Naive Social Learning, Mislearning, and Unlearning

Tristan Gagnon-Bartsch
Harvard University

Matthew Rabin*
Harvard University

April 19, 2016

Abstract

We study social learning in several natural, yet under-explored, environments among people who naively think each predecessor's action reflects solely that person's private information. Such naivete leads to striking forms of mislearning, yielding a set of states of the world that agents always come to disbelieve even when true. When the truth lies in this set, society will "unlearn": although an early generation learns the truth, later generations come to believe something false. Society remains confident only in those hypotheses such that the behavior chosen by people certain of the hypothesis most closely resembles the behavior of privately informed agents if that hypothesis were true. We demonstrate manifestations of these principles in a variety of settings. In cases where options such as assets or medicines have quality independent of one another, naive inference polarizes perceptions: people overestimate the quality of the best option and underestimate the quality of all lesser options. When choosing a portfolio of investments, such polarized perceptions lead to severe under-diversification. In markets with diverse tastes and a large set of options, these perceptions lead all consumers to use a product beneficial for only a small minority. When there is uncertainty over the information structure, people tend to overestimate the informativeness of their signals.

JEL Classification: B49, D82, D83.
Keywords: Observational learning, naive inference, herding.

*Department of Economics and Business School, Harvard University. E-mails: gagnonbartsch@fas.harvard.edu and matthewrabin@fas.harvard.edu. We are grateful to Ben Bushong, Erik Eyster, Josh Schwartzstein, Neil Thakral, He Yang, seminar participants at Harvard-MIT, London School of Economics, UC Berkeley and the 2015 CSEF-IGIER Symposium for helpful feedback.

1 Introduction

People often rely on the behavior of others to inform their decisions. Diners select restaurants based on the crowds gathered outside, investors adopt financial strategies based on their friends' portfolios, and both doctors and patients consider the popularity of a drug when choosing prescriptions. Understanding how knowledge spreads through observational learning is the premise of a large literature, starting with Banerjee (1992) and Bikchandani, Hirshleifer, and Welch (1992). They emphasize how rationality might lead people to imitate others. However, Eyster and Rabin (2014) show that, in virtually all settings, rationality demands limited imitation (and sometimes *anti*-imitation), since imitation by earlier people renders the information in group behavior massively redundant. Because the rationality needed to fully discount this redundancy when our neighbors select the same restaurant, stock, or drug seems extremely unrealistic, there is growing interest (e.g., DeMarzo, Vayanos, and Zwiebel 2003; Eyster and Rabin 2010, 2014) in understanding forms of "naive" social learning.[1] In these models, people neglect the redundancies in the information gleaned from others' behavior, which causes excessive imitation and induces overconfidence. This paper explores novel implications of naive social learning that emerge in an array of environments extending beyond those previously studied.

To first develop an intuition for naive inference, consider privately informed agents who sequentially choose from a set of options with unknown payoffs. Each person observes the choices of all those before her, but neglects that her predecessors are themselves learning from *their* predecessors. Thus, she treats each observed action as if it reflects solely that predecessor's private signal. For instance, a naive customer thinks *each* diner at a crowded restaurant has a positive signal about the restaurant, and thus forms overly strong beliefs about its quality. Eyster and Rabin (2010) (henceforth ER) formalize this logic and show that, in information-rich environments where rational players surely learn the truth, naive players likewise grow fully confident—but with positive probability in the wrong thing.

By examining a broader range of natural environments than earlier papers, we reveal further general principles of naive learning, and we use these principles to draw out welfare implications and comparative statics across several applications. Our primary insight is that naivete sharply limits the states of the world in which society may come to believe. Specifically, there can exist "abandoned states" that people *always* come to disbelieve even when they are true. In fact, even if an early generation knows for certain that such a state is true, later generations "unlearn": they become convinced of something false. The existence of abandoned states has clear economic

[1] In addition to DeMarzo et al. (2003), several other papers study, both theoretically and empirically, heuristic rules of thumb that seem to embed redundancy neglect in structures ubiquitously studied in the network literature. Such papers, which typically build from an early model by DeGroot (1974), include Golub and Jackson (2010), Chandrasekhar, Larreguy, and Xandri (2012), and Levy and Razin (2015).

implications. For instance, when inferring the payoff difference between two independent alternatives, naive learners maximally exaggerate this difference. Not only will they grow too confident a popular medicine is better than an unpopular one, they infer it is *much* better—even when in reality they are nearly identical. We show that such exaggeration can lead to severe under-diversification in investment settings, and to costly herding in markets with diverse tastes. People who naively exaggerate quality differences, for instance, may engage in wasteful queuing or spending to such a degree that they are worse off observing others than if they made decisions in isolation.[2]

Earlier models of naivete obscure these results by focusing on the canonical social-learning environment with just two states of the world and common preferences. While rational models assume this for analytical ease, the tractability of the naivete model allows us to study more general—and more natural—settings. In doing so, we reveal new implications of naivete. One such result is deterministic mislearning: while ER show that people mislearn in the canonical environment only when early signals are misleading, the existence of "abandoned states" in richer settings guarantees mislearning irrespective of early signal realizations. In fact, to articulate the different nature of our mislearning, we consider environments where a large generation of players take actions each round. This means early signals are not misleading. Even without the primacy effects that drive bounded learning in rational models and mislearning in ER, naive inference still leads society astray through agents' misunderstanding of the informational content of past actions.[3]

Section 2 introduces our general model. Every period, a new generation of players chooses from an identical set of actions and receive payoffs that depend on their own action and an unknown state of the world. Each player updates her beliefs based on predecessors' behavior and her conditionally independent private signal. Players naively infer as if others' actions reflect solely their private information and the common prior. While most of our results hold more generally, we focus on a setting with two key features: each generation (1) is infinitely large, and (2) observes only the preceding generation. Combined, these assumptions simplify analysis. And the first makes long-run mislearning more surprising: because the first generation consists of a large number people acting independently, Generation 2 always learns the true state.

Although Generation 2 learns correctly, naive inference can lead Generation 3 astray. Generation 3 wrongly assumes each person in Generation 2 acts "autarkically"—relying only on her own signal—rather than realizing each person in Generation 2 has perfect information. As such, Gen-

[2]While we identify new ways in which naive learning harms welfare, earlier models already demonstrate that naive observers can be made worse off in expectation by observing others' choices. This never happens with full rationality. In herd models such as Bikchandani, et al. (1992), information cascades prevent fully-rational players from learning the state of the world. However, rational players are not *harmed* by observing others—the informational externality simply prevents society from reaching the first best.

[3]Our focus is for clarity and tractability: our main results do not require large generations nor the elimination of primacy effects. Even if players act in single file, the "abandoned states" we identify still induce mislearning. We discuss the robustness of our results at relevant points throughout the paper.

eration 3 comes to believe in the state most likely to generate those signals necessary for autarkic behavior to resemble the behavior of Generation 2.[4] If this is the true state, then Generation 3 and all subsequent generations learn correctly. Otherwise, Generation 3 "unlearns", and long-run beliefs never settle on the truth. Whenever public beliefs do converge, they do so to a "fixed point" of this process: when interpreted as autarkic, the behavior of those certain in state ω is best predicted by ω.

To be concrete, imagine investors learning about the net return on two projects, A and B. It is common knowledge that the payoff of A, ω^A, is a fifty-fifty draw from $\{\$500, \$1000\}$, while ω^B is an independent fifty-fifty draw from $\{\$0, \$700\}$. Suppose the signal distributions conditional on (ω^A, ω^B) have the following intuitive properties: (1) the percent of investors who choose A based solely on their signal is strictly increasing in $\omega^A - \omega^B$, and (2) this percent exceeds 50% if and only if $\omega^A > \omega^B$. If in truth $(\omega^A, \omega^B) = (\$1000, \$700)$, a majority choose A in period 1. From this, investors in period 2 correctly infer that A outperforms B, so *all* choose A. In the third round, investors' best explanation for such a consensus is that A has the largest possible payoff advantage over B, as this state maximizes the likelihood an investor selects A using private information alone. Following this logic, Generation 3 comes to believe the state is $(\$1000, \$0)$ whenever they see all pick A—that is, whenever $(\omega^A, \omega^B) \in \{(\$1000, \$0), (\$500, \$0), (\$1000, \$700)\}$. Thus, in states $(\$500, \$0)$ and $(\$1000, \$700)$, investors inevitably mislearn by way of exaggerating the quality difference between projects.[5] Although this exaggeration does not lead to wrong choices in this simple example, we show that in settings where investors can act on the extremity of their beliefs—for instance, they can contribute more resources toward A—naivete leads to costly mistakes.

Section 3 discusses general implications of naive inference for long-run beliefs. First, based on the "fixed-point" logic above, we characterize the set of states on which public beliefs can settle. We draw out several implications of this characterization, revealing the extent to which naive inference limits the conclusions society might draw. For instance, the set of stable long-run beliefs may be a singleton—society draws the same conclusion no matter what is true—or it may be empty—in which case beliefs continually cycle.

Section 3 also presents a simple yet stark implication of naive inference in settings where, as in the investment example above, players have common preferences and options have independent payoffs. Namely, perceptions of payoffs grow "polarized": people come to believe the best option is as good as possible, and all lesser options are as bad as possible. Once a herd starts on option

[4] Generally, the confident behavior of Generation 2 will not match the action distribution predicted by autarkic play in any state. Because naive observers attribute such discrepancies to sampling variation, Generation 3 grows certain of the state that *best* predicts autarkic behavior. With naivete, this state minimizes the cross entropy between the realized action distribution and the one predicted by autarkic play.

[5] In state $(\$500, \$700)$ investors correctly learn that B is optimal, as $(\$500, \$700)$ is the only state in which a majority selects B.

A, people think that each of the many predecessors who took A received an independent signal indicating A is better than its alternatives. Under natural assumptions on the signal structure, this misconception leads to these extreme, polarized beliefs.[6]

We explore in Section 4 how these extreme beliefs distort investors' allocation of resources across risky assets. Perceptions of an asset's value continually grow more extreme over time, resulting in severe under-diversification. To illustrate, suppose investors know the expected return of asset A, but learn about B's average return, ω, from predecessors' allocations.[7] Although first-period allocations resolve this uncertainty, naive investors overreact to later allocations as if they reflect new information. When ω beats expectations, investors in period 2 correctly infer this and allocate more to B. However, because later investors neglect that this increase is the result of social learning, they wrongly attribute it to new, more optimistic information. Investment in B increases yet again. As this process plays out, investors eventually allocate all wealth to B. Worse, when ω is sufficiently far from priors, investors allocate everything to the wrong asset. In a richer model with prices, our prediction of over-reaction to private information accords with Glaeser and Nathanson's (2015) analysis of housing-price data, which suggests that medium-run momentum derives, in part, from naive inference based on past market prices. It may also help explain unwarranted extrapolation of returns in financial markets, as in Greenwood and Shleifer (2014).

Section 5 examines the consequences of extreme beliefs when people have diverse preferences over one set of options (e.g., medical treatments) relative to an outside option (e.g., no treatment). In such settings, naivete can cause excessive, costly herding where all people choose an option beneficial for only a minority. Since naivete leads people to exaggerate the quality of popular options, those with low valuations are enticed to follow the herd when they should in fact pursue their outside option. Interestingly, we show that such "over-adoption" can be triggered by large choice sets. For example, patients with a disease can either experiment with unproven treatments or abstain. Most patients abstain unless they are fairly confident a treatment works, while a minority in dire straits experiments no matter what. If the number of available treatments is sufficiently large, then society comes to believe one is universally beneficial even when none are. In such states, Generation 2 learns that none are fully effective. While those who are not desperate abstain, those in dire straits herd on the "least bad" treatment. Naive observers, who mistake the consensus among the desperate for independent decisions, grow convinced that this treatment works: they

[6]Because ER considers the binary-state model where the payoff difference between any two actions takes on just two possible values, their setting obscures our result that naive players maximally exaggerate the payoff difference between the herd action and its alternatives.

[7]An important feature of our model is that people observe their predecessors' behavior but not the outcomes of these decisions. For this reason, we have in mind assets that pay off long after the initial investment decision, such as real-estate development or education.

think that with so many options to pick from, it is unlikely that their autarkistic choices would coincide unless the treatment were truly effective. Hence, all patients, including those harmed by treatment, adopt this false cure.

We explore settings with uncertainty about the information structure in Section 6. Since there is very little variation in actions in a herd, naive observers conclude that the variation in signals is also minimal. Hence, when the precision of private signals is uncertain, naive observers will overestimate it.[8] This mistake of misidentifying the source of consensus as pervasive strong information rather than the aggregation of dispersed weak information may not matter much in environments where people only see consensus opinions. But in situations where a person sees only a few individuals' information or hears the advice of some would-be expert before she learns from others, it may cause problems. If the typical reliability of advice by doctors or financial advisors who follow others is misidentified as intrinsic wisdom, people are in danger of listening to misleading isolated thoughts from these folks.

We conclude in Section 7 by discussing the robustness of our results and proposing some extensions. Although we do not provide a formal analysis, we argue that our results hold under alternative observation structures, including canonical settings where a single player acts per period and each observes the complete history of play. Finally, we discuss a few applications that are somewhat beyond the scope of our particular solution concept. These include asset pricing and investment settings where agents endogenously choose when to act.

2 Model

This section presents our general model. Section 2.1 describes the social-learning environment, and Section 2.2 formally defines naive inference. To differentiate our setting from those previously considered, Section 2.3 discusses the framework and results of Eyster and Rabin (2010) and other models of naive inference.

[8] Because Section 6 considers cases where people try to figure out the extent of correlation between others' direct information, it also highlights the nature of the naivete in our model. As in Eyster and Rabin (2010, 2014), naive players in this model neglect only the correlation in others' behavior that is due specifically to social learning. This "redundancy neglect" in interpreting social actions contrasts with other formal models, such as DeMarzo, et al. (2003) and Levy and Razin (2015), that consider different forms of correlation neglect in specific contexts. Our concrete model rules out some realistic forms of correlation neglect (e.g., what seems to be found in an experiment by Enke and Zimmerman (2015)). But its focus on social-learning-induced redundancy neglect yields findings suggesting that naivete may lead people to *exaggerate* the correlation in others' signals.

2.1 Social-Learning Environment

In every period $t = 1, 2, ...$, a new set of N players enters, and each simultaneously takes an action. To ease exposition, we assume a finite set of actions $\mathscr{A} \equiv \{A_1, ..., A_M\}$, where $M \geq 2$.[9] Each player is labeled (n,t), where t is the period in which she acts and $n \in \{1, ..., N\}$ is her index within Generation t. Let $x_{(n,t)}$ denote Player (n,t)'s action and let $a_t(m)$ denote the fraction of players in period t who take action A_m. Vector $\mathbf{a}_t = (a_t(1), ..., a_t(M))$ is the action distribution played in period t.

Players wish to learn an unknown payoff-relevant state of the world $\omega \in \{\omega_1, \omega_2, ..., \omega_K\} \equiv \Omega$, $K < \infty$. Players share a common prior $\pi_1 \in \Delta(\Omega)$, where $\pi_1(k) > 0$ denotes the probability of state ω_k. A player's payoff from action A_m depends on ω and, in applications with heterogeneous preferences, her type $\theta \in \{\theta_1, \theta_2, ...\theta_J\} \equiv \Theta$; payoffs are denoted by $u(A_m|\omega, \theta)$. Preference types are private information and are i.i.d. across players according to a commonly known probability measure $\lambda \in \Delta(\Theta)$.

Players learn about the state from two information channels: private signals and social observation. Each Player (n,t) receives a random private signal $\mathbf{s}_{(n,t)} \in \mathbb{R}^d$, $d \geq 1$, about the state of the world. Conditional on state ω, private signals are i.i.d. across players with c.d.f. $F(\cdot|\omega)$ and density (or mass) function $f(\cdot|\omega)$. We assume these signal distributions have identical support $\mathscr{S} \subset \mathbb{R}^d$ for each $\omega \in \Omega$. The only additional restriction on distributions $\{F(\cdot|\omega)\}_{\omega \in \Omega}$ is an identification assumption introduced below.

For convenience, we will focus throughout the paper on a specific observation structure with two key features. First, we assume players observe the behavior of only the previous generation. Formally, Player (n,t)'s information set $I_{(n,t)} \equiv \{\mathbf{s}_{(n,t)}, \mathbf{a}_{t-1}\}$ consists of her private signal and the complete distribution of actions in $t-1$. Second, ensuring that each generation reaches a nearly confident consensus on the state, we focus on the limit as each generation grows large (i.e., $N \to \infty$). In fact, to simplify our analysis, we follow a "large-generations" convention similar to Banerjee and Fudenberg (2004): in each generation and each state ω, the fraction of players who receive signal $\mathbf{s}$ exactly equals $f(\mathbf{s}|\omega)$ and the fraction of players with type θ exactly equals $\lambda(\theta)$. This allows us to compute for any ω the fraction of players who strictly prefer each action A_m when best responding solely to private signals, and simplifies analysis by generating deterministic belief and action dynamics.[10]

[9] This environment closely resembles Jackson and Kalai's (1997) notion of a "recurring game": each period, a new set of players—randomly drawn according to a time-invariant type distribution—play an identical finite stage game. We extend our model to settings with continuous action spaces in some applications below.

[10] Our social-learning environment deviates somewhat from the canonical models developed by Banerjee (1992), Bikchandani et al. (1992), and Smith and Sørensen (2000). These models assume players observe the complete history of actions and take actions in "single file" ($N = 1$). Banerjee (1992) and Bikchandani et al. (1992) assume discrete signal distributions and common preferences ($|\Theta| = 1$), while Smith and Sørensen (2000) generalize the model by allowing continuous signals and multiple preference types. Our "large-generations" observation structure is

Our simple setting can be interpreted as large, overlapping generations. Each generation is present for two periods: in the first, they observe the actions of the preceding generation, and in the second, they take actions based on inference from this observation. Although this structure is convenient, we argue in Section 7 that it is not necessary. In fact, our main results hold under a range of observation structures, including the most common structure in the literature (e.g., Bikchandani et al. 1992; Smith and Sørensen 2000) in which players act in single file and observe *all* past actions in order. We focus on this particular "overlapping-generations" structure solely for tractability and to starkly highlight how the dynamics of rational and naive beliefs differ.

We call the belief Player (n,t) would form solely from observed actions her *public belief*. Upon observing $\mathbf{a}_{t-1}$, Generation t forms public belief $\pi_t \in \Delta(\Omega)$ using Bayes' rule:

$$\pi_t(j) = \frac{\Pr(\mathbf{a}_{t-1}|\omega_j)\pi_1(j)}{\sum_{k=1}^{K}\Pr(\mathbf{a}_{t-1}|\omega_k)\pi_1(k)}. \tag{1}$$

Our model of naivete will assume that players err only in their calculation of $\Pr(\mathbf{a}_{t-1}|\omega_k)$. Aside from this mistake, each Player (n,t) is rational: she uses Bayes' Rule to combine the public belief with her signal to form posterior $\mathbf{p}_{(n,t)} \in \Delta(\Omega)$, and chooses an action to maximize expected utility, $x_{(n,t)} \in \arg\max_{A\in\mathscr{A}} \sum_{k=1}^{K} p_{(n,t)}(k)u(A|\omega_k,\theta_{(n,t)})$. To avoid annoyances that may arise when the optimal action is not unique, we assume players follow a commonly known tie-breaking rule whenever indifferent.

2.2 Naive Social Inference

Following Eyster and Rabin (2010), we assume individuals naively think that any predecessor's action relies solely on that player's private information. This implies that a naive agent draws inference *as if* all her predecessors ignored the history of play and hence learned nothing from others' actions. That is, she infers as if all her predecessors acted in "autarky".[11] In Equation 1 above, a naive player miscalculates $\Pr(\mathbf{a}_{t-1}|\omega)$—the likelihood of action profile $\mathbf{a}_{t-1}$ in state ω. The true probability depends on the beliefs of Generation $t-1$, which in turn depend on what $t-1$ observed, and so on. Naive players neglect the social inference conducted by the preceding generations, and instead learn from $\mathbf{a}_{t-1}$ as if it reflects solely the private information of those

not typically studied because it is uninteresting under rationality—rational agents immediately learn the state after a single round of actions. With naive players, however, it implies that each generation grows confident in some state, but not necessarily the correct one—different generations may reach different conclusions across time.

[11] While this literal interpretation of the bias provides an intuition for the model, it does not match the motivating psychology. We believe people *neglect* that others use the informational content of public behavior, not that people actively believe others do not have *access* to this information. The definition of naivete formally proposed by Eyster and Rabin (2008), which they call "Best-Response Trailing Naive Inference" (BRTNI), assumes naive players best respond to the belief that all others are fully cursed in the sense of Eyster and Rabin's (2005) "cursed equilibrium". However, in social-learning settings, any of these assumptions yield identical results.

players acting in $t-1$.[12]

Our formalization of naivete rests on the concept of *autarkic action distributions*—the theoretical distribution of actions in state ω assuming players act solely on private signals and the prior.

Definition 1. *Conditional on ω, the* autarkic distribution $\mathbb{P}_\omega \in \Delta(\mathscr{A})$ *is the distribution of actions generated by autarkic play: $\mathbb{P}_\omega(m)$ is the probability that a player takes action A_m based solely on her realized signal* **s**.

In state ω, the observed actions taken by the first generation, regardless of whether players are rational or naive, will follow the autarkic distribution $\mathbb{P}_\omega$. A naive player, however, expects behavior to be autarkic in *every* period.

Definition 2. *An* inferentially naive *player infers from each action $x_{(n,t)}$ as if, conditional on state ω, it is an independent draw from $\mathbb{P}_\omega$.*

As we will show in Section 3, belief dynamics are characterized by the interplay between the autarkic distributions and the *aggregated-signal distributions*—the theoretical distribution of actions in state ω assuming each player observes an infinitely large collection of independent private signals drawn from $F(\cdot|\omega)$. In familiar settings where an infinite collection of signals perfectly reveals the state, the aggregated-signal distribution simply matches the distribution of actions observed among a generation certain (either rightly or wrongly) of state ω.

Definition 3. *The* aggregated-signal distribution $\mathbb{T}_\omega \in \Delta(\mathscr{A})$ *is the distribution of actions generated when all players put probability 1 on ω: $\mathbb{T}_\omega(m)$ is the probability that a player takes action A_m given certainty the state is ω.*[13]

Finally, we make two common assumptions about the underlying signal structure. First, autarkic distributions are distinct across states. Together with large generations, this implies that naive agents presume $\mathbf{a}_{t-1}$ perfectly identifies the state of the world. Second, autarkic distributions have

[12] In essence, naive players fail to realize that past behavior (in $t \geq 2$) already incorporates all useful private information. Eyster and Rabin (2014) refer to this as "redundancy neglect". In simple single-file settings, this directly generates over-counting of early signals.

[13] Both the aggregated-signal distributions and autarkic distributions are well-defined. Formally,

$$\mathbb{P}_\omega(m) = \int_{\mathscr{S}} \sum_{\theta \in \Theta} \lambda(\theta)\zeta(m|\theta,\mathbf{s},\pi_1)\,\mathrm{d}F(\mathbf{s}|\omega),$$

where $\zeta(m|\theta,\mathbf{s},\pi_1)$ is the probability that type θ takes A_m when relying solely on her private signal $\mathbf{s}$ and the prior π_1. Given a fixed tie-breaking rule, $\zeta(m|\theta,\mathbf{s},\pi_1)$, and hence $\mathbb{P}_\omega$, are well defined. Similarly, letting δ_ω denote a degenerate belief on state ω, $\mathbb{T}_\omega(m) = \sum_{\theta\in\Theta}\lambda(\theta)\zeta(m|\theta,\delta_\omega)$ where $\zeta(m|\theta,\delta_\omega)$ is the probability that type θ takes action A_m when certain the state is ω.

full support, which ensures that players never observe actions they thought were impossible.[14]

Assumption 1. *The collection of signal distributions* $\{F(\cdot|\omega)\}_{\omega\in\Omega}$ *is such that:*

1. (Uniqueness.) *For all* ω, $\omega' \in \Omega$, $\mathbb{P}_\omega \neq \mathbb{P}_{\omega'}$ *whenever* $\omega \neq \omega'$.

2. (Full Support.) *For all* $\omega \in \Omega$, $\mathbb{P}_\omega$ *has full support over* $\mathscr{A}$.

To motivate our analysis of long-run beliefs in Section 3, we first illustrate the implications of our model across the first few generations and discuss some of our assumptions. Suppose the state is ω^*. Since Generation 1 acts solely on private information, our "large-generations" assumption means that $\mathbf{a}_1 = \mathbb{P}_{\omega^*}$ in the limit as $N \to \infty$. Next, our uniqueness assumption implies that Generation 2, whether rational or naive, perfectly infers ω^* from $\mathbf{a}_1$. Hence, the distribution of actions in $t = 2$ is $\mathbf{a}_2 = T_{\omega^*}$. Rational players understand that $\mathbf{a}_2$ reflects the behavior of perfectly informed agents. Hence, outside of "confounded" environments where $\mathbb{T}_{\omega^*}$ is not unique, rational behavior converges by $t = 2$—all following generations, if rational, continue to play $\mathbb{T}_{\omega^*}$.

Naive followers may deviate from $\mathbb{T}_{\omega^*}$. Because they think predecessors act solely on private signals, they believe play in each round should resemble an autarkic distribution. Hence, Generation $t \geq 3$ thinks $\mathbf{a}_{t-1}$ reflects autarkic play. As such, players in Generation t will infer what distribution of signals Generation $t-1$ must have received in order to take actions $\mathbf{a}_{t-1}$ under autarkic play, and will then come to believe in the state most likely to generate those signals. However, since generations are in fact learning from one another, this reasoning is flawed. In actuality, Generation $t \geq 3$ observes the behavior of a Generation $t-1$ who is certain of some state—$\mathbf{a}_{t-1}$ matches an aggregated-signal distribution, not an autarkic one. This mistake is the crux of our model of naive inference: observers interpret *confident* behavior as if it were autarkic play. As such, the interplay between the collection of autarkic and aggregated-signal distributions determines what naive agents come to believe. Whether and when naive players will converge to optimal play $\mathbb{T}_{\omega^*}$—and what they converge to if not—is the basic premise of this paper. The next section explores this question generally.

The role of our "large-generations" convention in the arguments above warrants further discussion. As noted earlier, to simplify our analysis we in fact take a shortcut by assuming that the empirical distributions of actions realized in each round *exactly* match the underlying theoretical distributions (e.g., $\mathbf{a}_1 = \mathbb{P}_{\omega^*}$, $\mathbf{a}_2 = \mathbb{T}_{\omega^*}$, and so on) and that agents reach fully confident beliefs from these observations. Of course, short of the limit as $N \to \infty$, these identities will be approximately true with high probability. For instance, if $\mathbf{a}_1$ is close to $\mathbb{P}_{\omega^*}$, then $\mathbf{a}_2$ is far from $\mathbb{T}_{\omega^*}$

[14] We maintain these assumptions throughout the paper aside from two minor deviations. An application in Section 4 introduces aggregate uncertainty by assuming the payoff state has two dimensions, $(\omega, \xi) \in \Omega \times \Xi$, and that people receive private signals about the first dimension but not the second. Signals about ω are drawn from $F(\cdot|\omega)$ meeting our identifiability assumption, but beliefs about ξ are derived solely from the prior. Additionally, an example in Section 3.3 considers a signal structure that fails Part 2 of Assumption 1, but this is solely to simplify exposition.

only in the extremely unlikely event that a large fraction of players in Generation 2 receives strong signals that contradict the truth. We believe that assuming an "exact match" for distributions of actions and types even short of the limit is innocuous for all the analysis we focus on. For instance, it seems clear (but notationally cumbersome) that for any given finite number of generations, all dynamics of interest will happen with arbitrarily high probability for a sufficiently large N.[15] (We argue further in Section 7 that most of our results continue to hold even in environments where N is not large and each generation observes all those before it.)

2.3 Related Models

Before turning to long-run dynamics in Section 3, we briefly review how our paper compares to others that explore naive learning. Given that we adopt ER's model of naivete, our approach differs from their's primarily in the type of environments we consider. Other papers, however, offer entirely different approaches to modeling naive social learning.

ER explores a binary-state model with a continuum of actions, common preferences, and one player acting per round. Specifically, $\Omega = \{0,1\}$, $\mathscr{A} = [0,1]$, and $u(x|\omega) = -(x-\omega)^2$. With these preferences, a player optimally chooses $x \in [0,1]$ equal to her belief that $\omega = 1$, which implies that actions perfectly reveal an agent's posterior belief. Their main result is that, with positive probability, society grows confident in the wrong state. A naive player in period t treats the announced posterior of a player in $t-1$ as that player's independent signal, despite the fact that it also incorporates the signals of players in all periods $\tau < t-1$. As such, players vastly over-count early signals. If early signals are sufficiently misleading—which happens with positive probability—then players grow confident in the wrong state.[16]

Naivete leads society astray in ER's environment only when early signals are misleading. However, with our "large generations" assumption, early signals are never misleading. As such, when

[15] Our wording here is meant to be cautious about the order of the limits in the following sense. Infinite-horizon models along the lines of Kandori, Mailath, and Rob (1993) often converge to "selected" outcomes that happen 100% of the time as the likelihood of deviations become arbitrarily small, but do so despite having (in the very long run) an infinite number of arbitrarily long spells in other outcomes. The long-run frequency captures the fact that, in the limit, the selected equilibrium happens with infinitely greater duration than the others. We believe that our model would generate similar dynamics: assuming signals induce posteriors with full support on $[0,1]$ and fixing any arbitrarily large N, eventually there will be some generation t where an amazing number of players get extremely unlikely signals that cause enough of them to simultaneously deviate in a way that leads the following generation to reach a public belief different from the one held by Generation t. If this belief corresponds with another fixed point of the dynamics we develop below, society would stay at that point for quite a while before deviating back. As in Kandori, Mailath, and Rob (1993), society spends arbitrarily longer durations in our predicted equilibria. But unlike some of these models, our predicted equilibria are far most likely to happen early on. Hence, when N is large, our claims about what is far most likely to happen in all of a finite number of generations holds.

[16] Eyster and Rabin (2014) show that this intuition holds even when players over-count predecessors' signals by any arbitrarily small amount. These results stand in sharp contrast with the rational model, in which wrong herds are likely to occur only in settings where players remain relatively uncertain of the state. Rational models never lead to confident beliefs in a false state, and are thus not compelling models of society thinking it knows things it does not.

large populations act each round in ER's two-state setting, a naive society always converges to the truth. In contrast, we emphasize that in richer environments, naive observes may still converge to false beliefs even when early generations perfectly reveal the state.

Additionally, ER's two-state framework implies that if people learn the payoff of action $x = 1$, then they implicitly learn the payoff of all other actions. In contrast, we consider more natural settings where payoffs are independent across actions: knowledge that x is superior to x' does not reveal *by how much* x is preferred to x'. Such a distinction matters, for instance, when deciding how much to pay, or how long to wait, to obtain x over x'. In these settings, agents systematically overestimate the payoff of the herd action relative to those not chosen. In this sense, naive inference restricts which constellation of payoffs agents may come to believe.

Our model of naive inference—adopted from ER—is related to several other approaches motivated by a similar intuition that people neglect informational redundancies when learning from others. DeMarzo et al. (2003) propose a model of "persuasion bias" in which neighbors in a network communicate posterior beliefs. Building on DeGroot's (1974) model of consensus formation, they assume players form posteriors by taking the average of neighbors' beliefs as if they reflect independent signals with known precision. Since players neglect that stated beliefs already incorporate signals previously shared, they over-count early signals. Our model is also related to Level-k thinking (e.g., Crawford and Iriberri 2007). Naive agents act like Level-2 thinkers, as they best respond to the belief that others use only private information (Level-1).[17] Additionally, Bohren (2016) studies a variant of the canonical two-state model in which only a fraction $\alpha \in [0,1]$ of players can observe past actions, and people have wrong beliefs about α. Her model corresponds with naive inference when $\alpha = 1$ but all players think $\hat{\alpha} = 0$.

Empirically, very few experiments are designed with much power to identify naivete. ER discuss how findings from earlier experimental work, like Kübler and Weizsäcker (2004), suggest that people neglect informational redundancies in social learning. A recent experiment by Eyster, Rabin, and Weizsäcker (2015) find more direct evidence. They first tell each subject the difference in the number of heads and tails from 100 independent flips of a coin. Then, moving in sequence, each subject estimates the total difference in heads and tails across all predecessors—including herself—and announces this estimate.[18] A Bayesian Nash equilibrium strategy is to add one's own 100-trial sample to her immediate predecessors' estimate. However, they find a weak tendency towards redundancy neglect: subjects fail to fully understand that the most recent predecessor's

[17] The equivalence between naive inference as defined by ER and Level-2 thinking is not general, but happens to hold in typical formal models of social learning where players are concerned with others' irrationality only to the extent that it interferes with inference.

[18] In fact, the optimal behavior in this setting does not even require proper use of Bayes' Rule: the Bayesian Nash equilibrium corresponds to the unique Iterated-Weak-Dominance (IWD) outcome, and in this IWD strategy profile no players need to apply Bayes' Rule.

behavior incorporates the information of earlier predecessors. They also show severe redundancy neglect in a treatment where 4 independent players per round are asked to derive these sums.

3 Naive Long-Run Beliefs and Abandoned States

This section derives some general implications of naive inference on long-run learning. We first characterize the set of beliefs to which society may converge. This characterization reveals that there may exist states of the world that people always come to disbelieve, even if true. In these "abandoned states", society *unlearns*: although Generation 2 perfectly learns the true state, later generations become convinced of something false.

We then apply our characterization to a natural environment where actions have payoffs independent of one another. Naivete polarizes perceptions of payoffs: people grow confident that the best option achieves its highest possible payoff while all others attain their lowest. We also provide examples demonstrating two additional ways naivete constricts the set of beliefs on which society may settle. First, the set may be a singleton, implying society comes to the same conclusion no matter what is true. Second, the set may be empty, meaning beliefs continually cycle.

3.1 Characterization of Long-Run Beliefs

We now characterize the dynamics of naive public beliefs. If beliefs converge on some ω, then it must be that the behavior of people certain of ω most closely resembles the behavior we would see by privately informed agents if ω were true. The next lemma we develop shows that "closeness" is naturally measured by the cross entropy between the realized action distribution and the one predicted by autarkic play.

Beliefs transition from one generation to the next in a deterministic fashion. Our "large-generations" assumption implies that each Generation $t \geq 2$ grows confident in some state, which we denote by $\hat{\omega}_t$. Hence, we can express dynamics by a deterministic belief-transition function $\phi : \Omega \to \Omega$ that maps the belief of Generation t to that of Generation $t+1$—that is, $\hat{\omega}_{t+1} = \phi(\hat{\omega}_t)$.

Under naive inference, the transition function ϕ is characterized by the solution to a particular distance-minimization problem between the autarkic and aggregate-signal distributions. Specifically, suppose Generation t believes $\hat{\omega}_t$, so $\mathbf{a}_t = \mathbb{T}_{\hat{\omega}_t}$. Then Generation $t+1$ grows confident in the state $\hat{\omega}_{t+1}$ whose predicted autarkic distribution $\mathbb{P}_{\hat{\omega}_{t+1}}$ is closest to the observed distribution $\mathbb{T}_{\hat{\omega}_t}$ in terms of cross entropy.

Definition 4. *The* cross-entropy distance *between the observed distribution* $\mathbb{T} \in \Delta(\mathscr{A})$ *and the*

predicted distribution $\mathbb{P} \in \Delta(\mathscr{A})$ *is defined as*

$$H(\mathbb{T},\mathbb{P}) = -\sum_{m=1}^{M} \mathbb{T}(m)\log\mathbb{P}(m).^{19} \tag{2}$$

For simplicity, our next assumption rules out trivial complications that arise when an aggregated-signal distribution lies directly "in between" two distinct autarkic distributions.

Assumption 2. *For each* $\omega \in \Omega$, $H(\mathbb{T}_\omega,\mathbb{P}_{\omega'}) = H(\mathbb{T}_\omega,\mathbb{P}_{\omega''})$ *if and only if* $\omega' = \omega''$.

Assumption 2 ensures that our transition function ϕ is single valued, and, given that any setting where it fails is non-generic, we believe this assumption is mild. Finally, ϕ is characterized as follows.

Lemma 1. *The belief-transition function* ϕ *is defined by*

$$\phi(\hat{\omega}_t) = \arg\min_{\omega\in\Omega} H(\mathbb{T}_{\hat{\omega}_t},\mathbb{P}_\omega).$$

We are interested in what society eventually comes to believe—the limit of process $\langle\hat{\omega}_t\rangle$—as a function of what is true. The true state ω defines the initial condition of $\langle\hat{\omega}_t\rangle$: since Generation 1 does in fact act in autarky, Assumption 1 implies that Generation 2 correctly infers the true state. Thus, $\hat{\omega}_2 = \omega$.

Naivete has an effect starting in Generation 3: they neglect that Generation 2 learned from Generation 1, and hence think Generation 2 acted solely on private signals. As such, Generation 3 presumes $\mathbf{a}_2$ reflects autarkic play when in fact $\mathbf{a}_2 = \mathbb{T}_\omega$. As specified in Section 2, the third generation of naive agents will (1) infer what distribution of signals Generation 2 must have received in order to take actions $\mathbf{a}_2$ under autarkic play, and (2) come to believe in the state $\hat{\omega}_3 \in \Omega$ most likely to generate those signals. It is worth noting, however, that our model predicts that people will generally see a large number of predecessors behaving in a way they thought was extremely unlikely: there will not generically exist $\hat{\omega}$ such that $\mathbb{P}_{\hat{\omega}} = \mathbb{T}_\omega$. The updating rule ϕ treats any discrepancy between $\mathbf{a}_2$ and the predicted distribution as coming from sampling variation. While many models of errors (e.g., Barberis, Shleifer, and Vishny 1998 and Rabin 2002, among many others) share this feature, it raises the issue that the would-be error-makers might be able to deduce they are making a mistake. Yet there are reasons to believe this concern is less warranted than it

[19] We refer to the cross-entropy colloquially as a distance even though it is not symmetric and thus not a proper metric. The well-known Kullback-Leibler divergence (or "relative entropy") of $\mathbb{P}$ from $\mathbb{T}$ in terms of cross entropy is simply $H(\mathbb{T},\mathbb{P}) - H(\mathbb{T},\mathbb{T})$. These measures are common in information theory (see, for example, Cover and Thomas 1991) and have been applied in recent work in economics on learning with incorrect or uncertain models. Examples include Rabin (2002), Schwartzstein (2014), Acemoglu, Chernozhukov, and Yildiz (2016), and Esponda and Pouzo (2016). An older literature in statistics on misspecified models, starting with Berk (1966), follows a similar approach.

may first appear, and we illustrate in Section 6 cases where our model is "explicable" in the sense that naive agents find what they see no more surprising than rational agents would.

It is noteworthy, albeit immediate, that naive inference leads to "unlearning" across generations. Although Generation 2 necessarily learns the state, this correct belief is maintained by Generation 3 if and only if $\omega = \phi(\omega)$. When $\omega \neq \phi(\omega)$, society "unlearns" ω between Generations 2 and 3 and becomes convinced of something false. This in turn implies that society fails to converge on the truth: even if public beliefs return to ω at some later date, they will again move away the following period.

Turning to long-run beliefs, the limiting behavior of $\langle \hat{\omega}_t \rangle$ falls into one of two broad categories—"stationary" or "cyclic". If $\langle \hat{\omega}_t \rangle$ converges to a single absorbing state, then society settles on a stationary long-run belief. We let $\Omega^* \subseteq \Omega$ denote the set of all states on which stationary beliefs may settle.[20] Alternatively, $\langle \hat{\omega}_t \rangle$ may reach an absorbing *set*: beliefs forever (deterministically) cycle over the states in this set. To account for this weaker notion of convergence, we let $\Omega^{**} \supseteq \Omega^*$ consist of all states that society may continually assign positive probability. We refer to any state outside of Ω^{**} as *abandoned*. Although most of our applications will involve stationary convergence, for the sake of completeness we briefly describe with some generality the full spectrum of possible limit beliefs.

We focus on two questions: First, as a function of the true state of the world, what will society come to believe? And second, might there exist "abandoned states" that are *always* assigned zero probability in the long run, no matter what is true? That is, might Ω^* or Ω^{**} be strict subsets of Ω? Answering these questions requires identifying each of the unique absorbing sets of the belief process, denoted by $\widehat{\Omega}_1, \ldots, \widehat{\Omega}_L$, where $L \geq 1$ depends on the environment. These are easily derived from the transition function ϕ.[21] Hence, Ω^{**}—the set of states that continually receive positive weight in the long run—is simply the union of all absorbing sets: $\Omega^{**} = \cup_{l=1}^{L} \widehat{\Omega}_l$

We first consider what people come to believe as a function of the true state. For each absorbing set $\widehat{\Omega}_l$, there exists a distinct "basin of attraction", $\Omega_l \supseteq \widehat{\Omega}_l$, such that $\langle \hat{\omega}_t \rangle$ eventually settles in $\widehat{\Omega}_l$ if and only if $\omega \in \Omega_l$.[22] Given the collection of absorbing sets and their associated basins of attraction, we can define $\beta : \Omega \to \Delta(\Omega)$ as the long-run average of the public belief as a function of the true state. The k^{th} entry of $\beta(\omega)$ is the (long-run) faction of generations who are confident

[20] If the setting is such that $\mathbb{T}_\omega$ is distinct for each $\omega \in \Omega$—which is the case in most commonly studied social-learning environments—then rational learning trivially implies $\Omega^* = \Omega$.

[21] Formally, an absorbing set $\widehat{\Omega}_l$ is defined as follows: $\omega \in \widehat{\Omega}_l$ if and only if $\hat{\omega}_t = \omega$ implies that for all $\tau > t$ both $\hat{\omega}_\tau \in \widehat{\Omega}_l$ and $\hat{\omega}_\tau = \omega$ infinitely often. The singleton sets—those that correspond to stationary convergence—are simply the fixed points of ϕ. The remaining sets consist of the fixed points of ϕ^k for $k = 2, ..., K$, where $K = |\Omega|$. To see this, $\tilde{\omega} = \phi^k(\tilde{\omega})$ implies that if $\hat{\omega}_t = \tilde{\omega}$, then $\hat{\omega}_{t+k} = \tilde{\omega}$ as well. Hence, the belief process continually returns to $\tilde{\omega}$ every k^{th} period. The absorbing set containing $\tilde{\omega}$ is thus $\widehat{\Omega} = \{\tilde{\omega}, \phi(\tilde{\omega}), \ldots, \phi^{k-1}(\tilde{\omega})\}$.

[22] It is straightforward to construct the basin of attraction for a given absorbing set. Assume the true state is ω. Let $\Phi(\omega)$ be the state Generation t believes is true the first time $\hat{\omega}_t \in \Omega^{**}$; that is, $\Phi(\omega) \equiv \min_k \phi^k(\omega)$ subject to $\phi^k(\omega) \in \Omega^{**}$. Then $\Omega_l = \{\omega \in \Omega \mid \Phi(\omega) \in \widehat{\Omega}_l\}$.

in ω_k when the true state is ω. Hence, for any $\omega \in \Omega_L$, $\beta_k(\omega) = 1/|\widehat{\Omega}_l|$ if $\omega_k \in \widehat{\Omega}_l$ and $\beta_k(\omega) = 0$ otherwise.

For ease of reference, we summarize these observations in the following obvious lemma, which describes necessary conditions for a state to be in the support of long-run beliefs.

Lemma 2. *For all $\omega \in \Omega$, $\beta_k(\omega) > 0$ if and only if $\omega_k \in \Omega^{**}$. Furthermore, if $\langle \hat{\omega}_t \rangle$ converges to a stationary limit $\hat{\omega}$, then $\hat{\omega} \in \Omega^* = \{\omega \in \Omega \mid \phi(\omega) = \omega\}$.*

The second part of Lemma 2 highlights how naivete restricts the set of states on which public beliefs may settle. The fixed points of ϕ, which comprise the set Ω^*, are the only states on which beliefs may rest. More intuitively, the set Ω^* consists solely of those states ω such that the behavior of people certain of ω most closely resembles the behavior we would see by privately informed agents if ω were true. Furthermore, whenever $\Omega^{**} \neq \Omega$, there exist abandoned states that people continually disbelieve no matter what is true. Thus, in contrast with ER's canonical environment where society mislearns only when early signals are misleading, some states in our settings lead to mislearning with probability *one*.

Lemma 2 is robust to many generalizations of our environment. Our assumptions that players observe only the previous generation and that generations are arbitrarily large—which together imply deterministic dynamics—are imposed solely to make the logic transparent. The result holds without these assumptions so long as the number of players observed by Generation t grows large as $t \to \infty$. So, even when people act in single file, the only states on which beliefs can settle are the fixed points of the same transition function ϕ as specified in Lemma 1.[23] Section 7 discusses these and other robustness properties in greater detail.

The remainder of the paper studies several implications of Lemma 2. Although little can be said about Ω^{**} or Ω^*—and hence which states are abandoned—without specifying autarkic distributions, we note two immediate but general observations before turning to specific applications. First, in no environment are people forever wrong in every state of the world.

Proposition 1. *There exists some state $\omega_k \in \Omega$ such that $\beta_k(\omega_k) > 0$.*

While it is possible that some states of the world lead to perpetually false beliefs, this cannot happen in all of them. Still, as our applications will show, there are natural settings in which a *majority* of states will lead a naive society to false conclusions.

Second, if agents have common preferences and there are more states than actions, then there

[23] Loosely, if beliefs converge to certainty in $\hat{\omega}$, then the long-run distribution of behavior converges to $\mathbb{T}_{\hat{\omega}}$. Hence, in order for agents to remain confident in $\hat{\omega}$, it must be that $\hat{\omega} = \arg\min_{\omega \in \Omega} H(\mathbb{T}_{\hat{\omega}}, \mathbb{P}_\omega)$. Essentially, players compare autarkic distributions to the *long-run* action distribution instead of the per-period action distribution. For stationary convergence, the relationship between the autarkic and long-run distribution must satisfy exactly the same "fixed-point" condition as the one in Lemma 2.

necessarily exist states that society will continually disbelieve even when they are true. That is, Ω^{**} is a strict subset of Ω.

Proposition 2. *Suppose players have common preferences. If* $|\Omega| > |\mathscr{A}|$*, then* $\Omega^{**} \neq \Omega$.

Since Generation 2 identifies the commonly-preferred action, all players in $t = 2$ herd on this action. Furthermore, because there is a unique state that best predicts such a herd in autarky, there are at most $M = |\mathscr{A}|$ states that society can come to believe: one for each of the M possible herds. If $|\Omega| > M$, then there exist states that naive agents always assume false after observing a herd. This result is obscured by ER, who assume $|\Omega| \leq |\mathscr{A}|$. But $|\Omega| > |\mathscr{A}|$ is true in many environments. In particular, it is true whenever the unknown payoffs of each action are independent of one another. The following section (Section 3.2) provides a detailed application of this result.

3.2 Extremism

We now present a simple yet stark implication of naive learning in settings where players with common preferences choose among options with independent payoffs. Namely, people grow certain that one option is as good as possible and that all other options are as bad as possible. Under naive inference, people think a herd on option A results from (infinitely) many independent signals indicating A is better than all of its alternatives. Within a natural class of signal structures, this misinterpretation of the herd suggests a state where the payoff difference between A and all other actions is most extreme.

Suppose each option $A_m \in \{A_1, ..., A_M\}$ has unknown quality q^m independent of the others. For instance, imagine diners learning about the quality of unrelated restaurants around town, or investors learning about the returns to assets in different sectors. The payoff-relevant state is the "quality vector" $\omega = (q^1, ..., q^M)$. Players have common preferences represented by $u(A_m|\omega) = q^m$, so the payoff of action m depends only on the quality of m. Each q^m is drawn from a finite set Q^m, $|Q^m| \geq 2$, according to a known prior π^m with full support on Q^m, and for all $j \neq m$, q^j and q^m are independent. Thus, any quality profile $\omega \in \times_{m=1}^{M} Q^m$ is a feasible state of the world.

We focus on a natural class of signal structures that satisfy the familiar Monotone Likelihood Ratio Property (MLRP). Each player receives an independent signal $s^m \in \mathbb{R}$ about each q^m—news about q^m provides no information about q^j, $j \neq m$. Conditional on q^m, let $F^m(\cdot|q^m)$ and $f^m(\cdot|q^m)$ denote the c.d.f. and associated density (or mass) function of s^m, respectively. Let S^m denote the support of s^m, which is identical for all values of $q^m \in Q^m$.[24]

Definition 5. *F^m satisfies the strict* Monotone Likelihood Ratio Property (MLRP) *if for every $q > q'$, $f^m(s|q)/f^m(s|q')$ is strictly increasing in s.*

[24]Signals also satisfy the maintained assumptions of Section 2.1: (1) for each m, s^m is conditionally independent and identically distributed across all players, and (2) when $N \to \infty$, the autarkic distribution reveals ω.

MLRP means that higher signals unambiguously indicate higher expected quality. In this setting, MLRP naturally implies that the share of players who choose A_m in autarky strictly increases as its quality q^m increases. This suggests that naive social learners will conflate high demand with high quality. Indeed, our next proposition shows that they will come to believe in an "extreme state":

Definition 6. Extreme state *m, denoted by* ω_e^m*, is the state in which* $q^m = \max Q^m$ *and* $q^j = \min Q^j$ *for all* $j \neq m$*.*

Proposition 3. *Suppose* $\arg\max_j(q^1,\ldots,q^M)$ *is unique and equal to m. If signals are independent and satisfy MLRP, then beliefs converge to extreme state* ω_e^m*:* $\hat{\omega}_t = \omega_e^m$ *for all* $t > 2$*. Hence,* Ω^{**} *and* Ω^* *equal the set of extreme states.*

To demonstrate the logic, suppose $q^1 = \max(q^1,...,q^M)$. Because Generation 2 infers that A_1 is optimal, they unanimously choose A_1. Generation 3 comes to believe in the state most likely to induce a herd on A_1 under autarkic play. Intuitively, this state must maximize the chance of good news about A_1, but minimize the chance of good news about any other option. Because of MLRP, this happens in state ω_e^1. As such, Generation 3 infers $\hat{\omega}_3 = \omega_e^1$ and again herds on A_1, which implies that ω_e^1 is an absorbing state: $\hat{\omega}_t = \omega_e^1$ for all $t > 2$.[25] In essence, naive observers mistake herds caused by social learning as evidence that the solely chosen option has "maximally" superior quality.

This "extremism" result implies that naive learners exaggerate the quality difference between options. Imagine that restaurant A is only slightly better than its neighbor, B. When droves of patrons dine at A while B sits empty, the naive onlooker infers not only that A is better than B, but that it is *much* better. From the naive viewpoint, A's market share is strictly increasing in its quality advantage over B—a small advantage leads to roughly equal market shares, whereas a big one means A gets most of the business. This naive logic neglects the fact that, because early diners correctly infer that A is better than B, all subsequent patrons will choose A no matter how small is its quality advantage.[26]

Proposition 3 clearly demonstrates how naivete restricts the hypotheses society may come to believe. No matter which of the $\prod_{m=1}^{M} |Q^m|$ possible states is realized, beliefs eventually settle on one of only M extreme states.[27] This setting also exemplifies the prediction of Proposition 2: since there are more states than actions, some states are surely abandoned.

[25] In terms of the general treatment in Section 3.1, this environment has M basins of attraction, $\Omega_1,\ldots,\Omega_M$, where each Ω_m contains all states that lead to a long-run belief in extreme state ω_e^m. Hence $\Omega_m = \{\omega \in \Omega | q^m = \max(q^1,\ldots,q^M)\}$. Given Assumption 1, we implicitly restrict attention to cases where $q^m = \max(q^1,\ldots,q^M)$ is unique.

[26] Of course, this is a highly simplified model of markets. In reality, we do not expect *all* consumers to patronize the better business—both pricing and various frictions, like search and queuing costs, would prevent this. The point of this example is to demonstrate a more general, qualitative prediction: under social learning, the better business captures *more* of the market than if consumers were to make decisions in autarky. Since naive learners neglect this social determinant of market shares, they misattribute larger-than-expected market shares to a firm's quality.

[27] The canonical setting assumed in Eyster and Rabin (2010) precludes exaggerated perceptions of quality differ-

The force generating extreme beliefs is quite robust across environments. Extremism arises in any environment where the autarkic frequency of action m is increasing in q^m and decreasing in q^j for all $j \neq m$. While MLRP implies this property, it is stronger than necessary. Furthermore, beliefs will converge on an extreme state even if people act in single file—large generations only guarantees convergence to the *optimal* extreme state. That said, our result is sensitive to our assumption that, conditional on ω, signals about each option are independent of one another. Namely, extremism need not occur if signals are positively correlated: if people are more likely to receive good news about A the better is B, then a herd on A may indicate that *both* options have high quality.

While in this particular setting naive learning has no welfare consequences—people herd at the best restaurant despite exaggerating its relative quality—"extreme beliefs" directly harm welfare in many of our remaining applications. For instance, Section 4 shows how extreme beliefs in an investment setting lead to costly under-diversifcation. And in settings with queuing costs, extreme beliefs may generate inefficiently high congestion: since people exaggerate the quality difference between the best option and all others, they are less willing—relative to rational players—to switch to the next-best option. In the restaurant example above, diners would be willing to wait in long lines or pay relatively high prices to patronize restaurant A despite B being nearly as good.

3.3 Additional Examples

While the setting above drives public beliefs toward an extreme state, the examples here demonstrate two other interesting forms of restrictive mislearning that arise in somewhat different environments. The first provides a setting where Ω^* is a singleton, meaning society settles on the same belief no matter what is true. The second illustrates non-convergence: Ω^* is empty, and thus beliefs continually cycle.

3.3.1 State-Independent Beliefs

Imagine farmers deciding whether to adopt a new hybrid seed (A) or stick with a well-known variety (B). Option B yields a payoff normalized to zero for all farmers. The payoff from A, however, is sensitive to a farmer's soil type and is positive only if it matches well with one's plot.[28] Suppose there are two types of soil, high salinity ($\theta = H$) and low ($\theta = L$). Each farmer knows both her own type and that fraction $\lambda > \frac{1}{2}$ of farmers are high types. Additionally, there are two equally likely states of nature: seed A is compatible with high salinity ($\omega = H$) or it is compatible

ences because the only states they consider are "extreme" to begin with. Only when we first allow non-extreme states to arise with positive probability do we see how naive inference restricts long-run beliefs.

[28]This example is inspired by Munshi (2003) and Foster and Rosensweig (1995), who study social learning among Indian farmers trying to deduce the optimal inputs for new "high-yield" strains of rice and wheat. Munshi (2003) notes that rice is quite sensitive to soil characteristics, but wheat is not.

with low salinity ($\omega = L$). A farmer with compatible soil earns $v > 0$ by planting A, but one with incompatible soil earns $-v$. Thus, option A is optimal for Farmer (n,t) if and only if $\theta_{(n,t)} = \omega$.

We consider a signal structure where only a fraction of individuals have private information. This will imply that, no matter the state, adoption of the new seed increases as social learning takes place. Concretely, suppose a known fraction $\psi \in (0,1)$ of farmers have private information about ω which, say, comes from past experience with seed A; the remaining farmers have no private information. Each informed farmer receives an i.i.d. signal $s \in \{H,L\}$ with mass function $f(s = \omega|\omega) = \rho > \frac{1}{2}$—her signal matches the true state with chance $\rho > \frac{1}{2}$. Finally, we imagine a string of villages $t = 1,2,\ldots$ such that village t observes what village $t-1$ has planted.

Determining what farmers come to believe requires a comparison of $\mathbb{P}_\omega$ and $\mathbb{T}_\omega$ across states. From the setup above, the autarkic distributions are $\mathbb{P}_L(A) = \psi[(1-\lambda)\rho + \lambda(1-\rho)]$ and $\mathbb{P}_H(A) = \psi[\lambda\rho + (1-\lambda)(1-\rho)]$. Among a generation certain of ω, all those with $\theta_{(n,t)} = \omega$ choose A; this implies $\mathbb{T}_L(A) = 1-\lambda$ and $\mathbb{T}_H(A) = \lambda$. Hence, if sufficiently few have private information—that is, if ψ is not too large—then social learning always leads to higher adoption rates relative to autarky: even those who are initially uninformed might adopt based on information gleaned from their neighbors.[29] Naive observers, however, misattribute this learning-based increase in adoption to preferences. When a large share adopts, they conclude the new seed must be optimal for the majority type. Essentially, people may mistake "confident" behavior in the low state for autarkic play in the high state.

To illustrate, suppose $\lambda = 0.7$, $\rho = 0.8$, $\psi = 0.5$, and $\omega = L$—only the less-common low types should adopt. In the first round, the adoption rate is $\mathbb{P}_L(A) = 19\%$, which reveals $\omega = L$ to Generation 2. Thus in round 2, all of those with $\theta_{(n,t)} = L$ choose A, yielding adoption rate $\mathbb{T}_L(A) = 30\%$. A naive Generation 3, however, expects to see either the autarkic rate $\mathbb{P}_L(A) = 19\%$ in state L or $\mathbb{P}_H(A) = 31\%$ in state H. They come to believe in whichever state is most likely to yield 30% as a result of sampling variation. Since 30% is "closer" (in terms of cross entropy) to 31% than it is to 19%, Generation 3 wrongly becomes convinced that $\omega = H$.[30]

In fact, $\hat{\omega}_3 = H$ whenever ψ is not too large relative to λ and ρ. In such cases, all high types—the majority of farmers—adopt A in round 3. Since this rate is higher than predicted in either state, all subsequent farmers will continue to believe $\omega = H$. Mislearning results from people intuitively (but wrongly) reasoning that high adoption rates indicate that the new technology is best for the majority type. The following proposition summarizes these results.

Proposition 4. *Given the setup above:*

[29] The relevant condition on ψ is $\psi < (1-\lambda)/[(1-\lambda)\rho + \lambda(1-\rho)]$. This threshold monotonically increases to 1 as the precision of signals, ρ, increases from .5 to 1.

[30] This "distance" calculation follows from Equation 2: $H(\mathbb{T}_L, \mathbb{P}_L) = -0.3\log(0.19) - 0.7\log(0.81) \approx 0.6457 > 0.6111 \approx H(\mathbb{T}_L, \mathbb{P}_H) = -0.3\log(0.31) - 0.7\log(0.69)$.

1. *If $\omega = H$, then $\hat{\omega}_t = H$ for all $t \geq 2$.*

2. *There exists a threshold $\bar{\psi}(\lambda, \rho) \in (0,1)$ continuously decreasing in both λ and ρ such that if $\omega = L$, then $\psi < \bar{\psi}(\lambda, \rho)$ implies $\hat{\omega}_t = H$ for all $t > 2$.*

Hence, if $\psi < \bar{\psi}(\lambda, \rho)$, then $\Omega^{**} = \Omega^* = \{H\}$.

3.3.2 Non-Convergence

Our next example shows that Ω^* may be empty: social beliefs never settle on a fixed belief and instead perpetually cycle. While the example is admittedly contrived, it clearly demonstrates the possibility of non-convergence.

On a fixed day each week, a fresh catch of fish arrives at the local market. Shoppers must choose a day to visit the market, and thus seek to learn the delivery day $\omega \in \{S, M, Tu, W, Th, F, Sa\}$. Conditional on ω, shoppers earn a payoff $u(x|\omega)$ from going to the market on day x. Letting $\omega + 1$ denote the day after ω and so forth, we assume $u(x = \omega|\omega) = 1$, $u(x = \omega + 1|\omega) = \frac{9}{10}$, and $u(x|\omega) = 0$ for all $x \notin \{\omega, \omega + 1\}$. That is, fish are best on the delivery day, slightly worse the next day, and either sold out or spoiled thereafter. Additionally, conditional on state ω, each shopper receives a signal $s \in \{\omega - 1, \omega\}$ with mass function $f(s = \omega|\omega) = \frac{2}{3}$. Given this setup, an uncertain autarkic shopper prefers to risk arriving a day late rather than arriving a day early. More precisely, following signal s, a shopper prefers to go the day *after* her signal suggests: $\mathbb{E}[u(s|\omega)|s] = \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot 0 = \frac{2}{3}$ and $\mathbb{E}[u(s+1|\omega)|s] = \frac{2}{3} \cdot \frac{9}{10} + \frac{1}{3} \cdot 1 = \frac{14}{15}$. This structure generates autarkic distributions with $\mathbb{P}_\omega(\omega) = \frac{1}{3}$, $\mathbb{P}_\omega(\omega + 1) = \frac{2}{3}$, and $\mathbb{P}_\omega(x) = 0$ otherwise. Hence, in autarky, shoppers expect to see crowds the day *after* delivery—not the day of.

To see how naive social learning evolves, suppose $\omega = Th$. Initially, most shoppers arrive on Friday, and observers correctly deduce the delivery day. The following week, the crowd arrives on Thursday: since there is no longer uncertainty, it is optimal to go on the precise day of delivery. However, when people interpret the Thursday crowd as if it were based solely on private information, they think the fish must have arrived on *Wednesday*. So, in the third week, the crowd shows up on Wednesday. Rolling forward, it is clear that for all $t \geq 2$, $\hat{\omega}_{t+1}$ is the day before $\hat{\omega}_t$. Beliefs about the delivery day continually cycle through the days of the week.[31]

While contrived, the punchline of this example does not depend on our assumption that each generation is large and observes only the actions of the previous generation. Even if people were to act in single file and observe all past predecessors, social beliefs would still fail to converge to a stationary point belief. Intuitively, a herd on one action always suggests a state in which it is optimal to take an action *different* from the herd.

[31] Notice that $\Omega^{**} = \Omega$ and that there is a single basin of attraction, Ω. Like the example in Section 3.3.1, the long-run distribution of beliefs is the same no matter the true state.

4 Extremism in Investment Decisions

This section shows how naivete generates extreme beliefs about the returns to investments, which consequently lead to severe under-diversification. When naive investors use predecessors' allocations to predict the returns on two risky investments, they succumb to two forms of inefficiency: (1) even when it is optimal to diversify, they inevitably allocate all resources to a single investment, and (2) when the true returns differ sufficiently from expectations, they allocate all resources to the worse investment. While the logic parallels the extremism results in Section 3.2, the investment setting studied here highlights additional implications of naivete. First, extreme perceptions lead to costly mistakes. Second, the dynamics exhibit forms of over-extrapolation and momentum observed in asset markets, where perceptions of an asset's value continually grow more extreme over time.

Suppose there are two risky investments that pay off in terms of a consumption good in some final period T.[32] We refer to one investment as "safer"—its expected payoff is known, and is equal to one unit of the consumption good. The other, which we call "riskier", has an unknown expected payoff equal to $1+\omega$ units; we refer to the unknown value ω as the "fundamental". Investors learn about the fundamental from private signals and predecessors' allocations.[33] Additionally, both assets are subject to aggregate uncertainty: payoffs are distorted by i.i.d. mean-zero random shocks about which there is no available information. We include these aggregate shocks to model scenarios where rational risk-averse investors diversify even when ω is perfectly known. Summarizing, the payoffs of the safer and riskier investments are $d_t^s = 1+\eta_t^s$ and $d_t^r = 1+\omega+\eta_t^r$, respectively, where the random shocks η_t^m are i.i.d. normal with mean zero and precision ρ_η (i.e., variance $1/\rho_\eta$) for both $m = r,s$. Finally, we assume for simplicity that investors' prior over the fundamental ω is normally distributed with mean $\overline{\omega}$ and precision ρ_ω and that each Investor (n,t) receives a signal $s_{(n,t)} = \omega + \varepsilon_{(n,t)}$, where $\varepsilon_{(n,t)}$ is i.i.d. normal with mean zero and precision ρ_ε. Investor (n,t)'s information set $I_{(n,t)} \equiv \{s_{(n,t)}, \bar{x}_{t-1}\}$ consists of her private signal and the aggregate share of wealth invested in the riskier asset during the preceding period, denoted $\bar{x}_{t-1}$.[34]

[32] While we focus on the limit case where T is arbitrarily large, the dynamics we describe are identical for finite T. We consider assets that pay off long after the initial investment decision to ensure that the current generation does not observe the outcomes of the previous generation's choices. Reasonable examples include investments in real estate or education. The model alternatively applies to cases where each investor's outcome is realized immediately but is privately observed.

[33] To keep the model simple and squarely within our framework, we ignore asset prices. Hence, rather than stocks, imagine people allocating resources like time or effort across various risky projects. We discuss in the conclusion how our results naturally extend to a model with prices.

[34] Our qualitative results do not depend on the specific assumptions made here. First, to be consistent with our general model in Section 2, we maintain that each player participates in the market for a single period and observes only the behavior of the preceding generation. However, we believe our results extend to a market where the same participants repeatedly interact. Additionally, we assume normally distributed states and signals simply to reach a closed-form solution for the optimal investment.

Each Investor (n,t) has initial resources $W_0 \in \mathbb{R}$ and allocates a fraction $x_{(n,t)} \in [0,1]$ to the riskier asset. For simplicity, we assume investors have exponential utility $u(W) = -\exp(-\alpha W)$, where α is an individual's coefficient of absolute risk aversion. It is well-known that, with these preferences and normally distributed wealth, Investor (n,t) chooses

$$x_{(n,t)} = \arg\max_x \widehat{\mathbb{E}}[W|x, I_{(n,t)}] - \frac{1}{2}\alpha \mathrm{Var}[W|x, I_{(n,t)}],$$

which implies

$$x_{(n,t)} = \frac{1}{2 + \rho_\eta \mathrm{Var}[\omega|I_{(n,t)}]} \left\{ 1 + \frac{\rho_\eta}{\alpha W_0} \widehat{\mathbb{E}}[\omega|I_{(n,t)}] \right\}. \tag{3}$$

Importantly, the expectation and variance above denote those *perceived* by naive Investor (n,t) according to her incorrect autarkic model. Since we assume a large market, the aggregate share allocated to the riskier asset in period t is $\bar{x}_t = \mathbb{E}[x_{(n,t)}|\omega]$, where $\mathbb{E}[\cdot|\omega]$ is the expectation with respect to the true model.

We now derive the dynamic process of beliefs and allocations among naive investors. In the first period, investors act solely on private signals. Because these initial investors have no opportunity to mislearn from previous investments, it follows from Equation 3 that the first-period allocation satisfies a linear "autarkic demand function" that fully reveals the true fundamental:

Lemma 3. *Given the fundamental ω, the aggregate share invested in the riskier asset in $t = 1$ is $\bar{x}_1 = \nu_0 + \nu_1 \omega$, where the coefficients $\nu_0, \nu_1 \in \mathbb{R}$ are strictly positive and depend solely on the commonly known parameters $(\overline{\omega}, \rho_\omega, \rho_\varepsilon, \rho_\eta, \alpha)$.*

The first-period allocation efficiently aggregates all private signals, and, because the market is large, $\bar{x}_1$ perfectly reveals ω. Since Generation 2 correctly thinks that Generation 1 acted in autarky, each investor n in Generation 2 correctly inverts $\bar{x}_1$ to learn $\widehat{\mathbb{E}}[\omega|I_{(n,2)}] = \omega$ with certainty; i.e., $\mathrm{Var}[\omega|I_{(n,2)}] = 0$. Demand in $t = 2$ properly adjusts to these new perceptions.

As usual, the inferential error beings among Generation 3. Since they too think demand in $t = 2$ is based solely on private information, they presume $\bar{x}_2$ satisfies the autarkic demand function (Lemma 3). Inverting this relation, they infer $\hat{\omega}_3 = (\bar{x}_2 - \nu_0)/\nu_1$.[35] And because each subsequent Generation t makes this same mistake, each grows certain that

$$\hat{\omega}_t = (\bar{x}_{t-1} - \nu_0)/\nu_1. \tag{4}$$

[35] Unlike previous applications, the fact that we have assumed a continuum of states implies that, for every possible observation $\bar{x}_{t-1}$, there exists a unique false belief $\hat{\omega}_t$ that perfectly rationalizes $\bar{x}_{t-1}$ among naive investors. Hence, the updating rule of Lemma 1 trivially reduces to the simple inversion described here.

From Equation 3, the aggregate allocation to the riskier investment in period t is thus

$$\bar{x}_t = x_{(n,t)} = \frac{1}{2}\left(1 + \frac{\rho_\eta}{\alpha W_0}\hat{\omega}_t\right), \tag{5}$$

so long as this value lies in $[0,1]$. Taken together, Equations 4 and 5 recursively define the allocation process, $\langle \bar{x}_t \rangle$.

As a first step in describing these investment dynamics, let us contrast naive and rational allocations. Rational investors allocate a constant amount x^r in all $t > 1$: they realize that predecessors efficiently use all available information, and thus perfectly infer ω from the previous generation's behavior. They rationally allocate more than 50% to the riskier asset if and only if they learn $\omega > 0$. Furthermore, because of aggregate uncertainty, rational investors diversify—$0 < x^r < 1$—so long as ω is not large in absolute value. Naive allocations, however, are not static. Investors in t form beliefs as if Generation $t-1$ used new, independent information in forming demand $\bar{x}_{t-1}$. As such, naive investors always think that past demand reflects information not yet accounted for.

There are two ways in which naivete leads investors astray over time. First, they inevitably invest all resources in a single asset, which implies under-diversification. Second, if the true value of the fundamental ω is sufficiently far from initial expectations, they allocate all resources to the worse asset—that is, they choose the one that, in reality, provides lower expected utility.

These results stem from the fact that naive inference polarizes investors' perceptions of the payoff difference between the two assets. As t grows large, beliefs about ω diverge to positive or negative infinity. As such, $\langle \bar{x}_t \rangle$ converges to either 1 or 0.[36]

Proposition 5. *Given the fundamental ω:*

1. *Perceptions of the riskier asset's return $\langle \hat{\omega}_t \rangle$ and allocations to the riskier asset $\langle \bar{x}_t \rangle$ are strictly increasing in t if $\omega > \omega^*$, where*

$$\omega^* \equiv \frac{1}{2\rho_\omega + \rho_\eta}\left(2\rho_\omega \overline{\omega} - \alpha W_0\right).$$

   If $\omega < \omega^*$, then $\langle \hat{\omega}_t \rangle$ and $\langle \bar{x}_t \rangle$ are strictly decreasing in t.

2. *$\langle \bar{x}_t \rangle$ monotonically converges to 1 or 0. Hence, players eventually invest in a single asset.*

Whether beliefs about ω increase or decrease over time depends on whether Generation 2—who learns the true value of ω—invests more or less in the riskier asset than Generation 1. This initial

[36] Since we have assumed a prior on ω with full support on $\mathbb{R}$, $\hat{\omega}_t$ converges to $\pm\infty$. Of course, this prediction is unrealistic—the presence of some rational investors in the market would moderate this result. Instead, we think the more important feature of this result is the qualitative prediction that perceptions tend to change monotonically over time. The extreme outcomes of "infinite" perceptions and zero diversification simply help make the logic clear.

revision in allocations creates momentum that propagates through all future periods. To illustrate, suppose the riskier asset's allocation increases from period 1 to 2 (i.e., $\bar{x}_2 > \bar{x}_1$). Since Generation 3 treats the revised split as if it reflects autarkic demand, Generation 3 must infer a higher value of ω than Generation 2. As such, the riskier allocation increases yet again. This logic plays out across all periods: each new generation observes a larger "autarkic demand" than the last, which leads them to allocate even more to the riskier asset. Likewise, whenever the initial revision in allocations is downward (i.e., $\bar{x}_2 < \bar{x}_1$), demand and perceived payoffs decrease over time.

The inferential error is driven by investors continually using past demand as if it reflects new information.[37] Investors neglect that observed demand already incorporates all information in the economy, and hence attribute any changes to new private information. When the current generation incorporates this "new" information, the allocation moves yet again in the same direction as the initial (rational) adjustment. Hence, naivete predicts momentum even when no new information is realized, offering a plausible explanation for the sort of unwarranted swings in group beliefs that appear to be a hallmark of financial markets. This qualitative prediction accords with the empirical findings of Glaeser and Nathanson (2015), who suggest that momentum in the housing market derives, in part, from naive inference based on past market prices. Indeed, an extension of our model that incorporates pricing predicts momentum in prices, which in turn leads to price bubbles.

The direction of momentum, which is triggered by the initial revision in allocations, depends on how the true value of ω compares with expectations. The critical value ω^* from Proposition 5 such that $\bar{x}_2 > \bar{x}_1 \Leftrightarrow \omega > \omega^*$ is somewhat below the prior mean $\overline{\omega}$. Intuitively, allocations initially increase when investors learn that ω exceeds exceptions. Due to risk aversion, there is also a range of values $\omega \in (\omega^*, \overline{\omega})$ for which allocations initially increase even though ω falls short of expectations: the decrease in uncertainty upon learning ω offsets the low return.[38]

Part 2 of Proposition 5 shows that aggregate allocations increase or decrease until investors devote either all or no wealth to the riskier asset. Hence, naive risk-averse investors are worse off relative to their rational counterparts whenever it is optimal to diversify. With exponential utility, this happens whenever the variance in returns (i.e., $1/\rho_\eta$) is large enough that $\omega \in (-\alpha W_0/\rho_\eta, \alpha W_0/\rho_\eta)$.[39]

[37] In this investment context, the implications of naive inference seemingly run opposite those of cursedness (Eyster and Rabin 2005). As explored in Eyster, Rabin, and Vayanos (2015), cursed traders in asset markets fail to infer from price. Investors in our model, however, *over*-infer from past behavior—they revise their beliefs even when demand provides no new information.

[38] Interestingly, the speed at which perceptions diverge from the fundamental value is increasing in the level of risk aversion, α. As risk aversion increases, observers expect greater conservatism among the (supposedly) autarkic investors acting before them. Hence, observers think previous investors require stronger signals in order to allocate a majority of wealth to the riskier asset. Fixing $\bar{x}_t > 1/2$, Generation $t+1$ thus infers a greater expected return the larger is α.

[39] If we instead assumed log utility, then an investor diversifies for all finites values of $\hat{\omega}$. Hence, unlike the exponential-utility case where $\bar{x}_t$ reaches 0 or 1 in finite time, log utility implies $\bar{x}_t$ reaches 0 or 1 only in the limit as $t \to \infty$.

Even more damning, naive investors sometimes pick the "wrong" asset. That is, they allocate all wealth to the one that provides lower expected utility. For instance, if the optimal strategy is to invest 80% in the riskier asset, naive investors may end up investing 0%!

Corollary 1. *For any collection of parameters* $(\overline{\omega}, \rho_\omega, \rho_\varepsilon, \rho_\eta, \alpha)$*, there exists an open interval* $\Omega' \subset \mathbb{R}$ *such that whenever* $\omega \in \Omega'$*, naive investors eventually allocate all resources to the asset that provides lower expected utility.*

The intuition is straightforward in light of Proposition 5. Consider the case where $\overline{\omega} > \omega^* > 0$—people expect the riskier asset to outperform the safe asset. While this expectation is fulfilled whenever $\omega \in (0, \omega^*)$, the fact that ω falls sufficiently short of expectations implies that perceptions of ω decrease over time (Proposition 5, part 2). Eventually, $\hat{\omega}_t$ will fall so low that investors allocate no wealth to the riskier asset. Similarly, investors wrongly come to believe ω is large when $\omega^* < \omega < 0$.

5 Heterogeneous Preferences and Costly Herds

This section extends our "extremism" result of Section 3.2 to settings where agents have heterogeneous valuations over the set of available options. In such environments, naive inference can cause excessive, costly herding. Consider again the example in Section 3.2 where a town updates its beliefs about the quality of two new restaurants, A and B. We extend the model solely by introducing an outside option: each person either selects a restaurant or stays home. The payoff from staying home varies across agents. Suppose, in truth, both restaurants have low quality and A is only slightly better than B. Because the second round of diners form correct beliefs about quality, those with bad outside options go to A while the rest stay home. This pattern of behavior may lead the third round of diners astray. On the one hand, the fact that everybody dining out chooses A reflects positively on A's quality—naive observers expect A's market share to dwarf B's only when A is significantly better. On the other hand, the fact that many stay home can reflect *bad* news about the restaurants—if A were truly good, naive observes might expect few to stay home.

If the good news dominates, then everybody in the following rounds, including those who would be better off staying home, end up choosing A. Recall that Proposition 3 shows that, with common preferences, society comes to believe A has the highest possible quality and B has the lowest. Although in that case society overestimates the quality of A, people choose optimally. Here, however, such exaggeration of quality entices low-valuation consumers to sub-optimally choose A when they should in fact abstain. The remainder of this section discusses when society is likely to inefficiently herd on an action that is not universally beneficial, and emphasizes how large choice sets in particular can lead to such costly herding.

We focus on choice environments where action sets contain a menu of M "risky" options with uncertain payoffs, $\mathscr{A}^r \equiv \{A_1, \ldots, A_M\}$, and an outside option, A_0, with a known payoff that depends on a player's type. Such a combination of common values over actions with more or less known heterogeneity in outside options captures many natural situations. Abstracting away from the canonical binary-restaurant example, we frame our more general results in terms of a medical application: patients with a disease can either experiment with unproven treatments $\mathscr{A}^r$ or choose not to treat, A_0. Using the setup of Section 3.2, treatment $A_m \in \mathscr{A}^r$ yields payoff $q^m \in \{\underline{q}^m, 0\}$. We call a treatment "effective" if $q^m = 0$; otherwise it yields $\underline{q}^m < 0$, meaning it is only partially effective or harmful. To avoid technicalities, assume $\underline{q}^m \neq \underline{q}^j$ for all $m \neq j$, and, without loss of generality, $\underline{q}^1 = \max_m \{\underline{q}^m\}_{m=1}^M$. That is, A_1 is the best option whenever none are fully effective.

The payoff of the outside option depends on a player's type θ, and is denoted q_θ^0. For simplicity, we assume just two types, $\theta \in \{L, H\}$, such that $q_H^0 < \underline{q}^1 < q_L^0 < 0$. Type $\theta = H$ represents a "dire" patient who highly values intervention: her outside option is worse than the worst-case scenario under treatment, and thus she always chooses some $A_m \in \mathscr{A}^r$. A type $\theta = L$ patient has a relatively good outside option, so she selects a treatment only when sufficiently confident it will work. Let $\lambda \in (0,1)$ denote the fraction of high-valuation patients.[40]

Our primary concern is the public belief in states where the "risky" options are so bad that low-valuation types ($\theta = L$) should optimally choose their outside option. In such states, the second generation's behavior bifurcates based on type. All high-value types select the "least bad" risky option, A_1, while low types abstain. This generates the tension described above, where the third generation weights two potentially conflicting pieces of news: (1) the "consensus news"—all high types choose the *same* risky option, and (2) the "abstention news"—all low types stick with their outside option. Under the assumption that Generation 2 acts in autarky, consensus news unambiguously increases society's expectations of A_1, but abstention news potentially decreases those expectations.

Low types are forever lured into using A_1 if and only if the consensus news dominates. The most basic instance of this happens when the abstention news reveals no information to naive learners. For instance, suppose low-valuation patients require implausibly strong signals in order to experiment with risky treatments. In autarky, all low types abstain irrespective of their information. Because Generation 3 thinks that only the actions of high types reveal information, they update their beliefs solely on the fact that these types herd on A_1.

Any modification of the environment that increases the strength of the consensus news relative to the abstention news increases the chance that all types will inefficiently choose A_1 in the long run. One such environmental factor is the size of the choice set. The remainder of this section shows that for any λ, there exists a threshold $\overline{M}$ such that if the number of risky options exceeds $\overline{M}$, then

[40] Our results in this section hold regardless of whether a player's type is publicly observable or not.

all types eventually herd on A_1 irrespective of whether this is optimal. Loosely put, increasing the number of available options makes the consensus news more surprising: with more options to pick from, it becomes less likely that autarkistic choices would coincide unless A_1 were truly effective.

5.1 Binary Signals

We first examine the case of binary signals where the effect of choice-set cardinality is most transparent. We then allow for continuous signals for sake of robustness. As in Section 3.2, each player receives a conditionally independent signal s^m about each treatment. For all m, suppose $s^m \in \{0,1\}$ with conditional mass functions such that $f(s^m = 1|q^m = 0) = f(s^m = 0|q^m = \underline{q}^m) = \rho > 1/2$. Parameter ρ again denotes the precision of private signals. Let ω^0 denote the state in which no treatment is effective and ω^m denote that in which only Treatment m is effective. To ensure that increasing the number of options does not mechanically increase the likelihood of an effective treatment, we fix the probability of ω^0 at $\chi \in (0,1)$ independent of M. Finally, whether A_m is effective is independent of all other treatments, implying $\Pr(q^m = 0) = 1 - \chi^{\frac{1}{M}}$.

We show that if no treatment is effective, then society mislearns if and only if M is sufficiently large. To see this, consider inference and behavior among Generations 2 and 3. Since Generation 2 correctly infers the state, *all* patients in Generation 2 choose Treatment m whenever it is truly effective. This herd clearly indicates to future generations that A_m is effective, and society correctly learns. However, when no treatment works, Generation 2 sends a more opaque message to followers: dire types all use the treatment with the fewest side effects—Treatment 1—but low-value patients abstain.

What Generation 3 infers from this mixed behavior depends on the size of M. Recall that, in autarky, abstention by a low-value patient reveals bad news about the treatments only when her behavior depends on her signal realization. Importantly, the number of treatments to choose from, M, governs the informativeness a private signal. A patient's expected value of Treatment m conditional on good news,

$$\mathbb{E}[q^m|s^m = 1] = \Pr(q^m = \underline{q}^m|s^m = 1)\underline{q}^m = \left(\left(\frac{1-\chi^{\frac{1}{M}}}{\chi}\right)\left(\frac{\rho}{1-\rho}\right)+1\right)^{-1}\underline{q}^m, \tag{6}$$

is strictly decreasing in M: with more treatments available, good news is more likely to be a false positive. Essentially, the larger is M, the less informative is a positive signal. When M is sufficiently small, a low type chooses among those treatments for which she receives a good signal. But facing larger choice sets, a good signal alone is not strong enough to induce her to experiment. This happens when M is large enough that $\mathbb{E}[q^m|s^m = 1] < q_L^0$; that is,

$$M > \frac{\log \chi}{\log\left(1-\chi\left(\frac{q^1-q_L^0}{q_L^0}\right)\left(\frac{1-\rho}{\rho}\right)\right)} \equiv M^*. \tag{7}$$

First, suppose that $M > M^*$ (Condition 7), which implies low types never experiment in autarky. Since Generation 3 assumes that low types reveal no information, they infer only from the choices of dire types. But since Generation 3 neglects that dire types choose Treatment 1 as a result of social learning, they conclude it must be effective—this best explains a disproportionate use of Treatment 1 based on signals alone.

When $M < M^*$, Generation 3 treats the actions of low types as informative. The fact that these types abstain in $t = 2$ now provides news suggesting that no treatment is effective. This countervailing information limits the promising inference about A_1's efficacy. If it is strong enough, then Generation 3 correctly learns that no treatment works. Intuitively, the abstention news prevails when the fraction of low-valuation patients—those who abstain—is sufficiently large.

Proposition 6. *If signals about each option are independent and binary with precision ρ, then:*

1. *If Treatment m is solely effective (i.e., $\omega = \omega^m$), then $\hat{\omega}_t = \omega^m$ for all $t \geq 2$.*
2. *If no treatment is effective (i.e., $\omega = \omega^0$) and $M > M^*$, then $\hat{\omega}_t = \omega^1$ for all $t > 2$.*
3. *Suppose no treatment is effective (i.e., $\omega = \omega^0$) and $M < M^*$. If $\lambda < 1/2$, then $\hat{\omega}_t = \omega^0$ for all $t \geq 2$. If $\lambda > 1/2$, then there exits a threshold $\overline{M}(\lambda)$ strictly decreasing in λ such that $M > \overline{M}(\lambda)$ implies $\hat{\omega}_t = \omega^1$ for all $t > 2$.*

In summary, Proposition 6 states that when all treatments are ineffective, a large choice set ($M > M^*$) leads people to believe that one of them works. Furthermore, the threshold on M above which society mislearns is weakly decreasing in the prevalence of dire types (λ).

While this binary-signal environment reveals the basic intuition for why the number of available options affects inference, it has the "knife-edge" feature that with large M, no low-valuation patients experiment in autarky. Although this feature plays a central role in the intuition behind Proposition 6, it is not necessary. The next section generalizes this result by allowing for continuous signals such that, for all values of M, low types experiment with positive probability in autarky.

5.2 Continuous Signals

As in Section 3.2, patients receive conditionally independent signals $s^m \sim F^m(\cdot|q^m)$ about each treatment where each F^m satisfies MLRP (Definition 5). Additionally, suppose that each s^m is

continuously distributed and "unbounded" in the terminology of Smith and Sørensen (2000). This means that, for all $r \in (0,1)$, there is positive probability of both $p^m(s^m) < r$ and $p^m(s^m) > r$, where $p^m(s^m)$ is the posterior belief that Treatment m is effective conditional on signal s^m. With unbounded signals, a low-value patient selects no treatment in autarky only if she receives sufficiently bad news about all treatments. There exist signal thresholds $\bar{s}^m \in \mathbb{R}$, $m = 1, \ldots, M$, such that $\mathbb{E}[q^m | s^m] < q_L^0$ if and only if $s^m < \bar{s}^m$. Hence, a low type chooses A_0 only if $s^m < \bar{s}^m$ for all m. As in the binary-signal case, dire types experiment no matter their signal realizations.

Proposition 7. *Assume the state is ω^0. For all $\lambda \in (0,1)$, there exists a finite value $\overline{M} \geq 2$ such that $\hat{\omega}_t = \omega^0$ for all $t > 2$ if and only if $M < \overline{M}$. If $M \geq \overline{M}$, then $\hat{\omega}_t = \omega^1$ for all $t > 2$.*

Again, increasing the number of options makes mislearning more likely when the state is ω^0. Like the binary case, a naive Generation 3 observes conflicting consensus and abstention news. The behavior they observe—dire types select A_1 while all others abstain—is more likely in ω^1 than ω^0 if and only if $\mathbb{P}_{\omega^1}(0)^{1-\lambda}\mathbb{P}_{\omega^1}(1)^{\lambda} > \mathbb{P}_{\omega^0}(0)^{1-\lambda}\mathbb{P}_{\omega^0}(1)^{\lambda}$. Rewriting this condition in terms of likelihood ratios, ω^1 is naively more likely than ω^0 if and only if

$$\underbrace{\left(\frac{\mathbb{P}_{\omega^1}(0)}{\mathbb{P}_{\omega^0}(0)}\right)^{1-\lambda}}_{\text{Abstention News}} \underbrace{\left(\frac{\mathbb{P}_{\omega^1}(1)}{\mathbb{P}_{\omega^0}(1)}\right)^{\lambda}}_{\text{Consensus News}} > 1. \tag{8}$$

Condition 8 is more apt to hold the larger is M. To see the intuition, we first argue that the likelihood ratio of the dire herd on A_1—$\mathbb{P}_{\omega^1}(1)/\mathbb{P}_{\omega^0}(1)$—is increasing in M. Recall that a patient chooses A_1 in autarky only when s^1 is sufficiently large relative to each of her remaining $M-1$ signals and her outside option.[41] Consider how the probability of this event changes when moving from state ω^0 to ω^1: the signal distributions are unchanged for options $m = 2, \ldots, M$, yet s^1 now first-order stochastically dominates its previous distribution. This implies each of the M threshold conditions on s^1 is more likely to hold. In this sense, the change from ω^0 to ω^1 has an "M-fold" effect toward increasing the likelihood a patient chooses A_1 based solely on her signals. Hence, the larger is M, the higher is the likelihood of observing A_1 in ω^1 relative to ω^0. Turning to the likelihood of abstention, recall that a low type chooses A_0 only if *all* her signals are sufficiently low—$s^m < \bar{s}^m$ for all m. Out of these M conditions, only one—$s^1 < \bar{s}^1$—is less likely satisfied in state ω^1 relative to ω^0. Hence, the relative likelihood of observing no treatment in ω^1 versus ω^0 is independent of M.

These arguments imply that the strength of the consensus news increases in M, but the abstention news is independent of M. Consequently, the overall likelihood of ω^1 relative to ω^0 (the left-hand

[41] More specifically, for each $m = 2, \ldots, M$, there exist increasing functions $k^m : S^m \to S^1$ such that A_1 is chosen in autarky by Player n only if $s_n^1 > k^m(s_n^m)$ for each $m = 2, \ldots, M$ and $s^1 > \bar{s}^1$. The functions k^m are implicitly defined by $s_n^1 > k^m(s_n^m) \Leftrightarrow \mathbb{E}[q^1 | s_n^1] > \mathbb{E}[q^m | s_n^m]$.

side of Condition 8) increases in M. In fact, there always exists an M large enough such that the consensus among dire types will dominate inference. When this happens, society wrongly concludes A_1 is effective and all patients of either type choose A_1 from Generation 3 onward. Since it appears as if dire patients have strong private information, their consensus behavior lures the rest of society to imitate them.[42] Interestingly, this mistake never happens when there is only one treatment available: naive observers expect all dire types to choose the lone treatment no matter the state, so they infer nothing from their herd. It is the availability of many viable options that leads naive observers to over-infer from consensus behavior.

While this section focused on a particular application, the logic easily extends to other contexts. First, the assumption of two types was solely for simplicity. Our result holds for any finite number of types who differ in their outside option. Second, the setting naturally captures other applications, such as learning about the return on risky investments among agents who vary in risk aversion. Our result implies that, with many potential investments, very risk-averse agents may wrongly follow the investment strategies of the less risk averse.

6 Learning the Distribution of Information

Until now, we have assumed that players know the distribution of private signals conditional on the payoff-relevant state. This section relaxes that assumption, and explores scenarios where players simultaneously infer which payoff state and information structure best explains past behavior. Naive inference distorts player's perceptions of the informativeness of private signals. Namely, with common preferences and uncertainty over the precision of signals, we find an informational analog of our extremism result in Section 3.2: people conclude that signals are as precise as possible. This follows from the fact that social learning leads people to herd on a single action. In familiar settings where the degree of consensus in autarkic behavior increases in the precision of signals, maximal precision is the best (naive) explanation for why all predecessors make the same choice.[43]

To clearly demonstrate this point, first consider the simplest variant of the "extremism" environ-

[42] Gagnon-Bartsch (2015), who studies social learning among people who exaggerate the extent to which others share their tastes, predicts a similar form of costly herding. Like here, society might settle on a single action when in fact those with different tastes are better off choosing distinct options.

[43] We have additionally worked out a model (excluded from the paper) showing implications of naive herding in an environment with *aggregate* uncertainty—after combining all information in the economy, a rational player remains uncertain about payoffs. In such settings, a naive player rightfully expects to remain uncertain, but inevitably grows confident in some (perhaps false) payoff state. For example, imagine investors learning about the probability that an asset will yield positive returns. Naive traders conclude that the most popular asset will surely pay off, while the remaining assets never will. In contrast, a rational observer correctly infers the probabilities with which each asset pays off, but remains uncertain about the particular payoff realization. Hence, relative to rational inference, naive inference predicts overconfidence about the payoff state.

ment presented in Section 3.2. Each action $m \in \{1,\ldots,M\}$ has a payoff q^m independently drawn from $\{0,1\}$. Players receive an i.i.d. binary signal $s^m \in \{0,1\}$ about each action m with mass function $f(s^m = q^m|q^m) = \rho$. Nature draws the precision of signals, ρ, according to a known distribution h with support $\mathscr{R} \subset [.5,1]$.[44] The state ω is thus $(q^1,...,q^M,\rho)$.

In autarky, each Player (n,t) chooses an action that delivered a positive signal (i.e., $s^m_{(n,t)} = 1$). If she receives good news about more than one option (or none), suppose she breaks the tie by choosing that with the lowest index.[45] As usual, Generation 2 infers the optimal action—denoted by A_m—and rationally herds on A_m. Hence, Generations $t > 2$ will come to believe in the state that maximizes the autarkic probability of A_m, $\mathbb{P}_{(\mathbf{q},\rho)}(m)$. Proposition 3 implies that, for a fixed ρ, $(\mathbf{q},\rho)$ maximizes $\mathbb{P}_{(\mathbf{q},\rho)}(m)$ if and only if $\mathbf{q} = \omega^m_e$, where ω^m_e is an "extreme state" in the sense of Definition 6. Consequently, all Generations $t > 2$ will believe ω^m_e. Fixing this belief, $\mathbb{P}_{(\omega^m_e,\rho)}(m) = (\rho)^m$, which is maximized at $\rho = 1$.[46] The logic is straightforward. For any value of $\rho < 1$, a naive observe would expect her autarkic predecessors to occasionally receive misleading signals and take actions other than A_m. An autarkic society behaves identically only when misleading signals are impossible—that is, $\rho = 1$.

The logic extends to more-general environments with continuous signals. Suppose each option $m \in \{1,\ldots,M\}$ has a payoff q^m drawn uniformly from a finite set $Q \subset \mathbb{R}$. Signals about each q^m are distributed as $s^m = q^m + \varepsilon^m/\sqrt{\rho}$, where $\{\varepsilon^m\}_{m=1}^M$ are i.i.d. mean-zero random variables with full support over $\mathbb{R}$ and variance normalized to 1. Since $\text{Var}[s^m] = 1/\rho$, the parameter ρ measures precision. Again, nature draws ρ from a distribution h with finite support $\mathscr{R} \subset \mathbb{R}_{++}$.[47]

Proposition 8. *Let $\bar{\rho} \equiv \max \mathscr{R}$. If $q^m = \max(q^1,...,q^M)$, then each Generation $t > 2$ grows certain that $\mathbf{q} = \omega^m_e$ and that $\rho = \bar{\rho}$.*

Since naive observers assume actions are based solely on private signals, they believe the dispersion in behavior reflects the variance in private information. And because they eventually observe a herd, they conclude this variance is minimal.

This false belief has no consequences in this particular setting: people are right to infer from the wisdom of the crowd even if their understanding of its source is wrong. But there can exist contexts where exaggerating the quality of individuals' information may lead people astray. For instance,

[44] We restrict attention to $\rho \in [.5,1]$ so that signals satisfy MLRP (Definition 5) and uniquely reveal the payoff state when aggregated: $s^m = 1$ occurs more often than $s^m = 0$ if and only if $q^m = 1$. If $\rho < 1/2$, then this is no longer true. The case of $\rho < 1/2$ is explored by Acemoglu et al. (2016) in the context of a fully Bayesian model.

[45] This tie-breaking rule is unnecessary for our result, and is simply imposed for the sake of exposition.

[46] The probability that Player (n,t) chooses A_m in autarky equals the probability that she receives signal $s^j_{(n,t)} = 0$ about each $j < m$ and $s^m_{(n,t)} = 1$ about m. In state ω^m_e, this happens with probability $(\rho)^m$.

[47] We assume that signals have a common precision across options for simplicity. An analog of Proposition 8 holds when the precision may vary across options, but it requires additional assumptions on the prior distribution from which these precisions are drawn.

consider the investment setting of Section 4, and suppose that the precision of signals ρ_ε is initially unknown. Proposition 8 shows that investors will overestimate ρ_ε. If a new asset is introduced and investors assume its information structure is similar to those already in the market, then they will overweight their own signals when assessing the value of this new asset. And they will also exaggerate the precision of *others'* signals. A naive investor who consults a financial advisor about a new prospect before the market reaches a consensus may rely too heavily on that advice.

This environment has the important feature that, whenever perfectly revealing signals are possible, herds are consistent with a naive agent's model of the world. Some of our previous applications lacked this property. For instance, in Section 3.2, naive observers best explain a herd by assuming an extreme state. Although such a state minimizes the cross entropy between the predicted and observed play, the long-run distribution of actions never matches what a naive observer *expects* to see when signals are noisy. Naive players expect disperse behavior, yet they see predecessors choose identically. Allowing agents to simultaneously draw inference about the signal distribution and payoff structure eliminates this issue of "inexplicable" observations by permitting naive agents to make sense of what they see.

Naivete has interesting implications when agents are additionally uncertain about the interpretation of a signal. In the binary-signal example above, the assumption that $\rho > 1/2$ implies a high signal ($s = 1$) unambiguously indicates higher expected quality than a low signal ($s = 0$). If we instead allow ρ to take any value in $[0,1]$, this interpretation is lost: whether $s = 1$ suggests high quality depends on $\rho \lessgtr 1/2$. Interestingly, after observing a herd, each player will use her own signal realization to form her belief about the meaning of signals. To illustrate, suppose that Generation 2 herds on A_1. Generation 3 comes to believe that option 1 has high quality, but what do they infer about ρ? All players conclude signals are perfectly precise (i.e., $\rho \in \{0,1\}$). However, they disagree on signal interpretation: those who received $s = 1$ think that high signals are associated with high quality (i.e., they believe $\rho = 1$), but those who observe $s = 0$ think *low* signals imply high quality (they believe $\rho = 0$).

In the discussion above, players rationalize consensus behavior by assuming each predecessor received a very precise, yet independent, signal. However, there are other plausible ways to rationalize such a consensus. For instance, perhaps all predecessors observed the same realization of a noisy signal—that is, signals are perfectly correlated. In environments where the correlation between players' signals is uncertain, herding will cause naive agents to infer that private information is perfectly correlated. Since herds are rather uninformative when they result from perfectly-correlated information, naive observers in this case may form *less* confident beliefs than their rational counterparts. This counters the usual intuition that naive inference leads to overconfident beliefs.

7 Discussion and Conclusion

This paper explores new predictions of Eyster and Rabin's (2010) model of naive inference that emerge in a broader array of environments than previously studied. In many natural environments, there exist "abandoned states" that naive agents continually disbelieve even when they are true. For instance, when payoffs are independent across actions, naive agents systematically overestimate the payoff of the best option relative to its alternatives. As we have shown, these extreme beliefs can lead to severe under-diversification in investment settings, or to the costly over-adoption of goods beneficial for only a minority of consumers. This logic also suggests that, relative to more direct methods of information transmission, consumers may be more susceptible to overpay for products when information spreads through observational learning. Finally, the same force driving these extreme perceptions leads people to overestimate the quality of their private information. We now conclude by discussing the robustness of our results and proposing some extensions.

Although our model assumes a particular observation structure, most of our results hold more broadly. Our assumption that each generation is arbitrarily large and observes only the previous generation allowed us to characterize both a deterministic path of beliefs and the set of states on which beliefs may concentrate. While the deterministic path depends crucially on the assumption of large generations, the characterization of limit beliefs (i.e., Lemma 2) holds irrespective of the observation structure so long as the number of predecessors each agent observes grows large in t. The canonical "herding model" serves as a primary example of such a structure: one agent acts per round and each agent observes the complete history of play (e.g., Bikchandani, et al. 1992; Smith and Sørensen 2000).

The following argument provides a simple intuition for why the conclusion of Lemma 2 is robust to the observation structure—that is, why any stable long-run belief must be a fixed point of the transition function ϕ introduced in Lemma 1. Let $O_{(n,t)}$ be the set of all predecessors who Player (n,t) observes; our only assumption on the observation structure is that for all $n = 1,\ldots,N$, $\lim_t |O_{(n,t)}| = \infty$. If beliefs converge to ω_k, then the aggregate behavior observed by a Player (n,t) late in the sequence will eventually resemble $\mathbb{T}_{\omega_k}$. Additionally, because a naive learner thinks each predecessor acts independently, the observed order of actions does not influence her inference—only the *aggregate* distribution of behavior matters. Hence, as $t \to \infty$, Player (n,t) observes an arbitrarily large population taking actions distributed according to $\mathbb{T}_{\omega_k}$. Following the same logic as Lemma 1, the state that maximizes the "autarkic" likelihood of this observation is $\hat{\omega} = \arg\min_{\omega\in\Omega} H(\mathbb{T}_{\omega_k}, \mathbb{P}_\omega) = \phi(\omega_k)$. Thus, in order for Player (n,t) to remain confident in ω_k, it must be that $\omega_k = \phi(\omega_k)$. This implies that the only states on which public beliefs may settle—no matter the specific observation structure—are characterized by ϕ. As such, all of our results characterizing environments where a naive society fails to learn the truth (in the long-run)

are robust to alternative structures.

While relaxing our large-generation assumption preserves our result that beliefs converge only to states in Ω^*, we can no longer perfectly predict on *which* $\omega \in \Omega^*$ public beliefs will eventually concentrate. With finite generations, limit beliefs will depend on the sample path of signals. No matter the true state, for each $\omega \in \Omega^*$, there is a positive probability that society grows certain of ω. Hence, finite generations only broaden the scope for mislearning: there is a chance society mislearns even when the truth lies in Ω^*, which was impossible with large generations.[48]

As discussed in Section 2.3, this is not the first paper to study "redundancy neglect" in environments richer than the canonical model. However, it does provide predictions distinct from earlier work. As mentioned, DeMarzo et al. (2003) consider a variant of DeGroot's (1974) model in which players share their signals about a normally distributed state with their neighbors in a network. Each round, a player observes her neighbors' mean posterior beliefs and updates her own belief. In doing so, she treats her neighbors' reports as independent signals, neglecting that their posteriors already incorporate information previously shared amongst each other. Since agents over-count signals, they grow confident in some false state whenever initial signals are misleading.

The implications of this error differ from ours in two ways. First, agents in DeMarzo et al. do not gravitate toward extreme perceptions over time. Since players use a naive averaging rule, beliefs converge to a weighted average of initial signals instead of tending to extreme values. Second, when the number of agents observed grows large, players in DeMarzo et al. (2003) learn correctly. The law of large numbers implies that the first round of communication sends players directly to a confident and correct posterior. In our setting, even if players correctly learn the state after one round of observation, later generations mislearn by reinterpreting confident behavior as if it were autarkic. Relative to existing models, extremism and unlearning are distinct predictions of Eyster and Rabin's formulation of naive inference.[49]

There are several interesting applications of naive learning beyond the scope of our simple framework. Since our model of naive inference specifies only how a player best responds to her observations, we lack a solution concept allowing us to analyze settings where a player's payoff depends directly on others' actions.[50] For instance, a natural extension of our extremism results would analyze how firms set prices in order to exploit, or undermine, the fact that naive consumers come to believe one product is superior to all its competitors on the market. Similarly, in the in-

[48]The logic follows directly from Eyster and Rabin's (2010) Proposition 4: so long as $\omega \in \Omega^*$, there exists a sample path of signals realized with positive probability such that beliefs settle on ω.

[49]Like DeMarzo et al. (2003), many of the other papers on naive inference in network settings build from DeGroot's (1974) framework (e.g., Golub and Jackson 2010; Chandrasekhar et al. 2012). As such, the results of those papers—insofar as they relate to ours—are in line with DeMarzo et al. (2003).

[50]Eyster and Rabin (2008) define an equilibrium concept called "Inferentially-Naive Information Transmission" (INIT) that incorporates the form of naivete assumed in this paper. INIT reduces to naivete in social-learning environments where players care about others' actions solely for their informational content.

vestment setting of Section 4, an equilibrium model incorporating prices would predict a bubble: as the perceived returns to an asset continually increase, so will its price.

Finally, interesting predictions likely arise with endogenous timing and costly delay.[51] Instead of moving in sequence, suppose each agent can invest as soon as she sees fit. Although the benefit from investment shrinks over time, there is incentive to strategically delay in order to glean information about returns from others' decisions. But a naive investor, who wrongly thinks others rely solely on private signals, expects those with optimistic signals to invest immediately and those with pessimistic signals to forever abstain. That is, a naive investor neglects others' incentive to delay, expecting many to invest immediately whenever returns are high. Because investors do in fact exercise the option to delay, naive observes wrongly attribute limited initial investment to low returns. We conjecture that such environments systematically promote pessimism among naive investors. Analyzing the equilibria of these various extensions is part of an on-going research agenda incorporating both naive inference and cursed thinking (Eyster and Rabin 2005) into formal models of dynamic learning.

References

[1] ACEMOGLU, D., V. CHERNOZHUKOV, AND M. YILDIZ (2016): "Fragility of Asymptotic Agreement under Bayesian Learning," *Theoretical Economics*, 11(1), 187–225

[2] BANERJEE, A. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797–817.

[3] BANERJEE, A. AND D. FUDENBERG (2004): "Word-of-Mouth Learning," *Games and Economic Behavior*, 46, 1–22.

[4] BARBERIS, N., A. SHLEIFER, AND R. VISHNY (1998): "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, 307–343.

[5] BERK, R. (1966): "Limiting Behavior of Posterior Distributions When the Model Is Incorrect," *Annals of Mathematical Statistics*, 37, 51–58.

[6] BIKCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992–1026.

[7] BOHREN, A. (2016): "Informational Herding with Model Misspecification," *Journal of Economic Theory*, forthcoming.

[51] Such settings have been analyzed with rational agents by Chamely and Gale (1994) and Gul and Lundholm (1995).

[8] CRAWFORD, V. AND N. IRIBERRI (2007): "Level-K Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica*, 75(6), 1721–1770.

[9] CHAMLEY, C. AND D. GALE (1994): "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica*, 62(5), 1065–1085.

[10] CHANDRASEKHAR, A., H. LARREGUY, AND J.P. XANDRI (2012): "Testing Models of Social Learning on Networks: Evidence From a Framed Field Experiment," Mimeo.

[11] COVER, T. AND J. THOMAS (1991): *Elements of Information Theory*, John Wiley & Sons, New York.

[12] DEGROOT, M. (1970): *Optimal Statistical Decisions*, McGraw-Hill, New York.

[13] DEGROOT, M. (1974): "Reaching a Consensus," *Journal of the American Statistical Association*, 69(345), 118–121.

[14] DEMARZO, D. VAYANOS, AND J. ZWIEBEL (2003): "Persuasion Bias, Social Influence, and Uni-Dimensional Opinions," *Quarterly Journal of Economics*, 118, 909–968.

[15] ESPONDA, I. AND D. POUZO (2016): "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica*, forthcoming.

[16] EYSTER, E. AND M. RABIN (2005): "Cursed Equilibrium," *Econometrica*, 73(5), 1623–1672.

[17] EYSTER, E. AND M. RABIN (2008): "Naive Herding," Mimeo.

[18] EYSTER, E. AND M. RABIN (2010): "Naive Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2(4), 221–243.

[19] EYSTER, E. AND M. RABIN (2014): "Extensive Imitation is Irrational and Harmful," *Quarterly Journal of Economics*, 129(4), 1861–1898.

[20] EYSTER, E., M. RABIN, AND D. VAYANOS (2015): "Financial Markets where Traders Neglect the Informational Content of Asset Prices," Mimeo.

[21] EYSTER, E., M. RABIN, AND G. WEIZSÄCKER (2015): "An Experiment on Social Mislearning," Mimeo.

[22] FOSTER, A., AND M. ROSENZWEIG (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103(6), 1176–1209.

[23] GAGNON-BARTSCH, T. (2015): "Taste Projection in Models of Social Learning," Mimeo.

[24] GLAESER, E., AND C. NATHANSON (2015): "An Extrapolative Model of House Price Dynamics," Mimeo.

[25] GOLUB, B. AND M. JACKSON (2010): "Naive Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2(1), 112–49.

[26] GREENWOOD, R. AND A. SHLEIFER (2014): "Expectations of Returns and Expected Returns," *Review of Financial Studies*, 27(3): 714–746.

[27] GUL, F. AND R. LUNDHOLM (1995): "Endogenous Timing and the Clustering of Agents' Decisions," *Journal of Political Economy*, 103, 1039–1066.

[28] JACKSON, M. AND E. KALAI (1997): "Social Learning in Recurring Games," *Games and Economic Behavior*, 21, 102–134.

[29] KÜBLER, D., AND G. WEIZSÄCKER (2004): "Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory," *Review of Economic Studies*, 71(2), 425–441.

[30] LEVY, G. AND R. RAZIN (2015): "Correlation Neglect in Group Communication," Mimeo.

[31] MILGROM, P. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), 380–391.

[32] MUNSHI, K. (2003): "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 73(1), 185–213.

[33] RABIN, M. (2002): "Inference by Believers in the Law of Small Numbers", *Quarterly Journal of Economics*, 117(3): 775–816.

[34] SCHWARTZSTEIN, J. (2014): "Selective Attention and Learning," *Journal of the European Economic Association*, 12(6), 1423–1452.

[35] SMITH, L. AND P. SØRENSEN (2000): "Pathological outcomes of observational learning," *Econometrica*, 68(2), 371–398.

A Proofs

Proof of Lemma 1.

Proof. Let N_t be the number of players in Generation t. The proof follows by induction while taking $N_t \to \infty$ sequentially for each $t = 1, 2, \ldots$. Note that this sequential approach requires no additional assumptions on how the limiting ratio $N_t/N_{t'}$ behaves for any two periods t and $t' \neq t$.

Naive observers in any period $t+1$ think actions in t conditional on ω are independent draws from $\mathbb{P}_\omega$: they think $N_t \mathbf{a}_t \sim \text{Multinomial}(N_t, \mathbb{P}_\omega)$ in state ω, implying

$$\Pr(\mathbf{a}_t|\omega) = C(N_t, \mathbf{a}_t) \prod_{m=1}^{M} \mathbb{P}_\omega(m)^{N_t a_t(m)},$$

where $C(N,\mathbf{a}) \equiv N!/\prod_{m=1}^{M} Na(m)!$ is a normalization constant independent of ω. Thus

$$\frac{\pi_{t+1}(j)}{\pi_{t+1}(k)} = \frac{\Pr(\mathbf{a}_t|\omega_j)\pi_1(j)}{\Pr(\mathbf{a}_t|\omega_k)\pi_1(k)} = \left(\frac{\prod_{m=1}^{M}\mathbb{P}_{\omega_j}(m)^{a_t(m)}}{\prod_{m=1}^{M}\mathbb{P}_{\omega_k}(m)^{a_t(m)}}\right)^{N_t}\frac{\pi_1(j)}{\pi_1(k)}. \tag{9}$$

Suppose the state is ω^*. Recall that $a_1(m) = \frac{1}{N_1}\mathbb{1}\{x_{(n,1)} = A_m\}$. By the Strong Law of Large Numbers, $a_1(m) \to \mathbb{P}_{\omega^*}(m)$ a.s. for each $m = 1,\ldots,M$ as $N_1 \to \infty$. Uniqueness of $\mathbb{P}_\omega$ (Assumption 1) implies $\hat{\omega}_2 = \omega^*$. Now consider updating in $t = 3$ fixing $\pi_2 = \delta_{\hat{\omega}_2}$. Since $N_2 \to \infty$ implies $a_2(m) \to \mathbb{T}_{\hat{\omega}_2}(m)$ a.s., the likelihood ratio $\pi_3(j)/\pi_3(k)$ converges to 0 in N_2 for all $\omega_j \neq \omega_k \Leftrightarrow \omega_k = \arg\max_{\omega\in\Omega}\prod_{m=1}^{M}\mathbb{P}_\omega(m)^{\mathbb{T}_{\hat{\omega}_2}(m)}$. Letting δ_ω denote a degenerate belief on state ω, $\pi_3 \to \delta_{\hat{\omega}_3}$ a.s. where $\hat{\omega}_3 \equiv \arg\max_{\omega\in\Omega}\prod_{m=1}^{M}\mathbb{P}_\omega(m)^{\mathbb{T}_{\hat{\omega}_2}(m)}$; this state is unique for all $\mathbb{T}_\omega$ by Assumption 2.

Given that the updating rule in Equation 9 is independent of t, the inductive step is identical to the base case: if Generation $t > 2$ is certain of $\hat{\omega}_t$, then $a_t(m) \to \mathbb{T}_{\hat{\omega}_t}(m)$ a.s. in N_t, and thus $\pi_{t+1}(j)/\pi_{t+1}(k)$ converges to 0 in N_t for all $\omega_j \neq \omega_k \Leftrightarrow \omega_k = \arg\max_{\omega\in\Omega}\prod_{m=1}^{M}\mathbb{P}_\omega(m)^{\mathbb{T}_{\hat{\omega}_t}(m)}$. Hence, for each $t \geq 2$ and any $\hat{\omega}_t \in \Omega$, $\pi_{t+1} \to \delta_{\hat{\omega}_{t+1}}$ a.s., where

$$\hat{\omega}_{t+1} = \arg\max_{\omega\in\Omega}\prod_{m=1}^{M}\mathbb{P}_\omega(m)^{\mathbb{T}_{\hat{\omega}_t}(m)} = \arg\min_{\omega\in\Omega}\left(-\sum_{m=1}^{M}\mathbb{T}_{\hat{\omega}_t}(m)\log\mathbb{P}_\omega(m)\right) = \arg\min_{\omega\in\Omega}H(\mathbb{T}_{\hat{\omega}_t},\mathbb{P}_\omega).$$

The proof is completed by defining $\phi(\hat{\omega}_t) \equiv \arg\min_{\omega\in\Omega}H(\mathbb{T}_{\hat{\omega}_t},\mathbb{P}_\omega)$.

■

Proof of Proposition 1.

Proof. First consider the case where $\Omega^* \neq \varnothing$. If the true state is $\omega_k \in \Omega^*$, then $\hat{\omega}_2 = \omega_k$ and $\hat{\omega}_3 = \phi(\omega_k) = \omega_k$. Rolling forward, $\hat{\omega}_t = \omega_k$ for all $t \geq 2$, so $\beta_k(\omega_k) = 1$. Now suppose $\Omega^* = \varnothing$, so $\langle\hat{\omega}_t\rangle$ is necessarily cyclic. Consider $\omega_k \in \Omega^{**}$, and let $\widehat{\Omega}_k$ denote the absorbing set containing ω_k. Note that $\omega_k \in \widehat{\Omega}_k$ implies $\omega_k = \phi^c(\omega_k)$ where $c \equiv |\widehat{\Omega}_k| \geq 2$. Thus in state ω_k, $\hat{\omega}_2 = \omega_k$ and $\hat{\omega}_{2+c\tau} = \omega_k$ for all $\tau = 1,2,\ldots$. Thus $\beta_k(\omega_k) = 1/c$. ■

Proof of Proposition 2.

Proof. Since each Generation $t > 1$ has degenerate beliefs, common preferences imply that for all $t > 1$, the action distribution $\mathbf{a}_t$ is degenerate (a "herd"). Denote by $\mathbf{a}^m$ the action distribution degenerate on A_m, and let $\mathscr{A}^h = \{\mathbf{a}^m\}_{m=1}^{M}$. For each $\mathbf{a}^m \in \mathscr{A}^h$, let $\hat{\omega}^m = \arg\min_{\omega\in\Omega}H(\mathbb{P}_\omega,\mathbf{a}^m)$ denote the updated belief upon observing $\mathbf{a}^m$, and let Ω^h be the set of distinct values of $\hat{\omega}^m$ across $m = 1,\ldots,M$. Since $\hat{\omega}^m$ is unique fixing $\mathbf{a}^m$, $|\Omega^h| \leq M$. Because $\mathbf{a}_t \in \mathscr{A}^h$ for all $t > 1$, $\hat{\omega}_t \in \Omega^h$ for all $t > 1$, and thus $\Omega^{**} \subseteq \Omega^h$. It follows that $|\Omega^{**}| \leq M$, which implies $\Omega^{**} \subsetneq \Omega$ whenever $|\Omega| > M$.

■

Proof of Proposition 3.

Proof. We first prove a lemma that we use to prove both this proposition some that follow.

Lemma 4. *If each F^m satisfies strict MLRP, then $\mathbb{P}_\omega(m)$ is strictly increasing in q^m and strictly decreasing in q^j for all $j \neq m$.*

Proof of Lemma. Without loss of generality, we prove the result for $\mathbb{P}_\omega(1)$. We make use of well-known implicatiosn of MLRP (see Milgrom 1981, Proposition 2):

Remark 1. *Suppose F^m satisfies strict MLRP.*

1. *$\mathbb{E}[q^m|s^m]$ is strictly increasing in $s^m \in S^m$.*
2. *$F^m(s|q^m)$ satisfies first-order stochastic dominance in s: if $q^m > \tilde{q}^m$, then for all $s \in S^m$, $F^m(s|q^m) \leq F^m(s|\tilde{q}^m)$.*

In autarky (i.e., $t = 1$), Player n with signal realization $\mathbf{s} = (s^1, \ldots, s^M)$ chooses A_1 if $1 = \arg\max_{m \in \{1,\ldots,M\}} \mathbb{E}[q^m|\mathbf{s}]$. Since signals are independent across options, $\mathbb{E}[q^m|\mathbf{s}] = \mathbb{E}[q^m|s^m]$ for each m, and MLRP implies that each $\mathbb{E}[q^m|s^m]$ is strictly increasing in s^m. (All remaining instances of "increasing" and "decreasing" within this proof are meant in the strict sense.) For each m, let $k_m(\cdot)$ be the increasing function implicitly defined by $\mathbb{E}[q^1|s^1] > \mathbb{E}[q^m|s^m]$ iff $s^1 > k(s^m)$. Thus, action A_1 is chosen iff $s^1 > k_m(s^m)$ for all m. This implies that in state $\omega = (q^1, \ldots, q^M)$, the autarkic probability of choice A_1 is

$$\mathbb{P}_\omega(1) = \prod_{m=2}^{M} \Pr\left(s^1 > k_m(s^m)\middle|\omega\right) = \prod_{m=2}^{M} \int_{S^m} 1 - F^1\left(k_m(s^m)\middle|q^1\right) \mathrm{d}F^m\left(s^m\middle|q^m\right). \tag{10}$$

We first show that $\mathbb{P}_\omega(1)$ is (strictly) increasing in q^1. From Remark 1, $F^1\left(k_m(s^m)\middle|q^1\right)$ is decreasing in q^1, which implies that each term $\int_{S^m} 1 - F^1\left(k_m(s^m)\middle|q^1\right) \mathrm{d}F^m\left(s^m\middle|q^m\right)$ of the product in Equation 10 is increasing in q^1, and thus $\mathbb{P}_\omega(1)$ is increasing in q^1. Next, we show $\mathbb{P}_\omega(1)$ is decreasing in q^m for all $m \geq 2$. For any arbitrary $m \geq 2$, note that each term of the product in Equation 10 can be expressed as $\mathbb{E}[h(s^m)|q^m]$ where $h(s^m) = 1 - F^1\left(k_m(s^m)|q^1\right)$ is a decreasing function of s^m independent of q^m. It is well known that if random variable X first-order stochastically dominates X', then $\mathbb{E}[h(X)] < \mathbb{E}[h(X')]$ for any decreasing function $h(\cdot)$ provided these expectations are finite. Since s^m conditional on q^m first-order stochastically dominates s^m conditional on $\tilde{q}^m$ iff $q^m > \tilde{q}^m$, $\mathbb{E}[h(s^m)|\tilde{q}^m] > \mathbb{E}[h(s_m)|q^m]$ iff $q^m > \tilde{q}^m$, which implies that $\mathbb{E}[h(s^m)|q^m] = \int_{S^m} 1 - F^1\left(k_m(s^m)\middle|q^1\right) \mathrm{d}F^m\left(s^m\middle|q^m\right)$ is decreasing in q^m. Thus, from Equation 10, $\mathbb{P}_\omega(1)$ is decreasing in q^m. Since $\mathbb{P}_\omega(1)$ is increasing in q^1 and decreasing in q^m for all $m \geq 2$, it follows that ω_e^1 uniquely maximizes $\mathbb{P}_\omega(1)$. This concludes the proof of the lemma.

We now use Lemma 4 to prove Proposition 3. Without loss of generality, index options such that $q^1 = \arg\max_m\{q^m\}$. By Assumption 1, $\mathbf{a}_1$ reveals ω to Generation 2, implying $a_2(1) = 1$ and $a_2(m) = 0$ for $m \geq 2$. From Lemma 1, $\hat{\omega}_3 = \phi(\hat{\omega}_2) = \arg\max_{\tilde{\omega}\in\Omega}\prod_{m=1}^{M}\mathbb{P}_{\tilde{\omega}}(m)^{a_2(m)} = \arg\max_{\tilde{\omega}\in\Omega}\mathbb{P}_{\tilde{\omega}}(1)$. Hence, Lemma 4 implies $\hat{\omega}_3 = \omega_e^1$. Since $\hat{\omega}_3 = \omega_e^1$, all players in Generation 3 choose A_1, implying $\hat{\omega}_4 = \phi(\omega_e^1) = \arg\max_{\omega\in\Omega}\mathbb{P}_{\omega}(1) = \omega_e^1$. Since ω_e^1 is a fixed point of ϕ, $\hat{\omega}_t = \omega_e^1$ for all $t > 2$.

■

Proof of Proposition 4.

Proof. Part 1. Suppose $\omega = H$, so $\hat{\omega}_2 = H$ and $\mathbf{a}_2 = \mathbb{T}_H$, where $\mathbb{T}_H(A) = \lambda$. Using Lemma 1, $\hat{\omega}_3 = H \Leftrightarrow \mathbb{P}_H(A)^{\lambda}\mathbb{P}_H(B)^{1-\lambda} > \mathbb{P}_L(A)^{\lambda}\mathbb{P}_L(B)^{1-\lambda}$. Note that $\mathbb{P}_H(A) = \psi[\lambda\rho + (1-\lambda)(1-\rho)]$ and $\mathbb{P}_L(A) = \psi[(1-\lambda)\rho + \lambda(1-\rho)]$. Letting $\ell \equiv \lambda\rho + (1-\lambda)(1-\rho)$, the autarkic frequencies simplify to $\mathbb{P}_H(A) = \psi\ell$ and $\mathbb{P}_L(A) = \psi(1-\ell)$. Thus $\mathbb{P}_H(A)^{\lambda}\mathbb{P}_H(B)^{1-\lambda} > \mathbb{P}_L(A)^{\lambda}\mathbb{P}_L(B)^{1-\lambda}$ iff

$$k_H(\ell,\psi) \equiv \left(\frac{\ell}{1-\ell}\right)^{\lambda}\left(\frac{1-\psi\ell}{1+\psi\ell-\psi}\right)^{1-\lambda} > 1.$$

Since $k_H(\ell,\psi)$ is decreasing in ψ, $k_H(\ell,\psi) > 1$ for all $\psi \in (0,1)$ if $k_H(\ell,1) > 1$. Since $\lambda > 1/2$ and $\rho > 1/2$ imply $\ell \in (1/2,1)$, it follows that $k(\lambda,1) = (\frac{\ell}{1-\ell})^{2\lambda-1} > 1$. Hence $\hat{\omega}_3 = H$, which implies $\omega = H$ is a fixed point of ϕ, and thus $\hat{\omega}_t = H$ for all $t > 2$.

Part 2. Suppose $\omega = L$, so $\hat{\omega}_2 = L$ and $\mathbf{a}_2 = \mathbb{T}_L$, where $\mathbb{T}_L(A) = 1-\lambda$. Following the same logic as Part 1, $\hat{\omega}_3 = H$ iff

$$k_L(\ell,\psi) \equiv \left(\frac{\ell}{1-\ell}\right)^{1-\lambda}\left(\frac{1-\psi\ell}{1+\psi\ell-\psi}\right)^{\lambda} > 1.$$

Fixing $\lambda > 1/2$, $k_L(\ell,\psi) > 1 \Leftrightarrow \psi < \frac{1-\Lambda}{\lambda-\Lambda(1-\ell)} \equiv \bar{\psi}(\ell)$, where $\Lambda \equiv \left(\frac{1-\ell}{\ell}\right)^{\frac{1-\lambda}{\lambda}}$. Note that $\bar{\psi}(\ell)$ is decreasing in ℓ: $\frac{\partial}{\partial\ell}\bar{\psi}(\ell) < 0 \Leftrightarrow \left(\frac{2\ell-1}{\ell^2}\right)\left(\log\left(\frac{1-\lambda}{\lambda}\right)+1\right)\Lambda < 1$. This holds for any $\ell \in (1/2,1)$ because: (i) $\frac{2\ell-1}{\ell^2} < 1$, (ii) $\log\left(\frac{1-\lambda}{\lambda}\right)+1 < 1$, and (iii) $\Lambda < 1$. Finally, it's straightforward to verify that $\bar{\psi}(.5) = 1$ and $\bar{\psi}(1) = 0$. Thus, for any $\ell \in (1/2,1)$, $\bar{\psi}(\ell) \in (0,1)$ and $\psi < \bar{\psi}(\ell) \Leftrightarrow \hat{\omega}_3 = H$. Thus, if $\psi > \bar{\psi}$, then $\hat{\omega}_3 = L$. This implies $\omega = L$ is a fixed point of ϕ and thus $\hat{\omega}_t = L$ for all $t > 1$. If $\psi < \bar{\psi}(\ell)$, then $\hat{\omega}_3 = H$. As shown in Part 1, $\omega = H$ is a fixed point of ϕ for all values of ψ, meaning $\hat{\omega}_t = H$ for all $t > 1$.

■

Proof of Lemma 3.

Proof. Consider Player $(n,1)$ in $t=1$ with signal $s_{(n,1)}$. Since priors and signals about ω are normal, it follows (e.g., DeGroot 1970) that Player $(n,1)$ has a normal posterior with mean $\widehat{\mathbb{E}}[\omega|I_{(n,1)}] = \frac{\rho_\varepsilon}{\rho_\varepsilon+\rho_\omega}s_{(n,1)} + \frac{\rho_\omega}{\rho_\varepsilon+\rho_\omega}\overline{\omega}$ and variance $\mathrm{Var}[\omega|I_{(n,1)}] = \frac{1}{\rho_\varepsilon+\rho_\omega}$. From Equation 3, individual demand is

$$x_{(n,1)} = \frac{\rho_\varepsilon+\rho_\omega}{2(\rho_\varepsilon+\rho_\omega)+\rho_\eta}\left\{1+\frac{\rho_\eta}{\alpha W_0}\left(\frac{\rho_\varepsilon}{\rho_\varepsilon+\rho_\omega}s_{(n,1)}+\frac{\rho_\omega}{\rho_\varepsilon+\rho_\omega}\overline{\omega}\right)\right\}.$$

Since $\mathbb{E}[s_{(n,1)}] = \omega$, aggregating over all Generation 1 yields

$$\bar{x}_1 = \mathbb{E}[x_{(n,1)}] = \frac{\rho_\varepsilon+\rho_\omega}{2(\rho_\varepsilon+\rho_\omega)+\rho_\eta}\left\{1+\frac{\rho_\eta}{\alpha W_0}\left(\frac{\rho_\varepsilon}{\rho_\varepsilon+\rho_\omega}\omega+\frac{\rho_\omega}{\rho_\varepsilon+\rho_\omega}\overline{\omega}\right)\right\}. \tag{11}$$

It follows that $\bar{x}_1 = \nu_0 + \nu_1\omega$ where

$$\begin{aligned}\nu_0 &= \frac{1}{2(\rho_\varepsilon+\rho_\omega)+\rho_\eta}\left\{\rho_\omega\left(\frac{\rho_\eta}{\alpha W_0}\overline{\omega}+1\right)+\rho_\varepsilon\right\},\\ \nu_1 &= \frac{\rho_\eta}{\alpha W_0}\left(\frac{\rho_\varepsilon}{2(\rho_\varepsilon+\rho_\omega)+\rho_\eta}\right).\end{aligned}$$

■

Proof of Proposition 5.

Proof. The proof makes use of the following lemma, which follows trivially from combining Equations 4 and 5 and using Lemma 3.

Lemma 5. *Starting from the initial condition $\bar{x}_1 = \nu_0 + \nu_1\omega$, aggregate allocations $\langle\bar{x}_t\rangle$ evolve as follows:*

$$\bar{x}_t = \begin{cases} 0 & \text{if} \quad \kappa_0+\kappa\bar{x}_{t-1} \le 0,\\ \kappa_0+\kappa\bar{x}_{t-1} & \text{if} \quad \kappa_0+\kappa\bar{x}_{t-1} \in (0,1),\\ 1 & \text{if} \quad \kappa_0+\kappa\bar{x}_{t-1} \ge 1,\end{cases}$$

where

$$\begin{aligned}\kappa_0 &= -\frac{\rho_\omega}{2\rho_\varepsilon}\left(\frac{\rho_\eta}{\alpha W_0}\overline{\omega}+1\right),\\ \kappa &= 1+\frac{2\rho_\omega+\rho_\eta}{2\rho_\varepsilon}.\end{aligned}$$

Part 1. From Equations 4 and 5, $\hat{\omega}_t$ is monotonic in t if and only if $\bar{x}_t$ is. From Lemma 5, $\bar{x}_t > \bar{x}_{t-1} \Leftrightarrow \kappa_0+\kappa\bar{x}_{t-1} > \bar{x}_{t-1} \Leftrightarrow \bar{x}_{t-1} > -\kappa_0/(\kappa-1) \equiv \bar{x}^* > 0$, where the final inequality follows from Lemma 5. Thus, we need only check whether the initial value $\bar{x}_1 > \bar{x}^*$. If so, then $\bar{x}_2 > \bar{x}_1 > \bar{x}^*$,

and by induction, all $\bar{x}_t > \bar{x}^*$. This implies $\langle \bar{x}_t \rangle$ is increasing. Similarly, if $\bar{x}_1 < \bar{x}^*$, then $\langle \bar{x}_t \rangle$ is decreasing in t. From Lemma 5,

$$\bar{x}^* = \frac{-\kappa_0}{1-\kappa} = \frac{\rho_\omega}{2\rho_\omega + \rho_\eta}\left(\frac{\rho_\eta}{\alpha W_0}\overline{\omega} + 1\right). \tag{12}$$

It follows that $\langle \bar{x}_t \rangle$ is increasing iff $\bar{x}_1 > \bar{x}^* \Leftrightarrow \nu_0 + \nu_1 \omega > \bar{x}^*$. Using Equations 11 and 12, this condition reduces to $\omega > \frac{1}{2\rho_\omega + \rho_\eta}\left(2\rho_\omega \overline{\omega} - \alpha W_0\right) \equiv \omega^*$.

Part 2. We show that the change in allocation across periods $\Delta_x(t) \equiv \bar{x}_t - \bar{x}_{t-1}$ is increasing in magnitude in t. From Lemma 5, $\Delta_x(t) = \bar{x}_t - \bar{x}_{t-1} = \kappa_0 + (\kappa - 1)\bar{x}_{t-1}$, and $\Delta_x(t+1) = \bar{x}_{t+1} - \bar{x}_t = \kappa_0 + (\kappa-1)(\kappa_0 + \kappa \bar{x}_{t-1}) = \kappa[\kappa_0 - (\kappa-1)\bar{x}_{t-1}] = \kappa \Delta_x(t)$. Since $\kappa > 1$, $|\Delta_x(t+1)| > |\Delta_x(t)|$. And because t is arbitrary, $|\Delta_x(t)|$ is increasing in t. As such, $\langle \bar{x}_t \rangle$ converges to either boundary value 0 or 1, depending on whether $\langle \bar{x}_t \rangle$ is increasing or decreasing. Furthermore, because $\Delta_x(t)$ is increasing in t, $\bar{x}_t$ reaches either boundary value 0 or 1 in finite time. Consider the smallest t large enough that the system reaches the upper boundary. Hence, t is the first instance in which $\hat{\omega}_t > \alpha W_0/\rho_\eta$, so $\bar{x}_t = 1$. We show that $\hat{\omega}_{t+1} = \infty$. Note $\bar{x}_t = 1$ implies each player (n,t) selects $x_{(n,t)} = 1$. From Equation 3, if acting in autarky, $x_{(n,t)} = 1 \Leftrightarrow s_{(n,t)} > c$, where $c = (1-\nu_0)/\nu_1$. Hence, $\mathbb{P}_\omega(x=1) = 1 - \Phi\big((c-\omega)/\sqrt{\rho_\varepsilon^{-1}}\big)$, where $\Phi(\cdot)$ is the standard-normal c.d.f.. Since Lemma 1 assumes a finite action and state space, we cannot directly invoke it here. However, the logic naturally extends. The continuous analog of the cross entropy between $\mathbb{P}_\omega(x)$ and $\mathbb{T}_{\hat{\omega}}(x)$ is $-\int_0^1 T_{\hat{\omega}}(x) \log \mathbb{P}_\omega(x)\,\mathrm{d}x$. Note that $\mathbb{T}_{\hat{\omega}_t}(x) = \delta(x-1)$, where $\delta(\cdot)$ is the Dirac delta function. Hence $-\int_0^1 T_{\hat{\omega}_t}(x) \log \mathbb{P}_\omega(x)\,\mathrm{d}x = -\log \mathbb{P}_\omega(1)$, which is minimized at the state in which $\mathbb{P}_\omega(1) = 1 - \Phi\big((c-\omega)/\sqrt{\rho_\varepsilon^{-1}}\big)$ is maximized. Thus $\hat{\omega}_{t+1} = \sup \Omega = \infty$. Furthermore, given this belief, $\mathbb{T}_{\hat{\omega}_{t+1}}(x) = \delta(x-1)$, implying again that $\hat{\omega}_{t+2} = \infty$. Hence, $\hat{\omega}_\tau = \infty$ for all $\tau > t$. The proof that $\hat{\omega}_\tau \to -\infty$ when the system reaches its lower boundary ($\bar{x}_t = 0$) is analogous and omitted.

■

Proof of Corollary 1.

Proof. Since the aggregate shocks are identically distributed for both assets, Asset r dominates Asset s when ω is known if and only if $\mathbb{E}[d_t^r|\omega] > \mathbb{E}[d_t^s|\omega] \Leftrightarrow \omega > 0$. There are two cases to consider: $\omega^* > 0$ and $\omega^* < 0$. First, suppose $\overline{\omega} > \alpha W_0/2\rho_\omega$, so $\omega^* > 0$. Define $\Omega' = (0, \omega^*)$. From Proposition 5, for any value of $\omega \in \Omega'$, $\langle \hat{\omega}_t \rangle$ diverges to $-\infty$. Hence, whenever $\omega \in \Omega'$, investors believe Asset s provides infinitely better payoff than Asset r despite Asset s being the dominated asset. Now consider the case where $\omega < \alpha W_0/2\rho_\omega$, so $\omega^* < 0$. Define $\Omega' = (\omega^*, 0)$. For any realization of $\omega \in \Omega'$, $\langle \hat{\omega}_t \rangle$ diverges to $+\infty$. Hence investors believe it provides infinitely higher payoff than Asset s, despite the fact it is dominated by Asset s.

■

Proof of Proposition 6.

Proof. Part 1. Suppose the state is ω^m. By Assumption 1, $\mathbf{a}_1$ reveals ω^m to Generation 2. Hence, $\mathbf{a}_2$ is such that $a_2(m) = 1$. Following the logic of Proposition 3, $\hat{\omega}_3$ is the state that maximizes $\mathbb{P}_\omega(m)$. Lemma 4 implies that this state is ω^m. Since $\hat{\omega}_3 = \omega^m$ leads Generation 3 to play $a_3(m) = 1$, ω^m is an absorbing state.

Parts 2 and 3. Suppose the state is ω^0. So long as M is finite, which implies that the prior likelihood ratio $\pi_1(\omega^0)/\pi_1(\omega^m) = \frac{\chi^{\frac{1}{M}}}{1-\chi^{\frac{1}{M}}}$ is finite, we can invoke Lemma 1 to determine $\langle\hat{\omega}_t\rangle$. By Assumption 1, $\mathbf{a}_1$ reveals ω^0 to Generation 2. Hence, $\mathbf{a}_2 = \mathbb{T}_{\omega^0} = (1-\lambda, \lambda, 0, ..., 0)$. From Lemma 1, $\hat{\omega}_3 = \phi(\hat{\omega}_2) = \arg\max_{\omega\in\Omega} \prod_{m=0}^{M} \mathbb{P}_\omega^{a_2(m)} = \arg\max_{\omega\in\Omega} \mathbb{P}_\omega(0)^{1-\lambda}\mathbb{P}_\omega(1)^{\lambda}$. Hence, $\hat{\omega}_3$ is the unique state satisfying

$$\left(\frac{\mathbb{P}_\omega(0)}{\mathbb{P}_{\hat{\omega}_3}(0)}\right)^{1-\lambda}\left(\frac{\mathbb{P}_\omega(1)}{\mathbb{P}_{\hat{\omega}_3}(1)}\right)^{\lambda} < 1 \tag{13}$$

for all $\omega \in \Omega \setminus \{\hat{\omega}_3\}$. Given that actions $m = 2, \ldots, M$ are not chosen, it is immediate that $\hat{\omega}_3 \in \{\omega^0, \omega^1\}$. Hence, we simply need to consider when

$$\left(\frac{\mathbb{P}_{\omega^0}(0)}{\mathbb{P}_{\omega^1}(0)}\right)^{1-\lambda}\left(\frac{\mathbb{P}_{\omega^0}(1)}{\mathbb{P}_{\omega^1}(1)}\right)^{\lambda} < 1; \tag{14}$$

if 14 holds, then $\hat{\omega}_3 = \omega^1$, otherwise $\hat{\omega}_3 = \omega^0$. We now consider the cases $M \gtrless \overline{M}$.

Case (i). $M > \overline{M}$. If $M > \overline{M}$, then $\mathbb{P}_\omega(0) = 1-\lambda$ for all ω. Thus, from 14, $\hat{\omega}_3 = \hat{\omega}^1 \Leftrightarrow \mathbb{P}_{\omega^1}(1) > \mathbb{P}_{\omega^0}(1)$, which is true by Lemma 4.

Case (ii). Now suppose $M < \overline{M}$. This means that there exist some actions A_m such that a low type prefers A_m over A_0 conditional on $s^m = 1$. For any $\mathscr{A}^r$, we can index the options such that $\underline{q}^i > \underline{q}^j \Leftrightarrow i < j$. Let m^* be the largest integer $m \leq M$ such that $\mathbb{E}[q^m | s^m = 1] > q_L^0$ (see Equation 6). A low type chooses $A_0 \Leftrightarrow$ she receives $\mathbf{s} = (s^1, \ldots, s^M)$ such $s^m = 0$ for all $m \leq m^*$. Thus $\mathbb{P}_{\omega^0}(0) = (1-\lambda)\rho^{m^*}$ and $\mathbb{P}_{\omega^1}(0) = (1-\lambda)(1-\rho)\rho^{m^*-1}$. Hence $\mathbb{P}_{\omega^0}(0)/\mathbb{P}_{\omega^1}(0) = \rho/(1-\rho)$. To complete the proof, we must determine the relative likelihoods of A_1. Note that type $\theta = L$ takes A_1 iff $s^1 = 1$; type $\theta = H$ takes A_1 iff s^1 or $s^m = 0$ for all $m = 1, \ldots, M$. Hence,

$$\frac{\mathbb{P}_{\omega^0}(1)}{\mathbb{P}_{\omega^1}(1)} = \frac{(1-\rho) + \lambda\rho^M}{\rho + \lambda(1-\rho)\rho^{M-1}}. \tag{15}$$

Since $\rho > 1/2$, it is straightforward that $\mathscr{L}(1|M) \equiv \mathbb{P}_{\omega^0}(1)/\mathbb{P}_{\omega^1}(1)$ is strictly decreasing in M, and converges to $(1-\rho)/\rho$ as $M \to \infty$. From $\mathbb{P}_{\omega^0}(0)/\mathbb{P}_{\omega^1}(0) = \rho/(1-\rho)$ and 14, $\hat{\omega}_3 = \omega^1$ iff

$$\mathscr{L}(1|M) < \left(\frac{1-\rho}{\rho}\right)^{\frac{1-\lambda}{\lambda}}. \tag{16}$$

If $\lambda < 1/2$, then the right-hand side of condition 16 is below $(1-\rho)/\rho$, and thus there exists no M for which 16 holds. If $\lambda > 1/2$, then the right-hand side of Condition 16 exceeds $(1-\rho)/\rho$. Since the sequence $(\mathscr{L}(1|M))_{M=1}^{\infty}$ converges from above to $(1-\rho)/\rho$, there exists an integer $\overline{M}(\lambda)$ such that $M > \overline{M}(\lambda)$ implies $\mathscr{L}(1|M) < \left(\frac{1-\rho}{\rho}\right)^{\frac{1-\lambda}{\lambda}}$. Since $\left(\frac{1-\rho}{\rho}\right)^{\frac{1-\lambda}{\lambda}}$ is strictly increasing in λ, $\overline{M}(\lambda)$ is strictly decreasing in λ.

■

Proof of Proposition 7.

Proof. Suppose the state is ω^0. As in Part 2 of the proof of Proposition 6, if Condition 14 holds, then $\hat{\omega}_3 = \omega^1$, otherwise $\hat{\omega}_3 = \omega^0$. To ease the analysis, we define $\mathscr{L}(m|M) \equiv \mathbb{P}_{\omega^0}(m)/\mathbb{P}_{\omega^1}(m)$ as an explicit function of M. The remainder of the proof characterizes those values of M for which $\mathscr{L}(0|M)^{1-\lambda}\mathscr{L}(1|M)^{\lambda} < 1$.

We first derive $\mathscr{L}(0|M)$. Type $\theta = H$ never chooses A_0. Type $\theta = L$ chooses $A_0 \Leftrightarrow q_L^0 > \mathbb{E}[q^m|s^m]$ for all m. $\mathbb{E}[q^m|s^m] = (1-p^m)\underline{q}^m$ where p^m is the posterior belief that $q^m = 0$ conditional on s^m. Note that

$$p^m = \Pr(q^m = 0|s^m) = \left[1 + \left(\frac{\chi^{\frac{1}{M}}}{1-\chi^{\frac{1}{M}}}\right)\frac{f(s^m|\underline{q}^m)}{f(s^m|0)}\right]^{-1}.$$

Let $\widetilde{F}^m(p|q^m)$ denote the distribution of posterior beliefs p^m conditional on q^m induced by the underlying signal distribution $F^m(s|q^m)$. Since $q_L^0 > \mathbb{E}[q^m|s^m]$ iff p^m is less than threshold $\bar{p}_L^m \equiv \frac{|\underline{q}^m - q_L^0|}{|\underline{q}^m|}$, type L chooses A_0 in state ω^0 with probability

$$\mathbb{P}_{\omega^0}(0) = \prod_{m=1}^{M} \widetilde{F}^m\left(\bar{p}_L^m|\underline{q}^m\right).$$

In state ω^1, this probability is

$$\mathbb{P}_{\omega^1}(0) = \widetilde{F}^1\left(\bar{p}_L^1|0\right)\prod_{m=2}^{M} \widetilde{F}^m\left(\bar{p}_L^m|\underline{q}^m\right).$$

Hence, the likelihood ratio of observing A_0 in ω^0 relative to ω^1 is

$$\mathscr{L}(0|M) = \frac{\mathbb{P}_{\omega^0}(0)}{\mathbb{P}_{\omega^1}(0)} = \frac{\widetilde{F}^1\left(\bar{p}_L^1|\underline{q}^1\right)}{\widetilde{F}^1\left(\bar{p}_L^1|0\right)} > 1,$$

where the inequality follows from MLRP (see Remark 1). Because $\mathscr{L}(0|M)$ is independent of M, we write it simply as $\mathscr{L}(0)$.

We now derive $\mathscr{L}(1|M)$. Type θ chooses A_1 if both $\mathbb{E}[q^1|s^1] > \mathbb{E}[q^m|s^m]$ for all $m > 1$ and

$\mathbb{E}[q^1|s^1] > q^0_\theta$. Note that $\mathbb{E}[q^1|s^1] > \mathbb{E}[q^m|s^m] \Leftrightarrow p^1 > 1 - \left(\frac{q^m}{q^1}\right)(1-p^m) \equiv k_m(p^m)$. This happens with probability

$$\int_0^1 1 - \widetilde{F}^1\big(k_m(p)\big|q^1\big)\,\mathrm{d}F^m(p|q^m),$$

which implies that in state ω^0, type $\theta = H$ chooses A_1 with probability

$$\prod_{m=2}^{M}\int_0^1 1 - \widetilde{F}^1\big(k_m(p)\big|\underline{q}^1\big)\,\mathrm{d}F^m\big(p\big|\underline{q}^m\big),$$

and type $\theta = L$ chooses A_1 with probability

$$\Big(1 - \widetilde{F}^1\big(\bar{p}^1_L\big|\underline{q}^1\big)\Big)\prod_{m=2}^{M}\int_0^1 1 - \widetilde{F}^1\big(k_m(p)\big|\underline{q}^1\big)\,\mathrm{d}F^m\big(p\big|\underline{q}^m\big).$$

Hence,

$$\mathscr{L}(1|M) = \frac{\mathbb{P}_{\omega^0}(1)}{\mathbb{P}_{\omega^1}(1)} = \left(\frac{1-(1-\lambda)\widetilde{F}^1\big(\bar{p}^1_L\big|\underline{q}^1\big)}{1-(1-\lambda)\widetilde{F}^1\big(\bar{p}^1_L\big|0\big)}\right)\prod_{m=2}^{M}\frac{\int_0^1 1-\widetilde{F}^1\big(k_m(p)\big|\underline{q}^1\big)\,\mathrm{d}F^m\big(p\big|\underline{q}^m\big)}{\int_0^1 1-\widetilde{F}^1\big(k_m(p)\big|0\big)\,\mathrm{d}F^m\big(p\big|\underline{q}^m\big)}. \tag{17}$$

Strict MLRP implies that each term in the product above (Equation 17) is bounded below 1. This implies that $\mathscr{L}(1|M)$ is decreasing in M and the sequence $(\mathscr{L}(1|M))_{M=1}^{\infty}$ converges to 0. Let $\overline{M}$ be the smallest integer such that $\mathscr{L}(0)^{1-\lambda}\mathscr{L}(1|\overline{M})^{\lambda} < 1$, and note that $\overline{M}$ is finite since $\mathscr{L}(0)^{1-\lambda}$ is constant in M and $\lim_{M\to\infty}\mathscr{L}(1|M) = 0$.

It follows that whenever $M \geq \overline{M}$, $\hat{\omega}_3 = \omega^1$, which we now show is an absorbing state. If $\hat{\omega}_3 = \omega^1$, then $a_3(1) = 1$. By Lemma 4, such autarikic behavior is most likely in ω^1. Hence, $\hat{\omega}_4 = \omega^1$, implying ω^1 is absorbing. Thus, $M \geq \overline{M}$ implies $\hat{\omega}_t = \omega^1$ for all $t > 2$. If $M < \overline{M}$, then $\hat{\omega}_3 = \omega^0$. Since ω^0 is likewise an absorbing state, $M < \overline{M}$ implies $\hat{\omega}_t = \omega^0$ for all $t > 2$.

■

Proof of Proposition 8.

Proof. Denote the state by $\omega = (\mathbf{q}, \rho)$. Because of uniqueness (Assumption 1), $a_2(m^*) = 1$ where $m^* \equiv \arg\max_m(q^1, \ldots, q^M)$. Following the logic of Proposition 3, $\hat{\omega}_t$ converges to the state that maximizes the autarkic probability of action A_{m^*}, $\mathbb{P}_{(\mathbf{q},\rho)}(m^*)$. Fixing any ρ, Lemma 4 implies that $\mathbb{P}_{(\mathbf{q},\rho)}(m^*)$ is maximized at $\mathbf{q} = \omega_e^{m^*}$. Thus, we need only show that $\mathbb{P}_{(\mathbf{q},\rho)}(m^*)$ is increasing in ρ conditional on $\mathbf{q} = \omega_e^{m^*}$. Without loss of generality, let $m^* = 1$, and let $F^m(s|q^m,\rho)$ denote the conditional c.d.f. of s^m. Since A_1 is selected in autarky iff $s^1 > s^m$ for all $m \neq 1$, A_1 is chosen in state (ω_e^1, ρ) with probability

$$\mathbb{P}_{(\omega_e^1,\rho)}(1) = \prod_{m=2}^{M} \Pr\left(s^1 > s^m \middle| \omega_e^1, \rho\right) = \prod_{m=2}^{M} \int_{-\infty}^{\infty} 1 - F^1\left(s^m \middle| \overline{q}, \rho\right) \mathrm{d}F^m\left(s^m \middle| \underline{q}, \rho\right),$$

where $\overline{q} \equiv \max Q$ and $\underline{q} \equiv \min Q$. Hence $\mathbb{P}_{(\mathbf{q},\rho)}$ is increasing in ρ if for every $m \geq 2$,

$$\int_{-\infty}^{\infty} F^1\left(s^m \middle| \overline{q}, \rho\right) \mathrm{d}F^m\left(s^m \middle| \underline{q}, \rho\right) \tag{18}$$

is decreasing in ρ. Letting F^{ε} denote the c.d.f. of random variable ε and noting that $s^m = \underline{q} + \varepsilon/\sqrt{\rho}$ for each $m \geq 2$ conditional on ω_e^1, the term in 18 can be written as

$$\int_{-\infty}^{\infty} F^{\varepsilon}\left(\varepsilon - \sqrt{\rho}(\overline{q} - \underline{q})\right) \mathrm{d}F^{\varepsilon}(\varepsilon).$$

This expression is decreasing in ρ since $F^{\varepsilon}(\cdot)$ is increasing and $\overline{q} - \underline{q} > 0$.

■