PENN INSTITUTE for ECONOMIC RESEARCH
UNIVERSITY of PENNSYLVANIA

The Ronald O. Perelman Center for
Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 18-017

Social Learning with Model Misspeciification: A Framework and a Robustness Result

J. AISLINN BOHREN
University of Pennsylvania
& Carnegie Mellon University

DANIEL N. HAUSER
Aalto University

July 1, 2018

https://ssrn.com/abstract=3236842

Social Learning with Model Misspecification: A Framework and a Robustness Result*

J. Aislinn Bohren[†] Daniel N. Hauser[‡]

July 2018

We explore how model misspecification affects long-run learning in a sequential social learning setting. Individuals learn from diverse sources, including private signals, public signals and the actions and outcomes of others. An agent's type specifies her model of the world. Misspecified types have incorrect beliefs about the signal distribution, how other agents draw inference and/or others' preferences. Our main result is a simple criterion to characterize long-run learning outcomes that is straightforward to derive from the primitives of the misspecification. Depending on the nature of the misspecification, we show that learning may be correct, incorrect or beliefs may not converge. Multiple degenerate limit beliefs may arise and agents may asymptotically disagree, despite observing the same sequence of information. We also establish that the correctly specified model is robust – agents with approximately correct models almost surely learn the true state. We close with a demonstration of how our framework can capture three broad categories of model misspecification: strategic misspecification, such as level-k and cognitive hierarchy, signal misspecification, such as partisan bias, and preference misspecification from social perception biases, such as the false consensus effect and pluralistic ignorance. For each case, we illustrate how to calculate the set of asymptotic learning outcomes and derive comparative statics for how this set changes with the parameters of the misspecification.

KEYWORDS: Social learning, model misspecification, bounded rationality

*We thank Nageeb Ali, Alex Imas, Shuya Li, George Mailath, Margaret Meyer, Ali Polat, Andrew Postlewaite, Andrea Prat, Yuval Salant, Larry Samuelson, Joel Sobel, Ran Spiegler and conference and seminar participants for helpful comments and suggestions. Aislinn would like to thank the briq Institute for research support while writing this paper.

[†]Email: abohren@sas.upenn.edu; Carnegie Mellon University and University of Pennsylvania
[‡]Email: daniel.hauser@aalto.fi; Aalto University

1 Introduction

Faced with a new decision, individuals gather information from many diverse sources before choosing an action. This can include the choices or outcomes of peers, the announcements of public institutions, such as a government or health agency, as well as private sources, such as past experiences in similar situations. For example, when deciding whether to enroll in a degree program, an individual may read pamphlets and statistics about the opportunities the program provides, discuss the merits of the program with faculty, observe the enrollment choices and job placement of other students, and consider her own prior education experiences. Learning from these sources requires a model of how to interpret such signals and how to infer from the actions and outcomes of others.

A rich literature in psychology and economics documents that individuals can be systematically biased when processing information and interpreting the decisions of others. Depending on the context, individuals have been shown to overreact or underreact to new information (over- and under-confidence), slant information towards a preferred state (motivated reasoning, partisan bias), differentially weight information based on their prior beliefs (confirmation bias), incorrectly aggregate correlated information (correlation neglect), misunderstand strategic interaction (level-k, cognitive hierarchy), and miscalculate the extent to which others' preferences are similar to their own (false consensus effect, pluralistic ignorance).[1,2]

In this paper, we develop a framework to represent these cognitive biases as forms of model misspecification where individuals have incorrect models of the informational environment and how others make decisions. Importantly, the framework is not specific to a given set of biases and can be used to model broad classes of systematic deviations from Bayesian learning with a correctly specified model. It can be used to represent non-Bayesian learning rules, such as the counting heuristic (Ungeheuer and Weber 2017),

[1]Overconfidence: Moore and Healy (2008); Ortoleva and Snowberg (2015); Motivated reasoning / partisan bias: Bartels (2002); Bénabou and Tirole (2011); Brunnermeier and Parker (2005); Jerit and Barabas (2012); Kőszegi (2006); Kunda (1990); Confirmation bias: Darley and Gross (1983); Lord, Ross, and Lepper (1979); Plous (1991); Correlation neglect: Enke and Zimmermann (Forthcoming); Eyster and Weizsacker (2011); Kallir and Sonsino (2009); Level-k / cognitive hierarchy: Kübler and Weizsäcker (2004, 2005); Penczynski (2017) Social perception bias: Gilovich (1990); Grebe, Schmid, and Stiehler (2008); Marks and Miller (1987); Miller and McFarland (1987, 1991); Ross, Greene, and House (1977). Theories of cognitive limitations provide a foundation for such biases. For example, bounded memory leads to behavior consistent with many documented behavioral biases, including belief polarization, confirmation bias and stickiness (Wilson 2014), while selective awareness leads to confirmation bias and conservatism bias (Gottlieb 2015).

[2]The context of the social learning setting will determine which biases are of first order relevance. For example, pluralistic ignorance often arises in contexts where agents believe that a negative trait affects their own behavior, while the false consensus effect arises for non-normative behaviors such as sexual promiscuity or smoking.

within a Bayesian framework. This representation provides substantial added structure and tractability for analysis.

We explore how model misspecification affects learning in a sequential social learning setting. Each agent is faced with a decision problem: she selects an action, and her payoff depends on her action choice as well as an unknown state of the world. Prior to making this decision, the agent learns about the state from a diverse set of sources: she may observe the actions or outcomes of her predecessors, a private signal and/or a sequence of public signals. An agent's type – her model of the world – specifies her preferences and how she interprets signals, as well as her beliefs about others' preferences and how they interpret signals. A model of the signal process is a subjective belief about the signal distribution in each state, while a model of how others draw inference is a subjective distribution over the types of other agents. Model misspecification refers to the case where these subjective distributions *differ* from the true distributions. To maintain structure, we assume that all types have a common understanding of certain key features of the environment. In particular, they have a common interpretation of the relative informativeness of signals, and they have the same ordinal preferences over their undominated actions when they know the state. This framework captures the information-processing biases cited above, and nests several previously developed behavioral models of learning.[3]

We study asymptotic beliefs and behavior to determine whether individuals with misspecified models adopt the desirable action, and whether those with different models agree or disagree. We know from correctly specified observational learning models that individuals asymptotically adopt the desirable action when there are arbitrarily precise private signals (Smith and Sorensen 2000), actions perfectly reveal beliefs (Ali 2018; Lee 1993), a subset of agents do not observe others' actions (Acemoglu, Dahleh, Lobel, and Ozdaglar 2011), or there is an infinite sequence of public signals. We show that misspecification opens the door to learning outcomes – long run beliefs about the state – that do not occur in the correctly specified model. In particular, asymptotic learning may be *incorrect*, where beliefs converge to the wrong state; agents may perpetually *disagree*, with beliefs converging to different states, despite observing the same information; learning may be *cyclical*, with beliefs that never converge; and *multiple* learning outcomes may arise with positive probability, which leads to path-dependent learning (for example, the same agent may have correct or incorrect learning, depending on initial signals). This represents another distinction from correctly specified social learning models. In such settings, when informational herds are possible, convergence to multiple limit beliefs also

[3]Appendix D demonstrates how our framework nests Bohren (2016); Epstein, Noor, and Sandroni (2010); Rabin and Schrag (1999).

occurs (Banerjee 1992; Bikhchandani, Hirshleifer, and Welch 1992; Smith and Sorensen 2000).[4] However, in contrast to the misspecified setting, all but at most one of these limit beliefs must be non-degenerate. This difference is economically important. Informational herds are fragile and easy to overturn, whereas degenerate beliefs – such as the ones that arise with misspecified models – are not. Lastly, we show that cyclical learning and multiple learning outcomes are distinct features of misspecified *social* learning settings: these learning outcomes do not arise in learning models in which agents are misspecified solely about exogenous sources of information (i.e. public signals).

The asymptotic beliefs that arise in misspecified social learning models have important consequences for behavior. When learning is incorrect, agents make inefficient choices, while when learning is cyclical, action choices oscillate between efficient and inefficient choices infinitely often. For example, when agents learn about the riskiness of a behavior, e.g. binge drinking, and have a misspecified model of others' preferences, they may perpetually oscillate between risky and safe behavior. When multiple learning outcomes arise, relatively sophisticated agents with different life experiences can become very certain that different states of the world are true. For example, an initial signal that a medical technology is dangerous or a new restaurant is low quality, when in fact the opposite is true, can lead to the mistaken belief becoming entrenched. In contrast, if the initial signal had been positive, agents would have learned the correct state. Path dependent learning can explain why different populations with similar models can come to have very different entrenched views.

Our first main result (Theorem 1) characterizes how the set of asymptotic beliefs that arise with positive probability depends on the form of misspecification. We show that this set is determined by two expressions that are straightforward to derive from the primitives of the model: (i) the *expected change in the log likelihood ratio* for each type at each candidate learning outcome; and (ii) an ordering over the type space, which we refer to as *maximal accessibility*. The first condition is used to determine whether a learning outcome is locally stable, in that beliefs converge to this limit belief with positive probability, from a neighborhood of the limit belief. We show that a learning outcome is locally stable if and only if the expected change in the log likelihood ratio moves toward this learning outcome from nearby beliefs.[5] Maximal accessibility determines when beliefs converge to a disagreement outcome with positive probability, starting from a common prior. It establishes that it is possible to separate the beliefs

[4] Banerjee (1992) and Bikhchandani et al. (1992) first studied the sequential observational learning framework with a binary signal space. They demonstrate that incomplete learning may arise when the action space is coarser than the belief space.

[5] This condition relates to the relative entropy of a type's model in each state. Intuitively, a type's beliefs move towards the state that is more likely to generate the observed pattern of actions, outcomes and signals. As discussed below, limit beliefs are a Berk-Nash equilibrium (Esponda and Pouzo 2016).

of different types and push them to a neighborhood of the disagreement outcome. An analogous condition is not necessary for *agreement* outcomes, where types have the same (possibly incorrect) limit beliefs. Beliefs converge to an agreement outcome with positive probability, starting from a common prior, if and only if the agreement outcome is locally stable. These conditions are straightforward to verify from the primitives of the misspecification, and characterize which types of misspecification lead to which patterns of long-run behavior.

To establish Theorem 1, we use results from Markov dynamic systems to characterize the set of limit beliefs for each type. A challenge in social learning settings is that the informational content of actions and outcomes depends on the current belief for each type. Therefore, in principle, the asymptotic properties of beliefs could depend on the behavior of beliefs across the infinite belief space.An important feature of our characterization is that the conditions we outline only need to be verified at a *finite* set of beliefs: that is, the set of beliefs in which all types have degenerate beliefs on one of the states. This significantly simplifies the characterization of asymptotic learning.

Our characterization also establishes a robustness property (Theorems 3 and 4): regardless of the form of misspecification, agents almost surely learn the correct state when they have approximately correct models.[6] This shows that agents do not have to know exactly how their peers behave in order to learn from their choices. Complete learning obtains even in the presence of model heterogeneity, as long as none of these models are too misspecified. This may not seem surprising, since Bayes rule is continuous. But in an infinite horizon setting, a small mistake in each period could sum to a large aggregate mistake. If biases aggregate in this manner, then even arbitrarily small misspecification could interfere with learning, which would in principle limit the applicability of rational learning models. Our results establish when this does not occur.

We close with a demonstration of how our framework can capture three forms of model misspecification prominent in the empirical literature: strategic misspecification, signal misspecification and preference misspecification. For the first category, we demonstrate how our framework can be used to capture learning with *level-k* and *cognitive hierarchy* models of reasoning. Agents correctly interpret signals and payoffs, but have a misspecified model of how others draw inference. In the level-k parameterization, a level-k agent believes that other agents are level-(k-1). We show that, depending on the true distribution over types, either learning is cyclical or agents with different levels of

[6]The robustness result in Bohren (2016) is a special case of this result. In contrast, Madarász and Prat (2016) find a lack of robustness in a mechanism design setting where a principal's model of an agent's preferences is misspecified. Using the optimal mechanism with respect to a misspecified model can lead to non-vanishing losses, even when the level of misspecification is small. This contrasts with our robustness results, in which the losses from misspecification vanish as the misspecified model approaches the correctly specified model.

reasoning come to be arbitrarily certain of different states – they asymptotically disagree. In the cognitive hierarchy parameterization, a level-k agent believes that other agents are level-1 through level-(k-1). In this case, enough information arrives to ensure belief convergence, but depending on the severity of the misspecification, each type may learn the correct or incorrect state. A surprising finding in both cases is that a higher level of reasoning may perform strictly *worse* than a lower level of reasoning.

For the second category of misspecification, we apply our framework to a setting where agents slant their beliefs towards a preferred state, exhibiting *partisan bias*. Nonpartisan types correctly interpret signals, but do not account for the slant of the partisan types. We characterize how the severity of the partisan bias and the frequency of agents who exhibit partisan bias affects asymptotic learning for partisan and non-partisan types. We show that the partisan type's bias, coupled with the nonpartisan type's failure to account for it, can impede the convergence of beliefs or lead to incorrect learning for both types.

Finally, we explore preference misspecification by studying social perception biases. When agents overestimate the similarity between their own preferences and the preferences of others – exhibiting the *false consensus effect* – they may learn the incorrect state, while when agents systematically underestimate this similarity – exhibiting *pluralistic ignorance* – beliefs may not converge. For example, suppose agents are learning about the return to choosing a risky action. If a risk-averse agent overestimates the share of other agents who are also risk-averse (false consensus), she will underestimate the share who choose the risky action. As a consequence, she will observe a higher than expected failure rate. This reinforces her choice of the safe action. In contrast, if a risk-averse agent underestimates the share of other agents who are also risk-averse (pluralistic ignorance) – for example, a college student overestimates the share of students who enjoy heavy drinking despite the risk of failing their classes – then the outcomes of agents who are actually engaged in safe behavior will be misperceived as evidence that the risky choices are actually safe. Here, pluralistic ignorance will prevent agents from learning the negative consequences of high risk behaviors. In both cases, effective interventions to change behavior will require information about the *choices* of others, rather than information about the outcomes of these choices.

Related Literature. A rich literature explores when model misspecification interferes with learning in both individual and social learning settings. The results are mixed: in some cases, misspecification impedes learning about the state or leads to inefficient behavior, while in other cases, misspecified agents still learn the correct state asymptotically. For example, overweighting information (Epstein et al. 2010; Rabin and Schrag 1999), failing to account for redundant information (Bohren 2016; Eyster and Rabin

2010) and selective attention (Schwartzstein 2014) can lead to incorrect learning, while misspecified prior beliefs (Fudenberg, Romanyuk, and Strack 2017; Nyarko 1991), overestimating redundant information (Bohren 2016) and underestimating the similarity of others' preferences (Gagnon-Bartsch 2017) can lead to non-convergence. Correlation neglect can lead to inefficient risk-taking (Levy and Razin 2015), while overconfidence can lead to inefficiently low effort (Heidhues, Koszegi, and Strack 2018) and ideological extremeness (Ortoleva and Snowberg 2015). In contrast, underweighting information (Epstein et al. 2010), using coarse reasoning (Guarino and Jehiel 2013) or using a linear updating heuristic that puts sufficient weight on agents' own signals (Jadbabaie, Molavi, Sandroni, and Tahbaz-Salehi 2012) leads to correct learning almost surely. By providing a general characterization of when misspecification interferes with learning and when it does not, our framework unifies insights from different forms of misspecification. Molavi, AlirezaTahbaz-Salehi, and Jadbabaie (2018) engage in a similar exercise when agents share their beliefs on a network and have imperfect recall. They nest common learning rules that agents use to aggregate *beliefs* (as opposed to actions and outcomes), including the canonical Degroot model, and show how this impacts long-run information aggregation.

Esponda and Pouzo (2016, 2017) explore the implications of model misspecification for solution concepts. In a Berk-Nash equilibrium, agents have a set of (possibly misspecified) models of the world. They play optimally with respect to the model that is the best fit, i.e. the model that minimizes relative entropy with respect to the true distribution of outcomes under the equilibrium strategy profile. Our paper corresponds to the case in which each agent has a single misspecified model in each state.[7] The belief about the state for each type can converge to any limit belief, such that at that limit belief, each type's model in the corresponding state is the best fit, given the observed frequency of actions and signals when each type is choosing the optimal action with respect to this limit belief. This is equivalent to a Berk-Nash equilibrium in a dynamic game with infinitely many players.

A related class of papers explore the foundations of non-Bayesian updating and model misspecification. Ortoleva (2012) axiomatizes a non-Bayesian updating rule in which agents switch models when they observe a sufficiently low probability event. Cripps (2018) axiomatizes a class of non-Bayesian updating processes that are independent of how individuals partition information. Frick, Iijima, and Ishii (2018) show that the false consensus effect can arise when agents' beliefs are derived only from local interactions in an assortative society.

[7] In our framework, an agent's type and a state corresponds to a model of the world in the Esponda and Pouzo (2016) framework.

An older statistics literature on model misspecification complements recent work. Berk (1966) and Kleijn and van der Vaart (2006) show that when an agent with a misspecified model is learning from i.i.d. draws of a signal, her beliefs will converge to the distribution that minimizes relative entropy with respect to the true model. Shalizi (2009) extends these result to a class of non-i.i.d. signal processes. Our environment does not fall into this class of processes. In particular, the asymptotic-equipartition property, which describes the long-run behavior of the sample entropy, is generally not satisfied in social learning environments with model misspecification.

The paper proceeds as follows. Section 2 sets up the model. Section 3 outlines the agent's decision problem and defines a recursive representation of beliefs. Section 4 presents the asymptotic learning characterization and robustness results. Section 5 develops three applications to explore specific forms of misspecification, while Section 6 concludes. All proofs are in the Appendix.

2 The Common Framework

2.1 The Model

States and Actions. There are two payoff-relevant states of the world, $\omega \in \{L, R\}$, with common prior belief $Pr(\omega = L) = p_0$. Nature selects one of these states at the beginning of the game. A countably infinite set of agents $t = 1, 2, ...$ act sequentially and make a single decision $a_t \in \mathcal{A}$, where $\mathcal{A}$ is a finite set with $M \equiv |\mathcal{A}| \geq 2$ actions.

Signals and Histories. Agents learn from private information, public information and the actions of other agents.[8] Before choosing an action, each agent t observes the ordered history of past actions $(a_1, ..., a_{t-1})$, the ordered history of public signals $(y_1, ..., y_t)$, where $y \in \mathcal{Y}$ and $\mathcal{Y}$ is a finite signal space, and a private signal $z_t \in \mathcal{Z}$, where $\mathcal{Z}$ is an arbitrary signal space. Let $h_t = (a_1, ..., a_{t-1}, y_1, ..., y_{t-1})$ denote the action and public signal history.

Suppose signals $\langle z_t \rangle$ and $\langle y_t \rangle$ are i.i.d. across time, conditional on the state, jointly independent, and drawn according to probability measures $\mu_z^\omega \in \Delta(\mathcal{Z})$ and $\mu_y^\omega \in \Delta(\mathcal{Y})$ in state ω. Assume that no private or public signal perfectly reveals the state, which implies that both μ_z^L, μ_z^R and μ_y^L, μ_y^R are mutually absolutely continuous with common supports, which we assume to be $\mathcal{Z}$ and $\mathcal{Y}$. Finally, assume that some signals are *informative*, which rules out the trivial case where both $d\mu_z^L/d\mu_z^R = 1$ almost surely and $d\mu_y^L/d\mu_y^R = 1$ almost surely.

Given private signal z, the private belief that the state is L is $s(z) = 1/(1 +$

[8] In the introduction, we reference the possibility that agents also learn from the stochastic outcomes of other agents. To maintain clarity, we present the case in which agents learn solely from actions here, and present the extension in which agents also learn from outcomes in Appendix B. The analysis and characterization is almost identical.

$d\mu_z^R/d\mu_z^L(z))$. Let c.d.f. $F^\omega(s) \equiv \mu_z^\omega(z|s(z) \leq s)$ denote the distribution of s, and $\mathcal{S} \subseteq [0,1]$ denote the convex hull of the common support of s. Private beliefs are bounded if $\inf \mathcal{S} > 0$ and $\sup \mathcal{S} < 1$, and unbounded if $\mathcal{S} = [0,1]$. Similarly, given public signal y, the belief that that state is L is $\sigma(y) = 1/(1 + d\mu_y^R/d\mu_y^L(y))$. Let c.d.f. $G^\omega(\sigma) \equiv \mu_y^\omega(y|\sigma(y) \leq \sigma)$ denote the distribution of σ, and $\Sigma \subset [0,1]$ denote the common support of σ. Let $\sigma_L \equiv \max_{y \in \mathcal{Y}} \sigma(y)$ denote the *maximal public signal* in state L, i.e. the posterior belief corresponding to the public signal that is the strongest evidence for state L, and analogously, let $\sigma_R \equiv \min_{y \in \mathcal{Y}} \sigma(y)$ denote the maximal public signal in state R. By the finiteness of the public signal space, these signals exist.

Timing. At time t, agent t observes the history h_t and the private signal s_t, then chooses action a_t. Then public signal y_t is realized and the history h_{t+1} is updated to include (a_t, y_t).[9]

Types Framework. Agent t has privately observed type $\theta_t \in \Theta$ drawn from distribution $\pi \in \Delta(\Theta)$, where $\Theta \equiv (\theta_1, ..., \theta_n)$ is a non-empty finite set. An agent's type specifies her model of inference and preferences. A model of inference determines how a type interprets information from signals and prior actions and forms its belief about the state. Preferences determine which action this type chooses, given its belief about the state.

Models of Inference. For each type θ_i, a model of inference includes (i) a subjective private signal distribution $\hat{\mu}_{z,i}^\omega$ in each state $\omega \in \{L, H\}$, (ii) a subjective public signal distribution $\hat{\mu}_{y,i}^\omega$ in each state $\omega \in \{L, H\}$, and (iii) a subjective distribution of types $\hat{\pi}^i \in \Delta(\Theta)$. Assume that each type θ_i believes that no private or public signal perfectly reveals the state, and does not observe a signal that is inconsistent with its model of inference, which implies that $(\hat{\mu}_{z,i}^L, \hat{\mu}_{z,i}^R)$ and $(\hat{\mu}_{y,i}^L, \hat{\mu}_{y,i}^R)$ are mutually absolutely continuous and have full support on $\mathcal{Z}$ and $\mathcal{Y}$, respectively. Given private signal z, type θ_i's subjective private belief that the state is L is $\hat{s}_i(z) = 1/(1 + d\hat{\mu}_{z,i}^R/d\hat{\mu}_{z,i}^L(z))$. Similarly, we can write type θ_i's subjective belief that the state is L, given public signal y, as $\hat{\sigma}_i(y) = 1/(1 + d\hat{\mu}_{y,i}^R/d\hat{\mu}_{y,i}^L(y))$.

We focus on forms of misspecification in which agents have a common understanding of some aspects of the private and public signals. Define two pairs of signal measures as *aligned* if they have the same ordinal ranking over the informativeness of signals. In other words, for any two signals z and z', if z is stronger evidence for state L than z' under one measure, then z is also stronger evidence for state L than z' under the other measure.

[9] Allowing agent t to observe y_t before choosing an action does not affect the analysis, but complicates the notation.

Definition 1 (Aligned Signals). *Given signal space $\mathcal{Z}$, the mutually absolutely continuous probability measures $\mu^L, \mu^R \in \Delta(\mathcal{Z})^2$ and $\nu^L, \nu^R \in \Delta(\mathcal{Z})^2$ are* aligned *if for any $z, z' \in \operatorname{supp}\mu \cap \operatorname{supp}\nu$ such that $\frac{d\mu^L}{d\mu^R}(z) \geq \frac{d\mu^L}{d\mu^R}(z')$, then $\frac{d\nu^L}{d\nu^R}(z) \geq \frac{d\nu^L}{d\nu^R}(z')$, with equality iff $\frac{d\mu^L}{d\mu^R}(z) = \frac{d\mu^L}{d\mu^R}(z')$.*

We assume that each type's subjective private and public signal distributions are aligned with the true private and public signal distributions. We allow one exception for the possibility that a type believes signals are completely uninformative.

Assumption 1 (Aligned Subjective Signals). *For all $\theta_i \in \Theta$, either the subjective private signal distributions $\hat{\mu}^L_{z,i}, \hat{\mu}^R_{z,i}$ are aligned with the true distributions μ^L_z, μ^R_z or the subjective private signal distributions are uninformative $\hat{\mu}^L_{z,i} = \hat{\mu}^R_{z,i}$, and similarly for the subjective public signal distributions $\hat{\mu}^L_{y,i}, \hat{\mu}^R_{y,i}$.*

Under Assumption 1, all types ordinally rank signals in the same way, in terms of which signals are more or less indicative of state L. Types may differ in the degree to which a signal influences their belief about the state – both relative to other types and relative to the true distribution. When public signals are aligned, the maximal public signals σ_L and σ_R are also maximal with respect to each agent's subjective public signal distribution.

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 a belief $p \in [0, 1]$ that the state is L, the expected payoff from choosing action a is $pu_i(a, L) + (1 - p)u_i(a, R)$. An agent chooses the action that maximizes her expected payoff. For each type, 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 that no action is dominated for all types.

We focus on settings in which agents generate information in a common way, in terms of their action choices. This restricts how preferences vary across types. Define a set of utility functions as aligned if, under complete information, each utility function has the same ordinal ranking over undominated actions.

Definition 2. *Utility functions $u_1, ..., u_n$ are* aligned *if there exists a complete order $\succ$ on $\mathcal{A}$ such that if $a \succ a'$, then for all $i = 1, ..., n$, either $u_i(a, L) > u_i(a', L)$ or a is dominated.*[10]

The definition places no restrictions on how to order actions that are optimal for a single type or how a type ranks its dominated actions. Smith and Sorensen (2000) establish

[10] For any undominated actions a and a', if $u_i(a, L) > u_i(a', L)$, then $u_i(a, R) < u_i(a', R)$. Therefore, this definition implies that utility functions also have the same ordinal ranking over undominated actions when the state is R, i.e. if $a \succ a'$, then for all $i = 1, ..., n$, either $u_i(a, R) < u_i(a', R)$ or a' is dominated.

that confounded learning can arise when types have preferences that are not aligned, such as $u_1 = \mathbb{1}_{a=\omega}$ and $u_2 = \mathbb{1}_{a \neq \omega}$. The same is true with misspecification. Therefore, we restrict attention to settings in which confounded learning does not arise in the correctly specified model by assuming that preferences are aligned.

Assumption 2 (Aligned Preferences)**.** *The set of types Θ have aligned preferences.*

Assumption 2 implies common knowledge that preferences are aligned, since all agents believe that other agents have a type in Θ, and so on.

Given Assumption 2, we maintain a complete order over the action space $\mathcal{A}$ by relative preference in state L. Fixing an order $\succ$ that satisfies Definition 2, index actions to correspond to this order, i.e. $\mathcal{A} \equiv (a_1, ..., a_M)$, where $a_m \succ a_l$ iff $m > l$.[11] Under this order, a_M denotes the *maximal action* in state L, and a_1 denotes the *maximal action* in state R.

Categories of Types. We can broadly group types into four categories based on their models of inference: noise, autarkic, sociable and correct. A *noise* type does not use its private signal or the history to learn about the state. We can model this using the types framework by defining a noise type to believe that private and public signals are uninformative, $\hat{\mu}^L_{z,i} = \hat{\mu}^R_{z,i}$ and $\hat{\mu}^L_{y,i} = \hat{\mu}^R_{y,i}$. Noise types also believe that actions reflect no information about the state, which is modeled as the belief that 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 private signal, but not the history. It believes that its private signal is informative, $\hat{\mu}^L_{z,i} \neq \hat{\mu}^R_{z,i}$, the public signal is uninformative, $\hat{\mu}^L_{y,i} = \hat{\mu}^R_{y,i}$, and all agents are noise types, $\hat{\pi}_i(\Theta_N) = 1$. To avoid the trivial case in which an autarkic type is observationally equivalent to a noise type, we assume that an autarkic type has preferences such that it has at least two undominated actions on the set of posterior beliefs that arise from its subjective private signal distribution. A type is *sociable* if it uses the history to learn about the state. These types believe that either actions or the public signal are informative. Finally, a *correct* type has a correct model of inference, $\hat{\mu}^\omega_{z,i} = \mu^\omega_z$, $\hat{\mu}^\omega_{y,i} = \mu^\omega_y$ and $\hat{\pi}_i = \pi$.

Let Θ be ordered such that the first k types are sociable and the remaining $n - k$ types are noise and autarkic types. Let $\Theta_S = (\theta_1, ..., \theta_k)$ denote the set of sociable types, Θ_A denote the set of autarkic types and Θ_N denote the set of noise types.

[11] This order is not necessarily unique. Definition 2 places no restriction on how actions that are optimal for a single type are ordered. For example, if one type chooses action L_1 when $p \geq 1/2$, and otherwise chooses R_1, and a second type chooses action L_2 when $p \geq 1/2$, and otherwise chooses R_2, then both the orders $L_1 \succ L_2 \succ R_1 \succ R_2$ and $L_2 \succ L_1 \succ R_1 \succ R_2$ satisfy Definition 2. This is not a problem, as any order that satisfies Definition 2 can be used.

Adequate Consistent Information. We focus on settings in which adequate information arrives for agents to learn the state in a correctly specified model, and study whether and how misspecification interferes with such learning. We know from Smith and Sorensen (2000) that incomplete learning arises in correctly specified models when there are no public signals or autarkic types and private signals are uniformly bounded in strength. The same is true for misspecified models: if actions and public signals cease to reveal information, and all types are aware of this, then learning will be incomplete. Assumption 3 rules out such settings by assuming that either public signals are informative or autarkic types occur with positive probability. Since autarkic types do not observe the history, their actions are always informative.

Assumption 3 (Adequate Information)**.** *Either (i) public signals are informative, $d\mu_y^L/d\mu_y^R \neq 1$, and all sociable types $\theta_i \in \Theta_S$ believe that public signals are informative, $d\hat{\mu}_{y,i}^L/d\hat{\mu}_{y,i}^R \neq 1$, or (ii) there exists an autarkic type $\theta_j \in \Theta_A$ with $\pi(\theta_j) > 0$ that plays actions a_1 and a_M with positive probability, and each sociable type $\theta_i \in \Theta_S$ believes this autarkic type exists, $\hat{\pi}_i(\theta_j) > 0$.*

This assumption ensures that actions or public signals are informative, and sociable types believe that actions or public signals are informative.

We also focus on settings in which the observed history is consistent with each type's model of inference, in that types do not observe what they believe to be zero-probability histories. In the case of a single type, the type trivially has a correct model of the type distribution, and consistency is not an issue.[12] With multiple types, a type may have a model of inference that places probability zero on an action that occurs with positive probability. To rule out this possibility, we assume that sociable types believe that there is an autarkic or noise type that plays each action with positive probability (this probability can be arbitrarily small).

Assumption 4 (Consistent Information)**.** *When there are multiple types, $|\Theta| \geq 2$, then for each $a \in \mathcal{A}$ and for each sociable 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 sociable type believes that all histories are on the equilibrium path, and we do not need to model how a type reacts to zero probability events.

Any misspecified model can be slightly perturbed so that it satisfies Assumptions 3 and 4 by either (i) perturbing an uninformative public signal distribution so that it is

[12] When $|\Theta| = 1$, it must be that this type has a correct belief about the distribution of types, $\hat{\pi}^1(\theta_1) = \pi(\theta_1) = 1$, and all observed actions will be consistent. Note that even if only one type actually exists, $\pi(\theta_1) = 1$, if this type believes that there is another type θ_2, i.e. $\hat{\pi}^1(\theta_2) > 0$, then $\Theta = (\theta_1, \theta_2)$ and we are in the case with more than one type.

slightly informative, or (ii) perturbing the type distribution to add an autarkic or noise type that occurs with arbitrarily small probability.

2.2 Examples

This types framework can capture many information-processing biases and models of reasoning about others' action choices that have been studied empirically, as illustrated in the following examples.

Partisan Bias. Individuals have an information-processing bias that systematically slants signals towards one state. A parameterization that slants private signals towards state L is $\hat{s}(z) = s(z)^{\nu}$, where $\nu < 1$ (Bartels 2002; Jerit and Barabas 2012).

Under/Overconfidence. Agents underweight or overweight signals. For example, $\frac{\hat{s}(z)}{1-\hat{s}(z)} = (\frac{s(z)}{1-s(z)})^{\nu}$, where $\nu \in [0,1)$ corresponds to underweighting and $\nu \in (1,\infty)$ corresponds to overweighting (Angrisani, Guarino, Jehiel, and Kitagawa 2018; Moore and Healy 2008).

Correlation Neglect. Agents underestimate the correlation in the actions of prior agents: the true share of autarkic types is $\pi(\Theta_A)$, but sociable types believe that the share of autarkic types is $\hat{\pi}(\Theta_A) > \pi(\Theta_A)$ (Enke and Zimmermann Forthcoming).

Level-k/Cognitive Hierarchy. All types have a misspecified distribution of types. Level-1 believes all other agents are noise types, and behaves as an autarkic type. Level-2 believes all other agents are level-1, and interprets all prior actions as independent private signals. Level-3 believes all other agents are level-2, and so on. The cognitive hierarchy model is similar, but allows agents to have a richer belief structure over the types of other agents: a level-k type places positive probability on levels 0 through k-1 (Penczynski 2017).

False Consensus Effect: Payoffs. Agents overweight the likelihood that others have similar preferences. For example, there are two types with preferences $u_1 \neq u_2$. Both types believe that others have the same preferences as their own, $\hat{\pi}^1(\theta_1) = 1$ and $\hat{\pi}^2(\theta_2) = 1$ (Marks and Miller 1987; Ross et al. 1977).

False Consensus Effect: Signals. Agents overweight the likelihood that others have similar models of inference. For example, there are two types. Type θ_1 has a correct model of the signal process, $r_1(s) = s$, and believes that other agents do as well, $\hat{\pi}^1(\theta_1) = 1$. Type θ_2 has partisan bias, $r_2(s) = s^{0.5}$, and believes other agents interpret information in the same way, $\hat{\pi}^2(\theta_2) = 1$ (Gilovich 1990).

Pluralistic Ignorance. Agents underweight the likelihood that others have similar preferences or models of inference. For example, all agents have preferences u_1, but believe that others have preferences u_2. Alternatively, all agents correctly interpret private signals, but believe that others are overconfident (Miller and McFarland 1987, 1991).

Limited Recall. A single long-run agent has limited memory: she can recall past actions, but not past signals. The agent may also be misspecified in how she recalls these actions.

2.3 Discussion of Model

We briefly comment on several notable features of the model.

Types as Models. An agent's type captures her model of the world, which she uses to learn about the environment. Our types framework implicitly restricts this learning to features of the environment that are directly payoff-relevant, i.e. the state. We can divide type θ_i's model into a model of the world in state L, which consists of the type distribution and the signal distributions in state L, $(\hat{\mu}^L_{z,i}, \hat{\mu}^L_{y,i}, \hat{\pi}_i)$, and analogously, a model of the world in state R, $(\hat{\mu}^R_{z,i}, \hat{\mu}^R_{y,i}, \hat{\pi}_i)$.[13] We take this model in each state as fixed, and explore long-run learning about the state.

Hierarchies of Beliefs and Aligned Type Spaces. When agents have heterogenous models, agents may be misspecified about how other agents learn.[14] This can lead to complicated higher-order beliefs. For example, when an agent believes that other agents have partisan bias, we also need to model what this agent believes that these partisan bias agents believe about others. In our framework, these higher-order beliefs are fully captured by the subjective type distributions. If type θ_i believes that all agents are type θ_j, then type θ_j's subjective 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, the type space of models Θ determines the set of belief hierarchies that we consider, and hence, determines the belief type space.

We place two restrictions on Θ, which structures the forms of model misspecification we consider. First, we focus on *aligned* type spaces in which all agents generate and interpret information in a common way. Aligned signals (Assumption 1) guarantee that agents have a common interpretation of the relative order of signals as evidence for state L, and aligned preferences (Assumption 2) guarantee that the action choices of agents are ordered in a way that reflects the same relative strength of evidence for state L.[15] For example, it is common knowledge that lung cancer is stronger evidence that smoking has a negative impact on the lungs, relative to shortness of breath, but agents may differ in their beliefs about the magnitude of these two signals. Or agents have the

[13]This set-up implicitly assumes that the type distribution is the same in both states. It is a straightforward extension to allow the true and/or subjective type distributions to depend on ω.

[14]This is not possible with a single type, $|\Theta| = 1$. Trivially, a single type has a correct model of the type distribution, $\hat{\pi}_1(\theta_1) = 1$.

[15]Common knowledge of the same signal distribution is nested as a special case of our setting, in which all types have the same model of inference. This does not preclude misspecification: the model of the signal distribution may be incorrect.

same preferences between a risky and a safe asset when they are certain about the state, but differ in their risk preferences, and therefore, the threshold belief about the state at which they are willing to start investing in the risky asset.

Second, we assume that Θ is finite. This limits the number of models that each type can attribute to other agents. It also rules out infinite chains of models of the form: type θ_i believes all agents are type θ_{i+1}, type θ_{i+1} believes all agents are θ_{i+2}, etc. for $i = 1, 2, ...$

Individual Learning Model. Our framework nests an individual learning model in which a long-run agent learns from a sequence of exogenous signals. To model such settings, suppose that the public signal is informative, the private signal is uninformative, and there is a single type with an uninformative subjective private signal distribution. Then actions contain no private information. This is isomorphic to a model in which a single long-run agent observes a sequence of signals.[16]

Extensions. We assume that agents have a common prior about the state and that private signals are drawn from the same distribution for all types. It is straightforward to extend the types framework to allow for heterogenous prior beliefs about the state, i.e. type θ_i has a prior belief $p_{i,0}$, and to allow private signals to be drawn from different distributions, i.e. type θ_i has signals drawn from $\mu^{\omega}_{z,i}$. Now, an agent's belief about a type captures both what this agent believes is the type's signal distribution, as well as what this agent believes that the type believes is its signal distribution. The analysis carries through unchanged using this augmented definition of a type (albeit with more burdensome notation).

We assume that the action and public signal spaces are finite, and the state space is binary for technical convenience. Allowing for a continuous action and public signal space, or a finite state space would not qualitatively change the analysis. Generalizing to an infinite dimensional state space would introduce significant technical overhead.

3 Action Choices and Beliefs

A Signal Representation. It is convenient to work with the private beliefs $s(z)$ and $\hat{s}(z)$, rather than the underlying private signal z, as the private belief space is ordered. We show that when signals are aligned (Assumption 1), we can define a mapping between the true private belief and the misspecified private belief for each type that is sufficient for the underlying signal. This allows us to work in the private belief space, and also provides an interpretation for the misspecification.

When signals are aligned, if z and z' lead to the same true belief, $s(z) = s(z')$, then

[16] It is also isomorphic to a model in which multiple long-run agents observe the same sequence of public signals.

they also lead to the same subjective belief for each type θ_i, $\hat{s}_i(z) = \hat{s}_i(z')$. Similarly, if z leads to a higher true belief than z', $s(z) > s(z')$, then it also leads to a higher subjective belief for each type θ_i that believes private signals are informative, $\hat{s}_i(z) > \hat{s}_i(z')$. Therefore, it is possible to represent type θ_i's subjective belief following signal z as a function of the true private belief, $\hat{s}_i(z) = r_i(s(z))$, where $r_i : \mathcal{S} \to [0, 1]$. For each type that believes signals are informative, r_i is a strictly increasing function. Similarly, we can represent type θ_i's belief following public signal y as $\hat{\sigma}_i(y) = \rho_i(\sigma(y))$, where $\rho_i : \Sigma \to [0, 1]$ has the same properties as r_i. Lemma 15 in Appendix C derives this result.[17]

Given these representations, we can work directly with the processes $\langle s_t \rangle$ and $\langle \sigma_t \rangle$ as signals, where $s_t \equiv s(z_t)$ is referred to as the private signal and $\sigma_t \equiv \sigma(y_t)$ is referred to as the public signal. The functions $r_i(s)$ and $\rho_i(\sigma)$ determine type θ_i's posterior beliefs following signals s and σ. Let $\hat{F}_i^\omega(s) \equiv \hat{\mu}_{z,i}^\omega(z|s(z) \le s)$ denote type θ_i's subjective distribution of s, and $\hat{G}_i^\omega(\sigma) \equiv \hat{\mu}_{y,i}^\omega(y|\sigma(y) \le \sigma)$ denote type θ_i's subjective distribution of σ. The tuple $\{\hat{F}_i^L, \hat{F}_i^R, \hat{G}_i^L, \hat{G}_i^R\}$ is sufficient for representing type θ_i's subjective signal distribution, and we do not need to keep track of the underlying measures on $\mathcal{Y}$ or $\mathcal{Z}$.

The 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, $P_i(h|\omega)$, and applies Bayes rule to form the likelihood ratio

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

that the state is L versus R. For autarkic or noise types, $\lambda_i(h) = p_0/(1-p_0)$ for all h, since these types believe that the history is uninformative. Let $\boldsymbol{\lambda}(h) \equiv (\lambda_1(h), ..., \lambda_k(h))$ denote the vector of likelihood ratios for sociable types $(\theta_1, ..., \theta_k)$.

To construct $\lambda_i(h)$ for sociable types, we first consider each type's decision rule. Given a belief $\lambda_i(h)$, the agent observes signal s, uses her model of inference to compute the likelihood of s in state L versus R, and applies Bayes rule again to form the private posterior likelihood ratio

$$\frac{p_i(\lambda_i(h), s)}{1 - p_i(\lambda_i(h), s)} \equiv \frac{P_i(L|h, s)}{P_i(R|h, s)} = \lambda_i(h)\left(\frac{r_i(s)}{1 - r_i(s)}\right) \tag{2}$$

that the state is L versus R.

[17] Alternatively, we can define a form of misspecification relative to the relationship between $\hat{s}$ and s, and provide a foundation for this misspecification on an underlying signal space. In Appendix C, we also show that for any strictly increasing function $r : \mathcal{S} \to [0, 1]$ with $r(\inf \mathcal{S}) < 1/2$ and $r(\sup \mathcal{S}) > 1/2$, there exists a pair of mutually absolutely continuous probability measures with full support on $\Delta(\mathcal{Z})$ that are represented by r. The same holds for the public signal.

Given $p_i(\lambda_i, s)$, type θ_i chooses the action that maximizes its expected payoff. Since no two actions yield the same payoff in both states, no action is optimal at a single belief, preferences are aligned and actions $(a_1, ...a_M)$ are ordered by relative preference in state L, there exist private belief thresholds $0 = p_{i,0} \leq p_{i,1} \leq ... \leq p_{i,M} = 1$ such that we can partition the belief space into a finite set of closed intervals, with action a_m optimal at beliefs $p \in [p_{i,m-1}, p_{i,m}]$ iff $p_{i,m-1} \neq p_{i,m}$, and a_m dominated iff $p_{i,m-1} = p_{i,m}$. Without loss of generality, assume the tie-breaking rule is to choose the undominated action with the lower index at each interior cut-off $p_{i,m} \in (0, 1)$, i.e. if $p_{i,m-1} \neq p_{i,m}$, choose a_m at belief $p_{i,m}$. Since there are at least two undominated actions, there are at least two intervals with a non-empty interior, $p_{i,m-1} \neq p_{i,m}$. The right hand side of (2) is strictly increasing in s at any interior belief $\lambda_i \in (0, \infty)$. In turn, the posterior belief $p_i(\lambda_i, s)$ is also strictly increasing in s. Therefore, there exist signal cut-offs $0 = \overline{s}_{i,0}(\lambda_i) \leq \overline{s}_{i,1}(\lambda_i) \leq ... \leq \overline{s}_{i,M}(\lambda_i) = 1$ such that type θ_i chooses action a_m at belief λ_i iff the realized private signal $s \in (\overline{s}_{i,m-1}(\lambda_i), \overline{s}_{i,m}(\lambda_i)]$ and $\overline{s}_{i,m-1}(\lambda_i) \neq \overline{s}_{i,m}(\lambda_i)$. For any belief λ_i, this decision rule maps the private signal s into an action choice.

The Likelihood Ratio From the decision rules characterized above, we can determine how each sociable type interprets the history to compute $P_i(h|\omega)$ and $\lambda_i(h)$. Misspecification introduces a wedge between the subjective and true probability of observing each action. An agent's model of inference determines how she interprets each action, while the true signal and type distributions determine the true probability. Suppose an agent of type θ_i has likelihood ratio λ_i. The probability that she chooses action a_m is equal to the probability of observing a private signal in the interval $(\overline{s}_{i,m-1}(\lambda_i), \overline{s}_{i,m}(\lambda_i)]$. This is determined by the true signal distribution, $F^\omega(\overline{s}_{i,m}(\lambda_i)) - F^\omega(\overline{s}_{i,m-1}(\lambda_i))$. Therefore, given $\boldsymbol{\lambda}$ and ω, the true probability of observing action a_m is

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

Type θ_j uses its subjective signal distribution to calculate the probability that θ_i chooses a_m, $\hat{F}_j^\omega(\overline{s}_{i,m}(\lambda_i)) - \hat{F}_j^\omega(\overline{s}_{i,m-1}(\lambda_i))$, and its subjective type distribution to calculate the probability of each type. Therefore, type θ_j believes that action a_m occurs with probability

$$\hat{\psi}_j(a_m|\omega, \boldsymbol{\lambda}) \equiv \sum_{i=1}^{n} \hat{\pi}_j(\theta_i)(\hat{F}_j^\omega(\overline{s}_{i,m}(\lambda_i)) - \hat{F}_j^\omega(\overline{s}_{i,m-1}(\lambda_i))). \tag{4}$$

Similarly, there is a wedge between the subjective probability $d\hat{G}_i^\omega(\sigma)$ and the true probability $dG^\omega(\sigma)$ of each public signal σ. In a slight abuse of notation, let $\hat{\psi}_i(a, \sigma|\omega, \boldsymbol{\lambda}) \equiv \hat{\psi}_i(a|\omega, \boldsymbol{\lambda})d\hat{G}_i^\omega(\sigma)$ and $\psi(a_t, \sigma_t|\omega, \boldsymbol{\lambda}_t) \equiv \psi(a_t|\omega, \boldsymbol{\lambda}_t)dG^\omega(\sigma_t)$ denote θ_i's subjective probability and the true probability of (a, σ) in state ω, respectively. From these expressions, we can construct $\lambda_i(h)$. Following h_t,

$$\lambda_i(h_t) = \left(\frac{p_0}{1-p_0}\right) \prod_{\tau=1}^{t-1} \frac{\hat{\psi}_i(a_\tau, \sigma_\tau | L, \boldsymbol{\lambda}(h_\tau))}{\hat{\psi}_i(a_\tau, \sigma_\tau | R, \boldsymbol{\lambda}(h_\tau))}, \tag{5}$$

where the second term on the right hand side of (5) captures $\frac{P_i(h|L)}{P_i(h|R)}$. As $\boldsymbol{\lambda}_t \equiv \boldsymbol{\lambda}(h_t)$ is sufficient for the history, we suppress the dependence on h_t going forward.

From (5), we can define the likelihood ratio recursively. It begins at $\lambda_{i,1} = p_0/(1-p_0)$ for each $\theta_i \in \Theta_S$. Given $\boldsymbol{\lambda}_t$, following action a_t and public signal σ_t, the likelihood ratio in period $t+1$ updates to

$$\lambda_{i,t+1} = \lambda_{i,t} \left(\frac{\hat{\psi}_i(a_t, \sigma_t | L, \boldsymbol{\lambda}_t)}{\hat{\psi}_i(a_t, \sigma_t | R, \boldsymbol{\lambda}_t)} \right). \tag{6}$$

That is, each type's model of inference determines the new *value* of the likelihood ratio. In contrast, the true probability $\psi(a_t, \sigma_t | \omega, \boldsymbol{\lambda}_t)$ of each action and public signal determines the *probability* that the likelihood ration transitions to this value. In correctly specified models, $\hat{\psi}_i(a, \sigma | \omega, \boldsymbol{\lambda}) = \psi(a, \sigma | \omega, \boldsymbol{\lambda})$, and the likelihood ratio is a martingale in state R.

The behavior of $\langle \boldsymbol{\lambda}_t \rangle_{t=1}^{\infty}$ determines the learning dynamics for each type. Characterizing the behavior of this process is challenging. It is an equilibrium object with nonlinear state-dependent transition probabilities – due to the dependence of equilibrium actions on the current value of $\boldsymbol{\lambda}$, the transition probabilities also depend on $\boldsymbol{\lambda}$. This presents a technical challenge, as the process fails to satisfy standard conditions from the existing literature on Markov chains.

4 Asymptotic Learning

We study the asymptotic learning outcomes – the long-run beliefs about the state – for sociable types. Autarkic and noise types do not learn from the history; therefore, their beliefs following the history are constant across time and their behavior is stationary.

4.1 Asymptotic Learning Outcomes

Without loss of generality, we define asymptotic learning outcomes relative to state R. Let *correct* learning (for type θ_i) denote the event where $\boldsymbol{\lambda}_t \to 0^k$ ($\lambda_{i,t} \to 0$), *incorrect* learning (for type θ_i) denote the event where $\boldsymbol{\lambda}_t \to \infty^k$ ($\lambda_{i,t} \to \infty$), and *cyclical* learning (for type θ_i) denote the event where $\boldsymbol{\lambda}_t$ ($\lambda_{i,t}$) does not converge. Learning is *complete* if correct learning occurs almost surely. Agents asymptotically *agree* when all sociable types have the same limit beliefs, $\boldsymbol{\lambda}_t \to \{0^k, \infty^k\}$, and agents *disagree* when some sociable types have incorrect learning and others have correct learning, $\boldsymbol{\lambda}_t \to \{0, \infty\}^k \setminus \{0^k, \infty^k\}$. Learning is *mixed* if some sociable types have correct or incorrect learning and others have cyclical learning, while learning is *stationary* if beliefs converge for all sociable types.[18]

4.2 Asymptotic Learning Characterization.

Our main result characterizes the asymptotic learning outcomes in misspecified models. In correctly specified models, the likelihood ratio is a martingale, and the Martingale Convergence Theorem can be used to rule out cyclical and incorrect learning. This is not the case in a misspecified model. With even the slightest misspecification, the likelihood ratio is no longer a martingale, as any perturbation of a correctly specified model breaks the equality condition. Therefore, an alternative approach is necessary. We use results on the stability of nonlinear dynamic systems to characterize the limiting behavior of the likelihood ratio.

The characterization we develop depends on two expressions that are straightforward to derive from the primitives of the model, i.e. the type space and the signal distributions. We first characterize the behavior of the likelihood ratio when it is in a neighborhood of a learning outcome. We establish necessary and sufficient conditions for the likelihood ratio to converge to this outcome with positive probability. Second, we determine when the likelihood ratio converges to a given learning outcome with positive probability, from any initial belief. This ensures that our characterization holds independent of the prior belief. Finally, we use these expressions to determine when the likelihood ratio almost surely converges to a stationary learning outcome.

To simplify notation, we present the case for two sociable types in this section, i.e. $k = 2$, and present an analogous derivation for more than two sociable types in Appendix A.2.

Local Stability. A learning outcome is *locally stable* if the likelihood ratio converges to this limit belief with positive probability, from a neighborhood of the belief.

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

The first expression for the characterization, the expected change in the log likelihood ratio, determines whether a learning outcome is *locally stable*. For type θ_i, the expected

[18]We show in Lemmas 1 and 2 that Assumption 3 rules out $\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}$ for any $\boldsymbol{\lambda} \notin \{0, \infty\}^k$. Therefore, we do not define this *incomplete* learning outcome. Regarding disagreement, agents' beliefs will differ when beliefs do not converge, converge for some types but not others, and converge to different limit beliefs. We use the term disagreement to refer to the case in which beliefs converge to different limit beliefs.

change in the log likelihood ratio at belief $\boldsymbol{\lambda}$ depends on the subjective and true probability of each action,

$$\gamma_i(\boldsymbol{\lambda}, \omega) \equiv \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi(a, \sigma|\omega, \boldsymbol{\lambda}) \log\left(\frac{\hat{\psi}_i(a, \sigma|L, \boldsymbol{\lambda})}{\hat{\psi}_i(a, \sigma|R, \boldsymbol{\lambda})}\right). \tag{7}$$

Equation (7) has a natural interpretation. Suppose the true state is R and fix a belief $\boldsymbol{\lambda}$. Then $\gamma_i(\boldsymbol{\lambda}, R)$ is the difference between (i) the Kullback-Liebler divergence from type θ_i's subjective model in state R, $\hat{\psi}_i(\cdot|R, \boldsymbol{\lambda})$ to the true model in state R, $\psi(\cdot|R, \boldsymbol{\lambda})$ and (ii) the Kullback-Liebler divergence from θ_i's subjective model in state L, $\hat{\psi}_i(\cdot|L, \boldsymbol{\lambda})$, to the true model in state R, $\psi(\cdot|R, \boldsymbol{\lambda})$. At a given belief $\boldsymbol{\lambda}$, 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 positive, $\gamma_i(\boldsymbol{\lambda}, R) > 0$, and $\log \lambda_i$ moves towards state L in expectation. Otherwise, $\log \lambda_i$ moves towards state R in expectation.

The sign of each component of $\boldsymbol{\gamma}(\boldsymbol{\lambda}, \omega) \equiv (\gamma_1(\boldsymbol{\lambda}, \omega), \gamma_2(\boldsymbol{\lambda}, \omega))$ determines local stability. Let

$$\Lambda_i(\omega) \equiv \{\boldsymbol{\lambda} \in \{0, \infty\}^2 | \gamma_i(\boldsymbol{\lambda}, \omega) < 0 \text{ if } \lambda_i = 0 \text{ and } \gamma_i(\boldsymbol{\lambda}, \omega) > 0 \text{ if } \lambda_i = \infty\} \tag{8}$$

denote the set of stationary learning outcomes in which the expected change in the log likelihood ratio decreases if $\lambda_i = 0$ and increases if $\lambda_i = \infty$, and let

$$\Lambda(\omega) \equiv \Lambda_1(\omega) \cap \Lambda_2(\omega) \tag{9}$$

denote the set that satisfies this property for both sociable types. We show that a stationary learning outcome $\boldsymbol{\lambda}^*$ is locally stable in state ω if and only if $\boldsymbol{\lambda}^* \in \Lambda(\omega)$ (Lemma 3 in Appendix A.1). In other words, if $\langle \boldsymbol{\lambda}_t \rangle_{t=1}^{\infty}$ converges for all sociable types, then it must converge to a limit random variable whose support lies in $\Lambda(\omega)$. Intuitively, in order for the likelihood ratio to converge to a given learning outcome with positive probability, in expectation, the log likelihood ratio must move towards this learning outcome from nearby beliefs. This also implies that if $\Lambda(\omega)$ is empty, then almost surely at least one type has cyclical learning.[19]

This result significantly simplifies the set of possible limit beliefs. It is straightforward to compute $\Lambda(\omega)$ from the primitives of the model, as we will illustrate in the applications in Section 5.

[19] Local stability for $k > 2$ sociable types is identical, substituting $\{0, \infty\}^k$ as the set of candidate limit beliefs, and $\Lambda(\omega) \equiv \cap_{i=1}^{k} \Lambda_i(\omega)$ as the locally stable set for all types.

Mixed Learning. Next we consider the behavior of the likelihood ratio in the neighborhood of a mixed learning outcome. Consider the mixed outcome in which sociable type θ_1 has correct learning, $\lambda_1^* = 0$, and sociable type θ_2 has cyclical learning. This outcome will almost surely not arise if at $\lambda_1^* = 0$, it is possible for the beliefs of θ_2 to converge, i.e. either $(0, 0)$ or $(0, \infty)$ is locally stable for θ_2. Intuitively, if $\langle \lambda_{2,t} \rangle$ converges with positive probability when $\lambda_1^* = 0$, then almost surely $\langle \lambda_{2,t} \rangle$ 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 $(0, \infty) \notin \Lambda_2(\omega)$. This ensures that in a neighborhood of $(0, 0)$ or $(0, \infty)$, θ_2's beliefs drift away from this outcome.

Generalizing this intuition, let $\Lambda_M(\omega)$ denote the set of mixed learning outcomes in which there are no locally stable beliefs for the non-convergent type,

$$\Lambda_M(\omega) \equiv \{\lambda_i^* \in \{0, \infty\}, i \in \{1, 2\} | \forall \lambda_{-i} \in \{0, \infty\}, (\lambda_i^*, \lambda_{-i}) \notin \Lambda_{-i}(\omega)\}. \tag{10}$$

We establish that if a mixed learning outcome is not in $\Lambda_M(\omega)$, then almost surely it does not occur (Lemma 6 in Appendix A.1). Therefore, if $\Lambda_M(\omega)$ is empty, mixed learning almost surely does not arise. If both agreement outcomes are locally stable, $(0, 0) \in \Lambda(\omega)$ and $(\infty, \infty) \in \Lambda(\omega)$, or both disagreement outcomes are locally stable, $(0, \infty) \in \Lambda(\omega)$ and $(\infty, 0) \in \Lambda(\omega)$, then $\Lambda_M(\omega)$ is empty and no mixed learning outcomes are locally stable.

It is straightforward to compute $\Lambda_M(\omega)$ from $\Lambda_i(\omega)$. Doing so allows us to study whether mixed learning is likely to arise in specific forms of model misspecification. In Section 5, we show that $\Lambda_M(\omega)$ is empty for three commonly studied forms of model misspecification; specifically, mixed learning outcomes almost surely do not arise in these applications.[20]

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. A learning outcome is *globally stable* if the likelihood ratio converges to this outcome with positive probability, from *any* initial belief.

Definition 4 (Global Stability)**.** $\boldsymbol{\lambda}^*$ *is* globally stable *if for any initial belief* $\boldsymbol{\lambda}_1 \in (0, \infty)^2$, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) > 0$.

For an agreement outcome, we show that local stability is necessary and sufficient for global stability (Lemma 4 in Appendix A.1). Aligned signals and preferences (Assump-

[20]When $k > 2$, an analogous condition rules out mixed learning outcomes in which a single type has cyclical learning. We also need to rule out mixed learning outcomes in which more than one type has cyclical learning. This requires joint conditions on $\Lambda_i(\omega)$ for the non-convergent types. See Appendix A.2.

tions 1 and 2) guarantee that there exist signal and action pairs that move the beliefs of all types in the same direction. Therefore, we can construct sequences of actions and signals that occur with positive probability and move the beliefs of all types to a neighborhood of an agreement outcome. Given this, computing $\Lambda(\omega)$ is the only calculation necessary to determine whether correct or incorrect learning occurs with positive probability in state ω. These learning outcomes occur with positive probability if and only if the corresponding limit beliefs, $(0,0)$ or (∞,∞), are in $\Lambda(\omega)$.[21]

Global stability does not immediately follow from local stability for disagreement outcomes. In contrast to agreement outcomes, it is not always possible to construct a sequence of action and public signal realizations that push the likelihood ratio arbitrarily close to the disagreement outcome. There may exist initial values of the likelihood ratio such that a locally stable disagreement outcome is reached with probability zero. For example, if two types are sufficiently close to each other, then disagreement may arise if their initial beliefs are very far apart, but may not be possible if their initial beliefs are close together. Therefore, a failure of local stability is sufficient to ensure that a disagreement outcome does not occur, but local stability does not guarantee that the outcome occurs with positive probability.

For a disagreement outcome to be globally stable, it must be possible to separate the beliefs for the type converging to $\lambda_i = 0$ and the type converging to $\lambda_i = \infty$, starting from any initial belief. The second expression for the learning characterization is a sufficient condition to separate beliefs using maximal actions and signals. Recall from Section 2 that there exists a maximal action and public signal in each state. The maximal action and public signal in state R, denoted (a_1, σ_R), decrease the likelihood ratio, and the maximal action and public signal in state L, denoted (a_M, σ_L), increase the likelihood ratio. The *maximal L-order* partially orders how types update their beliefs following each maximal action and signal.

Definition 5 (Maximal L-Order)**.** *The maximal L-order $\succeq_{\boldsymbol{\lambda}}$ at likelihood ratio $\boldsymbol{\lambda}$ is defined by $\theta_i \succeq_{\boldsymbol{\lambda}} \theta_j$ iff θ_i interprets both maximal action and public signal pairs as stronger evidence of state L than θ_j,*

$$\frac{\hat{\psi}_i(a,\sigma|L,\boldsymbol{\lambda})}{\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda})} \geq \frac{\hat{\psi}_j(a,\sigma|L,\boldsymbol{\lambda})}{\hat{\psi}_j(a,\sigma|R,\boldsymbol{\lambda})} \tag{11}$$

for $(a,\sigma) \in \{(a_1,\sigma_R),(a_M,\sigma_L)\}$. Define the corresponding strict order $\succ_{\boldsymbol{\lambda}}$ if (11) holds with strict inequality for either (a_1,σ_R) or (a_M,σ_L).

Consider disagreement outcome $(0,\infty)$. Suppose that at $(0,0)$, θ_2 interprets both

[21] This also directly follows from local stability for $k > 2$ sociable types.

maximal action and signal pairs as stronger evidence of state L than θ_1, i.e. $\theta_2 \succ_{(0,0)} \theta_1$. Then in a neighborhood of $(0,0)$, we can construct a finite sequence of maximal actions and signals that decrease θ_1's beliefs and increase θ_2's beliefs. Such a sequence occurs with positive probability, since it is finite. Therefore, $\theta_2 \succ_{(0,0)} \theta_1$ is sufficient to separate beliefs in the direction of $(0,\infty)$. Similarly, $\theta_2 \succ_{(\infty,\infty)} \theta_1$ is also sufficient. The following definition formalizes this condition for each disagreement outcome.

Definition 6 (Maximal Accessibility)**.** *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* $\theta_1 \succ_{(\infty,\infty)} \theta_2$*.*

As discussed above, from any initial belief, the likelihood enters a neighborhood of each agreement outcome with positive probability. Maximal accessibility establishes that a neighborhood of a disagreement outcome is reached with positive probability from a neighborhood of an agreement outcome. Local stability then establishes convergence. Therefore, maximal accessibility is a sufficient condition for the global stability of a disagreement outcome (Lemma 5 in Appendix A.1).[22] Once again, this condition is straightforward to verify from the primitives of the model. As we illustrate in Section 5.1, one needs to verify (11) at (a_1, σ_R) and (a_M, σ_L) for either beliefs $(0,0)$ or (∞,∞).[23]

Learning Results. Given the sets $\Lambda(\omega)$ and $\Lambda_M(\omega)$ and the disagreement outcomes that are maximally accessible, we can now complete the asymptotic learning characterization. Theorem 1 uses these expressions to characterize the set of asymptotic learning outcomes in each state.

Theorem 1. *Assume there are two sociable types. Given Assumptions 1, 2, 3 and 4 and* $\omega = R$*:*

1. **Agreement.** Correct *learning occurs with positive probability iff* $(0,0) \in \Lambda(R)$ *and* incorrect *learning occurs with positive probability iff* $(\infty,\infty) \in \Lambda(R)$*.*

2. **Disagreement.** *Sociable types* disagree *with positive probability if* $\Lambda(R)$ *contains a maximally accessible disagreement outcome, and sociable types almost surely do not* disagree *if* $\Lambda(R)$ *contains no disagreement outcomes. Each maximally accessible* disagreement *outcome in* $\Lambda(R)$ *occurs with positive probability.*

[22]For $k > 2$ sociable types, we define an analogous notion of maximal accessibility and show that it is sufficient for global stability. Separating the beliefs of more than two types near an agreement outcome requires using the maximal L-order to decrease the beliefs of all types that converge to zero and increase the beliefs of all types that converge to infinity. See Appendix A.2.

[23]An alternative sufficient condition for the global stability of $(0,\infty)$ is $(0,0) \in \Lambda_1(\omega) \setminus \Lambda_2(\omega)$ or $(\infty,\infty) \in \Lambda_2(\omega) \setminus \Lambda_1(\omega)$ (i.e. $\gamma_1(\boldsymbol{\lambda},\omega) < 0$ and $\gamma_2(\boldsymbol{\lambda},\omega) > 0$ for either agreement outcome $\boldsymbol{\lambda} \in \{(0,0),(\infty,\infty)\}$). This condition can be directly verified from the local stability construction, but it will not be satisfied in applications in which both agreement outcomes are locally stable. An analogous condition holds for $(\infty,0)$.

3. **Cyclical Learning.** Cyclical *learning occurs almost surely for all sociable types if $\Lambda(R) \cup \Lambda_M(R)$ is empty, and cyclical learning occurs almost surely for at least one sociable type if $\Lambda(R)$ is empty. Cyclical learning almost surely does not occur for any sociable type if $\Lambda(R)$ contains an agreement outcome or maximally accessible disagreement outcome and $\Lambda_M(R)$ is empty.*

An analogous result holds for $\omega = L$.[24]

An important feature of this characterization is that it requires calculations at a *finite* set of beliefs. In principle, the asymptotic behavior of the likelihood ratio could depend on its behavior across the infinite belief space $[0, \infty]^2$. Since action choices depend on beliefs, $\boldsymbol{\gamma}(\boldsymbol{\lambda}, R)$ may vary with $\boldsymbol{\lambda}$ and each type's beliefs may behave differently at different points in the belief space. However, determining the sign of each component of $\boldsymbol{\gamma}(\boldsymbol{\lambda}, R)$ at all $\boldsymbol{\lambda} \in [0, \infty]^2$ is not necessary to characterize the set of asymptotic learning outcomes. Our characterization establishes that we only need to determine the sign at a *finite* set of beliefs: that is, the set of stationary beliefs $\boldsymbol{\lambda} \in \{0, \infty\}^2$. Deriving $\Lambda(\omega)$ and $\Lambda_M(\omega)$ and verifying maximal stability requires calculating the updates to the likelihood ratio at these four stationary beliefs.[25] Therefore, Theorem 1 significantly simplifies the characterization of asymptotic behavior.

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 disagreement outcomes, we establish a sufficient condition for the outcome to occur (maximal accessibility), and a sufficient condition for outcome not to occur ($\Lambda(\omega)$ empty). In many applications, as we demonstrate in Section 5, all locally stable disagreement outcomes are maximally accessible. Therefore, there is no wedge between the sufficient conditions for disagreement to occur and not to occur – a disagreement outcome arises if and only if it is in $\Lambda(\omega)$. However, this is not always the case. In particular, when a disagreement outcome is locally stable but not maximally accessible, whether disagreement arises can depend on initial beliefs.

If $\Lambda_M(\omega)$ is not empty, then mixed learning may arise. Mixed learning presents a challenge, as we need to consider the movement of the convergent type's likelihood ratio across all possible beliefs for the non-convergent type (in contrast to Theorem 1, where we could restrict attention to stationary beliefs for both types). The following theorem characterizes sufficient conditions for mixed learning to occur with positive probability.

[24] The statement of the theorem is identical for more than two sociable tyeps, using the modified definitions of $\Lambda(\omega)$, $\Lambda_M(\omega)$ and maximal accessibility.

[25] For $k > 2$, the calculations will be at the finite set of beliefs $\{0, \infty\}^k$.

Theorem 2. *Suppose there are two sociable types, Assumptions 1, 2, 3, 4 are satisfied, and the true and subjective private signal distributions have a finite number of discontinuities. If mixed outcome* $\lambda_i \in \Lambda_M(\omega)$ *and (i)* $\sup_{\lambda_{-i}} \gamma_i((0, \lambda_{-i}), \omega) < 0$ *if* $\lambda_i = 0$*, or (ii)* $\inf_{\lambda_{-i}} \gamma_i((\infty, \lambda_{-i}), \omega) > 0$ *if* $\lambda_i = \infty$*, then the mixed outcome occurs with positive probability, and if* $\lambda_i \notin \Lambda_M(\omega)$*, then the mixed outcome almost surely does not occur.*

The intuition is similar to that for convergent learning outcomes. Consider the mixed learning outcome in which $\lambda_1 \to 0$. A sufficient condition for the likelihood ratio of type θ_1 to converge to zero, independently of λ_2, is that the expected change in $\log \lambda_1$ is negative at zero for all possible beliefs of type θ_2, i.e. $\sup_{\lambda_2} \gamma_1((0, \lambda_2), \omega) < 0$. We also need to ensure that type θ_2's beliefs do not converge. By definition of $\Lambda_M(\omega)$, no limit beliefs with $\lambda_1 = 0$ are locally stable for θ_2, i.e. $(0, 0) \notin \Lambda_2(\omega)$ and $(0, \infty) \notin \Lambda_2(\omega)$. Therefore, θ_2's beliefs do not converge.

Action Convergence. Belief convergence forces action convergence: each type eventually settles on an action if and only if its beliefs converge. The limit action choice is efficient if learning is correct, and otherwise is inefficient. If learning is cyclical for a type, then that type chooses all undominated actions in the limit – both efficient and inefficient actions will be chosen infinitely often.

4.3 Robustness of Complete Learning.

An immediate consequence of Theorem 1 is that in correctly specified models, learning is *complete* – correct learning occurs almost surely (this also holds for the case of $k > 2$ in Appendix A.2). Suppose $\omega = R$. When all types are correctly specified, the likelihood ratio is a martingale. Due to the concavity of the log function, this means that the expected change in the log likelihood ratio for each type is negative at all beliefs, $\gamma_i(\boldsymbol{\lambda}, R) < 0$ for all $\boldsymbol{\lambda} \in [0, \infty]^k$. Therefore, 0^k is the unique locally stable belief, $\Lambda(R) = \{0^k\}$, and $\Lambda_M(R)$ is empty.

But from Theorem 1, $\gamma_i(\boldsymbol{\lambda}, R) < 0$ at all interior beliefs is not a necessary condition for complete learning. If $\gamma_i(\boldsymbol{\lambda}, R) < 0$ at all *stationary* beliefs $\boldsymbol{\lambda} \in \{0, \infty\}^k$, then learning is complete. Further, complete learning may obtain even if $\gamma_i(\boldsymbol{\lambda}, R)$ is positive for some types at some stationary beliefs. Therefore, Theorem 1 provides much weaker conditions for complete learning.

Corollary 1. *Complete learning obtains if* $\Lambda(R) = \{0^k\}$ *and* $\Lambda_M(R)$ *is empty.*

These weaker conditions are important for establishing the robustness of complete learning in misspecified models, as with even an arbitrarily small amount of misspecification, the likelihood ratio is no longer a martingale.

The next two theorems establish that correctly specified models are robust to some misspecification, in that learning is complete when sociable types have approximately correct models. This may not seem surprising, since Bayes rule is continuous. But in an infinite horizon setting, a small bias in each period has the potential to sum to a large bias in aggregate. For example, in misaligned learning environments, nearby models with small per-period differences in belief updating can lead to very different limit beliefs. When small biases aggregate to large differences, then even arbitrarily small departures from the correctly specified model will interfere with learning. In principle, this would substantially limit the applicability of rational learning models to real-world settings. Theorems 3 and 4 establish that this does not occur in the aligned environments we consider in this paper.

We first establish that complete learning obtains for any form of misspecification in which each sociable type's subjective model of how to interpret actions and public signals is close enough to the true model. This result depends on the equilibrium probabilities of actions and public signals at stationary beliefs.

Theorem 3. *Given Assumptions 1, 2, 3 and 4, there exists a* $\delta > 0$ *such that if* $|\hat{\psi}_i(a, \sigma|\omega, \boldsymbol{\lambda}) - \psi(a, \sigma|\omega, \boldsymbol{\lambda})| < \delta$ *at stationary beliefs* $\boldsymbol{\lambda} \in \{0, \infty\}^k$ *for all* $(a, \sigma) \in \mathcal{A} \times \Sigma$ *and* $\theta_i \in \Theta_S$*, then learning is complete in state* ω*.*

Theorem 4 presents a sufficient condition on the type space for complete learning to obtain. If all sociable types have subjective type and signal distributions close enough to the true distributions, then learning is complete.[26]

Theorem 4. *Given Assumptions 1, 2, 3 and 4, there exists a* $\delta > 0$ *such that if* $||\hat{\pi}_i - \pi|| < \delta$*,* $||\hat{F}_i^\omega - F^\omega|| < \delta$ *and* $||\hat{G}_i^\omega - G^\omega|| < \delta$ *for all* $\theta_i \in \Theta_S$*, then learning is complete in state* ω*, where* $||\cdot||$ *denotes the supremum metric.*

These robustness results follow from the continuity of $\boldsymbol{\gamma}(\boldsymbol{\lambda}, \omega)$ in each type's subjective signal and type distributions. In any correctly-specified model, $\Lambda(R) = \{0^k\}$ and $\Lambda_M(R) = \emptyset$. By continuity, these sets don't change when some misspecification is introduced. For many important forms of misspecification, including those developed in Section 5, the parameter space in which complete learning obtains is quite large.

Finally, when some types of agents have misspecified models and other types have correctly specified models, these misspecified types do not interfere with the learning of the correctly specified types. A correctly specified type is able to probabilistically parse out the accurate information conveyed by actions.

[26]It is also possible for robustness to hold when agents agents are very wrong about the type distribution, as long as the types that they do believe to occur are "close" to the actual types. For example, suppose that all sociable types are type θ, but believe all types are type $\theta' \neq \theta$. If types θ and θ' have similar preferences and subjective signal distributions, then learning is complete.

Theorem 5. *Given Assumptions 1, 2, 3 and 4, learning is complete for all correctly specified types.*

Similar robustness results hold for individual types that have models close to the correctly specified model.

4.4 Discussion of Results

We briefly discuss the notable features and economic implications of misspecified social learning models.

Misspecified Social Learning. In social learning settings, model misspecification interacts with the endogenous informativeness of actions to give rise to several distinct learning outcomes that do not arise in misspecified learning models with exogenous information. In particular, incorrect learning and disagreement can arise in both settings, but cyclical learning and multiple learning outcomes (for example, both incorrect and correct learning arise with positive probability) are distinct features of misspecified social learning settings.[27] This means that when agents learn from the action choices of their peers, beliefs and actions may not settle down in the long-run, or long-run learning may be path-dependent.

Consider an *exogenous information* setting in which agents learn from the public signal, but not from the actions of others. Model misspecification takes the form of an incorrect model of the signal process. For example, agents overreact to new information or have partisan bias.[28] The informational content of the public signal is exogenous, since the distribution of the public signal is independent of the current belief. Therefore, $\gamma_i(\boldsymbol{\lambda}, \omega)$ is independent of $\boldsymbol{\lambda}$ and its sign is constant across the belief space. It follows from Theorem 1 that there is exactly one locally stable learning outcome.[29]

Corollary 2. *In exogenous information settings, $|\Lambda(\omega)| = 1$ and beliefs almost surely converge to the unique belief in $\Lambda(\omega)$.*

This immediately rules out cyclical learning or convergence to a limit random variable with multiple learning outcomes in its support. Additionally, there is no wedge between local and global stability for disagreement outcomes. Since the expected change in the log likelihood ratio moves in the same direction across the entire belief space, if the

[27] Similarly, cyclical learning may arise in misspecified experimentation models (Fudenberg et al. 2017; Nyarko 1991). In such settings, information is also endogenous, since the signal depends on the action choice of the agent.

[28] As discussed in Section 2.3, this setting is isomorphic to a model with a single long-run agent.

[29] When information is exogenous, the conditions in Theorem 1 collapse to the standard result that beliefs converge to the state that minimizes the relative entropy from the misspecified model in this state to the correct model in the realized state (Berk (1966)).

unique locally stable learning outcome is a disagreement outcome, then it is maximally accessible and occurs almost surely.

In contrast, in social learning settings, agents can also be misspecified about how other agents learn, and their models of how to interpret actions vary with beliefs. Belief convergence requires that when agents are almost certain of a state, the action frequencies they observe confirm their model in that state. If agents are "surprised" when they are almost certain of either state, this leads to cyclical learning (i.e. $|\Lambda(\omega)| = 0$). In contrast, if there are multiple learning outcomes near which the action frequencies confirm each type's model at that outcome, then multiple learning outcomes occur with positive probability (i.e. $|\Lambda(\omega)| > 1$).

Economic Implications of Misspecification. These distinct features have important economic implications. Cyclical learning is a failure of beliefs (and actions) to settle down, even after an arbitrarily long period of time. This means that in the long-run, action choices oscillate infinitely often between efficient and inefficient actions. For example, in Section 5.3, we present a setting in which agents learn about the riskiness of a behavior, e.g. binge drinking. Their misspecified model of the preferences of others leads to perpetual oscillation between risky and safe behavior.

When multiple learning outcomes arise, agents become arbitrarily certain about multiple states. An initial signal that, for instance, a medical technology is dangerous or a new restaurant is low quality when in fact the opposite is true can lead to this misconception becoming entrenched and beliefs converging to the incorrect state. In contrast, if the initial signal had been positive, agents would have learned the correct state. Therefore, a multiplicity of learning outcomes leads to path dependent learning. This can explain why different populations with similar models can come to have very different entrenched views.

Focus on Asymptotic Learning. When incorrect learning, non-convergence, disagreement or multiple learning outcomes arise asymptotically, this illustrates that even if there is an infinite amount of information, we should still expect to observe these "negative" learning outcomes. Therefore, we should also expect to observe inefficient choices, disagreement and belief cycles in finite time. Importantly, the asymptotic results establish that the source of these inefficiencies does not solely stem from a lack of sufficient information to learn the state. Agents are bounded away from efficiency, *irrespective* of the amount of information that they observe.

Characterizing the speed of learning is also an interesting question. The expression $\gamma_i(\boldsymbol{\lambda}, \omega)$ that we use to characterize the locally stable set also determines the asymptotic speed of convergence. The larger this term is in magnitude, the faster the rate of convergence to (or, depending on the sign, the faster the rate of divergence from) the candidate limit belief from a neighborhood of this belief.

Relation to Informational Herding. Convergence to multiple limit beliefs also occurs in correctly specified social learning settings with informational herds (Banerjee 1992; Bikhchandani et al. 1992; Smith and Sorensen 2000). In contrast to misspecified settings, all but at most one of these limit beliefs must be non-degenerate (i.e. incomplete learning). This difference is economically important. In correctly specified models, informational herds are fragile (Bikhchandani et al. 1992). Even though all agents are playing the same action, they remain uncertain about the state. Therefore, a herd of any length can be overturned by a relatively uninformative public signal or other piece of new information. In contrast, when an incorrect herd persists in our setting, beliefs almost surely converge to the incorrect state. This implies that longer herds will become increasingly difficult to overturn.

Learning from Outcomes. In many situations, agents learn from observing the outcomes of others' choices, rather than directly observing their actions. Our learning characterization directly extends to such settings. In Appendix B, we develop the analogue of Theorem 1 for a setting in which agents observe stochastic outcomes, rather than actions. The application in Section 5.3 illustrates the implications of model misspecification in this setting.

5 Applications

Next, we demonstrate how our framework can capture three forms of model misspecification prominent in the empirical literature: strategic misspecification, signal misspecification and preference misspecification. For the first category, we demonstrate how our framework can be used to model learning with cognitive hierarchy and level-k models. For the second category, we apply our framework to a setting in which agents slant their beliefs towards a preferred state, exhibiting partisan bias. For the third category, we explore social perception biases, which include pluralistic ignorance and the false consensus effect. For each category, we illustrate how to calculate the set of asymptotic learning outcomes and derive comparative statics for how this set changes with the parameters of the misspecification. Similar to our robustness results for correctly specified models, our results for these misspecified models are robust: the insights are not sensitive to the exact parameterization used to pin down each bias.

5.1 Strategic Misspecification: Level-k and Cognitive Hierarchy

Cognitive hierarchy and level-k models describe how boundedly rational agents draw inference in strategic settings (Camerer, Ho, and Chong 2004; Costa-Gomes, Crawford, and Iriberri 2009). Agents in these models are characterized by their "depth" of reasoning. The most unsophisticated type, level-0, chooses an action without learning from signals or the actions of others. The other levels correspond to the number of iterated best responses to uninformative level-0 play. Level-1 learns from its own signal but believes that others are level-0, and therefore, does not learn from their actions. Higher levels use progressively more sophisticated reasoning. These models can be viewed as a form of misspecification in which agents have an incorrect model of the strategic link between prior actions.[30]

Penczynski (2017) analyzes experimental data on social learning to determine whether it is consistent with level-k reasoning. He finds evidence of model heterogeneity and inferential naivety. Most agents are level-1, 2 or 3 types, with a modal type of level-2. In this application, we study a setting with level-1, 2 and 3 types. In the level-k parameterization, level-3 believes all agents are level-2, while in the cognitive hierarchy parameterization, level-3 places positive probability on level-1 and level-2 types. We characterize asymptotic learning in both parameterizations.

Types Framework. In the framework of this paper, each level corresponds to a type, $\Theta = \{\theta_0, \theta_1, \theta_2, \theta_3\}$. Level-0 believes that private signals and prior actions are uninformative, i.e. $r_0(s) = 1/2$ and $\hat{\pi}_0(\theta_0) = 1$. The level-1, 2 and 3 types correctly interpret private information, $r_i(s) = s$, but have a misspecified model of how others draw inference. This is captured by a misspecified type distribution. Level-1 types believe that prior actions are uninformative, i.e. all other agents are level-0, $\hat{\pi}_1(\theta_0) = 1$. Level-2 types believe prior actions solely reflect private information from level-1 types, i.e. $\hat{\pi}_2(\theta_1) = 1$. They do not understand that prior actions reflect both private information and redundant information from the prior actions of others. Therefore, they fail to parse out this redundant information. Level-3 types are the most sophisticated: they understand that prior actions contain redundant information, but they do not allow for the possibility that other agents also account for this. They believe that some agents act solely based on their private information, i.e. $\hat{\pi}_3(\theta_1) = 1 - q$ for $q \in (0, 1)$, and other agents do not account for redundant information, i.e. $\hat{\pi}_3(\theta_2) = q$, but they do not account for the presence of other level-3 types, i.e. $\hat{\pi}_3(\theta_3) = 0$. In this set-up, level-0 is a noise type, level-1 is an autarkic type, and level-2 and level-3 are sociable types. Note that the correctly specified model is not a special case of this set-up for either level-2 or level-3 types, as neither type is aware of the level-3 type.

To close the model, assume that level-0 types do not actually exist in the population,

[30] The level-2 type corresponds to the "BRTNI" agents in Eyster and Rabin (2010) and the "naive Bayesians" in Hung and Plott (2001). In Eyster and Rabin (2010), all agents are level-2, but believe that all other agents are level-1. Bohren (2016) can be interpreted as a modified level-k model in which agents are level-1 or level-2, and level-2 agents have a misspecified distribution about the share of other level-2 agents.

$\pi(\theta_0) = 0$, level-1 types occur with positive probability, $\pi(\theta_1) \in (0, 1)$, there are no public signals, and all types have common prior $p_0 = 1/2$. We consider a binary action setting in which all types earn a payoff of one from choosing the action that matches the state, $u(a, \omega) = \mathbb{1}_{a=\omega}$, where $\mathcal{A} = \{L, R\}$.[31]

Trivially, preferences and signals are aligned (Assumptions 1 and 2), since all agents have the same preferences and a correctly specified model of private signals. Level-1 types occur with positive probability, $\pi(\theta_1) > 0$, so adequate information arrives (Assumption 3). Level-2 and level-3 types believe that level-1 types occur with positive probability, $\hat{\pi}_2(\theta_1) > 0$ and $\hat{\pi}_3(\theta_1) > 0$, so all action histories are consistent with these types' models of inference (Assumption 4).

Action Choices and Beliefs. We first construct the action choices and likelihood ratio for each type. Let $\boldsymbol{\lambda} = (\lambda_2, \lambda_3)$ denote the vector of likelihood ratios for the sociable types, θ_2 and θ_3.

A level-1 type incorporates solely its private information into its decision, and its likelihood ratio is constant across time, $\lambda_{1,t} = 1$ for all t. When agent t is level-1, she chooses $a_t = L$ iff $s_t \geq 1/2$. The informational content of her action is independent of the history: she chooses action R with probability $F^\omega(1/2)$, i.e. the probability that a private signal is less than $1/2$, and action L with probability $1 - F^\omega(1/2)$.

A level-2 type believes all past actions are from level-1 types. Its subjective probability of each R action in the history is the probability that a level-1 type chooses action R, $\hat{\psi}_2(R|\omega, \boldsymbol{\lambda}) = F^\omega(1/2)$. Analogously, its subjective probability of each L action is $\hat{\psi}_2(L|\omega, \boldsymbol{\lambda}) = 1 - F^\omega(1/2)$. These probabilities are independent of the action history. Therefore, the number of R and L actions is a sufficient statistic for the likelihood ratio of a level-2 type. One way to view the level-2 type is an agent who uses a simple heuristic: count the number of L and R actions in the history, and use this number to form beliefs. When agent t is level-2, she chooses $a_t = L$ iff $s_t \geq 1/(\lambda_{2,t} + 1)$. The informational content of an action from a level-2 type *does* depend on the history, through belief $\lambda_{2,t}$.

A level-3 type believes past actions are from either level-1 or level-2 types. Its subjective probability of an R action is a weighted average of the probability that level-1 and level-2 types choose action R. When level-2 has belief λ_2, she chooses an R action with probability $F^\omega(1/(\lambda_2 + 1))$. Therefore, level 3's subjective probability of an R action is

$$\hat{\psi}_3(R|\omega, \boldsymbol{\lambda}) = (1 - q)F^\omega(1/2) + qF^\omega\left(\frac{1}{\lambda_2 + 1}\right). \tag{12}$$

[31] These assumptions are made for expositional simplicity. The results from Section 4 apply to any level-k model in which the level-1 type occurs with positive probability, $\pi(\theta_1) > 0$.

The subjective probability of an L action is analogous. Both of these probabilities depend on the current belief of a level-2 type. Therefore, how a level-3 type updates its belief depends on the current belief of the level-2 types. When agent t is level-3, she chooses $a_t = L$ iff $s_t \geq 1/(\lambda_{3,t} + 1)$.

The *true* probability of an R action at time t depends on the correct distribution over types and the current belief of each type,

$$\psi(R|\omega, \boldsymbol{\lambda}) = \pi(\theta_1)F^\omega(1/2) + \pi(\theta_2)F^\omega\left(\frac{1}{\lambda_2 + 1}\right) + \pi(\theta_3)F^\omega\left(\frac{1}{\lambda_3 + 1}\right). \tag{13}$$

This is the distribution that governs the transition probabilities of $\langle \boldsymbol{\lambda}_t \rangle$.

We are interested in the asymptotic learning outcomes of level-2 and level-3 types, since these types learn from the action choices of others. There are four candidate learning outcomes: correct learning for both types, ($(0,0)$ when $\omega = R$), incorrect learning for both types, ((∞, ∞) when $\omega = R$), and disagreement, $(0, \infty)$ or $(\infty, 0)$.

Asymptotic Learning Characterization: Level-k. We study an approximation of the level-k parameterization in which the level-3 type places an arbitrarily small probability on the level-1 type and the remaining probability on the level-2 type, $q \approx 1$.[32] Using Theorem 1, we characterize asymptotic learning outcomes in three steps: (i) construct the set of locally stable learning outcomes $\Lambda(\omega)$, (ii) show both disagreement outcomes are maximally accessible, and (iii) show $\Lambda_M(\omega)$ is empty. It follows from (ii) and (iii) that $\Lambda(\omega)$ fully characterizes the set of asymptotic learning outcomes. This establishes how asymptotic learning depends on the true distribution over types, and illustrates how our framework of model misspecification can rationalize entrenched disagreement.

To construct $\Lambda(\omega)$, we use the ψ and $\hat{\psi}_i$ expressions derived above to determine the sign of $\gamma_2(\boldsymbol{\lambda}, \omega)$ and $\gamma_3(\boldsymbol{\lambda}, \omega)$ at each learning outcome. Suppose the true state is $\omega = R$ and consider the correct learning outcome, $(0,0)$. Both level-2 and level-3 types choose action R for all signals. Therefore, the level-3 type believes that R actions are uninformative, $\lim_{q\to 1} \frac{\hat{\psi}_3(R|L,(0,0))}{\hat{\psi}_3(R|R,(0,0))} = 1$ and L actions are from level-1 types, $\lim_{q\to 1} \frac{\hat{\psi}_3(L|L,(0,0))}{\hat{\psi}_3(L|R,(0,0))} = \frac{1-F^L(1/2)}{1-F^R(1/2)}$. Since only level-1 types play action L, the true probability of an L action is $\pi(\theta_1)(1 - F^R(1/2))$. Therefore,

$$\gamma_3((0,0), R) \approx \pi(\theta_1)(1 - F^R(1/2)) \log\left(\frac{1 - F^L(1/2)}{1 - F^R(1/2)}\right) > 0$$

and correct learning is not locally stable for level-3 types, $(0,0) \notin \Lambda_3(R)$. Intuitively, near correct learning, level-3 types underestimate the informational content of confirming

[32] The exact parameterization of the level-k model, i.e. $q = 1$, violates Assumption 3.

R actions, since they do not account for level-1 types also playing R, but they correctly infer the informational content of contrary L actions, since they correctly attribute them to level-1 types. This pulls their belief towards state L in expectation (i.e. away from 0). Similarly,

$$\gamma_3((\infty,\infty),R) \approx \pi(\theta_1)F^R(1/2)\log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) < 0$$

and incorrect learning is not locally stable for level-3 types, $(\infty,\infty) \notin \Lambda_3(R)$. Now, level-3 types underestimate the informational content of confirming L actions, but correctly infer the informational content of contrary R actions. This establishes that correct learning and incorrect learning almost surely do not occur, as these outcomes are not locally stable for level-3 types.

This leaves the disagreement outcomes as candidate learning outcomes. Consider the disagreement outcome $(0,\infty)$ in which level-2 has correct learning and level-3 has incorrect learning. In this outcome, level-3 believes that R actions are from level-2 types and are approximately uninformative, while L actions are from level-1 types and are therefore informative. The level-3 type is now misspecified about the informativeness of both L and R actions, as it does not account for informative R actions from level-1 types nor uninformative L actions from other level-3 types. The misspecification about the contrary R action dominates, pushing level-3's belief towards state L in expectation. Therefore, the disagreement outcome is locally stable for level-3,

$$\gamma_3((0,\infty),R) \approx (\pi(\theta_1)(1-F^R(1/2)) + \pi(\theta_3))\log\left(\frac{1-F^L(1/2)}{1-F^R(1/2)}\right) > 0.$$

The intuition is similar for disagreement outcome $(\infty,0)$. It is locally stable for level-3,

$$\gamma_3((\infty,0),R) \approx (\pi(\theta_1)F^R(1/2) + \pi(\theta_3))\log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) < 0.$$

Next, we determine whether the disagreement outcomes are locally stable for level-2 types. Level-2 types believe that all actions are from level-1 types. Therefore, they interpret L and R actions in the same way at both disagreement outcomes. However, at $(0,\infty)$, the true probability an R action is $\pi(\theta_1)F^R(1/2) + \pi(\theta_2)$, while at $(\infty,0)$, it is $\pi(\theta_1)F^R(1/2) + \pi(\theta_3)$. Therefore,

$$\gamma_2((0,\infty), R) = (\pi(\theta_1)F^R(1/2) + \pi(\theta_2)) \log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) + (\pi(\theta_1)(1 - F^R(1/2)) + \pi(\theta_3)) \log\left(\frac{1 - F^L(1/2)}{1 - F^R(1/2)}\right)$$

and

$$\gamma_2((\infty,0), R) = (\pi(\theta_1)F^R(1/2) + \pi(\theta_3)) \log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) + (\pi(\theta_1)(1 - F^R(1/2)) + \pi(\theta_2)) \log\left(\frac{1 - F^L(1/2)}{1 - F^R(1/2)}\right).$$

The signs of these expressions vary with the true distribution of types. When

$$\pi(\theta_1)(2F^R(1/2) - 1) + \pi(\theta_2) - \pi(\theta_3) > 0,$$

$\gamma_2((0,\infty), R) < 0$ and $(0,\infty) \in \Lambda_2(R)$. Given $(0,\infty) \in \Lambda_1(R)$, this implies $(0,\infty) \in \Lambda(R)$. When $\pi(\theta_1)(2F^R(1/2) - 1) - \pi(\theta_2) + \pi(\theta_3) < 0$, $\gamma_2((\infty,0), R) > 0$ and $(\infty,0) \in \Lambda_2(R)$. This implies $(\infty,0) \in \Lambda(R)$. Otherwise, neither disagreement outcome is locally stable for level-2 and $\Lambda(R)$ is empty. The construction of $\Lambda(L)$ is identical.

When $\Lambda(\omega)$ contains a disagreement outcome, we need to check whether the disagreement outcome is maximally accessible to determine whether it occurs with positive probability from any initial belief. At beliefs $(0,0)$, θ_2 believes that more R actions are arriving from level-1 types than θ_3, and therefore, $\theta_3 \succ_{(0,0)} \theta_2$. Similarly, at beliefs (∞,∞), θ_2 believes more L actions are arriving from level-1 types than θ_3, and $\theta_2 \succ_{(\infty,\infty)} \theta_3$. Therefore, both disagreement outcomes are maximally accessible.

Finally, we need to rule out mixed learning outcomes in which θ_2's beliefs converge, but θ_3's beliefs cycle, or vice versa. Given the signs of γ_2 and γ_3 characterized above, this would imply that one of the agreement vectors is locally stable, a contradiction. Therefore, $\Lambda_M(\omega)$ is empty and mixed learning does not arise.

Proposition 1 summarizes this construction. Since $\Lambda_M(R)$ is empty and both disagreement outcomes are maximally accessible, $\Lambda(\omega)$ fully determines the set of asymptotic learning outcomes.

Proposition 1 (Level-k)**.** *In the level-k model ($q \approx 1$), either (i) almost surely learning is cyclical and $\Lambda(\omega) = \emptyset$; or (ii) disagreement occurs almost surely and $\Lambda(\omega)$ is a nonempty subset of $\{(0,\infty), (\infty,0)\}$. A disagreement outcome occurs with positive probability iff it is in $\Lambda(\omega)$.*

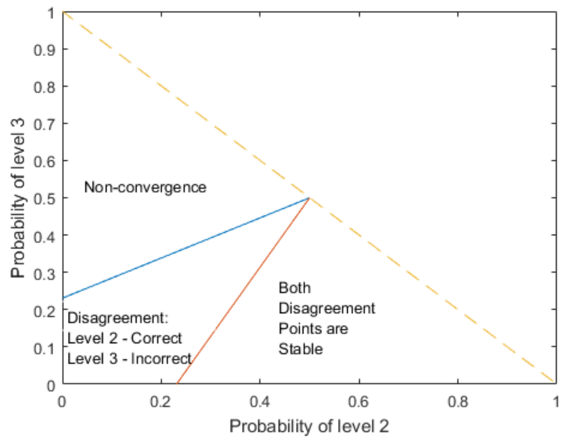

FIGURE 1. Level-k Learning Outcomes
($\omega = R$, $F^L(s) = \frac{5}{3}(s^2 - .04)$, $F^R = \frac{10}{3}(s - \frac{1}{2}s^2 - 3/5)$)

1. *Cyclical learning: 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.*

2. *Disagreement: 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.*

3. *Suppose the private signal distribution is symmetric. If $(\infty, 0) \in \Lambda(R)$, then $(0,\infty) \in \Lambda(R)$ and if $(0,\infty) \in \Lambda(L)$, then $(\infty, 0) \in \Lambda(L)$.*[33]

There are three distinct regions of learning, which depend on the true distribution over types. If a large share of agents are level-3 types, then learning is cyclical. Level-3 types underweight confirmatory actions and are too responsive to contrary actions. This causes both types to doubt their current beliefs and never become certain of either state. As the share of level-2 types increases, convergence becomes possible. For an intermediate share of level-2 types, the disagreement outcome in which level-2 learns the correct state and level-3 learns the incorrect state almost surely arises. A surprising finding is that a higher level of reasoning may perform strictly *worse* than a lower level of reasoning. For a large share of level-2 types, both disagreement outcomes arise. In this case, learning is path dependent. Two different populations who learn about the same state from different action histories may converge to different long-run beliefs. In both cases, disagreement is driven by the level-2 type's desire to imitate and the level-3 type's desire to anti-imitate. Penczynski (2017) estimates that most agents are level-2

[33]More generally, $(\infty, 0) \in \Lambda(R) \Rightarrow (0,\infty) \in \Lambda(R)$ when (i) $\log \frac{1-F^L(1/2)}{1-F^R(1/2)} \leq -\log \frac{F^L(1/2)}{F^R(1/2)}$; (ii) $\log \frac{1-F^L(1/2)}{1-F^R(1/2)} > -\log \frac{F^L(1/2)}{F^R(1/2)}$ and $\pi(\theta_1) > \bar{\pi}_1$ for some cutoff $\bar{\pi}_1 \in (0,1)$. A similar result holds for $\omega = L$.

and very few are level-3. His estimates of the population distribution lead to learning outcomes that lie in the disagreement region. Figure 1 illustrates these learning regions.

Asymptotic Learning Characterization: Cognitive Hierarchy. In the cognitive hierarchy parameterization of this set-up, we study how asymptotic learning varies with the level-3 type's belief q about the frequency of level-2 types. To simplify notation, assume that the types are evenly distributed, $\pi(\theta_1) = \pi(\theta_2) = \pi(\theta_3) = 1/3$, and private signals are symmetrically distributed across states, $F^L(1/2) = 1 - F^R(1/2)$. Similar to the level-k parameterization, we construct $\Lambda(\omega)$ for each q, show that $\Lambda_M(\omega)$ is empty and that both disagreement outcomes are maximally accessible. Again, $\Lambda(\omega)$ fully characterizes the set of asymptotic learning outcomes.

Proposition 2 characterizes how asymptotic learning outcomes depend on q.

Proposition 2 (Cognitive Hierarchy)**.** *The likelihood ratio almost surely converges to a limit random variable with support $\Lambda(\omega) \neq \emptyset$. When $\omega = R$, there exist unique cutoffs $0 < q_1 < q_2 < q_3 < 1$ such that:*

1. *If $q < q_1$, then incorrect and correct learning occur with positive probability, $\Lambda(R) = \{(0,0), (\infty,\infty)\}$.*
2. *If $q \in (q_1, q_2)$, then incorrect learning, correct learning and disagreement occur with positive probability, $\Lambda(R) = \{(0,0), (\infty,\infty), (0,\infty)\}$.*
3. *If $q \in (q_2, q_3)$, then correct learning and disagreement occur with positive probability, $\Lambda(R) = \{(0,0), (0,\infty)\}$.*
4. *If $q > q_3$, then disagreement occurs almost surely, $\Lambda(R) = \{(0,\infty)\}$.*

An analogous result holds for $\omega = L$.

When q is low, level-3 types believe most agents are level-1 and they behave similarly to level-2 types. Both types overweight confirming actions and underweight contrary actions. Initial actions have an outsize effect on asymptotic beliefs, as the information from these actions is amplified in every subsequent action. Therefore, whether initial actions are correct or incorrect will influence whether beliefs build momentum on the correct or incorrect state, leading to either correct or incorrect learning. The models of level-2 and level-3 types are very close, and asymptotic disagreement is not possible.

As q increases, level-2 and level-3 types interpret the action history in an increasingly different way, and disagreement becomes possible. Further, as q increases, level-3 types move closer to the level-k model in which they anti-imitate the more frequent action. Even though level-2's model does not change, the shift in level-3's model leads to behavior that moves level-2's model closer to the correctly specified model. Therefore,

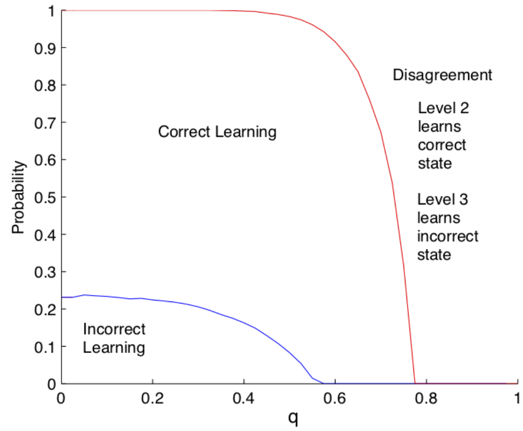

FIGURE 2. Cognitive Hierarchy Learning Outcomes
($\omega = R$, $F^L(s) = \frac{5}{3}(s^2 - .04)$, $F^R(s) = \frac{10}{3}(s - \frac{1}{2}s^2 - 3/5)$)

disagreement takes a specific form: level-2 learns the correct state, while level-3 learns the incorrect state. Once q is sufficiently large, this disagreement outcome becomes the unique learning outcome, and level-3 almost surely learns the incorrect state, while level-2 almost surely learns the correct state.

Figure 2 plots the probability of each learning outcome, as a function of q. Increasing q monotonically increases the probability that level-2 learns the correct state, as level-3's behavior mitigates level-2's bias. However, increasing q has a non-monotonic effect on the probability that level-3 learns the correct state. At first, raising q moves level-3's model closer to the true model, as they become aware of level-2 types. This increases the probability of complete learning. But above $q = .55$, increasing q moves level-3's model further from the true model, as they begin to overestimate the frequency of level-2 types. In this specification, $q_1 = .01$, $q_2 = .55$ and $q_3 = .76$.

While this example focuses on a particular distribution of types, $\pi = (0, 1/3, 1/3, 1/3)$, a robustness result that is similar in spirit to Theorem 3 establishes that Proposition 2 holds for nearby type distributions.

5.2 Signal Misspecification: Partisan Bias

A literature in economics, psychology and political science has documented settings in which individuals systematically slant information towards a particular state. Motivated reasoning (Kunda 1990) leads individuals to systematically slant information towards a preferred state (i.e. personal intelligence) due to self-image concerns (Bénabou and Tirole 2011), ego utility (Kőszegi 2006) or optimism (Brunnermeier and Parker 2005). Party affiliation impacts information-processing: individuals are better at recalling facts that support their political position (Jerit and Barabas 2012), and individuals update their evaluations of candidates in response to new information in a way that is favorable towards their political position (Bartels 2002).

We show that when agents systematically slant their interpretation of signals towards a particular state – that is, they exhibit what we refer to as *partisan bias* – this can impede the convergence of beliefs or lead to incorrect learning. We model partisan bias as a form of misspecification about the signal distributions in each state, and remain agnostic as to its source. We characterize how the severity of the partisan bias and the frequency of agents who exhibit the bias affects asymptotic learning for both partisan and non-partisan types.

Types Framework. Suppose that there are two ways in which agents process information. Some individuals, who we refer to as partisan types, systematically slant information towards state L. Following any private signal, these partisan types believe that state L is more likely than it actually is, given the true measure over signals. We model this as a misspecified private signal distribution $\hat{F}_i^\omega(s) = F^\omega(s^\nu)$, where $\nu \in (0,1)$ parameterizes the level of partisan bias and the true distribution of signals has support $\mathcal{S} = [0,1]$. Given any private signal s, the partisan type's subjective private belief is greater than the true private belief, $r_i(s) = s^\nu > s$.[34] Other individuals, who we refer to as nonpartisan types, correctly interpret private information, $\hat{F}_i^\omega(s) = F^\omega(s)$. The nonpartisan type's subjective private belief is equal to the true private belief, $r_i(s) = s$.

Suppose that some partisan and nonpartisan agents observe the history, and others do not. Therefore, there are four types, $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$. Types θ_1 and θ_3 are partisan sociable and autarkic types, respectively, with $\hat{F}_1^\omega(s) = \hat{F}_3^\omega(s) = F^\omega(s^\nu)$. Types θ_2 and θ_4 are nonpartisan sociable and autarkic types, respectively, with $\hat{F}_2^\omega(s) = \hat{F}_4^\omega(s) = F^\omega(s)$. Let $q \equiv \pi(\theta_3) + \pi(\theta_1) \in (0,1)$ denote the share of partisan types. Suppose that an equal share $\alpha \in (0,1)$ of partisan and nonpartisan types are autarkic, $\pi(\theta_3) = \alpha q$ and $\pi(\theta_4) = \alpha(1-q)$.

In the presence of partisan types, there is an additional challenge to learn from the actions of others, relative to a model in which all agents correctly interpret private signals. To accurately interpret actions, an agent must be aware of the partisan types, and have a correct model of both their level of bias (i.e. ν) and their frequency in the population (i.e. q). We assume that agents are not this sophisticated. In particular, both partisan and nonpartisan types exhibit a false consensus effect: they believe that all agents interpret private information in the same manner as themselves (Marks and Miller 1987). Although sociable nonpartisan types have a correct model of the signal

[34] An alternative interpretation for partisan bias is a type who believes that signals are manipulated towards state R. For instance, suppose vaccines are dangerous in state L and safe in state R. Then a type who believes that the government exaggerates safety information will look at a study of vaccine safety and believe that the results are exaggerated to some degree, i.e. a signal of reported strength s was actually a signal of strength s^ν before it was manipulated.

distribution, they incorrectly assume that other agents do as well. Therefore, they do not invert the bias of the partisan types when learning from actions. This corresponds to believing that no types have partisan bias, $\hat{\pi}_2(\theta_1) = \hat{\pi}_2(\theta_3) = 0$. In contrast, partisan types have a correct model of how other partisan types interpret information, but they have an incorrect model of the signal distribution driving this process and an incorrect model of how nonpartisan types interpret information. This corresponds to believing that all types have partisan bias, $\hat{\pi}_1(\theta_2) = \hat{\pi}_1(\theta_4) = 0$.

To close the model, assume that both partisan and nonpartisan sociable types correctly understand how to account for redundant information in actions – that is, they have correct beliefs about the share of autarkic types in the population. Consider a binary action setting $\mathcal{A} = \{L, R\}$ in which all types earn a payoff of one from choosing the action that matches the state, $u(a, \omega) = \mathbb{1}_{a=\omega}$. Although partisan and nonpartisan agents agree on the optimal action when the state is known, they will potentially disagree on the optimal action following imperfect signals, as the partisan types will believe that signals are more favorable towards state L than nonpartisan types. Assume that there are no public signals and all types have common prior $p_0 = 1/2$.[35]

In this set-up, signals are aligned (Assumption 1), since partisan types order signals in the same way as nonpartisan types, i.e. s^ν is increasing in s. Trivially, preferences are aligned (Assumption 2), since all agents have the same preferences. Autarkic types occur with positive probability, $\alpha > 0$, so adequate information arrives (Assumption 3), and sociable types have a correct belief about the share of autarkic types, so all action histories are consistent (Assumption 4).

Action Choices and Beliefs. We first construct the action choices and likelihood ratios for each type. At belief λ and signal s, each type θ_i chooses action R iff $\lambda \left(\frac{r_i(s)}{1-r_i(s)}\right) \leq 1$. Therefore, given λ, the sociable partisan type plays action R following signals $s \leq \overline{s}_1(\lambda) = 1/(1+\lambda)^{1/\nu}$, while the sociable nonpartisan type plays action R following signals $s \leq \overline{s}_2(\lambda) = 1/(1+\lambda)$. Similarly, the autarkic partisan type plays action R following signals $s \leq (1/2)^{1/\nu}$, while the autarkic nonpartisan type plays action R following signals $s \leq 1/2$. The partisan type's cut-off to choose action L is lower than the nonpartisan type's, $\overline{s}_1(\lambda) < \overline{s}_2(\lambda)$ – at any belief, the partisan type chooses action L for a larger interval of signals, and therefore, with higher frequency.

A partisan type believes that other agents interpret information in the same way. Therefore, it believes that all other agents use cut-off $\overline{s}_1(\lambda)$, whereas fraction $1-q$ of agents are actually using cut-off $\overline{s}_2(\lambda)$. Additionally, it has a misspecified model of the signal distribution – it believes that signals are below $\overline{s}_1(\lambda)$ in state ω with probability

[35] The results from Section 4 apply to any partisan bias model in which $\alpha > 0$, the sociable types believe that $\alpha > 0$ and preferences are aligned.

$\hat{F}_1^\omega(\overline{s}_1(\lambda))$, which is greater than the true probability $F^\omega(\overline{s}_1(\lambda))$. Therefore, the partisan type underestimates the range of signals for which other agents choose action R and overestimates the probability of these signals. Its subjective probability of an R action is

$$\begin{aligned}\hat{\psi}_1(R|\omega,(\lambda_1,\lambda_2)) &= (1-\alpha)\hat{F}_1^\omega(\overline{s}_1(\lambda_1)) + \alpha\hat{F}_1^\omega((1/2)^{1/\nu}) \\ &= (1-\alpha)F^\omega(\overline{s}_2(\lambda_1)) + \alpha F^\omega(1/2),\end{aligned}$$

where the second equality follows from $\overline{s}_1(\lambda) = \overline{s}_2(\lambda)^{1/\nu}$ and $\hat{F}_1^\omega(s) = F^\omega(s^\nu)$. This is greater than the true probability that a *partisan* type plays an R action, due to the signal misspecification.

A nonpartisan type believes that other agents are also nonpartisan and use cut-off $\overline{s}_2(\lambda)$, and it has a correctly specified model of the signal distribution. Therefore, it overestimates the range of signals for which other agents choose action R, since partisan types are using cut-off $\overline{s}_1(\lambda) < \overline{s}_2(\lambda)$, but it correctly estimates the probability of these signals. The nonpartisan type's subjective probability of an R action is

$$\hat{\psi}_2(R|\omega,\lambda_1,\lambda_2) = (1-\alpha)F^\omega(\overline{s}_2(\lambda_2)) + \alpha F^\omega(1/2).$$

This is equal to the true probability that a *nonpartisan* type plays an R action, but is strictly greater than the true probability of an R action, due to the failure to account for partisan types.

If $\lambda_1 = \lambda_2$, then $\hat{\psi}_1(R|\omega,\lambda_1,\lambda_2) = \hat{\psi}_2(R|\omega,\lambda_1,\lambda_2)$. Therefore, if the partisan and nonpartisan sociable types start with a common prior, both types update their likelihood ratio in the same way following each action, and after any history h_t, $\lambda_{1,t} = \lambda_{2,t}$. Although these types have different models of the world, their models collapse to the same subjective probability of each action. For different reasons, they both update too much towards state L following L actions and update too little towards state R following R actions. This means that we can consider the partisan and nonpartisan sociable types as a single type to characterize asymptotic learning.[36] It also rules out the possibility of disagreement or mixed learning, since the two likelihood ratios move in unison.

Asymptotic Learning Characterization. When partisan bias slants towards the incorrect state (i.e. when $\omega = R$), then the learning outcome depends on the severity of the partisan bias. Given each type's model of actions derived above, $\hat{\psi}_1$ and $\hat{\psi}_2$, both types overweight L actions and underweight R actions. If a large share of agents have

[36] It does not imply that a partisan and nonpartisan type with belief λ and private signal s will choose the same action, as they have different private signal cut-offs.

partisan bias (i.e. q close to one) or partisan bias is severe (i.e. ν close to zero), then this misspecification pulls the beliefs of both types towards state L and they almost surely learn the incorrect state. If partisan bias is not severe (i.e. ν close to one) or few agents have partisan bias (i.e. q close to zero), overweighting L actions is not significant enough to interfere with learning and both types learn the correct state. For intermediate values of ν and q, learning is cyclical. Agents believe L actions are not very informative when beliefs are close to state L, as most agents are choosing L for a large range of signals. This prevents incorrect learning. But these agents also believe R actions are not very informative when beliefs are close to state R, and therefore, L actions pull beliefs away from state R and prevent correct learning.

When partisan bias slants towards the correct state (i.e. when $\omega = L$), then learning is complete regardless of the frequency of partisan types or their level of bias. Partisan bias simply speeds up the rate at which beliefs converge. Proposition 3 formalizes these results.

Proposition 3 (Partisan Bias)**.** *When $\omega = R$, there exists an $\overline{q} \in (0,1)$ such that for $q > \overline{q}$, there exist unique cutoffs $0 < \nu_1(q) < \nu_2(q) < 1$ such that:*

1. *If $\nu > \nu_2(q)$, then almost surely learning is correct, $\Lambda(R) = \{(0,0)\}$.*
2. *If $\nu \in (\nu_1(q), \nu_2(q))$, then almost surely learning is cyclical, $\Lambda(R) = \emptyset$.*
3. *If $\nu < \nu_1(q)$, then almost surely learning is incorrect, $\Lambda(R) = \{(\infty, \infty)\}$.*

and there exists a $\underline{q} < \overline{q}$ such that for $q < \underline{q}$, almost surely learning is correct. When $\omega = L$, almost surely learning is correct, $\Lambda(L) = \{(0,0)\}$.

Figure 3 illustrates how the asymptotic learning outcomes depend on the frequency q of partisan types and the degree of their bias ν. Proposition 3 and Figure 3 also illustrate the robustness of the correctly specified model (Theorem 3), in which $q = 0$ and $\nu = 1$: for (q, ν) close enough to $(0, 1)$, learning is complete. The size of the robust region is quite large: when the degree of partisan bias is small, then correct learning obtains even if all agents have partisan bias ($q = 1$), and when the share of partisan types is small, then correct learning obtains even if these partisan types have a very severe bias ($\nu \approx 0$).

5.3 Payoff Misspecification: Social Perception Biases

Research on social perception has documented settings in which individuals overestimate the population prevalence of their preferences, opinions or behaviors – that is, they perceive a *false consensus* (Ross et al. 1977). False consensus effects are generally

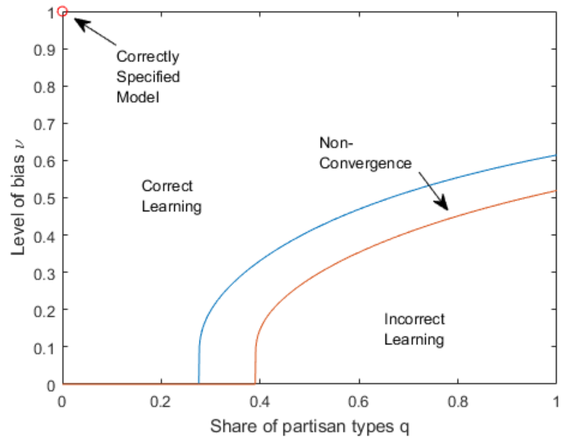

FIGURE 3. Partisan Bias Learning Outcomes
($\omega = R$, $\alpha = .1$, $F^L(s) = s^2$, $F^R(s) = 2s - s^2$)

found for non-normative behaviors. For example, adolescents have been found to exhibit a false consensus effect for estimating peers' smoking choices (Sherman, Presson, Chassin, Corty, and Olshavsky 1983), the prevalence of excessive drinking (Suls, Wan, and Sanders 1988) and peer sexual activity Whitley (1998).

In other social settings, individuals perceive a discrepancy between their preferences and behavior, and the preferences and behavior of others – that is, they exhibit *pluralistic ignorance*. For example, pluralistic ignorance has been documented with respect to perceptions of gender stereotypes (Prentice and Miller 1996), the extent of others' social inhibition (people underestimate it, relative to their own inhibition) and the inclination of others to choose a beneficial action that may have embarrassing consequences (people overestimate it, relative to their own inclination) (Miller and McFarland 1987). In contrast to the false consensus, pluralistic ignorance often arises in contexts where there is widespread behavioral adherence to a social norm, or where individuals believe that a *negative* trait affects their own behavior but not others' behavior.

We show that when agents systematically overestimate the similarity between their own preferences and the preferences of others – exhibiting the *false consensus effect* – this can lead to incorrect learning; when agents systematically underestimate this similarity – exhibiting *pluralistic ignorance* – this can prevent beliefs from converging. We model these social perception biases as a form of misspecification about the preferences of other agents. We utilize the outcomes framework discussed in Section 4.4 to characterize how the severity of social misperception affects asymptotic learning when individuals learn from the outcomes of others.[37]

[37]See Appendix B for the formal presentation of the learning from outcomes framework.

Types Framework. Suppose agents choose between engaging in a safe or risky behavior, denoted by actions a_1 or a_2, respectively. An agent's action affects her payoff-relevant outcome: she either fails or passes, denoted by outcomes x_1 or x_2, respectively. The risky behavior is enjoyable, but increases the probability of failure. If the agent chooses the safe action, then she passes with probability one, $Pr(x_2|a_2) = 1$. If the agent chooses the risky action, then she passes with probability $q^\omega \in (0, 1)$ in state ω, i.e. $Pr(x_2|a_1, \omega) = q^\omega$. The probability of passing when choosing the risky action is higher in state L than in state R, $q^L > q^R$. For example, a college student decides whether to study or go to a party. Partying is fun, but decreases the likelihood that she passes her classes. The student is uncertain about the extent to which partying decreases her probability of passing.

An agent derives utility from choosing the risky action and from passing. There are two types of agents, $\Theta = \{\theta_1, \theta_2\}$, and these types differ in the intensity of their preference for the risky action. An agent of type θ_i receives payoff $\upsilon_i(a, x) = \upsilon_i \mathbb{1}_{a=a_1} + \mathbb{1}_{x=x_2}$, where $\upsilon_i > 0$ is the utility derived from the risky action. Type θ_2 has a stronger preference for the risky action, $\upsilon_2 > \upsilon_1$. Given the distributions over outcomes and fixing state ω, type θ_i receives expected utility $u_i(a_1, \omega) = \upsilon_i + q^\omega$ when she chooses the risky action and $u_i(a_2, \omega) = 1$ when she chooses the safe action. Assume that the risky action is dominant for type θ_2, $\upsilon_2 + q^R > 1$, and type θ_1 prefers the risky action in state L and the safe action in state R, $\upsilon_1 + q^L > 1 > \upsilon_1 + q^R$.

Before making a decision, an agent learns about the state by observing the outcomes of her peers. We assume that there are no private or public signals, and agents do not observe action choices – they learn solely from the outcome history. The distribution of an agent's outcomes depends on her action choice. Therefore, in order to learn from prior outcomes, an agent needs a model of how agents before her chose their actions. Type θ_2's model is irrelevant, as this type has a dominant action. Let $\hat{\pi}_1(\theta_1) \in [0, 1)$ denote type θ_1's belief about the frequency of type θ_1, and let $\pi(\theta_1) \in (0, 1)$ denote the true frequency of type θ_1 in the population. We can model type θ_1's social perception bias as a misspecified type distribution in which the agent has a misspecified model of others' preferences.

Definition 7 (Social Perception Bias)**.** *Type θ_1 exhibits* pluralistic ignorance *when $\hat{\pi}_1(\theta_1) < \pi(\theta_1)$, and exhibits the* false consensus effect *when $\hat{\pi}_1(\theta_1) > \pi(\theta_1)$.*

For example, the college student is unsure if her peers have the same preferences for studying versus partying. She exhibits pluralist ignorance if she overestimates the likelihood that others derive more pleasure from partying, and she exhibits the false consensus effect if she overestimates the likelihood that others share her preference for partying. Other examples include a patient who is unsure about other patients' willingness to pay for health insurance or a parent who is unsure if other parents have similar preferences for vaccinating their children.

Type θ_1 also needs a model of how actions influence outcomes. To focus on social perception biases, we assume that θ_1 has a correct model of the outcome distribution conditional on each action and state, i.e. θ_1 correctly believes that the risky action leads to the pass outcome with probability q^ω in state ω. Therefore, θ_1's only source of misspecification is with respect to the type distribution.

To close the model, assume that a positive share of agents are type θ_2, $\pi(\theta_2) > 0$, and all types have common prior $p_0 = 1/2$. Type θ_1 has a correctly specified model of the outcome distribution, conditional on the action and state. Therefore, trivially, the subjective outcome distributions are aligned (Assumption 1*).[38] Only one action (the risky action) is informative, so trivially, outcomes are aligned (Assumption 2*). Further, the risky action has full support over the outcome space (Assumption 5). Since θ_2 chooses an action with full support over the outcome distribution and $\pi(\theta_2) > 0$, adequate information arrives to learn the state. Type θ_1 believes that θ_2 are present with positive probability, $\hat{\pi}_1(\theta_2) > 0$, so θ_1's model is consistent (Assumption 3*).

Action Choices and Beliefs. We focus on the learning of type θ_1. Since θ_2 has a dominant action, its beliefs do not influence its action choice. Therefore, the informational content of its outcomes are independent of its beliefs and we do not need to keep track of λ_2. Therefore, $\boldsymbol{\lambda} = (\lambda_1)$.

In order to learn from outcomes, an agent needs to infer the equilibrium action that each type chooses. Type θ_2 always chooses the risky action and passes with probability q^ω. The optimal action of type θ_1 depends on its current belief about the state. If beliefs λ_1 are such that type θ_1 chooses the safe action, then it almost surely passes and the probability of observing a pass outcome is

$$\psi_x(x_2|\lambda_1, \omega) = \pi(\theta_1) + (1 - \pi(\theta_1))q^\omega,$$

where ψ_x denotes the analogue of ψ for outcomes, while type θ_1's subjective probability of observing a pass outcome is

$$\hat{\psi}_{x,1}(x_2|\lambda_1, \omega) = \hat{\pi}_1(\theta_1) + (1 - \hat{\pi}_1(\theta_1))q^\omega.$$

Therefore, if θ_1 has the false consensus effect, it overestimates the probability of passing, and if θ_1 has pluralistic ignorance, it underestimates the probability of passing. If beliefs λ_1 are such that both types choose the risky action, then the true probability of passing

[38] These assumptions are outlined in the learning from outcomes framework in Appendix B.

is q^ω for both types, $\psi_x(x_2|\lambda_1, \omega) = q^\omega$. In this case, type θ_1's subjective probability of observing a pass outcome is correct $\hat{\psi}_{x,1}(x_2|\lambda_1, \omega) = q^\omega$.

Asymptotic Learning Characterization. We next characterize how the asymptotic learning outcomes for type θ_1 depend on its social perception bias.

First consider pluralistic ignorance. When $\hat{\pi}_1(\theta_1)$ is less than $\pi(\theta_1)$, type θ_1 displays a form of pluralistic ignorance in which it overestimates how many other agents have a dominant preference for the risky action. When θ_1 is almost certain that the safe action is optimal (state R), it overestimates the share of agents choosing the risky action. This leads it to observe outcomes that are on average better than it expects in state R, as it attributes passes from agents choosing the safe action to passes from agents choosing the risky action. This pushes its beliefs away from state R. In state L, this prevents incorrect learning, while in state R, this rules out correct learning when the bias is severe enough. In contrast, when θ_1 is almost certain that the risky action is optimal (state L), it correctly interprets outcomes, since both types choose the risky action. This allows correct learning when the state is L and prevents incorrect learning when the state is R. Proposition 4 summarizes this result.

Proposition 4 (Pluralistic Ignorance). *Suppose $\hat{\pi}_1(\theta_1) < \pi(\theta_1)$.*

1. *When $\omega = L$, almost surely learning is correct and θ_1 converges to choosing the risky action.*

2. *When $\omega = R$, there exists a cutoff $\overline{\pi}^R \in [0, \pi(\theta_1))$ such that if $\hat{\pi}_1(\theta_1) < \overline{\pi}^R$, then almost surely learning is cyclical, and otherwise, almost surely learning is correct and θ_1 converges to choosing the safe action. The cutoff $\overline{\pi}^R$ is increasing in $\pi(\theta_1)$.*

Proposition 4 shows that if the degree of pluralistic ignorance is severe enough, then type θ_1 will never learn to take the safe action when it is optimal, but it also won't herd on the risky action – it will choose both the safe and the risky action infinitely often. Behavior will cycle between the two actions. When a lot of agents choose the risky action, a high frequency of negative outcomes convinces agents that they need to make a safer choice. But in periods when agents are choosing the safe action, agents underestimate the negative consequences of the risky action. In contrast, θ_1 will learn to take the risky action when it is optimal.

To illustrate the intuition for this result, suppose a college student overestimates the share of students who enjoy drinking heavily despite the risk of failing. Then she believes many of her peers are partying, when in fact they are studying. If this student only observes how well her peers perform in class, she will attribute the high frequency of passes to students who drink and succeed. This causes her to believe that drinking is not very risky, and she will choose to drink inefficiently often in the state where drinking does in fact significantly reduce the probability of passing. But she won't converge to the incorrect belief that drinking has a minor impact on the probability of passing, as the failure rate becomes too high as beliefs approach this incorrect state.

Information programs designed to discourage risky behavior by providing information about the outcomes of choosing a risky action, such as "just say no" programs (e.g. DARE), will be ineffective in the presence of pluralistic ignorance. The outcomes of agents engaging in efficient, safe behavior will be misperceived as evidence that the risky behavior is actually safe. This prevents agents from learning the negative consequences of high risk behaviors. When agents have pluralistic ignorance, effective interventions require information about the *choices* of others, rather than information about the outcomes of these choices.

Next consider the false consensus effect. When $\hat{\pi}_1(\theta_1)$ is greater than $\pi(\theta_1)$, type θ_1 displays a form of the false consensus effect in which it underestimates how many other agents have a dominant preference for the risky action. When type θ_1 is almost certain that the safe action is optimal (state R), it underestimates the frequency of the risky action and observes a higher than anticipated failure rate. This reinforces choosing the safe action and move beliefs towards state R. It facilitates correct learning if the state is R, and the bias is severe enough, allows incorrect learning if the state is L. In contrast, when θ_1 is almost certain that the risky action is optimal (state L), it correctly interprets outcomes, since both types choose the risky action. This allows correct learning when the state is L and rules out incorrect learning when the state is R. Proposition 5 shows that the false consensus effect can lead to both correct and incorrect learning when the risky action is optimal. In contrast, θ_1 will learn to take the safe action when it is optimal.

Proposition 5 (False Consensus Effect)**.** *Suppose* $\hat{\pi}_1(\theta_1) > \pi(\theta_1)$.

1. *When* $\omega = R$*, almost surely learning is correct and* θ_1 *converges to choosing the safe action.*

2. *When* $\omega = L$*, there exists a cutoff* $\overline{\pi}^L \in (\pi(\theta_1), 1)$ *such that if* $\hat{\pi}_1(\theta_1) > \overline{\pi}^L$*, then correct and incorrect learning arise with positive probability, and otherwise, almost surely learning is correct and* θ_1 *converges to choosing the risky action. The cutoff* $\overline{\pi}^L$ *is increasing in* $\pi(\theta_1)$.

This result suggests that the false consensus effect will stymie information programs that encourage risk-taking, such as a campaign to encourage investing savings in the stock market. As more agents start to choose the risky action, if a type does not account for this change, the observed higher failure rate will reinforce its choice of the safe action. Policies that subsidize risk-taking will effectively mitigate the false consensus effect for the agents who receive subsidies, but will have a perverse effect on the unsubsidized agents.

6 Conclusion

We develop a general framework for social learning with model misspecification. Agents learn from the actions or outcomes of others, and may have misspecified models of how to interpret signals, how other agents learn, and/or other agents' preferences. When agents are misspecified, complete learning – where individuals eventually place probability one on the correct state – is no longer guaranteed. Our main result characterizes how asymptotic learning outcomes depend on the form of misspecification. We show that asymptotic learning may be incorrect, individuals may perpetually disagree, or beliefs may not converge at all. This characterization also establishes a robustness property: regardless of the form of misspecification, agents almost surely learn the correct state when they have approximately correct models.

We use this characterization to illustrate how model misspecification impacts long-run learning in three applications: a model of strategic misspecification in which agents use the level-k/cognitive hierarchy framework to learn from others, a model of signal misspecification in which some agents slant information towards a favored state, and a model of preference misspecification where agents exhibit social perception bias and are misspecified about the preferences of other agents. These results yield new insights about how misspecification impacts social learning, and provide a unified framework to study forms of misspecification that have been previously studied.

References

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

Ali, N. (2018): "On the Role of Responsiveness in Rational Herds," *Economic Letters*, 163, 79–82.

Angrisani, M., A. Guarino, P. Jehiel, and T. Kitagawa (2018): "Information redundancy neglect versus overconfidence: a social learning experiment," *mimeo.*

Banerjee, A. V. (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.

Bénabou, R. and J. Tirole (2011): "Identity, morals, and taboos: Beliefs as assets," *The Quarterly Journal of Economics*, 126, 805–855.

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.

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

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. A., V. P. Crawford, and N. Iriberri (2009): "Comparing Models of Strategic Thinking in Van Huyck, Battalio, and Beil's Coordination Games," *Journal of the European Economic Association*, 7, 365–376.

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

Darley, J. M. and P. H. Gross (1983): "A hypothesis-confirming bias in labeling effects." *Journal of Personality and Social Psychology*, 44, 20.

Enke, B. and F. Zimmermann (Forthcoming): "Correlation neglect in belief formation," *The Review of Economic Studies.*

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.

——— (2017): "Equilibrium in Misspecified Markov Decision Processes," *Mimeo.*

Eyster, E. and M. Rabin (2010): "Naïve Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2, 221–43.

Eyster, E. and G. Weizsacker (2011): "Correlation neglect in financial decision-making," .

Frick, M., R. Iijima, and Y. Ishii (2018): "Dispersed Behavior and Perceptions in Assortative Societies," *mimeo.*

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

Gagnon-Bartsch, T. (2017): "Taste Projection in Models of Social Learning," *Mimeo.*

Gilovich, T. (1990): "Differential construal and the false consensus effect," *Journal of personality and social psychology*, 59, 623.

Gottlieb, D. (2015): "Will You Never Learn? Self Deception and Biases in Information Processing," *Mimeo.*

Grebe, T., J. Schmid, and A. Stiehler (2008): "Do individuals recognize cascade behavior of others? – An experimental study," *Journal of Economic Psychology*, 29, 197 – 209.

Guarino, A. and P. Jehiel (2013): "Social Learning with Coarse Inference," *American Economic Journal: Microeconomics*, 5, 147–74.

Heidhues, P., B. Koszegi, and P. Strack (2018): "Unrealistic Expectations and Misguided Learning," *Econometrica*, forthcoming.

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.

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

Kallir, I. and D. Sonsino (2009): "The Neglect of Correlation in Allocation Decisions," *Southern Economic Journal*, 75, 1045–1066.

Kleijn, B. J. and A. W. van der Vaart (2006): "Misspecification in infinite-dimensional Bayesian statistics," *The Annals of Statistics*, 837–877.

Köszegi, 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.

——— (2005): "Are longer cascades more stable?" *Journal of the European Economic Association*, 3, 330–339.

Kunda, Z. (1990): "The case for motivated reasoning." *Psychological bulletin*, 108, 480.

Lee, I. H. (1993): "On the Convergence of Infomational Cascades," *Journal of Economic Theory*, 61, 395–411.

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

Lord, C. G., L. Ross, and M. R. Lepper (1979): "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence." *Journal of personality and social psychology*, 37, 2098.

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.

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., AlirezaTahbaz-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.

Nyarko, Y. (1991): "Learning in mis-specified 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, 2410–36.

Ortoleva, P. and E. Snowberg (2015): "Overconfidence in political behavior," *American Economic Review*, 105, 504–35.

Penczynski, S. P. (2017): "The nature of social learning: Experimental evidence," *European Economic Review*, 94, 148–165.

Plous, S. (1991): "Biases in the assimilation of technological breakdowns: Do accidents make us safer?" *Journal of Applied Social Psychology*, 21, 1058–1082.

Prentice, D. A. and D. T. Miller (1996): "Pluralistic Ignorance and the Perpetuation of Social Norms by Unwitting Actors," 28, 161 – 209.

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.

Sherman, S. J., C. C. Presson, L. Chassin, E. Corty, and R. Olshavsky (1983): "The False Consensus Effect in Estimates of Smoking Prevalence: Underlying Mechanisms," *Personality and Social Psychology Bulletin*, 9, 197–207.

Smith, L. and P. Sorensen (2000): "Pathological Outcomes of Observational Learning," *Econometrica*, 68, 371–398.

Suls, J., C. K. Wan, and G. S. Sanders (1988): "False consensus and false uniqueness in estimating the prevalence of health-protective behaviors," *Journal of Applied Social Psychology*, 18, 66–79.

Ungeheuer, M. and M. Weber (2017): "The Perception of Dependence, Investment Decisions, and Stock Prices," *Mimeo*.

Whitley, B. E. (1998): "False Consensus on Sexual Behavior among College Women: Comparison of Four Theoretical Explanations," *The Journal of Sex Research*, 35, 206–214.

Wilson, A. (2014): "Bounded Memory and Biases in Information Processing," *Econometrica*, 82, 2257–2294.

A Appendix: Proofs

A.1 Proof of Theorem 1

We establish Theorem 1 through a series of lemmas. In Lemma 1, we characterize the set of stationary beliefs, which are candidate limit points of $\langle \boldsymbol{\lambda}_t \rangle$. Lemma 2 rules out convergence to non-stationary beliefs. Next, Lemma 3 establishes when a stationary belief is locally stable. Lemma 4 establishes that global stability immediately follows from local stability for agreement outcomes, while Lemma 5 establishes that maximal accessibility is a sufficient condition for global stability of disagreement outcomes. In Lemma 6, we show that locally stable mixed learning outcomes must be in $\Lambda_M(\omega)$. Therefore, if $\Lambda_M(\omega)$ is empty, almost surely mixed learning does not arise. Finally, Lemma 7 establishes that when there is at least one globally stable stationary outcome and no locally stable mixed outcomes, the likelihood ratio converges almost surely for all sociable types.

We present Lemmas 1 - 4 for an arbitrary number of sociable types $k \geq 1$, as the constructions of local stability and the global stability of agreement outcomes are identical for $k \leq 2$ and $k > 2$. Establishing the global stability of disagreement outcomes and ruling out mixed learning is more involved for more than two sociable types, as the number of possible outcomes increases with k. Therefore, we present Lemmas 5 - 7 for $k \leq 2$ types, and present the analogues for $k > 2$ in Appendix A.2.

Throughout this section, assume Assumptions 1, 2, 3 and 4. Given $\varepsilon > 0$, define a neighborhood $B_\varepsilon(\boldsymbol{\lambda})$ of $\boldsymbol{\lambda} \in \{0, \infty\}^k$ as $\lambda_i \in [0, \varepsilon)$ if $\lambda_i = 0$ and $\lambda_i \in (1/\varepsilon, \infty]$ if $\lambda_i = \infty$.

A.1.1 Statement of Lemmas

In this section, we state Lemmas 1-7 outlined above. The proofs follow in Appendix A.1.2. At a stationary belief, the likelihood ratio remains constant for any action and signal pair that occurs with positive probability.

Definition 8 (Stationary). *$\boldsymbol{\lambda}^* \in [0, \infty]^k$ is* stationary *if for all $(a, \sigma) \in \mathcal{A} \times \Sigma$, either (i) $\psi(a, \sigma|\omega, \boldsymbol{\lambda}^*) = 0$ or (ii) $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^* \left(\frac{\hat{\psi}_i(a,\sigma|L,\boldsymbol{\lambda}^*)}{\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda}^*)} \right)$ for all $\theta_i \in \Theta_S$.*

By Assumption 3, actions and/or public signals are informative at any interior belief. Therefore, the set of stationary beliefs correspond to each type placing probability one on either state R ($\lambda = 0$) or state L ($\lambda = \infty$).

Lemma 1 (Stationary Beliefs). *The set of stationary beliefs are $\{0, \infty\}^k$.*

Further, the likelihood ratio almost surely does not converge to non-stationary beliefs.

Lemma 2 (Non-Stationary Beliefs). *If $\boldsymbol{\lambda}^* \in (0, \infty)^k$, then $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$.*

Therefore, if the likelihood ratio converges for all types, then it must converge to a stationary belief $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$.

Next, we determine when the likelihood ratio converges with positive probability. Recall that $\boldsymbol{\lambda}^*$ is *locally stable* if the process $\langle \boldsymbol{\lambda}_t \rangle$ converges to $\boldsymbol{\lambda}^*$ with positive probability from a neighborhood of $\boldsymbol{\lambda}^*$, and that $\gamma_i(\boldsymbol{\lambda}, \omega)$ is the expected change in the log likelihood ratio for type θ_i at belief $\boldsymbol{\lambda}$. Lemma 3 establishes the relationship between the local stability of stationary belief $\boldsymbol{\lambda}^*$ and the sign of $\gamma_i(\boldsymbol{\lambda}^*, \omega)$.

Lemma 3 (Locally Stable Beliefs)**.** *Let $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ be a stationary belief.*

1. *If $\gamma_i(\boldsymbol{\lambda}^*, \omega) < 0$ for all $\theta_i \in \Theta_S$ such that $\lambda_i^* = 0$ and $\gamma_i(\boldsymbol{\lambda}^*, \omega) > 0$ for all $\theta_i \in \Theta_S$ such that $\lambda_i^* = \infty$, then $\boldsymbol{\lambda}^*$ is locally stable.*
2. *If there exists a $\theta_i \in \Theta_S$ such that $\lambda_i^* = 0$ and $\gamma_i(\boldsymbol{\lambda}^*, \omega) > 0$ or $\lambda_i^* = \infty$ and $\gamma_i(\boldsymbol{\lambda}^*, \omega) < 0$, then $\boldsymbol{\lambda}^*$ is not locally stable and $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$.*

Lemma 3 uses results on the local stability of nonlinear equations developed in Smith and Sorensen (2000) (Theorems C.1 and C.2). Given Lemma 3, the set $\Lambda(\omega)$ defined in (9) is generically the set of locally stable beliefs. If there are no locally stable beliefs, i.e. $\Lambda(\omega)$ is empty, then the likelihood ratio almost surely does not converge for at least one type, as Lemma 3 rules out convergence to stationary beliefs that are not in $\Lambda(\omega)$, and Lemma 2 results out convergence to non-stationary beliefs.

We are interested in determining whether convergence occurs with positive probability from any initial value of the likelihood ratio, i.e. global stability. Clearly, the set of globally stable learning outcomes is a subset of the set of locally stable learning outcomes. Therefore, it remains to establish when local stability implies global stability. For agreement outcomes, $\boldsymbol{\lambda}^* \in \{0^k, \infty^k\}$, global stability immediately follows from local stability.

Lemma 4 (Global Stability of Agreement)**.** *For $\boldsymbol{\lambda}^* \in \{0^k, \infty^k\}$, if $\boldsymbol{\lambda}^*$ is locally stable, then $\boldsymbol{\lambda}^*$ is globally stable, i.e. for any initial belief $\boldsymbol{\lambda}_1 \in (0, \infty)^k$, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) > 0$.*

All types update their beliefs in the same direction following either the maximal action and signal in favor of state L, (a_M, σ_L), or the maximal action and signal in favor of state R, (a_1, σ_R). Therefore, from any initial belief, it is possible construct a finite sequence of action and public signal pairs that occurs with positive probability and pushes 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.

Lemma 5 establishes that maximal accessibility is a sufficient condition for the global stability of a disagreement outcome in the case of two sociable types, $k = 2$.[39]

[39] By definition, disagreement or mixed learning outcomes require $k \geq 2$.

Lemma 5 (Global Stability of Disagreement)**.** *Suppose* $k = 2$*. If disagreement outcome* $\boldsymbol{\lambda}^* \in \{(0, \infty), (\infty, 0)\}$ *is locally stable and maximally accessible, then* $\boldsymbol{\lambda}^*$ *is globally stable.*

Maximal accessibility orders the way each type interprets maximal actions and public signals, which guarantees that there exists a finite sequence of maximal actions and public signals that separates the beliefs of each type in the direction of the disagreement outcome.[40] As before, once the likelihood ratio is sufficiently close to the disagreement outcome, local stability establishes convergence.

As discussed in Section 4.2, a sufficient condition for ruling out mixed outcomes is that $\Lambda_M(\omega)$ is empty.

Lemma 6 (Unstable Mixed Outcomes)**.** *Suppose* $k = 2$*. If mixed learning outcome* $\lambda_i^* \notin \Lambda_M(\omega)$*, then* $Pr(\lambda_{i,t} \to \lambda_i^*$ *and* $\lambda_{-i,t}$ *does not converge*$) = 0$*.*

Finally, if there is at least one locally stable agreement or maximally accessible disagreement outcome, and no locally stable mixed outcomes, then the likelihood ratio converges almost surely for all types.

Lemma 7 (Belief Convergence)**.** *Suppose* $k = 2$*,* $\Lambda(\omega)$ *contains an agreement outcome or maximally accessible disagreement outcome and* $\Lambda_M(\omega)$ *is empty. Then for any initial belief* $\boldsymbol{\lambda}_1 \in (0, \infty)^2$*, there exists a random variable* $\boldsymbol{\lambda}_\infty$ *with* $\operatorname{supp}(\boldsymbol{\lambda}_\infty) = \Lambda(\omega)$ *such that* $\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}_\infty$ *almost surely.*

Theorem 1 immediately follows. Part (1) follows from the local and global stability of agreement outcomes (Lemmas 3 and 4). Part (2) follows from the local and global stability of disagreement outcomes (Lemmas 3 and 5). For part (3), Lemmas 1 and 2 rule out convergence to non-stationary beliefs, Lemma 3 rules out convergence to unstable stationary outcomes, and Lemma 6 rules out convergence to a mixed learning outcome when $\Lambda_M(\omega)$ is empty. Therefore, if $\Lambda(\omega)$ is empty, there are no locally stable learning outcomes and almost surely the likelihood ratio does not converge for at least one sociable type, establishing the second statement in part (3). If $\Lambda_M(\omega)$ is also empty, then almost surely the likelihood ratio does not converge for any sociable type, establishing the first statement in part (3). The final statement in part (3) follows from Lemma 7, which establishes convergence when $\Lambda(\omega)$ contains an agreement outcome or maximally accessible disagreement outcome.

[40] While maximal accessibility is simple and easy to verify, it can be restrictive, especially in models with large action or public signal spaces. In Lemma 9 (Appendix A.1.2), we establish a more general sufficient condition to separate beliefs, which we call separability (Definition 9).

A.1.2 Proofs of Lemmas 1 - 7

Proof of Lemma 1 (Stationary Beliefs). At a stationary belief $\boldsymbol{\lambda}^* \in [0,\infty]^k$,

$$\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^* \left(\frac{\hat{\psi}_i(a,\sigma|L,\boldsymbol{\lambda}^*)}{\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda}^*)} \right) \tag{14}$$

for all (a,σ) such that $\psi(a,\sigma|\omega,\boldsymbol{\lambda}^*) > 0$. Trivially, (14) is satisfied for all $\boldsymbol{\lambda}^* \in \{0,\infty\}^k$, independent of $\psi(a,\sigma|\omega,\boldsymbol{\lambda}^*)$. Therefore, all $\boldsymbol{\lambda}^* \in \{0,\infty\}^k$ are stationary. It remains to be determined whether there exist any interior stationary beliefs $\boldsymbol{\lambda}^* \in (0,\infty)^k$.

Suppose $\boldsymbol{\lambda}^* \in (0,\infty)^k$ and Assumption 3.ii holds, i.e. there exists an autarkic θ_j with $\pi(\theta_j) > 0$ that plays a_1 with probability in $(0,1)$, and each sociable type θ_i believes this autarkic type occurs with positive probability, $\hat{\pi}_i(\theta_j) > 0$. Then $\psi(a_1|\omega,\boldsymbol{\lambda}^*) \in (0,1)$ and $\hat{\psi}_i(a_1|\omega,\boldsymbol{\lambda}^*) \in (0,1)$ for $\omega \in \{L,R\}$. Further, $\hat{\psi}_i(a_1|L,\boldsymbol{\lambda}^*) < \hat{\psi}_i(a_1|R,\boldsymbol{\lambda}^*)$, since $\hat{F}_i^L < \hat{F}_i^R$ when a type plays an action with an interior probability. Given $\rho_i(\sigma_R) \leq 1/2$, this implies $\hat{\psi}_i(a_1,\sigma_R|L,\boldsymbol{\lambda}^*) < \hat{\psi}_i(a_1,\sigma_R|R,\boldsymbol{\lambda}^*)$ and (14) does not hold for (a_1,σ_R). But (a_1,σ_R) occurs with positive probability in either state, $\psi(a_1,\sigma_R|\omega,\boldsymbol{\lambda}^*) > 0$. Therefore, $\boldsymbol{\lambda}^*$ cannot be stationary.

Suppose $\boldsymbol{\lambda}^* \in (0,\infty)^k$ and Assumption 3.i holds. Then $\rho_i(\sigma_R) < 1/2$ and $\rho_i(\sigma_L) > 1/2$ for all sociable types θ_i. Further, σ_R and σ_L occur with positive probability, independent of $\boldsymbol{\lambda}$. At least one action a occurs with positive probability at $\boldsymbol{\lambda}^*$. Since public signals are informative, it must be that $\frac{\hat{\psi}_i(a,\sigma_L|L,\boldsymbol{\lambda}^*)}{\hat{\psi}_i(a,\sigma_L|R,\boldsymbol{\lambda}^*)} \neq \frac{\hat{\psi}_i(a,\sigma_R|L,\boldsymbol{\lambda}^*)}{\hat{\psi}_i(a,\sigma_R|R,\boldsymbol{\lambda}^*)}$. Therefore, (14) cannot hold for both (a,σ_L) and (a,σ_R). But both action-signal pairs occur with positive probability in either state, $\psi(a,\sigma_R|\omega,\boldsymbol{\lambda}^*) > 0$ and $\psi(a,\sigma_L|\omega,\boldsymbol{\lambda}^*) > 0$. Therefore, $\boldsymbol{\lambda}^*$ cannot be stationary. □

Proof of Lemma 2 (Non-Stationary Beliefs). Suppose beliefs converge to a non-stationary belief $\boldsymbol{\lambda}^* \in [0,\infty]^k \setminus \{0,\infty\}^k$ with positive probability. After action and public signal $(a_t,\sigma_t) = (a_M,\sigma_L)$, by Lemma 11, $\lambda_{i,t+1} - \lambda_{i,t}$ is bounded uniformly away from zero for all sociable types $\theta_i \in \Theta_S$. For sufficiently small $\varepsilon > 0$, if $\boldsymbol{\lambda}_t \in B_\varepsilon(\boldsymbol{\lambda}^*)$, then after observing $(a_t,\sigma_t) = (a_M,\sigma_L)$, $\lambda_{i,t+1} \notin B_\varepsilon(\boldsymbol{\lambda}^*)$ for any type with an interior belief $\lambda_{i,t} \in (0,\infty)$. The probability $Pr(\exists t < T|(a_t,\sigma_t) = (a_M,\sigma_L))$ converges to one as $T \to \infty$. Therefore, the likelihood ratio almost surely leaves $B_\varepsilon(\boldsymbol{\lambda}^*)$. □

Proof of Lemma 3 (Locally Stable Beliefs). Suppose $\omega = R$. The proof for $\omega = L$ is analogous.

Part 1. Consider $\boldsymbol{\lambda}^* = 0^k$ and suppose $\gamma_i(0^k,R) < 0$ for all sociable 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,

$$\sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi(a,\sigma|R,0^k) \sup_{\boldsymbol{\lambda}\in[0,\varepsilon]^k} \log \frac{\hat{\psi}_i(a,\sigma|L,\boldsymbol{\lambda})}{\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda})} < 0. \tag{15}$$

for all $\theta_i \in \Theta_S$. Let

$$g_i(a,\sigma) \equiv \sup_{\boldsymbol{\lambda}\in[0,\varepsilon]^k} \log \frac{\hat{\psi}_i(a,\sigma|L,\boldsymbol{\lambda})}{\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda})}$$

denote the maximal update from action and signal (a,σ) in the neighborhood $[0,\varepsilon]^k$, with $\mathbf{g}(a,\sigma) \equiv (g_1(a,\sigma),...,g_k(a,\sigma))$. Let

$$\bar{g}_i \equiv \max_{(a,\sigma)\in\mathcal{A}\times\Sigma} g_i(a,\sigma)$$

denote the maximal update across all action and signal pairs 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

$$\sup_{\boldsymbol{\lambda}\in[0,\varepsilon_\delta]^k} |\psi(a,\sigma|R,\boldsymbol{\lambda}) - \psi(a,\sigma|R,0^k)| < \delta.$$

By Lemma 12, $\psi(a,\sigma|R,\boldsymbol{\lambda})$ 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: $\boldsymbol{\lambda}_{\delta,1} = \boldsymbol{\lambda}_1$,

$$\log \boldsymbol{\lambda}_{\delta,t+1} = \log \boldsymbol{\lambda}_{\delta,t} + \mathbf{g}(a(\theta_t,s_t,0^k),\sigma_t),$$

when (θ_t,s_t) is such that $a(\theta_t,s_t,\boldsymbol{\lambda}) = a(\theta_t,s_t,0^k)$ 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 $\omega = R$, let $\psi_\delta(a,\sigma)$ be the probability of (a,σ) in the former event and let $\bar{\psi}_\delta$ be the probability of the latter event. Note $\psi_\delta(a,\sigma) \leq \inf_{\boldsymbol{\lambda}\in[0,\varepsilon_\delta]^k} \psi(a,\sigma|R,\boldsymbol{\lambda})$ and $\bar{\psi}_\delta + \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi_\delta(a,\sigma|R) = 1$. By Theorem C.1 of Smith and Sorensen (2000), if

$$\bar{\psi}_\delta \bar{g}_i + \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi_\delta(a,\sigma) g_i(a,\sigma) < 0 \tag{16}$$

for all $\theta_i \in \Theta_S$, then $\lim_{t\to\infty} \boldsymbol{\lambda}_{\delta,t} = 0^k$. Equation (16) holds for sufficiently small δ, since by (15), it is strictly less than zero at $\delta = 0$.

Let $\delta_1 > 0$ denote an upper bound such that (16) holds for all $\delta < \delta_1$. Whenever (θ_t, s_t) is such that $a(\theta_t, s_t, \boldsymbol{\lambda}) = a(\theta_t, s_t, 0^k)$ for all $\boldsymbol{\lambda} \in [0, \varepsilon_\delta]$, the process $\langle \log \boldsymbol{\lambda}_{\delta,t} \rangle$ updates by $\mathbf{g}(a, \sigma)$. 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 $\log \frac{\hat{\psi}_i(a,\sigma|L,\boldsymbol{\lambda}_t)}{\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda}_t)}$ 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 $\boldsymbol{\lambda}_{\delta,t} \geq \boldsymbol{\lambda}_t$ and $\boldsymbol{\lambda}_{\delta,t} \in [0, \varepsilon_\delta]^k$, then $\boldsymbol{\lambda}_{\delta,t+1} \geq \boldsymbol{\lambda}_{t+1}$. 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 $\delta < \delta_1$,

$$Pr(\cup_t \cap_{s \geq t} \{\boldsymbol{\lambda}_{\delta,s} \in [0, \varepsilon_\delta]^k\}) = 1.$$

Therefore, there exists a $t \geq 1$ such that $Pr(\forall s \geq t, \boldsymbol{\lambda}_{\delta,s} \in [0, \varepsilon_\delta]^k) > 0$. 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 $\boldsymbol{\lambda}_t \leq \boldsymbol{\lambda}_{\delta,t}$. Moreover, this holds for all $\boldsymbol{\lambda} \leq \boldsymbol{\lambda}_{\delta,1}$. 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 $\varepsilon^* = \inf \lambda_{\delta,i,1}$ 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 are analogous. If $\lambda_i^* = \infty$, substitute λ_i^{-1} for type θ_i and modify the transition rules accordingly.

Part 2. Let $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ be a stationary belief and suppose that there exists a type, which without loss of generality we denote θ_1, such that $\lambda_1^* = 0$ but $\gamma_1(\boldsymbol{\lambda}^*, R) > 0$. 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 $\gamma_1(\boldsymbol{\lambda}^*, R) > 0$, there exists a $\varepsilon > 0$ such that for neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*) \equiv [0, \varepsilon]^\kappa \times [1/\varepsilon, \infty]^{k-\kappa}$ of $\boldsymbol{\lambda}^*$,

$$\sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi(a, \sigma|R, \boldsymbol{\lambda}^*) \inf_{\boldsymbol{\lambda}\in B_\varepsilon(\boldsymbol{\lambda}^*)} \log \frac{\hat{\psi}_1(a, \sigma|L, \boldsymbol{\lambda})}{\hat{\psi}_1(a, \sigma|R, \boldsymbol{\lambda})} > 0. \tag{17}$$

Let $\tau_\varepsilon \equiv \min\{\tau | \boldsymbol{\lambda}_t \in B_\varepsilon(\boldsymbol{\lambda}^*) \ \forall t \geq \tau\}$ 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

$$g_1(a, \sigma) \equiv \inf_{\boldsymbol{\lambda}\in B_\varepsilon(\boldsymbol{\lambda}^*)} \log \frac{\hat{\psi}_1(a, \sigma|L, \boldsymbol{\lambda})}{\hat{\psi}_1(a, \sigma|R, \boldsymbol{\lambda})}$$

denote the minimal update for type θ_1 following action and signal (a, σ) in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*)$ and let

$$\underline{g}_1 \equiv \min_{(a,\sigma)\in\mathcal{A}\times\Sigma} g_1(a,\sigma)$$

denote the minimal update across all action and signal pairs 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 $\langle\tilde{\lambda}_t\rangle$ as follows: let $\tilde{\lambda}_\tau = \lambda_{1,\tau}$ and for $t > \tau$,

$$\log \tilde{\lambda}_{t+1} = \log \tilde{\lambda}_t + g_1(a(\theta_t, s_t, \boldsymbol{\lambda}^*), \sigma_t)$$

when (θ_t, s_t) is such that $a(\theta_t, s_t, \boldsymbol{\lambda}) = a(\theta_t, s_t, \boldsymbol{\lambda}^*)$ 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. When $\omega = R$, let $\psi(a, \sigma)$ be the probability of (a, σ) in the former event and let $\underline{\psi}$ be the probability of the latter event. Note $\underline{\psi} + \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi(a, \sigma) = 1$. Choose ε sufficiently small so that

$$\underline{\psi}\,\underline{g}_1 + \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi(a,\sigma) g_1(a,\sigma) > 0. \tag{18}$$

Given (17), (18) 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 (18). By the Law of Large Numbers, almost surely, $\frac{1}{t}(\log \tilde{\lambda}_{t+1} - \log \tilde{\lambda}_t)$ converges to (18), 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 $\lambda_{1,t} \geq \tilde{\lambda}_t$ and $\boldsymbol{\lambda}_t \in B_\varepsilon(\boldsymbol{\lambda}^*)$, then $\lambda_{1,t+1} \geq \tilde{\lambda}_{1,t+1}$. 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 $t > \tau$

$$\lim_{t\to\infty} \log \lambda_{1,t} \geq \lim_{t\to\infty} \log \tilde{\lambda}_t \to \infty.$$

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 $\gamma_1(\boldsymbol{\lambda}^*, R) < 0$, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) = 0$. □

Locally Stable Neighborhoods. The general definition of $\Lambda_i(\omega)$ for $k \geq 1$ is

$$\Lambda_i(\omega) \equiv \{\boldsymbol{\lambda} \in \{0, \infty\}^2 | \gamma_i(\boldsymbol{\lambda}, \omega) < 0 \text{ if } \lambda_i = 0 \text{ and } \gamma_i(\boldsymbol{\lambda}, \omega) > 0 \text{ if } \lambda_i = \infty\}, \tag{19}$$

with $\Lambda(\omega) \equiv \cap_{i=1}^k \Lambda_i(\omega)$. From Lemma 3, if $\boldsymbol{\lambda}^* \in \Lambda(\omega)$, then $\boldsymbol{\lambda}^*$ is locally stable, i.e. there exists an $\varepsilon > 0$ and a *stable* neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*)$ such that when $\boldsymbol{\lambda}_1 \in B_\varepsilon(\boldsymbol{\lambda}^*)$, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) > 0$. Also, generically, for each stationary belief $\boldsymbol{\lambda}^* \notin \Lambda(\omega)$, there exists an $\varepsilon > 0$ and an *unstable* neighborhood $B_\varepsilon(\boldsymbol{\lambda}^*)$ such that when $\boldsymbol{\lambda}_1 \in B_\varepsilon(\boldsymbol{\lambda}^*)$, $\langle \boldsymbol{\lambda}_t \rangle$ almost surely leaves this neighborhood.

Fix state ω and define $E > 0$ as the smallest constant such that if $\log \lambda_i \in \mathbb{R} \backslash [-E, E]$ for each $\theta_i \in \Theta_S$, then $\boldsymbol{\lambda}$ is contained in one of these stable or unstable neighborhoods, and let $B_E(\boldsymbol{\lambda}^*)$ denote the corresponding neighborhood for each stationary $\boldsymbol{\lambda}^*$.[41] Let $\mathcal{B}$ denote the union of the stable neighborhoods, $\mathcal{B} \equiv \cup_{\boldsymbol{\lambda}^* \in \Lambda(\omega)} B_E(\boldsymbol{\lambda}^*)$, and let $\mathcal{B}_U$ denote the union of the unstable neighborhoods, $\mathcal{B}_U = \cup_{\boldsymbol{\lambda}^* \in \{0,\infty\}^k \backslash \Lambda(\omega)} B_E(\boldsymbol{\lambda}^*)$. We will use these neighborhoods in the proofs of Lemmas 4 - 7.

Proof of Lemma 4 (Global Stability of Agreement). Suppose the agreement outcome is locally stable, $0^k \in \Lambda(\omega)$, and there are at least two types, $|\Theta| \geq 2$. By Assumption 4, a_1 occurs with positive probability, and by Lemma 10, observing (a_1, σ_R) 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, σ_R) actions and signals required for the likelihood ratio to reach the stable neighborhood, $\boldsymbol{\lambda}_{N+1} \in B_E(0^k)$. By Lemma 11, the change in the likelihood ratio following (a_1, σ_R) is bounded away from zero. Therefore, $N < \infty$. Let τ_1 be the first time that $\langle \boldsymbol{\lambda}_t \rangle$ enters $B_E(0^k)$, $\tau_1 \equiv \min\{t | \boldsymbol{\lambda}_t \in B_E(0^k)\}$, let τ_2 be the first that $\langle \boldsymbol{\lambda}_t \rangle$ leaves $B_E(0^k)$ after entering, $\tau_2 \equiv \min\{t > \tau_1 | \boldsymbol{\lambda}_t \notin B_E(0^k)\}$, and let τ_3 be the first time the likelihood ratio enters $B_E(0^k)$ and never leaves, $\tau_3 \equiv \min\{\tau | \boldsymbol{\lambda}_t \in B_E(0^k) \ \forall t \geq \tau\}$. 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, σ_R) action and signal pairs. 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 Lemma 3, if the likelihood ratio remains in $B_E(0^k)$ for all t, beliefs almost surely converge to 0^k. Therefore, if

[41] 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. This simplifies notation in subsequent lemmas.

$0^k \in \Lambda(\omega)$, then from any initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^k$, $Pr(\boldsymbol{\lambda}_t \to 0^k) > 0$.

Suppose the agreement outcome is locally stable, $0^k \in \Lambda(\omega)$, and there is a single type, $|\Theta| = 1$. Then Assumption 3.i must hold and public signals are informative. With a single type, action a_1 may occur with probability zero at some beliefs, and we need to adapt the proof for multiple types. Let $\underline{a}(\lambda)$ be the lowest action that type θ_1 plays at belief λ. When there is a single type θ_1, this type has a correctly specified model of the type distribution (this must be the case when $|\Theta| = 1$, as trivially, $\hat{\pi}^1(\theta_1) = 1$), and therefore, observing $\underline{a}(\lambda)$ at belief λ weakly decreases the likelihood ratio (by similar reasoning to Lemma 10). Therefore, observing $(\underline{a}(\lambda), \sigma_R)$ strictly decreases the likelihood ratio, since public signals are informative. Substituting the sequence $(a(\lambda_t), \sigma_R)_{t=1}^N$ for the sequence of N consecutive (a_1, σ_R) actions and signals, where λ_t is the updated belief following $(a(\lambda_{t-1}), \sigma_R)$, the remainder of the proof is the same as in the multiple types case.

The proof for agreement outcome ∞^k is analogous. □

Intermediate Results for Lemma 5. The following definitions and lemmas are intermediate results for the proof of Lemma 5, and they hold for $k \geq 1$. Order the public signals by relative likelihood of state L, $(\sigma_1, \sigma_2, ..., \sigma_{|\Sigma|})$ (note $\sigma_1 = \sigma_R$ and $\sigma_{|\Sigma|} = \sigma_L$). Order the action and signal pairs $((a_1, \sigma_1), (a_2, \sigma_1), ..., (a_M, \sigma_{|\Sigma|}))$ so that pair $M(l-1)+m$ corresponds to action a_m and signal σ_l. Define $A(\boldsymbol{\lambda})$ as the matrix of updates to the log likelihood ratio at beliefs $\boldsymbol{\lambda}$, where each row corresponds to the updates for sociable type θ_i, and each column corresponds to the update following action and signal pair j,

$$(A(\boldsymbol{\lambda}))_{ij} \equiv \log \frac{\hat{\psi}_i((a,\sigma)_j | L, \boldsymbol{\lambda})}{\hat{\psi}_i((a,\sigma)_j | R, \boldsymbol{\lambda})}. \tag{20}$$

Without loss of generality, we consider disagreement outcomes that are ordered so that the first $\kappa \in \{1, ..., k-1\}$ types have belief 0 and the remaining $k - \kappa$ types have belief ∞, i.e. $\boldsymbol{\lambda}^* = (0^\kappa, \infty^{k-\kappa})$. To consider other disagreement outcomes, simply reorder the types so that this holds.

Definition 9 (Separability)**.**

1. *Given $\kappa \in \{1, ..., k\}$, stationary likelihood ratio $\boldsymbol{\lambda}^* = (0^\kappa, \infty^{k-\kappa})$ is* separable at zero *if there exist vectors $c \in [0,\infty)^{|\mathcal{A}\times\Sigma|}$ and $G \in \mathbb{R}^k$ with $G_i > 0$ for all $i \geq \kappa$ and $G_i < 0$ for all $i < \kappa$, such that $A(\boldsymbol{\lambda}^*)c = G$.*

2. *Given $\kappa \in \{0, ..., k-1\}$, stationary likelihood ratio $\boldsymbol{\lambda}^* = (0^\kappa, \infty^{k-\kappa})$ is* separable at infinity *if there exist vectors $c \in (0,\infty)^{|\mathcal{A}\times\Sigma|}$ and $G \in \mathbb{R}^k$ with $G_i > 0$ for all $i > \kappa + 1$ and $G_i < 0$ for all $i \leq \kappa + 1$, such that $A(\boldsymbol{\lambda}^*)c = G$.*

Definition 10 (Adjacently Accessible)**.** *Stationary likelihood ratio*

$$\boldsymbol{\lambda}_2^* \in \{(0^{\kappa-1}, \infty^{k-\kappa+1}), (0^{\kappa+1}, \infty^{k-\kappa-1})\}$$

is adjacently accessible *from stationary likelihood ratio* $\boldsymbol{\lambda}_1^* = (0^\kappa, \infty^{k-\kappa})$ *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* $Pr(\boldsymbol{\lambda}_{t+\tau(\boldsymbol{\lambda})} \in B_{\varepsilon_2}(\boldsymbol{\lambda}_2^*)) > 0$.

Lemma 8 (Adjacently Accessible)**.** *If* $\boldsymbol{\lambda}_1^* = (0^\kappa, \infty^{k-\kappa})$ *is separable at zero, then* $\boldsymbol{\lambda}_2^* = (0^{\kappa-1}, \infty^{k-\kappa+1})$ *is adjacently accessible from* $\boldsymbol{\lambda}_1^*$*, and if* $\boldsymbol{\lambda}_1^*$ *is separable at infinity, then* $\boldsymbol{\lambda}_2^* = (0^{\kappa+1}, \infty^{k-\kappa-1})$ *is adjacently accessible from* $\boldsymbol{\lambda}_1^*$*.*

Proof. 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. 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 $Pr(\boldsymbol{\lambda}_{1+\tau(\boldsymbol{\lambda})} \in B_{\varepsilon_2}(\boldsymbol{\lambda}_2^*)) > 0$. Since the log likelihood ratio process is linear, this also holds for any $\boldsymbol{\lambda}_t = \boldsymbol{\lambda}$.

For $\varepsilon > 0$, recall $B_\varepsilon(\boldsymbol{\lambda}_1^*) \equiv [0, \varepsilon)^\kappa \times (1/\varepsilon, \infty]^{k-\kappa}$ denotes a neighborhood of $\boldsymbol{\lambda}_1^*$. 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

$$g_{\varepsilon,i}(a, \sigma) \equiv \inf_{\boldsymbol{\lambda} \in B_\varepsilon(\boldsymbol{\lambda}_1^*)} \log \frac{\hat{\psi}_i(a, \sigma | L, \boldsymbol{\lambda})}{\hat{\psi}_i(a, \sigma | R, \boldsymbol{\lambda})},$$

as the smallest update to the log likelihood ratio when type $i \geq \kappa$ observes (a, σ) and has likelihood ratio in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}_1^*)$, and

$$g_{\varepsilon,i}(a, \sigma) \equiv \sup_{\boldsymbol{\lambda} \in B_\varepsilon(\boldsymbol{\lambda}_1^*)} \log \frac{\hat{\psi}_i(a, \sigma | L, \boldsymbol{\lambda})}{\hat{\psi}_i(a, \sigma | R, \boldsymbol{\lambda})}$$

as the largest update to the log likelihood ratio when type $i < \kappa$ observes (a, σ) and has likelihood ratio in the neighborhood $B_\varepsilon(\boldsymbol{\lambda}_1^*)$. Finally, define

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

as the largest update to the log likelihood ratio when type κ observes (a, σ) 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 separability at zero, 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 $c_{a,\sigma} \geq 0$ for each $(a,\sigma) \in \mathcal{A} \times \Sigma$ such that

$$G_i \equiv \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} c_{a,\sigma} g_{\varepsilon_3,i}(a,\sigma), \tag{21}$$

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

$$\overline{G}_\kappa \equiv \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} c_{a,\sigma} \bar{g}_{\varepsilon_3,\kappa}(a,\sigma). \tag{22}$$

Next we define processes $\xi_{i,t} \equiv \sum_{s=1}^{t-1} g_{\varepsilon_3,i}(a_s,\sigma_s)$ and $\bar{\xi}_{\kappa,t} \equiv \sum_{s=1}^{t-1} \bar{g}_{\varepsilon_3,\kappa}(a_s,\sigma_s)$. Given a sequence with $c_{a,\sigma}$ realizations of each (a,σ), at time $\tau_1 \equiv \sum_{\mathcal{A}\times\Sigma} c_{a,\sigma} + 1$, the process $\xi_{i,\tau_1} = G_i$ by (21) and $\bar{\xi}_{\kappa,\tau_1} = \overline{G}_\kappa$ by (22). 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 $c_{a,\sigma}$ realizations of each (a,σ), at time $\tau_K \equiv N_K \sum_{\mathcal{A}\times\Sigma} c_{a,\sigma} + 1$,

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 $K > \bar{K}$. 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 $c_{a,\sigma}$ realizations of each (a, σ) 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 $c_{a,\sigma}$ realizations of each (a, σ) 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, σ) 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, σ_L), $\log \lambda_{\kappa,t+N_2} > K(\varepsilon_2)$. Let K_2 be the most any type $i < \kappa$'s log likelihood ratio increases after N_2 realizations of (a_M, σ_L) 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, σ_L), $\log \lambda_{i,t} < -K(\varepsilon_2) - K + \bar{K} + K_2$ 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, σ_L), 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 and signals 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 $\boldsymbol{\lambda}_1^*$ separable at infinity is analogous. □

Definition 11 (Accessible). *A belief $\boldsymbol{\lambda}^*$ 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 $Pr(\boldsymbol{\lambda}_\tau \in B_\varepsilon(\boldsymbol{\lambda}^*)) > 0$.*

Lemma 9 (Accessible Disagreement)**.** *Suppose $k = 2$. If $(0,0)$ is separable at zero or (∞,∞) is separable at infinity, then disagreement outcome $(0,\infty)$ is accessible.*

Proof. By Lemma 8, if $(0,0)$ is separable at zero, then $(0,\infty)$ is adjacently accessible from $(0,0)$. Fix initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^2$ and choose $\varepsilon_2 > 0$. Choose $\varepsilon_1 > 0$ such that for any $\boldsymbol{\lambda} \in B_{\varepsilon_1}((0,0))$, there exists a $\tau_2(\boldsymbol{\lambda}) < \infty$ such that if $\boldsymbol{\lambda}_t = \boldsymbol{\lambda}$, then $Pr(\boldsymbol{\lambda}_{t+\tau_2(\boldsymbol{\lambda})} \in B_{\varepsilon_2}((0,\infty))) > 0$. By adjacent accessibility, such an ε_1 exists. By Lemma 4, there exists a finite sequence ξ_1 of N_1 action and signal pairs that occurs with positive probability, such that following ξ_1, $\boldsymbol{\lambda}_{N_1+1} \in B_{\varepsilon_1}((0,0))$. By adjacent accessibility, there exists a finite sequence ξ_2 of N_2 action and signal pairs that occurs with positive probability, such that following sequences ξ_1 and ξ_2, $\boldsymbol{\lambda}_{N_1+N_2+1} \in B_{\varepsilon_2}((0,\infty))$. Since these sequences occur with positive probability, $Pr(\boldsymbol{\lambda}_{N_1+N_2+1} \in B_{\varepsilon_2}((0,\infty))) > 0$, which is the definition of accessible. The case where (∞,∞) is separable at infinity is analogous. $\square$

Proof of Lemma 5 (Global Stability of Disagreement). Suppose $k = 2$, $(0,\infty) \in \Lambda(\omega)$ and $\theta_2 \succ_{(0,0)} \theta_1$. We first show that $\theta_2 \succ_{(0,0)} \theta_1$ implies that $(0,0)$ is separable at zero. Define the submatrix

$$A_{max} \equiv \begin{pmatrix} \log \frac{\hat{\psi}_2(a_1,\sigma_R|L,(0,0))}{\hat{\psi}_2(a_1,\sigma_R|R,(0,0))} & \log \frac{\hat{\psi}_2(a_M,\sigma_L|L,(0,0))}{\hat{\psi}_2(a_M,\sigma_L|R,(0,0))} \\ \log \frac{\hat{\psi}_1(a_1,\sigma_R|L,(0,0))}{\hat{\psi}_1(a_1,\sigma_R|R,(0,0))} & \log \frac{\hat{\psi}_1(a_M,\sigma_L|L,(0,0))}{\hat{\psi}_1(a_M,\sigma_L|R,(0,0))} \end{pmatrix}.$$

Since $\theta_2 \succ_{(0,0)} \theta_1$, this has a positive determinant. Therefore, there exists a $c \in \mathbb{R}^2_+$ that solves

$$A_{max} c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

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

$$A_{max} \tilde{c} = \begin{pmatrix} G_2 \\ G_1 \end{pmatrix},$$

where $G_1 < 0$ and $G_2 > 0$. Therefore, by Definition 9, $(0,0)$ is separable at zero, since we can set values of c_j to zero for the remaining action and signal pairs in matrix (20). Therefore, by Lemma 9, $(0,\infty)$ is accessible.

We will next show that for any initial belief, $Pr(\boldsymbol{\lambda}_t \to (0,\infty)) > 0$. Fix initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^2$ and choose $\varepsilon < e^{-E}$. By accessibility, there exists a finite sequence ξ of N action and signal pairs that occurs with positive probability, such that following sequences ξ, $\boldsymbol{\lambda}_{N+1} \in B_\varepsilon((0,\infty))$. From $(0,\infty) \in \Lambda(\omega)$, $Pr(\boldsymbol{\lambda}_t \to (0,\infty)|\xi) > 0$. Therefore, from any initial belief $\boldsymbol{\lambda}_1 \in (0,\infty)^2$, $Pr(\boldsymbol{\lambda}_t \to (0,\infty)) > 0$, which implies that $(0,\infty)$ is globally stable. The case where $\theta_2 \succ_{(\infty,\infty)} \theta_1$ is analogous, as is the proof for

$(\infty, 0)$. □

Proof of Lemma 6 (Unstable Mixed Outcomes). Suppose $k = 2$ and consider the mixed learning outcome $\lambda_1 = 0$ in which θ_1's belief converges to zero and θ_2's belief doesn't converge. Suppose $\lambda_1 = 0 \notin \Lambda_M(\omega)$, 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 $\varepsilon < e^{-E}$. 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, the mixed outcome 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 Lemma 3, 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 $(0, 0)$. Suppose $(0, 0) \notin \Lambda_1(\omega)$. By the construction in Lemma 3, 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 Lemma 3, 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,

$$Pr(\tau_1 < \infty \text{ or } \boldsymbol{\lambda}_t \in B_\varepsilon((0, 0)) \cup B_\varepsilon((0, \infty)) \text{ i.o.}) = 1. \tag{23}$$

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. □

Proof of Lemma 7 (Belief Convergence). Suppose $k = 2$, $\Lambda(\omega)$ contains an agreement outcome or maximally accessible disagreement outcome and $\Lambda_M(\omega)$ is empty. Recall that $\mathcal{B}$ is the set of locally stable neighborhoods and $\mathcal{B}_U$ is the set of locally unstable neighborhoods. 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 11, there exists a finite sequence of actions and signals such that starting from any initial belief $\boldsymbol{\lambda}_1 \in (0, \infty)^2$, $\langle \boldsymbol{\lambda}_t \rangle$ 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 Lemma 3, $\langle \boldsymbol{\lambda}_t \rangle$ 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 6, $\langle \boldsymbol{\lambda}_t \rangle$ almost surely leaves this neighborhood or converges to a locally stable belief. Therefore, mixed learning outcomes almost surely do not arise. By Lemma 2, $\langle \boldsymbol{\lambda}_t \rangle$ 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. □

A.1.3 Intermediate Results

The following three lemmas are intermediate results used to prove Lemmas 1 - 7. They hold for any $k \geq 1$.

Lemma 10. *If $|\Theta| > 1$ then for all sociable types θ_j,*

$$\frac{\hat{\psi}_j(a_1, \sigma_R|L, \boldsymbol{\lambda})}{\hat{\psi}_j(a_1, \sigma_R|R, \boldsymbol{\lambda})} < 1 \text{ and } \frac{\hat{\psi}_j(a_M, \sigma_L|L, \boldsymbol{\lambda})}{\hat{\psi}_j(a_M, \sigma_L|R, \boldsymbol{\lambda})} > 1$$

at all beliefs $\boldsymbol{\lambda} \in [0, \infty]^k$. Further, for any $\boldsymbol{\lambda}_t \in (0, \infty)^k$, if $(a_t, \sigma_t) = (a_M, \sigma_L)$, then $\boldsymbol{\lambda}_{t+1} > \boldsymbol{\lambda}_t$, and if $(a_t, \sigma_t) = (a_1, \sigma_R)$, then $\boldsymbol{\lambda}_{t+1} < \boldsymbol{\lambda}_t$.

If $|\Theta = 1|$ then for any likelihood ratio $\boldsymbol{\lambda} \in (0, \infty)$ there exists actions $a_1(\boldsymbol{\lambda})$ and $a_M(\boldsymbol{\lambda})$ such that

$$\frac{\hat{\psi}_j(a_1(\boldsymbol{\lambda}), \sigma_R|L, \boldsymbol{\lambda})}{\hat{\psi}_j(a_1(\boldsymbol{\lambda}), \sigma_R|R, \boldsymbol{\lambda})} < 1 \text{ and } \frac{\hat{\psi}_j(a_M, \sigma_L|L, \boldsymbol{\lambda})}{\hat{\psi}_j(a_M, \sigma_L|R, \boldsymbol{\lambda})} > 1$$

at all beliefs $\boldsymbol{\lambda} \in [0,\infty]^k$*. Further, for any* $\boldsymbol{\lambda}_t \in (0,\infty)^k$*, if* $(a_t, \sigma_t) = (a_M, \sigma_L)$*, then* $\boldsymbol{\lambda}_{t+1} > \boldsymbol{\lambda}_t$*, and if* $(a_t, \sigma_t) = (a_1, \sigma_R)$*, then* $\boldsymbol{\lambda}_{t+1} < \boldsymbol{\lambda}_t$*.*

Proof. Fix $\boldsymbol{\lambda} \in [0,\infty]^k$ and consider the $\succ$-minimal action played with positive probability at $\boldsymbol{\lambda}$, where $\succ$ is the order from Definition 2 (if $|\Theta| > 1$ this is a_1) and in a slight abuse of notation denote this by a_1. Consider how sociable type θ_j updates its beliefs following a_1. Since preferences are aligned, all types $\theta_i \in \Theta$ choose a_1 at $\boldsymbol{\lambda}$ for any signal $s \leq \bar{s}_{i,1}(\lambda_i)$, where $\bar{s}_{i,1}(\lambda_i) \in [0,1]$. Type θ_j believes that type θ_i plays a_1 with probability $\hat{F}_j^{\omega}(\bar{s}_{i,1}(\lambda_i))$. By Lemma A.1 in Smith and Sorensen (2000), $F^L(s) \leq F^R(s)$, with strict equality for $s \in (0,1)$. Since signals are aligned, this is also true for the subjective beliefs, i.e. $\hat{F}_j^L(s) \leq \hat{F}_j^R(s)$. Therefore, $\hat{F}_j^L(\bar{s}_{i,1}(\lambda_i)) \leq \hat{F}_j^R(\bar{s}_{i,1}(\lambda_i))$ for each θ_i. Further, under Assumption 3.ii, at least one type θ_i with $\hat{\pi}_j(\theta_i) > 0$ has $\bar{s}_{i,1}(\lambda_i) \in (0,1)$ and $\hat{F}_j^L(\bar{s}_{i,1}(\lambda_i)) < \hat{F}_j^R(\bar{s}_{i,1}(\lambda_i))$. Therefore, $\hat{\psi}_j(a_1|L,\boldsymbol{\lambda}) < \hat{\psi}_j(a_1|R,\boldsymbol{\lambda})$, since it is a convex combination of the probability each type chooses a_1. Public signals are aligned, so it must be that $\rho_j(\sigma_R) \leq 1/2$, as the maximal public signal in state R is either uninformative or indicative of state R. Under Assumption 3.i, $\rho_j(\sigma_R) < 1/2$. Either Assumption 3.i or 3.ii holds, so it must be that $\hat{\psi}_j(a_1, \sigma_R|L,\boldsymbol{\lambda}) < \hat{\psi}_j(a_1, \sigma_R|R,\boldsymbol{\lambda})$. Therefore, if $\boldsymbol{\lambda}_t \in (0,\infty)^k$, following $(a_t, \sigma_t) = (a_1, \sigma_R)$, beliefs update toward state R,

$$\lambda_{j,t+1} = \lambda_{j,t}\left(\frac{\hat{\psi}_j(a_1, \sigma_R|L,\boldsymbol{\lambda}_t)}{\hat{\psi}_j(a_1, \sigma_R|R,\boldsymbol{\lambda}_t)}\right) < \lambda_{j,t}.$$

Similar logic holds for (a_M, σ_L). □

Lemma 11. *If* $|\Theta| > 1$ *then the minimal update to the likelihood ratio following* (a_M, σ_L) *is greater than one,*

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

and the maximal update to the log likelihood ratio following (a_1, σ_R) *is less than one,*

$$\sup_{\boldsymbol{\lambda}\in[0,\infty]^k} \frac{\hat{\psi}_i(a_1, \sigma_R|L,\boldsymbol{\lambda})}{\hat{\psi}_i(a_1, \sigma_R|R,\boldsymbol{\lambda})} < 1.$$

If $|\Theta| = 1$ *then there exists a function* $a_M : (0,\infty)^k \to \mathcal{A}$ *such that the minimal update to the likelihood ratio following* $(a_M(\boldsymbol{\lambda}), \sigma_L)$ *is greater than one,*

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

and there exists a function $a_1 : (0,\infty)^k \to \mathcal{A}$ *such that the maximal update to the log* *likelihood ratio following* $(a_1(\boldsymbol{\lambda}), \sigma_R)$ *is less than one,*

$$\sup_{\boldsymbol{\lambda} \in [0,\infty]^k} \frac{\hat{\psi}_i(a_1(\boldsymbol{\lambda}), \sigma_R | L, \boldsymbol{\lambda})}{\hat{\psi}_i(a_1(\boldsymbol{\lambda}), \sigma_R | R, \boldsymbol{\lambda})} < 1.$$

Proof. Let a_1 denote the action from Lemma 10.

$$\log \frac{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = L) + \hat{\pi}_i(\Theta_S)\hat{P}_i(a_1 | \theta \in \Theta_S, \omega = L)}{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = R) + \hat{\pi}_i(\Theta_S)\hat{P}_i(a_1 | \theta \in \Theta_S, \omega = R)}.$$

It must be that $\hat{P}_i(a_1|\theta, \omega = L) \leq \hat{P}_i(a_1|\theta, \omega = R)$ in equilibrium, so this is bounded above by

$$\log \frac{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = L) + \hat{\pi}_i(\Theta_S)\hat{P}_i(a_1 | \theta \in \Theta_S, \omega = R)}{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = R) + \hat{\pi}_i(\Theta_S)\hat{P}_i(a_1 | \theta \in \Theta_S, \omega = R)},$$

which in turn is bounded above by

$$\log \frac{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = L) + \hat{\pi}_i(\Theta_S)}{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = R) + \hat{\pi}_i(\Theta_S)} \leq 0,$$

Since by construction $\hat{P}_i(a_1|\theta \in \Theta_A, \omega = R) \geq \hat{P}_i(a_1|\theta \in \Theta_A, \omega = L)$ for all possible a_1)) with this inequality strict under assumption 3.ii. A similar construction holds for a_M. Moreover, $\rho_i(\sigma_R) \leq 1/2$, and $\rho_i(\sigma_R) < 1/2$ if Assumption 3.i holds. Therefore, under Assumption 3.ii

$$\frac{\hat{\psi}_i(a_1, \sigma_R | L, \boldsymbol{\lambda})}{\hat{\psi}_i(a_1, \sigma_R | R, \boldsymbol{\lambda})} < \frac{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = L) + \hat{\pi}_i(\Theta_S)}{\hat{\pi}_i(\Theta_A)\hat{P}_i(a_1 | \theta \in \Theta_A, \omega = R) + \hat{\pi}_i(\Theta_S)} \frac{\rho_i(\sigma_1)}{1 - \rho_i(\sigma_R)} < 1.$$

□

Lemma 12 (Continuity)**.** $\boldsymbol{\lambda} \mapsto \psi(a, \sigma | \omega, \boldsymbol{\lambda})$ *and* $\boldsymbol{\lambda} \mapsto \hat{\psi}_i(a, \sigma | \omega, \boldsymbol{\lambda})$ *are continuous at each stationary* $\boldsymbol{\lambda}^* \in \{0, \infty\}^k$ *for all* $(a, \sigma) \in \mathcal{A} \times \Sigma$ *and* $\omega \in \{L, R\}$*.*

Proof. 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 sociable types who play a at 0^k. Fix $\varepsilon > 0$. Let

$$\delta_1 \equiv \min_{a \in \mathcal{A}} \frac{\varepsilon}{\max\{\pi(\Theta_S \setminus \Theta_a), \pi(\Theta_a)\}}$$

and

$$\delta_2 \equiv \min_{a\in\mathcal{A},\theta_i\in\Theta_S} \frac{\varepsilon}{\max\{\hat{\pi}^i(\Theta_S \setminus \Theta_a), \hat{\pi}^i(\Theta_a)\}}.$$

and $\delta \equiv \min\{\delta_1, \delta_2\}$. Signals are not perfectly informative, so there exists a $\bar{s}$ such that $1 - \hat{F}_i^{\omega}(\bar{s}) < \delta$ and $1 - \hat{F}_i^{\omega}(\bar{s}) < \delta$ for all $\theta_i \in \Theta_S$ and $\omega \in \{L, R\}$. Define

$$\varepsilon_1(\delta) \equiv \frac{\varepsilon_1}{\max_{\theta_i\in\Theta_s} r_i(\bar{s})/(1 - r_i(\bar{s}))}.$$

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 playing action a is bounded above by

$$\psi(a|\omega, \boldsymbol{\lambda}) \leq \pi(\Theta_a) + \delta\pi(\Theta_S \setminus \Theta_a) + q_a$$

and bounded below by

$$\psi(a|\omega, \boldsymbol{\lambda}) \geq \pi(\Theta_a)(1 - \delta) + q_a.$$

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 public signal distribution is independent of $\boldsymbol{\lambda}$. Therefore, this continuity extends to $\psi(a, \sigma|\omega, \boldsymbol{\lambda})$ and $\hat{\psi}_i(a, \sigma|\omega, \boldsymbol{\lambda})$ for all $\theta_i \in \Theta_S$. The proof for other stationary beliefs is identical. □

A.2 Analogue of Theorem 1 for $k > 2$.

This section proves the analogue of Theorem 1 for more than two sociable types, $k > 2$. The statement of the result is identical to Theorem 1, using the modified definitions of $\Lambda(\omega)$ (defined in (19)), $\Lambda_M(\omega)$ and maximal accessibility (defined for $k > 2$ below). Recall that Lemmas 1 - 4 hold for all $k \geq 1$. Therefore, we prove analogues of Lemmas 5 - 7.

Global Stability of Disagreement. As above, without loss of generality, order a disagreement outcome so that the first κ types have belief zero, and the remaining $k - \kappa$ types have belief infinity, i.e. $\boldsymbol{\lambda} = (0^{\kappa}, \infty^{k-\kappa})$. As in the case of two sociable types, we can use the maximal action and signal pairs to define a sufficient condition for global stability, and use this to prove an analogue of Lemma 5.

Definition 6′ (Maximal Accessibility). *Disagreement outcome* $(0^{\kappa}, \infty^{k-\kappa})$ *is* maximally accessible *if either:*

(i) for all $\kappa' = 0, ..., \kappa - 1$*, given* $\boldsymbol{\lambda} = (0^{\kappa'}, \infty^{k-\kappa'})$*,* $\theta_i \succ_{\boldsymbol{\lambda}} \theta_{\kappa'+1}$ *for all* $i > \kappa' + 1$ *and* $\theta_{\kappa'+1} \succeq_{\boldsymbol{\lambda}} \theta_i$ *for all* $i < \kappa' + 1$*;*

(ii) for all $\kappa' = \kappa+1, ..., k$*, given* $\boldsymbol{\lambda} = (0^{\kappa'}, \infty^{k-\kappa'})$*,* $\theta_i \succeq_{\boldsymbol{\lambda}} \theta_{\kappa'}$ *for all* $i > \kappa'$ *and* $\theta_{\kappa'} \succ_{\boldsymbol{\lambda}} \theta_i$ *for all* $i < \kappa'$*,*

where $\succ_{\boldsymbol{\lambda}}$ *is the maximal L-order defined in Definition 5.*

Note that this definition is equivalent to Definition 6 when $k = 2$. If the belief of the type with $\lambda_i^* = 0$ that interprets maximal action and signal pairs as the weakest evidence of state R decreases at a faster rate than the belief of the type with $\lambda_j^* = \infty$ that interprets maximal action and signal pairs as the strongest evidence of state R, then it is possible to find a finite sequence of maximal action and public signal pairs that separate beliefs. Once again, this condition is straightforward to verify from the primitives of the model. As in the case of $k = 2$, for any disagreement outcome in $\Lambda(\omega)$, maximal accessibility is a sufficient condition for global stability. Given this revised definition of maximal accessibility, the statement of Lemma 5 is identical.[42]

Lemma 5′ (Global Stability of Disagreement)**.** *If disagreement outcome* $\boldsymbol{\lambda}^* = (0^\kappa, \infty^{k-\kappa})$ *is locally stable and maximally accessible, then* $\boldsymbol{\lambda}^*$ *is globally stable.*

Mixed Learning Outcomes. Consider the mixed outcome in which beliefs converge to $\boldsymbol{\lambda}_I^* \in \{0, \infty\}^{|I|}$ for some subset of sociable types $I \subset \Theta_S$, where $\boldsymbol{\lambda}_I^*$ denotes the likelihood ratio vector restricted to set I, and beliefs do not converge for the remaining sociable types $N \equiv \Theta_S \setminus I$. This outcome is not locally stable if it is possible for the beliefs of the non-convergent types to converge. For example, suppose there are three sociable types. If $(0, 0, 0) \in \Lambda_1(\omega)$, then the mixed learning outcome in which θ_1 has cyclical learning, and θ_2 and θ_3 have correct learning is not locally stable, since if the beliefs of θ_2 and θ_3 converge to zero, then the beliefs of θ_1 will also almost surely converge to zero. For mixed learning outcomes in which two or more types have cyclical learning, the argument is more involved. To rule out mixed learning, we also need to show that a locally stable outcome for the non-convergent types is accessible from other points in the mixed outcome belief space. For example, if $(0, 0, 0) \in \Lambda_1(\omega) \cap \Lambda_2(\omega)$, then to rule out the mixed learning outcome in which θ_1 and θ_2 have cyclical learning and θ_3 has correct learning, we need to show that for $\boldsymbol{\lambda} \in \{(0, \infty, 0), (\infty, 0, 0), (\infty, \infty, 0)\}$, either (i) beliefs will almost surely enter a neighborhood of $(0, 0, 0)$ from a neighborhood of $\boldsymbol{\lambda}$, or (ii) $\boldsymbol{\lambda} \in \Lambda_1(\omega) \cap \Lambda_2(\omega)$. The following definition formalizes this notion.

Definition 12 (Mixed Accessible)**.** *Given mixed outcome* $\boldsymbol{\lambda}_I^*$ *for* $I \subset \Theta_S$ *and* $N = \Theta_S \setminus I$, $\boldsymbol{\lambda}'_N$ *is* mixed accessible *from* $\boldsymbol{\lambda}_N$ *if* $\boldsymbol{\lambda}'_N \neq \boldsymbol{\lambda}_N$ *and* $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N) \notin \Lambda_i(\omega)$ *for each* $i \in N$ *such that* $\lambda'_i \neq \lambda_i$, *and* $\boldsymbol{\lambda}'_N$ *is* strongly mixed accessible *from* $\boldsymbol{\lambda}_N$ *if* $\boldsymbol{\lambda}'_N$ *is mixed accessible from* $\boldsymbol{\lambda}_N$ *and for each distinct* $i, j \in N$ *with* $\lambda_i \neq \lambda'_i$ *and* $\lambda_j \neq \lambda'_j$, *then* $\lambda_i = \lambda_j$.

[42] The proof of Lemma 5′ shows that Definition 6′ is sufficient for a much weaker, but more complicated to verify, condition called separability (Definition 9) that utilizes the entire set of actions to separate beliefs.

Given mixed outcome $\boldsymbol{\lambda}_I^*$, we construct a graph $\mathcal{G}(\boldsymbol{\lambda}_I^*)$ between the nodes $\boldsymbol{\lambda}_N \in \{0,\infty\}^{|N|}$ to represent which nodes $\boldsymbol{\lambda}_N$ are mixed accessible from other nodes $\boldsymbol{\lambda}'_N$ for the non-convergent types.

Definition 13 (Accessible Graph)**.** *Given $\boldsymbol{\lambda}_I^*$, define the directed graph $\mathcal{G}(\boldsymbol{\lambda}_I^*)$ with nodes $\boldsymbol{\lambda}_N \in \{0,\infty\}^{|N|}$ as: there is an edge from $\boldsymbol{\lambda}_N$ to $\boldsymbol{\lambda}'_N$ iff $\boldsymbol{\lambda}'_N$ is strongly mixed accessible from $\boldsymbol{\lambda}_N$.*

A *terminal node* $\boldsymbol{\lambda}_N$ is a node with no edges leaving it.

Definition 14 (Reducible)**.** *A mixed outcome $\boldsymbol{\lambda}_I^*$ is* reducible *if $\mathcal{G}(\boldsymbol{\lambda}_I^*)$ has no cycles.*

If a mixed outcome is reducible, then conditional on the convergent types I remaining in a neighborhood of $\boldsymbol{\lambda}_I^*$, almost surely, the beliefs of the non-convergent types converge. This is a contradiction. Therefore, almost surely, this mixed outcome will not arise. Let $\Lambda_M(\omega)$ denote the set of mixed learning outcomes that are not reducible,

$$\Lambda_M(\omega) \equiv \{\boldsymbol{\lambda}_I^* \in \{0,\infty\}^{|I|}, I \subset \Theta_S | \boldsymbol{\lambda}_I^* \text{ is not reducible}\}. \tag{24}$$

As in the case of two sociable types, if a mixed learning outcome arises with positive probability, it must be in $\Lambda_M(\omega)$.

Reducibility is always satisfied in some important cases and is relatively straightforward to verify. For instance, it is satisfied in models near the correctly specified model, in which $\gamma_i(\boldsymbol{\lambda},\omega) < 0$ at all stationary $\boldsymbol{\lambda} \in \{0,\infty\}^k$ for all sociable types θ_i. In this model, each node in the graph is connected to all nodes with fewer ∞'s than it, and is connected to no other nodes. Therefore, every path in the graph terminates at $\boldsymbol{\lambda}_N = 0^{|N|}$. This is a convergent point for mixed outcome $\boldsymbol{\lambda}_I^* = 0^{|I|}$. Otherwise, it is a point at which λ_i moves towards zero in expectation for all $i \in \Theta_S$, so therefore, some $i \in I$'s beliefs must eventually exit a neighborhood of $(\boldsymbol{\lambda}_I^*, 0^{|N|})$.[43]

Given the modified definition of $\Lambda_M(\omega)$, the statement of Lemma 6 is identical.

Lemma 6′ (Unstable Mixed Outcomes)**.** *Given $I \subset \Theta_S$ and $N = \Theta_S \setminus I$, if mixed learning outcome $\boldsymbol{\lambda}_I^* \notin \Lambda_M(\omega)$, then $Pr(\boldsymbol{\lambda}_{I,t} \to \boldsymbol{\lambda}_I^*$ and $\boldsymbol{\lambda}_{N,t}$ does not converge$) = 0$.*

Belief Convergence. Finally, if there is at least one globally stable agreement or maximally accessible disagreement outcome, and no locally stable mixed outcomes, then the likelihood ratio converges almost surely for all types. Given the modified definitions of $\Lambda(\omega)$, $\Lambda_M(\omega)$ and maximal accessibility, the statement of Lemma 7 is identical.

[43] An alternative condition involves bounding $\gamma_i(\boldsymbol{\lambda},\omega)$ across the belief space for $i \in N$ can also be used to rule out mixed learning.

Lemma 7′ (Belief Convergence). *Suppose $\Lambda(\omega)$ contains an agreement outcome or maximally accessible disagreement outcome and $\Lambda_M(\omega)$ is empty. 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.*

A.2.1 Proofs of Lemmas 5′ - 7′.

We first state an intermediate result, and then prove Lemma 5′. Using the definition of accessible (Definition 11), the analogue of Lemma 9 for $k > 2$ is as follows.

Lemma 9′ (Accessible Disagreement). *Disagreement outcome $\boldsymbol{\lambda}^*_J = (0^\kappa, \infty^{k-\kappa})$ is accessible if there exists a sequence of stationary likelihood ratios $\boldsymbol{\lambda}^*_1, \boldsymbol{\lambda}^*_2 \dots \boldsymbol{\lambda}^*_J$, with $\boldsymbol{\lambda}^*_1 \in \{0^k, \infty^k\}$ and $\boldsymbol{\lambda}^*_j$ adjacently accessible from $\boldsymbol{\lambda}^*_{j-1}$ for $j = 2, ..., J$.*

Proof. The proof follows almost directly from Lemma 8. Each element of the sequence $\boldsymbol{\lambda}^*_j$ is adjacently accessible from the previous element of the sequence $\boldsymbol{\lambda}^*_{j-1}$. Starting with $\boldsymbol{\lambda}^*_J$ and any $\varepsilon_J > 0$, there exists an $\varepsilon_{J-1} > 0$ and $\tau_J < \infty$ such that if $\boldsymbol{\lambda}_t \in B_{\varepsilon_{J-1}}(\boldsymbol{\lambda}^*_{J-1})$, then $Pr(\boldsymbol{\lambda}_{t+\tau_J} \in B_{\varepsilon_J}(\boldsymbol{\lambda}^*_J)) > 0$. Iterating back to $\boldsymbol{\lambda}^*_1$, for any $\varepsilon_J > 0$, there exists an $\varepsilon_1 > 0$ and $\tau_2 < \infty$ such that if $\boldsymbol{\lambda}_t \in B_{\varepsilon_1}(\boldsymbol{\lambda}^*_1)$, then $Pr(\boldsymbol{\lambda}_{t+\sum_{j=2}^J \tau_j} \in B_{\varepsilon_J}(\boldsymbol{\lambda}^*_J)) > 0$. Consider agreement outcome $\boldsymbol{\lambda}^*_1 \in \{0^k, \infty^k\}$. By Lemma 4, 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 and public signals such that following this sequence, $\boldsymbol{\lambda}_{\tau_1+1} \in B_{\varepsilon_1}(\boldsymbol{\lambda}^*_1)$. Therefore, from any initial beliefs, $Pr(\boldsymbol{\lambda}_{\tau_1+1} \in B_{\varepsilon_1}(\boldsymbol{\lambda}^*_1)) > 0$. Therefore, for any $\varepsilon_J > 0$ and initial beliefs $\boldsymbol{\lambda}_1 \in (0,\infty)^k$, $Pr(\boldsymbol{\lambda}_\tau \in B_{\varepsilon_J}(\boldsymbol{\lambda}^*_J)) > 0$, where $\tau \equiv \sum_{j=1}^J \tau_j + 1$. Since each $\tau_j < \infty$, $\tau < \infty$. □

Proof of Lemma 5′. Consider $\boldsymbol{\lambda}^* = (0^\kappa, \infty^{k-\kappa})$. Suppose $\boldsymbol{\lambda}^* \in \Lambda(\omega)$ and $\boldsymbol{\lambda}^*$ is maximally accessible. Consider the sequence of stationary likelihood ratios $\boldsymbol{\lambda}^*_j = (0^{k-j+1}, \infty^{j-1})$ for $j = 1, \dots k-\kappa+1$, and suppose part (ii) of Definition 6′ holds. We first show that this implies separability at zero (Definition 9) for each likelihood ratio in the sequence. For each $j = 1, \dots k - \kappa + 1$, define the submatrix

$$A_j \equiv \begin{pmatrix} \log \frac{\hat{\psi}_{k-j+1}(a_1,\sigma_R|L,\boldsymbol{\lambda}^*_j)}{\hat{\psi}_{k-j+1}(a_1,\sigma_R|R,\boldsymbol{\lambda}^*_j)} & \log \frac{\hat{\psi}_{k-j+1}(a_M,\sigma_L|L,\boldsymbol{\lambda}^*_j)}{\hat{\psi}_{k-j+1}(a_M,\sigma_L|R,\boldsymbol{\lambda}^*_j)} \\ \log \frac{\hat{\psi}_{k-j}(a_1,\sigma_R|L,\boldsymbol{\lambda}^*_j)}{\hat{\psi}_{k-j}(a_1,\sigma_R|R,\boldsymbol{\lambda}^*_j)} & \log \frac{\hat{\psi}_{k-j}(a_M,\sigma_L|L,\boldsymbol{\lambda}^*_j)}{\hat{\psi}_{k-j}(a_M,\sigma_L|R,\boldsymbol{\lambda}^*_j)} \end{pmatrix}.$$

Since $\theta_{k-j+1} \succ \theta_{k-j}$, this has a positive determinant. Therefore, there exists a $c \in \mathbb{R}^2_+$ that solves

$$A_j c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

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

$$A_j \tilde{c} = \begin{pmatrix} G_{k-j+1} \\ G_{k-j} \end{pmatrix},$$

where $G_{k-j+1} > 0$ and $G_{k-j} < 0$. Moreover, by maximal accessibility, for any $j' > k-j+1$,

$$\left(\log \frac{\hat{\psi}_{j'}(a_M, \sigma_L | L, \boldsymbol{\lambda}^*_j)}{\hat{\psi}_{j'}(a_M, \sigma_L | R, \boldsymbol{\lambda}^*_j)}, \log \frac{\hat{\psi}_{j'}(a_1, \sigma_R | L, \boldsymbol{\lambda}^*_j)}{\hat{\psi}_{j'}(a_1, \sigma_R | R, \boldsymbol{\lambda}^*_j)} \right) \cdot \tilde{c} > 0$$

and for any $j' < k-j$,

$$\left(\log \frac{\hat{\psi}_{j'}(a_M, \sigma_L | L, \boldsymbol{\lambda}^*_j)}{\hat{\psi}_{j'}(a_M, \sigma_L | R, \boldsymbol{\lambda}^*_j)}, \log \frac{\hat{\psi}_{j'}(a_1, \sigma_R | L, \boldsymbol{\lambda}^*_j)}{\hat{\psi}_{j'}(a_1, \sigma_R | R, \boldsymbol{\lambda}^*_j)} \right) \cdot \tilde{c} < 0.$$

Therefore, $\boldsymbol{\lambda}^*_j$ is separable at zero, since we can set the elements of c to zero for the remaining action and signal pairs in matrix (20). Therefore, by Lemma 8, $\boldsymbol{\lambda}^*_{j+1}$ is adjacently accessible from $\boldsymbol{\lambda}^*_j$. Since this holds for each element of the sequence, starting at $\boldsymbol{\lambda}^*_1 = 0^k$ and ending at $\boldsymbol{\lambda}^*_J = \boldsymbol{\lambda}^*$, by Lemma 9′, $\boldsymbol{\lambda}^*$ is accessible. Similar to the proof of Lemma 9, we can choose $\varepsilon < e^{-E}$, so that the likelihood ratio reaches the locally stable neighborhood of $\boldsymbol{\lambda}^*$ with positive probability. From here, local stability implies that $P(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}^*) > 0$. The case where part (i) of Definition 6′ holds is analogous. □

We first state an intermediate result, then prove Lemma 6′.

Lemma 13. *Given mixed outcome $\boldsymbol{\lambda}^*_I$ and $\mathcal{G}(\boldsymbol{\lambda}^*_I)$, if $\boldsymbol{\lambda}_N$ is a terminal node, then $(\boldsymbol{\lambda}^*_I, \boldsymbol{\lambda}_N) \in \cap_{i \in N} \Lambda_i(\omega)$.*

Proof. Let $\boldsymbol{\lambda}_N$ be a terminal node in $\mathcal{G}(\boldsymbol{\lambda}^*_I)$. By definition of terminal node, no nodes are strongly mixed accessible from $\boldsymbol{\lambda}_N$. If any node is mixed accessible from $\boldsymbol{\lambda}_N$, then there exists an $i \in N$ such that $(\boldsymbol{\lambda}^*_I, \boldsymbol{\lambda}_N) \notin \Lambda_i(\omega)$. Then the node $\boldsymbol{\lambda}'_N$ where $\lambda'_j = \lambda_j$ for all $j \neq i$ is strongly mixed accessible, so $\boldsymbol{\lambda}_N$ is not a terminal node. This is a contradiction. Therefore, if $\boldsymbol{\lambda}_N$ is a terminal node, then no nodes $\boldsymbol{\lambda}'_N$ are mixed accessible from $\boldsymbol{\lambda}_N$. Therefore, by definition of mixed accessibility, $(\boldsymbol{\lambda}^*_I, \boldsymbol{\lambda}_N) \in \cap_{i \in N} \Lambda_i(\omega)$. □

Proof of Lemma 6′. Suppose mixed outcome $\boldsymbol{\lambda}^*_I$ is reducible, i.e. $\boldsymbol{\lambda}^*_I \notin \Lambda_M(\omega)$. We will show that this implies that $\boldsymbol{\lambda}^*_I$ almost surely does not occur. Fix $\varepsilon < e^{-E}$ and suppose $\boldsymbol{\lambda}_{I,1} \in B_\varepsilon(\boldsymbol{\lambda}^*_I)$. We will show that almost surely, either (i) there exists a time $\tau < \infty$ such that $\boldsymbol{\lambda}_{I,\tau} \notin B_\varepsilon(\boldsymbol{\lambda}^*_I)$ or (ii) $\langle \boldsymbol{\lambda}_t \rangle$ converges for all sociable types.

By reducibility, at every $\boldsymbol{\lambda}_N \in \{0, \infty\}^{|N|}$, either $(\boldsymbol{\lambda}^*_I, \boldsymbol{\lambda}_N) \in \cap_{i \in N} \Lambda_i(\omega)$ or there exists a $\boldsymbol{\lambda}'_N \in \{0, \infty\}^{|N|}$ that is strongly mixed accessible from $\boldsymbol{\lambda}_N$ such that $(\boldsymbol{\lambda}^*_I, \boldsymbol{\lambda}'_N) \in$

$\cap_{i\in N}\Lambda_i(\omega)$. First consider $\boldsymbol{\lambda}_N \in \{0,\infty\}^{|N|}$ such that $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N) \in \cap_{i\in N}\Lambda_i(\omega)$. By the construction in Lemma 3, if beliefs enter $B_\varepsilon((\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N))$, then $\langle \boldsymbol{\lambda}_{N,t}\rangle$ is bounded above by a process that converges to $\boldsymbol{\lambda}_N$ with positive probability, and this probability is uniformly bounded away from zero for any belief in $B_\varepsilon((\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N))$. If $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N) \in \cap_{i\in I}\Lambda_i(\omega)$, then $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N)$ is locally stable, so with positive probably, $\boldsymbol{\lambda}_t \to (\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N)$. Otherwise, if $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N) \notin \cap_{i\in I}\Lambda_i(\omega)$, then for some $i \in I$, $\langle \lambda_{i,t}\rangle$ is bounded below by a process that almost surely leaves $B_\varepsilon(\boldsymbol{\lambda}_I^*)$. Therefore, in the event that $\langle \boldsymbol{\lambda}_{N,t}\rangle \to \boldsymbol{\lambda}_N$, $\langle \boldsymbol{\lambda}_{I,t}\rangle$ almost surely leaves $B_\varepsilon(\boldsymbol{\lambda}_I^*)$.

Next consider $\boldsymbol{\lambda}_N \in \{0,\infty\}^{|N|}$ such that $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N) \notin \cap_{i\in N}\Lambda_i(\omega)$. Fix $0 < \varepsilon' < e^{-E}$. We want to show that there exists a $\varepsilon_2 > 0$ such that if initial belief $\boldsymbol{\lambda}_{N,1} \in B_{\varepsilon_2}(\boldsymbol{\lambda}_N)$, then there exists a $\boldsymbol{\lambda}'_N$ that is strongly mixed accessible from $\boldsymbol{\lambda}_N$ such that with probability uniformly bounded away from zero in initial belief $\boldsymbol{\lambda}_{N,1}$, beliefs enter a neighborhood $B_{\varepsilon'}(\boldsymbol{\lambda}'_N)$. Given $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N)$, let λ_i denote the component for type $i \in N$ and λ_i^* denote the component for type $i \in I$. By the construction in Lemma 3, there exists an $i \in N$ such that $\langle \lambda_{i,t}\rangle$ is bounded below by a process that almost surely leaves $B_\varepsilon(\lambda_i)$. Let N_U be the set of types $i \in N$ such that $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N) \notin \Lambda_i(\omega)$, with $N_{U,0}$ the set of $i \in N_U$ such that $\lambda_i = 0$ and $N_{U,\infty}$ the set of $i \in N_U$ such that $\lambda_i = \infty$. We now argue that starting from a neighborhood $B_{\varepsilon_2}(\boldsymbol{\lambda}_N)$ for $i \in N$ and $B_\varepsilon(\boldsymbol{\lambda}_I^*)$ for $i \in I$, with positive probability, either $\langle \boldsymbol{\lambda}_{I,t}\rangle$ leaves $B_\varepsilon(\boldsymbol{\lambda}_I^*)$ or $\langle \boldsymbol{\lambda}_{N,t}\rangle$ reaches $B_{\varepsilon'}(\boldsymbol{\lambda}'_N)$ for some strongly mixed accessible point $\boldsymbol{\lambda}'_N$. For $i \in N_{U,0}$, let N_i be the minimum number of (a_M, σ_L) actions it takes for any $\lambda_{i,t} \in [\varepsilon', 1/\varepsilon']$ to hit $1/\varepsilon'$. Similarly, for $i \in N_{U,\infty}$, let N_i be the minimum number of (a_1, σ_R) actions it takes for any $\lambda_{i,t} \in [\varepsilon', 1/\varepsilon']$ to hit ε'. By the construction in Lemma 3, there exists an $\varepsilon_2 > 0$ such that if $\boldsymbol{\lambda}_{N,1} \in B_{\varepsilon_2/2}(\boldsymbol{\lambda}_N)$, with positive probability there exists a finite t such that $\boldsymbol{\lambda}_{N\setminus N_U,t} \in B_{\varepsilon_2}(\boldsymbol{\lambda}_{N\setminus N_U})$, and $\boldsymbol{\lambda}_{N_U,t} \notin B_{\varepsilon'}(\boldsymbol{\lambda}_{N_U})$.

Choose ε_2 such that if $\boldsymbol{\lambda}_{N\setminus N_U,1} \in B_{\varepsilon_2}(\boldsymbol{\lambda}_{N\setminus N_U})$, then after $\sum_{i\in N_{U,0}} N_i$ action and signal realizations (a_M, σ_L), $\lambda_{i,t} \in B_{\varepsilon'}(\lambda_i)$ for all $i \in N \setminus N_{U,0}$, and after $\sum_{i\in N_{U,\infty}} N_i$ action and signal realizations (a_R, σ_R), $\lambda_{i,t} \in B_{\varepsilon'}(\lambda_i)$ for all $i \in N \setminus N_{U,\infty}$. Therefore, if $\boldsymbol{\lambda}_{N,1} \in B_{\varepsilon_2/2}(\boldsymbol{\lambda}_N)$ and $\boldsymbol{\lambda}_{I,1} \in B_\varepsilon(\boldsymbol{\lambda}_I^*)$, then with positive probability either (i) there exists a $t < \infty$ such that $\boldsymbol{\lambda}_{I,t} \notin B_\varepsilon(\boldsymbol{\lambda}_I^*)$, or (ii) there exists $t < \infty$ such that for some $i \in N_U$, $\lambda_{i,t} \notin B_{\varepsilon'}(\lambda_i)$ and for all $i \in N \setminus N_U$, $\lambda_{i,t} \in B_{\varepsilon_2}(\lambda_i)$. First consider case (ii) and suppose that a type $i \in N_{U,0}$ leaves. After N_i actions and signals (a_M, σ_L), if $\boldsymbol{\lambda}_{N,t} \in B_{\varepsilon'}(\boldsymbol{\lambda}'_N)$ for some $\boldsymbol{\lambda}'_N$ that is strongly mixed accessible from $\boldsymbol{\lambda}_N$, then stop. Otherwise, there exists an $i_2 \in N_{U,0}$ such that $\lambda_{i_2,t} \notin B_{\varepsilon'}(\lambda_{i_2})$. Repeat N_{i_2} realizations (a_M, σ_L). After these $N_{i_1} + N_{i_2}$ realizations of (a_M, σ_L), if $\boldsymbol{\lambda}_{N,t} \in B_{\varepsilon'}(\boldsymbol{\lambda}'_N)$ for some $\boldsymbol{\lambda}'_N$ that is strongly mixed accessible from $\boldsymbol{\lambda}_N$, then stop. Otherwise, there is an $i_3 \in N_{\boldsymbol{\lambda},0}$ such that $\lambda_{i_3,t} \notin B_{\varepsilon'}(\lambda_{i_3})$. Repeat N_{i_3} realizations (a_M, σ_L), and so on. Therefore, after at most $\sum_{i\in N_{U,0}} N_i$ realizations of (a_M, σ_L), beliefs have entered the ε' ball around some other stationary point $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N')$ such that $\boldsymbol{\lambda}_N'$ is strongly mixed accessible from $\boldsymbol{\lambda}_N$. Therefore, the probability of either $\langle \boldsymbol{\lambda}_{N,t} \rangle$ reaching a neighborhood $B_{\varepsilon'}(\boldsymbol{\lambda}_N')$ of some $\boldsymbol{\lambda}_N'$ that is strongly mixed accessible from $\boldsymbol{\lambda}_N$ or $\langle \boldsymbol{\lambda}_{I,t} \rangle$ leaving the neighborhood $B_\varepsilon(\boldsymbol{\lambda}_I^*)$ is bounded below by the probability of $\sum_{i \in N_{U,0}} N_i$ realizations of (a_M, σ_L), which is strictly positive. The argument for a type $i \in N_{U,\infty}$ is analogous.

Consider the graph $\mathcal{G}(\boldsymbol{\lambda}_I^*)$. We will choose an $\varepsilon(\boldsymbol{\lambda}_N)$ to correspond to each node $\boldsymbol{\lambda}_N$. At any terminal node $\boldsymbol{\lambda}_N$, define $\varepsilon(\boldsymbol{\lambda}_N) = \varepsilon$. For any node $\boldsymbol{\lambda}_N'$ that only has edges going to terminal nodes, by the above construction, there exists an $\varepsilon(\boldsymbol{\lambda}_N')$ such that if $\boldsymbol{\lambda}_{N,t} \in B_{\varepsilon(\boldsymbol{\lambda}_N')}(\boldsymbol{\lambda}_N')$, then with positive probability, either $\langle \boldsymbol{\lambda}_{N,t} \rangle$ reaches $B_{\varepsilon(\boldsymbol{\lambda}_N)}(\boldsymbol{\lambda}_N)$ for terminal node $\boldsymbol{\lambda}_N$ or $\langle \boldsymbol{\lambda}_{I,t} \rangle$ exits $B_\varepsilon(\boldsymbol{\lambda}_I^*)$. Repeat this process for each node in the graph.

Let $\tau_1 = \min\{t | \boldsymbol{\lambda}_{I,t} \notin B_\varepsilon(\boldsymbol{\lambda}_I^*)\}$. Then almost surely, $\tau_1 < \infty$ or $\langle \boldsymbol{\lambda}_{N,t} \rangle$ enters the neighborhood of a node on the graph constructed above infinitely often,

$$Pr(\tau_1 < \infty \text{ or for some } \boldsymbol{\lambda}_N \in \{0, \infty\}^{|N|},\ \boldsymbol{\lambda}_{N,t} \in B_{\varepsilon(\boldsymbol{\lambda}_N)}(\boldsymbol{\lambda}_N) \text{ i.o.}).$$

If $\langle \boldsymbol{\lambda}_{N,t} \rangle$ enters the neighborhood of a terminal node $\boldsymbol{\lambda}_N$ infinitely often, then $\boldsymbol{\lambda}_N \in \cap_{i \in N} \Lambda_i(\omega)$, so either $\boldsymbol{\lambda}_{N,t} \to \boldsymbol{\lambda}_N$ or $\langle \boldsymbol{\lambda}_{I,t} \rangle$ leaves $B_\varepsilon(\boldsymbol{\lambda}_I^*)$. Otherwise. $\langle \boldsymbol{\lambda}_{N,t} \rangle$ enters the neighborhood of some $\boldsymbol{\lambda}_N'$ that is strongly mixed accessible infinitely often. Since any path of this form ends at a terminal node, this implies that almost surely, either $\langle \boldsymbol{\lambda}_{N,t} \rangle$ converges or $\langle \boldsymbol{\lambda}_{I,t} \rangle$ leaves $B_\varepsilon(\boldsymbol{\lambda}_I^*)$. Therefore, the mixed outcome $\boldsymbol{\lambda}_I^*$ almost surely does not arise. □

Proof of Lemma 7′. Suppose $\Lambda(\omega)$ contains an agreement vector or maximally accessible disagreement vector and $\Lambda_M(\omega)$ is empty. Recall that $\mathcal{B}$ is the set of locally stable neighborhoods and $\mathcal{B}_U$ is the set of locally unstable neighborhoods. 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 11, there exists a finite sequence of actions and signals 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 Lemma 3, $\langle \boldsymbol{\lambda}_t \rangle$ almost surely leaves $\mathcal{B}_U$. If $\langle \boldsymbol{\lambda}_t \rangle$ enters the neighborhood of a mixed outcome $\boldsymbol{\lambda}_I$, by Lemma 6′, $\langle \boldsymbol{\lambda}_t \rangle$ almost surely leaves this neighborhood or converges to a locally stable point. By Lemma 2, $\langle \boldsymbol{\lambda}_t \rangle$ does not converge to a non-stationary belief. Therefore, almost surely, either $\langle \boldsymbol{\lambda}_t \rangle$ does not converge for all types or $\langle \boldsymbol{\lambda}_t \rangle$ converges to a learning outcome in $\Lambda(\omega)$.

Suppose with positive probability $\langle \boldsymbol{\lambda}_t \rangle$ exits and never re-enters the interior of the belief space, $[e^{-E}, e^E]^k$. Then either $\langle \boldsymbol{\lambda}_t \rangle$ enters the neighborhood of each mixed outcome where $|I| = 1$ infinitely often, in which case with probability one they visit a locally stable set, or there exists some i such that λ_i is constant across all neighborhoods that $\langle \boldsymbol{\lambda}_t \rangle$ enters. But then $\langle \boldsymbol{\lambda}_t \rangle$ is in the neighborhood of the mixed outcome λ_i, and by Lemma 6′, almost surely, $\langle \boldsymbol{\lambda}_t \rangle$ must leave this neighborhood or converge to a locally stable point. So almost surely, beliefs either return to $[e^{-E}, e^{E}]^k$ or converge to a locally stable point.

Let $\tau_2 \equiv \min\{\tau | \boldsymbol{\lambda}_t \in \mathcal{B} \ \forall t > \tau\}$ be the time at which $\langle \boldsymbol{\lambda}_t \rangle$ enters $\mathcal{B}$ and never leaves. From Lemma 3, $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}_\infty | \tau_2 < \infty) = 1$, where $\boldsymbol{\lambda}_\infty$ is a random variable with $\text{supp}(\boldsymbol{\lambda}_\infty) \subset \Lambda(\omega)$. Suppose $\tau_2 = \infty$. Then 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. This is a contradiction, as we supposed $\tau_2 = \infty$. Therefore, $Pr(\tau_2 < \infty) = 1$. This implies $Pr(\boldsymbol{\lambda}_t \to \boldsymbol{\lambda}_\infty) = 1$, where $\boldsymbol{\lambda}_\infty$ is a random variable with $\text{supp}(\boldsymbol{\lambda}_\infty) \subset \Lambda(\omega)$. □

A.3 Proof of Theorem 2

Suppose $\omega = R$ and suppose the mixed outcome $\lambda_2 = 0 \in \Lambda_M$. As in the proof of Lemma 3, we can construct neighborhoods $(0, e^{-E}]^2$ and $[e^{E}, \infty) \times (0, e^{-E}]$ such that in each of these neighborhoods, there exists an i.i.d. process that bounds θ_1's updates above as long as beliefs remain in the neighborhood and almost surely converges to zero, and a process that bounds θ_2's updates below (above) in the nbhd of 0 (∞) and eventually leaves the nbhd.

Consider the interior of the belief space, $[e^{-E}, e^{E}]^2$. This space can be partitioned into finitely many closed, convex sets $D_1, D_2, \ldots D_N$ where $\gamma_2(\cdot, R)$ is continuous on the interior of these sets. Consider the set D_j and define the function $\hat{\gamma}_{D_j} : D_j \to \mathbb{R}$ as

$$\hat{\gamma}_{i,D_j}(\boldsymbol{\lambda}) \equiv \begin{cases} \gamma_2(\boldsymbol{\lambda}, R) \text{ if } \boldsymbol{\lambda} \in \text{ interior of } D_j \\ \lim_{x \to \boldsymbol{\lambda}} \gamma_2(x, R) \text{ otherwise.} \end{cases}$$

This is a continuous function. So, for each $(\lambda, 0) \in D_j$, we can construct an open, convex set $B(\lambda, 0)$ such that if $\boldsymbol{\lambda}_t$ is in this set, then $\log \lambda_{2,t+1} - \log \lambda_{2,t}$ is bounded above by

$$g_j(a, \sigma) \equiv \sup_{\boldsymbol{\lambda} \in B(\lambda,0)} \log \frac{\hat{\psi}_2(a, \sigma | L, \boldsymbol{\lambda})}{\hat{\psi}_2(a, \sigma | R, \boldsymbol{\lambda})}.$$

Let

$$\bar{g}_j \equiv \max_{(a,\sigma) \in \mathcal{A} \times \Sigma} g_j(a, \sigma).$$

Define the process

$$\xi_{D_j,t+1} = \xi_{D_j,t} + g_j(a(\theta_t, s_t, (\lambda, 0)), \sigma_t),$$

when (θ_t, s_t) is such that $a(\theta_t, s_t, \boldsymbol{\lambda}) = a(\theta_t, s_t, (\lambda, 0))$ for all beliefs $\boldsymbol{\lambda} \in D_j$ (note this includes all autarkic types), and

$$\xi_{D_j,t+1} = \xi_{D_j,t} + \bar{g}_j$$

otherwise. When $\omega = R$, let $\psi_j(a,\sigma)$ be the probability of (a,σ) in the former event and let $\bar{\psi}_j$ be the probability of the latter event. By construction, $\log\lambda_{2,t+1} - \log\lambda_{2,t} < \xi_{D_j,t+1} - \xi_{D_j,t}$ if $\boldsymbol{\lambda}_t \in D_j$. Moreover, choose D_j sufficiently small so that

$$\bar{\psi}_j\bar{g}_j + \sum_{(a,\sigma)\in\mathcal{A}\times\Sigma} \psi_j(a,\sigma)g_j(a,\sigma) < 0. \tag{25}$$

As in Lemma 3, this sequence converges to $-\infty$ almost surely. But now the process that bounds the updates changes as the likelihood ratio moves across the state space even if λ_2 stays in a neighborhood of 0, so this is insufficient to conclude that λ_2 converges.

By compactness, we can find a finite collection of open sets $B_{D_j,1}\ldots B_{D_j,n}$ that contain all $(\lambda,0)\in D_j$. Since this procedure can be done for each D_j, there exists a disjoint, finite collection of sets $\mathcal{C} = (C_j)_{i=1}^N$ and an $\varepsilon > 0$ such that these sets contain each point $\boldsymbol{\lambda} \in [e^{-E}, e^{E}]\times[0,\varepsilon]$, each set is contained in exactly one $B_{D_j,j}$ for some i and j, and each discontinuity point $\boldsymbol{\lambda}$ is contained in a subset of the D_j where $\gamma_2(\boldsymbol{\lambda}, R) = \hat{\gamma}_{D_j}(\boldsymbol{\lambda})$. Within each set in this cover, $\log\lambda_{2,t} - \log\lambda_{2,1}$ is bounded above by an i.i.d. process $\xi_{C_j,t} - \xi_{C_j,t-1}$, and with positive probability, $\sup_{C_j} E(\xi_{C_j,t} - \xi_{C_j,t-1}) < 0$. It remains to show that with positive probability, λ_2 remains below ε and converges. For each C_j, there exists a sequence of actions and public signals such that $\xi_{C_j} \to -\infty$ and $\sup\xi_t < \log\varepsilon/2N$. Let N_j be the set of realizations of the process $(\xi_{C_j,t} - \xi_{C_j,t-1})_{t=1}^{\infty}$ for set C_j. Let Ξ_j^T be the set of these sequences truncated after the first T terms.

Let $\tau_{j,k}$ be the kth time the $\boldsymbol{\lambda}$ process enters C_j, and let $\tau_{j,0} = 0$ and $\xi_{C_j,0} = 0$. Let n be the n that satisfies where $\tau_{j,n} \leq t$ and $\tau_{j,n+1} > t$ if it exists. Let A_t be the event that for all C_j, the sequence $N_j^T = ((\xi_{C_j,\tau_{j,k}+1} - \xi_{j,\tau_{j,k}}))_{k=1}^n$ is in $\Xi_j^{\tau_{j,n}}$ for each j. Elements in this sequence bound the change in the log-likelihood ratio above at each time $\tau_{j,k}$ when the likelihood ratio is in set C_j. Finally let P_{j,N_j^T} be the probability that the process realization of the process $\xi_{C_j,s}$ $s \geq T$ satisfies $(N_j^T, (\xi_{C_j,s+1} - \xi_{C_j,s})_{s=T}^{\infty})) \in \Xi_j$, and let P_j

be the probability that $(\xi_{C_j,s+1} - \xi_{C_j,s})_{s=T}^{\infty} \in \Xi_j$. Let c_t be the set C that $\boldsymbol{\lambda}_t$ is in. Then

$$\begin{aligned}
Pr(A_2|c_1) &= Pr(A_1|c_1)Pr(A_2|A_1,c_1) \\
&= Pr(A_1|c_1)E(Pr(A_2|\xi_{c_1,1},A_1,c_1)|A_1,c_1) \\
&= Pr(A_1|c_1)E[Pr(c_1 \neq c_2|c_1,A_1,\xi_{c_1,1})Pr(A_2|c_1,A_1,\xi_{c_1,1},c_1 \neq c_2) \\
&+ Pr(c_1 = c_2|c_1,A_1,\xi_1)Pr(A_2|c_1,A_1,\xi_{c_1,1},c_1 = c_2)|A_1,c_1] \\
&\geq Pr(A_1|c_1)E[Pr(c_1 \neq c_2|c_1) \\
&\quad \sum_{c \neq c_1} Pr(c|c_1,A_1,\xi_{c_1},c \neq c_1)P_c + Pr(c_1 = c_2|c_1,A_1,\xi_{c_1})P_{c_1,N_c^2}|A_1,c_1] \\
&\geq Pr(A_1|c_1)E(P_{c_1,N_c^2}\prod_{c \neq c_1} P_{c,1}|A_1,c_1) = \prod_{j=1}^{N} P_j > 0
\end{aligned}$$

where the first inequality follows from the fact that at time t, given the current neighborhood beliefs are in c, the probability that the next realization is consistent with the desired sequence is at least the probability that all subsequent realizations of ξ are in that sequence. Now suppose we start at time t, and condition on the current set and the past realizations of the sequences. Then

$$\begin{aligned}
Pr(A_{t+1}|(N_j^t)_{j=1}^n,c_t) &= Pr(A_t|(N_j^t)_{j=1}^N,c_t)Pr(A_{t+1}|A_t,(N_j^t)_{j=1}^N,c_t) \\
&= Pr(A_t|N_j^t)_{j=1}^n,c_t)E[Pr(c_t \neq c_{t+1}|(N_j^{t+1})_{j=1}^N,c_t,A_t) \\
&\quad *Pr(A_{t+1}|(N_j^{t+1})_{j=1}^N,c_t,A_t,c_t \neq c_{t+1}) \\
&+ Pr(c_t = c_{t+1}|(N_j^{t+1})_{j=1}^N,c_t,A_t) \\
&\quad *Pr(A_{t+1}|(N_j^{t+1})_{j=1}^N,c_t,A_t,c_t = c_{t+1})|A_t,c_t,(N_j^t)_{j=1}^N] \\
&\geq \Pr(A_t|(N_c)_{j=1}^N,c_t)E(P_{c_t,N_{c_t}^{t+1}}\prod_{j \neq c_t} P_{j,N_j^t}|A_t,c_t,(N_j^t)_{j=1}^N) \\
&= \prod_{j=1}^{N} P_{j,N_j^t},
\end{aligned}$$

where the inequality follows from similar logic to the previous case. Conditional on knowing the current set c_t and the previous realizations of the sequence when $\boldsymbol{\lambda}_t$ was in c_t, the current realization being consistent does not depend on anything else. Moreover, the current realization being consistent is bounded below by the probability that all future realizations are consistent. Finally

$$\begin{aligned} Pr(A_{t+1}) &= E(Pr(A_{t+1}|(N_j^t)_{j=1}^N c_t)) \geq E(\Pi_{j=1}^N P_{j,N_j^t}) \\ &= E(E(\Pi_{j=1}^N P_{j,N_j^t}|(N_j^{t-1})_{j=1}^N, c_{t-1})) \\ &= E(P_{j,N_j^{t-1}}) \dots \prod_{j=1}^n P_j > 0. \end{aligned}$$

Therefore, $\lim_{T\to\infty} P(A_T) > 0$. By the dominated convergence theorem, $\lim_{T\to\infty} Pr(A_T) = Pr(A)$. Moreover, at any time T, if the event A_T has occurred and $\lambda_{2,1} < \varepsilon/2$, the likelihood ratio updates are bounded above by

$$\log \lambda_{2,T} - \log \lambda_{2,1} \leq \sum_{t=1}^{T-1} (\xi_{c_t,t+1} - \xi_{c_t,t}) < N\varepsilon/2N = \varepsilon/2.$$

So λ_2 never leaves the ε-ball.

Finally, since the mixed outcome $\lambda_2 = 0 \in \Lambda_M$, beliefs cannot converge to $(0,0)$ or $(\infty,0)$. Otherwise, beliefs would eventually enter either $(0,e^{-E}]^2$ or $[e^E,\infty) \times (0,e^{-E}]$ and never leave. But there exists a process that bounds type θ_1's belief updates below and leaves the neighborhood almost surely, which is a contradiction. Therefore, the mixed outcome occurs with positive probability. □

A.4 Proofs of Theorems 3, 4 and 5 (Robustness)

Proof of Theorem 3. Assume Assumptions 1, 2, 3 and 4 and suppose $\omega = R$. For any sociable type $\theta_i \in \Theta_S$, the mapping $\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda}) \mapsto \gamma_i(\boldsymbol{\lambda},R)$ is continuous. By the concavity of the log operator, $\gamma_i(\boldsymbol{\lambda},R)$ is negative when $||\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda}) - \psi(a,\sigma|R,\boldsymbol{\lambda})|| = 0$. Therefore, there exists a $\delta > 0$ such that if $||\hat{\psi}_i(a,\sigma|R,\boldsymbol{\lambda}) - \psi(a,\sigma|R,\boldsymbol{\lambda})|| < \delta$ for $(a,\sigma,\boldsymbol{\lambda}) \in \mathcal{A} \times \Sigma \times \{0,\infty\}^k$ and $\theta_i \in \Theta_S$, then $\gamma_i(\boldsymbol{\lambda},R) < 0$ for all $\boldsymbol{\lambda} \in \{0,\infty\}^k$ and $\theta_i \in \Theta_S$. Therefore, any locally stable point must have $\lambda_i = 0$ for each sociable type. Therefore, 0^k is the unique locally stable point.

We also need to show that $\Lambda_M(R)$ is empty, i.e. all mixed outcomes are reducible. Consider the mixed outcome $\boldsymbol{\lambda}_I^*$ with convergent types I and non-convergent types $N \equiv \Theta_S \setminus I$. For any node $\boldsymbol{\lambda}_N$ in the graph $\mathcal{G}(\boldsymbol{\lambda}_I^*)$ (as defined in Definition 13), it follows from the choice of δ that for each $i \in N$, $(\boldsymbol{\lambda}_I^*, \boldsymbol{\lambda}_N) \in \Lambda_i(R)$ iff $\lambda_i = 0$. Therefore, each $\boldsymbol{\lambda}_N'$ that is mixed accessible from $\boldsymbol{\lambda}_N$ has fewer $i \in N$ with $\lambda_i' = 0$. Therefore, each path terminates at $0^{|N|}$ and the graph has no cycles, i.e. $\boldsymbol{\lambda}_I^*$ is reducible. Therefore, $\Lambda_M(R)$ is empty. By Theorem 1, if $\Lambda(R) = \{0^k\}$ and $\Lambda_M(R)$ is empty, then the likelihood ratio almost surely converges to 0^k and learning is complete.

Similar logic holds for $\omega = L$. □

Proof of Theorem 4. Assume Assumptions 1, 2, 3 and 4, and fix state ω. For any sociable type θ_i, the mapping $(\hat{\pi}_i, \hat{F}_i^\omega, \hat{G}_i^\omega) \mapsto \hat{\psi}_i(a, \sigma|\omega, \boldsymbol{\lambda})$ is continuous. By continuity, for any $\delta_2 > 0$, there exists a $\delta > 0$ such that if $||\hat{\pi}_i - \pi|| < \delta$, $||\hat{F}_i^\omega - F^\omega|| < \delta$ and $||\hat{G}_i^\omega - G^\omega|| < \delta$ for all $\theta_i \in \Theta_S$, then $|\hat{\psi}_i(a, \sigma|\omega, \boldsymbol{\lambda}) - \psi(a, \sigma|\omega, \boldsymbol{\lambda})| < \delta_2$ for all $(a, \sigma) \in \mathcal{A} \times \Sigma$, $\boldsymbol{\lambda} \in \{0, \infty\}^k$ and $\theta_i \in \Theta_S$. Choose δ_2 sufficiently small so that Theorem 3 holds. Similar logic holds for $\omega = L$. □

Proof of Theorem 5. Assume Assumptions 1, 2, 3 and 4, and suppose $\omega = R$. Under these assumptions, if $\langle \lambda_{i,t} \rangle$ converges for any type θ_i, then the support of the limit belief λ_∞ is a subset of $\{0, \infty\}$, i.e. $\text{supp}(\lambda_\infty) \subset \{0, \infty\}$. Let θ_i be a correctly specified type. Then its subjective probability of each action is equal to the true probability, $\hat{\psi}_i = \psi$. Therefore, given $\omega = R$, $\langle \lambda_{i,t} \rangle$ is a martingale for any learning environment $\{\Theta, \pi, F^R, F^L\}$. By the Martingale Convergence Theorem, $\langle \lambda_{i,t} \rangle$ converges almost surely to a limit random variable λ_∞ with $\text{supp}(\lambda_\infty) \subset [0, \infty)$. This rules out incorrect and cyclical learning. Therefore, zero is the only candidate limit belief for the correctly specified type, $\text{supp}(\lambda_\infty) = \{0\}$, and it must be that $\lambda_{i,t} \to 0$ almost surely. □

A.5 Proofs from Section 5 (Applications)

A.5.1 Proofs from Section 5.1 (Level-k/Cognitive Hierarchy)

The following lemma implies that a disagreement outcome arises with positive probability iff it is in $\Lambda(\omega)$.

Lemma 14. *In the level-k/cognitive hierarchy model, both disagreement outcomes* $(0, \infty)$ *and* $(\infty, 0)$ *are maximally accessible for all distributions of types* $\pi \in \Delta(\Theta)$ *and* $q \in (0, 1)$.

Proof. At $\boldsymbol{\lambda} = (0, 0)$, type θ_2 perceives R actions as stronger evidence of state R than type θ_3,

$$\frac{\hat{\psi}_2(R|L, (0,0))}{\hat{\psi}_2(R|R, (0,0))} = \frac{F^L(1/2)}{F^R(1/2)} < \frac{q + (1-q)F^L(1/2)}{q + (1-q)F^R(1/2)} = \frac{\hat{\psi}_3(R|L, (0,0))}{\hat{\psi}_3(R|R, (0,0))},$$

and both types perceive L actions in the same way,

$$\frac{\hat{\psi}_2(L|L, (0,0))}{\hat{\psi}_2(L|R, (0,0))} = \frac{\hat{\psi}_3(L|L, (0,0))}{\hat{\psi}_3(L|R, (0,0))} = \frac{1 - F^L(1/2)}{1 - F^R(1/2)}. \tag{26}$$

Therefore, $\theta_3 \succeq_{(0,0)} \theta_2$. From Definition 5, this implies that $(0, \infty)$ is maximally accessible. At $\boldsymbol{\lambda} = (\infty, \infty)$, type θ_2 perceives L actions as stronger evidence of state L than type θ_3,

$$\frac{\hat{\psi}_2(L|L,(\infty,\infty))}{\hat{\psi}_2(L|R,(\infty,\infty))} = \frac{1-F^L(1/2)}{1-F^R(1/2)} > \frac{q+(1-q)(1-F^L(1/2))}{q+(1-q)(1-F^R(1/2))} = \frac{\hat{\psi}_3(L|L,(\infty,\infty))}{\hat{\psi}_3(L|R,(\infty,\infty))},$$

and both types perceive R actions in the same way,

$$\frac{\hat{\psi}_2(R|L,(\infty,\infty))}{\hat{\psi}_2(R|R,(\infty,\infty))} = \frac{\hat{\psi}_3(R|L,(\infty,\infty))}{\hat{\psi}_3(R|R,(\infty,\infty))} = \frac{F^L(1/2)}{F^R(1/2)}. \tag{27}$$

Therefore, $\theta_2 \succeq_{(\infty,\infty)} \theta_3$. From Definition 5, this implies that $(\infty, 0)$ is maximally accessible. □

Proof of Proposition 1. As $q \to 1$, as argued in the text, only disagreement outcomes can be locally stable. So either $\Lambda(\omega) = \emptyset$ or $\Lambda(\omega) \subseteq \{(0,\infty),(\infty,0)\}$. By Lemma 14, if a disagreement outcome is locally stable, then it is globally stable. We must also rule out mixed outcomes. Suppose $\omega = R$ and consider the four possible mixed outcomes.

1. Consider the mixed outcome in which λ_2 does not converge and $\lambda_3 \to 0$. By the concavity of the log operator,

$$F^R(1/2)\log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) + (1-F^R(1/2))\log\left(\frac{1-F^L(1/2)}{1-F^R(1/2)}\right) < 0.$$

Therefore, since $\frac{F^L(1/2)}{F^R(1/2)} < 0$,

$$\begin{aligned}\gamma_2((0,0),R) &= (\pi(\theta_1)F^R(1/2) + \pi(\theta_2) + \pi(\theta_3))\log\left(\frac{F^L(1/2)}{F^R(1/2)}\right)\\ &+ \pi(\theta_1)(1-F^R(1/2))\log\left(\frac{1-F^L(1/2)}{1-F^R(1/2)}\right) < 0.\end{aligned}$$

Therefore, $(0,0) \in \Lambda_2(R)$. By the definition of $\Lambda_M(R)$, this implies that $\lambda_3 = 0 \notin \Lambda_M(R)$ and this mixed learning outcome almost surely does not arise.

2. Consider the mixed outcome in which λ_2 does not converge and $\lambda_3 \to \infty$. This outcome is in $\Lambda_M(R)$ if $(\infty,\infty) \notin \Lambda_2(R)$ and $(0,\infty) \notin \Lambda_2(R)$, which is equivalent to $\gamma_2((\infty,\infty),R) < 0$ and $\gamma_2((0,\infty),R) > 0$. However, $\gamma_2((\lambda_2,\infty),R)$ is increasing in λ_2, so this is not possible. Therefore, $\lambda_3 = \infty \notin \Lambda_M(R)$ and this mixed learning outcome almost surely does not arise.

3. Consider the mixed outcome in which $\lambda_2 \to 0$ and λ_3 does not converge. This outcome is in $\Lambda_M(R)$ if $(0,0) \notin \Lambda_3(R)$ and $(0,\infty) \notin \Lambda_3(R)$. From the text, we

know that $(0,\infty) \in \Lambda_3(R)$. Therefore, $\lambda_2 = 0 \notin \Lambda_M(R)$ and this mixed learning outcome almost surely does not arise.

4. Consider the mixed outcome in which $\lambda_2 \to \infty$ and λ_3 does not converge. This outcome is in $\Lambda_M(R)$ if $(\infty,0) \notin \Lambda_3(R)$ and $(\infty,\infty) \notin \Lambda_3(R)$. From the text, we know that $(\infty,0) \in \Lambda_3(R)$. Therefore, $\lambda_2 = \infty \notin \Lambda_M(R)$ and this mixed learning outcome almost surely does not arise.

Therefore, $\Lambda_M(R) = \emptyset$ and mixed outcomes almost surely do not arise if the state is R. Similar logic rules out mixed outcomes if the state is L.

Given $\Lambda_M(\omega) = \emptyset$ and both disagreement outcomes are maximally accessible, by Theorem 1, $\Lambda(\omega)$ determines the set of asymptotic learning outcomes. Either $\Lambda(\omega) = \emptyset$, in which case learning is cyclical for both types, or $\Lambda(\omega) \subseteq \{(0,\infty),(\infty,0)\}$ and $\Lambda(\omega) \neq \emptyset$, in which case beliefs almost surely converge to a limit random variable with support $\Lambda(\omega)$.

Part 1: As $\pi(\theta_3) \to 0$,

$$\begin{aligned}\gamma_2((0,\infty),R) \to \quad & (\pi(\theta_1)F^R(1/2) + \pi(\theta_2)) \log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) \\ + \quad & \pi(\theta_1)(1 - F^R(1/2)) \log\left(\frac{1 - F^L(1/2)}{1 - F^R(1/2)}\right) < 0,\end{aligned}$$

so $(0,\infty) \in \Lambda_2(R)$, and as $\pi(\theta_3) \to 1$,

$$\gamma_2((0,\infty),\omega) \to \log\left(\frac{1 - F^L(1/2)}{1 - F^R(1/2)}\right) > 0,$$

so $(0,\infty) \notin \Lambda_2(\omega)$. Moreover $\pi \mapsto \gamma_2$ is continuous and $\gamma_2((0,\infty),\omega)$ is increasing in $\pi(\theta_3)$. Therefore, there exists an interior cut-off above which $(0,\infty) \notin \Lambda(R)$. Similarly, as $\pi(\theta_3) \to 1$,

$$\gamma_2((\infty,0),\omega) \to \log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) < 0,$$

so $(\infty,0) \notin \Lambda_2(\omega)$. Therefore, there exists a cut-off above which $(\infty,0) \notin \Lambda(\omega)$. From the text, we know that $(0,0) \notin \Lambda(\omega)$ and $(\infty,\infty) \notin \Lambda(\omega)$. Therefore, there exists an interior cutoff $\bar{\pi}_3 \in (0,1)$ such that if $\pi(\theta_3) > \bar{\pi}_3$, then $\Lambda(R) = \emptyset$. The case for $\omega = L$ is analogous.

Part 2: From the text, we know that for any π, $\Lambda_3(\omega) = \{(0,\infty),(\infty,0)\}$. As $\pi(\theta_2) \to 0$,

$$\begin{aligned}\gamma_2((\infty,0),R) \to \quad & (\pi(\theta_1)F^R(1/2) + \pi(\theta_3)\log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) \\ + \quad & \pi(\theta_1)(1-F^R(1/2))\log\left(\frac{1-F^L(1/2)}{1-F^R(1/2)}\right) < 0,\end{aligned}$$

so $(\infty,0) \notin \Lambda_2(R)$, and as $\pi(\theta_2) \to 1$,

$$\begin{aligned}\gamma_2((0,\infty),\omega) &\to \log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) < 0 \\ \gamma_2((\infty,0),\omega) &\to \log\left(\frac{1-F^L(1/2)}{1-F^R(1/2)}\right) > 0,\end{aligned}$$

so $\Lambda_2(\omega) = \{(0,\infty),(\infty,0)\}$. Further, $\gamma_2((0,\infty),\omega)$ is decreasing in $\pi(\theta_2)$ and $\gamma_2((\infty,0),\omega)$ is increasing in $\pi(\theta_2)$. Therefore, there exists an interior cutoff $\bar{\pi}_2 \in (0,1)$ such that if $\pi(\theta_2) > \bar{\pi}_2$, then $\Lambda(R) = \{(0,\infty),(\infty,0)\}$. The case for $\omega = L$ is analogous.

Part 3: As argued above, $\gamma_2((0,\infty),\omega)$ is strictly increasing and $\gamma_2((\infty,0),\omega)$ is strictly decreasing in $\pi(\theta_3)$. Therefore, for fixed $\pi(\theta_1)$ or $\pi(\theta_2)$, it is sufficient to characterize the unique value of $\pi(\theta_3)$ at which $\gamma_2((0,\infty),\omega) = 0$ or $\gamma_2((\infty,0),\omega) = 0$. Fix $\pi(\theta_1)$ and consider $\omega = R$. At $\pi(\theta_2) = \pi(\theta_3) = 0$,

$$\gamma_2((\infty,0),R) = F^R(1/2)\log\left(\frac{F^L(1/2)}{F^R(1/2)}\right) + (1-F^R(1/2))\log\left(\frac{1-F^L(1/2)}{1-F^R(1/2)}\right) < 0.$$

Since $\gamma_2((\infty,0),R)$ is decreasing in $\pi(\theta_3)$, $(\infty,0)$ is never locally stable when $\pi(\theta_2) = 0$. On the other hand, at $\pi(\theta_2) = \pi(\theta_3) = 0$, $\gamma_2((0,\infty),R) < 0$, and at $\pi(\theta_2) = 0$, $\pi(\theta_3) = 1$, $\gamma_2((0,\infty),R) > 0$. Therefore, there exists an interior cutoff $\bar{\pi}_3^R \in (0,1)$ such that at $\pi(\theta_2) = 0$ and $\pi(\theta_3) = \bar{\pi}_3^R$, $\gamma_2((0,\infty),R) = 0$. If $\omega = L$, this condition reverses, so $(0,\infty)$ is never locally stable, while there is an interior cutoff $\bar{\pi}_3^L \in (0,1)$ such that at $\pi(\theta_2) = 0$ and $\pi(\theta_3) = \bar{\pi}_3^L$, $\gamma_2((\infty,0),L) = 0$. Moreover, as $\pi(\theta_1) \to 0$, the cutoff for $(0,\infty)$ is

$$\pi(\theta_3) = \frac{\log\frac{F^R(1/2)}{F^L(1/2)}}{\log\frac{F^R(1/2)}{F^L(1/2)} - \log\frac{1-F^R(1/2)}{1-F^L(1/2)}},$$

while at $(\infty,0)$,

$$\pi(\theta_3) = \frac{\log\frac{1-F^R(1/2)}{1-F^L(1/2)}}{\log\frac{1-F^R(1/2)}{1-F^L(1/2)} - \log\frac{F^R(1/2)}{F^L(1/2)}}.$$

Finally, note that $\gamma_2(\boldsymbol{\lambda}, \omega)$ is linear in $\pi(\theta_3)$. Therefore, if

$$\frac{\log \frac{F^R(1/2)}{F^L(1/2)}}{\log \frac{F^R(1/2)}{F^L(1/2)} - \log \frac{1-F^R(1/2)}{1-F^L(1/2)}} \geq \frac{\log \frac{1-F^R(1/2)}{1-F^L(1/2)}}{\log \frac{1-F^R(1/2)}{1-F^L(1/2)} - \log \frac{F^R(1/2)}{F^L(1/2)}},$$

then in the simplex, the line at which $\gamma_2((0,\infty), R) = 0$ lies above the line at which $\gamma_2((\infty,0), R) = 0$ (and the reverse for $\omega = L$) and $(\infty, 0) \in \Lambda(R) \Rightarrow (0,\infty) \in \Lambda(R)$. Otherwise, the lines cross exactly once, and there exists a cutoff $\bar{\pi}_1$ such that if $\pi(\theta_1) \geq \bar{\pi}_1$, then $(\infty, 0) \in \Lambda(R) \Rightarrow (0,\infty) \in \Lambda(R)$, and if $\pi(\theta_1) \leq \bar{\pi}_1$, then $(\infty, 0) \in \Lambda(R) \Rightarrow (0,\infty) \in \Lambda(R)$. If the inequality is switched, then analogous properties hold with the states switched.

□

Proof of Proposition 2. Suppose $\omega = R$. Let $\alpha \equiv F^R(1/2)$ be the probability a level-1 type plays action R. Consider the level-2 type. Since $\alpha > 1/2$,

$$\begin{aligned}
\gamma_2((0,0), R) &= -\left(\frac{1+2\alpha}{3}\right)\log\left(\frac{\alpha}{1-\alpha}\right) < 0 \\
\gamma_2((\infty,0), R) &= \left(\frac{1-2\alpha}{3}\right)\log\left(\frac{\alpha}{1-\alpha}\right) < 0 \\
\gamma_2((0,\infty), R) &= \left(\frac{1-2\alpha}{3}\right)\log\left(\frac{\alpha}{1-\alpha}\right) < 0 \\
\gamma_2((\infty,\infty), R) &= \left(\frac{3-2\alpha}{3}\right)\log\left(\frac{\alpha}{1-\alpha}\right) > 0.
\end{aligned}$$

Therefore, $\Lambda_2(R) = \{(0,0), (0,\infty), (\infty,\infty)\}$. Consider the level-3 type.

$$\begin{aligned}
\gamma_3((\infty,\infty), R) &= \left(\frac{\alpha}{3}\right)\log\left(\frac{1-\alpha}{\alpha}\right) + \left(\frac{3-\alpha}{3}\right)\log\left(\frac{q+(1-q)\alpha}{q+(1-q)(1-\alpha)}\right) \\
\gamma_3((0,\infty), R) &= \left(\frac{1+\alpha}{3}\right)\log\left(\frac{q+(1-q)(1-\alpha)}{q+(1-q)\alpha}\right) + \left(\frac{2-\alpha}{3}\right)\log\left(\frac{\alpha}{1-\alpha}\right) \\
\gamma_3((0,0), R) &= \left(\frac{2+\alpha}{3}\right)\log\left(\frac{q+(1-q)(1-\alpha)}{q+(1-q)\alpha}\right) + \left(\frac{1-\alpha}{3}\right)\log\left(\frac{\alpha}{1-\alpha}\right).
\end{aligned}$$

If $\gamma_3((\infty,\infty), R) > 0$, then $(\infty,\infty) \in \Lambda(R)$. From these expressions, $\gamma_3((\infty,\infty), R)$ is positive at $q = 0$, decreasing in q, and negative at $q = 1$. Therefore, there exists a q_2 such that for $q < q_2$, $(\infty,\infty) \in \Lambda(R)$, and for $q > q_2$, $(\infty,\infty) \notin \Lambda(R)$. If $\gamma_3((0,\infty), R) > 0$, then $(0,\infty) \in \Lambda(R)$ and if $\gamma_3((0,0), R) < 0$, then $(0,0) \in \Lambda(R)$. The expressions $\gamma_3((0,0), R) < \gamma_3((0,\infty), R)$ are both negative at $q = 0$, increasing in q, and positive at $q = 1$. Therefore, there exists $q_1 < q_3$ such that $(0,0) \in \Lambda(R)$ for $q < q_3$ and

$(0,0) \notin \Lambda(R)$ for $q > q_3$, while $(0,\infty) \notin \Lambda(R)$ for $q < q_1$ and $(0,\infty) \in \Lambda(R)$ for $q > q_1$. This yields the characterization of $\Lambda(R)$ as a function of q.

It immediately follows from Theorem 1 that the agreement outcomes $(0,0)$ and (∞,∞) are globally stable iff they are in $\Lambda(R)$. By Lemma 14, both disagreement outcomes are maximally accessible. Therefore, $(0,\infty)$ is globally stable iff $(0,\infty) \in \Lambda(R)$.

Finally we have to rule out mixed outcomes. In the region where both correct learning and incorrect learning are locally stable (parts 1 and 2), it immediately follows that $\Lambda_M(R)$ is empty and mixed outcomes almost surely do not arise. Given $\gamma_2((\infty,0),R) < 0$ and

$$\begin{aligned} \gamma_3((\infty,0),R) &= \left(\frac{1+\alpha}{3}\right)\log\left(\frac{1-\alpha}{\alpha}\right) + \left(\frac{2-\alpha}{3}\right)\log\left(\frac{q+(1-q)\alpha}{q+(1-q)(1-\alpha)}\right) \\ &< \gamma_3((\infty,\infty),R) < 0, \end{aligned}$$

$\lambda_2 = \infty \notin \Lambda_M(R)$. If disagreement and correct learning are locally stable (part 3), then $\lambda_2 = \infty$ is the only candidate mixed outcome and therefore, $\Lambda_M(R)$ is empty. If only disagreement is locally stable (part 4), we also have to rule out $\lambda_3 = 0$. But since $(0,0) \notin \Lambda(R)$, $\gamma_3((0,0),R) > 0$. Also, $\gamma_2((0,0),R) < 0$. Therefore, $\lambda_3 = 0 \notin \Lambda_M(R)$. Therefore, $\Lambda_M(R)$ is empty for all $q \in (0,1]$.

Given this characterization, by Theorem 1, beliefs almost surely converge to a limit random variable $\boldsymbol{\lambda}_\infty$ with $\operatorname{supp} \boldsymbol{\lambda}_\infty = \Lambda(R)$. □

A.5.2 Proofs from Section 5.2 (Partisan Bias)

Proof of Proposition 3. Both partisan and nonpartisan types believe that share α of agents are autarkic. Partisan types think these autarkic types are also partisan, while nonpartisan types think these autarkic types are also nonpartisan. Let $x_1^\omega(\nu) \equiv F^\omega(0.5^{1/\nu})$ be the probability that the partisan autarkic type plays action R and $x_2^\omega \equiv F^\omega(0.5)$ be the probability that the nonpartisan autarkic type plays action R in state ω. Then $x_1^R(\nu) \leq x_2^R$ and $x_1^L(\nu) \leq x_2^L$ for all $\nu \in (0,1)$, since partisan types slant information in favor of state L. Moreover, action R occurs more often in state R, so $x_2^R > x_2^L$ and $x_1^R(\nu) > x_1^L(\nu)$ for all $\nu \in (0,1)$. Nonpartisan types believe that autarkic types play action R with probability x_2^ω, and partisan types believe that autarkic types play action R with probability $\hat{F}_1^\omega(0.5^{1/\nu}) = F^\omega(0.5) = x_2^\omega$.

Let $\gamma_1^{\nu,q}(\boldsymbol{\lambda},\omega)$ be the value of $\gamma_1(\boldsymbol{\lambda},\omega)$ in the model with partisan bias level ν and frequency q, with an analogous definition of $\gamma_2^{\nu,q}(\boldsymbol{\lambda},\omega)$. Since partisan and nonpartisan sociable types have the same subjective probability of each action, beliefs can never separate. Therefore, asymptotic disagreement and mixed learning is not possible. Additionally, $\gamma_1^{\nu,q} = \gamma_2^{\nu,q}$, and therefore, we only need to check the sign of $\gamma_1^{\nu,q}((0,0),\omega)$ to determine whether $(0,0)$ is locally stable, and the sign of $\gamma_1^{\nu,q}((\infty,\infty),\omega)$ to determine whether (∞,∞) is locally stable. Recall that global stability immediately follows for agreement outcomes.

Suppose $\omega = R$. To determine whether $(\infty,\infty) \in \Lambda(R)$ at (ν,q), we need to determine the sign of

$$\begin{aligned}\gamma_1^{\nu,q}((\infty,\infty),R) &= \psi^{\nu,q}(R|R,(\infty,\infty))\log\frac{\hat{\psi}_1(R|L,(\infty,\infty))}{\hat{\psi}_1(R|R,(\infty,\infty))} \\ &+ \psi^{\nu,q}(L|R,(\infty,\infty))\log\frac{\hat{\psi}_1(L|L,(\infty,\infty))}{\hat{\psi}_1(L|R,(\infty,\infty))},\end{aligned}$$

where

$$\begin{aligned}\hat{\psi}_1(R|\omega,(\infty,\infty)) &= \alpha x_2^\omega \\ \hat{\psi}_1(L|\omega,(\infty,\infty)) &= \alpha(1-x_2^\omega)+1-\alpha \\ \psi^{\nu,q}(R|R,(\infty,\infty)) &= \alpha q x_1^R(\nu)+\alpha(1-q)x_2^R \\ \psi^{\nu,q}(L|R,(\infty,\infty)) &= \alpha q(1-x_1^R(\nu))+\alpha(1-q)(1-x_2^R)+1-\alpha.\end{aligned}$$

If $\nu = 1$, then $x_1^R(1) = x_2^R$, so

$$\psi^{1,q}(R|R,(\infty,\infty)) = \hat{\psi}_1(R|R,(\infty,\infty))$$

and

$$\psi^{1,q}(L|R,(\infty,\infty)) = \hat{\psi}_1(L|R,(\infty,\infty)).$$

Therefore, $\gamma_1^{1,q}((\infty,\infty),R) < 0$ by the concavity of the log operator, for any q. At $\nu = 0$ and $q = 1$, $x_1^R(0) = 0$ and therefore $\psi^{0,1}(R|R,(\infty,\infty)) = 0$. Note that R actions decrease the likelihood ratio, $\log\frac{\hat{\psi}_1(R|L,(\infty,\infty))}{\hat{\psi}_1(R|R,(\infty,\infty))} < 0$, while L actions increase the likelihood ratio, $\log\frac{\hat{\psi}_1(L|L,(\infty,\infty))}{\hat{\psi}_1(L|R,(\infty,\infty))} > 0$, independently of q and ν. Therefore, $\gamma_1^{0,1}((\infty,\infty),R) > 0$. Also, $\psi^{\nu,q}(R|R,(\infty,\infty))$ is strictly decreasing in q and strictly increasing in ν, since $x_1^R(\nu)$ is strictly increasing in ν. Therefore, $\gamma_1^{\nu,q}((\infty,\infty),R)$ is strictly decreasing in ν and increasing in q. Therefore, there exists a cutoff q_1 such that for $q > q_1$, there exists a cutoff $\nu_1(q) > 0$ such that for $\nu < \nu_1(q)$, $\gamma_1^{\nu,q}((\infty,\infty),R) > 0$ and (∞,∞) is locally stable, while for $\nu > \nu_1(q)$, $\gamma_1^{\nu,q}((\infty,\infty),R) < 0$ and (∞,∞) is not locally stable.

To determine whether $(0,0) \in \Lambda(R)$ at (ν,q), we need to determine the sign of

$$\gamma_1^{\nu,q}((0,0),R) = \psi^{\nu,q}(R|R,(0,0))\log\frac{\hat{\psi}_1(R|L,(0,0))}{\hat{\psi}_1(R|R,(0,0))}+\psi^{\nu,q}(L|R,(0,0))\log\frac{\hat{\psi}_1(L|L,(0,0))}{\hat{\psi}_1(L|R,(0,0))},$$

where

$$\begin{aligned}
\hat{\psi}_1(R|\omega,(0,0)) &= \alpha x_2^{\omega} + 1 - \alpha \\
\hat{\psi}_1(L|\omega,(0,0)) &= \alpha(1 - x_2^{\omega}) \\
\psi^{\nu,q}(R|R,(0,0)) &= \alpha q x_1^R(\nu) + \alpha(1-q)x_2^R + 1 - \alpha \\
\psi^{\nu,q}(L|R,(0,0)) &= \alpha q(1 - x_1^R(\nu)) + \alpha(1-q)(1 - x_2^R).
\end{aligned}$$

If $\nu = 1$, then $x_1^R(1) = x_2^R$, so $\psi^{1,q}(R|R,(0,0)) = \hat{\psi}_1(R|R,(0,0))$ and $\psi^{1,q}(L|R,(0,0)) = \hat{\psi}_1(L|R,(0,0))$. Therefore, $\gamma_1^{1,q}((0,0),R) < 0$ by the concavity of the log operator. At $\nu = 0$ and $q = 1$, then $x_1^R(0) = 0$, and therefore $\psi^{0,1}(R|R,(0,0)) = 1 - \alpha$. Therefore, $\gamma_1^{0,1}((0,0),R) > 0$. Moreover, $\gamma_1^{\nu,q}((0,0),R)$ is strictly increasing in q and strictly decreasing in ν, since $x_1^R(\nu)$ is strictly increasing in ν. Therefore, there exists a cut-off $q_2 < 1$ such that for any $q > q_2$, there exists a cutoff $\nu_2(q)$ such that for $\nu < \nu_2(q)$, $\gamma_1^{\nu,q}((0,0),R) > 0$ and $(0,0)$ is not locally stable, and for $\nu > \nu_2(q)$, $\gamma_1^{\nu,q}((0,0),R) < 0$ and $(0,0)$ is locally stable.

Suppose $\omega = L$. Then $\gamma^{1,q}((\infty,\infty),L) > 0$ and $\gamma^{1,q}((0,0),L) > 0$ for all $q \in [0,1]$, since only correct learning can occur for $\nu = 1$. The only change in the above expressions is that now the true measures are taken for state L, rather than state R. Therefore, all of the comparative statics on γ are preserved. As above, for any q, $\gamma^{\nu,q}((0,0),L)$ is decreasing in ν. Therefore, $\gamma^{\nu,q}((0,0),L) > 0$ for all ν and q, and incorrect learning is never locally stable. Also, for any q, $\gamma^{\nu,q}((\infty,\infty),L)$ is decreasing in ν. Therefore, $\gamma^{\nu,q}((\infty,\infty),L) > 0$ for all ν and q, and correct learning is always locally stable. ☐

A.5.3 Proofs from Section 5.3 (Social Perception Bias)

Proof of Propositions 4 and 5. Recall

$$\gamma_{x,1}(\lambda_1,\omega) = \sum_a \psi_x(a|\omega,\lambda_1) \log \frac{\hat{\psi}_{x,1}(a|L,\lambda_1)}{\hat{\psi}_{x,1}(a|R,\lambda_1)}.$$

When there is a single type, $0 \in \Lambda_x(\omega)$ iff $\gamma_{x,1}(0,\omega) < 0$ and $\infty \in \Lambda_x(\omega)$ if $\gamma_{x,1}(\infty,\omega) > 0$. If $\lambda_1 = \infty$, both types choose the risky action a_1 and the misspecification over types is irrelevant. Therefore, $\hat{\psi}_{x,1}(a|\omega,\infty) = \psi_x(a|\omega,\infty)$ and θ_1 is correctly specified at $\lambda_1 = \infty$. This implies $\gamma_{x,1}(\infty,R) < 0$ and $\gamma_{x,1}(\infty,L) > 0$, which establishes that $\infty \notin \Lambda_x(R)$ and $\infty \in \Lambda_x(L)$. Therefore, almost surely, $\lambda_1 \not\to \infty$ in state R, and $\lambda_1 \to \infty$ with positive probability in state L.

Suppose $\omega = R$. Then

$$\gamma_{x,1}(0,R) = ((1-\pi(\theta_1))q^R + \pi(\theta_1)) \log \frac{(1-\hat{\pi}_1(\theta_1))q^L + \hat{\pi}_1(\theta_1)}{(1-\hat{\pi}_1(\theta_1))q^R + \hat{\pi}_1(\theta_1)} + (1-\pi(\theta_1))(1-q^R) \log \frac{1-q^L}{1-q^R}. \tag{28}$$

This expression is strictly decreasing in $\hat{\pi}_1(\theta_1)$ and is strictly negative at $\hat{\pi}_1(\theta_1) = \pi(\theta_1)$. Let $\overline{\pi}^R \in [0, \pi(\theta_1))$ be either the value of $\hat{\pi}_1(\theta_1)$ at which (28) is equal to zero, or $\overline{\pi}^R = 0$ if (28) is negative for all $\hat{\pi}_1(\theta_1)$. As $\pi(\theta_1) \to 1$, the limit of (28) is strictly positive, so $\overline{\pi}^R$ is not always equal to zero. Therefore, $0 \notin \Lambda_x(R)$ for $\hat{\pi}_1(\theta_1) < \overline{\pi}^R$ and $0 \in \Lambda_x(R)$ for $\hat{\pi}_1(\theta_1) > \overline{\pi}^R$. As shown above, $\infty \notin \Lambda_x(R)$. Therefore, if $\hat{\pi}_1(\theta_1) < \overline{\pi}^R$, $\Lambda_x(R) = \emptyset$ and almost surely, beliefs do not converge. If $\hat{\pi}_1(\theta_1) > \overline{\pi}^R$, $\Lambda_x(R) = \{0\}$ and learning is complete.

Next we show that $\overline{\pi}^R$ is increasing in $\pi(\theta_1)$. The first term on the right hand side of (28) is positive and the second is negative. Increasing $\pi(\theta_1)$ increases the weight placed on the positive term and decreases the weight placed on the negative term. Therefore, $\gamma_{x,1}(0,R)$ is increasing in $\pi(\theta_1)$. Fixing $\hat{\pi}_1(\theta_1)$, if at $\pi(\theta_1)$, $\gamma_{x,1}(0,R) = 0$, increasing $\pi(\theta_1)$ makes $\gamma_{x,1}(0,R)$ strictly positive. Since $\gamma_{x,1}(0,R)$ is decreasing in $\hat{\pi}_1(\theta_1)$, the new cut-off is larger than the original cut-off.

Suppose $\omega = L$. Then

$$\gamma_{x,1}(0,L) = ((1-\pi(\theta_1))q^L + \pi(\theta_1)) \log \frac{(1-\hat{\pi}_1(\theta_1))q^L + \hat{\pi}_1(\theta_1)}{(1-\hat{\pi}_1(\theta_1))q^R + \hat{\pi}_1(\theta_1)} + (1-\pi(\theta_1))(1-q^L) \log \frac{1-q^L}{1-q^R}. \tag{29}$$

This expression is strictly decreasing in $\hat{\pi}_1(\theta_1)$, is strictly positive at $\hat{\pi}_1(\theta_1) = \pi(\theta_1)$ and is strictly negative at $\hat{\pi}_1(\theta_1) = 1$, since the second term on the right hand side of (29) is strictly negative. Let $\overline{\pi}^L \in (\pi(\theta_1), 1)$ be the value of $\hat{\pi}_1(\theta_1)$ at which (29) is equal to zero. Therefore, $0 \notin \Lambda_x(L)$ for $\hat{\pi}_1(\theta_1) < \overline{\pi}^L$ and $0 \in \Lambda_x(L)$ for $\hat{\pi}_1(\theta_1) > \overline{\pi}^L$. As shown above, $\infty \in \Lambda_x(L)$. Therefore, if $\hat{\pi}_1(\theta_1) < \overline{\pi}^L$, $\Lambda_x(L) = \{\infty\}$ and almost surely, learning is complete. If $\hat{\pi}_1(\theta_1) > \overline{\pi}^L$, $\Lambda_x(L) = \{0, \infty\}$ and both correct and incorrect learning occur with positive probability.

Finally, we show that $\overline{\pi}^L$ is increasing in $\pi(\theta_1)$. Equation (29) is increasing in $\pi(\theta_1)$. Therefore, as $\pi(\theta_1)$ increases, if $\gamma_{x,1}(0,L) = 0$ at $\hat{\pi}_1(\theta_1)$ before, it is positive there now.
□

B Learning from Outcomes

States and Actions. The state and action space are the same as in Section 2.

Outcomes and Histories. Each agent's choice of action leads to a stochastic outcome $x \in \mathcal{X}$, where $2 \leq |\mathcal{X}| < \infty$ is the outcome space. When the agent chooses action a, the outcomes are distributed according to $H^\omega(x|a)$. Assume that no outcome perfectly reveals the state, which implies that $(H^L(\cdot|a), H^R(\cdot|a))$ are mutually absolutely continuous with common support for all $a \in \mathcal{A}$. An action has *full support* if all outcomes occur with positive probability following this action i.e. $\operatorname{supp} H(\cdot|a) = \mathcal{X}$. An action a is *informative* if, following this action, at least one realized outcome reveals information about the state i.e. $\frac{dH^L}{dH^R}(\cdot|a) \neq 1$.

Assumption 5 (Informative Action)**.** *There exists at least one informative full support action.*

We restrict attention attention to environments in which outcomes are aligned, in that we can order the outcome space so that for any two actions a and a', if outcome x is stronger evidence for state L than outcome x'', conditional on a, then it is also stronger evidence for state L, conditional on a'. This is analogous to aligned preferences (Assumption 2).

Assumption 2* (Aligned Outcomes)**.** *Outcomes are aligned for all actions i.e. for all actions $a, a' \in \mathcal{A}$, $H^L(\cdot|a), H^R(\cdot|a)$ and $H^L(\cdot|a'), H^R(\cdot|a')$ are aligned.*

Outcomes, but not actions, are observed by subsequent agents. To keep the notation simple, assume that there are no public or private signals. Therefore, subsequent agents learn solely from the information that outcomes convey about the state. The history at time t is the sequence of past outcome realizations $h_t = \{x_1, x_2, \ldots, x_{t-1}\}$.

Types Framework. An agent's type specifies her model of inference and her preferences. Agent t has privately observed type $\theta_t \in \Theta$, where Θ is a non-empty finite set and $\pi \in \Delta(\Theta)$ is the distribution over types. Now, a model of inference determines how a type learns from prior outcomes. For each type θ_i, this includes (i) a subjective belief about the outcome distribution, $\hat{H}_i^\omega(x|a)$ for each $\omega \in \{L, R\}$ and $a \in \mathcal{A}$, and (ii) a subjective belief about the likelihood of other types, $\hat{\pi}_i \in \Delta(\Theta)$. Assume that each type believes that no outcome perfectly reveals the state, $(\hat{H}_i^L(\cdot|a), \hat{H}_i^R(\cdot|a))$ are mutually absolutely continuous for all $a \in \mathcal{A}$ and $\theta \in \Theta$.

A type's payoffs depend on the realized outcome, in addition to the action and the state. Let $v_i : A \times \mathcal{X} \times \Omega \to \mathbb{R}$ denote the preferences for type θ_i. We work directly with the reduced form preferences, $u_i(a, \omega) = \hat{E}_i(v_i(a, x, \omega)|a, \omega)$, where the expectation is taken with respect to the type's potentially misspecified probability distribution over outcomes, conditional on the state, and its action choice. Assume that each type has a unique optimal action when the state is known.

We restrict attention to environments in which the way that each type interprets outcomes is aligned with the true distribution. This is analogous to aligned signals (Assumption 1).

Assumption 1* (Aligned Subjective Outcomes)**.** *Outcomes are aligned for all informative types θ_i i.e. $\hat{H}_i^L(\cdot|a), \hat{H}_i^R(\cdot|a)$ and $H^L(\cdot|a), H^R(\cdot|a)$ are aligned for all $a \in \mathcal{A}$.*

Fixing an action a, this assumption ensures that for any two outcomes x and x', if x leads to a higher true belief that the state is L than x', then it also leads to a higher subjective belief that the state is L for all types. We make one exception to continue to allow some types to perceive all outcomes as being uninformative.

We also need an assumption to ensure that outcomes are always informative and perceived as informative. The following assumption requires that at least one action that occurs with positive probability is informative and generates every possible outcome. This is the analogue of Assumptions 3 and 4.

Assumption 3* (Adequate Consistent Information)**.** *Assume that there exists a type θ_i that plays an informative full support action at all beliefs $\lambda_i \in [0, \infty)$, and all types believe that there is a type that plays an informative full support action at all beliefs $\lambda_i \in [0, \infty]$.*

This ensures that adequate information arrives for complete learning in correctly specified models. It also rules out outcome realizations that are inconsistent with a type's model of inference.

The Individual Decision-Problem. Given λ_i, let $a(\theta, \lambda_i)$ denote the optimal action for type θ at belief λ_i,

$$a(\theta, \lambda_i) \equiv \arg\max_{a \in \mathcal{A}} \sum_{x \in \mathcal{X}} \left(\left(\frac{1}{1+\lambda_i} \right) v_i(a, x, L) \hat{H}_i^L(x|a) + \left(\frac{\lambda_i}{1+\lambda_i} \right) v_i(a, x, R) \hat{H}_i^R(x|a) \right)$$

A type θ_i is indifferent between actions at finitely many interior beliefs. How ties are broken is irrelevant for the results, so for notational convenience, actions will be treated as if they are unique.

The Likelihood Ratio. For any equilibrium we can define the true and subjective probabilities of each outcome as

$$\psi_x(x|\omega, \boldsymbol{\lambda}) \equiv \sum_{j=1}^{n} \pi(\theta_j) dH^\omega(x|a(\theta_j, \lambda_j)), \tag{30}$$

and

$$\hat{\psi}_{x,i}(x|\omega,\boldsymbol{\lambda}) = \sum_{j=1}^{n} \hat{\pi}_i(\theta_j) d\hat{H}_i^{\omega}(x|a(\theta_j,\lambda_j)), \tag{31}$$

where $d\hat{H}_i^{\omega}(x|a)$ denotes type θ_i's subjective probability of outcome x in state ω, conditional on action a. We can use these probabilities to construct the likelihood ratio, as in (5) and (6).

Asymptotic Learning Characterization. We construct analogous expressions to charactertize asymptotic learning outcomes when agents learn from outcomes. Let

$$\gamma_{x,i}(\boldsymbol{\lambda},\omega) \equiv \sum_{x\in\mathcal{X}} \psi_x(x|\omega,\boldsymbol{\lambda}) \log \frac{\hat{\psi}_{x,i}(x|L,\boldsymbol{\lambda})}{\hat{\psi}_{x,i}(x|R,\boldsymbol{\lambda})}. \tag{32}$$

Define $\Lambda_{x,i}(\omega)$ as the analogue of (8) for $\gamma_{x,i}(\boldsymbol{\lambda},\omega)$, $\Lambda_{x,M}(\omega)$ as the analogue of (10), and let $\Lambda_x(\omega) \equiv \cap_{i=1}^{k} \Lambda_{x,i}(\omega)$.

By Assumption 5, there exists an informative full support action. Let a^* denote one such action. Let x_L denote the maximal outcome in state L when a^* is chosen, i.e. the outcome that maximizes $d\hat{H}_i^L(x|a^*)/d\hat{H}_i^R(x|a^*)$. Observing this outcome leads to the largest increase in the likelihood ratio. Similarly, let x_R denote the minimal outcome in state L when a^* is chosen, i.e. the outcome that minimizes $d\hat{H}_i^L(x|a^*)/d\hat{H}_i^R(x|a^*)$. Observing this outcome leads to the largest decrease in the likelihood ratio. By finiteness of the outcome space, these exist. By Assumption 2*, for any action a such that x_L (x_R) is in the support of the outcomes that arise following a, x_L (x_R) is the maximal (minimal) outcome, conditional on a. By Assumption 1*, x_L and x_R are also perceived to be the maximal and minimal outcomes. Given x_L and x_R, define a maximal L-order analogous to Definition 5, using the subjective probabilities of the maximal *outcomes* in each state, $\hat{\psi}_{x,i}(x_R|\omega,\boldsymbol{\lambda})$ and $\hat{\psi}_{x,i}(x_L|\omega,\boldsymbol{\lambda})$. Maximal accessibility for outcomes remains identical to Definition 6.

An analogue of Theorem 1 holds with respect to $\Lambda_x(\omega)$, and we can fully characterize asymptotic learning when agents learn from outcomes.

Theorem 1*. *Assume Assumptions 1*, 2*, 3* and 5. Suppose there are two sociable types, $k=2$, and $\omega = R$.*

1. **Agreement.** Correct *learning occurs with positive probability iff* $(0,0) \in \Lambda_x(R)$ *and* incorrect *learning occurs with positive probability iff* $(\infty,\infty) \in \Lambda_x(R)$.

2. **Disagreement.** *Sociable types* disagree *with positive probability if $\Lambda_x(R)$ contains a maximally accessible disagreement outcome, and sociable types almost surely do*

not disagree *if* $\Lambda_x(R)$ *contains no disagreement outcomes. Each maximally accessible* disagreement *outcome in* $\Lambda_x(R)$ *occurs with positive probability*

3. **Cyclical Learning.** Cyclical *learning occurs almost surely for all sociable types if* $\Lambda_x(R) \cup \Lambda_{x,M}(R)$ *is empty, and cyclical learning occurs almost surely for at least one sociable type if* $\Lambda_x(R)$ *is empty. Cyclical learning almost surely does not occur for any sociable type if* $\Lambda_x(R)$ *contains an agreement outcome or maximally accessible disagreement outcome and* $\Lambda_{x,M}(R)$ *is empty.*

An analogous result holds for $\omega = L$.[44]

The outcomes $x \in \mathcal{X}$ function effectively the same way as actions in the original model. When beliefs are in a sufficiently small neighborhood around the extreme beliefs, the probability of each action, and therefore, the distribution of realized outcomes, are distributed independently, conditional on the state. Similar arguments can be used to establish local stability. Further, outcomes are aligned across actions and types, and the maximal and minimal outcomes x_L and x_R serve the same purpose as the analogous action and signal pairs (a_L, σ_L) and (a_R, σ_R). Therefore, similar arguments establish global convergence. We omit a formal proof of Theorem 1*, as it directly mirrors the proof of Theorem 1.

Robustness of Complete Learning. Under these modified assumptions, a direct analogue of Theorem 3 holds with respect to ψ_x and $\hat{\psi}_{x,i}$. Additionally, a second robustness result similar to Theorem 4 holds. By the continuity of $\hat{\psi}_{i,x}$, if all sociable types have subjective type and outcome distributions close to the true distributions, i.e. there exists a $\delta > 0$ such that $||\hat{\pi}_i - \pi|| < \delta$ and $||\hat{H}_i^\omega(\cdot|a) - H^\omega(\cdot|a)|| < \delta$ for all $a \in \mathcal{A}$ and $\omega \in \{L, R\}$, then learning is complete. Further, agents can be very wrong about the type distribution, as long as the types that they do believe to occur are "close" to the actual types. We define a measure of closeness with respect to the optimal action for each type at belief $\boldsymbol{\lambda}$, denoted $a_i(\boldsymbol{\lambda})$ for type θ_i. Theorem 4* states this general robustness result. The case in which the subjective type distributions are close to the true distribution, $||\hat{\pi}_i - \pi|| < \delta$, is a special case of Theorem 4*.

Theorem 4*. *Assume Assumptions 1*, 2*, 3* and 5, and fix state* ω. *Let* $\Theta(a, \boldsymbol{\lambda}) \equiv \{\theta_i | a_i(\boldsymbol{\lambda}) = a\}$. *There exists a* $\delta > 0$ *such that if* $||\hat{\pi}_i(\Theta(\cdot, \boldsymbol{\lambda})) - \pi_i(\Theta(\cdot, \boldsymbol{\lambda}))|| < \delta$ *for* $\boldsymbol{\lambda} \in \{0, \infty\}^k$ *and* $||\hat{H}_i^\omega(\cdot|a) - H^\omega(\cdot|a)|| < \delta$ *for all* $a \in \mathcal{A}$ *and for all sociable types* θ_i, *then learning is complete in state* ω.

[44] As in the main paper, the statement of the theorem is identical for $k > 2$, modifying the definitions of $\Lambda_x(\omega)$, $\Lambda_{x,M}(\omega)$ and maximal accessibility in an analogous way to Appendix A.2.

Proof. Let $\boldsymbol{\lambda} \in \{0, \infty\}^k$. Recall that

$$\hat{\psi}_{x,i}(x|\omega, \boldsymbol{\lambda}) = \sum_{j=1}^{n} \hat{\pi}_i(\theta_j) d\hat{H}_i^{\omega}(x|a_j(\boldsymbol{\lambda})).$$

Since no type mixes at beliefs $\boldsymbol{\lambda} \in \{0, \infty\}^k$, we can group types by the action they play at $\boldsymbol{\lambda}$. grouping types by the action they play at $\boldsymbol{\lambda}$. Therefore,

$$\begin{aligned}
\hat{\psi}_{x,i}(x|\omega, \boldsymbol{\lambda}) &= \sum_{a \in \mathcal{A}} \sum_{\theta_j \in \Theta(a, \boldsymbol{\lambda})} \hat{\pi}_i(\theta_j) d\hat{H}_i^{\omega}(x|a) \\
&= \sum_{a \in \mathcal{A}} d\hat{H}_i^{\omega}(x|a) \sum_{\theta_j \in \Theta(a, \boldsymbol{\lambda})} \hat{\pi}_i(\theta_j) \\
&= \sum_{a \in \mathcal{A}} d\hat{H}_i^{\omega}(x|a) \hat{\pi}_i(\Theta(a, \boldsymbol{\lambda})).
\end{aligned}$$

So for all $\theta_i \in \Theta_s$, $\hat{\psi}_{x,i}$ varies continuously in $\hat{H}_i^{\omega}(\cdot|a)$ and $\hat{\pi}_i$. This implies that $\gamma_{x,i}(\boldsymbol{\lambda}, \omega)$ varies continuously. In any correctly specified model, $\gamma_{x,i}(\boldsymbol{\lambda}, R)$ is strictly negative at all stationary beliefs $\boldsymbol{\lambda} \in \{0, \infty\}^k$. Therefore, there exists a sufficiently small δ such that if $||\hat{H}_i(\cdot|a) - H(\cdot|a)|| < \delta$ and $||\hat{\pi}_i(\Theta(\cdot, \boldsymbol{\lambda})) - \pi(\Theta(\cdot, \boldsymbol{\lambda}))|| < \delta$, then $\gamma_{x,i}(\boldsymbol{\lambda}, R) < 0$ at all stationary beliefs. As argued in Appendix A.2, if $\gamma_{x,i}(\boldsymbol{\lambda}, R) < 0$ for all stationary $\boldsymbol{\lambda}$, then $\Lambda_{x,M}(R)$ is empty. Similarly, δ can be chosen to be sufficiently small so that $\gamma_{x,i}(\boldsymbol{\lambda}, L)$ is strictly positive at all stationary beliefs $\boldsymbol{\lambda} \in \{0, \infty\}^k$. Therefore, correct learning is robust to some misspecification. □

C Posterior Representation.

Let $\mathcal{Z}$ be a signal space. Let $\mu^{\omega} \in \Delta(\mathcal{Z})$ and $\nu^{\omega} \in \Delta(\mathcal{Z})$ be probability measures on $\mathcal{Z}$ in state ω. Assume μ^L, μ^R and ν^L, ν^R are mutually absolutely continuous. Let $s(z) \equiv 1/(1 + \frac{d\mu^R}{d\mu^L}(z))$ and $p(z) \equiv 1/(1 + \frac{d\nu^R}{d\nu^L}(z))$ denote the posterior belief that the state is L. The c.d.f.s $F_s^{\omega}(x) \equiv \mu^{\omega}(z|s(z) \leq x)$ and $G_p^{\omega}(x) \equiv \nu^{\omega}(z|p(z) \leq x)$ are the distributions of the posterior belief s and p under measure μ^{ω} and ν^{ω}, respectively. Given these two measures, we can also define the distribution of p under measure μ^{ω} as $F_p^{\omega}(x) \equiv \mu^{\omega}(z|p(z) \leq x)$, and the distribution of s under measure ν^{ω} as $G_s^{\omega}(x) \equiv \nu^{\omega}(z|s(z) \leq x)$.

Multiple signals z and z' can lead to the same posterior beliefs. Therefore, two distributions can map to the same distribution over posterior beliefs. This means that these distributions over $\mathcal{Z}$ can map the same signals to the same posterior belief, but have different measures over these signals z and z'. The following property describes an equivalence class of probability measures. These measures have the same ordinal ranking of signals and the same distribution over posterior beliefs.

Definition 15 (Equivalent Measures)**.** *Measures μ^L, μ^R and ν^L, ν^R are* equivalent *iff they are aligned,* $\operatorname{supp}\mu = \operatorname{supp}\nu$*, and* $\mu^\omega(z|1/(1+\frac{d\mu^R}{d\mu^L}(z)) \leq x) = \nu^\omega(z|1/(1+\frac{d\nu^R}{d\nu^L}(z)) \leq x)$ *for all* $x \in [0,1]$ *and* $\omega \in \{L, R\}$*.*

Lemma 15 establishes that when a probability measure $\nu^L, \nu^R \in \Delta(\mathcal{Z})^2$ is aligned with probability measure $\mu^L, \mu^R \in \Delta(\mathcal{Z})^2$, there is a unique representation of ν^L, ν^R as (r, G_s^L), where $r : \operatorname{supp} F_s \to [0,1]$ is a strictly increasing function mapping the posterior s to the posterior p and G_s^L is the distribution of s under measure ν^L.

Lemma 15. *Suppose μ^L, μ^R have full support and signals are informative, $\frac{d\mu^R}{d\mu^L}(z) \neq 1$.*

1. *For any mutually absolutely continuous probability measures $\nu^L, \nu^R \in \Delta(\mathcal{Z})^2$ that have full support and are aligned with μ^L, μ^R, there exists a unique (r, G_s^L), where $r : \operatorname{supp} F_s \to [0,1]$ is a strictly increasing function with $r(\inf \operatorname{supp} F_s) < 1/2$ and $r(\sup \operatorname{supp} F_s) > 1/2$, such that $r(s(z)) = 1/(1 + \frac{d\nu^R}{d\nu^L}(z))$ for all $z \in \mathcal{Z}$ and G_s^L is the distribution of s under measure ν^L.*

2. *For any strictly increasing function $r : \operatorname{supp} F_s \to [0,1]$ and any c.d.f. G_s^L with $\operatorname{supp} G_s^L = \operatorname{supp} F_s$ and $\int_0^1 \left(\frac{1-r(s)}{r(s)}\right) dG_s^L = 1$, there exist unique (up to an equivalent pair of measures) mutually absolutely continuous probability measures $(\nu^L, \nu^R) \in \Delta(\mathcal{Z})^2$ that have full support, are aligned with μ^L, μ^R, and satisfy $r(s(z)) = 1/(1 + \frac{d\nu^R}{d\nu^L}(z))$ for all $z \in \mathcal{Z}$. The measures ν^L, ν^R are aligned with μ^L, μ^R.*[45]

3. *For any strictly increasing function $r : \operatorname{supp} F_s \to [0,1]$ such that $r(\inf \operatorname{supp} F_s) < 1/2$ and $r(\sup \operatorname{supp} F_s) > 1/2$, there exists mutually absolutely continuous probability measures $\nu^L, \nu^R \in \Delta(\mathcal{Z})^2$ that have full support, are aligned with μ^L, μ^R, and satisfy $r(s(z)) = 1/(1 + \frac{d\nu^R}{d\nu^L}(z))$ for all $z \in \mathcal{Z}$.*

The first part of Lemma 15 implies that $F_p^\omega(r(s)) = F_s^\omega(s)$ for all $s \in \operatorname{supp}(F_s)$ and $\operatorname{supp} F_p = r(\operatorname{supp} F_s)$. Similarly, $G_p^\omega(r(s)) = G_s^\omega(s)$ for all $s \in \operatorname{supp} G_s$ and $\operatorname{supp} G_p = r(\operatorname{supp} G_s)$.

Proof. First establish part (i). Let $(\nu^L, \nu^R) \in \Delta(\mathcal{Z})^2$ be mutually absolutely continuous probability measures that have full support and are aligned with (μ^L, μ^R). Define the mapping $r : \operatorname{supp} F_s \to [0,1]$ as $r(s(z)) = p(z)$, where $p(z) \equiv 1/(1 + \frac{d\nu^R}{d\nu^L}(z))$. This is a function since if $s(z) = s(z')$, then $p(z) = p(z')$, which establishes existence. For any

[45] Note that if G_s^L is a c.d.f. and $\int_0^1 \left(\frac{1-r(s)}{r(s)}\right) dG_s^L = 1$, then it must be that $r(\sup \operatorname{supp} F_s) > 1/2$ and $r(\inf \operatorname{supp} F_s) < 1/2$.

z such that $s(z) > s(z')$, $p(z) = r(s(z)) > p(z') = r(s(z'))$ since (ν^L, ν^R) is aligned. Therefore, r is strictly increasing on $\operatorname{supp} F_s$.

By the Bayesian constraint, it must be that $E_\nu[p(z)] = 1/2$, where the expectation is taken with respect to (ν^L, ν^R). Given that (μ^L, μ^R) are informative and aligned with (ν^L, ν^R), it cannot be that $p(z) = 1/2$ for all $z \in \mathcal{Z}$. Therefore, there exist $z, z' \in \mathcal{Z}$ such that $p(z) > 1/2$ and $p(z') < 1/2$, which implies that there exist $s, s' \in \operatorname{supp} F_s$ such that $r(s) > 1/2$ and $r(s') < 1/2$. Given that r is strictly increasing in s, it immediately follows that $r(\inf \operatorname{supp} F_s) < 1/2$ and $r(\sup \operatorname{supp} F_s) > 1/2$. Define $G_s^L(x) \equiv \nu^L(z|s(z) \leq x)$. Then G_s^L is the distribution of s under measure ν^L. Given $\{r, G_s^L\}$, G_s^R is uniquely pinned down by

$$G_s^R(x) = \int_0^x \left(\frac{1-r(s)}{r(s)}\right) dG_s^L(s)$$

for any $x \in \operatorname{supp} F_s$.

Next, show part (ii). Let $r : \operatorname{supp} F_s \to [0,1]$ be a strictly increasing function and let c.d.f. G_s^L be the distribution of s under measure ν^L, with $\operatorname{supp} G_s^L = \operatorname{supp} F_s$ and $\int_0^1 \left(\frac{1-r(s)}{r(s)}\right) dG_s^L = 1$. By Lemma A.1 in Smith and Sorensen (2000), the distribution of s under measure ν^R is uniquely determined by

$$G_s^R(x) = \int_0^x \left(\frac{1-r(s)}{r(s)}\right) dG_s^L(s).$$

Since G_s^R has Radon-Nikodym derivative $\frac{1-r(s)}{r(s)}$, it induces posterior belief $r(s)$ after observing signal z from set of signals $Z = \{z|s(z) = s\}$ that lead to correctly specified posterior s, for any $s \in \operatorname{supp} F_s$. If any other distribution induced the same posterior beliefs, then it would also have Radon-Nikodym derivative $\frac{1-r(s)}{r(s)}$, so it would be equivalent to G_s^R. Since $\frac{1-r(s)}{r(s)} > 0$ and $G_s^R(1) = 1$, G_s^R is a probability distribution.

Define the random variable $S = s(z)$. G_s^ω defines a probability measure over this random variable in state ω. For any measurable set $A \subseteq \mathcal{Z}$, define

$$\nu^\omega(A) = \int E(\mathbb{1}_A|S) dG_s^\omega,$$

where E is the conditional expectation defined with respect to μ^L. By the uniqueness and additivity of conditional expectation, for any disjoint, measurable sets $A, B \subseteq \mathcal{Z}$,

$$\nu^\omega(A \cup B) = \int E(\mathbb{1}_{A\cup B}|S) dG_s^\omega = \int (E(\mathbb{1}_A|S) + E(\mathbb{1}_B|S)) dG_s^\omega = \nu^\omega(A) + \nu^\omega(B),$$

so ν^ω is a measure. For any set A, if $\nu^L(A) = 0$, then $\nu^R(A) = 0$ and vice versa, since the integrand used to define ν^R is strictly positive. Therefore, the distributions (ν^L, ν^R)

are mutually absolutely continuous with common support $\operatorname{supp}\nu$. Also, $\operatorname{supp}\nu = \operatorname{supp}\mu$ by construction, so the measures have full support on $\mathcal{Z}$. Moreover, since F_s^ω is unique, ν^ω is unique up to the probability measure that is used to evaluate $E(\cdot|S)$. For any measurable set $A \subseteq \mathcal{Z}$,

$$\nu^R(A) = \int E(\mathbb{1}_A|S)\left(\frac{1-r(S)}{r(S)}\right) dG_s^L = \int_A \left(\frac{1-r(s(z))}{r(s(z))}\right) d\nu^L(z),$$

where the first equality follows from the definition of G_s^R and the second equality follows from the definition of ν^L, so these distributions induce the correct posterior beliefs. Finally, $\nu^L(\mathcal{Z}) = \int_0^1 dG_s^L(s) = 1$ and $\nu^R(\mathcal{Z}) = \int_0^1 dG_s^R(s) = 1$, so these are indeed probability measures.

Finally, show part (iii). Suppose $r : \operatorname{supp} F_s \to [0,1]$ is a strictly increasing function with $r(\inf \operatorname{supp} F_s) < 1/2$ and $r(\sup \operatorname{supp} F_s) > 1/2$. Fix any distribution G with support $\operatorname{supp} F_s \cap \{s|r(s) < 1/2\}$. Then $\int_0^1 \left(\frac{1-r(s)}{r(s)}\right) dG_s(s) < 1$. Similarly, fix a distribution $\hat{G}$ with support $\operatorname{supp} F_s \cap \{s|r(s) \geq 1/2\}$. Then $\int_0^1 \left(\frac{1-r(s)}{r(s)}\right) d\hat{G}(s) > 1$. For any $\lambda \in [0,1]$, let G_λ be the distribution of the compound lottery $G_\lambda = \lambda G + (1-\lambda)\hat{G}$. This lottery draws signals from G with probability λ and $\hat{G}$ with probability $(1-\lambda)$. The function $H(\lambda) \equiv \int \left(\frac{1-r(s)}{r(s)}\right) dG_\lambda$ is a continuous mapping from $[0,1]$ to $\mathbb{R}$, so by the intermediate value theorem, there exists a $\lambda^* \in (0,1)$ such that $\int \left(\frac{1-r(s)}{r(s)}\right) dG_{\lambda^*} = 1$. Let $G^L = G_{\lambda^*}$. Then G^L is a probability distribution, since it is the convex combination of two distributions. By construction, $\operatorname{supp} G_s^L = \operatorname{supp} F_s$ and $\int_0^1 \left(\frac{1-r(s)}{r(s)}\right) dG_s^L = 1$. Therefore, from part(ii), it is possible to construct the desired probability measures (ν^L, ν^R). □

D Supplemental Appendix: Examples of Nested Models

This framework directly nests Bohren (2016). There are two types: a sociable type θ_1 and an autarkic type θ_2. The sociable type has a misspecified model of the frequency of the autarkic type. Mapping the notation from Bohren (2016) to this paper, $p = \pi(\theta_2)$ is the true share of autarkic types, and $\hat{p} = \hat{\pi}_1(\theta_2)$ is θ_1's belief about the share of autarkic types. Both agents have a correct model of the signal distributions and preferences.

It is simple to extend the model in this paper to allow agent's misspecified models to depend on their own beliefs. In particular, an agent's misspecified model can be a continuous function $(F^L, F^R, G^L, G^R, u, \pi) : [0,1] \to \Delta([0,1])^4 \times \mathbb{R}^{|\mathcal{A}|} \times \Delta(\Theta)$ that maps from type i's current belief to the misspecified model they use for updating at this belief. This allows us to apply the techniques developed in this paper to analyze boundedly rational models of several other papers, including Rabin and Schrag (1999) and Epstein et al. (2010).

D.1 Confirmation Bias: Rabin and Schrag (1999)

Rabin and Schrag (1999) examines individual learning with confirmation bias. Agents receive a binary signal, but if they receive a signal that goes against their prior beliefs then with probability q they misinterpret that signal as the other signal (which agrees with their prior belief). In order to nest this model, a slight extension must be made to the framework we've outlined. In particular, this the mapping ρ to be able to map two public signals that induce the same posterior to different misspecified beliefs. It is straightforward to extend all arguments made in this paper to this case.

This is a misspecified model with one type θ. There are 4 public signals y_{L_1}, y_{L_2}, y_{R_1}, y_{R_2}. All L signals induce the same posterior and all R signals induce the same posterior. Conditional on seeing an L signal, y_{L_2} is draw with probability q. Similarly, conditional on seeing an R signal y_{R_1} is drawn with probability q. Assume $Pr(y_{L_1} \text{ or } y_{L_2} | \omega = L) = Pr(y_{R_1} \text{ or } y_{R_2} | \omega = R) = \bar{\sigma} > 1/2$. As before, let $\sigma(y)$ be the posterior belief a correctly specified type would have if they received signal y. The misspecification is as follows, if $\lambda < 1$, then $\rho(y_{L_2}) = \sigma(y_{R_1})$ and all other signals are interpreted correctly. If $\lambda > 1$ then $\rho(y_{R_1}) = \sigma(y_{L_2})$ and all other signals are interpreted correctly.

To complete the model, assume that public signals are the only source of information and they are informative (i.e. private signals are uninformative, are believed to be uninformative, and $\sigma_L > 1/2$). An agent's beliefs about how other agents interpret public signals is irrelevant, as there is no additional information contained in actions. Agents choose actions $a \in \{L, R\}$ and receive utility $u(a, \omega) = 1_{a=\omega}$.

This misspecification captures a model where an agent skews evidence that goes against his prior towards his prior belief. If an agent believes that it is more likely that the state is L than R, whenever he receives a signal that favors state R with probability

q it is signal y_{R_1}, which he interprets as being an L signal instead of an R signal.

The parameter q indexes the degree of confirmation bias. Higher q means it is more likely that agents misinterpret signals that go against their prior. Under this specification,

$$\gamma(0,R) = (1-q)\left(\bar{\sigma}\log\left(\frac{1-\bar{\sigma}}{\bar{\sigma}}\right) + (1-\bar{\sigma})\log\left(\frac{\bar{\sigma}}{1-\bar{\sigma}}\right)\right) + q\log\left(\frac{\bar{\sigma}}{1-\bar{\sigma}}\right).$$

and

$$\gamma(\infty,R) = (1-q)\left(\bar{\sigma}\log\left(\frac{1-\bar{\sigma}}{\bar{\sigma}}\right) + (1-\bar{\sigma})\log\left(\frac{\bar{\sigma}}{1-\bar{\sigma}}\right)\right) + q\log\left(\frac{1-\bar{\sigma}}{\bar{\sigma}}\right).$$

As q increases, more weight is placed on the last term, which is negative when $\lambda = 0$ and positive when $\lambda = \infty$. So when the degree of confirmation bias is sufficiently high agents learn the correct state and incorrect state with positive probability. If the degree of confirmation bias is sufficiently severe agent's are susceptible to falling into traps, initial signals that go against the true state are hard to overturn due to confirmation bias. As soon as the prior favors one state or the other, most of the signals start to confirm that prior. Beliefs become entrenched and are hard to overturn.

In addition to this model presented originally in Rabin and Schrag (1999), our framework allows us to analyze potentially richer forms of confirmation bias. For instance, the confirmation bias in Rabin and Schrag (1999) is relatively extreme, agents misinterpret information in exactly the same way if they are almost convinced that the state is L or if they only believe state L is ε more likely than state R. Perhaps in smoother models of confirmation bias, incorrect learning would be impossible.

Suppose there are 4 public signals y_{L_1}, y_{L_2}, y_{R_1} and y_{R_2} that are drawn with the same frequency as in the preceding example. But now the misspecification is as follows

$$\rho(y,p) = \begin{cases} \sigma(y) & \text{if } y \in \{y_{L_1}, y_{R_2}\} \\ \sigma(y) & \text{if } y = y_{L_2} \text{ and } p \geq 1/2 \\ \sigma(y) & \text{if } y = y_{R_1} \text{ and } p \leq 1/2 \\ \sigma(y) + \varrho(p)(\sigma(y_{R_1}) - \sigma(y)) & \text{if } y = y_{L_2} \text{ and } p < 1/2 \\ \sigma(y) + \varrho(p)(\sigma(y_{L_2}) - \sigma(y)) & \text{if } y = y_{R_1} \text{ and } p > 1/2 \end{cases}$$

where $\varrho : [0,1] \to [0,1]$ is a continuous function with $\varrho(0) = \varrho(1) = 1$. As before, confirmation bias is more severe the higher q is, the probability that signals in favor of one state are missinterpreted as signals in favor of the other. But now $\varrho(\cdot)$ allows the degree of confirmation bias to depend the degree to which the prior favors the state, weighting how much the signal is shifted towards the signal that favors the other state of the world. If ϱ is strictly decreasing on $[0, 1/2]$ and strictly increasing on $[1/2, 1]$, then the bias becomes more severe as the agent's prior becomes more extreme. The model of Rabin and Schrag (1999) corresponds to the case where ϱ is constant and equal to 1.

Using the tools developed in this paper, one can verify that the possibility of the long-run incorrect learning is independent of the shape of $\varrho(\cdot)$, so the extreme form of confirmation bias present in Rabin and Schrag (1999) is not driving the possibility of this long run outcome. No matter the form ϱ takes, confirmation bias leads to entrenchment. Agents underweight information that does not confirm their current beliefs. This can lead to traps; if agents reach a sufficiently extreme belief in favor of one state, they'll skew information so that much it becomes almost impossible to overturn their preconception.

Proposition 6. *Suppose $\omega = R$. There exists a unique cutoff $\bar{q} = 1 - 1/(2\bar{\sigma}) < 1$ such that for any $\varrho : [0, 1] \to [0, 1]$, continuous with $\varrho(0) = \varrho(1) = 1$, if*

1. *If $q < \bar{q}$ then learning is correct almost surely.*
2. *If $q > \bar{q}$ then both correct and incorrect learning occur with positive probability, and beliefs converge almost surely.*

This cutoff is the same as the cutoff in Rabin and Schrag (1999) for all $\varrho(\cdot)$, which is driven by a combination of the continuity of ϱ at 0 and 1 and $\varrho(0) = \varrho(1) = 1$. Whenever $\varrho(p) = 1$, at that belief p agents update exactly like they would update in Rabin and Schrag (1999), so there will always be a neighborhood of belief near each stationary point where the model behaves sufficiently similarly to the model of Rabin and Schrag (1999). So, as in that model, this neighborhood becomes a trap, beliefs that enter it become entrenched no matter what the true state of the world is and incorrect learning occurs with positive probability.

While possibility of the long-run incorrect learning is independent of ϱ, the probability of each long-run outcome depends crucially on it. Intuitively, a larger ϱ amplifies the impact of early signals, which in turn makes it harder to overturn beliefs that even mildly favor the incorrect state.

In Figure 4 smaller x corresponds to more severe confirmation bias, since decreasing x increases the weight placed on the signal that favors the other state of the world. A higher x reduces the degree to which missperceived signals are skewed, and thus makes incorrect learning less likely. $x = 0$ corresponds to the extreme form of confirmation bias present in Rabin and Schrag (1999).

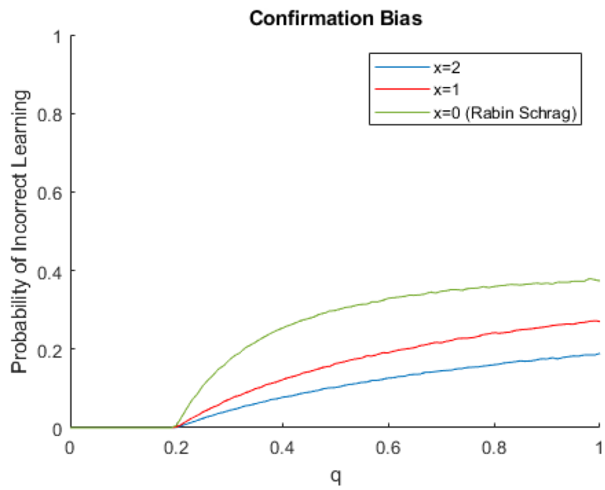

FIGURE 4. $\varrho(p) = (2|p-1/2|)^x$, $\bar{\sigma} = 3/8$

Proof of Proposition 6. Suppose $\omega = R$. Then

$$\gamma(0,R) = (1-q)\left(\bar{\sigma}\log\left(\frac{1-\bar{\sigma}}{\bar{\sigma}}\right) + (1-\bar{\sigma})\log\left(\frac{\bar{\sigma}}{1-\bar{\sigma}}\right)\right) + q\log\left(\frac{\bar{\sigma}}{1-\bar{\sigma}}\right)$$

and

$$\gamma(\infty,R) = (1-q)\left(\bar{\sigma}\log\left(\frac{1-\bar{\sigma}}{\bar{\sigma}}\right) + (1-\bar{\sigma})\log\left(\frac{\bar{\sigma}}{1-\bar{\sigma}}\right)\right) + q\log\left(\frac{1-\bar{\sigma}}{\bar{\sigma}}\right)$$

since $\varrho(1) = \varrho(0) = 1$ so $\rho(y_{L_2}, 0) = 1-\bar{\sigma}$ and $\rho(y_{R_1}, 1) = \bar{\sigma}$. As q increases, $\gamma(0,R)$ decreases and $\gamma(\infty,R)$ increases, At $q=0$, agent's have the correctly specified model, so $\gamma(0,R) < 0$ and $\gamma(\infty,R) < 0$. As $q \to 1$,

$$\gamma(0,R) \to \log\left(\frac{\bar{\sigma}}{1-\bar{\sigma}}\right) < 0$$

and

$$\gamma(0,R) \to \log\left(\frac{1-\bar{\sigma}}{\bar{\sigma}}\right) > 0$$

so the desired cutoffs exist.

D.2 Over/underweighting: Epstein et al. (2010)

Epstein et al. (2010) considers an individual learning model where agents overweight beliefs towards the prior or towards the posterior. Specifically, an agent with prior p who would update her beliefs to $BU(p)$ instead updates to

$$(1-\alpha)BU(p) + \alpha p$$

for some $\alpha \leq 1$. When $\alpha = 0$, this is the correct model, for $\alpha > 0$ agents overweight the prior and for $\alpha < 0$, agents overweight new information. For simplicity of notation, suppose that $Pr(\sigma_L|\omega = R) = Pr(\sigma_R|\omega = L) = \sigma < 0.5$. In our framework, this is a model with a single agent type who only receives public signal σ and maps this signal to

$$\rho(\sigma, p) = \frac{\frac{\sigma(1-\alpha)}{(1-\sigma)(1-p)+ps} + \alpha}{\frac{1}{1-p} + \frac{1-2p}{1-p}\left(\frac{\sigma(1-\alpha)}{(1-\sigma)(1-p)+p\sigma} + \alpha\right)},$$

with

$$\rho(\sigma, 1) = \frac{\sigma}{(1-\alpha)(1-\sigma) + (1+\alpha)\sigma},$$

which implies that $\rho(\sigma, 1) = \lim_{p\to 1} \rho(\sigma, 1)$.[46]

Under this misspecification, whenever an agent with prior p_t updates their beliefs, the likelihood ratio becomes

$$\lambda_{t+1} = \frac{\frac{p_t\sigma(1-\alpha)}{(1-\sigma)(1-p_t)+p_t\sigma} + \alpha p_t}{1 - \frac{p_t\sigma(1-\alpha)}{(1-\sigma)(1-p_t)+p_t\sigma} - \alpha p_t}.$$

Therefore, the Bayes update is

$$p_{t+1} = \frac{p_t\sigma(1-\alpha)}{(1-\sigma)(1-p_t) + p_t\sigma} + \alpha p_t.$$

Therefore, the update rule from Epstein et al. (2010) can be represented in our framework.

Under this specification, the likelihood ratio update is

$$\lambda_{t+1}/\lambda_t = \frac{\frac{\sigma(1-\alpha)}{(1-\sigma)(1-p_t)+p_t\sigma} + \alpha}{\frac{(1-\alpha)(1-\sigma)}{(1-\sigma)(1-p_t)+p_t\sigma} + \alpha}$$

As $p \to 1$, the likelihood ratio update conveges to

$$\frac{1}{(1-\alpha)\frac{1-\sigma}{\sigma} + \alpha}$$

[46] Epstein et al. (2010) does not identify how signals are interpreted at 0 or 1, since beliefs are stationary at these points. In order to characterize asymptotic outcomes, the tools developed in this paper show ho the limit of the update rule as $p \to 0$ or 1 can be used to characterize asymptotic outcomes of the model in Epstein et al. (2010).

and as $p \to 0$, the likelihood ratio update converges to

$$\frac{\sigma(1-\alpha)}{1-\sigma} + \alpha$$

In an environment with symmetric binary signals,

$$\gamma(0, R) = \sigma \log[(1-\alpha)\frac{1-\sigma}{\sigma} + \alpha] + (1-\sigma)\log[(1-\alpha)\frac{\sigma}{1-\sigma} + \alpha],$$

and

$$\gamma(\infty, R) = \sigma \log \frac{1}{(1-\alpha)\frac{\sigma}{1-\sigma} + \alpha} + (1-\sigma)\log \frac{1}{(1-\alpha)\frac{1-\sigma}{\sigma} + \alpha}.$$