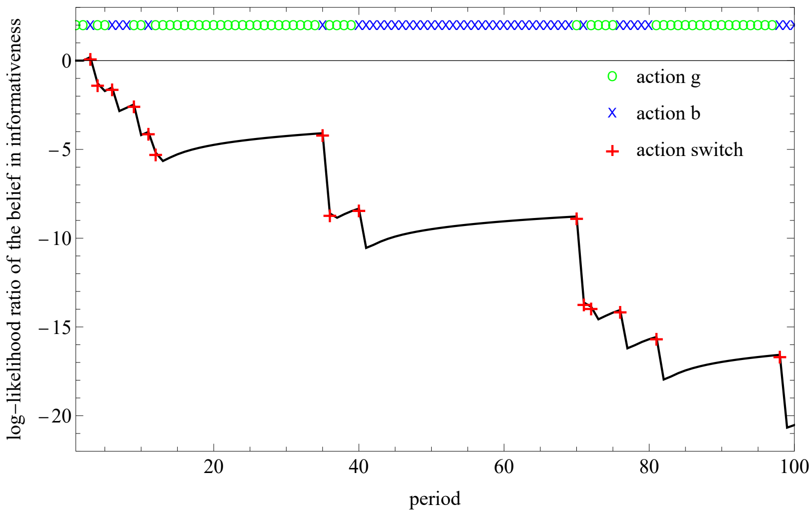

FIGURE 1. This simulation tracks an outside observer's belief in the informativeness of the source over 100 periods. It is generated under uninformative signals with a distribution $\mu_0 = \mathcal{N}(0, 2)$ and the corresponding informative signals are normally distributed with means of -1 and $+1$, both with a unit variance. Red crosses mark the periods when action switches occur.

sequence of identical actions, this conveys a substantial amount of information to the outside observer, as it allows the observer to infer that this agent must have received an extreme signal. Given that extreme signals are more likely under uninformative signals with fatter tails, such a deviation makes the observer more convinced that the source is uninformative. In contrast, when the tails are thinner, an uninformative source tends to produce more moderate signals. Thus, in this case, the presence of extreme signals (and the corresponding irregular behavior) no longer suggests an uninformative source.

We illustrate this intuition in Figure 1, which depicts one simulation of the belief process of an outside observer under uninformative signals with fatter tails. First, it is worth noting that every action switch, following an extended period of identical actions, results in a significant decrease in the outside observer's belief (in terms of the log-likelihood ratio) that the source is informative. In other words, as discussed before, observing these unusual action switches makes the observer more convinced that the source is uninformative. Second, even though the belief gradually increases in the absence of these action switches, it would eventually converge to zero, as suggested by our main theorem. In contrast, if the simulation were generated under uninformative signals with thinner tails, then extreme signals would be less likely. As a result, action switches would eventually stop, leaving the outside observer uncertain about the source's informativeness.

A key assumption underlying our main result is the uniform prior assumption regarding the payoff-relevant state. This assumption ensures that despite the presence of uncertainty, rational agents always act as if the signals they receive are informative (Lemma 1). This is simply because, in the absence of any useful information—conditioned on an uninformative source—each agent with a uniform prior is indifferent between the available actions. We make the uniform prior assumption to capture 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 (Samuelson, 1965; Fama, 1965) implies that investors should be close to indifference.[4]

The mechanism behind our main result is as follows. Since agents always act as if signals are informative, when the source is indeed informative and generates unbounded signals, agents will eventually converge to a consensus on the correct action. On the one hand, suppose that the source is uninformative and generates signals with thinner tails. In this case, agents are unlikely to receive an extreme signal, so they usually mimic their predecessors. As a result, agents will also eventually reach a consensus, thus exhibiting behavior similar to that under an informative source. It then becomes impossible for an outside observer who only observes agents' actions to be certain about the informativeness of the source. On the other hand, suppose that the source is uninformative and generates signals with fatter tails. In this case, extreme signals are more likely, preventing agents from reaching a consensus. In fact, they will perpetually change their actions. Hence, an outside observer who observes an infinite number of action switches learns that the source is uninformative.

For general private belief distributions where the relative tail thickness is incomparable, we show that the same holds: Asymptotic learning holds if and only if, conditioned on the source being uninformative, both actions are taken infinitely often (Proposition 1). In terms of behavioral implications, as mentioned before, when the source happens to be informative, all agents eventually take the correct action, regardless of whether asymptotic learning is achieved. However, in such cases, unless asymptotic learning holds, agents can never be certain that they are taking the correct action. Conversely, if agents are eventually certain that they are taking the correct action, then asymptotic learning must hold (Lemma 2). In contrast, when the source happens to be uninformative, agents are clearly not guaranteed to converge to the correct action; in fact, their actions may or may not converge at all. Proposition 1 demonstrates that an outside observer learns the informativeness of the source if and only if the agents' actions do not converge.

Related Literature. Our paper contributes to a rich literature on sequential social learning. The primary focus of this literature has been on determining whether agents can eventually learn to choose the correct payoff-relevant state when provided with an informative source. Various factors, such as the information structure (Banerjee, 1992;

[4]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 information that is not yet reflected in market prices, thereby challenging the hypothesis.

Bikhchandani, Hirshleifer, and Welch, 1992; Smith and Sørensen, 2000) and the observational networks (Çelen and Kariv, 2004; Acemoglu, Dahleh, Lobel, and Ozdaglar, 2011; Lobel and Sadler, 2015), have been extensively studied to analyze their impacts on information aggregation, including efficiency (Rosenberg and Vieille, 2019) and the speed of learning (e.g., Vives, 1993; Hann-Caruthers, Martynov, and Tamuz, 2018). However, the question of learning about the informativeness of the source that agents use to make decisions—which is the focus of this paper—remains largely unexplored.[5]

A few papers explore the idea of agents having access to multiple sources of information in the context of social learning. For example, Chen (2022) examines a sequential social learning model in which ambiguity-averse agents have access to different information sources. Instead of having uncertainty regarding the informativeness of a common source, in his model, uncertainty arises since 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. In a different setting, 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. They focus on the externality in the information acquisition decisions and show that information complementarity can result in either efficient information aggregation or "learning traps," in which society becomes stuck in choosing suboptimal information structures.

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, a recent work by Arieli and Mueller-Frank (2021) demonstrates that pairwise unbounded signals are necessary and sufficient for learning, when the decision problem that the agents face includes, for each state, a distinct action that is the only optimal action in that state. This is not the case in our model, because the same action is optimal in different states, 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 more general setting that subsumes our model as a special case. Their Theorem 2 implies that agents in our model do not always learn the payoff-relevant state, which is—in our very specific setting—easy to see, because when signals are uninformative, agents do not learn the payoff-relevant state. 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

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

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.

2. Model

2.1. **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 $t \in \mathbb{N} = \{1, 2, \ldots\}$ arrive in order, each acting once. The action of agent t is $a_t \in A = \{\mathfrak{g}, \mathfrak{b}\}$, and the utility function is $u : \Theta \times A \to \{0, 1\}$ where $u(\theta, a) = 1$ if $a = \theta$ and zero otherwise. Before agent t chooses an action, she observes the history of all past actions $H_t = \{a_1, \ldots, a_{t-1}\}$ made by her predecessors and receives a private signal s_t, taking values in a measurable space (S, Σ). Let $\mathcal{I}_t = A^{t-1} \times S$ be the set of all possible pieces of information available to agent t prior to her decision, so $I_t = (H_t, s_t)$ is an element of $\mathcal{I}_t$.

A pure strategy of agent t is a measurable function $\sigma_t : \mathcal{I}_t \to A$ that selects an action for each possible information set. 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 utility 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 to maximize her expected payoff conditional on the information available to her:

$$a_t \in \arg\max_{a\in A} \mathbb{E}[u(\theta, a)|I_t].$$

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

2.2. **The Informativeness of the Source.** Unlike the classic sequential social learning model, where all private signals are generated from an informative source, in our model, the informativeness of this source remains uncertain. Specifically, at time 0, nature independently chooses an additional binary state $\Omega \in \{0, 1\}$ with equal probability. When $\Omega = 1$, the source is informative and sends out conditionally i.i.d. signals, with distribution μ_Θ. When $\Omega = 0$, the source is uninformative and sends out i.i.d. signals, with distribution μ_0. The realization of Ω determines the signal-generating process for all agents. Throughout, we denote by $\mathbb{P}_\omega[\cdot] := \mathbb{P}[\cdot|\Omega = \omega]$ and $\mathbb{E}_\omega := \mathbb{E}[\cdot|\Omega = \omega]$ the probability and expectation conditioned on the realization of Ω. Similarly, we denote by $\mathbb{P}_{(\omega,\theta)}[\cdot] := \mathbb{P}[\cdot|\Omega = \omega, \Theta = \theta]$ and $\mathbb{E}_{(\omega,\theta)} := \mathbb{E}[\cdot|\Omega = \omega, \Theta = \theta]$ the probability and expectation conditioned on the pair of state realizations of Ω and Θ.

We first observe that, despite the uncertainty regarding the source's informativeness, the equilibrium action for each agent is to choose the action that is most likely to match the state, *conditioned* on the source being informative.

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

$$a_t \in \arg\max_{a \in A} \mathbb{P}_1[\Theta = a | I_t]. \tag{1}$$

In other words, agents always act as if signals are informative, irrespective of the underlying signal-generating process. This is 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.

2.3. **Information Structure.** The distributions μ_0, $\mu_\mathfrak{g}$, and $\mu_\mathfrak{b}$ 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 s_t

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

is well-defined. Since agents always act as if the source is informative (Lemma 1), it captures how agents update their beliefs regarding the relative likelihood of the good state over the bad state upon receiving their private signals. We denote the cumulative distribution function (CDF) of ℓ_t conditioned on $\Omega = 0$ by F_0, and denote its CDF conditioned on $\Omega = 1$ and $\Theta = \theta$, for any $\theta \in \{\mathfrak{g}, \mathfrak{b}\}$ by F_θ. All distributions F_0, $F_\mathfrak{g}$ and $F_\mathfrak{b}$ are also mutually absolutely continuous as $\mu_0, \mu_\mathfrak{g}$ and $\mu_\mathfrak{b}$ are. Let f_0, $f_\mathfrak{g}$ and $f_\mathfrak{b}$ denote the corresponding density functions of F_0, $F_\mathfrak{g}$ and $F_\mathfrak{b}$ whenever they are differentiable.

We consider *unbounded* private signals in the sense that ℓ_t 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 ℓ_t that 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 state realizations (ω, θ). An extreme signal for the Gaussian case is a signal that is, for example, 5-σ away from the mean $m_{(\omega,\theta)}$.

We make two assumptions for expository simplicity. First, we assume that the pair $(F_\mathfrak{g}, F_\mathfrak{b})$ is symmetric around zero, i.e., $F_\mathfrak{g}(x) + F_\mathfrak{b}(-x) = 1$. This implies that the model is invariant with respect to a relabeling of the payoff-relevant state. Second, we assume that $F_\mathfrak{g}$ and $F_\mathfrak{b}$ are continuous so that 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 density function $f_{\mathfrak{g}}$ that 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.

2.4. **Asymptotic Learning.** Denote by $p_t^\Omega := \mathbb{P}[\Omega = 1|H_t]$ the *public belief* at time t that the source is informative given the history of actions. This reflects an outside observer's belief about the source being informative based on agents' actions. As this observer collects more information over time, his belief p_t^Ω converges almost surely, as it is a bounded martingale. To capture the idea that the outside observer 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 $\omega \in \{0, 1\}$,

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

That is, conditioned on an informative source, the public belief that the source is informative converges to one almost surely. Conversely, conditioned on an uninformative source, the public belief that the source is informative converges to zero almost surely. As we explain below in Section 4.1, asymptotic learning is always attainable, regardless of the underlying information structure, if all private signals are publicly observable.

3. Relative Tail Thickness

To study the conditions for asymptotic learning, it is crucial to understand the concept of relative tail thickness, i.e., the comparison between the tails of the distributions of the private beliefs induced by different signals. This is important as it captures the relative likelihood of generating extreme signals under different information sources. Formally, for a fixed pair of cumulative distribution functions (F_0, F_θ) where $\theta \in \{\mathfrak{g}, \mathfrak{b}\}$ and some $x \in \mathbb{R}_+$, we denote by

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

the left and right tail ratio of F_0 over F_θ evaluated at x, respectively. To assess whether extreme signals are more or less likely to occur under an uninformative source compared to an informative one, we introduce the following concepts of fatter and thinner tails.

Definition 2 (Fatter and Thinner-Tailed Uninformative Signals)**.**

(i) The uninformative signals have *fatter tails* than the informative signals if there exists an $\varepsilon > 0$ such that $L_{\mathfrak{b}}(x) \geq \varepsilon$ for all x large enough and $R_{\mathfrak{g}}(x) \geq \varepsilon$ for all x large enough.

(ii) The uninformative signals have *thinner tails* than the informative signals if there exists an $\varepsilon > 0$ such that either $L_{\mathfrak{g}}(x) \leq 1/\varepsilon$ for all x large enough or $R_{\mathfrak{b}}(x) \leq 1/\varepsilon$ for all x large enough.

That is, for uninformative signals to have fatter tails, it requires both their left tail ratio with respect to the bad state and right tail ratio with respect to the good state to be eventually bounded from below. Conversely, if either the left tail ratio with respect to the good state or the right tail ratio with respect to the bad state is eventually bounded from above, then uninformative signals have thinner tails.[6]

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., Smith and Sørensen, 2000; Chamley, 2004; Rosenberg and Vieille, 2019). Similar statements apply to the remaining three conditions. In addition, notice that uninformative signals cannot have both fatter and thinner tails simultaneously, as $F_{\mathfrak{g}}$ and $F_{\mathfrak{b}}$ represent distributions of the log-likelihood ratio.[7] Finally, for differentiable distributions, one can equivalently define the tail ratios in terms of the corresponding density ratios, if they exist.

Intuitively, uninformative signals with fatter tails have a higher probability of being extreme compared to informative signals. By Bayes' Theorem, this means that under fatter-tailed uninformative signals, when an agent observes an extreme signal, it suggests that the source is uninformative. Conversely, under thinner-tailed uninformative signals, an extreme signal becomes less suggestive of the source being uninformative. Next, we provide three examples of uninformative signals with fatter or thinner tails.

Example 1 (Gaussian Signals)**.** Consider the case that $F_{\mathfrak{g}}$ is normal with mean $+1$ and unit standard deviation and that $F_{\mathfrak{b}}$ is 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 (positive or negative) signal, such as anything above 11, is observed, it is much more likely that the source is uninformative than that an informative 10-σ signal was 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 becomes an indication that the source is informative. A graphical example of Gaussian signals with fatter-tailed and thinner-tailed uninformative signals is depicted in Figure 2.

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

[7] 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 $x \in \mathbb{R}$.

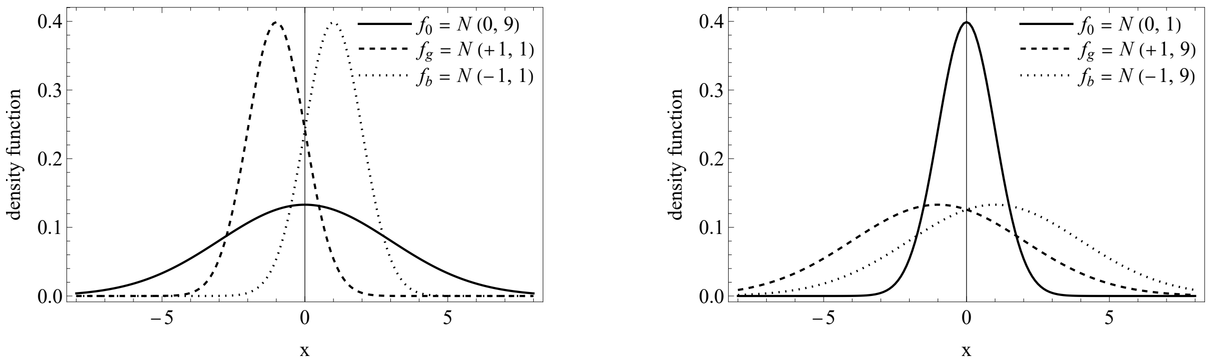

FIGURE 2. An example of Gaussian signals with fatter-tailed uninformative signals (on the left) and an example of Gaussian signals with thinner-tailed uninformative signals (on the right).

Example 2 (First-Order Stochastically Dominated (FOSD) Signals)**.** Suppose that either F_0 first-order stochastically dominates $F_{\mathfrak{g}}$ or that $F_{\mathfrak{b}}$ first-order stochastically dominates F_0. In either case, the uninformative signals have thinner tails than informative signals due to their heavy bias toward one of the states.

Suppose it is the former case and an extreme signal of a large positive value is observed. Then, it is unlikely that the source is informative and associated with the bad state. Instead, the source is likely to be either uninformative or informative but associated with the good state. Therefore, observing an extreme positive signal indicates that the state is good but does not indicate the informativeness of the source as such an extreme signal is likely under both F_0 and $F_{\mathfrak{g}}$.

Example 3 (Mixture Signals)**.** Fix any pair of distributions $(F_{\mathfrak{g}}, F_{\mathfrak{b}})$ and suppose that $F_0 = \alpha F_{\mathfrak{g}} + (1-\alpha)F_{\mathfrak{b}}$ for any $\alpha \in (0,1)$. In other words, to draw signals from the uninformative source, one would draw a signal from $F_{\mathfrak{g}}$ with probability α and draw a signal from $F_{\mathfrak{b}}$ with probability $1-\alpha$. Observe that the uninformative signals have fatter tails than the informative signals.[8] In particular, when $\alpha = 1/2$, the corresponding mixture distribution coincides with the unconditional distribution of the private signals given an informative source. This suggests that, in this case, an uninformative source is a priori indistinguishable from an informative one.

Alternatively, the mixture distribution F_0 can be viewed as generating conditionally i.i.d. signals, but these signals are now conditioned on a different state, say $\tilde{\Theta}$, rather than the true payoff-relevant state Θ: In each period, the state $\tilde{\Theta}$ is randomly drawn from the set $\{\mathfrak{g}, \mathfrak{b}\}$, independently of the state Θ. If the event $\tilde{\Theta} = \mathfrak{g}$ occurs, which happens with probability α, then the source draws a signal from $F_{\mathfrak{g}}$, otherwise from $F_{\mathfrak{b}}$.

[8] To see this, consider any fixed pair of distributions $(F_{\mathfrak{g}}, F_{\mathfrak{b}})$ and some constant $\alpha \in (0,1)$. Let $F_0 = \alpha F_{\mathfrak{g}} + (1-\alpha)F_{\mathfrak{b}}$. Since CDFs always take nonnegative values, for any $x \in \mathbb{R}$, $F_0(x) \geq (1-\alpha)F_{\mathfrak{b}}(x)$ and 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 hence, by definition, F_0 generates uninformative signals with fatter tails.

4. Results

4.1. **The Public Signals Benchmark.** As a benchmark, we briefly discuss the case of public signals in which an outside observer observes all private signals received by the agents. Equivalently, one can think of this scenario as a case of a single agent who receives i.i.d. signals in every period. These signals either come from μ_0, $\mu_{\mathfrak{g}}$ or $\mu_{\mathfrak{b}}$, and since these measures are distinct, this observer will eventually learn which distribution is being sampled. Formally, at any time t, an outside observer can calculate the empirical distribution of the signals, which assigns to a measurable set $B \subseteq S$

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

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

This benchmark result shows that, regardless of the underlying signal-generating process, any uncertainty concerning the informativeness of the source will eventually be resolved if all signals are publicly observable. However, if signals are private and only actions are publicly observable, an outside observer is no longer able to determine the exact signals received by the agents, preventing him from learning about the informativeness of the source.

4.2. **Public Actions.** We now turn to the setting where only past actions are observable but not the private signals. Our main result shows that in contrast to the public-signals benchmark, the achievement of asymptotic learning crucially depends on the relative tail thickness between the uninformative and informative signals (see 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.*

For normal distributions, the relative thickness of the tail is determined solely by the variances: a larger variance corresponds to fatter tails, while a smaller variance corresponds to thinner tails (by the lemma 9 in the appendix). An immediate consequence of Theorem 1 is the following result.

Corollary 1. *Suppose all private signals are Gaussian 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 $\tau < \sigma$.*

Theorem 1 demonstrates that an outside observer can eventually learn to identify an uninformative source if it generates widely dispersed extreme noise, as opposed to an informative source. In contrast, when the uninformative signals are relatively concentrated and have thinner tails, this observer is unable to recognize their lack of informativeness.

To understand why relative tail thickness is crucial for achieving asymptotic learning, let us first consider the case where the source is informative. In this case, the probability of producing extreme signals that could overturn a long sequence of identical actions decreases rapidly. Since agents always treat signals as if they are informative (Lemma 1), they would eventually converge to a consensus on the correct action. Hence, for an outsider observer to effectively distinguish an uninformative source from an informative one, there need to be noticeable differences in agents' behavior under different sources. We refer to the event in which both actions are taken infinitely many times as "perpetual disagreement". We show that, in fact, achieving asymptotic learning is equivalent to having perpetual disagreement among agents, provided that the source is uninformative (Proposition 1).

Building on the connection between perpetual disagreement and asymptotic learning, the proof idea behind Theorem 1 is as follows. Suppose that the source is uninformative and generates signals with fatter tails. This implies that the probability that later agents receive an extreme signal that would overturn a long sequence of identical actions diminishes relatively slowly. As a result, agents would never settle on a particular action. Hence, in the presence of fatter-tailed uninformative signals, perpetual disagreement will persist, leading to asymptotic learning. Conversely, if the uninformative signals have thinner tails, the probability of agents receiving these extreme, overturning signals declines rapidly enough so that agents will eventually settle on an action. As a consequence, perpetual disagreement is not guaranteed under thinner-tailed uninformative signals, resulting in the failure of asymptotic learning.

4.3. **Gaussian Private Signals.** While Theorem 1 provides valuable insights into the role of relative tail thickness in determining asymptotic learning, there are situations where the uninformative signals have neither fatter nor thinner tails compared to the 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 the left tail ratio $L_{\mathfrak{b}}(x)$ converging to zero. As a result, F_0 does not have fatter tails. Similarly, as x approaches infinity, both the left tail ratio $L_{\mathfrak{g}}(x)$ and the right tail ratio $R_{\mathfrak{b}}(x)$ tend to infinity, implying that F_0 does not have thinner tails either.

To complement the findings of Theorem 1, we focus on the canonical Gaussian environment where all signals are normal and have the same variance σ^2. Suppose, without loss of generality, that the informative signals are symmetric with mean $+1$ and -1, respectively, while the uninformative signals have mean $\mu \in (-1, 1)$. In this setting, a simple calculation shows that the log-likelihood ratio of the private belief, conditioned on informative signals, is directly proportional to the private signal: $\ell_t = 2s_t/\sigma^2$. As a consequence, F_0, $F_\mathfrak{g}$, and $F_\mathfrak{b}$ are also normal distributions. The following result shows that asymptotic learning is only achieved if the uninformative signals have mean zero, which is equal to the ex-ante expected value of the payoff-relevant state.

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 $\mu \in (-1, 1)$. Asymptotic learning holds if and only if $\mu = 0$.*

Together with Corollary 1, Theorem 2 provides a complete characterization of asymptotic learning in Gaussian environments with symmetric informative signals. Interestingly, Theorem 2 shows that even at the threshold ($\tau = \sigma$) where the Gaussian uninformative signals have neither fatter nor thinner tails, asymptotic learning still holds as long as the distribution F_0 is symmetric around zero. However, asymptotic learning fails if F_0 exhibits any bias toward a specific state. Intuitively, any shifts in the mean make the uninformative signals slightly more similar to the corresponding informative signals, thus making it more challenging to differentiate between them.

In an extreme case where F_0 shifts entirely to the right of $F_\mathfrak{g}$ (or to the left of $F_\mathfrak{b}$), we have already seen that asymptotic learning fails in Example 2. This type of learning failure corresponds to the non-falsifiability of the hypothesis that the signals are uninformative when they happen to be biased toward the true payoff-relevant state. We discuss the connection between asymptotic learning and hypothesis testing in Section 6.

4.4. **Numerical Simulation.** To further illustrate our main result, we use Monte Carlo simulations to numerically simulate both the belief process and the resulting action switches. For a fixed pair of distributions of informative signals, $\mu_\mathfrak{g} = \mathcal{N}(+1, 2)$ and $\mu_\mathfrak{b} = \mathcal{N}(-1, 2)$, we simulate these processes under uninformative signals with fatter tails with distribution $\mu_0 = \mathcal{N}(0, 3)$ and with thinner tails with distribution $\mu_0 = \mathcal{N}(0, 1)$. Note that in this case, the log-likelihood ratio $\ell_t = s_t$, so that $F_0, F_\mathfrak{g}$ and $F_\mathfrak{b}$ have the same distribution as $\mu_0, \mu_\mathfrak{g}$ and $\mu_\mathfrak{b}$, respectively. We conduct these simulations 1,000 times and calculate the average for each period. This yields the expected belief and the expected total number of action switches, conditioned on the source being uninformative. Figure 3 displays the results of these simulations.

What immediately stands out is that under fatter-tailed uninformative signals, the outside observer's belief that the source is informative decreases much faster compared to thinner-tailed uninformative signals. By period 60, his belief that the source is informative is approximately 0.3 under fatter-tailed uninformative signals, which is less than two-thirds of the belief observed with thinner-tailed uninformative signals. This finding aligns with the predictions of Theorem 1, which suggests that in the presence of fatter-tailed

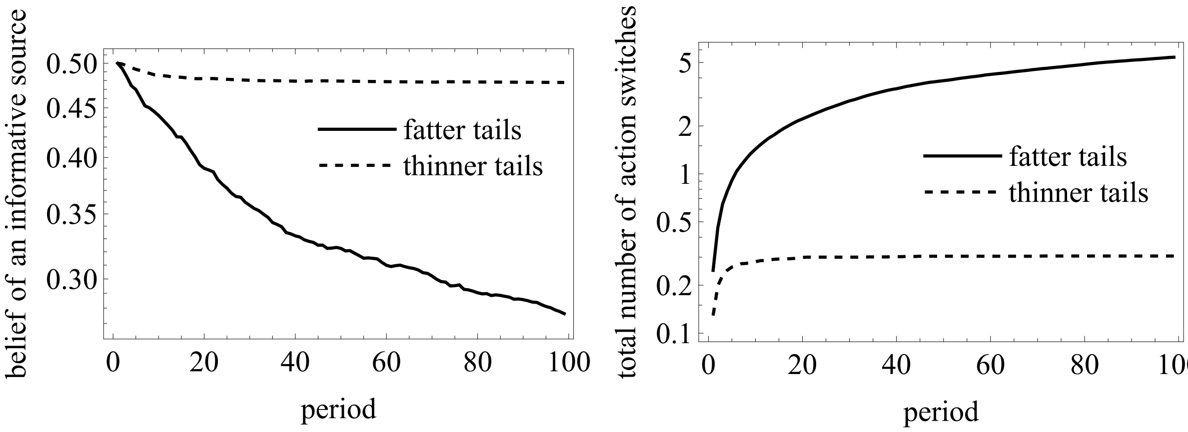

FIGURE 3. The belief that the source is informative (on the left) and the total number of action switches (on the right) under uninformative signals with fatter and thinner tails.

uninformative signals, agents would eventually learn that the source is uninformative. Consequently, the public belief in the informativeness of the source converges to zero.

As mentioned before, the main mechanism driving our main result is the persistence of perpetual disagreement under an uninformative source. The intuition is that the uninformative source with fatter tails has a higher probability of generating extreme signals, which in turn prevents convergence to a consensus. This is evident in the right plot of Figure 3, where the average total number of action switches under fatter-tailed uninformative signals increases over time. In contrast, under thinner-tailed uninformative signals, the total number of switches plateaus in a short amount of time, indicating that perpetual disagreement does not occur in this case.

4.5. **Asymptotic Learning and Other Learning Concepts.** We end this section by discussing the connections between asymptotic learning and other learning concepts in the literature. When the source is always informative, one common learning concept, concerning the convergence of actions, is called *correct herding*, i.e., $\lim_{t\to\infty} a_t = \Theta$ almost surely. Another notion of learning, related to the convergence of beliefs, is called *complete learning*. To formally define this notion, we denote by $p_t^\Theta := \mathbb{P}[\Theta = \mathfrak{g}|H_t]$ the *public belief* at time t that the state is good given the history of actions.

Definition 3. We say that *complete learning* occurs if conditioned on $\Theta = \mathfrak{g}$, $\lim_{t\to\infty} p_t^\Theta = 1$ almost surely and conditioned on $\Theta = \mathfrak{b}$, $\lim_{t\to\infty} p_t^\Theta = 0$ almost surely.

That is, learning is complete if the public belief about the payoff-relevant state eventually converges to the truth. Recall that by Lemma 1, agents always act as if signals are informative, regardless of the informativeness of the source. Thus, when the source is indeed informative and generates unbounded signals, it follows from Smith and Sørensen

(2000) that correct herding still occurs, i.e.,

$$\lim_{t\to\infty} a_t = \Theta \quad \mathbb{P}_1\text{-almost surely.} \tag{2}$$

That is, in our model, when the source is informative, all but finitely many agents take the correct action, regardless of whether asymptotic learning is achieved. However, even though agents do herd on the correct action, they cannot be completely certain that it is the correct action unless asymptotic learning occurs, as shown in the following lemma.

Lemma 2. *Asymptotic learning holds if and only if conditioned on an informative source, complete learning occurs, i.e.,*

$$\lim_{t\to\infty} p_t^{\Theta} = 1 \quad \mathbb{P}_{1,\mathfrak{g}}\text{-almost surely, and } \lim_{t\to\infty} p_t^{\Theta} = 0 \quad \mathbb{P}_{1,\mathfrak{b}}\text{-almost surely.}$$

This lemma states that if the source is informative, then for the public belief to converge to the truth, agents must also learn that the source is informative and vice versa. Intuitively, if agents are unsure about the informativeness of the source, they cannot fully trust the public information. Conversely, once they learn that the source is informative, the public information becomes highly accurate so that the public belief will converge to the truth. Note that if the source is uninformative, achieving asymptotic learning is equivalent to having the public belief about the payoff-relevant state converging to the uniform prior almost surely.

5. Analysis

In this section we first examine how agents update their beliefs and present some standard yet useful properties of their belief-updating process, namely, the *overturning principle* and *stationarity*. We then discuss two important phenomena: *perpetual disagreement* and *immediate herding*, and explain how they are related to asymptotic learning. These relationships allow us to characterize asymptotic learning in terms of immediate herding, which simplifies the problem. We provide a proof sketch of our main result (Theorem 1) at the end of this section.

5.1. 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. Let p_t denote the *posterior belief* of agent t assigned to the good state after observing $I_t = (H_t, s_t)$ and conditioned on $\Omega = 1$, i.e., $p_t := \mathbb{P}_1[\Theta = \mathfrak{g}|I_t]$. The log-likelihood ratio (LLR) of agent t's posterior beliefs of the good state over the bad state is

$$L_t := \log \frac{p_t}{1-p_t} = \log \frac{\mathbb{P}_1[\Theta = \mathfrak{g}|H_t, s_t]}{\mathbb{P}_1[\Theta = \mathfrak{b}|H_t, s_t]}.$$

Similarly, we denote by $\pi_t := \mathbb{P}_1[\Theta = \mathfrak{g}|H_t]$ the *public belief* of agent t assigned to the good state conditioned on $\Omega = 1$ and only the history of actions H_t. Upon observing the history of actions, this forms the prior belief of agent t about the payoff-relevant state before receiving her private signal. The corresponding LLR of π_t is thus

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

Recall that ℓ_t is the LLR of the private beliefs induced by signal s_t conditioned on $\Omega = 1$. By Bayes' rule, we can write the posterior LLR of agent t as

$$L_t = r_t + \ell_t,$$

which is the sum of the LLRs of her public belief and her private belief. It follows from (1) that agent t chooses action $\mathfrak{g}$ when $\ell_t \geq -r_t$, and chooses $\mathfrak{b}$ otherwise. Thus, conditioned on $\Omega = 1$ and $\Theta = \theta$, the probability of agent t choosing $\mathfrak{g}$ is $1 - F_\theta(-r_t)$. As a consequence, the public LLR of the agent $\{r_t\}$ evolves as follows:

$$r_{t+1} = r_t + D_{\mathfrak{g}}(r_t) \quad \text{if } a_t = \mathfrak{g}, \tag{3}$$

$$r_{t+1} = r_t + D_{\mathfrak{b}}(r_t) \quad \text{if } a_t = \mathfrak{b}, \tag{4}$$

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 likewise, $D_{\mathfrak{b}}$ always takes negative values. Intuitively, seeing an action $\mathfrak{g}$ makes the agent more convinced that the state is good, so her public belief that the state is good increases. Likewise, seeing an action $\mathfrak{b}$ makes the agent more convinced that the state is bad, and thus her public belief that the state is good decreases.

Overturning Principle and Stationarity. One important property of agents' public beliefs is called the *overturning principle* (Sørensen, 1996; Smith and Sørensen, 2000). It states that observing one action that is different from the previous actions is enough to change the verdict of agents' public beliefs.

Lemma 3 (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 $\pi_{t+1} \leq 1/2$.*

Proof. Fix any $t \geq 1$. By the law of the iterated expectation,

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

Thus, if $a_t = \mathfrak{g}$, then by (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. □

Another important property held by agents' public beliefs is *stationarity*—the value of π_t captures all past information about the payoff-relevant state, independent of time. This property holds in our model, since regardless of the informativeness of the source, when agents observe an action, they always update their public beliefs (in terms of LLR) using (3) or (4). We write $\mathbb{P}_{\omega,\theta,\pi}$ to emphasize the value of the agent's public belief while continuing to omit it when it is uniform.

Lemma 4 (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 $(\omega,\theta) \in \{0,1\} \times \{\mathfrak{g},\mathfrak{b}\}$*

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

That is, regardless of the informativeness of the source, if the public belief at a given time t 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 at time 1, when the prior is π.

5.2. **Perpetual Disagreement and Immediate Herding.** In our model, when the source is informative, the public belief π_t about the payoff-relevant state evolves as in the standard model, so it is a martingale. One can then apply the martingale convergence argument to show that correct herding occurs with unbounded signals, regardless of the achievement of asymptotic learning. However, when the source is uninformative, as π_t still evolves as in the standard model, it ceases to be a martingale under the incorrect measure $\mathbb{P}_0$. Thus, we need to employ a different analytical approach to understand what ensures asymptotic learning in this case.[9]

In the following proposition, we first establish an important relationship between asymptotic learning and *perpetual disagreement*—the event where agents never converge to any action. In other words, perpetual disagreement occurs if both actions are taken infinitely often. We denote by $S = \sum_{t=1}^\infty \mathbb{1}(a_t \neq a_{t+1})$ the total number of action switches and thus $S = \infty$ is the event of perpetual disagreement.

Proposition 1. *Asymptotic learning holds if and only if conditioned on the source being uninformative, perpetual disagreement occurs almost surely, i.e., $\mathbb{P}_0[S = \infty] = 1$.*

This proposition states that an outside observer eventually learns the informativeness of the source if and only if given an uninformative source, agents never converge to any action. Intuitively, since the outside observer knows that agents would eventually herd under informative signals, he can thus infer that the source must be uninformative if perpetual disagreement persists. Conversely, if agents do converge to an action (correct or wrong) under an uninformative source, then he can no longer be sure that the source is uninformative, as such an action convergence is also possible under an informative source.

Next, we focus on the opposite of perpetual disagreement, namely, the phenomenon of *immediate herding*—the event in which all agents immediately herd on an action, i.e., $\{a_1 = a_2 = \ldots = a\}$ for some $a \in \{\mathfrak{g},\mathfrak{b}\}$. We denote such an event by $\{\bar{a} = a\}$. First, we observe that conditioned on an informative source, immediate herding on the wrong

[9]A similar approach can be found in Arieli et al. (2023), where the public belief also ceases to be a martingale under the correct measure because agents who are overconfident have misspecified beliefs.

action is impossible, whereas immediate herding on the correct action is possible, at least for some prior. This observation is summarized in the following lemma.

Lemma 5. *Conditioned on the source being informative and the payoff-relevant state $\Theta = \theta$ for any $\theta \in \{\mathfrak{g}, \mathfrak{b}\}$, the following two conditions hold:*

(i) The event $\{\bar{a} \neq \theta\}$ occurs with probability zero.

(ii) The event $\{\bar{a} = \theta\}$ occurs with positive probability for some prior $\pi \in (0, 1)$.

The first part of this lemma holds since all but finitely many agents take the correct action conditioned on an informative source, it clearly implies that agents cannot immediately herd on the wrong action. Similarly, the second part of this lemma holds because if correct herding eventually occurs, then by the stationarity of this process, agents can also do so immediately for some prior.[10]

Building upon Proposition 1, we now characterize asymptotic learning in terms of immediate herding. Specifically, the next proposition establishes that conditioned on an uninformative source, if asymptotic learning holds, then immediate herding is impossible for any action initiated at the uniform prior. In fact, it is also a sufficient condition for asymptotic learning.

Proposition 2. *Asymptotic learning holds if and only if conditioned on the source being uninformative, immediate herding initiated at the uniform prior occurs with probability zero, i.e., for any action $a \in \{\mathfrak{g}, \mathfrak{b}\}$, $\mathbb{P}_0[\bar{a} = a] = 0$.*

To see why this condition is equivalent to asymptotic learning, let us first consider a stronger condition. That is, given an uninformative source, suppose it is impossible for agents to immediately herd on any action for *all* priors. If this were true, then by stationarity again, agents would never herd on any action starting from any prior. Hence, perpetual disagreement occurs with probability one, and it follows from Proposition 1 that this is equivalent to the achievement of asymptotic learning. The rest of the proof of Proposition 2 aims to show the equivalence between the stronger condition and the condition stated in the proposition, and it relies on the eventual monotonicity of the belief updating process (see Lemma 7 in the appendix).

Therefore, we have reduced the problem of asymptotic learning to the problem of immediate herding, which is much easier to analyze. This is because, conditioned on immediate herding, the process of the public LLR $\{r_t\}$ evolves deterministically according to either (3) or (4). Conditioned on the event $\bar{a} = \mathfrak{g}$, we denote the deterministic process of r_t based on (3) by $r_t^{\mathfrak{g}}$. Recall that agent t chooses $a_t = \mathfrak{g}$ if $\ell_t \geq -r_t$, and $a_t = \mathfrak{b}$ otherwise. Therefore, the probability of $\bar{a} = \mathfrak{g}$ is equal to the probability that $\ell_t \geq -r_t^{\mathfrak{g}}$ for all $t \geq 1$. Since conditioned on an uninformative source, private signals are i.i.d., and

[10] In fact, part (ii) of Lemma 5 holds not only for some prior, but also for all priors. One can see this by applying a similar argument as in the proof of Lemma 8 in the appendix. We omit the stronger statement here as it is not necessary for the proof of our main result.

so are the private log-likelihood ratios. Thus, the probability of immediate herding on action $\mathfrak{g}$ conditioned on an uninformative source 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, this is equivalent to examining whether the sum of the probabilities of the following events is finite or infinite:

$$\mathbb{P}_0[\bar{a} = \mathfrak{g}] > 0 \ (= 0) \ \Leftrightarrow \ \sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}) < \infty \ (= \infty). \tag{5}$$

Similarly, by the symmetry of the model, one has

$$\mathbb{P}_0[\bar{a} = \mathfrak{b}] > 0 \ (= 0) \ \Leftrightarrow \ \sum_{t=1}^{\infty} \big(1 - F_0(r_t^{\mathfrak{g}})\big) < \infty \ (= \infty). \tag{6}$$

In summary, asymptotic learning occurs if and only if, conditioned on the source being uninformative, as signals become more extreme, the probability of generating these extreme signals decreases slowly enough so that both sums in (5) and (6) are infinite. As we discuss below, these sums diverge for uninformative signals with fatter tails and converge for signals with thinner tails.

5.3. **Proof Sketch of Theorem 1.** We conclude this section by providing a sketch of the proof of Theorem 1. Suppose that the source is uninformative and generates signals with fatter tails. Then, by part (i) of Lemma 5, the probability of generating extreme signals decreases relatively slowly under informative signals. Since, under fatter-tailed uninformative signals, this probability declines even more slowly, it implies that the sums of these probabilities in (5) and (6) are infinite. That is, no immediate herding is possible, and thus, by Proposition 2, asymptotic learning holds. In contrast, suppose that the source is uninformative but generates signals with thinner tails. Then, by part (ii) of Lemma 5, the probability of generating extreme signals already decreases relatively fast under informative signals. Since, under thinner-tailed uninformative signals, this probability decreases even more rapidly for either positive or negative extreme signals, this implies that at least one of the sums of these probabilities in (5) and (6) is finite. That is, it is possible to immediately herd on some action, and hence, by Proposition 2, asymptotic learning fails.

6. Discussion

6.1. **Hypothesis Testing.** The question of asymptotic learning in our model is closely related to the problem of hypothesis testing. To illustrate, consider our motivating example where an outside evaluator aims to assess the informativeness of a recommendation system, but only has access to some correlated samples of investment decisions. Based on these samples, he forms the null hypothesis $\mathcal{H}_0$ that the system is informative ($\Omega = 1$) and the alternative hypothesis $\mathcal{H}_1$ that the system is uninformative ($\Omega = 0$). Suppose that the evaluator employs a log-likelihood ratio test:

$$\Lambda(H_t) = \log \frac{\mathbb{P}[\Omega = 1|H_t]}{\mathbb{P}[\Omega = 0|H_t]},$$

where $H_t = (a_1, \ldots, a_t)$ represents a collection of t investment decisions. He then sets a threshold $c > 0$ for his assessment and accepts the null hypothesis if $\Lambda(H_t) \geq c$ and rejects otherwise.

In this context, achieving asymptotic learning (Definition 1) is equivalent to the following two conditions: (i) $\Lambda(H_t) \to \infty$ almost surely under $\mathcal{H}_0$, and (ii) $\Lambda(H_t) \to -\infty$ almost surely under $\mathcal{H}_1$. In words, when the null hypothesis is true, the log-likelihood ratio test $\Lambda(H_t)$ tends to infinity, and meanwhile, it approaches negative infinity when the alternative hypothesis is true. Hence, the question of whether asymptotic learning holds is the same as the question of whether the evaluator can eventually make the correct assessment, regardless of the decision rule employed. In other words, as the size of the correlated sample increases, can the evaluator eventually achieve both zero type I error (false positive) and zero type II error (false negative) for any decision rule? Theorem 1 provides an answer to this question by characterizing the conditions under which it is possible or impossible to achieve both zero error rates asymptotically.

Falsifiability. The idea of falsifiability, introduced by Popper (1959), highlights the importance of being able to test and potentially refute hypotheses.[11] We discuss the case in our model where the failure of asymptotic learning corresponds to the non-falsifiability of the hypothesis that the source is uninformative.

Recall that in Example 2, when F_0 first-order stochastically dominates $F_{\mathfrak{g}}$, the uninformative signals have thinner tails and thus asymptotic learning fails. Suppose that the hypotheses are the same as before, i.e., $\mathcal{H}_0 : \Omega = 1$ and $\mathcal{H}_1 : \Omega = 0$, but the true payoff-relevant state Θ is equal to $\mathfrak{g}$. Then, based on the history of past actions, the alternative hypothesis that the source is uninformative is non-falsifiable because agents would eventually converge to the correct action $\mathfrak{g}$, regardless of whether the null or the alternative hypothesis is true.[12] In other words, if the uninformative source is biased toward the correct state, it becomes impossible to distinguish between the uninformative and the informative source, and so asymptotic learning fails.

6.2. **Bounded Belief Under Unbounded Signals.** One interesting feature introduced by the possibility of an uninformative source is that the agent's *unconditional* private

[11] His famous example that "all swans are white" serves as a null hypothesis that can be easily falsified by observing a single black swan. However, the alternative hypothesis that "there exists a non-white swan" is much harder to falsify as it would require checking every swan to confirm that each of them is white.

[12] Clearly, when the payoff-relevant state Θ is equal to $\mathfrak{b}$, the alternative hypothesis is falsifiable, as agents would eventually herd on action $\mathfrak{b}$ if the null hypothesis is true and on action $\mathfrak{g}$ if the alternative hypothesis is true.

belief about the payoff-relevant state can remain bounded with unbounded signals, as illustrated in the following example. Intuitively, this is because the uncertainty regarding the informativeness of the source reduces the informational value that a single private signal can provide about the payoff-relevant state.

Example 4 (Bounded Belief Induced by Unbounded Signals)**.** Suppose that the informative signals follow distributions $\mu_{\mathfrak{g}} = \mathcal{N}(1, 1)$ and $\mu_{\mathfrak{b}} = \mathcal{N}(-1, 1)$, respectively. The uninformative signals follow the Laplace distribution centered at zero with scale one, i.e., $\mu_0 = \text{Laplace}(0, 1)$ with probability density function $f(x) = \frac{1}{2}\exp(-|x|)$. Note that in this case, $\ell_t = 2s_t$, so all private signals are unbounded.

By Bayes' rule, the log-likelihood ratio of the unconditional private belief of the good state over the bad state induced by a private signal $s \in \mathbb{R}$ is equal to

$$\log \frac{\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}(s-1)^2) + \frac{1}{2}\exp(-|s|)}{\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}(s+1)^2) + \frac{1}{2}\exp(-|s|)}.$$

As the absolute value of the private signal s goes to infinity, the above expression converges to zero, implying that the unconditional private belief is bounded.[13] Since the Laplace distribution has fatter tails compared to the Gaussian informative signals, by Theorem 1, asymptotic learning holds. That is, an outside observer eventually learns the informativeness of the source, even with bounded unconditional private belief.

7. 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. Specifically, it hinges on the relative tail distribution between signals: If uninformative signals have fatter tails compared to their informative counterparts, asymptotic learning holds; conversely, with thinner-tailed uninformative signals, asymptotic learning fails. We also characterize the conditions for asymptotic learning in the canonical case of Gaussian private signals.

More generally, our economic insight suggests that irregular behavior, such as an action switch following a prolonged sequence of identical actions, is the driving force behind asymptotic learning. Contrary to the public-signal benchmark case in which learning is always achieved, with private signals, an outsider observer can only learn that the source is uninformative from observing these action switches. We show that when action switches accumulate indefinitely, the observer can eventually differentiate between an uninformative source and an informative one. In the context of scientific paradigms,

[13] If we replace the Laplace distribution with the standard Gaussian distribution, this private belief remains unbounded.

this insight is reminiscent of Kuhn (1962)'s idea that the accumulation of anomalies may trigger scientific revolutions and paradigm shifts.

A limitation of our results is that they only apply asymptotically. Thus, it would be interesting to understand the speed at which an outside observer learns about the informativeness of the source. Our simulation results suggest that in the case of Gaussian uninformative signals with fatter tails, the asymptotics can already kick in relatively early in the process. Another interesting direction is to allow the information source, e.g., a financial advisor of unknown quality, to be strategic about what advice to give in order to maximize their chances of being employed. In such a setting, one could ask whether society can eventually learn the true quality of this advisor. We leave these questions for future research.

References

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

Arieli, I., Y. Babichenko, S. Müller, F. Pourbabaee, and O. Tamuz (2023): "The Hazards and Benefits of Condescension in Social Learning," *arXiv preprint arXiv:2301.11237*.

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

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

Bikhchandani, S., D. Hirshleifer, O. Tamuz, and I. Welch (2021): "Information cascades and social learning," Tech. rep., National Bureau of Economic Research.

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

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

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

Çelen, B. and S. Kariv (2004): "Observational learning under imperfect information," *Games and Economic behavior*, 47, 72–86.

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

Chen, J. Y. (2022): "Sequential learning under informational ambiguity," *Available at SSRN 3480231*.

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

Frick, M., R. Iijima, and Y. Ishii (2020): "Misinterpreting others and the fragility of social learning," *Econometrica*, 88, 2281–2328.

——— (2023): "Belief convergence under misspecified learning: A martingale approach," *The Review of Economic Studies*, 90, 781–814.

Golub, B. and E. Sadler (2017): "Learning in social networks," *Available at SSRN 2919146*.

Hann-Caruthers, W., V. V. Martynov, and O. Tamuz (2018): "The speed of sequential asymptotic learning," *Journal of Economic Theory*, 173, 383–409.

Herrera, H. and J. Hörner (2012): "A necessary and sufficient condition for information cascades," *Columbia and Yale mimeo*.

Kartik, N., S. Lee, T. Liu, and D. Rappoport (2022): "Beyond Unbounded Beliefs: How Preferences and Information Interplay in Social Learning," Tech. rep.

Kuhn, T. S. (1962): *The structure of scientific revolutions*, University of Chicago press.

Lehmann, E. L. (1988): "Comparing location experiments," *The Annals of Statistics*, 521–533.

Liang, A. and X. Mu (2020): "Complementary information and learning traps," *The Quarterly Journal of Economics*, 135, 389–448.

Lobel, I. and E. Sadler (2015): "Information diffusion in networks through social learning," *Theoretical Economics*, 10, 807–851.

Parzen, E. (1979): "Nonparametric statistical data modeling," *Journal of the American statistical association*, 74, 105–121.

Popper, K. (1959): *The logic of scientific discovery*, Routledge.

Rojo, J. (1992): "A pure-tail ordering based on the ratio of the quantile functions," *The Annals of Statistics*, 570–579.

Rosenberg, D. and N. Vieille (2019): "On the efficiency of social learning," *Econometrica*, 87, 2141–2168.

Samuelson, P. (1965): "Proof that properly anticipated prices fluctuate randomly," *Industrial Management Review*, 6, 41–49.

Smith, L. and P. Sørensen (2000): "Pathological outcomes of observational learning," *Econometrica*, 68, 371–398.

Smith, L., P. N. Sørensen, and J. Tian (2021): "Informational herding, optimal experimentation, and contrarianism," *The Review of Economic Studies*, 88, 2527–2554.

Sørensen, P. N. (1996): "Rational social learning," Ph.D. thesis, Massachusetts Institute of Technology.

Vives, X. (1993): "How fast do rational agents learn?" *The Review of Economic Studies*, 60, 329–347.

Proof of Lemma 1. By Bayes' rule, conditioned on the information $I_t = (H_t, s_t)$, the relative likelihood between the good state and the bad state is equal to:

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

Note that $\mathbb{P}_0[\Theta = \mathfrak{g}|H_t, s_t] = \mathbb{P}_0[\Theta = \mathfrak{g}]$ and $\mathbb{P}_0[\Theta = \mathfrak{b}|H_t, s_t] = \mathbb{P}_0[\Theta = \mathfrak{b}]$. This is because, conditioned on $\Omega = 0$, neither the public history H_t nor the private signal s_t contains information regarding the state Θ. Since the states Ω and Θ are independent of each other, and the prior on Θ is uniform, it follows from $\mathbb{P}_0[\Theta = \mathfrak{g}] = \mathbb{P}_0[\Theta = \mathfrak{b}]$ that $\mathbb{P}_0[\Theta = \mathfrak{g}|H_t, s_t] = \mathbb{P}_0[\Theta = \mathfrak{b}|H_t, s_t]$. Consequently,

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

That is, the equilibrium action for agent t is to choose the most likely action conditioned on the source being informative. □

Proof of Lemma 2. Recall that $H_t = (a_1, \ldots, a_{t-1})$ is the history of actions at time t. Let $H_\infty = \cup_t H_t$, which is the information available to an outside observer who observes all agents' actions at time infinity. Denote by $p_\infty^\Omega := \mathbb{P}[\Omega = 1|H_\infty]$ and $p_\infty^\Theta := \mathbb{P}[\Theta = \mathfrak{g}|H_\infty]$ the belief that he holds about $\Omega = 1$ and $\Theta = \mathfrak{g}$ at time infinity, respectively. By the law of total probability,

$$\begin{aligned} p_\infty^\Theta &= \mathbb{P}_1[\Theta = \mathfrak{g}|H_\infty] \cdot p_\infty^\Omega + \mathbb{P}_0[\Theta = \mathfrak{g}|H_\infty] \cdot (1 - p_\infty^\Omega) \\ &= \pi_\infty \cdot p_\infty^\Omega + \frac{1}{2} \cdot (1 - p_\infty^\Omega), \end{aligned} \tag{A.1}$$

where $\pi_\infty := \mathbb{P}_1[\Theta = \mathfrak{g}|H_\infty]$ and the second equality holds since $\mathbb{P}_0[\Theta = \mathfrak{g}|H_\infty] = \mathbb{P}_0[\Theta = \mathfrak{g}] = 1/2$. Furthermore, since agents always act as if the signals are informative (Lemma 1), conditioned on $\Omega = 1$, the public belief π_t is still a martingale. By a standard martingale convergence argument with unbounded signals (Smith and Sørensen, 2000), $\mathbb{P}_{1,\mathfrak{g}}[\pi_\infty = 1] = \mathbb{P}_{1,\mathfrak{b}}[\pi_\infty = 0] = 1$. By definition, asymptotic learning holds if and only if $\mathbb{P}_{1,\mathfrak{g}}[p_\infty^\Omega = 1] = \mathbb{P}_{1,\mathfrak{b}}[p_\infty^\Omega = 1] = 1$. Thus, by (A.1), asymptotic learning holds if and only if conditioned on $\Omega = 1$, complete learning holds: $\mathbb{P}_{1,\mathfrak{g}}[p_\infty^\Theta = 1] = \mathbb{P}_{1,\mathfrak{b}}[p_\infty^\Theta = 0] = 1$. □

Proof of Lemma 5. The proof idea is similar to the proof of Lemma 10 in Arieli et al. (2023). By the symmetry of the model, we can assume without loss of generality that $\Theta = \mathfrak{g}$. Since conditioned on $\Omega = 1$, correct herding occurs, (2) implies 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), since conditioned on $\Omega = 1$ and $\Theta = \mathfrak{g}$, all agents eventually herd on $\mathfrak{g}$ almost surely, we can define $\tau < \infty$ as the last random time at which the agent chooses the wrong action $\mathfrak{b}$. That is, $1 = \mathbb{P}_{1,\mathfrak{g}}[\lim_t a_t = \mathfrak{g}] = \sum_{k=1}^\infty \mathbb{P}_{1,\mathfrak{g}}[\tau = k]$. By the overturning principle, $a_\tau = \mathfrak{b}$ implies that $\pi_{\tau+1} \le 1/2$. As a consequence,

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

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

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 $p_t^\Omega = \mathbb{P}[\Omega = 1 | H_t]$ is the belief of an outside observer at time t that the source is informative.

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

$$\mathbb{P}_0[p_t^\Omega = \pi] \le \frac{1-a}{a} \cdot \mathbb{P}_1[p_t^\Omega = \pi], \text{ for all } \pi \in [a, 1/2];$$
$$\mathbb{P}_0[p_t^\Omega = \pi] \ge \frac{1-b}{b} \cdot \mathbb{P}_1[p_t^\Omega = \pi], \text{ for all } \pi \in [1/2, b].$$

Proof of Claim 1. Fix any $\pi \in (0,1)$ and let $\tilde{H}_t$ be a realization of the history of actions such that the associated belief p_t^Ω is equal to π. By the law of total expectation,

$$\mathbb{P}[\Omega = 1 | p_t^\Omega = \pi] = \mathbb{E}[\mathbb{E}[\mathbb{1}(\Omega = 1) | \tilde{H}_t] | p_t^\Omega = \pi] = \mathbb{E}[\pi | p_t^\Omega = \pi] = \pi.$$

It follows from Bayes' rule that

$$\frac{\mathbb{P}_0[p_t^\Omega = \pi]}{\mathbb{P}_1[p_t^\Omega = \pi]} = \frac{\mathbb{P}[\Omega = 0 | p_t^\Omega = \pi]}{\mathbb{P}[\Omega = 1 | p_t^\Omega = \pi]} \cdot \frac{\mathbb{P}[\Omega = 1]}{\mathbb{P}[\Omega = 0]} = \frac{1-\pi}{\pi}.$$

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

$$\mathbb{P}_0[p_t^\Omega = \pi] = \frac{1-\pi}{\pi} \cdot \mathbb{P}_1[p_t^\Omega = \pi] \le \frac{1-a}{a} \cdot \mathbb{P}_1[p_t^\Omega = \pi].$$

The second inequality follows from a symmetric argument. $\square$

Proof of Proposition 1. Recall that $S = \sum_{t=1}^{\infty} \mathbb{1}(a_t \neq a_{t+1})$. We show the if-direction first. Consider an outside observer x who observes everyone's actions. The information available to him at time t is $H_t = (a_1, \ldots, a_{t-1})$ and at time infinity is $H_\infty = \cup_t H_t$. He gets a utility of one if his action matches Ω and zero otherwise. Denote by a_t^x the action that x chooses to maximize his probability of matching Ω at time t. For any fixed 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 let $A_t^c(k)$ denote the complementary event.

Consider the following strategy at time infinity for x, for which we denote by $\tilde{a}_\infty^x(k)$: $\tilde{a}_\infty^x(k) = 0$ if $A_\infty(k)$ occurs and $\tilde{a}_\infty^x(k) = 1$ otherwise. For any fixed k, the probability of x matching the state Ω at time infinity under this strategy is

$$\begin{aligned}\mathbb{P}[\tilde{a}_\infty^x(k) = \Omega] &= \mathbb{P}[\Omega = 0, A_\infty(k)] + \mathbb{P}[\Omega = 1, A_\infty^c(k)] \\ &= \mathbb{P}_0[A_\infty(k)] \cdot \mathbb{P}[\Omega = 0] + \mathbb{P}_1[A_\infty^c(k)] \cdot \mathbb{P}[\Omega = 1]. \end{aligned} \tag{A.2}$$

Sinc conditioned on $\Omega = 1$, all agents eventually herd with probability one, it follows that for all k large enough,

$$\mathbb{P}_1[A_\infty^c(k)] = 1. \tag{A.3}$$

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 (A.2) and (A.3) that for all k large enough, $\mathbb{P}[\tilde{a}_\infty^x(k) = \Omega] = 1$.

Hence, the optimal strategy for x at time infinity, i.e., $a_\infty^x = 1$ if $p_\infty^\Omega = \mathbb{P}[\Omega = 1|H_\infty] \geq 1/2$ and $a_\infty^x = 0$ otherwise, must also match the state Ω with probability one:

$$1 = \mathbb{P}[a_\infty^x = \Omega] = \mathbb{P}_1[p_\infty^\Omega \geq 1/2] \cdot \mathbb{P}[\Omega = 1] + \mathbb{P}_0[p_\infty^\Omega < 1/2] \cdot \mathbb{P}[\Omega = 0]. \tag{A.4}$$

The above equation implies that $\mathbb{P}_0[p_\infty^\Omega < 1/2] = \mathbb{P}_1[p_\infty^\Omega \geq 1/2] = 1$. It remains to show that $\mathbb{P}_0[p_\infty^\Omega = 0] = 1$ and $\mathbb{P}_1[p_\infty^\Omega = 1] = 1$. To this end, note that $\mathbb{P}_0[p_\infty^\Omega < 1/2] = 1$ implies that $\mathbb{P}_0[p_\infty^\Omega \geq 1/2] = 0$. Thus, by Claim 1, one has that for any $b \in (1/2, 1)$, $\mathbb{P}_1[p_\infty^\Omega = \pi] = 0$ for all $\pi \in [1/2, b]$. It then follows from $\mathbb{P}_1[p_\infty^\Omega \geq 1/2] = 1$ that $\mathbb{P}_1[p_\infty^\Omega = 1] = 1$. The case that $\mathbb{P}_0[p_\infty^\Omega = 0] = 1$ follows from an identical argument.

Next, we show the only-if direction by contraposition. Suppose $\mathbb{P}_0[S < \infty] > 0$. Since conditioned on $\Omega = 1$, all agents eventually herd almost surely, it implies that $\mathbb{P}_1[S < \infty] = 1$. Thus, there exists a history at time infinity $\tilde{H}_\infty$ that has a positive probability under both probability measures $\mathbb{P}_0$ and $\mathbb{P}_1$: $\mathbb{P}_0[\tilde{H}_\infty] > 0$ and $\mathbb{P}_1[\tilde{H}_\infty] > 0$. Thus, by Bayes' rule, $0 < \mathbb{P}[\Omega = 1|\tilde{H}_\infty] < 1$ and similarly, $0 < \mathbb{P}[\Omega = 0|\tilde{H}_\infty] < 1$.

Assume without loss of generality that under this history $\tilde{H}_\infty$, $\tilde{p}_\infty^\Omega \geq 1/2$, so conditioned on $\tilde{H}_\infty$, the outside observer x should choose $a_\infty^x = 1$. As a consequence, the probability of x matching the state Ω is strictly less than one:

$$\begin{aligned}\mathbb{P}[a_\infty^x = \Omega] &= \mathbb{P}[a_\infty^x = \Omega, \tilde{H}_\infty] + \mathbb{P}[a_\infty^x = \Omega, \tilde{H}_\infty^c] \\ &= \mathbb{P}[\Omega = 1|\tilde{H}_\infty] \cdot \mathbb{P}[\tilde{H}_\infty] + \mathbb{P}[a_\infty^x = \Omega, \tilde{H}_\infty^c] \\ &< \mathbb{P}[\tilde{H}_\infty] + \mathbb{P}[\tilde{H}_\infty^c] = 1.\end{aligned}$$

But the above inequality implies that asymptotic learning fails since if otherwise, it would follow from the definition of asymptotic learning and (A.4) that $\mathbb{P}[a_\infty^x = \Omega] = 1$, a contradiction. $\square$

Appendix B. Proof of Theorem 1

In this section, we prove Proposition 2 and Theorem 1. To prove Proposition 2, we will first prove the following proposition. Then, together with Proposition 1, they jointly imply Proposition 2. The proof of Theorem 1 is presented at the end of this section.

Proposition 3. *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 public log-likelihood ratio r_t, conditioned on immediate herding. These results lead to Lemma 8, a crucial part in establishing the equivalence between no immediate herding and perpetual disagreement. We present the proof of Proposition 3 towards the end of this section.

Preliminaries. By the symmetry of the model, we can focus without loss of generality on the case of immediate herding on action $\mathfrak{g}$. Recall that, conditioned on $\{\bar{a} = \mathfrak{g}\}$, the process of the public log-likelihood ratio r_t evolves deterministically according to (3). We denote such a process by $r_t^{\mathfrak{g}}$, and denote the corresponding updating function by

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

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 is specified, we denote the sequence $r_t^{\mathfrak{g}}$ with an initial value of r by $r_t^{\mathfrak{g}}(r)$. Thus, we can 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 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 over time, and the second part shows that it takes only some bounded time for $r_t^{\mathfrak{g}}$ to reach any positive value.

Lemma 6 (The Long-Run and Short-Run Behaviors of $r_t^{\mathfrak{g}}$)**.**

(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 $r_t^{\mathfrak{g}}$ eventually approaches infinity, it may not do so monotonically without additional assumptions on the distributions $F_{\mathfrak{g}}$ and $F_{\mathfrak{b}}$.[14] 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.

[14]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.

Lemma 7 (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, $D_g(x) = \log \frac{1-F_{\mathfrak{g}}(-x)}{1-F_{\mathfrak{b}}(-x)}$ and taking the derivative of $D_{\mathfrak{g}}$, one has

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

Observe that the log-likelihood ratio of the log-likelihood ratio of the private belief is the log-likelihood ratio itself (see, e.g., Chamley (2004)):

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

It thus follows that

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

Hence, for some x large enough, there exists $\varepsilon > 0$ small enough such that $-D'_{\mathfrak{g}}(x) \leq (1+\varepsilon)f_{\mathfrak{b}}(-x)$ and $(1+\varepsilon)\rho \leq 1$. Hence, for all $x \leq x'$,

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

Rearranging the above equation, one has

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

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

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

Lemma 8. *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 $\bar{a} = \mathfrak{g}$ with $\pi_1 = \tilde{\pi}$ 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., we can write

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

Thus, $\mathbb{P}_{0,\tilde{\pi}}[\bar{a} = \mathfrak{g}] > 0$ 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 6), $\log\left(1 - F_0(-r_t^{\mathfrak{g}}(\tilde{r}))\right) \approx -F_0(-r_t^{\mathfrak{g}}(\tilde{r}))$. Thus, the above sum is finite if and only if

$$\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}(\tilde{r})) < M. \tag{B.2}$$

By the overturning principle, it suffices to show that (B.2) 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 7) and the fact that $r_t^{\mathfrak{g}}(\tilde{r}) \to \infty$, we can find a large enough $\bar{t}$ such that $\phi(r) \geq \phi(\bar{r})$ for all $r \geq \bar{r}$, where $\bar{r} = r_{\bar{t}}^{\mathfrak{g}}(\tilde{r}) \geq 0$. By part (ii) of Lemma 6, 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 for all $\tau \geq 1$

$$F_0(-r_{\tau+t_0}^{\mathfrak{g}}(r)) \leq F_0(-r_{\tau+\bar{t}}^{\mathfrak{g}}(\tilde{r})), \quad \text{for any } r \geq 0.$$

Thus, (B.2) implies that for any $r \geq 0$, $\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}(r)) < \infty$, as required. The case of action $\mathfrak{b}$ follows from a symmetric argument. $\square$

Now, we are ready to prove Proposition 3.

Proof of Proposition 3. 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 for all prior $\pi \in (0,1)$, $\mathbb{P}_{0,\pi}[\bar{a} = a] > 0$. In particular, it holds for the uniform prior so that $\mathbb{P}_0[\bar{a} = a] > 0$. Since the event $\{\bar{a} = a\}$ is contained in the event $\{S < \infty\}$,

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

which implies that $\mathbb{P}_0[S = \infty] < 1$, a contradiction to (ii). Finally, by taking the negation of the statements in Lemma 8, it follows that (iii) implies (i). This concludes the proof of Proposition 3. $\square$

Proof of Proposition 2. By the equivalence between (ii) and (iii) in Proposition 3 and its equivalence to asymptotic learning (Proposition 1), we have shown Proposition 2. $\square$

Given Proposition 2, we are now ready to prove our main result.

Proof of Theorem 1. Suppose that the uninformative signals have fatter tails than the informative signals. That is, there exists some $\varepsilon > 0$ such that for all x large enough, $F_0(-x) \geq \varepsilon \cdot F_{\mathfrak{b}}(-x)$ and $1 - F_0(x) \geq \varepsilon \cdot (1 - F_{\mathfrak{g}}(x))$. By part (i) of Lemma 5, $\mathbb{P}_{1,\mathfrak{b}}[\bar{a} = \mathfrak{g}] = 0$ and $\mathbb{P}_{1,\mathfrak{g}}[\bar{a} = \mathfrak{b}] = 0$. Since conditioned on $\Omega = 1$ and $\Theta = \mathfrak{b}$, signals are i.i.d., we can write

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

This implies that $-\sum_{t=1}^{\infty} \log(1 - F_{\mathfrak{b}}(-r_t^{\mathfrak{g}})) = \infty$. 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

$$\sum_{t=1}^{\infty} F_{\mathfrak{b}}(-r_t^{\mathfrak{g}}) = \infty. \tag{B.3}$$

Similarly, applying the same logic that led to (B.3), one has

$$\mathbb{P}_{1,\mathfrak{g}}[\bar{a} = \mathfrak{b}] = 0 \Leftrightarrow \sum_{t=1}^{\infty}(1 - F_{\mathfrak{g}}(-r_t^{\mathfrak{b}})) = \infty, \tag{B.4}$$

where $r_t^{\mathfrak{b}}$ is the deterministic process of r_t that evolves according to (4). By the symmetry of the model, one has $r_t^{\mathfrak{b}} = -r_t^{\mathfrak{g}}$ for all $t \geq 1$. Since the uninformative signals have fatter tails, it follows from (B.3) and (B.4) that $\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}) = \infty$ and $\sum_{t=1}^{\infty}(1 - F_0(r_t^{\mathfrak{g}})) = \infty$. By the same logic that we use to deduce (B.3), these are equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{g}] = 0$ and $\mathbb{P}_0[\bar{a} = \mathfrak{b}] = 0$. Thus, by Proposition 2, asymptotic learning holds.

Now suppose that the uninformative signals have thinner tails than the informative ones. That is, there exists some $\varepsilon > 0$ such that either $F_0(-x) \leq (1/\varepsilon) \cdot F_{\mathfrak{g}}(-x)$ for all x large enough, or $1 - F_0(x) \leq (1/\varepsilon) \cdot (1 - F_{\mathfrak{b}}(x))$ for all x large enough. By part (ii) of Lemma 5, 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 the same argument that led to (B.2), these are equivalent to

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

Since the uninformative signals have thinner tails, the above equations imply that either $\sum_{t=1}^{\infty} F_0(-r_t^{\mathfrak{g}}(r)) < \infty$ or $\sum_{t=1}^{\infty}(1 - F_0(r_t^{\mathfrak{g}}(r')))$, which means that either $\mathbb{P}_{0,\pi}[\bar{a} = \mathfrak{g}] > 0$ for some $\pi \in (0,1)$ or $\mathbb{P}_{0,\pi'}[\bar{a} = \mathfrak{b}] > 0$ for some $\pi' \in (0,1)$. It follows from Lemma 8 that these also hold for all prior $\pi, \pi' \in (0,1)$, including the uniform prior, so we have either $\mathbb{P}_0[\bar{a} = \mathfrak{g}] > 0$ or $\mathbb{P}_0[\bar{a} = \mathfrak{b}] > 0$. Therefore, Proposition 2 implies that asymptotic learning fails. □

Appendix C. Gaussian Private Signals

In this section, we prove Corollary 1 and Theorem 2. We say that private signals are Gaussian when $\mu_{\mathfrak{g}}$ and $\mu_{\mathfrak{b}}$ are normal distributions with the same variance σ^2 and mean $+1$ and -1, respectively. In addition, μ_0 is also a normal distribution with mean $m_0 \in (-1, 1)$ and variance τ^2. Notice that in this case, conditioned on $\Omega = 1$, the log-likelihood ratio induced by a private signal s_t is

$$\ell_t = \log \frac{f_{\mathfrak{g}}(s_t)}{f_{\mathfrak{b}}(s_t)} = \frac{2}{\sigma^2} s_t. \tag{C.1}$$

Since ℓ_t is proportional to s_t, its corresponding distributions F_0 and F_θ are also normal, with variances $4\tau^2/\sigma^4$ and $4/\sigma^2$, respectively.

The following lemma demonstrates that, for Gaussian signals, the relative thickness of the tails depends solely on the relative variances. According to Definition 2, for a pair of distributions (F, G), we say that F has a fatter left (or right) tail than G if the corresponding tail ratio is eventually bounded below. Conversely, we say that F has a thinner left (or right) tail than G if the corresponding tail ratio is eventually bounded above.

Lemma 9. *Suppose F and G are two Gaussian cumulative distribution functions with mean μ_1, μ_2 and variance σ_1^2 and σ_2^2, respectively. If $\sigma_1 > \sigma_2$, then F has fatter (left and right) tails than G, and conversely, G has thinner (left and right) tails than F, regardless of the means.*

Proof. Let f and g denote the probability density functions of F and G. Then, for any $x \in \mathbb{R}$

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

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 implies that for all x large enough, $F(-x) \geq G(-x)$ and $1 - F(x) \geq 1 - G(x)$. Thus, we conclude that 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 9 and Theorem 1.

Proof of Corollary 1. If $\tau > \sigma$, then by (C.1), F_0 has a strictly higher variance than both $F_\mathfrak{g}$ and $F_\mathfrak{b}$. Consequently, Lemma 9 implies that F_0 has fatter tails than both $F_\mathfrak{g}$ and $F_\mathfrak{b}$, and so, it follows from Theorem 1 that asymptotic learning holds. A similar argument applies to the case where $\tau < \sigma$. $\square$

We henceforth focus on the case where $\tau = \sigma$. With Gaussian private signals, 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):

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

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

Proof of Theorem 2. In the 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 $|f(t)| \le M \cdot g(t)$ for all $t \ge t_0$.

Note that by (C.1), conditioned on $\Omega = 0$, we can write

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

Thus, for t large enough, it follows from (C.2) that

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

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 $\mu_0(x) = \frac{1}{2}\operatorname{erfc}(-\frac{x-m_0}{\sigma\sqrt{2}})$, where $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty \mathrm{e}^{-t^2}dt$ is the complementary error function.

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

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

where $\delta > 0$ is a constant.

Case (i): suppose $m_0 = 0$. Then (C.3) 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 $\frac{1}{t\log t} \le \frac{1}{t\sqrt{\log t}}$ for all $t \ge 2$, the sum of the first term also diverges. As a consequence, one has $\sum_{t=1}^\infty F_0(-r_t^\mathfrak{g}) = \infty$, and by the same argument used in the proof of Theorem 1 (see, e.g., (B.3)), this is equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{g}] = 0$. Since μ_0 is symmetric around 0, the case $\mathbb{P}_0[\bar{a} = \mathfrak{b}] = 0$ follows from a symmetric argument. Thus, we conclude that for any action $a \in \{\mathfrak{g}, \mathfrak{b}\}$, $\mathbb{P}_0[\bar{a} = a] = 0$ and by Proposition 2, 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 $c = \frac{m_0\sqrt{2}}{\sigma} > 0$. From the integral test it follows that the sum in (C.3) converges and, consequently, $\sum_t^\infty F_0(-r_t^{\mathfrak{g}}) < \infty$. Again, this is equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{g}] > 0$. If, on the other hand, $m_0 < 0$, then it follows from a symmetry argument that $\sum_t^\infty (1 - F_0(r_t^{\mathfrak{g}})) < \infty$, which is equivalent to $\mathbb{P}_0[\bar{a} = \mathfrak{b}] > 0$. Therefore, according to Proposition 2, asymptotic learning fails for any $m_0 \in (-1, 1)$ such that $m_0 \neq 0$. $\square$