Learning with Heterogeneous Misspecified Models: Characterization and Robustness*

J. Aislinn Bohren† Daniel N. Hauser‡

February 3, 2021
First version: May 15, 2017

This paper develops a general framework to study how misinterpreting information impacts learning. Our main result is a simple criterion to characterize long-run beliefs based on the underlying form of misspecification. We present this characterization in the context of social learning, then highlight how it applies to other learning environments, including individual learning. A key contribution is that our characterization applies to settings with model heterogeneity and provides conditions for entrenched disagreement. Our characterization can be used to determine whether a representative agent approach is valid in the face of heterogeneity, study how differing levels of bias or unawareness of others' biases impact learning, and explore whether the impact of a bias is sensitive to parametric specification or the source of information. This unified framework synthesizes insights gleaned from previously studied forms of misspecification and provides novel insights in specific applications, as we demonstrate in settings with partisan bias, overreaction, naive learning, and level-k reasoning.

KEYWORDS: Model misspecification, Social learning
JEL: C73, D83

*This paper was previously circulated under the titles "Bounded Rationality and Learning: A Framework and a Robustness Result" and "Social Learning with Model Misspecification: A Framework and a Characterization." We thank Nageeb Ali, Mira Frick, Drew Fudenberg, Alex Imas, Brian Kovak, Ryota Iijima, Shuya Li, George Mailath, Margaret Meyer, Pauli Murto, Wojciech Olszewski, Pietro Ortoleva, Ali Polat, Andrew Postlewaite, Andrea Prat, Yuval Salant, Larry Samuelson, Joel Sobel, Ran Spiegler, Philipp Strack, Juuso Toikka, Juuso Välimäki, Rakesh Vohra, Yuichi Yamamoto and conference and seminar participants for helpful comments and suggestions. Cuimin Ba provided excellent research assistance. Bohren gratefully acknowledges financial support from NSF grant SES-1851629 and the briq Institute.

†University of Pennsylvania; Email: abohren@sas.upenn.edu
‡Aalto University and Helsinki GSE; Email: daniel.hauser@aalto.fi

1 Introduction

How do individuals learn when they misinterpret information? The literature on misspecified learning typically takes the following approach: fix an incorrect, or *misspecified*, model—such as overreaction to signals or a failure to account for correlated information—and explore how it impacts the long-run beliefs about the state. We know from this literature that a misspecified model may lead to *incorrect learning*, where beliefs converge to the wrong state, *cyclical learning*, where beliefs do not converge, *entrenched disagreement*, where agents with different models become certain of different states, and *path-dependent learning*, where multiple limit beliefs can arise—for example, correct and incorrect learning.[1]

In deriving these insights, studies usually consider a parameterized misspecified model that captures the cognitive bias or heuristic of interest and assume that all individuals have an identical level of this same bias.[2] But multiple parameterizations can often capture a given cognitive error, and these different parameterizations may yield different predictions about asymptotic learning. Further, there may be model heterogeneity, either because agents exhibit varying levels of the same bias, have fundamentally distinct biases, or use different heuristics. This raises the question of whether it is valid to use a representative agent approach or consider a single bias in isolation. Finally, a given form of misspecification may have a different impact on learning depending on whether the source of information is private, public, or social.

This paper develops a general framework to study how misinterpreting information impacts learning. A central contribution of this framework is the ability to allow for model heterogeneity. Our main result is a simple criterion to characterize long-run beliefs and behavior based on the underlying form of misspecification. We present this characterization in the context of a social learning environment in which individuals observe private signals and the action choices of predecessors and critically, have misspecified models of how to interpret these sources. We then highlight how our characterization applies to other learning environments, including active individual learning settings and settings with different sources of information (e.g. public signals, social outcomes).

This characterization provides a deeper understanding of how misspecification influences learning and can be used to address the issues raised above. Specifically, it can determine

[1] Berk (1966) showed that model misspecification can lead to incorrect learning. When information depends on beliefs—as is the case when agents learn from their peers, their models vary with the history, or their actions influence future signals—misspecification can also give rise to cyclical learning (e.g. Nyarko (1991)) and path-dependent learning (e.g. (Rabin and Schrag 1999)), while when agents have heterogeneous models, entrenched disagreement can emerge (e.g. Gagnon-Bartsch (2016)). See the Related Literature section for additional references.

[2] Notable exceptions include Ortoleva and Snowberg (2015), which allows for heterogeneous levels of correlation neglect, and Gagnon-Bartsch (2016); Frick, Iijima, and Ishii (2020a), in which agents exhibit the false consensus effect about other agents' preferences or population characteristics.

whether a class of misspecified models are *robust*—in that the learning predictions are not sensitive to parametric specification and similar levels of a bias lead to similar learning outcomes—without needing to analyze variations on a case-by-case basis. Such robustness ensures that knowledge of the exact parametric form and level of a bias are not necessary to accurately predict its impact on learning. When agents exhibit varying levels of the same bias, the characterization can be used to evaluate whether a representative agent model is a good approximation, and in the case where it is not, to determine how such heterogeneity impacts learning. When multiple biases coexist, it can be used to study how agents with different models influence each others' learning. Using a common framework to encompass multiple learning environments provides insight into how the impact of misspecification varies with the source of information.

Our framework captures a rich array of ways in which individuals are biased when processing information and interpreting others' choices. Depending on the context, the empirical literature in psychology and economics has documented that individuals exhibit behavioral biases such as systematically overreacting or underreacting to new information, slanting information towards a preferred state (e.g. partisan bias), selectively weighting information (e.g. confirmation bias), incorrectly aggregating correlated information, or misperceiving others' preferences and beliefs (e.g. false consensus effect, pluralistic ignorance), and use simplifying decision rules such as the counting and social-circle heuristics.[3] Our framework represents these and other cognitive biases and heuristics as misspecified models where individuals have incorrect models of the signal distribution, others' preferences, and how others interpret information. By studying the impact of different biases within a unified framework, our characterization can synthesize the insights gleaned from papers that focus on a single bias and can link seemingly distinct biases that have a similar effect on behavior.[4]

The details of our framework are as follows. In the social learning environment, a sequence of individuals learn about a binary state from the actions of their predecessors and a private signal, then select an action; each agent's payoff depends on this action and the state. In the individual learning environment, a single myopic agent observes a sequence of signals

[3]See Section 2.3 for relevant citations and the details of how our framework models these biases. Biases may be due to systematic errors or may emerge endogenously due to cognitive limitations (e.g. bounded memory (Wilson 2014)) or belief-based utility (e.g. desire to appear competent, anticipatory utility (Brunnermeier and Parker 2005; Kőszegi 2006; Gottlieb 2015)). The context of the learning setting determines which biases are of first order relevance.

[4]Our framework nests several prior behavioral models of learning. Section 4.2 shows how our framework nests naive learning in Bohren (2016). Online Appendix E demonstrates how it nests the parameterization of confirmation bias in Rabin and Schrag (1999) and the non-Bayesian learning rule in Epstein, Noor, and Sandroni (2010). More generally, our framework can be used to study heuristics and biases that reduce to Markovian updating rules. It cannot nest heuristics that reduce to non-Markovian updating rules or are calibrated based on equilibrium objects (e.g. the analogy-based expectation equilibria in Guarino and Jehiel (2013)).

and acts repeatedly; future information can depend on the history, which can capture, for example, active learning or a history-dependent model of inference. In both settings, an agent's type specifies preferences and a model of inference. This includes a subjective belief about the signal distribution, which determines how she interprets signals, and a subjective belief about the type distribution, which determines how she interprets actions via her beliefs about others' preferences and models of inference. Agents are Bayesian learners with respect to their subjective distributions. Model misspecification refers to the case where these subjective distributions *differ* from the true distributions.

In order to derive meaningful predictions, our framework requires additional structure on how agents interpret signals and actions. First, we focus on *aligned* type spaces in which all agents share a common ordinal ranking of signals and actions in terms of which are stronger evidence for a given state. Second, we focus on environments that have *uniformly informative actions*, in that for each state, there is an action that occurs with higher probability in this state at all possible beliefs. When all agents have correctly specified models, these assumptions ensure that learning is almost surely *correct*, in that beliefs converge to the realized state—they rule out confounded learning and informational herds (Banerjee 1992; Bikhchandani, Hirshleifer, and Welch 1992; Smith and Sørensen 2000).

This framework resolves several challenges that arise when incorporating model heterogeneity. First, model heterogeneity can lead to complicated higher-order beliefs. For example, when an agent believes that others have a misspecified signal distribution, it is also necessary to specify what the agent believes these misspecified agents believe about others, and so on. In our framework, types serve as a modeling tool to represent agents' hierarchies of beliefs; higher order beliefs are fully determined by the subjective type distributions. Second, heterogeneity leads to a more complex learning process, as agents with different models have different beliefs following the same history. These beliefs determine their action choices. Therefore, when agents learn from others' actions, the informational content of the history depends on a vector of beliefs. Our framework provides substantial added structure and tractability for analyzing this multidimensional belief process.

Using this framework, we explore when the asymptotic learning outcomes described above—correct, incorrect, cyclical, entrenched disagreement and path-dependent—arise. Our main result (Theorem 4) characterizes the set of learning outcomes that arise with positive probability based on two expressions that are straightforward to derive from the primitives of the misspecification: (i) the difference between the Kullback-Leibler divergence from a type's perceived action distribution in each state at a candidate learning outcome to the true action distribution in the realized state at this learning outcome; and (ii) an ordering over the type space—*maximal accessibility*—based on each type's perceived action distributions at certain beliefs (i.e. all types have a degenerate belief on one of the states).

To establish Theorem 4, we first determine whether a belief is locally stable in that the belief process converges to it with positive probability from nearby beliefs. Building on techniques used in Smith and Sørensen (2000) and Bohren (2016), we use the first expression described above to derive a necessary and sufficient condition for local stability (Theorem 1). Intuitively, at a given vector of beliefs, a type's belief process moves towards the state that is more likely to generate the observed pattern of actions at that vector. The difference between the Kullback-Leibler divergences at a candidate degenerate belief determines whether this is the case in a neighborhood of the belief, and therefore, whether the belief is locally stable. We also show that non-degenerate beliefs cannot be locally stable—this stems from our focus on aligned environments with uniformly informative actions. Therefore, each type is certain about the state at any locally stable belief. As discussed below, the set of locally stable beliefs correspond to the set of strict Berk-Nash equilibria (Esponda and Pouzo 2016).

We next determine whether a locally stable learning outcome is globally stable, in that beliefs converge to it with positive probability from any initial belief. This step follows immediately from local stability for the learning outcomes in which all types have correct or incorrect learning (Theorem 2). But in learning outcomes in which types have different limit beliefs, i.e. disagreement outcomes, even if the outcome is locally stable, it may not be possible to separate the beliefs of different types and push them to a neighborhood of the disagreement outcome. Maximal accessibility—the second expression described above—is sufficient to do this (Theorem 3).

Taken together, these local and global stability results establish Theorem 4. An important feature of our characterization is that the two expressions we outline only need to be verified at a *finite* set of beliefs—that is, the set of certain beliefs. When the informational content of the history depends on the belief for each type, in principle, the asymptotic properties of beliefs could depend on the dynamics of beliefs across the entire belief space. Therefore, this feature significantly simplifies the calculations required to use the characterization in specific settings. Given a particular form of misspecification, the expressions are straightforward to verify.

From this characterization, we see that model heterogeneity has several important implications for learning that are distinct from settings with a single type. First, entrenched disagreement can arise within a population that observes a common history (see Section 4.3 for an application). This arises despite our focus on aligned type spaces, which ensures that agents have a common interpretation of the relative informational content of signals and actions. Therefore, model heterogeneity provides an explanation for how connected populations observing shared sources can perpetually disagree. Second, cognitive biases impact the learning of agents who are not inherently biased but are misspecified due to their unawareness of others' biases. Such unawareness can have an equally severe impact on learning as the bias itself (see Example 2).

We use our characterization to explore whether learning predictions are robust. We show that, except for knife-edge cases, nearby misspecified environments will have the same set of learning outcomes (Theorem 5). Further, learning is almost surely correct when agents have approximately correct models (Theorem 6). These results strengthen the applicability of correctly specified environments to real-world settings with mild biases. They also establish that small errors on the part of a researcher in modeling or measuring biases will not significantly alter the predicted learning outcomes. In contrast, Frick, Iijima, and Ishii (2020b,c) show that correctly specified environments are not robust in settings with either private actions and an infinite state space or that violate our uniformly informative actions assumption.[5]

Given our characterization, we return to the issues raised in the second paragraph in the context of specific biases. In Section 4.1, we demonstrate that overreaction has a qualitatively different impact based on whether agents learn from a private or social source: when agents learn from their peers, it can lead to cyclical learning, while when individuals learn directly from signals, learning is almost surely correct. In contrast, Epstein, Noor, and Sandroni (2008) find that a different parameterization of overreaction leads to incorrect learning when agents observe signals direcly, suggesting that overreaction is sensitive to modeling choice. In Section 4.2, we show that a representative agent model is a good approximation when agents fail to account for redundant information at a similar level, but when there is sufficient heterogeneity in their bias, the representative agent model will underestimate the set of parameters that lead to correct learning. In Section 4.3, we show how agents using different levels of reasoning impact each others' learning outcomes. The presence of higher level agents can lead to different learning outcomes for level-2 agents (i.e. naive learners) relative to settings that consider the impact of naive learning in isolation (e.g. Eyster and Rabin (2010); Bohren (2016)).

Related Literature. As discussed above, the set of locally stable beliefs in our learning characterization are strict Berk-Nash equilibria, the solution concept for agents with misspecified models developed in Esponda and Pouzo (2016). In a Berk-Nash equilibrium, agents play optimally with respect to the model that is the best fit, in terms of the Kullback-Leibler divergence under the equilibrium strategy profile.[6]

A rich literature explores when the learning outcomes discussed above arise for specific forms of misspecification. In an individual learning setting, selective attention (Schwartzstein

[5] Madarász and Prat (2016) find a failure of robustness in an agency setting, which stems from the interaction between misspecification and incentives.

[6] Arrow and Green (1973) provided the first equilibrium framework that explicitly distinguished between the true model and agents' subjective models in a setting with a misspecified oligopolist. Other solution concepts for specific forms of misspecification include Eyster and Rabin (2005); Jehiel (2005); Spiegler (2016).

2014) and misattribution of reference dependence (Bushong and Gagnon-Bartsch 2019) lead to incorrect learning almost surely, confirmation bias (Rabin and Schrag 1999) and overreaction to signals (Epstein et al. 2010) lead to path-dependent learning, overconfidence in one's ability leads to inefficiently low effort (Heidhues, Kőszegi, and Strack 2018), and misspecified prior beliefs (Nyarko 1991; Fudenberg, Romanyuk, and Strack 2017) lead to cyclical learning. In contrast, underreaction to signals leads to correct learning almost surely (Epstein et al. 2010).

Turning to social learning, in the canonical binary state sequential environment, underestimating redundant information leads to correct and incorrect learning (Eyster and Rabin 2010; Bohren 2016), while overestimating redundant information leads to cyclical learning (Bohren 2016). In a variation of this canonical environment, underestimating redundant information leads to incorrect learning almost surely or cyclical learning (Gagnon-Bartsch and Rabin 2016). Misinterpreting others' preferences (Frick et al. 2020b) and the gambler's fallacy (He 2020) lead to incorrect learning almost surely; misinterpreting others' preferences can also lead to cyclical learning (Gagnon-Bartsch 2016; Bohren and Hauser 2019a) or entrenched disagreement (Gagnon-Bartsch 2016).[7] In contrast, coarse reasoning (Guarino and Jehiel 2013)—which also results in underestimating redundant information—or a linear updating heuristic that puts sufficient weight on agents' own signals (Jadbabaie, Molavi, Sandroni, and Tahbaz-salehi 2012) lead to correct learning almost surely. By capturing multiple forms of model misspecification and learning environments within the same framework, our analysis provides a tool to unify some of these insights.

A recent set of papers explore convergence in more general misspecified learning environments.[8] For the most part, this work focuses on active individual learning settings. Fudenberg et al. (2017) characterize long-run beliefs for an agent who learns about a binary state from a diffusion process with drift that depends on the state and current action. They use this characterization to illustrate how learning outcomes can differ for myopic and patient misspecified agents. Esponda and Pouzo (2019) characterize steady state behavior for a class of Markov decision problems, while Heidhues, Kőszegi, and Strack (2019) derive convergence results in a setting with Gaussian signals and state. Fudenberg, Lanzani, and Strack (2020); Esponda, Pouzo, and Yamamoto (2019) characterize properties of the limiting action distribution when the agent is non-myopic and the state space is infinite. The former show that if actions converge, then they must converge to a refinement of Berk-Nash equilibrium. The latter characterize the long-run action distribution in terms of the solutions to

[7] In other settings, correlation neglect leads to inefficient risk-taking (Levy and Razin 2015) and ideological extremeness (Ortoleva and Snowberg 2015)

[8] An older statistics literature on model misspecification characterizes limiting beliefs in terms of the Kullback-Leibler divergence (Berk 1966; Shalizi 2009). These papers do not apply to active and social learning settings, as the signal process is exogenous, or to settings where an agent's model varies with the history, as the model is fixed across time.

a generalization of a differential equation, providing insight into which action distributions arise when actions fail to converge. In a social learning setting, Molavi, Tahbaz-Salehi, and Jadbabaie (2018) study information aggregation when agents share beliefs on a network and treat neighbors' current beliefs as sufficient statistics for the history. They nest common rules to aggregate beliefs on a network, including the canonical DeGroot model. Our paper complements this work by focusing on the asymptotic properties of social learning environments in which agents use heuristics or have biases that can be captured by misspecified Bayesian updating and these misspecified models may differ across agents. Frick et al. (2020c) build on our results to explore convergence in settings that have a finite number of states and can violate our uniformly informative property. The technical challenges that arise when there are more than two states are similar to those that arise from model heterogeneity in our setting. Relaxing uniform informativeness necessitates new methods to characterize local stability.

A complementary literature proposes explanations for how misspecification can persist despite infinite data that contradicts the model. Gagnon-Bartsch, Rabin, and Schwartzstein (2018) show that limited attention causes agents to ignore information that would overturn their models. Kominers, Mu, and Peysakhovich (2020) consider a setting where updating via Bayes rule is costly, and therefore, agents do not always update after observing new information. In Ba (2021), an agent switches to an alternative model only if the alternative model fits the data significantly better than the status quo model.[9]

The paper proceeds as follows. Section 2 sets up the model, Section 3 presents the analysis, Section 4 develops several applications, and Section 5 concludes. Proofs for Section 3 are in Appendix A and proofs for Section 4 are in Online Appendix C.

2 A General Framework

We first introduce a general framework for social learning, and then discuss how to adapt it to individual learning. We conclude the section with several examples of settings that our framework captures. A reader who prefers to skip the microfoundation for the learning environment can jump to the reduced form stochastic process we analyze in Section 3.

2.1 The Model: Social Learning

States and Actions. Nature selects one of two payoff-relevant states of the world $\omega \in \{L, R\}$ at the beginning of the game according to prior $p_0 \equiv Pr(\omega = R)$. A countably infinite set of agents $t = 1, 2, ...$ act sequentially and choose an action $\tilde{a}_t$ from a finite set $\mathcal{A}$ with

[9] A related set of papers provide a foundation for non-Bayesian updating and model misspecification. Ortoleva (2012) axiomatizes non-Bayesian updating rules in which agents deviate from Bayes rule when reacting to "unexpected" news and Cripps (2018) axiomatizes rules that are independent of how information is partitioned. Frick et al. (2020a) show that the false consensus effect can arise when agents' beliefs are derived from local interactions in an assortative society.

$M \equiv |\mathcal{A}| \geq 2$ actions.[10] Let $h_t \equiv (\tilde{a}_1, ..., \tilde{a}_{t-1})$ denote the publicly observable action history.

Signals. Agents learn about the state from private information and the actions of other agents. Given state ω, agent t observes signal realization $\tilde{s}_t \in [0, 1]$ governed by conditional c.d.f. F^ω, independently of the signal realizations of other agents. No signal perfectly reveals the state: F^L and F^R are mutually absolutely continuous with common support $\mathcal{S}$. Therefore, there exists a positive finite Radon-Nikodym derivative dF^R/dF^L. At least some signals are *informative*, which rules out $dF^R/dF^L = 1$ almost surely. As is conventional, normalize the signal realization to be the posterior probability that the state is R following a neutral prior, i.e. $s = 1/(1 + dF^L/dF^R(s))$ for all $s \in \mathcal{S}$.

Types. Each agent has a privately observed type $\tilde{\theta}_t \in \Theta$ drawn independently from distribution $\pi \in \Delta(\Theta)$, where $\Theta \equiv (\theta_1, ..., \theta_n)$ is a non-empty finite set. A type specifies preferences and a model of inference.[11]

Preferences. Type θ_i earns payoff $u_i(a, \omega)$ from choosing action a in state ω, where $u_i : \mathcal{A} \times \{L, R\} \to \mathbb{R}$. Given probability p that the state is R, the type chooses the action that maximizes its expected payoff $(1 - p)u_i(a, L) + pu_i(a, R)$. Assume that at least two actions are not weakly dominated, no two actions yield the same payoff in both states, and no action is optimal at a single belief. Without loss of generality, assume no action is dominated for all types.

Model of Inference. A type interprets signals and actions using its subjective model of the world. Type θ_i has a subjective signal distribution in each state, represented as conditional c.d.f. $\hat{F}_i^\omega$ in state ω, and subjective type distribution $\hat{\pi}_i \in \Delta(\Theta)$.[12] It believes that no signal perfectly reveals the state: $\hat{F}_i^L$ and $\hat{F}_i^R$ are mutually absolutely continuous, with full support on $\mathcal{S}$ to ensure realized signals are consistent with θ_i's model. Given signal s, let $\hat{s}_i(s) \equiv 1/(1 + d\hat{F}_i^L/d\hat{F}_i^R(s))$ denote θ_i's subjective belief that the state is R following a neutral prior.[13] All types share common prior p_0 that the state is R. It is straightforward to allow for type-specific prior beliefs or a model of inference that varies with the belief about the state (see Online Appendix E for the latter extension). Agents do not update their models of inference—we take these models as fixed and explore how they impact learning about features that are directly payoff-relevant, i.e. the state.

[10] We maintain the convention that a_i or a corresponds to an arbitrary element of $\mathcal{A}$ and $\tilde{a}_t$ corresponds to a random variable with support $\mathcal{A}$, with analogous notation for subsequent random variables.

[11] While we assume types are private, suitably defining the action space and preferences so that each type chooses distinct actions can render types observable.

[12] We implicitly restrict attention to forms of signal misspecification in which signals that map to the same true posterior also map to the same subjective posterior. In this case, it is without loss of generality to define subjective signal distributions with respect to signals normalized as private beliefs (Bohren and Hauser 2020)).

[13] We can also take $\hat{s}$ as a primitive: for any strictly increasing function $\hat{s} : \mathcal{S} \to [0, 1]$ with $\hat{s}(\inf \mathcal{S}) < 1/2$ and $\hat{s}(\sup \mathcal{S}) > 1/2$, there exists a pair of mutually absolutely continuous probability measures with full support on $\mathcal{S}$ that are represented by $\hat{s}$ (see (Bohren and Hauser 2020)).

A *correctly specified* type has a correct model of inference, $(\hat{F}_i^L, \hat{F}_i^R) = (F^L, F^R)$ and $\hat{\pi}_i = \pi$, while a *misspecified* type has an incorrect model, $(\hat{F}_i^L, \hat{F}_i^R) \neq (F^L, F^R)$ and/or $\hat{\pi}_i \neq \pi$. We group types into three additional categories. A *noise* type does not learn: it believes that signals and actions are uninformative, $\hat{F}_i^L = \hat{F}_i^R$ and all agents are noise types, $\hat{\pi}_i(\Theta_N) = 1$, where Θ_N denotes the set of noise types. An *autarkic* type learns from its signal but not the history: it believes that signals are informative, $\hat{F}_i^L \neq \hat{F}_i^R$ and all agents are noise types, $\hat{\pi}_i(\Theta_N) = 1$. To avoid the case in which an autarkic type is observationally equivalent to a noise type, assume that an autarkic type has preferences such that there are at least two strictly optimal actions on the set of posterior beliefs that arise from its subjective signal distribution. A *social* type learns from the history: it believes that actions are informative. Let Θ be ordered such that the first k types are social, denoted by $\Theta_S \equiv (\theta_1, ..., \theta_k)$, and the remaining $n-k$ types are noise or autarkic, denoted by Θ_A and Θ_N, respectively. A learning environment (Θ, π) is *correctly specified* if all social types are correctly specified and *misspecified* if at least one social type is misspecified.

When agents have different models of inference, this can lead to complicated higher-order beliefs. For example, when an agent believes that other agents have a misspecified signal distribution, we also need to model what the agent believes these misspecified agents believe about others. In our framework, higher-order beliefs are fully determined by the subjective type distributions. If type θ_i believes that all agents are type θ_j, then θ_j's subjective type distribution $\hat{\pi}_j$ captures θ_i's second order beliefs, the subjective type distributions of the types in the support of $\hat{\pi}_j$ capture third order beliefs, and so on. Therefore, in addition to describing the types that actually exist, Θ may contain types that serve as a tool to represent hierarchies of beliefs—in other words, types that occur with positive probability under a type's subjective distribution but with probability zero under the true distribution.[14]

Aligned Environments. In order to derive meaningful predictions, our framework requires some structure on how agents interpret signals and actions. We focus on environments that are *aligned*, in that it is common knowledge that agents have the same ordinal ranking of signals and actions in terms of which are stronger evidence for state R. For signals, this corresponds to subjective signal distributions that satisfy the following assumption.

Assumption 1 (Aligned Subjective Signals)**.** *For each $\theta_i \in \Theta$, the subjective signal distribution is either* aligned *with the true signal distribution, i.e. for any $s, s' \in \mathcal{S}$ such that $s > s'$, then $\hat{s}_i(s) > \hat{s}_i(s')$ or uninformative, i.e. $\hat{s}_i(s) = 1/2$ for all $s \in \mathcal{S}$.*

In other words, for any two signals s and s', if s is stronger evidence for state R than s' under the true measure, then s is also stronger evidence for state R than s' under the

[14] Mertens and Zamir (1985) construct the universal type space, which is the set of hierarchies of beliefs that satisfy certain consistency requirements. Finiteness combined with subsequent restrictions we impose on Θ restrict the set of belief hierarchies we analyze to a subset of the universal type space.

subjective measure. We make one exception to allow for noise types who believe signals are uninformative. Types can differ in the perceived *strength* of signals—both relative to other types and to the true distribution. For example, all agents can believe that lung cancer is stronger evidence that smoking is harmful than shortness of breath, but differ in their perceived strength of these two signals. Assumption 1 implies common knowledge that signals are aligned, since all agents believe that other agents have a type in Θ, and so on.

For actions, we assume that, when the state is known, each type has the same ordinal ranking over its undominated actions. Types may have different sets of undominated actions, and therefore, choose different actions when the state is known.

Assumption 2 (Aligned Preferences)**.** *The set of types Θ have* aligned *preferences, in that there exists a complete order $\succ$ on $\mathcal{A}$ such that if $a \succ a'$, then for each $\theta_i \in \Theta$, either $u_i(a, R) > u_i(a', R)$ or a is dominated for θ_i.*[15]

For example, all agents prefer a risky asset in one state and a safe asset in the other, but differ in the belief at which they are willing to start investing in the risky asset or some agents prefer less risky assets across all beliefs about the state. Given Assumption 2, we maintain a complete order over the action space by relative preference in state R.[16] Fixing such an order, index actions $\mathcal{A} \equiv (a_1, ..., a_M)$ so that $a_m \succ a_l$ iff $m > l$.

While signal and preference alignment are not necessary, they are a simple yet general set of restrictions that allow us to derive sharp predictions in a broad class of learning environments. These restrictions do rule out some natural economic settings—for example, some versions of horizontally differentiated environments (e.g. the horizontally differentiated preferences in Gagnon-Bartsch (2016)). In Section 3.6, we discuss how our techniques can be applied to environments that are not aligned.

Informative Actions and Consistent Histories. We focus on environments that are *uniformly informative*, in that for each state there is an action that occurs and is perceived to occur with higher probability in this state regardless of the history. Since an autarkic type believes its signal is informative and does not observe the history, its action choice depends on its private signal at all histories. Therefore, as we show in Section 3.1, the following simple condition—combined with aligned signals and preferences—is sufficient to establish that a_1 is *uniformly informative* of state L—that is, it occurs with higher probability in state L at all possible beliefs—and similarly, a_M is uniformly informative of state R.

Assumption 3 (Informative Actions)**.** *For actions $a \in \{a_1, a_M\}$, there exists an autarkic type $\theta_j \in \Theta_A$ with $\pi(\theta_j) > 0$ that plays a with positive probability, and each social type*

[15] For any undominated actions a and a', if $u_i(a, R) > u_i(a', R)$, then $u_i(a, L) < u_i(a', L)$. Therefore, if $a \succ a'$, then for each $\theta_i \in \Theta$, either $u_i(a', L) > u_i(a, L)$ or a' is dominated for θ_i.

[16] This order may not be unique since Assumption 2 places no restriction on how a type ranks its dominated actions or actions that are optimal for a single type.

$\theta_i \in \Theta_S$ believes that such an autarkic type exists.

Alternative assumptions can also establish uniform informativeness. Our analysis carries through unchanged provided at least one action is uniformly informative of each state. Further, uniform informativeness does not need to hold with respect to actions—for example, our analysis also applies if there is an alternative source that is uniformly informative. We discuss this further in Section 3.1.

We also focus on settings in which the realized history is consistent with each type's model of inference. To rule out the possibility that a type observes what it believes to be a zero probability history, we assume that social types believe that there is an autarkic or noise type that plays each action with positive probability.

Assumption 4 (Consistent Histories). *For each $a \in \mathcal{A}$ and for each social type $\theta_i \in \Theta_S$, there exists an autarkic or noise type $\theta_j \in \Theta_A \cup \Theta_N$ with $\hat{\pi}_i(\theta_j) > 0$ that plays a with positive probability.*

This ensures that each social type believes that all histories are on the equilibrium path. Any learning environment can be slightly perturbed so that it satisfies Assumptions 3 and 4 by adding such an autarkic or noise type with arbitrarily small probability.

Timing. At time t, agent t realizes its type $\tilde{\theta}_t$ and observes the history h_t and private signal $\tilde{s}_t$, then chooses action $\tilde{a}_t$. Then the history is updated to h_{t+1}.

In Section 3.6, we discuss possible extensions to our framework, including misaligned type spaces, heterogeneous signal distributions, and state-dependent type distributions.

2.2 The Model: Individual Learning

Our framework can capture misspecified learning with a single long-run agent by modifying the learning environment so that the signal process is public. Suppose $\mathcal{S}$ is finite and otherwise maintain the same assumptions on the true signal process as in Section 2.1. A single type has subjective signal distributions $\hat{F}^L$ and $\hat{F}^R$ that are mutually absolutely continuous with full support on $\mathcal{S}$. Replace Assumption 3 with the assumption that signals are perceived as informative, $d\hat{F}^R/d\hat{F}^L \neq 1$. When there is a single type, Assumptions 1, 2 and 4 are unnecessary and the type distribution is trivial. Allowing the perceived signal distributions $\hat{F}^L$ and $\hat{F}^R$ to vary with the belief about the state captures cognitive biases such as *confirmation bias*, nesting Rabin and Schrag (1999), and certain forms of *under-/overreaction*, nesting Epstein et al. (2010) (see Online Appendix E). Allowing the true distributions F^L and F^R to vary with the belief about the state captures an active learning model in which action choices influence information. In Section 3.1, we show how this individual learning set-up reduces to a belief process that satisfies the same properties as the social learning set-up, and therefore, we can also apply our learning characterization to this setting.

2.3 Examples

The following examples demonstrate how our framework can capture many information-processing biases, heuristics, and other misspecified models of inference.

Partisan Bias. A type systematically slants signals towards one state (Bartels 2002; Jerit and Barabas 2012). For example, an R-partisan type interprets all signals as being stronger evidence for state R than is actually the case, $\hat{s}(s) = s^{\nu}$.

Under-/Overreaction. A type under- or overreacts to signals (Moore and Healy 2008; Moore, Tenney, and Haran 2015; Angrisani, Guarino, Jehiel, and Kitagawa 2020). For example, $\frac{\hat{s}(s)}{1-\hat{s}(s)} = (\frac{s}{1-s})^{\nu}$, where $\nu \in [0,1)$ corresponds to underreaction and $\nu \in (1,\infty)$ corresponds to overreaction.

Correlation Neglect/Naive Learning. A type underestimates the correlation in the actions of prior agents: the true share of autarkic types is $\pi(\Theta_A)$, but the type believes that the share of autarkic types is $\hat{\pi}(\Theta_A) > \pi(\Theta_A)$ (Eyster and Rabin 2010; Bohren 2016; Enke and Zimmermann 2019; Eyster, Rabin, and Weizsäcker 2020). The *counting heuristic*, where agents simply count actions to form beliefs, provides a foundation for this bias (Ungeheuer and Weber 2020).

Level-k/Cognitive Hierarchy. Level-1 is an autarkic type that believes all agents are noise types; level-2 believes all agents are level-1 and interprets each prior action as reflecting an independent private signal; level-3 believes all agents are level-2, and so on (Costa-Gomes and Crawford 2006; Penczynski 2017). The cognitive hierarchy model is similar, but allows for richer beliefs over types (Camerer, Ho, and Chong 2004).

False Consensus Effect. A type overweights the likelihood that others have similar preferences or models of inference (Ross, Greene, and House 1977; Marks and Miller 1987; Gagnon-Bartsch 2016). For example, there are two types with preferences $u_1 \neq u_2$. Both types believe all agents share their preferences, $\hat{\pi}_1(\theta_1) = 1$ and $\hat{\pi}_2(\theta_2) = 1$.

Pluralistic Ignorance. A type underweights the likelihood that others have similar preferences or models of inference (Miller and McFarland 1987, 1991). For example, all types correctly interpret signals but believe others overreact.

A model of inference that depends on the belief about the state can capture confirmation biases such as *selective exposure*, *selective perception* and *selective recall* (Nickerson 1998); see Online Appendix E for this extension. Type-specific signal distributions can capture biases that involve interpersonal comparisons of the quality of information, such as *overconfidence* in the accuracy of one's information relative to others (Moore and Healy 2008) and overestimating the precision of signals from agents who have similar preferences or models i.e. the *social circle heuristic* (Pachur, Rieskamp, and Hertwig 2004); see Section 3.6 for details.

3 Asymptotic Learning

We study asymptotic learning outcomes—the long-run beliefs about the state—for social types. Our main result characterizes how these long-run beliefs depend on *two expressions* that are straightforward to derive from the primitives of the model.

3.1 Belief Updating

We first characterize how each type updates its belief about the state.

Individual Decision Problem. Consider an agent of type θ_i who observes history h and private signal s. The agent uses her model of inference to compute the probability of h in each state, $\hat{P}_i(h|\omega)$, and applies Bayes rule to form the likelihood ratio

$$\lambda_i(h) \equiv \frac{\hat{P}_i(R|h)}{\hat{P}_i(L|h)} = \left(\frac{p_0}{1-p_0}\right)\frac{\hat{P}_i(h|R)}{\hat{P}_i(h|L)} \tag{1}$$

that the state is R versus L. By Assumption 4, θ_i believes that h occurs with positive probability and therefore Bayes rule can be used to update beliefs. If the agent is an autarkic or noise type, she believes that the history is uninformative, implying $\lambda_i(h) = p_0/(1-p_0)$ for all h. The agent then uses her subjective signal distribution $\hat{s}_i(s)$ to update to private posterior belief $\lambda_i(h)\hat{s}_i(s)/(1-\hat{s}_i(s))$ and chooses the action that maximizes her expected payoff with respect to this belief. The following lemma represents each type's decision rule as a set of signal cut-offs.

Lemma 1 (Decision Rule)**.** *Assume Assumptions 1 and 2. For each θ_i and $\lambda \in [0,\infty]$, there exist signal cut-offs $0 = \overline{s}_{i,0}(\lambda) \leq \overline{s}_{i,1}(\lambda) \leq ... \leq \overline{s}_{i,M}(\lambda) = 1$ such that an agent of type θ_i chooses action a_m at likelihood ratio λ iff $\overline{s}_{i,m-1}(\lambda) \neq \overline{s}_{i,m}(\lambda)$ and she observes private signal $s \in (\overline{s}_{i,m-1}(\lambda), \overline{s}_{i,m}(\lambda)]$, with a closed interval if $\overline{s}_{i,m-1}(\lambda) = 0$.*[17]

Interpreting Action Histories. These signal cut-offs determine how agents interpret past action choices. A social type $\theta_i \in \Theta_S$ believes type θ_j with likelihood ratio λ_j chooses a_m with probability $\hat{F}_i^\omega(\overline{s}_{j,m}(\lambda_j)) - \hat{F}_i^\omega(\overline{s}_{j,m-1}(\lambda_j))$ in state ω, i.e. the subjective probability θ_i assigns to θ_j observing a signal in the interval $(\overline{s}_{j,m-1}(\lambda_j), \overline{s}_{j,m}(\lambda_j)]$ that lead to choice a_m. Therefore, given likelihood ratios $\boldsymbol{\lambda} \equiv (\lambda_1, ..., \lambda_k) \in [0,\infty]^k$ for social types and $\lambda_j = p_0/(1-p_0)$ for autarkic or noise types, θ_i believes that a_m is chosen with probability

$$\hat{\psi}_i(a_m|\omega, \boldsymbol{\lambda}) \equiv \sum_{j=1}^{n} \hat{\pi}_i(\theta_j)(\hat{F}_i^\omega(\overline{s}_{j,m}(\lambda_j)) - \hat{F}_i^\omega(\overline{s}_{j,m-1}(\lambda_j))) \tag{2}$$

[17] Throughout the paper, we work with the extended real number line $\mathbb{R} \cup \{-\infty, \infty\}$ to allow $\lambda = \infty$ to denote the belief at which an agent places probability one on state R.

in state ω, i.e. the subjective probability θ_i assigns to each type choosing a_m weighted by θ_i's subjective type distribution. Note that the probability of a_m varies with both the realized state and the belief about the state. The true probability of a_m is analogous, substituting the true signal and type distributions,

$$\psi(a_m|\omega, \boldsymbol{\lambda}) \equiv \sum_{j=1}^{n} \pi(\theta_j)(F^\omega(\overline{s}_{j,m}(\lambda_j)) - F^\omega(\overline{s}_{j,m-1}(\lambda_j))). \tag{3}$$

When $\hat{\psi}_i(a_m|\omega, \boldsymbol{\lambda}) \neq \psi(a_m|\omega, \boldsymbol{\lambda})$, misspecification introduces a wedge between the subjective and true probability of observing each action.

The following lemma establishes two key properties of $\psi(a|\omega, \boldsymbol{\lambda})$ and $\hat{\psi}_i(a|\omega, \boldsymbol{\lambda})$. First, all social types perceive action a_1 as *uniformly informative* of state L—that is, more likely in state L at all values of the likelihood ratio—and perceive a_M as uniformly informative of state R. This follows from Assumptions 1 and 2, which rule out confounded learning, and Assumption 3, which rules out informational herds.[18] Second, $\hat{\psi}_i(a|\omega, \boldsymbol{\lambda})$ is continuous with respect to $\boldsymbol{\lambda}$ at certainty. The same properties hold for $\psi(a|\omega, \boldsymbol{\lambda})$.

Lemma 2. *Assume Assumptions 1 to 4. For all $\theta_i \in \Theta_S$, $a_1(a_M)$ is perceived as uniformly informative of state $L(R)$,*

$$\sup_{\boldsymbol{\lambda} \in [0,\infty]^k} \frac{\hat{\psi}_i(a_1|R, \boldsymbol{\lambda})}{\hat{\psi}_i(a_1|L, \boldsymbol{\lambda})} < 1 \text{ and } \inf_{\boldsymbol{\lambda} \in [0,\infty]^k} \frac{\hat{\psi}_i(a_M|R, \boldsymbol{\lambda})}{\hat{\psi}_i(a_M|L, \boldsymbol{\lambda})} > 1$$

and $\boldsymbol{\lambda} \mapsto \hat{\psi}_i(a|\omega, \boldsymbol{\lambda})$ is continuous at $\boldsymbol{\lambda} \in \{0, \infty\}^k$ for all $(a, \omega) \in \mathcal{A} \times \{L, R\}$. Analogous properties hold for $\psi(a|\omega, \boldsymbol{\lambda})$.

From this point forward, our analysis is derived in terms of $\psi(a|\omega, \boldsymbol{\lambda})$ and $\hat{\psi}_i(a|\omega, \boldsymbol{\lambda})$. Therefore, one could abstract from the microfoundations for individual and social learning presented in Section 2 and work directly with a reduced form learning model represented by a state and belief-dependent stochastic process on $\mathcal{A}$. Our subsequent analysis holds for any such process that satisfies the properties derived in Lemma 2 (uniform informativeness can hold at any two elements in $\mathcal{A}$) and a consistency requirement that all observed information is in the support of agents' subjective distributions. This allows our learning characteri-

[18]Confounded learning corresponds to an interior belief at which actions are uninformative even though each type acts based on private information. When preferences are not aligned, such as $u_1 = \mathbb{1}_{a=\omega}$ and $u_2 = \mathbb{1}_{a\neq\omega}$, the aggregate probability that an action is chosen can be the same in each state even though different types choose actions with different probabilities (Smith and Sørensen 2000). Model heterogeneity can lead to the same issue when it is not common knowledge that signals and preferences are aligned. An informational herd corresponds to an interior belief at which all types choose the same action regardless of their private information (Banerjee 1992; Bikhchandani et al. 1992). Unbounded private signals also rule out informational herds (Smith and Sørensen 2000). While signals may be unbounded in our setting, our learning characterization requires the stronger notion of uniformly informative actions.

zation to be immediately applied to large class of dynamic decision problems with model misspecification—including active individual learning models with myopic agents, learning from alternative sources of information (i.e. stochastic outcomes, public signals), learning from multiple sources of information, and models of inference that vary with agents' beliefs about the state.[19] While some common economic settings violate uniform informativeness—namely, social learning with unbounded private signals and individual learning with costly information acquisition—the property is restored if there is an additional source of information or if all realizations of the information process are a stochastic function of choices and the state (e.g. $\mathcal{A}$ corresponds to stochastic outcomes as in Bohren and Hauser (2019a)).

The Likelihood Ratio Process. From Eq. (2), we derive how the likelihood ratios for social types depend on the history. Each $\theta_i \in \Theta_S$ initially has likelihood ratio $\lambda_i(h_1) = p_0/(1-p_0)$. At any history h_t with $t > 1$,

$$\lambda_i(h_t) = \left(\frac{p_0}{1-p_0}\right) \prod_{\tau=1}^{t-1} \frac{\hat{\psi}_i(\tilde{a}_\tau | R, \boldsymbol{\lambda}(h_\tau))}{\hat{\psi}_i(\tilde{a}_\tau | L, \boldsymbol{\lambda}(h_\tau))}. \tag{4}$$

The process is recursive: given $\boldsymbol{\lambda}(h_t)$ and $\tilde{a}_t$, $\lambda_i(h_{t+1}) = \lambda_i(h_t) \frac{\hat{\psi}_i(\tilde{a}_t|R,\boldsymbol{\lambda}(h_t))}{\hat{\psi}_i(\tilde{a}_t|L,\boldsymbol{\lambda}(h_t))}$. Therefore, $\boldsymbol{\lambda}_t \equiv \boldsymbol{\lambda}(h_t)$ is sufficient for the history and we suppress the dependence on h. The behavior of $\langle \boldsymbol{\lambda}_t \rangle_{t=1}^{\infty}$ determines the learning dynamics for each social type—where $\hat{\psi}_i(a|L, \boldsymbol{\lambda})$ and $\hat{\psi}_i(a|R, \boldsymbol{\lambda})$ determine the *update* to the likelihood ratio, and $\psi(a|\omega, \boldsymbol{\lambda})$ in the realized state determines the *probability* of this update. Characterizing the behavior of $\langle \boldsymbol{\lambda}_t \rangle_{t=1}^{\infty}$ is challenging: the process is an equilibrium object with nonlinear transitions that depend on its current value, and in contrast to correctly specified environments, it is generally not a martingale.

We close with an example, which we use throughout the paper to illustrate our results.

Example 1 (Partisan Bias)**.** *Suppose agents systematically slant signals towards state R. There are two types of agents: θ_1 is social and θ_2 is autarkic, with share $\pi(\theta_1) \in (0,1)$ of social types. Both have an identical level of partisan bias parameterized by $\hat{F}_i^{\omega}(s) = F^{\omega}(s^{\nu})$ for $\nu \in (0,1)$ and $\mathcal{S} = [0,1]$, which results in private belief $\hat{s}_i(s) = s^{\nu}$. The social type has a correctly specified type distribution, $\hat{\pi}_1 = \pi$. Both types seek to choose the action that matches the state, $\mathcal{A} = \{L, R\}$ and $u_i(a,\omega) = \mathbb{1}_{a=\omega}$, and start with prior $p_0 = 0.5$. When the social type has likelihood ratio λ and observes private signal s, it updates to private belief $\lambda\left(\frac{s^{\nu}}{1-s^{\nu}}\right)$. It chooses action L if this private belief is less than one, which occurs for signals less than $\overline{s}_{1,1} = 1/(1+\lambda)^{1/\nu}$. Similarly, the autarkic type chooses L for signals less than $\overline{s}_{2,1} = 0.5^{1/\nu}$.*

[19]Note that when ψ and $\hat{\psi}_i$ are independent of $\boldsymbol{\lambda}$, our set-up is a special case of Berk (1966). While in principle, our analysis could be applied to non-myopic active learning problems, in practice, it would be a significant challenge to verify the properties derived in Lemma 2 as $\hat{\psi}(\cdot|\omega, \boldsymbol{\lambda})$ would depend on the solution to a dynamic optimization problem. Therefore, our framework does not easily apply to settings such as that studied in Fudenberg et al. (2017).

The social type believes social types choose L with probability $\hat{F}_1^{\omega}(1/(1+\lambda)^{1/\nu}) = F^{\omega}(1/(1+\lambda))$ *and autarkic types choose L with probability* $\hat{F}_1^{\omega}(0.5^{1/\nu}) = F^{\omega}(0.5)$*. At any* $\lambda \in (0,\infty)$*, this is greater than the true probabilities* $F^{\omega}(1/(1+\lambda)^{1/\nu})$ *and* $F^{\omega}(0.5^{1/\nu})$*, respectively, that each type chooses L. Therefore, the social type overestimates the frequency of L actions.*

3.2 Stationary Beliefs and Learning Outcomes

At a stationary belief, the likelihood ratio remains constant for any action that occurs with positive probability.

Definition 1 (Stationary)**.** *Belief* $\boldsymbol{\lambda}^* \in [0,\infty]^k$ *is* stationary *if for all* $a \in \mathcal{A}$*, either (i)* $\psi(a|\omega,\boldsymbol{\lambda}^*) = 0$ *or (ii)* $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^*\left(\frac{\hat{\psi}_i(a|R,\boldsymbol{\lambda}^*)}{\hat{\psi}_i(a|L,\boldsymbol{\lambda}^*)}\right)$ *for all* $\theta_i \in \Theta_S$*.*

From Lemma 2, actions a_1 and a_M are uniformly informative across the belief space. Therefore, the set of stationary beliefs corresponds to the set of certain beliefs in which each type places probability one on a single state.

Lemma 3 (Stationary Beliefs)**.** *Assume Assumptions 1 to 4. The set of stationary beliefs is* $\{0,\infty\}^k$*. Given likelihood ratio process* $\langle\boldsymbol{\lambda}_t\rangle_{t=1}^{\infty}$*, for any belief* $\boldsymbol{\lambda}^* \in [0,\infty]^k$*, if* $\lambda_i^* \in (0,\infty)$ *for some* $\theta_i \in \Theta_S$*, then* $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$*.*

These stationary beliefs are the candidate limit points of the likelihood ratio: if the likelihood ratio converges for all types, then it must converge to $\boldsymbol{\lambda}^* \in \{0,\infty\}^k$. This rules out incomplete learning, i.e. $\lambda_{i,t}$ converges to an interior belief for some type.

We define asymptotic learning outcomes relative to the set of stationary beliefs.

Definition 2 (Learning Outcomes)**.** *In state L,* correct *learning (for type* θ_i*) denotes the event where* $\boldsymbol{\lambda}_t \to 0^k$ *(*$\lambda_{i,t} \to 0$*),* incorrect *learning (for type* θ_i*) denotes the event where* $\boldsymbol{\lambda}_t \to \infty^k$ *(*$\lambda_{i,t} \to \infty$*),* entrenched disagreement *denotes the event where* $\boldsymbol{\lambda}_t \to \{0,\infty\}^k \setminus \{0^k,\infty^k\}$*,* cyclical *learning (for type* θ_i*) denotes the event where* $\boldsymbol{\lambda}_t$ *(*$\lambda_{i,t}$*) does not converge, and* mixed *learning denotes the event where* $\lambda_{i,t}$ *converges for some social types but not others. The definitions are analogous in state R.*

When all social types have the same limit belief, we refer to this as an *agreement* outcome. *Entrenched disagreement* occurs when different types converge to different limit beliefs; throughout the paper, when we say 'disagreement' we are referring to 'entrenched disagreement'.[20] Learning is *complete* if correct learning occurs almost surely and is *path-dependent* if multiple learning outcomes arise with positive probability—for example, correct and incorrect learning. In correctly specified environments, the Martingale Convergence Theorem rules out incorrect, cyclical, and mixed learning, and entrenched disagreement. This is not the case in misspecified environments.

[20]Types' beliefs can also differ when beliefs do not converge i.e. cyclical or mixed learning.

3.3 Stability of Learning Outcomes

In this section, we derive conditions for the likelihood ratio to converge to each stationary belief with positive probability. To do so, we first characterize the behavior of the likelihood ratio when it is in a neighborhood of a stationary belief. Building on results on the local stability of nonlinear stochastic difference equations developed in Smith and Sørensen (2000), we establish necessary and sufficient conditions for the likelihood ratio to converge to this stationary belief with positive probability from nearby beliefs, which we refer to as *local stability* (Theorem 1). We then determine when the likelihood ratio converges to a locally stable belief with positive probability from any initial belief, which we refer to as *global stability* (Theorems 2 and 3). This ensures that our characterization holds independently of the initial belief.

Our approach builds on techniques used in Bohren (2016) to characterize asymptotic learning outcomes when there is a single type with a misspecified model of the share of autarkic types. Our key technical innovations are to allow for multiple types and to characterize conditions for entrenched disagreement. Relative to Bohren (2016), establishing the global stability of disagreement outcomes and belief convergence with multiple types requires novel and different techniques.

Local Stability. A learning outcome is *locally stable* if the likelihood ratio converges to it with positive probability from nearby beliefs and is *unstable* if the likelihood ratio almost surely does not converge to it.

Definition 3 (Local Stability)**.** $\boldsymbol{\lambda}^*$ *is* locally stable *if there exists an* $\varepsilon > 0$ *and neighborhood* $B_\varepsilon(\boldsymbol{\lambda}^*)$ *such that* $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^* | \boldsymbol{\lambda}_1 \in B_\varepsilon(\boldsymbol{\lambda}^*)) > 0$ *and is* unstable *if* $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$.

Our characterization of local stability depends on the sign of the average action update weighted by the true probability of each action. In state ω, this is equal to

$$\gamma_i(\omega, \boldsymbol{\lambda}) \equiv \sum_{a \in \mathcal{A}} \psi(a|\omega, \boldsymbol{\lambda}) \log \left(\frac{\hat{\psi}_i(a|R, \boldsymbol{\lambda})}{\hat{\psi}_i(a|L, \boldsymbol{\lambda})} \right) \tag{5}$$

for social type θ_i at belief $\boldsymbol{\lambda}$. This expression has two natural interpretations. First, at interior beliefs, it corresponds to the expected change in the log likelihood ratio. Second, it is the difference between (i) the Kullback-Leibler divergence from type θ_i's subjective model in state L, $\hat{\psi}_i(\cdot|L, \boldsymbol{\lambda})$ to the true model in state ω, $\psi(\cdot|\omega, \boldsymbol{\lambda})$ and (ii) the Kullback-Leibler divergence from θ_i's subjective model in state R, $\hat{\psi}_i(\cdot|R, \boldsymbol{\lambda})$, to the true model in state ω, $\psi(\cdot|\omega, \boldsymbol{\lambda})$. If θ_i's subjective model in state L is closer to the true model than θ_i's subjective model in state R, then this difference is negative and θ_i's log likelihood ratio moves towards state L in expectation; otherwise, it moves towards state R.

We show that the sign of each component of $\boldsymbol{\gamma}(\omega, \boldsymbol{\lambda}) \equiv (\gamma_i(\omega, \boldsymbol{\lambda}))_{i=1}^k$ at a stationary belief determines whether the belief is locally stable. Let

$$\Lambda_i(\omega) \equiv \{\boldsymbol{\lambda} \in \{0,\infty\}^k | \gamma_i(\omega,\boldsymbol{\lambda}) < 0 \text{ if } \lambda_i = 0 \text{ and } \gamma_i(\omega,\boldsymbol{\lambda}) > 0 \text{ if } \lambda_i = \infty\}, \tag{6}$$

denote the set of certain beliefs at which $\gamma_i(\omega,\boldsymbol{\lambda})$ is negative if social type θ_i believes the state is L and positive if θ_i believes the state is R. This is the first expression for our learning characterization. Theorem 1 establishes that all beliefs in $\Lambda(\omega) \equiv \cap_{i=1}^k \Lambda_i(\omega)$ are locally stable and, subject to a minor technical condition, beliefs not in $\Lambda(\omega)$ are unstable.

Theorem 1 (Locally Stable Beliefs)**.** *Assume Assumptions 1 to 4. Then $\boldsymbol{\lambda}^* \in \{0,\infty\}^k$ is locally stable in state ω if $\boldsymbol{\lambda}^* \in \Lambda(\omega)$ and is unstable if $\boldsymbol{\lambda}^* \notin \Lambda(\omega)$ and $\gamma_i(\omega,\boldsymbol{\lambda}^*) \neq 0$ for some θ_i with $\boldsymbol{\lambda}^* \notin \Lambda_i(\omega)$. All $\boldsymbol{\lambda}^* \in [0,\infty]^k \setminus \{0,\infty\}^k$ are unstable.*

The intuition is as follows. Consider certain belief $\boldsymbol{\lambda}^* \in \{0,\infty\}^k$ and suppose $\boldsymbol{\lambda}^* \in \Lambda(\omega)$. By the continuity of $\gamma_i(\omega,\boldsymbol{\lambda})$ at $\boldsymbol{\lambda}^*$, if $\lambda_i^* = 0$, then $\gamma_i(\omega,\boldsymbol{\lambda}) < 0$ in a neighborhood of $\boldsymbol{\lambda}^*$ and if $\lambda_i^* = \infty$, then $\gamma_i(\omega,\boldsymbol{\lambda}) > 0$. Therefore, in expectation, the log likelihood ratio moves towards $\boldsymbol{\lambda}^*$ from nearby beliefs. By similar reasoning, if $\boldsymbol{\lambda}^* \notin \Lambda(\omega)$, the log likelihood ratio moves away from $\boldsymbol{\lambda}^*$ at nearby beliefs (provided $\gamma_i(\omega,\boldsymbol{\lambda}^*) \neq 0$ for some θ_i with $\boldsymbol{\lambda}^* \notin \Lambda_i(\omega)$).

The set $\Lambda(\omega)$ has a natural interpretation: it corresponds to the set of strict Berk-Nash equilibria (Esponda and Pouzo 2016). At each $\boldsymbol{\lambda}^* \in \Lambda(\omega)$, each social type places probability one on the state that generates a perceived action distribution closest to the observed action distribution—that is, the state that minimizes the Kullback-Leibler divergence from the type's perceived action distribution to the true action distribution when all types act optimally with respect to $\boldsymbol{\lambda}^*$. Given the strict inequalities in Eq. (6), this state is unique for each type and the equilibrium is strict. Therefore, Theorem 1 establishes that all strict Berk-Nash equilibria are locally stable. Esponda et al. (2019); Fudenberg et al. (2020) establish a similar result in an individual learning setting that allows for non-myopic agents and a more general state space.

When $\gamma_i(\omega,\boldsymbol{\lambda}^*) = 0$ for a type, the perceived action distributions in each state are equally close to the true action distribution. At a stationary belief $\boldsymbol{\lambda}^* \in \{0,\infty\}^k$, if $\gamma_i(\omega,\boldsymbol{\lambda}^*) = 0$ for all social types with $\boldsymbol{\lambda}^* \notin \Lambda_i(\omega)$, we cannot determine whether $\boldsymbol{\lambda}^*$ is stable or unstable from the sign of $\gamma_i(\omega,\boldsymbol{\lambda}^*)$ (these correspond to weak Berk-Nash equilibria). Conditions for stability in this case significantly complicate the analysis without adding much economic insight. Going forward, we focus on learning environments in which this does not occur.

Definition 4 (Identified at Certainty)**.** *A learning environment is* identified at certainty *if $\gamma_i(\omega,\boldsymbol{\lambda}) \neq 0$ for all $\theta_i \in \Theta_S$, $\boldsymbol{\lambda} \in \{0,\infty\}^k$ and $\omega \in \{L,R\}$.*

The set of learning environments that are identified at certainty is *generic* in the set of environments that satisfy Assumptions 1 to 4. Note that all correctly specified environments that satisfy Assumptions 1 to 4 are identified at certainty. When the learning environment is identified at certainty, Theorem 1 simplifies to the following corollary.

Corollary 1. *Assume Assumptions 1 to 4. If the learning environment is identified at certainty, then $\boldsymbol{\lambda}^* \in \{0,\infty\}^k$ is locally stable in state ω if and only if $\boldsymbol{\lambda}^* \in \Lambda(\omega)$. All $\boldsymbol{\lambda}^* \in [0,\infty]^k \setminus \{0,\infty\}^k$ are unstable.*

In other words, if the likelihood ratio converges for all social types, then it must converge to a limit random variable whose support lies in $\Lambda(\omega)$. This reduces characterizing local stability to determining the sign of $\gamma_i(\omega, \boldsymbol{\lambda})$ for each social type at certain beliefs. It is straightforward to do so in applications, as demonstrated in the following example.

Example 1 (Partisan Bias, cont.)**.** *From the perceived and true probabilities of each action derived in Section 3.1, $\gamma_1(L, 0)$ is equal to*

$$\underbrace{(\pi(\theta_1)+\pi(\theta_2)F^L(.5^{\frac{1}{\nu}}))\log\frac{\pi(\theta_1)+\pi(\theta_2)F^R(.5)}{\pi(\theta_1)+\pi(\theta_2)F^L(.5)}}_{L-action}+\underbrace{\pi(\theta_2)(1-F^L(.5^{\frac{1}{\nu}}))\log\frac{1-F^R(.5)}{1-F^L(.5)}}_{R-action}.$$

The construction of $\gamma_1(L,\infty)$ is analogous. When the bias is small (ν is close to one), both $\gamma_1(L,0)$ and $\gamma_1(L,\infty)$ are negative, and $\Lambda(L)=\{0\}$. As the bias grows, R actions occur more frequently and both expressions increase. For intermediate levels of bias, $\gamma_1(L,0)$ is positive, $\gamma_1(L,\infty)$ is negative, and $\Lambda(L)=\emptyset$. When the bias is sufficiently large, both expressions are positive and $\Lambda(L)=\{\infty\}$. See Online Appendix B.1 for this derivation.

Global Stability. We are interested in a characterization of asymptotic learning that is independent of the initial belief. Therefore, we need a stronger notion of stability than local stability. We say a learning outcome is *globally stable* if the likelihood ratio converges to this outcome with positive probability, from *any* initial belief.

Definition 5 (Global Stability)**.** $\boldsymbol{\lambda}^*$ *is* globally stable *if for any initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^k$, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) > 0$.*

Clearly, the set of globally stable learning outcomes is a subset of the set of locally stable learning outcomes. It remains to establish when local stability implies global stability.

Aligned signals and preferences (Assumptions 1 and 2) guarantee that there exist actions that move the beliefs of all types in the same direction. Therefore, from any initial belief, it is possible to construct a finite sequence of actions that occur with positive probability and move the likelihood ratio arbitrarily close to an agreement outcome. Once the likelihood ratio is in a neighborhood of the agreement outcome, local stability establishes convergence. Therefore, global stability follows immediately from local stability for an *agreement* outcome.

Theorem 2 (Global Stability of Agreement)**.** *Consider a learning environment that is iden-*

tified at certainty and satisfies Assumptions 1 to 4. Agreement outcome $\boldsymbol{\lambda}^* \in \{0^k, \infty^k\}$ *is globally stable in state* ω *if and only if* $\boldsymbol{\lambda}^* \in \Lambda(\omega)$.[21]

Given Theorem 2, deriving $\Lambda(\omega)$ is the only calculation necessary to determine whether correct or incorrect learning occur with positive probability: these learning outcomes occur with positive probability if and only if the corresponding limit beliefs are in $\Lambda(\omega)$. Further, this result fully characterizes global stability when there is a single social type.

Global stability does not immediately follow from local stability for *disagreement* outcomes, as it may not be possible to separate the beliefs of different types and reach a neighborhood of the disagreement outcome. For example, the beliefs of two similar types can remain close together when starting from a common prior, even if disagreement is possible when they start with very different priors. The second expression for our learning characterization—*maximal accessibility*—provides a sufficient condition to separate beliefs and push the likelihood ratio arbitrarily close to a given disagreement outcome. We first define a partial order on how types update following a_1, which decreases the likelihood ratio, and a_M, which increases it.

Definition 6 (Maximal R-order)**.** *Define the order* $\succeq_{\boldsymbol{\lambda}}$ *over* Θ *at belief* $\boldsymbol{\lambda}$ *as* $\theta_i \succeq_{\boldsymbol{\lambda}} \theta_j$ *iff* θ_i *interprets* a_1 *and* a_M *as stronger evidence of state* R *than* θ_j,

$$\frac{\hat{\psi}_j(a|R, \boldsymbol{\lambda})}{\hat{\psi}_j(a|L, \boldsymbol{\lambda})} \leq \frac{\hat{\psi}_i(a|R, \boldsymbol{\lambda})}{\hat{\psi}_i(a|L, \boldsymbol{\lambda})} \tag{7}$$

for $a \in \{a_1, a_M\}$, *with strict order* $\succ_{\boldsymbol{\lambda}}$ *if Eq. (7) holds with strict inequality for at least one action* $a \in \{a_1, a_M\}$.

We use this order to define maximal accessibility. As the number of possible disagreement outcomes increases with the number of social types, so does the complexity of such a property; we present the case of two social types here and relegate the case of more than two social types to Online Appendix D.

Definition 7 (Maximal Accessibility ($k = 2$))**.** *Disagreement outcome* $(0, \infty)$ *is* maximally accessible *if* $\theta_2 \succ_{(0,0)} \theta_1$ *or* $\theta_2 \succ_{(\infty,\infty)} \theta_1$, *and disagreement outcome* $(\infty, 0)$ *is* maximally accessible *if* $\theta_1 \succ_{(0,0)} \theta_2$ *or* $\theta_1 \succ_{(\infty,\infty)} \theta_2$.

It is straightforward to verify maximal accessibility in applications by evaluating Eq. (7) at beliefs $(0, 0)$ or (∞, ∞).

When a disagreement outcome is maximally accessible, for any initial belief, there exists a finite sequence of actions that moves beliefs to a neighborhood of the disagreement outcome. To see this, suppose $\theta_2 \succ_{(0,0)} \theta_1$. As discussed above, the likelihood ratio enters a

[21] When a learning environment is not identified at certainty, our proof establishes that any agreement outcome $\boldsymbol{\lambda}^* \in \Lambda(\omega)$ is globally stable in state ω.

neighborhood of $(0,0)$ with positive probability from any initial belief. Given $\theta_2 \succ_{(0,0)} \theta_1$, we can construct a sequence of a_1 and a_M actions that decrease θ_1's beliefs and increase θ_2's beliefs in a neighborhood of $(0,0)$. We show that this guarantees that there exists a finite sequence of actions that occurs with positive probability and moves beliefs from a neighborhood of $(0,0)$ to a neighborhood of $(0,\infty)$. Once the likelihood ratio is sufficiently close to the disagreement outcome, local stability establishes convergence. Therefore, the global stability of a disagreement outcome follows from maximal accessibility and local stability.[22]

Theorem 3 (Global Stability of Disagreement ($k=2$))**.** *Consider a learning environment that satisfies Assumptions 1 to 4. If disagreement outcome $\boldsymbol{\lambda}^* \in \{(0,\infty),(\infty,0)\}$ is in $\Lambda(\omega)$ and maximally accessible, then $\boldsymbol{\lambda}^*$ is globally stable in state ω.*

See Section 4.3 for an application that uses maximal accessibility.

Mixed Learning. Next, we establish conditions to rule out mixed learning. Suppose $\omega = L$ and consider the mixed outcome in which θ_1 has correct learning and θ_2 has cyclical learning. If either $(0,0) \in \Lambda_2(\omega)$ or $(0,\infty) \in \Lambda_2(\omega)$, then $\langle \lambda_{2,t} \rangle$ converges with positive probability when $\lambda_1^* = 0$, and hence, almost surely cannot oscillate infinitely often. Therefore, in order for this mixed outcome to arise with positive probability, it must be that $(0,0) \notin \Lambda_2(\omega)$ and $(0,\infty) \notin \Lambda_2(\omega)$, i.e. in a neighborhood of $(0,0)$ or $(0,\infty)$, θ_2's beliefs drift away from the outcome. Generalizing this intuition, let $\Lambda_M(\omega)$ denote the set of mixed outcomes in which there are no locally stable beliefs for the non-convergent types. When $k=2$, this corresponds to

$$\Lambda_M(\omega) \equiv \{(\lambda_i^*, \theta_i) \in \{0,\infty\} \times \{\theta_1,\theta_2\} | (\lambda_i^*, 0) \notin \Lambda_{-i}(\omega) \text{ and } (\lambda_i^*, \infty) \notin \Lambda_{-i}(\omega)\}, \quad (8)$$

where (λ_i^*, θ_i) denotes the mixed outcome in which θ_i's beliefs converge to λ_i^* and θ_{-i}'s beliefs do not converge. As in the case of disagreement, mixed learning is more involved when there are more than two social types; we relegate this case to Online Appendix D and define $\Lambda_M(\omega) = \emptyset$ when $k=1$. The following result establishes that if a mixed outcome is not in $\Lambda_M(\omega)$, then it almost surely does not occur.

Lemma 4 (Unstable Mixed Outcomes ($k=2$))**.** *Consider a learning environment that is identified at certainty and satisfies Assumptions 1 to 4. If $(\lambda_i^*, \theta_i) \notin \Lambda_M(\omega)$, then $Pr(\lambda_{i,t} \to \lambda_i^*$ and $\lambda_{-i,t}$ does not converge$) = 0$.*

[22] Maximal accessibility is simple and tractable, but it can be restrictive, especially in models with large action spaces. In Theorem 7 in Appendix A.3, we establish a more general condition that uses all actions to separate beliefs, which we call *separability* (Definition 8). It is more cumbersome to verify but holds for a larger set of learning environments. Another simple sufficient condition to separate beliefs is: $(0,0) \in \Lambda_1(\omega) \setminus \Lambda_2(\omega)$ or $(\infty,\infty) \in \Lambda_2(\omega) \setminus \Lambda_1(\omega)$ for $(0,\infty)$, with an analogous condition for $(\infty,0)$. It can be directly verified from the construction of $\Lambda(\omega)$ but will not hold when both agreement outcomes are locally stable.

Therefore, $\Lambda_M(\omega) = \emptyset$ rules out mixed learning.[23] It is straightforward to derive $\Lambda_M(\omega)$ from $\Lambda_i(\omega)$. See Example 2 and Sections 4.2 and 4.3 for applications that show $\Lambda_M(\omega)$ is empty.

Lemma 4 does not determine whether a mixed outcome in $\Lambda_M(\omega)$ arises with positive probability. Doing so presents a novel challenge relative to stationary learning outcomes, as it requires characterizing the movement of the convergent type's belief across all possible beliefs for the non-convergent type. We leave this question for future research.

3.4 Learning Characterization

We use the stability results in the previous section to characterize the set of asymptotic learning outcomes in each state. The final piece of the characterization involves showing when the likelihood ratio converges almost surely for all social types (Lemma 7 in Appendix A.5). This establishes our main result.[24]

Theorem 4 (Learning Characterization ($k \leq 2$))**.** *Consider a learning environment that is identified at certainty and satisfies Assumptions 1 to 4. When the state is L:*

1. **Correct** *learning occurs with positive probability if and only if* $0^k \in \Lambda(L)$.
2. **Incorrect** *learning occurs with positive probability if and only if* $\infty^k \in \Lambda(L)$.
3. **Entrenched Disagreement** *occurs with positive probability if $\Lambda(L)$ contains a maximally accessible disagreement outcome and almost surely does not occur if $\Lambda(L)$ contains no disagreement outcome. Each maximally accessible* disagreement *outcome in $\Lambda(L)$ occurs with positive probability.*
4. **Cyclical Learning** *occurs almost surely if and only if $\Lambda(L) = \emptyset$ when $k = 1$. Cyclical learning occurs almost surely if $\Lambda(L) = \emptyset$ and $\Lambda_M(L) = \emptyset$, occurs almost surely for at least one social type if $\Lambda(L) = \emptyset$, and almost surely does not occur if $\Lambda(L)$ contains an agreement outcome or maximally accessible disagreement outcome when $k = 2$.*

An analogous result holds in state R.

The conditions for correct and incorrect learning are tight: these learning outcomes arise if and only if the respective limit beliefs are in $\Lambda(\omega)$. For a disagreement outcome, there is a wedge between the sufficient conditions for it to arise—maximal accessibility—and not arise—instability.[25] This wedge disappears if all locally stable disagreement outcomes are maximally accessible (see Section 4.3 for an example.) If multiple learning outcomes occur

---

[23] A simple sufficient condition for $\Lambda_M(\omega) = \emptyset$ is that both agreement outcomes or both disagreement outcomes are in $\Lambda(\omega)$. When $k > 2$, an analogous condition to Eq. (8) rules out mixed outcomes in which one type has cyclical learning. Ruling out mixed outcomes in which more than one type has cyclical learning requires joint conditions on $\Lambda_i(\omega)$ for the non-convergent types.

[24] The statement of the result for $k > 2$ social types is identical, using the expanded definitions of maximal accessibility and $\Lambda_M(\omega)$ in Online Appendix D.

[25] Theorem 7 in Appendix A.3 presents a weaker condition—separability—to establish the global stability of disagreement outcomes.

with positive probability, then learning is path-dependent—agents become certain of different states following different histories (again see Section 4.3). Theorem 4 also determines when learning is complete.

Corollary 2 (Complete Learning)**.** *If $\Lambda(L) = \{0^k\}$ and $\Lambda_M(L) = \emptyset$, correct learning occurs almost surely in state L. An analogous result holds in state R.*

It follows immediately from the martingale property of the likelihood ratio that these conditions are satisfied in a correctly specified environment.[26] As we show in Example 1 at the end of this subsection, they can also hold in misspecified environments.

An important feature of Theorem 4 is that the characterization requires calculations at a *finite* set of beliefs—in particular, determining $\Lambda(\omega)$, $\Lambda_M(\omega)$ and maximal accessibility only requires computing $\psi(a|\omega, \boldsymbol{\lambda})$ and $\hat{\psi}(a|\omega, \boldsymbol{\lambda})$ at stationary beliefs $\boldsymbol{\lambda} \in \{0, \infty\}^k$. Since action choices depend on beliefs, in principle, determining the asymptotic properties of the likelihood ratio could require characterizing its behavior across the entire belief space. The fact that this is not necessary makes Theorem 4 straightforward to use in applications.

Several economic insights emerge from Theorem 4. First, belief convergence forces action convergence: each type eventually settles on an action if and only if its beliefs converge. It follows that the limit action choice is efficient if and only if learning is correct. If learning is incorrect for a type, the type will choose inefficient actions infinitely often, and if learning is cyclical, the type will choose both efficient and inefficient actions infinitely often.[27]

Second, model misspecification gives rise to two potential sources of entrenched disagreement in society. Model heterogeneity can lead to entrenched disagreement *within* a population due to differing interpretations of a common history. A signal that, for instance, a vaccine is safe or a politician is corrupt can cause agents to update in opposite directions based on their belief about the credibility of the source. This arises even though preferences and signals are aligned, so that agents have a common interpretation of whether one signal or action is relatively more likely in a given state than another. Therefore, model heterogeneity can explain how connected populations observing shared sources can perpetually disagree.[28] Path-dependent learning can also lead to entrenched disagreement, but *across* populations that observe different histories rather than within populations that have different models.[29] This can explain how separate populations with similar models can come to have polarized ingrained views. When both sources are present, then within-population disagreement can

[26] The likelihood ratio is a martingale in state L and the log function is concave. Together with Assumptions 2 and 3, this implies $\gamma_i(L, \boldsymbol{\lambda}) < 0$ for all $\boldsymbol{\lambda} \in [0, \infty]^k$, so $\Lambda(L) = \{0^k\}$ and $\Lambda_M(L) = \emptyset$.

[27] In the proof of Theorem 4, we show that if the likelihood ratio for a type does not converge, then it enters a neighborhood of each certain belief infinitely often.

[28] For example, Gagnon-Bartsch (2016) show that taste projection can lead to entrenched disagreement.

[29] Earlier work establishing that path-dependent learning with multiple degenerate limit beliefs arises for specific misspecified models includes Rabin and Schrag (1999) (confirmation bias), Epstein et al. (2010) (overreaction) and Eyster and Rabin (2010); Bohren (2016) (naive social learning).

vary across populations depending on whether the observed history has a common or polarizing interpretation—in other words, the order in which information arrives will impact the level of disagreement within a population. For example, in a cognitive hierarchy learning model, agreement emerges following some histories while disagreement emerges following others (see our earlier working paper for this result (Bohren and Hauser 2019b)).

While path-dependent learning—and hence, across-population disagreement—can also occur in correctly specified social learning environments (Banerjee 1992; Bikhchandani et al. 1992), learning is incomplete at all but at most one possible limit beliefs (the degenerate belief on the correct state). Since agents remain uncertain about the state, disagreement between populations can be easily resolved by introducing a common source. In contrast, misspecified learning gives rise to path-dependent learning with multiple degenerate limit beliefs. As different populations come to place high probability on different states, it becomes increasingly difficult to reconcile prolonged disagreement with common information.

The following examples demonstrate how to apply Theorem 4.

Example 1 (Partisan Bias, cont.)**.** *Applying Theorem 4 to the characterization of $\Lambda(L)$ above establishes that correct learning occurs almost surely for mild partisan bias, cyclical learning arises almost surely for intermediate levels, and incorrect learning occurs almost surely for severe partisan bias (see Proposition 5 in Online Appendix B.1.)*

Example 2 (Partisan Bias and Unawareness)**.** *In the presence of model heterogeneity, agents have a more complex inference problem—in order to be correctly specified, an agent must know the form and frequency of different biases in the population. In this example, we show that when a type accurately interprets signals but does not account for others' partisan bias, it can be just as wrong as the partisan types. Augment Example 1 to include non-partisan agents who have correctly specified signal distributions. Analogous to the partisan types, there is a social and an autarkic non-partisan type. Neither social type is aware that some agents have a different subjective signal distribution. An analogous derivation of $\Lambda(L)$ establishes that the non-partisan social type has the same long-run learning outcome as the partisan social type—correct learning occurs almost surely for both types when there is a small share of partisan agents or mild bias, cyclical learning arises almost surely for intermediate levels, and incorrect learning occurs almost surely when there is a large share of very biased partisan agents. In other words, the presence of a large share of unaccounted for partisan types can prevent types who correctly interpreted signals from making efficient choices. On the other hand, the presence of many agents who correctly interpret signals can help even severely biased agents adopt the efficient action. Fig. 1 illustrates these three learning regions. See Online Appendix B.2 for the analysis.*

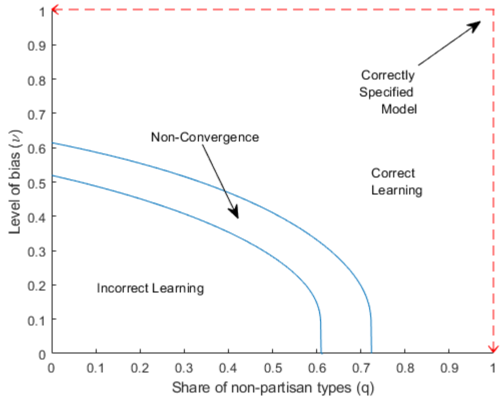

FIGURE 1. Partisan Bias and Unawareness
($\omega = L$, $F^L(s) = 2s - s^2$, $F^R(s) = s^2$, share autarkic types=0.1)

3.5 Robustness of Learning.

We next establish that learning is generally robust to the details of the environment. Specifically, correctly specified environments are robust to some misspecification in that learning is complete when social types have approximately correct models. Misspecified environments are also robust, in that except for knife-edge cases that separate different learning regions, nearby misspecified environments will have the same set of learning outcomes. To explore robustness, we first fix a learning environment and show that the same learning outcomes arise in environments with sufficiently similar equilibrium distributions of actions. We then use this result to establish that learning is complete when social types' models of inference are sufficiently close to correctly specified. Taken together, these results strengthen the applicability of correctly specified environments to real-world settings with mild biases and establish that small errors in measuring or forecasting more severe biases will not significantly alter the predicted learning outcomes.

For the first result, we focus on learning environments in which Theorem 4 fully characterizes learning outcomes—that is, environments that satisfy Assumptions 1 to 4, are identified at certainty, and in which all locally stable disagreement outcomes are maximally accessible and mixed learning almost surely does not arise, $\Lambda_M(L) = \Lambda_M(R) = \emptyset$. We refer to such environments as *regular*.[30] Fixing a regular learning environment, Theorem 5 establishes that any learning environment with sufficiently similar action distributions has the same set of long-run learning outcomes, where we use the total variation distance to measure the

[30] Within the set of environments that satisfy Assumptions 1 to 4 and are identified at certainty, this includes all correctly specified environments and all environments with a single social type, as well as the environments with multiple social types in Example 2 and Sections 4.2 and 4.3. Robustness also holds locally for environments with $\Lambda_M(\omega) \neq \emptyset$ or disagreement outcomes that are not globally stable.

closeness of distributions.[31]

Theorem 5 (Robustness). *Let (Θ^*, π^*) be a regular learning environment with set of stable learning outcomes $\Lambda^*(\omega)$ in state ω. There exists a $\delta > 0$ such that any learning environment (Θ, π) that satisfies Assumptions 1 to 4, has the same number of social types, and is sufficiently close to (Θ^*, π^*) in terms of the true and perceived distributions over actions in that $||\psi^*(\cdot|\omega, \boldsymbol{\lambda}) - \psi(\cdot|\omega, \boldsymbol{\lambda})|| < \delta$, $||\hat{\psi}_i^*(\cdot|L, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|L, \boldsymbol{\lambda})|| < \delta$ and $||\hat{\psi}_i^*(\cdot|R, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|R, \boldsymbol{\lambda})|| < \delta$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $i = 1, ..., k$, has the same set of long-run learning outcomes as (Θ^*, π^*) in state ω i.e. given sets $\Lambda(\omega)$ and $\Lambda_M(\omega)$ for (Θ, π), (i) $\Lambda(\omega) = \Lambda^*(\omega)$; (ii) $\Lambda_M(\omega) = \emptyset$; and (iii) all disagreement outcomes in $\Lambda(\omega)$ are maximally accessible.*

This result follows from the continuity of $\boldsymbol{\gamma}(\omega, \boldsymbol{\lambda})$ in ψ and $\hat{\psi}_i$, which implies that, provided $\gamma_i(\omega, \boldsymbol{\lambda}) \neq 0$, it does not change sign at belief $\boldsymbol{\lambda}$ when ψ and $\hat{\psi}_i$ are perturbed. Therefore, the locally stable set $\Lambda(\omega)$ remains the same, the mixed outcome set $\Lambda_M(\omega)$ remains empty, and locally stable disagreement outcomes remain globally stable. For example, in Fig. 1 we see that for any environment that is identified at certainty (i.e. does not lie on the two blue lines dividing the learning regions), nearby environments have the same learning outcome.

Correctly specified environments are regular and have complete learning. Therefore, fixing a correctly specified environment, Theorem 5 establishes that learning is complete in misspecified environments with similar action distributions. We use this to establish robustness when social types' models of inference are approximately correct. In order to compare models of inference without needing to define a complicated metric over types, we vary social types' models of inference while holding fixed other aspects of the learning environment—specifically, the preferences of social types and the preferences and models of autarkic and noise types. Formally, (Θ, π) is *structurally equivalent* to (Θ^*, π^*) if $|\Theta_S| = |\Theta^*_S|$, $u_i = u_i^*$ for $i = 1, ..., k$, $\Theta_A = \Theta^*_A$, $\Theta_N = \Theta^*_N$, and $\pi(\theta_i) = \pi^*(\theta_i^*)$ for $i = 1, ..., n$.[32] Theorem 6 establishes that learning is complete when social types have subjective type and signal distributions close enough to the true distributions, where again we use the total variation distance to measure the closeness of distributions.[33,34]

[31]The total variation distance between distributions ψ and ψ' is $||\psi(\cdot|\omega, \boldsymbol{\lambda}) - \psi'(\cdot|\omega, \boldsymbol{\lambda})|| \equiv \max_{A \subset \mathcal{A}} |\sum_{a \in A}(\psi(a|\omega, \boldsymbol{\lambda}) - \psi'(a|\omega, \boldsymbol{\lambda}))|$. Pinsker's inequality implies that our results also hold for the Kullback-Leibler divergence, which has been used to study robustness in subsequent work (Frick et al. 2020b,c).

[32]Recall that a correctly specified environment requires all social types to be correctly specified, but autarkic types may be misspecified. Therefore, misspecified environments have structurally equivalent correctly specified environments.

[33]In a slight abuse of notation, we simultaneously let $|| \cdot ||$ denote the total variation distance between two probability measures over the type space, i.e. $||\pi - \pi'|| \equiv \sup_{X \subset \{1,...,n\}} |\sum_{i \in X}(\pi(\theta_i) - \pi'(\theta_i'))|$, and the distance between two signal c.d.f.s with common support, i.e. $||F - F'|| \equiv \sup_{X \in \mathcal{B}_\mathcal{S}} |P_F(X) - P_{F'}(X)|$, where $\mathcal{B}_\mathcal{S}$ denotes the Borel σ-algebra on $\mathcal{S}$ and $P_F(X)$ denotes the probability of $X \in \mathcal{B}_\mathcal{S}$ under the measure described by c.d.f. F.

[34]An analogous result holds in settings where agents are very wrong about the type distribution, as long

Theorem 6. *Let (Θ^*, π^*) be a correctly specified environment that satisfies Assumptions 2 and 3. There exists a $\delta > 0$ such that in any structurally equivalent misspecified environment (Θ, π) that satisfies Assumptions 1 to 4 and in which social types have sufficiently correct models of inference in that $||\hat{\pi}_i - \pi|| < \delta$, $||\hat{F}_i^L - F^L|| < \delta$ and $||\hat{F}_i^R - F^R|| < \delta$ for all $\theta_i \in \Theta_S$, learning is complete.*

A similar result holds for an individual type that has an approximately correct model of inference, regardless of the degree to which other types are misspecified. Therefore, agents do not need to know exactly how their misspecified peers behave in order to accurately learn from their choices. Beyond Theorems 5 and 6, we can use Theorem 4 for a precise characterization of how large perturbations to the environment can be before they alter the set of learning outcomes. For example, in Fig. 1, we see that if the share of non-partisan types is 0.8, then learning is complete for any level of bias $\nu > 0$, whereas if the share of non-partisan types is 0.5, then learning is complete when $\nu > 0.4$.

These robustness results may not seem surprising, since Bayes rule is continuous. But in an infinite horizon setting, nearby models with small per-period differences in belief updating have the potential to aggregate to very different limit beliefs. For example, if agents' models are equidistant from the truth in either state at a certain belief—and therefore, the environment is not identified at certainty—then arbitrarily small perturbations alter the set of learning outcomes that arise. In Fig. 1, this is the case for the environments described by the parameters (q, ν) that trace out the boundaries dividing the different learning regions. At these parameters, there is an abrupt shift from complete learning to cyclical learning or cyclical learning to incorrect learning. But this corresponds to a measure zero set.

More generally, uniformly informative actions ensure that identification at certainty is a generic property of the learning environments we consider—and therefore, so is robustness. Further, it ensures that all correctly specified environments are identified at certainty, and therefore, all correctly specified environments are robust. The same holds true in Bohren (2016), whose robustness result is a special case of Theorem 6. In contrast, when actions are not uniformly informative, Frick et al. (2020c) show that a failure of robustness occurs in a correctly specified environment that is not identified at certainty. Identification at certainty fails because actions (or signals) are perceived to be uninformative at certainty—in contrast to our framework, in which individual actions are perceived to be informative but identification at certainty fails because the Kullback-Leibler divergence from a type's perceived action distribution to the true action distribution is the same in each state (a

as the types that they believe exist are "close" to the actual types. For example, neither type in Example 2 allows for the possibility of the other—their subjective type distributions have non-overlapping supports. But learning is complete when the partisan type is not too biased, as the partisan type is sufficiently close to the non-partisan type. By defining a suitable measure of distance between types, one could use Theorem 5 to derive an analogous result to Theorem 6 for such settings.

property that cannot arise in a correctly specified model).

Frick et al. (2020b) also demonstrate a failure of robustness in a social learning setting with an infinite state space and privately observed actions. In this environment, action choices are sensitive to small amounts of misspecification and agents with almost correct models can come to place probability one on an incorrect state. Together with our results, this demonstrates that the details of the learning environment are important to consider when exploring robustness. If one wishes to design a robust learning environment, then adjusting these details may be an important lever.

Robustness relates to whether a learning outcome is a weak or strict Berk-Nash equilibrium. In our setting, correct learning is a strict and unique Berk-Nash equilibrium in correctly specified environments. This ensures that correct learning is also the unique Berk-Nash equilibrium in nearby misspecified environments. In contrast, when a correctly specified environment is not identified at certainty—as in Smith and Sørensen (2000); Frick et al. (2020c)—correct learning is a weak Berk-Nash equilibrium and therefore is not robust.

3.6 Discussion

Focus on Asymptotic Learning. We focus on how misspecification affects long-run learning. When using this approach, an important question is whether the long-run is economically relevant. For incorrect learning, cyclical learning, or disagreement, showing these outcomes arise asymptotically establishes that agents are bounded away from efficiency or agreement, *irrespective* of the amount of information that agents observe or the rate of learning. Therefore, the source of these inefficiencies is not a lack of sufficient information to learn the state or the slow pace at which this information arrives. Long-run results also highlight an important distinction between inefficient action choices due to incomplete learning in correctly specified environments versus incorrect learning in misspecified environments. Incomplete learning is fragile and a herd can be overturned by a relatively uninformative additional information at any point in time. In contrast, when incorrect learning arises, more informative interventions are required to overturn longer incorrect herds.

When correct learning arises asymptotically leaves open important questions such as how quickly actions converge to efficiency. The expression $\boldsymbol{\gamma}(\boldsymbol{\lambda}, \omega)$ also determines the asymptotic rate of learning; we leave further study of this to future work.

Extensions. We outline several possible extensions to the learning framework.

Misaligned Type Spaces. Set $\Lambda(\omega)$ also characterizes locally stable beliefs in misaligned environments. Beyond ruling out confounded learning, the aligned assumptions are used to establish the global stability of agreement—they guarantee that there are actions that uniformly move social types' beliefs towards both agreement outcomes. If a misaligned environment has an action that uniformly moves all social types' beliefs towards a stationary belief, then our method can also be used to establish the global stability of this belief.

Signal and Type Distributions. It is a straightforward extension to allow types to receive signals from different distributions, to believe that other types receive signals from different distributions, or to allow the true and/or subjective type distributions to depend on the state. Simply augment the definition of a type to include the additional primitives. The extension to heterogeneous signal distributions allows us to model biases that involve interpersonal comparisons related to the quality of information. For instance, a natural way to model overconfidence is with a type that correctly interprets its own signals but believes other agents observe signals from a less informative distribution. It is also straightforward to allow multiple pieces of information to arrive at the same time—for example, a finite number of agents act simultaneously or all agents also observe a public signal process (for the latter, see an earlier working paper version of this article (Bohren and Hauser 2019b)). As long as the model reduces to a belief process that satisfies the conditions outlined in Lemma 2, our characterization applies.

Action and State Space. For technical convenience, we assume that the action space is finite. Allowing for a continuous action space would not qualitatively change the analysis. Similar techniques to those we use can be used to analyze a finite state space with more than two states, with the caveats that the definition of an aligned environment is more complicated and the notation is more cumbersome. We use results pertaining to stochastic difference equations in our analysis, which means that generalizing to an infinite state space requires different techniques.

4 Applications

We present three applications to demonstrate how our general framework can be used to address the issues raised in the introduction. First, we show that overreaction—a form of signal misspecification—has a fundamentally different impact when individuals learn from social versus private sources. Second, we examine whether a representative agent model is a good approximation in a setting with heterogeneous levels of naive learning. Finally, we show that entrenched disagreement arises and agreement almost surely does not in a level-k social learning model where different types have fundamentally distinct models of inference. All proofs for the results in this section are in Online Appendix C.

4.1 Overreaction: Individual versus Social Learning

This application demonstrates how our characterization can be used to determine whether the impact of a bias differs when agents learn from private versus social sources. We explore this question in a setting in which individuals overreact to signals.[35] We show that overreaction interacts with social learning to create long-run inefficiencies that are not present

[35] Overreaction has been widely documented empirically; see Section 2.3 for citations.

when agents learn directly from signals. In particular, cyclical learning arises for sufficiently severe levels of overreaction when agents learn from a social source, whereas learning is complete *regardless* of the severity of the bias when agents learn directly from signals. We conclude with a discussion of other models of overreaction and highlight how these different parameterizations differentially impact learning.

We model overreaction as a type who forms beliefs as if it has observed the same signal multiple times. The type believes private signals are distributed according to $\hat{F}^\omega(s) = F^\omega(\frac{s^\nu}{(1-s)^\nu+s^\nu})$ in state ω, where F^ω is continuous with support $\mathcal{S} = [0, 1]$ and $\nu \in (1, \infty)$ captures the degree of overreaction. This leads to private belief $\frac{\hat{s}(s)}{1-\hat{s}(s)} = \left(\frac{s}{1-s}\right)^\nu$ following signal s. For example, if $\nu = 2$, signals are double counted—the private belief following realization s corresponds to the correct belief following two realizations s—independent of the direction and strength of the signal.

Benchmark: Individual Learning. Given that the level of overreaction is independent of the realized signal, when type θ_1 learns directly from signals, the bias does not alter the sign of $\gamma_1(\omega, \boldsymbol{\lambda})$: as in the correctly specified model, $\gamma_1(\omega, \boldsymbol{\lambda}) < 0$ for all $\boldsymbol{\lambda}$ and learning is complete for any level of overreaction. This follows almost directly from Berk (1966).

Observation 1. *If type θ_1 observes signals directly, learning is complete for any $\nu \in [1, \infty)$.*

Social Learning. When signals are filtered through other agents' actions, the induced level of overreaction to an action depends on the observed action. Therefore, the bias can alter the sign of $\gamma_i(\omega, \boldsymbol{\lambda})$, and hence, the set of learning outcomes. We illustrate this in a learning environment with symmetric preferences and signal distributions, so that there are no inherent asymmetries in the underlying overreaction to signals or the mapping from beliefs to actions. Suppose there are two types of agents: θ_1 is social and θ_2 is autarkic, with $\pi(\theta_1) \in (0, 1)$. Both types have the same level of overreaction, as outlined above, face a decision problem with four actions, $a \in \{a_1, a_2, a_3, a_4\}$ (ordered according to increasing preference in state R) and a symmetric signal, $F^L(s) = 1 - F^R(1 - s)$, and have the same symmetric preferences, i.e. if a_1 is optimal at belief p that the state is R, then a_4 is optimal at belief $1 - p$, and similarly for a_2 and a_3. Given this symmetry, the agent's decision-rule can be represented as a cut-off rule $p^* \in (0, 1/2)$ such that the agent chooses a_1 if $p \leq p^*$, a_2 if $p \in (p^*, 0.5]$, a_3 if $p \in (0.5, 1 - p^*]$ and a_4 if $p \in (1 - p^*, 1]$. To ensure that moderate actions a_2 and a_3 are chosen for a sufficiently large window of beliefs, assume p^* is sufficiently small so that $\frac{F^L(p^*)-F^R(p^*)}{\log F^L(p^*)-\log F^R(p^*)} < F^R(.5)$. To close the model, assume that θ_1 has a correct subjective type distribution and both types have prior $p_0 = 1/2$.

The following result establishes that social learning causes overreaction to interfere with asymptotic learning. In particular, sufficiently severe overreaction leads to cyclical learning.

Proposition 1. *There exists a cutoff $\bar{\pi} \in (0, 1)$ on the share of social types such that: (i)*

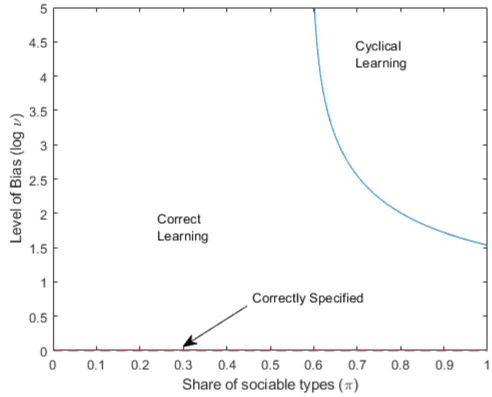

FIGURE 2. Overreaction to social information
($F^L = 2s - s^2$, $F^R = s^2$, $p^* = .25$.)

if $\pi(\theta_1) > \bar{\pi}$, then there exists a cut-off $\bar{\nu}(\pi(\theta_1)) \in (1, \infty)$, which is decreasing in $\pi(\theta_1)$, such that cyclical learning arises almost surely if $\nu > \bar{\nu}(\pi(\theta_1))$ and learning is complete if $\nu < \bar{\nu}(\pi(\theta_1))$; (ii) if $\pi(\theta_1) < \bar{\pi}$, then learning is complete for all $\nu \in [1, \infty)$.

When beliefs are close to zero, action a_1 is chosen for a larger set of signals than a_4. Therefore, the overreaction is asymmetric: it is stronger with respect to contradictory a_4 actions than confirmatory a_1 actions. This pulls beliefs away from zero. Similarly, when beliefs are close to infinity, overreaction is stronger with respect to contradictory a_1 actions. For sufficiently severe overreaction, this gives rise to cyclical learning, as illustrated in Fig. 2.

Other Parameterizations. Other approaches to modeling overreaction include Epstein et al. (2010), who model overreaction as a linear updating rule that places negative weight on the prior belief and a weight above one on the correctly specified posterior belief, and Bushong and Gagnon-Bartsch (2019), where an agent underestimates the extent of her reference dependence, which leads her to overreact when recalling past outcomes.[36] These different parameterizations are of consequence: when an agent learns directly from signals, Epstein et al. (2010)'s parameterization of overreaction alters updating in an asymmetric way and sufficiently severe overreaction leads to the possibility of incorrect learning. Similarly, incorrect learning can arise in Bushong and Gagnon-Bartsch (2019) when the agent is loss averse, so that she overreacts asymmetrically to losses and gains. In contrast, if the agent is not loss averse, the overreaction is symmetric and, as in our individual learning setting, this does not interfere with learning.

4.2 Naive Learning with Model Heterogeneity

As discussed in the introduction, papers that study model misspecification generally assume that all agents have the same form and level of misspecification. This can be viewed as

[36] In Online Appendix E, we show how our framework nests Epstein et al. (2010).

a representative agent approach, which significantly simplifies the analysis. However, the empirical literature has shown that even when agents have similar biases, they will exhibit it in differing degrees. As this application demonstrates, our characterization can determine whether a representative agent approach is valid in the face of heterogeneity in the sense that the long-run behavior of a representative agent approximates the long-run behavior of heterogeneous agents. This provides a template for evaluating the representative agent approach that is straightforward to apply to other forms of misspecification.

We explore this question in a setting in which agents are naive learners who overestimate the private information reflected in actions. We compare learning in a setting in which agents have heterogeneous levels of naivete to a representative agent setting in which a single type has a level of naivete equal to the average naivete of the population (the latter is a special case of Bohren (2016)). We show that when heterogeneity is small, this representative agent model is a good approximation of the underlying heterogeneous environment in that both environments have the same learning outcomes, whereas when heterogeneity is large, incorrect learning arises with positive probability in the corresponding representative agent model even though both types almost surely learn the correct state.

As in Bohren (2016), we model naive learning as a misspecified belief about the share of autarkic types. Let θ_A denote the autarkic type and assume $\pi(\theta_A) \in (0, 1)$. To capture model heterogeneity, suppose there are two social types, θ_1 and θ_2, that occur with equal probability, $\pi(\theta_1) = \pi(\theta_2)$. Both social types overestimate the share of autarkic types, with type θ_2 having a more severe bias, $\pi(\theta_A) < \hat{\pi}_1(\theta_A) \leq \hat{\pi}_2(\theta_A) \leq 1$. This leads agents to underestimate the correlation between prior actions. In the representative agent setting, a single social type believes that the autarkic type occurs with probability $\hat{\pi} \in (\pi(\theta_A), 1]$. We compare heterogeneous environments with biases $(\hat{\pi}_1(\theta_A), \hat{\pi}_2(\theta_A))$ to the corresponding representative agent environment with a bias equal to the average bias in the heterogeneous setting, i.e. $\hat{\pi} = (\hat{\pi}_1(\theta_A) + \hat{\pi}_2(\theta_A))/2$. To close the model, assume that each agent faces a binary decision problem in which she earns a payoff of one from choosing the action that matches the state, $\mathcal{A} = \{L, R\}$ and $u(a, \omega) = \mathbb{1}_{a=\omega}$, all types correctly interpret private signals, social types have correct beliefs about the relative frequency of each social type, and all types have common prior $p_0 = 1/2$.

We first show that the representative agent model is a good approximation when heterogeneity is sufficiently small.

Proposition 2. *Generically, for any average bias $\hat{\pi} \in (\pi(\theta_A), 1]$, there exists an $\varepsilon > 0$ such that if heterogeneity is sufficiently small, $|\hat{\pi}_1(\theta_A) - \hat{\pi}_2(\theta_A)| < \varepsilon$, then the heterogeneous and representative agent settings have the same set of long-run learning outcomes.*

This result illustrates the robustness of misspecified environments discussed in Theorem 5.

Next, we explore how heterogeneity affects learning. It is a priori unclear whether het-

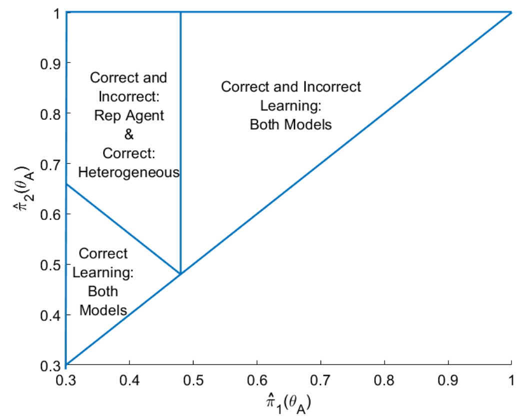

FIGURE 3. Naive Learning with a Representative Agent
$(\pi(\theta_A) = .3,\ F^L(s) = 2s - s^2,\ F^R(s) = s^2)$

erogeneity will facilitate or hinder learning, compared to the representative agent model. The type with milder misspecification may facilitate learning by counteracting the type with more severe misspecification, or the type with the more severe misspecification may distort information in a way that hinders learning for both types. The following result establishes that the first effect dominates and heterogeneity facilitates learning.

Proposition 3. *Suppose the signal distribution is symmetric, $F^L(s) = 1 - F^R(1-s)$. If learning is almost surely correct in the representative agent model with $\hat{\pi} \in (\pi(\theta_A), 1]$, then learning is almost surely correct in the heterogeneous model for all $\hat{\pi}_1(\theta_A) \in (\pi(\theta_A), 1]$ and $\hat{\pi}_2(\theta_A) \in (\hat{\pi}_1(\theta_A), 1]$ such that $(\hat{\pi}_1(\theta_A) + \hat{\pi}_2(\theta_A))/2 = \hat{\pi}$, and if incorrect learning occurs with positive probability in the heterogeneous model with $\hat{\pi}_1(\theta_A) \in (\pi(\theta_A), 1]$ and $\hat{\pi}_2(\theta_A) \in (\hat{\pi}_1(\theta_A), 1]$, then incorrect learning occurs with positive probability in the representative agent model with $\hat{\pi} \equiv (\hat{\pi}_1(\theta_A) + \hat{\pi}_2(\theta_A))/2$. For all $\hat{\pi}_1(\theta_A) \in (\pi(\theta_A), 1]$ and $\hat{\pi}_2(\theta_A) \in [\hat{\pi}_1(\theta_A), 1]$, almost surely learning is either correct or incorrect.*

Type θ_1 is more adept at correcting for correlated information, and as a result, asymptotically adopts the inefficient action with lower probability than θ_2. In turn, this helps θ_2 learn the true state. Actions from θ_1 confirm the state and θ_2 overestimates the private information reflected in these actions. This reduces the probability that θ_2 herds on an inefficient action.[37] Fig. 3 illustrates these learning regions.

This characterization allows us to precisely determine how much heterogeneity can be present before the representative agent model is no longer a good approximation. For the parameters considered in Fig. 3, when the average bias is low or high—for example, 0.4 or

[37]Heterogeneity does not always improve learning. If heterogeneity leads to fundamentally different biases—for example, if one type overestimates the correlation in prior actions and the other type underestimates it—then sufficient heterogeneity will interfere with long-run learning, even when the average bias is close to the truth (e.g. $\hat{\pi} \approx \pi(\theta_A)$) and learning is complete in the representative agent model.

0.8—any feasible level of heterogeneity results in the same set of learning outcomes as the corresponding representative agent model. At intermediate levels of the average bias—for example, 0.6—sufficient heterogeneity yields different learning outcomes than the representative agent model. In the knife-edge case of an average bias of 0.48, the representative agent model is not identified at certainty—representative agent models with $\hat{\pi} < 0.48$ are robust to any level of heterogeneity while representative agent models with $\hat{\pi}$ slightly above 0.48 are only robust to a very small level of heterogeneity.

Proposition 3 has important implications for policy interventions aimed at mitigating inefficient choices. Suppose a social planner wishes to intervene if and only if agents face the possibility of incorrect learning. The planner measures the average level of bias in the population and uses a representative agent approach to determine whether to intervene. Given Proposition 3, this method will result in *overintervention*, in that there are levels of bias at which the planner will intervene even though learning is almost surely correct. However, underintervention will not be an issue, as the planner will never fail to intervene when incorrect learning arises.

4.3 Entrenched Disagreement in a Level-k Learning Model

As discussed in the introduction, a central contribution of our framework is the ability to allow for model heterogeneity—either due to varying levels of the same bias, as illustrated in Section 4.2, or due to fundamentally distinct biases. This section demonstrates how our characterization can be used to explore whether agents with distinct models influence each others' learning and to determine when entrenched disagreement emerges. We illustrate this in the context of a social learning setting where agents use level-k reasoning. We show that entrenched disagreement emerges as a robust feature of this setting. Further, the presence of agents who use level-3 reasoning alters the learning outcomes of agents who use level-2 reasoning. This contrasts with settings that consider level-2 reasoning in isolation (Eyster and Rabin 2010; Bohren 2016), highlighting the potential error in predicting learning outcomes without properly accounting for the interaction between different models of inference.

Level-k models describe how boundedly rational agents draw inference in strategic settings (Costa-Gomes and Crawford 2006). Agents are characterized by their "depth" of reasoning, where higher levels use progressively more sophisticated reasoning. Applying this model of inference to a social learning setting, each level corresponds to a type with a misspecified model of the strategic link between prior actions and private signals. Level-0 is a noise type that chooses an action without learning from signals or the actions of others, i.e. $\hat{s}_0(s) = 1/2$ and $\hat{\pi}_0(\theta_0) = 1$. Higher levels accurately learn from signals but misinterpret actions, which is captured by a misspecified type distribution. Level-1 is an autarkic type who acts solely based on its private signal, i.e. it believes all agents are noise types,

$\hat{\pi}_1(\theta_0) = 1$. Level-2 fails to account for redundant information in prior actions, i.e. it believes all agents are autarkic types, $\hat{\pi}_2(\theta_1) = 1$.[38] Level-3 understands that prior actions contain redundant information, but does not allow for the possibility that other agents also account for this, i.e. it believes almost all other agents are level-2, $\hat{\pi}_3(\theta_2) = 1 - \varepsilon$ for some small $\varepsilon > 0$ (for technical reasons, we assume this type places arbitrarily small probability on the level-1 type, $\hat{\pi}_3(\theta_1) = \varepsilon$).[39] As is customary, the level-0 type anchors the model of level-1 but does not actually exist in the population, $\pi(\theta_0) = 0$. To close the model, assume that each agent faces a binary decision problem in which she earns a payoff of one from choosing the action that matches the state, $\mathcal{A} = \{L, R\}$ and $u(a, \omega) = \mathbb{1}_{a=\omega}$, the level-1 type occurs with positive probability, $\pi(\theta_1) \in (0, 1)$, there are no level-4 or higher types, and all types have common prior $p_0 = 1/2$.[40] Note that a correctly specified environment is not a special case of this set-up, as no type allows for the existence of its own type.

Although depth of reasoning models feature prominently in the empirical literature on social learning (Kübler and Weizsäcker 2004; Penczynski 2017), it has been relatively unexplored in the corresponding theoretical literature—despite the interest in naive learning—as characterizing learning outcomes is significantly more complex when agents learn in different ways.[41] The addition of level-3 adds two complications relative to a naive learning model with only level-2. First, it is necessary to characterize learning outcomes for multiple types simultaneously. Second, the presence of level-3 affects the learning of level-2, even though level-2 is not aware of level-3.

Proposition 4 establishes that either entrenched disagreement or cyclical learning almost surely arise, depending on the true distribution over types. Strikingly, agreement almost surely does not arise for any distribution over types.

Proposition 4. *There exists an $\overline{\varepsilon} > 0$ such that if $\varepsilon \in (0, \overline{\varepsilon})$, then either learning is cyclical almost surely or entrenched disagreement occurs almost surely. For $\varepsilon \in (0, \overline{\varepsilon})$, there exists a cutoff $\bar{\pi}_3 \in (0, 1)$ such that if $\pi(\theta_3) > \bar{\pi}_3$, then almost surely learning is cyclical, there exists a cutoff $\bar{\pi}_2 \in (0, 1)$ such that if $\pi(\theta_2) > \bar{\pi}_2$, then both disagreement outcomes arise with positive probability, and there exists a cutoff $\bar{\pi}_1 \in (0, 1)$ such that if $\pi(\theta_1) > \bar{\pi}_1$, then the*

[38] A level-2 type is analogous to the "BRTNI" agents in Eyster and Rabin (2010) and the "naive Bayesians" in Hung and Plott (2001). The naive learners in Bohren (2016) are a modified level-2 type that allows for the possibility of other level-2 agents. In Eyster and Rabin (2010), all agents have the same model—they are all level-2—while Bohren (2016), level-1 and level-2 types both occur with positive probability.

[39] The exact parameterization of the level-k model, i.e. $\varepsilon = 0$, violates Assumptions 3 and 4. In a cognitive hierarchy model (Camerer et al. 2004), level-3 places non-trivial probability on level-1 types. We explore this alternative parameterization in an earlier working paper version of this article (Bohren and Hauser 2019b).

[40] Our framework can allow for higher levels, but empirical studies rarely find evidence of such reasoning. For example, in a social learning experiment, Penczynski (2017) finds that most agents' behavior is consistent with level-1, 2 or 3 types, with a modal type of level-2.

[41] As far as we know, no theoretical papers characterize asymptotic learning in a level-k framework. While naive learners are analogous to the level-2 type, existing models focus on the case where all agents are level-2.

disagreement outcome in which level-2 learns the correct state and level-3 learns the incorrect state arises almost surely.

In contrast, cyclical learning does not arise in settings that examine level-2 reasoning in isolation (Eyster and Rabin 2010; Bohren 2016). Given the empirical evidence for both level-2 and level-3 reasoning in social learning settings (Penczynski 2017), the presence of both models of inference is important to take into account.

Disagreement is driven by level-2's imitation of the more frequent action and level-3's anti-imitation in order to correct for level-2's overreaction. If a large share of agents are level-1, then level-2's model is close to correct and almost surely level-2 agents learn the correct state. Therefore, the disagreement outcome in which level-2 agents learn the correct state and level-3 agents learn the incorrect state arises almost surely. Note that in this case, a higher level of reasoning performs strictly *worse* than a lower level of reasoning. Otherwise, if a large share of agents are level-2, then both disagreement outcomes emerge and learning is path-dependent. Therefore, two similar populations who learn from different action histories may converge to different forms of disagreement. If a large share of agents are level-3, then neither disagreement outcome is stable and learning is almost surely cyclical. Intuitively, near beliefs $(0, \infty)$ where level-2 agents choose L and level-3 agents choose R, a level-2 agent overreacts to the more frequently chosen R action, pulling her belief away from state L, and similarly for the other disagreement outcome. Correct and incorrect learning almost surely do not arise because the level-2 and level-3 agents have models that interpret actions in opposite ways, which prevents agreement.

Fig. 4 illustrates the learning regions for the level-k model, including the thresholds described in Proposition 4.[42] Penczynski (2017)'s estimate of the type distribution lies in the gray region in which both disagreement outcomes arise with positive probability.

5 Conclusion

We develop a general framework to study learning with model misspecification, which captures many information-processing biases and heuristics of interest in economic decision-making. A key contribution of our framework is the ability to allow for model heterogeneity, in which agents exhibit different levels of a bias or have distinct biases. Our main result characterizes the set of asymptotic learning outcomes based on two expressions that are straightforward to derive from the underlying form of misspecification. This characterization provides a unified way to compare different forms of misspecification that have been previously studied and yields new insights about forms of misspecification that have not been previously explored theoretically. The characterization also provides a rationale for

[42] In Online Appendix C.3, we show analytically that the qualitative features Fig. 4 hold for the full characterization of learning outcomes across $(\pi(\theta_1), \pi(\theta_2), \pi(\theta_3)) \in \Delta^2$; for expositional clarity, we state a partial characterization in Proposition 4.

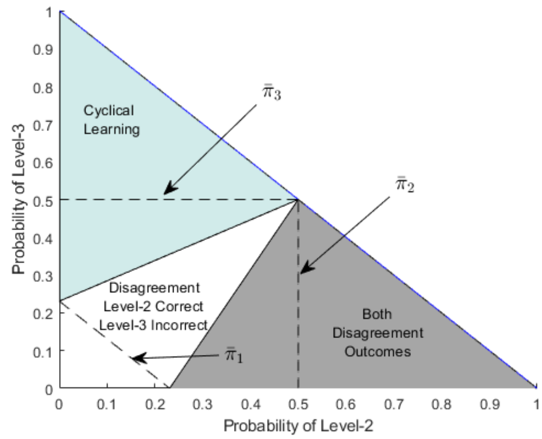

Figure 4. Entrenched Disagreement in Level-k Learning
($\omega = L$, $F^L = \frac{10}{3}(s - .5s^2) - .6$, $F^R(s) = \frac{5}{3}(s^2 - .04)$)

entrenched disagreement, in which agents with different models converge to different certain beliefs despite observing a common history. Our results yield insights into how the source of information (i.e. social versus private) impacts learning, whether learning predictions are sensitive to different parameterizations of a bias, and when a set of learning outcomes is robust to varying levels of misspecification. In the presence of model heterogeneity, our results can also be used to explore how different biases interact and determine whether a representative agent approach generates accurate learning predictions.

References

Angrisani, M., A. Guarino, P. Jehiel, and T. Kitagawa (2020): "Information Redundancy Neglect versus Overcondence: A Social Learning Experiment," *American Economic Journal: Microeconomics*.

Arrow, K. J. and J. R. Green (1973): "Notes on Expectations Equilibria in Bayesian Settings," *Institute for Mathematical Studies in the Social Sciences Working Papers*.

Ba, C. (2021): "Robust Model Misspecification and Paradigm Shift," .

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

Bartels, L. M. (2002): "Beyond the running tally: Partisan bias in political perceptions," *Political Behavior*, 24, 117–150.

Berk, R. H. (1966): "Limiting Behavior of Posterior Distributions When the Model is Incorrect," *The Annals of Mathematical Statistics*, 37, 51–58.

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

Bohren, A. (2016): "Informational Herding with Model Misspecification," *Journal of Economic Theory*, 222–247.

Bohren, J. A. and D. Hauser (2019a): "Misinterpreting Social Outcomes and Information Campaigns," .

Bohren, J. A. and D. N. Hauser (2019b): "Social Learning with Model Misspecification: A

Framework and a Characterization," Available at SSRN: https://ssrn.com/abstract=3236842 or http://dx.doi.org/10.2139/ssrn.3236842.

——— (2020): "Linking Heuristics and Misspecified Models," Mimeo.

Brunnermeier, M. K. and J. A. Parker (2005): "Optimal expectations," *The American Economic Review*, 95, 1092–1118.

Bushong, B. and T. Gagnon-Bartsch (2019): "Learning with Misattribution of Reference Dependence," .

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004): "A Cognitive Hierarchy Model of Games," *The Quarterly Journal of Economics*, 119, 861–898.

Costa-Gomes, M. and V. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 96, 1737 – 1768.

Cripps, M. W. (2018): "Divisible Updating," .

Enke, B. and F. Zimmermann (2019): "Correlation Neglect in Belief Formation," *The Review of Economic Studies*, 86, 313–332.

Epstein, L., J. Noor, and A. Sandroni (2008): "Non-Bayesian updating: a theoretical framework," *Theoretical Economics*, 3, 193–229.

Epstein, L. G., J. Noor, and A. Sandroni (2010): "Non-Bayesian Learning," *The B.E. Journal of Theoretical Economics*, 10.

Esponda, I. and D. Pouzo (2016): "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models," *Econometrica*, 84, 1093–1130.

——— (2019): "Equilibrium in Misspecified Markov Decision Processes," *Mimeo*.

Esponda, I., D. Pouzo, and Y. Yamamoto (2019): "Asymptotic Behavior of Bayesian Learners with Misspecified Models," .

Eyster, E. and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623–1672.

——— (2010): "Naive Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2, 221–243.

Eyster, E., M. Rabin, and G. Weizsäcker (2020): "An Experiment on Social Mislearning," .

Frick, M., R. Iijima, and Y. Ishii (2020a): "Dispersed Behavior and Perceptions in Assortative Societies," .

——— (2020b): "Misinterpreting Others and the Fragility of Social Learning," .

——— (2020c): "Stability and Robustness in Misspecified Learning Models," .

Fudenberg, D., G. Lanzani, and P. Strack (2020): "Limits Points of Endogenous Misspecified Learning," .

Fudenberg, D., G. Romanyuk, and P. Strack (2017): "Active learning with a misspecified prior," *Theoretical Economics*, 12, 1155–1189.

Gagnon-Bartsch, T. (2016): "Taste Projection in Models of Social Learning," *Mimeo*.

Gagnon-Bartsch, T. and M. Rabin (2016): "Naive Social Learning , Mislearning and Unlearning," .

Gagnon-Bartsch, T., M. Rabin, and J. Schwartzstein (2018): "Channeled attention and stable errors," .

Gottlieb, D. (2015): "Will you never learn? Self deception and biases in information processing," .

Guarino, A. and P. Jehiel (2013): "Social Learning with Coarse Inference," *American Economic Journal: Microeconomics*, 5, 147–174.

He, K. (2020): "Mislearning from Censored Data: The Gambler's Fallacy in Optimal-Stopping Problems," .

Heidhues, P., B. Kőszegi, and P. Strack (2018): "Unrealistic Expectations and Misguided

Learning," *Econometrica*, 86, 1159–1214.

——— (2019): "Convergence in Misspecified Learning Models with Endogenous Actions," .

Hung, A. A. and C. R. Plott (2001): "Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions," *American Economic Review*, 91, 1508–1520.

Jadbabaie, A., P. Molavi, A. Sandroni, and A. Tahbaz-salehi (2012): "Non-Bayesian social learning," *Games and Economic Behavior*, 76, 210–225.

Jehiel, P. (2005): "Analogy-based expectation equilibrium," *Journal of Economic Theory*, 123, 81–104.

Jerit, J. and J. Barabas (2012): "Partisan perceptual bias and the information environment," *The Journal of Politics*, 74, 672–684.

Kominers, S. D., X. Mu, and A. Peysakhovich (2020): "Paying (for) attention: The impact of information processing costs on bayesian inference," Mimeo.

Kőszegi, B. (2006): "Ego utility, overconfidence, and task choice," *Journal of the European Economic Association*, 4, 673–707.

Kübler, D. and G. Weizsäcker (2004): "Limited depth of reasoning and failure of cascade formation in the laboratory," *The Review of Economic Studies*, 71, 425–441.

Levy, G. and R. Razin (2015): "Correlation neglect, voting behavior, and information aggregation," *American Economic Review*, 105, 1634–45.

Madarász, K. and A. Prat (2016): "Sellers with misspecified models," *The Review of Economic Studies*.

Marks, G. and N. Miller (1987): "Ten years of research on the false-consensus effect: An empirical and theoretical review." *Psychological Bulletin*, 102, 72.

Mertens, J.-F. and S. Zamir (1985): "Formulation of Bayesian analysis for games with incomplete information," *International Journal of Game Theory*, 14, 1–29.

Miller, D. T. and C. McFarland (1987): "Pluralistic ignorance: When similarity is interpreted as dissimilarity." *Journal of Personality and Social Psychology*, 53, 298.

——— (1991): "When social comparison goes awry: The case of pluralistic ignorance," in *Social Comparison: Contemporary Theory and Research*, ed. by J. Suls and T. A. Wills, Lawrence Erlbaum Associates, Inc, 287–313.

Molavi, P., A. Tahbaz-Salehi, and A. Jadbabaie (2018): "A Theory of Non-Bayesian Social Learning," *Econometrica*, 86, 445–490.

Moore, D. A. and P. J. Healy (2008): "The trouble with overconfidence." *Psychological Review*, 115, 502–517.

Moore, D. A., E. R. Tenney, and U. Haran (2015): "Overprecision in Judgment," in *The Wiley Blackwell Handbook of Judgment and Decision Making*, John Wiley & Sons, Ltd, chap. 6, 182–209.

Nickerson, R. S. (1998): "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," *Review of General Psychology*, 2, 175–220.

Nyarko, Y. (1991): "Learning in Misspecified Models and the Possibility of Cycles," *Journal of Economic Theory*, 55, 416–427.

Ortoleva, P. (2012): "Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News," *American Economic Review*, 102, 2410–2436.

Ortoleva, P. and E. Snowberg (2015): "Overconfidence in political behavior," *American Economic Review*, 105, 504–535.

Pachur, T., J. Rieskamp, and R. Hertwig (2004): "The Social Circle Heuristic: Fast and Frugal Decisions Based on Small Samples," *Proceedings of the Annual Meeting of the Cognitive*

Science Society.
Penczynski, S. P. (2017): "The nature of social learning: Experimental evidence," *European Economic Review*, 94, 148–165.
Rabin, M. and J. L. Schrag (1999): "First Impressions Matter: A Model of Confirmatory Bias," *The Quarterly Journal of Economics*, 114, 37–82.
Ross, L., D. Greene, and P. House (1977): "The "false consensus effect": An egocentric bias in social perception and attribution processes," *Journal of experimental social psychology*, 13, 279–301.
Schwartzstein, J. (2014): "Selective Attention and Learning," *Journal of the European Economic Association*, 12, 1423–1452.
Shalizi, C. R. (2009): "Dynamics of Bayesian updating with dependent data and misspecified models," *Electronic Journal of Statistics*, 3, 1039–1074.
Smith, L. and P. N. Sørensen (2000): "Pathological Outcomes Observational Learning," *Econometrica*, 68, 371–398.
Spiegler, R. (2016): "Bayesian networks and boundedly rational expectations," *Quarterly Journal of Economics*, 131, 1243–1290.
Ungeheuer, M. and M. Weber (2020): "The Perception of Dependence, Investment Decisions, and Stock Prices," *Mimeo.*
Wilson, A. (2014): "Bounded Memory and Biases in Information Processing," *Econometrica*, 82, 2257–2294.

A Proofs from Section 3

Throughout this section, we use the following notation. Given $\varepsilon > 0$, define a neighborhood of $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ as $B_\varepsilon(\boldsymbol{\lambda}^*) \equiv \{\boldsymbol{\lambda} \in [0,\infty]^k | \lambda_i \in [0, \varepsilon)$ if $\lambda_i^* = 0$ and $\lambda_i \in (1/\varepsilon, \infty)$ if $\lambda_i^* = \infty\}$.

A.1 Proofs of Lemmas 1 to 3

Proof of Lemma 1. Let $\frac{p_i(\lambda_i(h),s)}{1-p_i(\lambda_i(h),s)} \equiv \frac{\hat{P}_i(R|h,s)}{\hat{P}_i(L|h,s)}$ denote the private belief of type θ_i following signal s and history h. Given $\lambda \in (0,\infty)$ and $s \in [0,1]$, consider type θ_i's choice at private belief $p_i(\lambda, s)$. Recall that actions $(a_1, ...a_M)$ are ordered by relative preference in state R. Since no two actions yield the same payoff in both states, no action is optimal at a single belief, and preferences are aligned (Assumption 2), there exist belief thresholds $0 = \overline{p}_{i,0} \leq \overline{p}_{i,1} \leq ... \leq \overline{p}_{i,M} = 1$ such that we can partition the belief space into a finite set of closed intervals, with action a_m optimal at $p_i(\lambda, s)$ if $p_i(\lambda, s) \in [\overline{p}_{i,m-1}, \overline{p}_{i,m}]$ and $\overline{p}_{i,m-1} \neq \overline{p}_{i,m}$, and a_m never optimal iff $\overline{p}_{i,m-1} = \overline{p}_{i,m}$. Without loss of generality, assume the tie-breaking rule is to choose the optimal action with the lower index at each interior cut-off $\overline{p}_{i,m} \in (0,1)$, i.e. if $\overline{p}_{i,m-1} \neq \overline{p}_{i,m}$, choose a_m over any other optimal action $a_{m+\kappa}$ at belief $\overline{p}_{i,m}$. Since there are at least two undominated actions, there are at least two intervals with a non-empty interior.

Since signals are aligned (Assumption 1), when $\theta_i \in \Theta_S \cup \Theta_A$, $p_i(\lambda, s)$ is strictly increasing in s for all $\lambda \in (0,\infty)$. Therefore, for each $\lambda \in (0,\infty)$, we can define the decision rule with respect to signal cut-offs $0 = \overline{s}_{i,0}(\lambda) \leq \overline{s}_{i,1}(\lambda) \leq ... \leq \overline{s}_{i,M}(\lambda) = 1$ such that the agent chooses action a_m at likelihood ratio λ iff $\overline{s}_{i,m-1}(\lambda) \neq \overline{s}_{i,m}(\lambda)$ and she observes private signal

$s \in (\bar{s}_{i,m-1}(\lambda), \bar{s}_{i,m}(\lambda)]$, with a closed interval if $\bar{s}_{i,m-1}(\lambda) = 0$. In the case of $\theta_i \in \Theta_N$, $p_i(\lambda, s)$ is constant with respect to s for all $\lambda \in (0, \infty)$. Therefore, given λ, a single action is optimal for θ_i at all signals. We define an analogous set of signal cut-offs: if a_m is optimal at likelihood ratio $\lambda \in (0, \infty)$, then the signal cut-offs are $\bar{s}_{i,0}(\lambda) = ... = \bar{s}_{i,m-1}(\lambda) = 0$ and $\bar{s}_{i,m}(\lambda) = ... = \bar{s}_{i,M}(\lambda) = 1$. Similarly, if $\lambda \in \{0, 1\}$, a single action is optimal for $\theta_i \in \Theta$ at all signals. We define the cut-offs for this case analogously to the noise type. □

Proof of Lemma 2. We first show that $\hat{\psi}_i(a_1|R, \boldsymbol{\lambda}) \leq \hat{\psi}_i(a_1|L, \boldsymbol{\lambda})$ for all $\boldsymbol{\lambda} \in [0, \infty]^k$ and social types $\theta_i \in \Theta_S$. Fix $\boldsymbol{\lambda} \in [0, \infty]^k$ and consider how social type θ_i updates its beliefs following a_1. By Lemma 1, θ_i believes that type θ_j plays a_1 with probability $\hat{F}_i^\omega(\bar{s}_{j,1}(\lambda_j))$. By Lemma A.1 in Smith and Sørensen (2000), $F^R(s) \leq F^L(s)$, with strict inequality for $s \in \text{int}(\mathcal{S})$. By Assumption 1, this is also true for the subjective signal distributions. Therefore, $\hat{F}_i^R(\bar{s}_{j,1}(\lambda_j)) \leq \hat{F}_i^L(\bar{s}_{j,1}(\lambda_j))$. This implies $\hat{\psi}_i(a_1|R, \boldsymbol{\lambda}) \leq \hat{\psi}_i(a_1|L, \boldsymbol{\lambda})$, since $\hat{\psi}_i(a|\omega, \boldsymbol{\lambda})$ is a convex combination of $\hat{F}_i^\omega(\bar{s}_{j,1}(\lambda_j))$ for each $\theta_j \in \Theta$.

We next establish the uniform bound. Recall that autarkic types have a likelihood ratio that is constant and equal to $p_0/(1-p_0)$. Therefore, for social type θ_i,

$$\begin{aligned}
\frac{\hat{\psi}_i(a_1|R, \boldsymbol{\lambda})}{\hat{\psi}_i(a_1|L, \boldsymbol{\lambda})} &= \frac{\sum_{\theta_j \in \Theta_A} \hat{\pi}_i(\theta_j) \hat{F}_i^R(\bar{s}_{j,1}(\frac{p_0}{1-p_0})) + \sum_{\theta_j \in \Theta_S} \hat{\pi}_i(\theta_j) \hat{F}_i^R(\bar{s}_{j,1}(\lambda_j))}{\sum_{\theta_j \in \Theta_A} \hat{\pi}_i(\theta_j) \hat{F}_i^L(\bar{s}_{j,1}(\frac{p_0}{1-p_0})) + \sum_{\theta_j \in \Theta_A} \hat{\pi}_i(\theta_j) \hat{F}_i^L(\bar{s}_{j,1}(\lambda_j))} \\
&\leq \frac{\sum_{\theta_j \in \Theta_A} \hat{\pi}_i(\theta_j) \hat{F}_i^R(\bar{s}_{j,1}(\frac{p_0}{1-p_0})) + \sum_{\theta_j \in \Theta_S} \hat{\pi}_i(\theta_j) \hat{F}_i^L(\bar{s}_{j,1}(\lambda_j))}{\sum_{\theta_j \in \Theta_A} \hat{\pi}_i(\theta_j) \hat{F}_i^L(\bar{s}_{j,1}(\frac{p_0}{1-p_0})) + \sum_{\theta_j \in \Theta_S} \hat{\pi}_i(\theta_j) \hat{F}_i^L(\bar{s}_{j,1}(\lambda_j))} \\
&\leq \frac{\sum_{\theta_j \in \Theta_A} \hat{\pi}_i(\theta_j) \hat{F}_i^R(\bar{s}_{j,1}(\frac{p_0}{1-p_0})) + \hat{\pi}_i(\Theta_S)}{\sum_{\theta_j \in \Theta_A} \hat{\pi}_i(\theta_j) \hat{F}_i^L(\bar{s}_{j,1}(\frac{p_0}{1-p_0})) + \hat{\pi}_i(\Theta_S)} < 1 \qquad (9)
\end{aligned}$$

where the first line follows by definition, the second line follows from $\hat{F}_i^R(s) \leq \hat{F}_i^L(s)$, the third line follows from $\sum_{\theta_j \in \Theta_S} \hat{\pi}_i(\theta_j) \hat{F}_i^L(s) \leq \hat{\pi}_i(\Theta_S)$ and $\hat{F}_i^R(s) \leq \hat{F}_i^L(s)$, and finally, the bound of one follows from Assumption 3, which ensures there exists at least one autarkic type $\theta_j \in \Theta_A$ with $\hat{\pi}_i(\theta_j) > 0$ and $\bar{s}_{j,1}(\frac{p_0}{1-p_0}) \in \text{int}(\mathcal{S})$, and therefore, $\hat{F}_i^R(\bar{s}_{j,1}(\frac{p_0}{1-p_0})) < \hat{F}_i^L(\bar{s}_{j,1}(\frac{p_0}{1-p_0}))$. Therefore, $\frac{\hat{\psi}_i(a_1|R,\boldsymbol{\lambda})}{\hat{\psi}_i(a_1|L,\boldsymbol{\lambda})}$ is uniformly bounded away from one. Similar logic holds for the case of a_M.

Finally, we establish continuity. Consider $\boldsymbol{\lambda}^* = 0^k$. Each type $\theta_i \in \Theta_S$ has a unique optimal action at 0^k, independent of the realization of the private signal. Moreover, since no action is optimal at a single belief, there exists an $\varepsilon_1 > 0$ such that if the posterior belief following the private signal is in $[0, \varepsilon_1)^k$, each type plays this action. Let Θ_a denote the set of social types who play a at 0^k. Fix $\varepsilon > 0$. Let $\delta_1 \equiv \min_{a \in \mathcal{A}} \frac{\varepsilon}{\max\{\pi(\Theta_S \backslash \Theta_a), \pi(\Theta_a)\}}$, $\delta_2 \equiv \min_{a \in \mathcal{A}, \theta_i \in \Theta_S} \frac{\varepsilon}{\max\{\hat{\pi}^i(\Theta_S \backslash \Theta_a), \hat{\pi}^i(\Theta_a)\}}$, and $\delta \equiv \min\{\delta_1, \delta_2\}$. Signals are not perfectly informative, so there exists a $\bar{s}$ such that $1 - F^\omega(\bar{s}) < \delta$ and $1 - \hat{F}_i^\omega(\bar{s}) < \delta$ for all $\theta_i \in \Theta_S$ and $\omega \in \{L, R\}$.

Define $\varepsilon_1(\delta) \equiv \frac{\varepsilon_1}{\max_{\theta_i \in \Theta_S} \hat{s}_i(\bar{s})/(1-\hat{s}_i(\bar{s}))}$. Fix an action $a \in \mathcal{A}$ and let q_a denote the probability that a type is autarkic and plays action a. If $\boldsymbol{\lambda} \in [0, \varepsilon_1(\delta))$, then the probability of action a in state ω is bounded above by $\pi(\Theta_a)+\delta\pi(\Theta_S \setminus \Theta_a)+q_a$ and bounded below by $\pi(\Theta_a)(1-\delta)+q_a$. So $|\psi(a|\omega, \boldsymbol{\lambda}) - \psi(a|\omega, 0^k)| \leq \varepsilon$ for all $\boldsymbol{\lambda} \in [0, \varepsilon_1(\delta))^k$. Similarly $|\hat{\psi}_i(a|\omega, \boldsymbol{\lambda}) - \hat{\psi}_i(a|\omega, 0^k)| \leq \varepsilon$ for all $\boldsymbol{\lambda} \in [0, \varepsilon_1(\delta))^k$ and $\theta_i \in \Theta_S$. The proofs for other stationary beliefs are identical. □

Proof of Lemma 3. At a stationary belief $\boldsymbol{\lambda}^* \in [0, \infty]^k$, $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^* \frac{\hat{\psi}_i(a|R, \boldsymbol{\lambda}^*)}{\hat{\psi}_i(a|L, \boldsymbol{\lambda}^*)}$ for all a such that $\psi(a|\omega, \boldsymbol{\lambda}^*) > 0$. Trivially, this is satisfied for all $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$, independent of $\psi(a|\omega, \boldsymbol{\lambda}^*)$, and these beliefs are stationary. It remains to be determined whether it is satisfied for any interior beliefs $\boldsymbol{\lambda}^* \in (0, \infty)^k$. Suppose $\boldsymbol{\lambda}^* \in (0, \infty)^k$. By Assumption 3, $\psi(a_1|\omega, \boldsymbol{\lambda}^*) > 0$ for each $\omega \in \{L, R\}$. By Lemma 2, $\frac{\hat{\psi}_i(a_1|R, \boldsymbol{\lambda}^*)}{\hat{\psi}_i(a_1|L, \boldsymbol{\lambda}^*)} < 1$. Therefore, this does not hold for a_1 and $\boldsymbol{\lambda}^*$ cannot be stationary.

Suppose beliefs converge to a non-stationary belief $\boldsymbol{\lambda}^* \in [0, \infty]^k \setminus \{0, \infty\}^k$ with positive probability. By Lemma 2, following action $\tilde{a}_t = a_M$, $\lambda_{i,t+1} - \lambda_{i,t}$ is bounded uniformly away from zero for all social types $\theta_i \in \Theta_S$. For sufficiently small $\varepsilon > 0$, when $\boldsymbol{\lambda}_t \in B_\varepsilon(\boldsymbol{\lambda}^*)$, $\lambda_{i,t+1} \notin B_\varepsilon(\boldsymbol{\lambda}^*)$ for any type with an interior belief $\lambda_{i,t} \in (0, \infty)$. The probability $Pr(\exists t < T\, s.t.\, \tilde{a}_t = a_M)$ converges to one as $T \to \infty$. Therefore, the likelihood ratio almost surely leaves $B_\varepsilon(\boldsymbol{\lambda}^*)$. □

A.2 Local Stability

Proof of Theorem 1. Without loss of generality, consider the case where $\omega = L$. The proof for $\omega = R$ is analogous.

Part 1. We first show that if $\boldsymbol{\lambda}^* \in \Lambda(L)$, then $\boldsymbol{\lambda}^*$ is locally stable. Consider $\boldsymbol{\lambda}^* = 0^k$ and suppose $\gamma_i(L, 0^k) < 0$ for all social types $\theta_i \in \Theta_S$. Then there exists a $\varepsilon > 0$ such that in the neighborhood $B_\varepsilon(0^k) \equiv [0, \varepsilon]^k$ of 0^k,

$$\sum_{a \in \mathcal{A}} \psi(a|L, 0^k) \sup_{\boldsymbol{\lambda} \in [0,\varepsilon]^k} \log \frac{\hat{\psi}_i(a|R, \boldsymbol{\lambda})}{\hat{\psi}_i(a|L, \boldsymbol{\lambda})} < 0. \tag{10}$$

for all $\theta_i \in \Theta_S$. Let $g_i(a) \equiv \sup_{\boldsymbol{\lambda} \in [0,\varepsilon]^k} \log \frac{\hat{\psi}_i(a|R, \boldsymbol{\lambda})}{\hat{\psi}_i(a|L, \boldsymbol{\lambda})}$ denote the maximal update from action a in the neighborhood $[0, \varepsilon]^k$, with $\mathbf{g}(a) \equiv (g_1(a), ..., g_k(a))$. Let $\bar{g}_i \equiv \max_{a \in \mathcal{A}} g_i(a)$ denote the maximal update across all actions in the neighborhood $[0, \varepsilon]^k$, with $\bar{\mathbf{g}} \equiv (\bar{g}_1, ..., \bar{g}_k)$.

For $\delta > 0$, choose a neighborhood $[0, \varepsilon_\delta]^k \subseteq [0, \varepsilon]^k$ with $\sup_{\boldsymbol{\lambda} \in [0,\varepsilon_\delta]^k} |\psi(a|L, \boldsymbol{\lambda}) - \psi(a|L, 0^k)| < \delta$. By Lemma 2, $\psi(a|L, \boldsymbol{\lambda})$ is continuous at $\boldsymbol{\lambda} = 0^k$, so such a neighborhood exists. Suppose $\boldsymbol{\lambda}_1 \in [0, \varepsilon_\delta]^k$. Let $a(\theta, s, \boldsymbol{\lambda})$ be the optimal action for type θ at beliefs $\boldsymbol{\lambda}$ after observing private signal s. Define the linear system $\langle \boldsymbol{\lambda}_{\delta,t} \rangle_{t=1}^\infty$ as follows: $\boldsymbol{\lambda}_{\delta,1} = \boldsymbol{\lambda}_1$, $\log \boldsymbol{\lambda}_{\delta,t+1} = \log \boldsymbol{\lambda}_{\delta,t} + \mathbf{g}(a(\tilde{\theta}_t, \tilde{s}_t, 0^k))$, when $(\tilde{\theta}_t, \tilde{s}_t)$ is such that $a(\tilde{\theta}_t, \tilde{s}_t, \boldsymbol{\lambda}) = a(\tilde{\theta}_t, \tilde{s}_t, 0^k)$ for all beliefs $\boldsymbol{\lambda} \in [0, \varepsilon_\delta]$ (note this includes all autarkic types), and $\log \boldsymbol{\lambda}_{\delta,t+1} = \log \boldsymbol{\lambda}_{\delta,t} + \bar{\mathbf{g}}$ otherwise. When $\omega = L$, let $\psi_\delta(a)$ be the probability of a in the former event and let $\bar{\psi}_\delta$ be the probability of the latter event. Note $\psi_\delta(a) \le \inf_{\boldsymbol{\lambda}\in[0,\varepsilon_\delta]^k} \psi(a|L,\boldsymbol{\lambda})$ and $\bar{\psi}_\delta + \sum_{a\in\mathcal{A}} \psi_\delta(a|L) = 1$. By Lemma C.1 of Smith and Sørensen (2000), if

$$\bar{\psi}_\delta \bar{g}_i + \sum_{a\in\mathcal{A}} \psi_\delta(a) g_i(a) < 0 \tag{11}$$

for all $\theta_i \in \Theta_S$, then almost surely $\lim_{t\to\infty} \boldsymbol{\lambda}_{\delta,t} = 0^k$. Eq. (11) holds for sufficiently small δ, since by Eq. (10), it is strictly less than zero as $\delta \to 0$.

Let $\delta_1 > 0$ denote an upper bound such that Eq. (11) holds for all $\delta < \delta_1$. Whenever $(\tilde{\theta}_t, \tilde{s}_t)$ is such that $a(\tilde{\theta}_t, \tilde{s}_t, \boldsymbol{\lambda}) = a(\tilde{\theta}_t, \tilde{s}_t, 0^k)$ for all $\boldsymbol{\lambda} \in [0, \varepsilon_\delta]$, the process $\langle \log \boldsymbol{\lambda}_{\delta,t} \rangle$ updates by $\mathbf{g}(a)$. When $\boldsymbol{\lambda}_t \in [0, \varepsilon_\delta]^k$, by construction this is larger than the update to the process $\langle \log \boldsymbol{\lambda}_t \rangle$, which is $\log \frac{\hat{\psi}_i(a|R,\boldsymbol{\lambda}_t)}{\hat{\psi}_i(a|L,\boldsymbol{\lambda}_t)}$ for each type $\theta_i \in \Theta_S$. Otherwise, $\langle \log \boldsymbol{\lambda}_{\delta,t} \rangle$ updates by $\bar{\mathbf{g}}$, which is also larger than the update to $\langle \log \boldsymbol{\lambda}_t \rangle$ when $\boldsymbol{\lambda}_t \in [0, \varepsilon_\delta]^k$. Therefore, for $\delta < \delta_1$, if $\boldsymbol{\lambda}_{\delta,t} \ge \boldsymbol{\lambda}_t$ and $\boldsymbol{\lambda}_{\delta,t} \in [0, \varepsilon_\delta]^k$, then $\boldsymbol{\lambda}_{\delta,t+1} \ge \boldsymbol{\lambda}_{t+1}$. Since $\boldsymbol{\lambda}_{\delta,1} \in [0, \varepsilon_\delta]^k$, as long as it remains in $[0, \varepsilon_\delta]^k$, $\langle \boldsymbol{\lambda}_t \rangle$ is bounded above by a stochastic process that converges to zero almost surely.

Since $\lim_{t\to\infty} \boldsymbol{\lambda}_{\delta,t} = 0^k$ almost surely for $\delta < \delta_1$, $Pr(\cup_t \cap_{s\ge t} \{\boldsymbol{\lambda}_{\delta,s} \in [0, \varepsilon_\delta]^k\}) = 1$. Therefore, there exists a $t \ge 1$ such that $Pr(\forall s \ge t, \boldsymbol{\lambda}_{\delta,s} \in [0, \varepsilon_\delta]^k) > 0$. Since the system is linear, if this holds at some $t > 1$, it must hold at $t = 1$. Therefore, there exists some $\boldsymbol{\lambda}_{\delta,1} \in [0, \varepsilon_\delta]^k$, with positive probability, $\boldsymbol{\lambda}_{\delta,t}$ remains in $[0, \varepsilon_\delta]^k$ for all $t > 1$ and $\boldsymbol{\lambda}_t \le \boldsymbol{\lambda}_{\delta,t}$. Moreover, this holds for all $\boldsymbol{\lambda} \le \boldsymbol{\lambda}_{\delta,1}$. When this happens, since $\lim_{t\to\infty} \boldsymbol{\lambda}_{\delta,t} = 0^k$, it must also be that $\lim_{t\to\infty} \boldsymbol{\lambda}_t = 0^k$. Let $\varepsilon^* = \min_{\theta_i\in\Theta_S} (\boldsymbol{\lambda}_{\delta,1})_i$ denote the minimum component of $\boldsymbol{\lambda}_{\delta,1}$. This establishes that when $\boldsymbol{\lambda}_1 \in [0, \varepsilon^*]^k$, with positive probability, $\lim_{t\to\infty} \boldsymbol{\lambda}_t = 0^k$ i.e. $\boldsymbol{\lambda}^* = 0^k$ is locally stable.

The proofs for the other stationary beliefs $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ are analogous. If $\lambda_i^* = \infty$, substitute λ_i^{-1} for type θ_i and modify the transition rules accordingly.

Part 2. We next show that if $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ and $\boldsymbol{\lambda}^* \notin \Lambda(L)$, then $\boldsymbol{\lambda}^*$ is not locally stable and $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$. Let $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ and suppose that there exists a type, which without loss of generality we denote θ_1, such that $\lambda_1^* = 0$ but $\gamma_1(L, \boldsymbol{\lambda}^*) > 0$. Without loss of generality, suppose the types are ordered so that the first κ types correspond to $\lambda_i^* = 0$ and the latter $k - \kappa$ types correspond to $\lambda_i^* = \infty$. Since $\gamma_1(L, \boldsymbol{\lambda}^*) > 0$, there exists a $\varepsilon > 0$ such that for neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*) \equiv [0, \varepsilon]^\kappa \times [1/\varepsilon, \infty]^{k-\kappa}$ of $\boldsymbol{\lambda}^*$,

$$\sum_{a\in\mathcal{A}} \psi(a|L, \boldsymbol{\lambda}^*) \inf_{\boldsymbol{\lambda}\in B_\varepsilon(\boldsymbol{\lambda}^*)} \log \frac{\hat{\psi}_1(a|R, \boldsymbol{\lambda})}{\hat{\psi}_1(a|L, \boldsymbol{\lambda})} > 0. \tag{12}$$

Let $\tau_\varepsilon \equiv \min\{\tau | \boldsymbol{\lambda}_t \in B_\varepsilon(\boldsymbol{\lambda}^*) \ \forall t \ge \tau\}$ be the first time at which beliefs enter $B_\varepsilon(\boldsymbol{\lambda}^*)$ and never exit. Suppose $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) > 0$. Then for all $\varepsilon > 0$, $\tau_\varepsilon < \infty$ with positive probability. We will reach a contradiction by showing that for small enough ε, $\tau_\varepsilon = \infty$ almost surely. Let $g_1(a) \equiv \inf_{\boldsymbol{\lambda}\in B_\varepsilon(\boldsymbol{\lambda}^*)} \log \frac{\hat{\psi}_1(a|R,\boldsymbol{\lambda})}{\hat{\psi}_1(a|L,\boldsymbol{\lambda})}$ denote the minimal update for type θ_1 following action

a in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*)$ and let $\underline{g}_1 \equiv \min_{a\in\mathcal{A}} g_1(a)$ denote the minimal update across all actions in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*)$. Suppose $\boldsymbol{\lambda}_\tau \in B_\varepsilon(\boldsymbol{\lambda}^*)$ for some time τ (if such a τ doesn't exist, then clearly $\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*$ is not possible along such a sample path). As above, let $a(\theta, s, \boldsymbol{\lambda})$ be the optimal action for type θ at beliefs $\boldsymbol{\lambda}$ after observing private signal s. Define a linear system $\langle\tilde{\lambda}_t\rangle_{t=1}^\infty$ as follows: let $\tilde{\lambda}_t$ denote the likelihood ratio for θ_1 at time $t \le \tau$, i.e. $\tilde{\lambda}_\tau = \lambda_{1,t}$ for all $t \le \tau$, and for $t > \tau$, $\log\tilde{\lambda}_{t+1} = \log\tilde{\lambda}_t + g_1(a(\tilde{\theta}_t, \tilde{s}_t, \boldsymbol{\lambda}^*))$ when $(\tilde{\theta}_t, \tilde{s}_t)$ is such that $a(\tilde{\theta}_t, \tilde{s}_t, \boldsymbol{\lambda}) = a(\tilde{\theta}_t, \tilde{s}_t, \boldsymbol{\lambda}^*)$ for all beliefs $\boldsymbol{\lambda} \in B_\varepsilon(\boldsymbol{\lambda}^*)$ (note this includes all autarkic types), and $\log\tilde{\lambda}_{t+1} = \log\tilde{\lambda}_t + \underline{g}_1$ otherwise. In other words, for $t > \tau$, following action a, the process is updated by type θ_1's minimum update for a across all beliefs in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*)$. When $\omega = L$, let $\psi(a)$ be the probability of a in the former event and let $\underline{\psi}$ be the probability of the latter event. Note $\underline{\psi} + \sum_{a\in\mathcal{A}}\psi(a) = 1$. Choose ε sufficiently small so that

$$\underline{\psi}\,\underline{g}_1 + \sum_{a\in\mathcal{A}}\psi(a)g_1(a) > 0. \tag{13}$$

Given Eq. (12), Eq. (13) is strictly greater than zero at $\varepsilon = 0$, so such an ε exists. Moreover, $(\log\tilde{\lambda}_{t+1} - \log\tilde{\lambda}_t)_{t=\tau}^\infty$ is an i.i.d. process with expectation equal to Eq. (13). By the Law of Large Numbers, almost surely, $\frac{1}{t}(\log\tilde{\lambda}_{t+1} - \log\tilde{\lambda}_t)$ converges to Eq. (13), which is positive. Therefore,

$$\lim_{t\to\infty}\log\tilde{\lambda}_t = \lim_{t\to\infty}\left(\log\lambda_{1,\tau} + \sum_{s=\tau}^{t}(\log\tilde{\lambda}_{s+1} - \log\tilde{\lambda}_s)\right) \to \infty.$$

By definition of $\langle\tilde{\lambda}_t\rangle$, if $\lambda_{1,t} \ge \tilde{\lambda}_t$ and $\boldsymbol{\lambda}_t \in B_\varepsilon(\boldsymbol{\lambda}^*)$, then $\lambda_{1,t+1} \ge \tilde{\lambda}_{1,t+1}$. Since $\boldsymbol{\lambda}_\tau \in B_\varepsilon(\boldsymbol{\lambda}^*)$, as long as $\langle\boldsymbol{\lambda}_t\rangle$ remains in $B_\varepsilon(\boldsymbol{\lambda}^*)$ for $t > \tau$, $\langle\lambda_{1,t}\rangle$ is bounded below by the stochastic process $\langle\tilde{\lambda}_t\rangle$. Therefore, if $\langle\boldsymbol{\lambda}_t\rangle$ remains in $B_\varepsilon(\boldsymbol{\lambda}^*)$ for all $t > \tau$ $\lim_{t\to\infty}\log\lambda_{1,t} \ge \lim_{t\to\infty}\log\tilde{\lambda}_t \to \infty$. But this implies that for small enough ε, $\boldsymbol{\lambda}_t \notin B_\varepsilon(\boldsymbol{\lambda}^*)$ for some $t > \tau$. This is a contradiction. So it must be that for small enough ε, $\tau_\varepsilon = \infty$ almost surely. Therefore, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$.

Similar logic establishes that for stationary $\boldsymbol{\lambda}^*$ such that $\lambda_1^* = \infty$ and $\gamma_1(L, \boldsymbol{\lambda}^*) < 0$, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$.

Part 3. Finally, Lemma 3 established that $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$ when $\boldsymbol{\lambda}^* \in (0,\infty)^k$. Therefore, if $\boldsymbol{\lambda}^* \notin \Lambda(L)$ and $\boldsymbol{\lambda}^* \in (0,\infty)^k$, then $\boldsymbol{\lambda}^*$ is not locally stable. □

A.3 Global Stability

Preliminary Notation. From Theorem 1, if $\boldsymbol{\lambda}^* \in \Lambda(\omega)$, then $\boldsymbol{\lambda}^*$ is locally stable, i.e. there exists an $\varepsilon(\boldsymbol{\lambda}^*) \in (0,1)$ and a *stable* neighborhood $B_{\varepsilon(\boldsymbol{\lambda}^*)}(\boldsymbol{\lambda}^*)$ such that when $\boldsymbol{\lambda}_1 \in B_{\varepsilon(\boldsymbol{\lambda}^*)}(\boldsymbol{\lambda}^*)$, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) > 0$. Also, for each stationary belief $\boldsymbol{\lambda}^* \notin \Lambda(\omega)$, there exists an $\varepsilon(\boldsymbol{\lambda}^*) \in (0,1)$ and an *unstable* neighborhood $B_{\varepsilon(\boldsymbol{\lambda}^*)}(\boldsymbol{\lambda}^*)$ such that when $\boldsymbol{\lambda}_1 \in B_{\varepsilon(\boldsymbol{\lambda}^*)}(\boldsymbol{\lambda}^*)$, $\langle\boldsymbol{\lambda}_t\rangle$ almost surely leaves this neighborhood. Fix state ω and define $E > 0$ as

$$E \equiv \min_{\boldsymbol{\lambda}^* \in \{0,\infty\}^k} -\log \varepsilon(\boldsymbol{\lambda}^*). \tag{14}$$

That is, if $\log \lambda_i \in \mathbb{R} \cup \{-\infty, \infty\} \setminus [-E, E]$ for each $\theta_i \in \Theta_S$, then $\boldsymbol{\lambda}$ is contained in one of these stable or unstable neighborhoods. Let

$$B_E(\boldsymbol{\lambda}^*) \equiv \{\boldsymbol{\lambda} \in [0,\infty]^k | \log \lambda_i < -E \text{ if } \lambda_i^* = 0 \text{ and } \log \lambda_i > E \text{ if } \lambda_i^* = \infty\} \tag{15}$$

denote the corresponding neighborhood for each stationary $\boldsymbol{\lambda}^*$, where in a slight abuse of notation, we switch from the neighborhood subscript denoting the bound for the likelihood ratio to denoting the bound for the log likelihood ratio to simplify notation in subsequent lemmas. Let $\mathcal{B} \equiv \cup_{\boldsymbol{\lambda}^* \in \Lambda(\omega)} B_E(\boldsymbol{\lambda}^*)$ denote the union of the stable neighborhoods and let $\mathcal{B}_U \equiv \cup_{\boldsymbol{\lambda}^* \in \{0,\infty\}^k \setminus \Lambda(\omega)} B_E(\boldsymbol{\lambda}^*)$ denote the union of the unstable neighborhoods. We will use these neighborhoods in subsequent proofs.

Proof of Theorem 2. Suppose the agreement outcome is locally stable, $0^k \in \Lambda(\omega)$. By Assumption 3, a_1 occurs with positive probability and by Lemma 2, observing a_1 decreases the likelihood ratio. Given initial likelihood ratio $\boldsymbol{\lambda}_1 \in (0,\infty)^k$, let N be the minimum number of consecutive a_1 actions required for the likelihood ratio to reach the stable neighborhood defined in Eq. (15), i.e. $\boldsymbol{\lambda}_{N+1} \in B_E(0^k)$. By Lemma 2, the change in the likelihood ratio following a_1 is bounded away from zero. Therefore, $N < \infty$. Further, given a_1 occurs with positive probability each period, the probability of a_1 occurring N times is strictly positive. Let $\tau_1 \equiv \min\{t | \boldsymbol{\lambda}_t \in B_E(0^k)\}$ be the first time that $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_E(0^k)$, $\tau_2 \equiv \min\{t > \tau_1 | \boldsymbol{\lambda}_t \notin B_E(0^k)\}$ be the first that $\langle \boldsymbol{\lambda}_t \rangle$ leaves $B_E(0^k)$ after entering, and $\tau_3 \equiv \min\{\tau | \boldsymbol{\lambda}_t \in B_E(0^k) \ \forall t \geq \tau\}$ be the first time the likelihood ratio enters $B_E(0^k)$ and never leaves. We know that $Pr(\tau_1 < \infty) > 0$, since the probability of transitioning from $\boldsymbol{\lambda}_1$ to $B_E(0^k)$ is bounded below by the probability of initially observing N consecutive a_1 actions. Also, $Pr(\tau_2 = \infty) > 0$, since by local stability, when the likelihood ratio is in $B_E(0^k)$, with positive probability, it never leaves. Therefore, $Pr(\tau_3 < \infty) > Pr(\tau_1 < \infty \wedge \tau_2 = \infty) > 0$. Therefore, with positive probability, the likelihood ratio eventually enters and remains in $B_E(0^k)$. By Theorem 1, if the likelihood ratio remains in $B_E(0^k)$ for all t, beliefs almost surely converge to 0^k. Therefore, if $0^k \in \Lambda(\omega)$, then from any initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^k$, $Pr(\boldsymbol{\lambda}_t \to 0^k) > 0$. The proof for agreement outcome ∞^k is analogous. □

Global Stability of Disagreement. We first state a result that uses a much weaker but more complicated to verify condition called *separability* to establish the global stability of a disagreement outcome (see Theorem 7 below). Starting from any initial belief, separability uses all of the actions to separate the beliefs of the different types and reach a neighborhood of the disagreement outcome. Therefore, together with local stability, the separability condition implies global stability. We then prove Theorem 3 by showing that maximal accessibility—which is easier to verify but can only use actions a_1 and a_M to separate beliefs—implies the separability condition.

Define $\Psi(\boldsymbol{\lambda})$ as the matrix consisting of the log of the ratios of the perceived action probabilities in each state at belief $\boldsymbol{\lambda}$, where each row corresponds to the ratios for social type θ_i and each column corresponds to the ratios for action a_m,

$$\Psi(\boldsymbol{\lambda})_{im} \equiv \log \frac{\hat{\psi}_i(a_m|R,\boldsymbol{\lambda})}{\hat{\psi}_i(a_m|L,\boldsymbol{\lambda})}, \tag{16}$$

and define the submatrix

$$\Psi[\theta_i,\theta_j;a_1,a_M](\boldsymbol{\lambda}) \equiv \begin{pmatrix} \log \frac{\hat{\psi}_i(a_1|R,\boldsymbol{\lambda})}{\hat{\psi}_i(a_1|L,\boldsymbol{\lambda})} & \log \frac{\hat{\psi}_i(a_M|R,\boldsymbol{\lambda})}{\hat{\psi}_i(a_M|L,\boldsymbol{\lambda})} \\ \log \frac{\hat{\psi}_j(a_1|R,\boldsymbol{\lambda})}{\hat{\psi}_j(a_1|L,\boldsymbol{\lambda})} & \log \frac{\hat{\psi}_j(a_M|R,\boldsymbol{\lambda})}{\hat{\psi}_j(a_M|L,\boldsymbol{\lambda})} \end{pmatrix} \tag{17}$$

as these ratios for social types θ_i and θ_j from actions a_1 and a_M. We use $\Psi(\boldsymbol{\lambda})$ to define separability (Definition 8) and use $\Psi[\theta_i,\theta_j;a_1,a_M](\boldsymbol{\lambda})$ to show that maximal accessibility implies separability.

Given a belief $\boldsymbol{\lambda}^* = (0^\kappa, \infty^{k-\kappa})$, we say $\boldsymbol{\lambda}^*$ is separable at zero for type θ_κ if there exists a sequence of actions that are on average more likely in state R than state L for θ_κ and the types with $\lambda_i^* = \infty$, and are on average more likely in state L than state R for the remaining $\kappa - 1$ types with $\lambda_i^* = 0$. In other words, in a neighborhood of $\boldsymbol{\lambda}^*$, there exists a sequence of actions that will decrease the beliefs of types $(\theta_1, ..., \theta_{\kappa-1})$ and increase the beliefs of types $(\theta_\kappa, ..., \theta_k)$. Separability at infinity is similar—in a neighborhood of $\boldsymbol{\lambda}^*$, there exists a sequence of actions that will on average decrease the beliefs of types $(\theta_1, ..., \theta_{\kappa+1})$ and increase the beliefs of types $(\theta_{\kappa+2}, ..., \theta_k)$. The following definition formalizes this notion for an arbitrary stationary belief.

Definition 8 (Separability ($k \geq 2$))**.** *(i) Belief $\boldsymbol{\lambda}^* \in \{0,\infty\}^k \setminus \infty^k$ is* separable at zero *for type θ_i with $\lambda_i^* = 0$ if there exist vectors $c \in [0,\infty)^{|\mathcal{A}|}$ and $G \in \mathbb{R}^k$ with $G_i > 0$, $G_j > 0$ for all j with $\lambda_j^* = \infty$ and $G_j < 0$ for all $j \neq i$ with $\lambda_j^* = 0$, such that $\Psi(\boldsymbol{\lambda}^*) \cdot c = G$; (ii) Belief $\boldsymbol{\lambda}^* \in \{0,\infty\}^k \setminus 0^k$ is* separable at infinity *for type θ_i with $\lambda_i^* = \infty$ if there exist vectors $c \in (0,\infty)^{|\mathcal{A}|}$ and $G \in \mathbb{R}^k$ with $G_i < 0$, $G_j > 0$ for all $j \neq i$ with $\lambda_j^* = \infty$ and $G_j < 0$ for all j with $\lambda_j^* = 0$, such that $\Psi(\boldsymbol{\lambda}^*) \cdot c = G$.*

The following result shows that separability can be used to establish the global stability of a disagreement outcome for the case of two social types. Theorem 7′ in Online Appendix D presents an analogous result for the case of more than two social types.

Theorem 7 (Global Stability of Disagreement ($k = 2$))**.** *Consider a learning environment that is identified at certainty and satisfies Assumptions 1 to 4. If $(0,\infty) \in \Lambda(\omega)$ and $(0,0)$*

is separable at zero for θ_2 or (∞,∞) is separable at infinity for θ_1, then $(0,\infty)$ is globally stable in state ω. Similarly, if $(\infty,0)\in\Lambda(\omega)$ and $(0,0)$ is separable at zero for θ_1 or (∞,∞) is separable at infinity for θ_2, then $(\infty,0)$ is globally stable in state ω.

We use Lemmas 5 and 6 to establish Theorem 7. We say a belief $\boldsymbol{\lambda}_2^*\in\{0,\infty\}^k$ is *adjacent* to a belief $\boldsymbol{\lambda}_1^*\in\{0,\infty\}^k$ if all but one component of the vectors are equal, i.e. there exists exactly one $i=1,...,k$ such that $(\boldsymbol{\lambda}_2^*)_i\neq(\boldsymbol{\lambda}_1^*)_i$. In other words, all but one social type have the same beliefs in $\boldsymbol{\lambda}_1^*$ and $\boldsymbol{\lambda}_2^*$. We say a belief $\boldsymbol{\lambda}_2^*\in\{0,\infty\}^k$ is adjacently accessible from adjacent belief $\boldsymbol{\lambda}_1^*\in\{0,\infty\}^k$ if, given the likelihood ratio process is in a neighborhood of $\boldsymbol{\lambda}_1^*$, with positive probability it enters a neighborhood of $\boldsymbol{\lambda}_2^*$ in finite time.

Definition 9 (Adjacently Accessible ($k\geq 2$))**.** *Belief $\boldsymbol{\lambda}_2^*\in\{0,\infty\}^k$ is* adjacently accessible *from adjacent belief $\boldsymbol{\lambda}_1^*\in\{0,\infty\}^k$ if for any $\varepsilon_2>0$, there exists an $\varepsilon_1>0$ such that for any $\boldsymbol{\lambda}\in B_{\varepsilon_1}(\boldsymbol{\lambda}_1^*)$, there exists a $\tau(\boldsymbol{\lambda})<\infty$ such that if $\boldsymbol{\lambda}_t=\boldsymbol{\lambda}$, then $Pr(\boldsymbol{\lambda}_{t+\tau(\boldsymbol{\lambda})}\in B_{\varepsilon_2}(\boldsymbol{\lambda}_2^*))>0$.*

The following lemma establishes that separability can be used to establish adjacent accessibility for any number of social types.

Lemma 5 (Adjacently Accessible ($k\geq 2$))**.** *Consider a learning environment that satisfies Assumptions 1 to 4. If $\boldsymbol{\lambda}_1^*\in\{0,\infty\}^k\setminus\infty^k$ with $(\boldsymbol{\lambda}_1^*)_i=0$ is separable at zero for θ_i, then adjacent belief $\boldsymbol{\lambda}_2^*$ with $(\boldsymbol{\lambda}_2^*)_i=\infty$ is adjacently accessible from $\boldsymbol{\lambda}_1^*$. If $\boldsymbol{\lambda}_1^*\in\{0,\infty\}^k\setminus 0^k$ with $(\boldsymbol{\lambda}_1^*)_i=\infty$ is separable at infinity for θ_i, then adjacent belief $\boldsymbol{\lambda}_2^*$ with $(\boldsymbol{\lambda}_2^*)_i=0$ is adjacently accessible from $\boldsymbol{\lambda}_1^*$.*

Proof. Given $\kappa\in\{1,...,k\}$, let $\boldsymbol{\lambda}_1^*=(0^\kappa,\infty^{k-\kappa})$, $\boldsymbol{\lambda}_2^*=(0^{\kappa-1},\infty^{k-\kappa+1})$ and suppose $\boldsymbol{\lambda}_1^*$ is separable at zero for θ_κ (where the subscript of $\boldsymbol{\lambda}^*$ is used to index different stationary beliefs). We will show that for any $\varepsilon_2>0$, there exists an $\varepsilon_1>0$ such that for any $\boldsymbol{\lambda}\in B_{\varepsilon_1}(\boldsymbol{\lambda}_1^*)$, there exists a $\tau(\boldsymbol{\lambda})<\infty$ such that if $\boldsymbol{\lambda}_1=\boldsymbol{\lambda}$, then $Pr(\boldsymbol{\lambda}_{1+\tau(\boldsymbol{\lambda})}\in B_{\varepsilon_2}(\boldsymbol{\lambda}_2^*))>0$ (where the subscript of $\boldsymbol{\lambda}$ is used to denote the likelihood ratio process at a given time i.e. $\boldsymbol{\lambda}_1$ is the value of the likelihood ratio at time $t=1$.) Since the log likelihood ratio process is linear, this also implies that if $\boldsymbol{\lambda}_t=\boldsymbol{\lambda}$, then $Pr(\boldsymbol{\lambda}_{t+\tau(\boldsymbol{\lambda})}\in B_{\varepsilon_2}(\boldsymbol{\lambda}_2^*))>0$.

Recall that $B_\varepsilon(\boldsymbol{\lambda}_1^*)\equiv[0,\varepsilon)^\kappa\times(1/\varepsilon,\infty]^{k-\kappa}$ denotes a neighborhood of $\boldsymbol{\lambda}_1^*$ for $\varepsilon>0$. Define $K(\varepsilon)\equiv-\log\varepsilon$, and let $[-\infty,-K(\varepsilon))^\kappa\times(K(\varepsilon),\infty]^{k-\kappa}$ denote the corresponding neighborhood of $\log\boldsymbol{\lambda}_1^*$. Define

$$g_{\varepsilon,i}(a)\equiv\begin{cases}\inf_{\boldsymbol{\lambda}\in B_\varepsilon(\boldsymbol{\lambda}_1^*)}\log\frac{\hat{\psi}_i(a|R,\boldsymbol{\lambda})}{\hat{\psi}_i(a|L,\boldsymbol{\lambda})} & i\geq\kappa\\ \sup_{\boldsymbol{\lambda}\in B_\varepsilon(\boldsymbol{\lambda}_1^*)}\log\frac{\hat{\psi}_i(a|R,\boldsymbol{\lambda})}{\hat{\psi}_i(a|L,\boldsymbol{\lambda})} & i<\kappa\end{cases}\tag{18}$$

as the smallest (largest) update to the log likelihood ratio when type $i\geq\kappa$ ($i<\kappa$) observes

a and has likelihood ratio in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}_1^*)$ and

$$\bar{g}_{\varepsilon,\kappa}(a) \equiv \sup_{\boldsymbol{\lambda}\in B_\varepsilon(\boldsymbol{\lambda}_1^*)} \log \frac{\hat{\psi}_\kappa(a|R,\boldsymbol{\lambda})}{\hat{\psi}_\kappa(a|L,\boldsymbol{\lambda})}$$

as the largest update to the log likelihood ratio when type κ observes a and has likelihood ratio in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}_1^*)$.

We construct a process that bounds $\langle\boldsymbol{\lambda}_t\rangle$ as long as it remains close to $\boldsymbol{\lambda}_1^*$, and use this process to show that we can separate the log likelihood ratios of types $1,...,\kappa-1$ and type κ by an arbitrary amount K while the beliefs of all types remain close to $\boldsymbol{\lambda}_1^*$. By $\boldsymbol{\lambda}_1^*$ separable at zero for θ_κ, there exist vectors $c \in [0,\infty)^k$ and $G \in \mathbb{R}^k$ that satisfy the separability condition. Moreover, since the rationals are dense in the reals, there exists vector $c \in [0,\infty)^k$ of rational numbers and vector $G \in \mathbb{R}^k$ that satisfies the separability condition. Therefore, there exists an $\varepsilon_3 > 0$ and integers $c_a \geq 0$ for each $a \in \mathcal{A}$ such that

$$G_i \equiv \sum_{a\in\mathcal{A}} c_a g_{\varepsilon_3,i}(a), \tag{19}$$

with $G_i > 0$ for all $i \geq \kappa$ and $G_i < 0$ for all $i < \kappa$. Let

$$\overline{G}_\kappa \equiv \sum_{a\in\mathcal{A}} c_a \bar{g}_{\varepsilon_3,\kappa}(a) \geq G_\kappa > 0. \tag{20}$$

Next we define processes $\xi_{i,t} \equiv \sum_{s=1}^{t-1} g_{\varepsilon_3,i}(a_s)$ and $\bar{\xi}_{\kappa,t} \equiv \sum_{s=1}^{t-1} \bar{g}_{\varepsilon_3,\kappa}(a_s)$. Given a sequence with c_a realizations of each a, at time $\tau_1 \equiv \sum_{\mathcal{A}} c_a + 1$, the process $\xi_{i,\tau_1} = G_i$ by Eq. (19) and $\bar{\xi}_{\kappa,\tau_1} = \overline{G}_\kappa$ by Eq. (20). For $i \geq \kappa$, $G_i > 0$, and therefore, $\xi_{i,\tau_1} > 0$, while for $i < \kappa$, $G_i < 0$, and therefore, $\xi_{i,\tau_1} < 0$. Moreover, there exists an $\underline{\mathrm{K}} > 0$ such that for all $i > \kappa$, $\xi_{i,t} \geq -\underline{\mathrm{K}}$ for all $t < \tau_1$, and there exists a $\bar{K} > 0$ such that for all $i < \kappa$, $\xi_{i,t} < \bar{K}$ for all $t < \tau_1$. Therefore, for any $K > 0$, there exists an N_K such that following N_K repetitions of the sequence of c_a realizations of each a, at time $\tau_K \equiv N_K \sum_{\mathcal{A}} c_a + 1$,

1. $\xi_{i,\tau_K} < -K$ for all $i < \kappa$,
2. $\xi_{i,\tau_K} > 0$ for all $i \geq \kappa$,
3. For all $t < \tau_K$, $\xi_{i,t} \leq \bar{K}$ for all $i < \kappa$ and $\xi_{i,t} \geq -\underline{\mathrm{K}}$ for all $i > \kappa$,
4. $\bar{\xi}_{\kappa,t} \leq N_K \overline{G}_\kappa$ for all $t \leq \tau_K$, with equality at $t = \tau_K$.

In summary, following N_K repetitions of the sequence, the processes $\langle\xi_{i,t}\rangle$ of types $i < \kappa$ and type κ are separated by at least K, and at all t during the repetitions, the process of type $i < \kappa$ is bounded above by $\bar{K}$ and the process of type $i > \kappa$ is bounded below by $-\underline{\mathrm{K}}$. As long as $\boldsymbol{\lambda}_s \in B_{\varepsilon_3}(\boldsymbol{\lambda}_1^*)$ for all $s \leq t$, the change in the log likelihood ratio of $i < \kappa$ is bounded above by $\xi_{i,t}$, $\log\lambda_{i,t} - \log\lambda_{i,1} \leq \xi_{i,t} \leq \bar{K}$, the change in the log likelihood ratio of $i = \kappa$ is bounded above by $\bar{\xi}_{\kappa,t}$, $\log\lambda_{\kappa,t} - \log\lambda_{\kappa,1} \leq \bar{\xi}_{\kappa,t}$, and the change in the log likelihood ratio of

$i > \kappa$ is bounded below by $\xi_{i,t}$, $\log \lambda_{i,t} - \log \lambda_{i,1} \geq \xi_{i,t} \geq -\underline{\mathrm{K}}$.

Fix $\varepsilon_2 \in (0, \varepsilon_3)$ and $K > \bar{K}$. Recall $K(\varepsilon) = -\log \varepsilon$. Choose an ε_1-neighborhood of $\boldsymbol{\lambda}_1^*$ such that $\log \lambda_{i,1} < -K(\varepsilon_2) - \max(\bar{K}, N_K \overline{G}_\kappa)$ for $i \leq \kappa$ and $\log \lambda_{i,1} > K(\varepsilon_2) + \underline{\mathrm{K}}$ for $i > \kappa$. Note $\varepsilon_1 < \varepsilon_2$. Suppose the initial likelihood ratio $\boldsymbol{\lambda}_1 \in B_{\varepsilon_1}(\boldsymbol{\lambda}_1^*)$. We establish local accessibility in three steps.

Step 1. Repeat N_K realizations of the sequence of c_a realizations of each a to separate the log likelihood ratio of types $i < \kappa$ and κ by K. It follows from items (3) and (4) that $\boldsymbol{\lambda}_t$ remains in $B_{\varepsilon_2}(\boldsymbol{\lambda}_1^*)$ for all $t \leq \tau_K$. Therefore, for each i and at all $t \leq \tau_K$, the process $\xi_{i,t}$ bounds the change in the log likelihood ratio, $\log \lambda_{i,t} - \log \lambda_{i,1} \leq \xi_{i,t}$. After N_K realizations of the sequence, $\log \lambda_{i,\tau_K} < -K(\varepsilon_2) - K$ for $i < \kappa$, and $\log \lambda_{i,\tau_K} > K(\varepsilon_2) + \underline{\mathrm{K}}$ for $i > \kappa$.

Step 2. Next, push type κ's log likelihood ratio to $-K(\varepsilon_3)$ as follows. Continue repeating the sequence of c_a realizations of each a until $\log \lambda_{\kappa,t} > -K(\varepsilon_3)$. By construction, the likelihood ratios of all types $i \neq \kappa$ remain in $B_{\varepsilon_2}(\boldsymbol{\lambda}_1^*)$ after every a in this sequence, since at any point in the sequence, $\log \lambda_{i,t} < -K(\varepsilon_2) - K + \bar{K}$ for all $i < \kappa$, and $\log \lambda_{i,t} > K(\varepsilon_2)$ for all $i > \kappa$.

Step 3. Finally, push type κ's log likelihood ratio from $-K(\varepsilon_3)$ to $K(\varepsilon_2)$, while keeping the log likelihood ratio of type $i < \kappa$ less than $-K(\varepsilon_2)$. Given ε_2, there exists an $N_2 < \infty$ such that if $\log \lambda_{\kappa,t} \in [-K(\varepsilon_3), K(\varepsilon_2)]$, then following N_2 realizations of a_M, $\log \lambda_{\kappa,t+N_2} > K(\varepsilon_2)$. Let K_2 be the most any type $i < \kappa$'s log likelihood ratio increases after N_2 realizations of a_M across all beliefs $\boldsymbol{\lambda} \in B_{\varepsilon_2}(\boldsymbol{\lambda}_1^*)$. Recall that when type κ hit the boundary of $-K(\varepsilon_3)$, $\log \lambda_{i,t} < -K(\varepsilon_2) - K + \bar{K}$ for all $i < \kappa$ and $\log \lambda_{i,t} > K(\varepsilon_2)$ for $i > \kappa$. Therefore, after N_2 realizations of a_M, $\log \lambda_{i,t} < -K(\varepsilon_2) - K + \bar{K} + K_2$ for all $i < \kappa$ and $\log \lambda_{i,t} > K(\varepsilon_2)$ for $i > \kappa$. In order to keep $i < \kappa$ in an ε_2-neighborhood of zero after N_2 realizations of a_M, we need to separate beliefs by at least $K = \bar{K} + K_2$. This determines the K we need to use in step 1.

Following these three steps with $K = \bar{K} + K_2$, the likelihood ratio is in neighborhood $B_{\varepsilon_2}(\boldsymbol{\lambda}_2^*)$. Each step required a finite number of actions that occur with positive probability. Therefore, given ε_1 and ε_2 defined above, for any $\boldsymbol{\lambda} \in B_{\varepsilon_1}(\boldsymbol{\lambda}_1^*)$, there exists a $\tau(\boldsymbol{\lambda}) < \infty$ such that if $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}$, then $Pr(\boldsymbol{\lambda}_{1+\tau(\boldsymbol{\lambda})} \in B_{\varepsilon_2}(\boldsymbol{\lambda}_2^*)) > 0$. The case of other disagreement outcomes $\boldsymbol{\lambda}_1^* \in \{0, \infty\}^k \setminus \infty^k$ that are separable at zero is analogous, as is the case of $\boldsymbol{\lambda}_1^* \in \{0, \infty\}^k \setminus 0^k$ separable at infinity. □

We say a belief $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ is accessible if, for any initial belief, with positive probability the likelihood ratio process enters a neighborhood of $\boldsymbol{\lambda}^*$ in finite time.

Definition 10 (Accessible ($k \geq 2$))**.** *A belief $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ is* accessible *if for any initial belief $\boldsymbol{\lambda}_1 \in (0, \infty)^k$ and any $\varepsilon > 0$, there exists a $\tau < \infty$ such that $Pr(\boldsymbol{\lambda}_\tau \in B_\varepsilon(\boldsymbol{\lambda}^*)) > 0$.*

We next show that adjacent accessibility can be used to establish accessibility.

Lemma 6 (Accessible Disagreement ($k \geq 2$)). *Consider a learning environment that satisfies Assumptions 1 to 4. Disagreement outcome $\boldsymbol{\lambda}^* \in \{0, \infty\}^k \setminus \{0^k, \infty^k\}$ is accessible if there exists a finite sequence of stationary beliefs $\boldsymbol{\lambda}_1^*, \boldsymbol{\lambda}_2^* \ldots \boldsymbol{\lambda}_L^*$, with $\boldsymbol{\lambda}_1^* \in \{0^k, \infty^k\}$, $\boldsymbol{\lambda}_{l+1}^*$ adjacently accessible from adjacent belief $\boldsymbol{\lambda}_l^*$ for $l = 1, ..., L-1$ and $\boldsymbol{\lambda}_L^* = \boldsymbol{\lambda}^*$.*

Proof. Given disagreement outcome $\boldsymbol{\lambda}^*$, suppose there exists a finite sequence of stationary beliefs $\boldsymbol{\lambda}_1^*, \boldsymbol{\lambda}_2^* \ldots \boldsymbol{\lambda}_L^*$, with $\boldsymbol{\lambda}_1^* \in \{0^k, \infty^k\}$, $\boldsymbol{\lambda}_{l+1}^*$ adjacently accessible from adjacent belief $\boldsymbol{\lambda}_l^*$ for $l = 1, ..., L-1$ and $\boldsymbol{\lambda}_L^* = \boldsymbol{\lambda}^*$. By definition of adjacently accessible, for any $\varepsilon_L > 0$, there exists an $\varepsilon_{L-1} > 0$ and $\tau_L < \infty$ such that if $\boldsymbol{\lambda}_t \in B_{\varepsilon_{L-1}}(\boldsymbol{\lambda}_{L-1}^*)$, then $Pr(\boldsymbol{\lambda}_{t+\tau_L} \in B_{\varepsilon_L}(\boldsymbol{\lambda}_L^*)) > 0$. Iterating back to $\boldsymbol{\lambda}_1^*$, for any $\varepsilon_L > 0$, there exists an $\varepsilon_1 > 0$ and $\tau_2 < \infty$ such that if $\boldsymbol{\lambda}_t \in B_{\varepsilon_1}(\boldsymbol{\lambda}_1^*)$, then $Pr(\boldsymbol{\lambda}_{t+\sum_{l=2}^{L}\tau_l} \in B_{\varepsilon_L}(\boldsymbol{\lambda}_L^*)) > 0$. Consider agreement outcome $\boldsymbol{\lambda}_1^* \in \{0^k, \infty^k\}$. By Theorem 2, for any initial belief $\boldsymbol{\lambda}_1 \in (0, \infty)^k$ and any $\varepsilon_1 > 0$, there exists a finite sequence of τ_1 actions that occur with positive probability such that following this sequence, $\boldsymbol{\lambda}_{\tau_1+1} \in B_{\varepsilon_1}(\boldsymbol{\lambda}_1^*)$. Therefore, starting from any initial belief, $Pr(\boldsymbol{\lambda}_{\tau_1+1} \in B_{\varepsilon_1}(\boldsymbol{\lambda}_1^*)) > 0$. Therefore, for any $\varepsilon_L > 0$ and initial belief $\boldsymbol{\lambda}_1 \in (0, \infty)^k$, $Pr(\boldsymbol{\lambda}_\tau \in B_{\varepsilon_L}(\boldsymbol{\lambda}_L^*)) > 0$, where $\tau \equiv \sum_{l=1}^{L} \tau_l + 1 < \infty$ since each $\tau_l < \infty$. By definition, this means that $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}_L^*$ is accessible. □

Accessibility and local stability together imply global stability. Therefore, by Lemmas 5 and 6, the separability of an agreement outcome combined with the local stability of an adjacently accessible disagreement outcome establishes that the disagreement outcome is globally stable.

Proof of Theorem 7. Suppose $(0, \infty) \in \Lambda(\omega)$ and either $(0, 0)$ is separable at zero for θ_2 or (∞, ∞) is separable at infinity for θ_1. By Lemma 5, $(0, \infty)$ is adjacently accessible from $(0, 0)$. It follows from Lemma 6 that $(0, \infty)$ is accessible. Fix initial belief $\boldsymbol{\lambda}_1 \in (0, \infty)^2$ and choose $\varepsilon < e^{-E}$, where E is defined in Eq. (14). By accessibility, there exists a finite sequence ξ of N actions that occurs with positive probability, such that following ξ, $\boldsymbol{\lambda}_{N+1} \in B_\varepsilon((0, \infty))$. From $(0, \infty) \in \Lambda(\omega)$, $Pr(\boldsymbol{\lambda}_t \to (0, \infty) | h = \xi) > 0$. Therefore, from any initial belief, $Pr(\boldsymbol{\lambda}_t \to (0, \infty)) > 0$, which implies that $(0, \infty)$ is globally stable. The proof of $(\infty, 0)$ is analogous. □

Finally, to prove Theorem 3, we show that maximal accessibility implies the conditions for separability outlined in Theorem 7.

Proof of Theorem 3. Suppose $\theta_2 \succ_{(0,0)} \theta_1$. We show that this implies $(0, 0)$ is separable at zero for θ_2. Since $\theta_2 \succ_{(0,0)} \theta_1$, the submatrix $\Psi[\theta_2, \theta_1; a_1, a_M]((0, 0))$ defined in Eq. (17) has a positive determinant. Therefore, there exists a $c \in \mathbb{R}_+^2$ that solves

$$\Psi[\theta_2, \theta_1; a_1, a_M]((0, 0)) \cdot c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

By continuity, there exists a perturbation of c to $\tilde{c} \in \mathbb{R}^2_+$ such that

$$\Psi[\theta_2, \theta_1; a_1, a_M]((0,0)) \cdot \tilde{c} = \begin{pmatrix} G_2 \\ G_1 \end{pmatrix},$$

where $G_1 < 0$ and $G_2 > 0$. Therefore, by Definition 8, $(0,0)$ is separable at zero for θ_2, since we can set values of c_j to zero for the remaining actions in matrix $\Psi((0,0))$. The case where $\theta_2 \succ_{(\infty,\infty)} \theta_1$ is analogous, as is the proof for $(\infty, 0)$. □

A.4 Mixed Learning

Proof of Lemma 4. Consider the mixed learning outcome $(0, \theta_1)$ in which θ_1's belief converges to zero and θ_2's belief doesn't converge. Suppose $(0, \theta_1) \notin \Lambda_M(\omega)$, i.e. $(0,0) \in \Lambda_2(\omega)$ or $(0,\infty) \in \Lambda_2(\omega)$. Without loss of generality, consider the case where $(0,0) \in \Lambda_2(\omega)$. Suppose the initial belief for type θ_1 is near zero, $\lambda_{1,1} \in B_\varepsilon(0)$ for any $\varepsilon < e^{-E}$, where E is defined in Eq. (14). We want to show that almost surely, either (i) there exists a $\tau < \infty$ such that $\lambda_{1,\tau} \notin B_\varepsilon(0)$; or (ii) $\langle \boldsymbol{\lambda}_t \rangle$ converges for both types. This will establish that almost surely, $(0, \theta_1)$ does not occur.

We first characterize how the behavior of $\langle \boldsymbol{\lambda}_t \rangle$ near $(0,0)$ and $(0,\infty)$ depends on $\Lambda_1(\omega)$ and $\Lambda_2(\omega)$. Suppose $(0,0) \in \Lambda_1(\omega)$ (recall by assumption, $(0,0) \in \Lambda_2(\omega)$). By the construction in Theorem 1, for $\varepsilon < e^{-E}$, if $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_\varepsilon(0,0)$, with positive probability, $\langle \boldsymbol{\lambda}_t \rangle$ converges to $(0,0)$. If $(0,0) \notin \Lambda_1(\omega)$, then by the construction in Theorem 1, for $\varepsilon < e^{-E}$, if $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_\varepsilon((0,0))$, then from any belief in $B_\varepsilon((0,0))$, (i) with positive probability uniformly bounded away from zero in the starting belief, $\langle \lambda_{1,t} \rangle$ exits $B_\varepsilon(0)$, and (ii) almost surely, $\langle \boldsymbol{\lambda}_t \rangle$ exits $B_\varepsilon((0,0))$. If $(0,\infty) \in \Lambda_2(\omega)$, the behavior of $\langle \boldsymbol{\lambda}_t \rangle$ in a neighborhood of $(0,\infty)$ is similar. If $(0,\infty) \notin \Lambda_2(\omega)$, then by the construction in Theorem 1, for $\varepsilon < e^{-E}$, if $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_\varepsilon((0,\infty))$, then almost surely, $\langle \boldsymbol{\lambda}_t \rangle$ exits $B_\varepsilon((0,\infty))$.

Let $\tau_1 \equiv \min\{t | \lambda_{1,t} \notin B_\varepsilon(0)\}$ be the first time that θ_1's belief leaves a neighborhood of zero. Then it must be that almost surely, $\tau_1 < \infty$ or $\langle \boldsymbol{\lambda}_t \rangle$ visits a neighborhood of $(0,0)$ or $(0,\infty)$ infinitely often,

$$Pr(\tau_1 < \infty \text{ or } \boldsymbol{\lambda}_t \in B_\varepsilon((0,0)) \cup B_\varepsilon((0,\infty)) \text{ i.o.}) = 1. \tag{21}$$

If $(0,0) \notin \Lambda_1(\omega)$, so $(0,0)$ is not locally stable, then λ_2 almost surely leaves $B_\varepsilon((0,0))$, and $Pr(\tau_1 < \infty \text{ or } \boldsymbol{\lambda}_t \in B_\varepsilon((0,\infty)) \text{ i.o.}) = 1$. Similarly, if $(0,\infty) \notin \Lambda(\omega)$, then λ_2 almost surely leaves $B_\varepsilon((0,\infty))$, and $Pr(\tau_1 < \infty \text{ or } \boldsymbol{\lambda}_t \in B_\varepsilon((0,0)) \text{ i.o.}) = 1$.

Case (i): Suppose $(0,0) \in \Lambda_1(\omega)$ or $(0,\infty) \in \Lambda_1(\omega) \cap \Lambda_2(\omega)$. If $\langle \boldsymbol{\lambda}_t \rangle$ enters a neighborhood of a locally stable belief infinitely often, then $\langle \boldsymbol{\lambda}_t \rangle$ almost surely converges for both types. Therefore, almost surely, $\tau_1 < \infty$ or $\langle \boldsymbol{\lambda}_t \rangle$ converges.

Case (ii): Suppose $(0,0) \notin \Lambda_1(\omega)$ and $(0,\infty) \in \Lambda_2(\omega) \setminus \Lambda_1(\omega)$. Each time $\langle \boldsymbol{\lambda}_t \rangle$ enters

$B_\varepsilon((0,0)) \cup B_\varepsilon((0,\infty))$, with positive probability uniformly bounded away from zero in the starting belief, $\langle \lambda_{1,t} \rangle$ exits $B_\varepsilon(0)$. Therefore, if $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_\varepsilon((0,0)) \cup B_\varepsilon((0,\infty))$ infinitely often, $\langle \lambda_{1,t} \rangle$ almost surely exits $B_\varepsilon(0)$. Therefore, almost surely $\tau_1 < \infty$.

Case (iii): Suppose $(0,0) \notin \Lambda_1(\omega)$, $(0,\infty) \notin \Lambda_2(\omega)$. Then $Pr(\tau_1 < \infty \text{ or } \boldsymbol{\lambda}_t \in B_\varepsilon((0,0)) \text{ i.o.}) = 1$. Each time $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_\varepsilon((0,0))$, with positive probability uniformly bounded away from zero in the starting belief, $\langle \lambda_{1,t} \rangle$ exits $B_\varepsilon(0)$. Therefore, if $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_\varepsilon((0,0))$ infinitely often, $\langle \lambda_{1,t} \rangle$ almost surely exits $B_\varepsilon(0)$. Therefore, almost surely $\tau_1 < \infty$. The proofs for the other mixed outcomes are analogous. □

A.5 Learning Characterization

Lemma 7 (Belief Convergence ($k \leq 2$))**.** *Consider a learning environment that is identified at certainty and satisfies Assumptions 1 to 4. If (i) $k = 1$ and $\Lambda(\omega) \neq \emptyset$ or (ii) $k = 2$, $\Lambda(\omega)$ contains an agreement outcome or maximally accessible disagreement outcome and $\Lambda_M(\omega) = \emptyset$, then for any initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^k$, there exists a random variable $\boldsymbol{\lambda}_\infty$ with* $\text{supp}(\boldsymbol{\lambda}_\infty) = \Lambda(\omega)$ *such that $\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}_\infty$ almost surely.*

Proof of Lemma 7. Suppose $\Lambda(\omega)$ contains an agreement outcome or maximally accessible disagreement outcome and $\Lambda_M(\omega)$ is empty. Recall that $\mathcal{B}$ is the set of stable neighborhoods and $\mathcal{B}_U$ is the set of unstable neighborhoods defined in Eq. (15). Let $\tau_1 \equiv \min\{t | \boldsymbol{\lambda}_t \in \mathcal{B}\}$ be the first time that the likelihood ratio enters the set of locally stable neighborhoods. By Lemma 2, there exists a finite sequence of actions such that starting from any initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^k$, $\langle \boldsymbol{\lambda}_t \rangle$ enters $\mathcal{B}$. This sequence occurs with positive probability. Therefore, the probability of entering $\mathcal{B}$ in finite time is bounded away from zero, $Pr(\tau_1 < \infty) > 0$. If $\langle \boldsymbol{\lambda}_t \rangle$ enters $\mathcal{B}_U$, then by Theorem 1, $\langle \boldsymbol{\lambda}_t \rangle$ almost surely leaves $\mathcal{B}_U$. Therefore, $\langle \boldsymbol{\lambda}_t \rangle$ does not converge to a stationary belief that is not locally stable. If $\langle \boldsymbol{\lambda}_t \rangle$ enters the neighborhood of a mixed outcome, by Lemma 4, $\langle \boldsymbol{\lambda}_t \rangle$ almost surely leaves this neighborhood or converges to a locally stable belief. Therefore, mixed learning outcomes almost surely do not arise. By Lemma 3, $\langle \boldsymbol{\lambda}_t \rangle$ does not converge to a non-stationary belief. Therefore, almost surely, either $\langle \boldsymbol{\lambda}_t \rangle$ does not converge for either type or $\langle \boldsymbol{\lambda}_t \rangle$ converges to a learning outcome in $\Lambda(\omega)$. Since $\langle \boldsymbol{\lambda}_t \rangle$ almost surely leaves the neighborhood of any mixed or unstable outcome, it must be that $\langle \boldsymbol{\lambda}_t \rangle$ enters $\mathcal{B}$ infinitely often, $Pr(\boldsymbol{\lambda}_t \in \mathcal{B} \text{ i.o.}) = 1$. But if $\langle \boldsymbol{\lambda}_t \rangle$ enters a neighborhood of a locally stable belief infinitely often, then almost surely $\langle \boldsymbol{\lambda}_t \rangle$ converges. □

Proof of Theorem 4. Parts (1) and (2) follow from the local and global stability of agreement outcomes (Theorems 1 and 2). Part (3) follows from the local and global stability of disagreement outcomes (Theorems 1 and 3). For part (4), Lemma 3 rules out convergence to non-stationary beliefs, Theorem 1 rules out convergence to stationary outcomes that are not locally stable, and Lemma 4 rules out convergence to a mixed learning outcome when $\Lambda_M(\omega) = \emptyset$. Therefore, if $\Lambda(\omega) = \emptyset$, there are no locally stable learning outcomes and almost surely the likelihood ratio does not converge for at least one social type. If $\Lambda_M(\omega) = \emptyset$, then almost surely the likelihood ratio does not converge for any social type. The final statement in part (4) follows from Lemma 7, which establishes when the likelihood ratio converges. $\square$

A.6 Robustness

Proof of Theorem 5. Fix a regular learning environment (Θ^*, π^*). Let $\psi^*(\cdot|\omega, \boldsymbol{\lambda})$ and $\hat{\psi}_i^*(\cdot|\omega, \boldsymbol{\lambda})$ denote the true and perceived action distributions in this environment, and analogously for $\boldsymbol{\gamma}^*(\omega, \boldsymbol{\lambda})$, $\Lambda^*(\omega)$, and $\Lambda_M^*(\omega)$. Throughout the proof, restrict attention to learning environments (Θ, π) that have the same number of social types as (Θ^*, π^*) and satisfy Assumptions 1 to 4. We first consider local stability of nearby learning environments. The mapping $(\psi(a|\omega, \boldsymbol{\lambda}), \hat{\psi}_i(a|\omega, \boldsymbol{\lambda})) \mapsto \gamma_i(\omega, \boldsymbol{\lambda})$ is continuous. By definition of identified at certainty, the sign of $\gamma_i^*(\omega, \boldsymbol{\lambda})$ is strictly positive or negative at $\boldsymbol{\lambda} \in \{0, \infty\}^k$. Therefore, there exists a $\delta_1(\omega) > 0$ such that in any learning environment (Θ, π) that is sufficiently close to (Θ^*, π^*) in that $||\psi^*(\cdot|\omega, \boldsymbol{\lambda}) - \psi(\cdot|\omega, \boldsymbol{\lambda})|| < \delta_1(\omega)$, $||\hat{\psi}_i^*(\cdot|L, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|L, \boldsymbol{\lambda})|| < \delta_1(\omega)$ and $||\hat{\psi}_i^*(\cdot|R, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|R, \boldsymbol{\lambda})|| < \delta_1(\omega)$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $i = 1, ..., k$, continuity implies that $\gamma_i(\omega, \boldsymbol{\lambda})$ has the same sign as $\gamma_i^*(\omega, \boldsymbol{\lambda})$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $i = 1, ..., k$. This implies $\Lambda_i(\omega) = \Lambda_i^*(\omega)$ for $i = 1, ..., k$, and therefore, $\Lambda(\omega) = \Lambda^*(\omega)$, so (Θ, π) has the same set of locally stable outcomes as (Θ^*, π^*) in state ω. Given $\Lambda_M(\omega)$ is constructed from $\Lambda_i(\omega)$, in any such learning environment, $\Lambda_M(\omega) = \Lambda_M^*(\omega)$ follows from $\Lambda_i(\omega) = \Lambda_i^*(\omega)$ for $i = 1, ..., k$. By definition of regular, given $\Lambda_M^*(\omega) = \emptyset$, this implies $\Lambda_M(\omega) = \emptyset$.

Finally, consider the global stability of locally stable disagreement outcomes in a learning environment with $\Lambda(\omega) = \Lambda^*(\omega)$. Suppose $k = 2$ and $(0, \infty) \in \Lambda(\omega)$. Then $(0, \infty) \in \Lambda^*(\omega)$ and by definition of regular, $(0, \infty)$ is maximally accessible in (Θ^*, π^*), i.e. either $\theta_2^* \succ_{(0,0)} \theta_1^*$ or $\theta_2^* \succ_{(\infty,\infty)} \theta_1^*$. Suppose $\theta_2^* \succ_{(0,0)} \theta_1^*$. By the proof of Theorem 3, this implies $(0, 0)$ is separable at zero for θ_2^*, so there exists a vector $G = (G_1, G_2)'$ with $G_1 < 0$ and $G_2 > 0$ and a vector $c \in \mathbb{R}_+^{|\mathcal{A}|}$ such that $\Psi^*((0, 0)) \cdot c = G$, where Ψ^* is the matrix defined in Eq. (16) for (Θ^*, π^*). Since $\Psi((0, 0)) \cdot c$ is continuous in $\hat{\psi}_i$, there exists a $\delta_{(0,\infty)} > 0$ such that in any learning environment (Θ, π) that is sufficiently close to (Θ^*, π^*) in that $||\hat{\psi}_i(\cdot|L, \boldsymbol{\lambda}) - \psi^*(\cdot|L, \boldsymbol{\lambda})|| < \delta_{(0,\infty)}$ and $||\hat{\psi}_i(\cdot|R, \boldsymbol{\lambda}) - \psi^*(\cdot|R, \boldsymbol{\lambda})|| < \delta_{(0,\infty)}$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^2$ and $i = 1, 2$, the expressions $(\Psi((0, 0)) \cdot c)_1 < 0$ and $(\Psi((0, 0)) \cdot c)_2 > 0$, where Ψ is the matrix defined in Eq. (16) for (Θ, π). Therefore, in any such learning environment, $(0, 0)$ is separable at zero for θ_2.[43] By Theorem 7, this implies $(0, \infty)$ is globally stable in state ω for (Θ, π). The case of $\theta_2^* \succ_{(\infty,\infty)} \theta_1^*$ is analogous, as is the proof for disagreement outcome $(\infty, 0)$ using some $\delta_{(\infty,0)} > 0$. When $k > 2$, a similar argument shows that a disagreement outcome $\boldsymbol{\lambda}^*$ that is locally stable and maximally accessible in (Θ^*, π^*) is locally stable and separable in sufficiently close learning environments, where close is defined relative to some

[43] Note that $\boldsymbol{\lambda}$ maximally accessible in (Θ^*, π^*) does not imply that it is maximally accessible in (Θ, π), as the strict maximal R-order can have one weak inequality.

$\delta_{\boldsymbol{\lambda}^*} > 0$, and therefore, globally stable. Let $\delta_2(\omega) \equiv \min_{\boldsymbol{\lambda}^* \in \{\Lambda^*(\omega) \setminus \{0^k, \infty^k\}} \delta_{\boldsymbol{\lambda}^*}$ denote the minimum $\delta_{\boldsymbol{\lambda}^*}$ across all locally stable disagreement outcomes in state ω for (Θ^*, π^*). Then any learning environment (Θ, π) that has $\Lambda(\omega) = \Lambda^*(\omega)$, $||\hat{\psi}_i^*(\cdot|L, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|L, \boldsymbol{\lambda})|| < \delta_2(\omega)$ and $||\hat{\psi}_i^*(\cdot|R, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|R, \boldsymbol{\lambda})|| < \delta_2(\omega)$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $i = 1, ..., k$ has the same set of globally stable outcomes as (Θ^*, π^*) in state ω.

Let $\delta \equiv \min\{\delta_1(\omega), \delta_2(\omega)\}$. Then any learning environment (Θ, π) with $||\psi^*(\cdot|\omega, \boldsymbol{\lambda}) - \psi(\cdot|\omega, \boldsymbol{\lambda})|| < \delta$, $||\hat{\psi}_i^*(\cdot|L, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|L, \boldsymbol{\lambda})|| < \delta$ and $||\hat{\psi}_i^*(\cdot|R, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|R, \boldsymbol{\lambda})|| < \delta$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $i = 1, ..., k$ has the same set of long-run learning outcomes as (Θ^*, π^*) in state ω. □

Proof of Theorem 6. Fix a correctly specified environment (Θ^*, π^*) that satisfies Assumptions 2 and 3. Let $\psi^*(a|\omega, \boldsymbol{\lambda})$ denote the distribution over actions in this environment. In a correctly specified environment, $\hat{\psi}_i^*(a|\omega, \boldsymbol{\lambda}) = \psi^*(a|\omega, \boldsymbol{\lambda})$ for $i = 1, ..., k$. By Corollary 2, learning is complete in (Θ^*, π^*). Further, correctly specified environments are regular. Throughout the proof, restrict attention to learning environments (Θ, π) that are structurally equivalent to (Θ^*, π^*) and satisfy Assumptions 1 to 4.

We first construct $\hat{\psi}_i(a|\omega, \boldsymbol{\lambda})$ and $\psi(a|\omega, \boldsymbol{\lambda})$ for such a (Θ, π). Let $\lambda_0 \equiv p_0/(1-p_0)$. From the decision rules constructed in Lemma 1, an autarkic or noise type $\theta_j \in \Theta_A \cup \Theta_N$ chooses action a_m if $\overline{s}_{j,m-1}(\lambda_0) \neq \overline{s}_{j,m}(\lambda_0)$ and it observes a signal $s \in (\overline{s}_{j,m-1}(\lambda_0), \overline{s}_{j,m}(\lambda_0)]$, with a closed interval if $\overline{s}_{i,m-1}(\lambda_0) = 0$.[44] Note that $\theta_j^* \in \Theta_A^* \cup \Theta_N^*$ has the same signal thresholds as θ_j, i.e. $\overline{s}_{j,m}^* = \overline{s}_{j,m}$ for $m = 1, ..., M$, since $\Theta_A \cup \Theta_N = \Theta_A^* \cup \Theta_N^*$ by definition of structurally equivalent. At any belief $\boldsymbol{\lambda} \in \{0, \infty\}^k$, social type $\theta_j \in \Theta_S$ has a unique optimal action that it plays for all signal realizations, independent of its model of inference $(\hat{F}_j^L, \hat{F}_j^R, \hat{\pi}_j)$. Let $\alpha_j(\boldsymbol{\lambda})$ denote this optimal action. Note that $\theta_j^* \in \Theta_S^*$ has the same optimal action, $\alpha_j^*(\boldsymbol{\lambda})$=$\alpha_j(\boldsymbol{\lambda})$, since it has the same preferences as θ_j by definition of structurally equivalent. Social type $\theta_i \in \Theta_S$ believes autarkic or noise type $\theta_j \in \Theta_A \cup \Theta_N$ chooses action a_m with probability $\hat{F}_i^\omega(\overline{s}_{j,m}(\lambda_0)) - \hat{F}_i^\omega(\overline{s}_{j,m-1}(\lambda_0))$. It believes $\theta_j \in \Theta_S$ chooses $\alpha_j(\boldsymbol{\lambda})$ with probability one independent of its model of inference $(\hat{F}_i^L, \hat{F}_i^R, \hat{\pi}_i)$. Therefore, for any $\boldsymbol{\lambda} \in \{0, \infty\}^k$, type θ_i believes a_m is chosen with probability

$$\hat{\psi}_i(a_m|\omega, \boldsymbol{\lambda}) = \sum_{\theta_j \in \Theta_A \cup \Theta_N} \hat{\pi}_i(\theta_j)(\hat{F}_i^\omega(\overline{s}_{j,m}(\lambda_0)) - \hat{F}_i^\omega(\overline{s}_{j,m-1}(\lambda_0))) + \sum_{\theta_j \in \Theta_S} \hat{\pi}_i(\theta_j)\mathbb{1}_{\alpha_j(\boldsymbol{\lambda})=a_m}. \quad (22)$$

This is continuous in $\hat{\pi}_i$ and $\hat{F}_i^\omega$ under the total variation norm, and it is independent of

[44] This proof maintains our convention for breaking indifference outlined in Lemma 1, i.e. the agent chooses the optimal action with the lowest index. Our robustness result holds across all equilibria, i.e. it also holds for equilibria that resolve indifference in different directions. A similar method to that outlined here establishes this for each of the finite possible ways to resolve indifference. We omit this analysis, as it requires cumbersome notation without added conceptual insight.

$(\hat{F}_j^L, \hat{F}_j^R, \hat{\pi}_j)$ for $\theta_j \in \Theta_S \setminus \{\theta_i\}$. Similarly, for any $\boldsymbol{\lambda} \in \{0, \infty\}^k$, the true probability of a_m is

$$\begin{aligned}\psi(a_m|\omega, \boldsymbol{\lambda}) &= \sum_{\theta_j \in \Theta_A \cup \Theta_N} \pi(\theta_j)(F^\omega(\overline{s}_{j,m}(\lambda_0)) - F^\omega(\overline{s}_{j,m-1}(\lambda_0))) + \sum_{\theta_j \in \Theta_S} \pi(\theta_j) \mathbb{1}_{\alpha_j(\boldsymbol{\lambda}) = a_m} \\ &= \sum_{\theta_j^* \in \Theta_A^* \cup \Theta_N^*} \pi^*(\theta_j^*)(F^\omega(\overline{s}_{j,m}^*(\lambda_0)) - F^\omega(\overline{s}_{j,m-1}^*(\lambda_0))) + \sum_{\theta_j^* \in \Theta_S^*} \pi^*(\theta_j^*) \mathbb{1}_{\alpha_j^*(\boldsymbol{\lambda}) = a_m} \\ &= \psi^*(a_m|\omega, \boldsymbol{\lambda}) \end{aligned} \tag{23}$$

where the second equality follows from $\Theta_A \cup \Theta_N = \Theta_A^* \cup \Theta_N^*$ and $\pi(\theta_j) = \pi^*(\theta_j^*)$ by definition of structurally equivalent, $\alpha_j(\boldsymbol{\lambda}) = \alpha_j^*(\boldsymbol{\lambda})$ for $j = 1, ..., k$ as shown above, and $\overline{s}_{j,m} = \overline{s}_{j,m}^*$ for $m = 1, ..., M$ and $j = k+1, ..., n$ as shown above.

From Eq. (22) continuous in $(\hat{\pi}_i, \hat{F}_i^\omega)$ and Eq. (23), it follows that for any $\varepsilon > 0$, there exists a $\delta(\varepsilon) > 0$ such that any environment (Θ, π) with $||\hat{\pi}_i - \pi|| < \delta(\varepsilon)$, $||\hat{F}_i^L - F^L|| < \delta(\varepsilon)$ and $||\hat{F}_i^R - F^R|| < \delta(\varepsilon)$ for all $\theta_i \in \Theta_S$ satisfies $||\psi^*(\cdot|L, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|L, \boldsymbol{\lambda})|| < \varepsilon$ and $||\psi^*(\cdot|R, \boldsymbol{\lambda}) - \hat{\psi}_i(\cdot|R, \boldsymbol{\lambda})|| < \varepsilon$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $i = 1, ..., k$. Further, from Eq. (23), $||\psi^*(\cdot|\omega, \boldsymbol{\lambda}) - \psi(\cdot|\omega, \boldsymbol{\lambda})|| = 0$ for all $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $\omega \in \{L, R\}$ in any such model. Given this, by Theorem 5, choosing ε small enough establishes that any such learning environment has the same set of long-run learning outcomes as (Θ^*, π^*) in both states. Since learning is complete in both states in (Θ^*, π^*), this establishes that learning is complete in both states in any learning environment with $||\hat{\pi}_i - \pi||$, $||\hat{F}_i^L - F^L||$ and $||\hat{F}_i^R - F^R||$ sufficiently small for all social types. □