Misspecified Learning and Evolutionary Stability*

Kevin He[†] Jonathan Libgober[‡]

First version: December 20, 2020
This version: September 4, 2025

Abstract

We extend the indirect evolutionary approach to the selection of (possibly misspecified) models. Agents with different models match in pairs to play a stage game, where models define feasible beliefs about game parameters and about others' strategies. In equilibrium, each agent adopts the feasible belief that best fits their data and plays optimally given their beliefs. We define the stability of the resident model by comparing its equilibrium payoff with that of the entrant model, and provide conditions under which the correctly specified resident model can only be destabilized by misspecified entrant models that contain multiple feasible beliefs (that is, entrant models that permit inference). We also show that entrants may do well in their matches against the residents only when the entrant population is large, due to the endogeneity of misspecified beliefs. Applications include the selection of demand-elasticity misperception in Cournot duopoly and the emergence of analogy-based reasoning in centipede games.

Keywords: misspecified Bayesian learning, endogenous misspecifications, evolutionary stability, analogy classes

*We thank Cuimin Ba, Thomas Chaney, Sylvain Chassang, In-Koo Cho, Krishna Dasaratha, Andrew Ellis, Ignacio Esponda, Mira Frick, Drew Fudenberg, Alice Gindin, Ryota Iijima, Yuhta Ishii, Philippe Jehiel, Pablo Kurlat, Elliot Lipnowski, Jonny Newton, Filippo Massari, Andy Postlewaite, Philipp Sadowski, Alvaro Sandroni, Grant Schoenebeck, Joshua Schwartzstein, Philipp Strack, Carl Veller, and various conference and seminar participants for helpful comments. Byunghoon Kim provided excellent research assistance. Kevin He thanks the California Institute of Technology for hospitality when some of the work on this paper was completed, and also acknowledges the University Research Foundation Grant at the University of Pennsylvania for financial support. Jonathan Libgober thanks Yale University and the Cowles Foundation for their hospitality.

[†]University of Pennsylvania. Email: hesichao@gmail.com
[‡]University of Southern California. Email: libgober@usc.edu

1 Introduction

In many economic settings, people draw *misspecified inferences* about the world: they learn from data but exclude the true data-generating process from consideration. Esponda and Pouzo (2016) introduce Berk-Nash equilibrium to accommodate this observation — a solution concept where agents use data to infer the best-fitting mapping from actions to outcomes, out of a set of mappings that are all wrong. A line of research (discussed below) uses this and related solution concepts to study the implications of Bayesian learning under particular misspecifications, with most papers treating misspecifications as exogenously given.

When should we expect misspecified inference to take hold in a rational society, as assumed in much of this literature? A defining feature that distinguishes misspecified inference from other kinds of errors and biases is the use of data to form beliefs — how does this *belief endogeneity* affect its viability? We develop a framework to answer these questions from an evolutionary perspective, studying the objective equilibrium payoffs of different agents. Unlike contemporaneous work in single-agent settings (Fudenberg and Lanzani, 2023; Frick, Iijima, and Ishii, 2024), we focus on strategic interactions where payoffs depend on equilibrium play.

1.1 Summary of the Setup

Each agent is endowed with a *model* that contains free parameters; the parameter values correspond to feasible beliefs about the stage game and about others' strategies. The model's adherents think that, for some parameter value, the instantiated model describes the true stage game and opponents' behavior. They estimate the best-fitting parameter value, which determines their subjective preference. Models rise and fall in prominence based on objective equilibrium payoffs of their adherents, as higher payoffs confer greater evolutionary success.

Society consists of the adherents of multiple competing models, who match up to play the stage game every period. Agents can identify which subpopulation their opponent belongs to, and (correctly) know that the game they play does not depend on opponent type.[1] Our framework assumes that agents may face one of several possible stage games, so richer models can in principle help agents by allowing them to adapt their behavior to the true game. Conditional on a realized stage game, each agent forms in equilibrium a Bayesian belief about

[1]If the players thought that the stage game could change with the opponent, then this would give additional channels for biases to invade a rational society. Our framework focuses on how the belief endogeneity that plays a distinctive role in misspecified learning affects the viability of errors.

the game and about others' strategies using data from all their interactions, and plays a subjective best response to each opponent subpopulation given this belief.

We say model A is *evolutionarily stable* against model B if, for all sufficiently small population shares of model B, model A's equilibrium payoff (averaged across the distribution of stage games) is weakly higher than that of model B. This criterion is familiar from the literature on the *indirect evolutionary approach*, which considers evolution acting on some trait that determines agents' strategy choices, as opposed to acting on these choices directly. Our stability concepts reduce to standard notions under this approach when models stipulate a dogmatic belief about the stage game. Our main contribution is to study flexible models that contain multiple feasible beliefs and, hence, the role of belief endogeneity in stability.

1.2 Belief Endogeneity and the Viability of Misspecifications

Consider an agent with a flexible model that contains multiple feasible beliefs. Equilibrium belief is endogenous because the agent infers model parameters from data, and this data depends on population composition and opponents' strategies. By contrast, equilibrium belief is exogenous for an agent whose model contains only one feasible belief — by construction, data plays no role in shaping this (possibly distorted) belief. Our formal results identify two novel stability phenomena that can only arise with belief endogeneity:

1. Endogenous beliefs allow agents to adopt different subjective best responses in different games, so misspecified inference may confer greater strategic benefits than dogmatic beliefs.

2. Agents with a fixed misspecification may be weak when rare but strong when common.

Section 4.1 discusses the former point by characterizing environments where the correctly specified model is only evolutionarily fragile against invading models that contain multiple feasible beliefs. One part of our argument constructs an optimal misspecified model for invading a rational society. This misspecification resembles an "illusion of control" bias, where agents think the game's outcome only depends on their own strategy and not on the opponent's strategy. Adherents of this model end up adopting the optimal commitment against a correctly specified opponent, game by game.

But in some environments, there exists a single commitment strategy that is beneficial in every game, so belief endogeneity is not necessary for the invading model. The second part of our argument in Section 4.1 identifies a geometric condition that ensures any model with a fixed dogmatic belief across all stage games cannot outperform the correctly specified model. Putting everything together: when the geometric condition holds and the rational model fails to achieve the best commitment payoff in some stage game, misspecified inference is necessary for the rational model's fragility.

Section 4.2 then turns to the latter point and identifies a type of fluidity in a misspecified model's performance based on the population proportions. Two models are said to exhibit *stability reversal* if:

- Whenever model A is dominant, its adherents strictly outperform model B's adherents not only on average, but even conditional on the opponent's type; and
- Whenever model B is dominant, its adherents strictly outperform model A's adherents on average.

In the absence of belief endogeneity, the first condition would imply that A outperforms B regardless of the two subpopulations' sizes. But this no longer holds when belief is endogenous due to misspecified inference. The reason is that the data from the B-vs-B matches may induce more evolutionarily advantageous beliefs than the data from the B-vs-A matches.

1.3 Applications

Extending the indirect evolutionary approach to accommodate models with belief endogeneity lets us analyze new applications. Our companion paper, He and Libgober (2025), illustrates this point by studying the selection of misspecified higher-order beliefs in an incomplete-information Cournot duopoly game. In the present paper, Section 3 presents a simpler complete information duopoly game, while Section 5 considers the selection of *analogy classes* in extensive-form games (Jehiel, 2005), based on the payoffs for players with different analogy classes. Under analogy-based reasoning, players (incorrectly) believe opponents choose the same action distribution at all nodes within an analogy class and infer that distribution from empirical frequencies. Our approach predicts not only that analogy-based reasoning may invade a correctly-specified society, but also that the two can coexist. By solving for the corresponding stable population composition, we obtain sharp predictions on the relative prominence of analogy-based reasoning as a function of the underlying stage game.

1.4 Related Literature

Our paper contributes to the literature on misspecified Bayesian learning by proposing a framework to assess which specifications are more likely to persist based on their objective payoffs. Misspecified inference in such a framework leads to endogenous beliefs, which in turn generate new phenomena in payoff-based selection of biases. Most prior works on misspecified Bayesian learning, by contrast, take the misspecification as exogenous, studying the subsequent implications in both single-agent decision problems[2] and multi-agent games.[3] Several papers establish general convergence properties of misspecified learning.[4]

Our interest in endogenizing misspecifications using objective payoffs[5] contrasts with alternatives using subjective expectations of payoffs[6] or goodness-of-fit tests.[7] There are several other papers that also use objective payoffs as the selection criterion. Our work differs in that we assume agents update beliefs using Bayes' rule and focus on the selection between various misspecifications under Bayesian learning. Massari and Newton (2020) show that learning rules departing from Bayes' rule can achieve higher objective payoffs, while Massari and Newton (2024) show that maximizing a combination of accuracy and payoffs can improve performance along *both* dimensions. Similarly, Heller and Winter (2020) and Berman and Heller (2024) study the evolution of different belief-formation processes under a reduced-form (and possibly non-Bayesian) approach, considering arbitrary inference rules.

Two independent and contemporaneous papers, Fudenberg and Lanzani (2023) and Frick, Iijima, and Ishii (2024), also consider payoff-based criteria under Bayesian inferences for selecting misspecifications, but restrict attention to single-agent decision problems.[8] We differ

[2]See Nyarko (1991); Fudenberg, Romanyuk, and Strack (2017); Heidhues, Koszegi, and Strack (2018); He (2022); Cho and Libgober (2025). Also related is Fudenberg et al. (2024) who show how memory limitations yield inferences resembling those in misspecified learning models.

[3]See Bohren (2016); Bohren and Hauser (2021); Jehiel (2018); Molavi (2019); Dasaratha and He (2020); Ba and Gindin (2023); Frick, Iijima, and Ishii (2020); Murooka and Yamamoto (2023).

[4]See Esponda and Pouzo (2016); Esponda, Pouzo, and Yamamoto (2021); Frick, Iijima, and Ishii (2023); Fudenberg, Lanzani, and Strack (2021).

[5]A separate line of work that has used objective payoffs to endogenize misspecified inference, restricting attention to financial markets (Sandroni, 2000; Massari, 2020), while our approach applies to general strategic environments.

[6]See Montiel Olea, Ortoleva, Pai, and Prat (2022); Levy, Razin, and Young (2022); Gagnon-Bartsch, Rabin, and Schwartzstein (2021)

[7]See Cho and Kasa (2015, 2017); Ba (2025); Schwartzstein and Sunderam (2021); Lanzani (2025).

[8]Fudenberg and Lanzani (2023) study a framework where a continuum of agents with heterogeneous misspecifications arrive each period and learn from their predecessors' data. Frick, Iijima, and Ishii (2024) assign a *learning efficiency index* to every misspecified signal structure and conduct a robust comparison of

in highlighting that belief endogeneity can *strictly* expand the possibility for misspecifications to invade rational societies in strategic settings (relative to biased invaders who do not draw inferences).

Our framework of competition between different specifications for Bayesian learning is inspired by the evolutionary game theory literature. Relative to this literature, our contribution is to accommodate misspecified inference. We follow past work that also uses objective payoffs as the selection criterion for subjective preferences in games and decision problems (e.g., Dekel, Ely, and Yilankaya (2007), see also the surveys Robson and Samuelson (2011) and Alger and Weibull (2019)) and the evolution of constrained strategy spaces (Heller, 2015; Heller and Winter, 2016). Like us, Güth and Napel (2006) allow for stage-game heterogeneity, studying the ability to discriminate between these games.

2 Environment and Stability Concept

We start with our formal stability concept, defining *equilibrium zeitgeist* to determine the evolutionary fitness of specifications that coexist in a society. Our general setup allows agents to both learn about the fundamentals and draw inferences about others' strategies: indeed, misinference about opponent's strategy is central to our application in Section 5. But most of our results and applications concern a special case of the setup where agents correctly know others' strategies in equilibrium, so the focus is on misinference about fundamental uncertainty and the role of such misinference on evolutionary selection. In the main text we primarily focus on the steady-state characterization of equilibrium zeitgeists, but we provide a learning foundation for this solution concept in Appendix C.

2.1 Objective Primitives

Agents in a population repeatedly match to play a stage game, which is a symmetric two-player game with a common, metrizable strategy space $\mathbb{A}$. There is a set of possible states of nature $G \in \mathcal{G}$, called *situations*. The strategy choices $a_i, a_{-i} \in \mathbb{A}$ of i and $-i$, together with the situation, stochastically generate consequences $y_i, y_{-i} \in \mathbb{Y}$ from a metrizable space $\mathbb{Y}$. Each agent i's consequence y_i determines their utility, according to a common utility function $\pi : \mathbb{Y} \to \mathbb{R}$, which we take to be Borel measurable with respect to the sigma algebra generated

welfare under different misspecifications.

by the topology on $\mathbb{Y}$. The objective distribution over consequences is $F^\bullet(a_i, a_{-i}, G) \in \Delta(\mathbb{Y})$, with an associated density or probability mass function denoted by $f^\bullet(a_i, a_{-i}, G)$, where $f^\bullet(a_i, a_{-i}, G)(y) \in \mathbb{R}_+$ for each $y \in \mathbb{Y}$. We suppress G from $f^\bullet$ and $F^\bullet$ when $|\mathcal{G}| = 1$. We allow for $\mathbb{Y}$ to be general outside of the previous technical restrictions.

This setup captures mixed strategies (if $\mathbb{A}$ is the set of mixtures over some pure actions), incomplete-information games (if S is a space of private signals, A a space of actions, and $\mathbb{A} = A^S$ is the set of signal-contingent actions), and even asymmetric games. For the latter, we consider the "symmetrized" version where each player is placed into each role with equal probability (see Section 5 for one application where agents play an asymmetric game).

In addition, at the end of a match where the strategy profile (a_i, a_{-i}) is played, each agent i observes a *monitoring signal* m_i about the opponent's strategy which only depends on a_{-i} and not on a_i or the situation. Let $\mathbb{M}$ be the space of monitoring signals, and let the objective distribution over monitoring signals when the opponent plays a_{-i} be given by the density or probability mass function $\varphi^\bullet(a_{-i})$, where $\varphi^\bullet(a_{-i}) : \mathbb{M} \to \mathbb{R}_+$. The monitoring signal m_i is payoff irrelevant and is generated independently of the consequences. Our framework separately defines the monitoring signal for the expositional simplicity of introducing the special case of environments with strategic certainty (where the monitoring signal is perfect) and for discussing the learning foundation of equilibrium in such environments (where we make the monitoring signal "almost perfect").

2.2 Models and Parameters

Throughout this paper, we will take the strategy space $\mathbb{A}$, the set of consequences $\mathbb{Y}$, the utility function over consequences π, the set of monitoring signals $\mathbb{M}$ and the strategy monitoring structure $\varphi^\bullet$ to be common knowledge among the agents. But, agents are unsure about how play in the stage game translates into consequences: that is, they have *fundamental uncertainty* about the function $(a_i, a_{-i}) \mapsto F^\bullet(a_i, a_{-i}, G)$. While we assume that the situation G is unobserved, we allow agents to draw inferences about it by observing the consequences from the matches they face. Agents may also be unsure about which strategies others use (*strategic uncertainty*), but they could get some information about others' play through their consequences and monitoring signals.

We focus on the case where society consists of two[9] observably distinguishable groups of

[9]We view the case of two groups of agents with different models as the natural starting point, though it is

agents, A and B, who may behave differently in the stage game due to different beliefs about how y is generated and about the strategies of their opponents. The two groups of agents entertain different *models* of the world that help resolve their fundamental uncertainty and strategic uncertainty. A model Θ is a collection of *parameters* (a_A, a_B, F) with $a_A, a_B \in \mathbb{A}$ and $F : \mathbb{A}^2 \to \Delta(\mathbb{Y})$. So, each parameter specifies conjectures a_A, a_B about how group A and group B opponents will act when playing against the agent. It also contains a conjecture F about how strategy profiles translate into consequences for the agent. So, we can view each model as a subset of $\mathbb{A}^2 \times (\Delta(\mathbb{Y}))^{\mathbb{A}^2}$. We assume the marginal of the model on $(\Delta(\mathbb{Y}))^{\mathbb{A}^2}$ is indexed by some $\gamma \in \Gamma$ for a metric space Γ acting as an indexing set, so this marginal can be written as $\{F_\gamma; \gamma \in \Gamma\}$. For each F_γ that is part of some parameter and for every $(a_i, a_{-i}) \in \mathbb{A}^2$, we suppose $F_\gamma(a_i, a_{-i})$ is a Borel measure on $\mathbb{Y}$ and it has associated with it a density or probability mass function $f_\gamma(a_i, a_{-i}) : \mathbb{Y} \to \mathbb{R}_+$. We also suppose that for every (a_i, a_{-i}), the map $\gamma \mapsto \mathbb{E}_{y \sim F_\gamma(a_i, a_{-i})}[\pi(y)]$ is Borel measurable.[10]

Each agent enters society with a persistent model, which depends entirely on whether she is from group A or group B. We refer to the agents who are endowed with a given model the *adherents* of that model. We call a model *correctly specified* if it is a superset of $\mathbb{A}^2 \times \{F^\bullet(\cdot, \cdot, G) : G \in \mathcal{G}\}$, so the agent can make unrestricted inferences about others' strategies and does not rule out the correct data-generating process $F^\bullet(\cdot, \cdot, G)$ for any situation G. We call $\Theta = \mathbb{A}^2 \times \{F^\bullet(\cdot, \cdot, G) : G \in \mathcal{G}\}$ the *minimal correctly specified* model. A model may exclude the true $F^\bullet(\cdot, \cdot, G)$ that produces consequences, at least in some situation G, or it may exclude some strategies as feasible conjectures of others' play. In this case, the model is *misspecified*.

An important special case of the setup focuses purely on misinference about fundamental uncertainty.

Definition 1. An environment has *strategic certainty* if

- $\mathbb{M} = \mathbb{A}$ and $\varphi^\bullet(a_{-i})$ puts probability 1 on a_{-i} for every $a_{-i} \in \mathbb{A}$,
- The model of each group $g \in \{A, B\}$ is of the form $\mathbb{A}^2 \times \mathcal{F}_g$ for some $\mathcal{F}_g \subseteq (\Delta(\mathbb{Y}))^{\mathbb{A}^2}$, and

straightforward to generalize Definition 2 to the case of more than two groups.

[10] Note that this measurability property would follow from the measurability of the mapping $\gamma \mapsto f_\gamma(a_i, a_{-i})(y)$ for each fixed (a_i, a_{-i}, y), under some further restrictions necessary to apply Fubini's theorem to the function $(a_i, a_{-i}, y) \mapsto \pi(y) f(a_i, a_{-i})(y)$.

- Every $y \in \mathbb{Y}$ with the property that $f^{\bullet}(a'_i, a'_{-i}, G)(y) > 0$ for some $a'_i, a'_{-i} \in \mathbb{A}$ and some $G \in \mathcal{G}$ also satisfies $f(a'_i, a''_{-i})(y) > 0$ for every $a''_{-i} \in \mathbb{A}$ and every f that is the density or probability mass function of some conjecture $F \in \mathcal{F}_A \cup \mathcal{F}_B$.

In environments with strategic certainty, monitoring signals perfectly reveal opponent's strategy and all agents can make unrestricted strategic inferences. Combined with the assumption that all consequences that agents can observe when they play a'_i have positive likelihood under any of their feasible conjectures about fundamental uncertainty, this will imply that agents hold correct beliefs about their opponents' strategies in the equilibrium concept that we define below. In such environments, the focus is on different groups' feasible conjectures about the fundamental uncertainty, $\mathcal{F}_A$ and $\mathcal{F}_B$. We will therefore sometimes omit mention of the monitoring signal when analyzing environments with strategic certainty.

In environments with strategic certainty, a model $\Theta = \mathbb{A}^2 \times \mathcal{F}$ with $|\mathcal{F}| = 1$ is called a *singleton* model. This terminology refers to the fact that such models stipulate a single dogmatic belief about the fundamental uncertainty (though agents still make flexible inferences about others' strategies). Adherents of singleton models do not draw inferences about the game from data. Since the situation is itself unobserved, this implies such agents also do not change their preferences with the situation, since this is only possible through drawing different inferences in different situations.

2.3 Zeitgeists

To study competition between two models, we must describe the social composition and interaction structure in the society where learning takes place. We have in mind a setting where each agent plays the stage game with a uniformly random opponent in every period and uses their personal experience in these matches to calibrate the most accurate parameter within their model. A zeitgeist describes the corresponding landscape.

Definition 2. Fix models Θ_A and Θ_B. A *zeitgeist* $\mathfrak{Z} = (\mu_A(G), \mu_B(G), p, a(G))_{G \in \mathcal{G}}$ consists of: (1) for each situation G, a belief over parameters for each model, $\mu_A(G) \in \Delta(\Theta_A)$ and $\mu_B(G) \in \Delta(\Theta_B)$; (2) relative sizes of the two groups in the society, $p = (p_A, p_B)$ with $p_A, p_B \geq 0$, $p_A + p_B = 1$; (3) for each situation G, each group's strategy when matched against each other group, $a = (a_{AA}(G), a_{AB}(G), a_{BA}(G), a_{BB}(G))$ where $a_{g,g'}(G) \in \mathbb{A}$ is the strategy that an adherent of Θ_g plays against an adherent of $\Theta_{g'}$ in situation G.

A zeitgeist outlines the beliefs and interactions among agents with heterogeneous models living in the same society. Part (1) captures the belief of each group. Part (2) determines the relative prominence of each model. Agents are matched uniformly at random across the entire society, so an agent from group g has probability p_g of being matched with an opponent from their own group and a complementary chance of being matched with an opponent from the other group.[11] Part (3) describes behavior in the society. Note that a zeitgeist describes each group's situation-contingent belief and behavior, since agents may infer different parameters and thus adopt different subjective best replies in different situations. However, we emphasize that since situations are not directly observed, they only influence strategies by changing the distribution of the agents' consequences (and hence their beliefs).

2.4 Equilibrium Zeitgeists

A model's fitness corresponds to the equilibrium payoff of its adherents. An equilibrium zeitgeist (EZ) requires behavior to be optimal given beliefs and beliefs to best fit the data given behavior. As we make clear in the learning foundation for EZs in Appendix C, this equilibrium concept relates to steady states in a society of long-lived Bayesian learners who use consequences and monitoring signals to make Bayesian inferences among parameters in their model, assuming there is convergence in beliefs and behavior. In the steady state, agents choose subjectively optimal strategies given their beliefs about others' strategies and about the stage game.

We now formalize this criterion. For two distributions over consequences and monitoring signals, $\Phi, \Psi \in \Delta(\mathbb{Y} \times \mathbb{M})$ with density or probability mass functions ϕ, ψ, define the Kullback-Leibler divergence (KL divergence) from Ψ to Φ as $D_{KL}(\Phi \parallel \Psi) := \int \phi(y,m) \ln\left(\frac{\phi(y,m)}{\psi(y,m)}\right) d(y,m)$. Recall that every data-generating process F, like the true fundamental $F^{\bullet}(\cdot,\cdot,G)$, outputs a distribution over consequences for every strategy profile, $(a_i, a_{-i}) \in \mathbb{A}^2$.

Definition 3. A zeitgeist $\mathfrak{Z} = (\mu_A(G), \mu_B(G), p, a(G))_{G \in \mathcal{G}}$ is an *equilibrium zeitgeist (EZ)* if, for every $G \in \mathcal{G}$ and $g, g' \in \{A, B\}$, $a_{g,g'}(G) \in \underset{\hat{a} \in \mathbb{A}}{\arg\max}\, \mathbb{E}_{(a_A, a_B, F) \sim \mu_g(G)} \left[\mathbb{E}_{y \sim F(\hat{a}, a_{g'})}(\pi(y)) \right]$

[11] He and Libgober (2025) contains an example that varies the matching assortativity and discusses how this affects the selection of biases.

and, for every $g \in \{A, B\}$, the belief $\mu_g(G)$ is supported on

$$\underset{(\hat{a}_A, \hat{a}_B, \hat{F}) \in \Theta_g}{\arg\min} \left\{ \begin{array}{c} (p_g) \cdot D_{KL}(F^\bullet(a_{g,g}(G), a_{g,g}(G), G) \times \varphi^\bullet(a_{g,g}(G)) \parallel \hat{F}(a_{g,g}(G), \hat{a}_g) \times \varphi^\bullet(\hat{a}_g)) \\ +(1-p_g) \cdot D_{KL}(F^\bullet(a_{g,-g}(G), a_{-g,g}(G), G) \times \varphi^\bullet(a_{-g,g}(G)) \parallel \hat{F}(a_{g,-g}(G), \hat{a}_{-g}) \times \varphi^\bullet(\hat{a}_{-g})) \end{array} \right\}$$

where $-g$ means the group other than g and $\times$ indicates the product between a distribution on $\mathbb{Y}$ and a distribution on $\mathbb{M}$. When $p_g = 0$ or $(1 - p_g) = 0$ but it is multiplied by infinity, we use the convention that $0 \cdot \infty = \infty$.

This definition requires agents from each group g to choose a subjective best response against their opponents, given the belief μ_g about the fundamental uncertainty and strategic uncertainty. No matter which group the agent is matched against, these choices are always made to selfishly maximize their (individual) subjective utility function. Each agent's belief μ_g is supported on the parameters in their model that minimize a weighted KL-divergence objective in situation G, with the data from each type of match weighted by the probability of confronting this type of opponent. The use of KL-divergence minimization as the inference procedure is standard in the misspecified Bayesian learning literature (such as in Esponda and Pouzo (2016)) and goes back to the basic result from Berk (1966) that the Bayesian posteriors under misspecification concentrate in the long run on the KL-divergence minimizers. We assume inference occurs separately across situations. This reflects situation persistence, with agents having enough data to establish new beliefs and behavior before the situation changes. Our learning foundation in Appendix C justifies this situation-by-situation updating, but we omit the details here as it otherwise plays no role in our results.

In general, agents choose their best-fitting model parameters based on two kinds of data: consequences and monitoring signals. In environments with strategic certainty, there is no equilibrium zeitgeist where the equilibrium belief $\mu_g(G)$ for any group g in any situation G puts positive weight on any parameter $(\hat{a}_A, \hat{a}_B, \hat{F})$ where $\hat{a}_{g'} \neq a_{g',g}(G)$ for any groups $g' \in \{A, B\}$. This is because any such parameter has a weighted KL divergence of infinity (given that $\varphi^\bullet$ is perfectly informative about the opponent's strategy), whereas parameters with $\hat{a}_A = a_{A,g}(G)$ and $\hat{a}_B = a_{B,g}(G)$ have finite weighted KL divergence. So, in environments of strategic certainty, we can view beliefs in equilibrium zeitgeists $\mu_g(G)$ as simply beliefs over the fundamental uncertainty $\mathcal{F}_g$, with $\mu_g(G)$ supported on

$$\underset{\hat{F}\in\mathcal{F}_g}{\arg\min}\left\{\begin{array}{c}(p_g)\cdot D_{KL}(F^{\bullet}(a_{g,g}(G),a_{g,g}(G),G)\parallel\hat{F}(a_{g,g}(G),a_{g,g}(G)))\\+(1-p_g)\cdot D_{KL}(F^{\bullet}(a_{g,-g}(G),a_{-g,g}(G),G)\parallel\hat{F}(a_{g,-g}(G),a_{-g,g}(G)))\end{array}\right\}.$$

In environments with strategic certainty, we will therefore omit reference to beliefs about others' strategies in describing zeitgeists and simply view $\mu_g(G)$ as an element in $\Delta(\mathcal{F}_g)$.

2.5 Evolutionary Stability of Models

Given a distribution $q\in\Delta(\mathcal{G})$ and an EZ, we define the *fitness* of each model as the expected objective payoff of its adherents in the EZ when G is drawn according to q. We have in mind an evolutionary story where the relative success of the two models depends on their relative fitness: for instance, agents may play a large number of games in different periods possibly facing different situations over time, and models of those agents with higher total objective payoffs are more likely to be adopted in the next generation.[12] Given this notion of fitness, our question of interest is: Can the adherents of a *resident model* Θ_A, starting at a position of social prominence, always repel an invasion from a small mass of agents who adhere to an *entrant model* Θ_B?

Evolutionary stability depends on the fitness of models Θ_A,Θ_B in EZs with $p_A=1-\epsilon, p_B=\epsilon$ for small $\epsilon>0$.

Definition 4. Say Θ_A is *evolutionarily stable [fragile]* against Θ_B if there exists some $\bar{\epsilon}>0$ so that for every $0<\epsilon\le\bar{\epsilon}$, there is at least one EZ with models Θ_A,Θ_B, $p=(1-\epsilon,\epsilon)$ and in all such EZs, Θ_A has a weakly higher [strictly lower] fitness than Θ_B.

Evolutionary stability is when Θ_A has higher fitness than Θ_B in all EZs, and evolutionary fragility is when Θ_A has lower fitness in all EZs.[13] These two cases give sharp predictions about whether a small share of entrant-model invaders might grow in size, across all equilibrium selections. We fix these rather stringent definitions of stability and fragility, and focus on showing in Section 4 how the belief endogeneity in models can generate new stability /

[12]One subtlety is that fitness maximization may require not maximizing expected payoffs, but rather some other function of the distribution of payoffs, if shocks can be correlated (Robson, 1996). However, our microfoundation in Appendix C posits that situations are fixed for long stretches of time, with no correlated shocks across matches, making the expectation an appropriate measurement of fitness.

[13]If the set of EZs is empty, then Θ_A is neither evolutionarily stable nor evolutionarily fragile against Θ_B.

fragility phenomena. A third possible case, where Θ_A has lower fitness than Θ_B in some but not all EZs, corresponds to a situation where the entrant model may or may not grow in the society, depending on the equilibrium selection.

2.6 Discussion

We clarify some important aspects of our framework before proceeding further.

2.6.1 Equilibrium Zeitgeist Existence

Implicit in our definition of evolutionary stability and fragility is that we do not have to worry about whether EZs exist in the first place. However, this is not a given, as we have not imposed the continuity and integrability conditions necessary to ensure existence and the well-definedness of expected utilities, KL divergences, and best responses. These issues are familiar from past work, so we do not belabor them in the main text. Appendix B provides sufficient conditions that guarantee existence, as well as that the set of EZs is upper hemicontinuous in population shares. This latter result is particularly useful for moving between the $\epsilon > 0$ case and the limit as $\epsilon \to 0$ in the definition of stability. Throughout, we focus on the case where expected utilities, KL divergences, and best responses are well defined and EZs exist, and refer interested readers to the appendix for the technical conditions that guarantee these.

2.6.2 Comparison with Other Evolutionary Frameworks

We apply the "indirect evolutionary approach" (see Robson and Samuelson (2011)) to settings where agents can draw inferences (especially misspecified inferences). In environments with strategic certainty with singleton models and $|\mathcal{G}| = 1$, our framework reduces to the setup studied by the literature on preference evolution Alger and Weibull (2019), since singleton models are equivalent to subjective preferences. But in general, models with multiple parameters allow agents to adapt their beliefs (which determine their subjective preferences) endogenously. Allowing for multiple situations is the most direct way for inference to be beneficial. With only a single situation, any steady-state outcome that emerges for some model can also emerge with a singleton model. That said, one could also study settings with multiple situations without inference (see Güth and Napel (2006) for an example of such an exercise).

2.6.3 Framework Assumptions

An important assumption is that agents (correctly) believe the economic fundamentals (represented by G) do not vary depending on which group they are matched against. That is, the mapping $(a_i, a_{-i}) \mapsto \Delta(\mathbb{Y})$ describes the stage game that they are playing, and agents know that they always play the same stage game even though opponents from different groups may use different strategies in the game. As a result, the agent's experiences in games against both groups of opponents jointly resolve the same fundamental uncertainty about the environment.[14] If adherents could believe that the fundamentals can change depending on their opponent, then this would give a trivial way for in-group preferences to emerge and also trivialize the question of which errors could invade. For expositional simplicity, we do not consider this elaboration.

Our framework assumes that agents can identify which group their matched opponent belongs to, though we do not assume that agents know the data-generating processes contained in other models or that they are capable of making inferences using other models. (In other words, models are primitives and cannot be changed even after agents see their opponents' actions. Players do not "read into" what others do when learning.) Some other works in the literature on the indirect evolutionary approach (e.g., Dekel et al. (2007)) consider a more general setup where agents only observe their opponent's group membership with some probability in each match, and receive no information about their opponent with the complementary probability. We expect the main insights to carry through when the probability of observation is high enough. At the other end of the spectrum, if agents never observe whether their matched opponent is from group A or group B, then results can change dramatically. For instance, consider an environment of strategic certainty with only one situation. There is no EZ where the minimal correctly specified resident model has strictly lower fitness than an entrant model. This is because both models face the same distribution of opponents' strategies and must play a single strategy against this distribution. If the correctly specified model has strictly lower fitness, then its adherents are not playing an objective best response, which contradicts the fact that they correctly know the game and have correct beliefs about others' strategies in EZ under strategic certainty.

[14]We note that play between two groups g and g' is not a Berk-Nash equilibrium Esponda and Pouzo (2016), since adherents from one group draw inferences about the game's parameters from the matches against the other group, which may adopt a different strategy. A Berk-Nash equilibrium between groups g and g' would require inferences to *only* be made from data generated in the match between g and g'.

Even as agents update their beliefs and optimize their behavior, population proportions p_A and p_B remain fixed. We imagine a world where the relative prominence of models changes much more slowly than the rate of convergence to an EZ. This assumption about the relative rate of change in the population sizes follows the previous work on evolutionary game theory (See Sandholm (2001) or Dekel et al. (2007)).

3 Illustrative Example

In this section, we use an example to illustrate our framework, explain the different stability implications of entrant models that expand the resident model versus entrant models that shift the resident model, and preview an "illusion of control" entrant model that will play a key role in our general results.

Consider an environment with strategic certainty and only one situation (so we omit mention of G). Suppose each player i is a Cournot duopolist with constant marginal cost c. Each player i simultaneously chooses a quantity a_i. A random market price $P = \beta^\bullet - r^\bullet(a_1 + a_2) + \varepsilon$ realizes and is observed by both players, where $r^\bullet > 0$, $\beta^\bullet > c$ are constants and ε is a mean-zero random variable with full support on $\mathbb{R}$. The utility of player i is $a_i \cdot (P - c)$.

Mapping back into the formalism from Section 2, we can let i's consequence be $y_i = (a_i, P)$ so that i's utility is a function of the consequence with $\pi(y_i) = \pi(a_i, P) = a_i \cdot (P - c)$. The true distribution over consequences $F^\bullet(a_i, a_{-i})$ given the strategy profile (a_i, a_{-i}) is such that the first dimension of y_i is always a_i, while the second dimension is distributed according to $\beta^\bullet - r^\bullet(a_i + a_{-i}) + \varepsilon$.

For $r > 0$ and $\beta \in \mathbb{R}$, let $F_{r,\beta}$ represent the conjecture about the stage game where for each strategy profile (a_i, a_{-i}), $F_{r,\beta}(a_i, a_{-i})$ is the distribution over consequences with the first dimension always being a_i and the second dimension being distributed according to $\beta - r(a_i + a_{-i}) + \varepsilon$. The residents (group A) are correctly specified, so their model has $\mathcal{F}_A = \{F_{r^\bullet,\beta} : \beta \in \mathbb{R}\}$. We say the entrants (group B) have a *slope perception* of $\hat{r}$ if their model has $\mathcal{F}_B = \{F_{\hat{r},\beta} : \beta \in \mathbb{R}\}$. The idea is that all agents hold dogmatic beliefs about the slope of the demand curve and use market-price data to make inferences about the intercept of the demand curve. The residents' belief about the slope is correct, while the entrants' belief is possibly wrong.

When entrants misperceive the slope to be $\hat{r} \neq r^\bullet$, they misunderstand how quantity choices affect market prices and thus misinfer the intercept of the demand curve. This in turn distorts their behavior: we can show that an agent who believes the slope and intercept of the demand curve to be r and β has a subjective best response of $\frac{\beta-c-ra_{-i}}{2r}$ when their opponent plays a_{-i}. Proposition 1 summarizes how the entrants' slope misperception affects their equilibrium behavior and welfare in the equilibrium zeitgeist with $p_A = 1$, $p_B = 0$.

Proposition 1. *Suppose $p_A = 1, p_B = 0$, and the residents are correctly specified.*

1. *In every equilibrium zeitgeist, we have $a_{AA} = \frac{\beta^\bullet - c}{3r^\bullet}$ and the fitness of the residents is $\frac{(\beta^\bullet - c)^2}{9r^\bullet}$.*
2. *In every equilibrium zeitgeist, the fitness of the entrants is a function of their equilibrium strategy: $\frac{1}{2}[(a_{BA}) \cdot (\beta^\bullet - c) - (a_{BA})^2 \cdot (r^\bullet)]$. In particular, the entrant's fitness strictly decreases in the distance between a_{BA} and the Stackelberg strategy,[15] $a_{stack} = \frac{\beta^\bullet - c}{2r^\bullet}$.*
3. *Suppose the entrants have slope perception $\hat{r}$. Then in every equilibrium zeitgeist, we have $a_{BA} = \frac{\beta^\bullet - c}{2\hat{r} + r^\bullet}$.*

We conclude with the following three observations, motivated by this result:

(1) Local vs. global mutations of the correctly specified model. Since $a_{BA} = \frac{\beta^\bullet - c}{2\hat{r} + r^\bullet}$, entrants who correctly perceive $\hat{r} = r^\bullet$ will play $a_{BA} = \frac{\beta^\bullet - c}{3r^\bullet}$ and have the same fitness as the residents. As this is lower than the Stackelberg strategy $a_{\text{stack}} = \frac{\beta^\bullet - c}{2r^\bullet}$, the fitness of the entrants strictly decreases as $\hat{r}$ increases above $r^\bullet$ and strictly increases as $\hat{r}$ decreases below $r^\bullet$ up until $\hat{r} = r^\bullet/2$. Suppose we consider the set of possible entrants that have a "local" mutation relative to the correctly specified model: that is, entrants who have a slope perception $\hat{r}$ with $|\hat{r} - r^\bullet| \leq \delta$ for some small $\delta > 0$ with $\delta < r^\bullet/2$. Among this family of possible entrants, the entrant with the slope perception $\hat{r} = r^\bullet - \delta$ has the highest fitness in equilibrium. But, if we do not limit the entrants to only have these local mutations, then the entrant who has the slope perception $\hat{r} = r^\bullet/2$ would have an even higher fitness.

(2) Shifts vs. expansions of the correctly specified model. The model of the entrants who have a slope misperception $\hat{r} \neq r^\bullet$ can be viewed as a "shift" of the correctly specified model. Indeed, the entrants have $\mathcal{F}_B = \{F_{\hat{r},\beta} : \beta \in \mathbb{R}\}$ while the residents have

[15] This strategy is the one chosen by the first-mover in the game where each player moves sequentially, with the second-mover observing the action of the first mover—in other words, in cases where one player can commit to an action and the other player chooses a best reply to that action.

$\mathcal{F}_A = \{F_{r^\bullet,\beta} : \beta \in \mathbb{R}\}$, two disjoint sets of conjectures about how consequences are generated in the stage game. Fudenberg and Lanzani (2023) study the stability of models under local mutations, but their notion of local mutation is that of a local *expansion*, where the mutated model contains all the parameters of the resident models and also some new parameters. In this Cournot example, if the entrants simply entertain more possible values for the slope of the demand curve than the residents, that is to say the entrants' model has $\mathcal{F}_B = \{F_{r,\beta} : |r - r^\bullet| < \delta, \beta \in \mathbb{R}\}$, then there is an EZ where the entrants and the residents come to the same (correct) beliefs about both the slope and intercept of the demand curve, play the same strategies, and have the same fitness.

Going beyond this example, it is not difficult to see that in any environment with strategic certainty, the correctly specified model is never evolutionarily fragile against any expansions of it for the same reason. If the resident model is evolutionarily fragile against an entrant model, some of the feasible beliefs under the former must be impossible under the latter.

(3) An "illusion of control" entrant model that maximizes entrant fitness. There is another entrant model outside of the slope misperception class discussed so far that also maximizes entrant fitness across all possible entrant models. This is a singleton model with $\mathcal{F}_B = \{F^*\}$, where the conjectured distribution of consequences $F^*(a_i, a_{-i})$ only depends on a_i and not on a_{-i}. In particular, $F^*(a_i, a_{-i})$ specifies that the first dimension of y_i is always a_i, and the second dimension is distributed according to $\beta^\bullet - r^\bullet(a_i + \frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet}) + \varepsilon$. The conjecture F^* features "illusion of control," for it stipulates that the distribution of market prices only depends on i's strategy and not on that of $-i$. In particular, it says that whatever strategy the opponent actually chooses, the realized market price distribution when i chooses a_i is the objective market price distribution when $-i$ plays the rational best response against it. It is easy to see that under the belief F^*, i's strictly dominant strategy is to choose the Stackelberg strategy $a_{BA} = a_{\text{stack}} = \frac{\beta^\bullet - c}{2r^\bullet}$ in every EZ. The rational residents must play the rational best response against it, so the entrants obtain the Stackelberg payoff as their fitness. In the next section, we construct a similar illusion of control model for more general environments to find the highest possible fitness among all entrants when the residents are correctly specified.

4 Stability Implications of Belief Endogeneity

In this section, we focus on environments with strategic certainty to illustrate some stability phenomena that distinguish misspecified inference from dogmatic beliefs in our framework. The main novelty of our framework relative to past work on the indirect evolutionary approach is that agents' beliefs about the game (and hence, subjective preferences) are endogenously determined. We showcase some of the unique implications of this belief endogeneity.

Belief endogeneity adds new ways for biased individuals to develop strategic commitments in games. First, unlike agents with fixed subjective preferences, misspecified learners with a fixed model can develop situation-specific commitments that are better tailored to the stage game. We show this mechanism expands the scope for invading rational societies. Second, misinference can induce different beliefs for a misspecified agent depending on who they most frequently interact with. This leads to new stability phenomena and adds nuance to extrapolations of the welfare implications of a misspecified model across different societies, relative to that of a distorted subjective preference.

4.1 When Is Misinference Necessary to Defeat Rationality?

Our first result characterizes when misspecified models can *only* invade a rational society when inference is possible. More precisely, when does there exist a distribution over situations such that the correctly specified model is not evolutionarily fragile against any singleton model, but it is evolutionarily fragile against some models with multiple parameters?

When the stage game is fixed, the preference evolution literature has long recognized that commitment to the game's Stackelberg strategy can allow entrants to outperform rational residents (see, for example, Section 2.5 of Robson and Samuelson (2011)). In our setting, if there is only one situation and the highest symmetric Nash equilibrium payoff is lower than the game's Stackelberg payoff, it is straightforward to show that the correctly specified model is evolutionarily fragile against any singleton model that misperceives the Stackelberg strategy to be strictly dominant. Adherents of this entrant model play the Stackelberg strategy against every opponent and enjoy strictly higher fitness than the rational residents when the entrant population share is close to zero.

When there are multiple situations, the analogous conclusion that the rational residents must have the Stackelberg payoff situation-by-situation to avoid invasion by a misspecified entrant requires the existence of an entrant who can behave differently in different situations, since the Stackelberg strategy can vary by situation. The proof of the first part of Theorem 1 constructs such an entrant, using the same "illusion of control" idea from Section 3. The entrant makes inferences among multiple parameters in their model, where the different parameters are different illusion-of-control beliefs meant to be adopted in different situations. Under the suitable beliefs for situation G, the agent thinks their consequence is solely controlled by their own action and views the Stackelberg strategy for situation G as strictly dominant. Also, the entrant's model is constructed so that there is no equilibrium zeitgeist where entrants adopt a belief meant for situation G' in a different situation $G \neq G'$, as inference from data would cause the entrants to revise their beliefs in favor of a better-fitting parameter in such a scenario.

The second part of Theorem 1 characterizes distributions over situations so that no singleton entrant model can invade the correctly specified residents. In some environments with multiple situations, singleton models that are unable to make inferences and adapt to the different situations can nevertheless still obtain higher fitness than the rational residents. In the Cournot duopoly setup from Section 3, for instance, we know that entrants with the slope misperception $\hat{r} = r^\bullet/2$ can outperform the residents. But now suppose there are two situations with true slope coefficients of $r^\bullet = 1$ and $r^\bullet = 1.001$, equally likely. Then the singleton entrant model with the slope misperception $\hat{r} = 1/2$ continues to outperform the residents on average, across the two situations. The content of the second part of Theorem 1 is to provide a condition that ensures the multiple situations are "sufficiently different" strategically so that the correctly specified resident model is not evolutionarily fragile against any singleton model.

Theorem 1 is stated for an environment with a finite strategy space. We require some notation. For $F : \mathbb{A}^2 \to \Delta(\mathbb{Y})$, let $U_i(a_i, a_{-i}, F)$ represent i's expected payoff under the strategy profile (a_i, a_{-i}) if consequences are generated by F, that is $U_i(a_i, a_{-i}, F) := \mathbb{E}_{y \sim F(a_i, a_{-i})}[\pi(y)]$. In each situation G, let $v_G^{\text{NE}} \in \mathbb{R}$ be the highest symmetric Nash equilibrium payoff in G, when agents choose strategies from $\mathbb{A}$. For each $a_i \in \mathbb{A}$, let $\underline{\text{BR}}(a_i, G)$ be a rational best response against the strategy a_i in situation G, breaking ties *against* the user of a_i. Let $\bar{v}_G \in \mathbb{R}$ be the Stackelberg equilibrium payoff in situation G, breaking ties against the Stackelberg leader, i.e.,

$$\bar{v}_G := \max_{a_i} U_i(a_i, \underline{\text{BR}}(a_i, G), F^\bullet(G)). \tag{1}$$

Call the strategy $\bar{a}_G$ that maximizes Equation (1) the Stackelberg strategy in situation G. We assume the Stackelberg strategy is unique in each situation, and furthermore that there is a unique rational best response to $\bar{a}_G$ in each situation G', where possibly $G \neq G'$. Finally, for $b : \mathbb{A} \rightrightarrows \mathbb{A}$ a subjective best-response correspondence (which may be induced by some belief about how consequences are distributed under different strategy profiles in the stage game), let v_G^b denote the worst equilibrium payoff of an agent with the best-response correspondence b when she plays against a rational opponent in situation G. That is, $v_G^b \in \mathbb{R}$ is i's lowest payoff across all strategy profiles (a_i, a_{-i}) such that $a_i \in b(a_{-i})$ and a_{-i} is a rational response to a_i in situation G.[16]

We impose two identifiability conditions:

Definition 5. *Situation identifiability* is satisfied if for every $a_i, a_{-i} \in \mathbb{A}$ and $G \neq G'$, we have $F^{\bullet}(a_i, a_{-i}, G) \neq F^{\bullet}(a_i, a_{-i}, G')$. *Stackelberg identifiability* is satisfied if whenever $G \neq G'$ and a_{-i}, a'_{-i} are rational best responses to $\bar{a}_G$ in situations G and G', we have $F^{\bullet}(\bar{a}_G, a_{-i}, G) \neq F^{\bullet}(\bar{a}_G, a'_{-i}, G')$.

Under situation identifiability, a minimal correctly specified agent can identify the true situation. Under Stackelberg identifiability, playing the Stackelberg strategy $\bar{a}_G$ for any situation G generates consequence data that can statistically distinguish whether the true situation is G or not, provided the opponent chooses a rational best response to the strategy for the true situation.

The following result presents our characterization of when misinference is required for misspecified models to outperform rationality, for some distribution over situations. The first part of the result says, under identifiability assumptions and other regularity conditions, the rational residents are always evolutionarily fragile against some entrants unless they are already getting the Stackelberg payoff in every situation. The second part of the result provides a condition for the rational residents to not be evolutionarily fragile against any singleton entrant. Whenever both conditions in Theorem 1 are satisfied, there is some distribution over situations so that the minimal correctly specified model is evolutionarily fragile against *some* entrant model, but not evolutionarily fragile against any *singleton* entrant model. In these environments, the ability to adapt preferences endogenously to the relevant situation (i.e., belief endogeneity) is a necessary condition for an invading entrant to displace

[16]If no such profile exists, let $v_G^b = -\infty$.

the rational resident. Hence, this result shows that entrants with misspecified models cannot in general be represented simply as entrants with fixed subjective preferences.

Theorem 1. *Suppose there are finitely many situations and there is a symmetric Nash equilibrium in* $\mathbb{A} \times \mathbb{A}$ *for every situation* G*.*

1. *If* $v_G^{NE} < \bar{v}_G$ *for some* G*, situation identifiability and Stackelberg identifiability hold, and there are finitely many strategies, then there exists a model* $\hat{\Theta}$ *such that the minimal correctly specified model is evolutionarily fragile against* $\hat{\Theta}$ *under any full-support distribution* $q \in \Delta(\mathcal{G})$*.*

2. *If there is no point* $(u_G)_{G\in\mathcal{G}}$ *in the convex hull of* $\{(v_G^b)_{G\in\mathcal{G}} \mid b : \mathbb{A} \rightrightarrows \mathbb{A}\}$ *with the property that* $u_G \geq v_G^{NE}$ *for every* $G \in \mathcal{G}$*, then there exists a full-support distribution* $q \in \Delta(\mathcal{G})$ *so that the minimal correctly specified model is not evolutionarily fragile against any singleton model.*

One environment where $v_G^{\text{NE}} = \bar{v}_G$ for every situation G is when agents face decision problems — that is, a player's payoff in every situation G is independent of the action of the matched opponent. When all situations are decision problems, the condition in the first part of Theorem 1 is violated: in fact, the correctly specified model is not evolutionarily fragile against any other model, regardless of whether such invaders infer from data.

For the second part of the theorem, if there is no subjective best-response correspondence b such that $v_G^b \geq v_G^{\text{NE}}$ for every $G \in \mathcal{G}$, then for every singleton entrant model there exists some distribution over situations so that the correctly specified model is not evolutionarily fragile against it. But the stronger condition we impose ensures that there exists one distribution over situations for which the correctly specified model is not evolutionarily fragile against *any* singleton entrant model.

Next, we use a numerical example to illustrate how we might verify the two sets of conditions from Theorem 1.

Example 1. Suppose $\mathbb{A} = \{a_1, a_2, a_3\}$, the consequences are $\mathbb{Y} = \{g, b\}$ with $u(g) = 1$ and $u(b) = 0$. Suppose there are two situations, G_A and G_B, and the probability a given player obtains g given a strategy profile and situation is determined by the table below.

G_A	a_1	a_2	a_3
a_1	0.1, 0.1	0.1, 0.1	0.1, 0.11
a_2	0.1, 0.1	0.3, 0.3	0.1, 0.1
a_3	0.11, 0.1	0.1, 0.1	0.2, 0.2

G_B	a_1	a_2	a_3
a_1	0.11, 0.11	0.5, 0.5	0.12, 0.4
a_2	0.5, 0.5	0.12, 0.12	0.14, 0.55
a_3	0.4, 0.12	0.55, 0.14	0.4, 0.4

It is easy to verify that the payoff-maximizing symmetric Nash equilibria in $\mathbb{A} \times \mathbb{A}$ are (a_2, a_2) for G_A and (a_3, a_3) for G_B, so $v_{G_A}^{\text{NE}} = 0.3$ and $v_{G_B}^{\text{NE}} = 0.4$. The unique Stackelberg strategy in G_A is a_2 and the unique Stackelberg strategy in G_B is a_1, and we have $\bar{v}_{G_A} = 0.3$ and $\bar{v}_{G_B} = 0.5$. There is a unique rational best response to every strategy in every situation.

Since $v_{G_B}^{\text{NE}} < \bar{v}_{G_B}$, the first set of conditions of Theorem 1 will be satisfied if we have situation identifiability and Stackelberg identifiability. By inspection, the probability of the g outcome under every strategy profile differs across the two situations, so situation identifiability holds. For Stackelberg identifiability, note that the unique rational best response to a_2 is a_2 in G_A and a_3 in G_B, and we have $F^\bullet(a_2, a_2, G_A) \neq F^\bullet(a_2, a_3, G_B)$. The unique rational best response to a_1 is a_2 in G_B and a_3 in G_A, and we have $F^\bullet(a_1, a_2, G_B) \neq F^\bullet(a_1, a_3, G_A)$. So, Stackelberg identifiability also holds.

To check the second set of conditions of Theorem 1, we consider three cases for the subjective best-response correspondence b.

- If $a_1 \in b(a_3)$, then since a_3 is a rational best response to a_1 in G_A and the highest possible payoff in G_B is 0.55, we get $v^b \leq (0.1, 0.55)$.
- If $a_2 \in b(a_3)$, then since a_3 is a rational best response to a_2 in G_B and the highest possible payoff in G_A is 0.3, we get $v^b \leq (0.3, 0.14)$.
- If $a_3 \in b(a_3)$, then since a_3 is a rational best response to a_3 in both G_A and G_B, we get $v^b \leq (0.2, 0.4)$.

These three cases are exhaustive since $b(a_3)$ cannot be empty. The half space in $\mathbb{R}^2$ below the line that runs through $(0.1, 0.55)$ and $(0.2, 0.4)$ contains all three points $(0.1, 0.55)$, $(0.3, 0.14)$, and $(0.2, 0.4)$. So, the convex hull of $\{(v_G^b)_{G \in \mathcal{G}} \mid b : \mathbb{A} \rightrightarrows \mathbb{A}\}$ is contained in this half space. But we have $v^{\text{NE}} = (0.3, 0.4)$, which is outside of the half space. Thus, the second set of conditions of Theorem 1 are satisfied.

We conclude, by Theorem 1, that there exists some full-support distribution over the two situations G_A and G_B such that the correctly specified model is evolutionarily fragile against some entrant model, but it is not evolutionarily fragile against any singleton model. In fact, we can take this distribution to be the one where the two situations are equally likely, and we can construct the invading entrant model as one that features illusion of control. This model has $\mathcal{F} = \{F_A, F_B\}$, where both F_A and F_B stipulate that the consequence only depends on the agent's own strategy and not on the opponent's strategy. Under F_A, a_1, a_2, and a_3 lead to consequence g with probabilities 0.1, 0.3, and 0.2 respectively (which are the probabilities of g if opponent plays a rational best response to these strategies in situation G_A). Under F_B, playing a_1, a_2, and a_3 lead to consequence g with probabilities 0.5, 0.14, and 0.4 respectively (which are the probabilities of g if opponent plays a rational best response to these strategies in situation G_B).

4.2 Stability Reversals

We now highlight another consequence of the endogeneity of misspecified beliefs: the potential for a greater indeterminacy in the emergence of stable biases. For expositional simplicity, we assume that $|\mathcal{G}| = 1$ throughout this section. We will refer to a model's *conditional fitness against group g*, i.e., the expected payoff of the model's adherents in matches against group g.

Definition 6. Two models Θ_A, Θ_B exhibit *stability reversal* if (i) in every EZ with $(p_A, p_B) = (1, 0)$, Θ_A has strictly higher conditional fitness than Θ_B against group A opponents and against group B opponents, but also (ii) in every EZ with $(p_A, p_B) = (0, 1)$, Θ_B has strictly higher fitness than Θ_A.

When $p_B = 0$, how Θ_A performs against Θ_B does not actually affect group A's fitness. Condition (i) encodes the strong requirement that Θ_A outperforms Θ_B even on the zero-probability event of being matched against a Θ_B opponent. A stability reversal occurs if this stronger requirement holds (when Θ_A dominates in society), and yet Θ_B still strictly outperforms Θ_A if Θ_B starts from a position of prominence.

We begin with two general results on when stability reversals *cannot* emerge. First, it cannot emerge without belief endogeneity:

Proposition 2. *Suppose* $|\mathcal{G}| = 1$. *Two singleton models (i.e., two subjective preferences in the stage game) cannot exhibit stability reversal.*

The reason is that for two singleton models, the conditional fitness of group g against group g' does not depend on the relative sizes of the groups. The subjective preference associated with a singleton model never changes with the social composition, so a strategy profile between groups g and g' that can be sustained in an EZ with $(p_A, p_B) = (1, 0)$ can also be sustained in an EZ with $(p_A, p_B) = (0, 1)$.

Stability reversals also cannot emerge in decision problems. We show this by introducing a class of models where agents always believe that strategic interactions do not matter:

Definition 7. A model Θ is *strategically independent* if for all $\mu \in \Delta(\Theta)$, $\arg\max_{a_i \in \mathbb{A}} U_i(a_i, a_{-i}; \mu)$ is the same for every $a_{-i} \in \mathbb{A}$.

The adherents of a strategically independent model believe that while an opponent's action may affect their utility, it does not affect their best response.

Proposition 3. *Suppose $|\mathcal{G}| = 1$, suppose Θ_A, Θ_B exhibit stability reversal and Θ_A is the correctly specified singleton model. Then, the beliefs that the adherents of Θ_B hold in all EZs with $p = (1, 0)$ and the beliefs they hold in all EZs with $p = (0, 1)$ form disjoint sets. Also, Θ_B is not strategically independent.*

The first claim of Proposition 3 underscores that stability reversal requires inference—it cannot happen if group B agents merely have a different subjective preference. The second claim shows that stability reversal can only happen if the misspecified agents respond differently to different rival play, immediately implying they cannot emerge in decision problems. The idea is that when the group B agents are prominent in the society, their misperception that the stage game is a decision problem implies that they will always choose the same strategy (say, $\hat{a}_i$) against both group A and group B opponents. But this means their fitness cannot be strictly higher than that of the rational group A agents, who play a rational best response against $\hat{a}_i$ when they match up against group B opponents.

We now show by example that stability reversal can emerge with models that allow for inference. Consider a two-player investment game where player i chooses an investment level $a_i \in \{1, 2\}$. A random productivity level P is realized according to $b^\bullet(a_i + a_{-i}) + \epsilon$ where ϵ is a zero-mean noise term, $b^\bullet > 0$. Player i's payoffs are $a_i \cdot P - (a_i - 1) \cdot c$. Consequences are $y = (a_i, a_{-i}, P)$. We record the payoff matrix of this investment game:

	1	2
1	$2b^\bullet, 2b^\bullet$	$3b^\bullet, 6b^\bullet - c$
2	$6b^\bullet - c, 3b^\bullet$	$8b^\bullet - c, 8b^\bullet - c$

Condition 1. $5b^\bullet < c < 6b^\bullet$.

In words, we assume that $a_i = 1$ is a strictly dominant strategy in the stage game, but the investment profile (2,2) Pareto dominates the investment profile (1,1) (so that the corresponding game is a prisoner's dilemma). Consider two models in the society. Take Θ_A to be a correctly specified singleton (thus knowing the true mapping from actions to payoffs), while Θ_B wrongly stipulates $P = b(a_i + a_{-i}) - m + \epsilon$, where $m > 0$ is fixed, while $b \in \mathbb{R}$ is a parameter that the adherents infer. We impose a condition on Θ_B, which holds whenever $m > 0$ is large enough:

Condition 2. $c < 4b^\bullet + \frac{1}{3}m$ and $c < 5b^\bullet + \frac{1}{4}m$.

We show that in this example models Θ_A and Θ_B exhibit stability reversal.

Example 2. In the investment game, under Condition 1 and Condition 2, Θ_A and Θ_B exhibit stability reversal.

The idea is that the Θ_B adherents hold endogenous beliefs about the value of b. They overestimate the complementarity of investments, and this overestimation is more severe when they face data generated from lower investment profiles. As a result, the match between Θ_A and Θ_B plays out differently depending on which model is resident: it results in the investment profile $(1, 2)$ when Θ_A is resident, but results in $(1, 1)$ when Θ_B is resident. (We relegate the formal argument to Appendix A.5.) Due to Propositions 2 and 3, we conclude that this example is possible due to the non-trivial strategic interactions and Θ_B's inference about b.

Stability reversals provide a clear demonstration of the endogeneity of beliefs and hence the fluidity of conditional fitness in models that permit inference. An entrant model may appear weak when present in small proportions, doing worse than the resident model conditional on every type of opponent. Yet, if the population share of the entrant model reaches a critical mass, its adherents infer a more evolutionarily advantageous model parameter based on their within-group interactions, change their best-response correspondence, and hence outperform the adherents of the resident model.

5 Evolutionary Stability of Analogy Classes

We apply the stability notions introduced in this paper to study coarse thinking in games. Jehiel (2005) introduced analogy-based expectation equilibrium (ABEE) in extensive-form games, where agents group opponents' nodes into *analogy classes* and only keep track of aggregate statistics of opponents' average behavior within each analogy class. An ABEE is a strategy profile where agents best respond to the belief that at all nodes in every analogy class, opponents behave according to the average behavior in the analogy class. The ensuing literature typically treats analogy classes as exogenously given, interpreted as arising from coarse feedback or agents' cognitive limitations.[17] We showcase the practical value of our approach by using the framework from Section 2 to endogenize analogy classes based on their objective expected payoffs in equilibrium.[18]

5.1 Defining Stable Population Shares

In this section, we will focus on an environment where agents know the stage game but may have misspecified beliefs about others' strategies. We will no longer work in the special case of strategic certainty, and in fact we turn off the monitoring signals by assuming that m_i is fully uninformative about the matched opponent's strategy a_{-i}. We will also be interested in stable population shares in a society that contains positive fractions of both rational and misspecified players. This is because the environment we analyze features a rational model and a misspecified model with neither model being evolutionarily stable against the other (as we will see later in Proposition 4).

We briefly introduce the following solution concept.

Definition 8. Call population share $(p, 1-p)$ with $p \in (0,1)$ a *stable population share* if there is an EZ with $(p, 1-p)$ where both models have the same fitness, and there exists $\bar{\epsilon}$ such that:

1. For any $0 < \epsilon < \bar{\epsilon}$, there is an EZ with population share $(p+\epsilon, 1-p-\epsilon)$ where Θ_A has strictly lower fitness than Θ_B

2. For any $0 < \epsilon < \bar{\epsilon}$, there is an EZ with population share $(p-\epsilon, 1-p+\epsilon)$ where Θ_A has strictly higher fitness than Θ_B.

[17] Section 6.2 of Jehiel (2005) mentions that if players could choose their own analogy classes, then the finest analogy classes need not arise, but also says "it is beyond the scope of this paper to analyze the implications of this approach." In a different class of games, Jehiel (1995) similarly observes that another form of bounded rationality (having a limited forecast horizon about opponent's play) can improve welfare.

[18] Other approaches to endogenizing analogy classes are pursued in Jehiel and Mohlin (2023); Jehiel and Weber (2025).

Whereas Definition 4's stability notion involves comparing the performance of the two models when one of them is present in an arbitrarily small fraction, stability with an interior population share as in Definition 8 refers to both models co-existing with equal fitness in a way that is robust to local perturbations of population sizes.

Another difference between Definition 4 and Definition 8 is that the former requires a uniform welfare comparison across all EZs and the latter just requires a welfare comparison in one EZ. Indeed, we will select a particular focal EZ, because the environment has trivial EZs where misspecified agents always "opt out" of playing the game, receive no information about how others play, and hold beliefs about others' strategies that make opting out subjectively optimal. In such EZs, misspecified agents have the same fitness as the rational agents, but not for any interesting reasons that relate to their misspecified models.

5.2 Centipede Games and Analogy-Based Reasoning

We now analyze analogy-based reasoning in the centipede game in Figure 1 (there is only one situation, given by the payoffs in this game). P1 and P2 take turns choosing Across (A) or Drop (D). The non-terminal nodes are labeled n^k, $1 \leq k \leq K$ where K is an even number. P1 acts at odd nodes and P2 acts at even nodes, where choosing Drop at n^k leads to the terminal node z^k. If Across is always chosen, then the terminal node z^{end} is reached. Every time a player i chooses Across, the sum of payoffs grows by $g > 0$. However, if the opponent chooses Drop next, i's payoff is $\ell > 0$ smaller than i's payoff had they chosen Drop, with $\ell > g$. Thus, if z^{end} is reached, both get $Kg/2$; if z^k is reached when k is odd, both players obtain $\frac{g(k-1)}{2}$; and if if z^k is reached when k is even, P1 obtains $\frac{k-2}{2}g - \ell$, and P2 obtains $\frac{k}{2}g + \ell$.

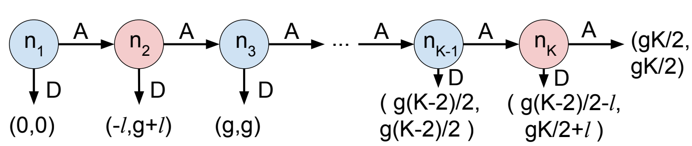

Figure 1: The centipede game. P1 (blue) and P2 (red) alternate in choosing Across (A) or Drop (D). Payoff profiles are shown at the terminal nodes.

While this is an asymmetric stage game, we study a symmetrized version where two matched agents are randomly assigned into the roles of P1 and P2. Let $\mathbb{A} = \{(d^k)_{k=1}^K \in [0,1]^K\}$, so each strategy is characterized by the probabilities of playing Drop at various nodes in the game tree. When assigned into the role of P1, the strategy (d^k) plays Drop with probabilities $d^1, d^3, ..., d^{K-1}$ at nodes $n^1, n^3, ...n^{K-1}$. When assigned into the role of P2, it plays Drop with probabilities $d^2, d^4, ..., d^K$ at nodes $n^2, n^4, ...n^K$. The set of consequences is $\mathbb{Y} = \{1, 2\} \times (\{z_k : 1 \leq k \leq K\} \cup \{z_{end}\})$, where the first dimension of the consequence returns the player role that the agent was assigned into, and the second dimension returns the terminal node reached. Let $F^\bullet : \mathbb{A}^2 \to \Delta(\mathbb{Y})$ be the objective distribution over consequences.

All agents know the game tree (i.e., $F^\bullet$), but some might adhere to a model which mistakenly assumes that their opponent plays Drop with the same probabilities at all of their nodes. Formally, define the restricted space of strategies $\mathbb{A}^{An} := \{(d^k) \in [0, 1]^K : d^k = d^{k'} \text{ if } k \equiv k' (\text{mod } 2)\} \subseteq \mathbb{A}$. The correctly specified model is $\Theta^\bullet := \mathbb{A} \times \mathbb{A} \times \{F^\bullet\}$. The misspecified model of interest is $\Theta^{An} := \mathbb{A}^{An} \times \mathbb{A}^{An} \times \{F^\bullet\}$, reflecting a dogmatic belief that opponents play the same mixed action at all nodes in the analogy class. We emphasize these restriction on strategies only exists in the subjective beliefs of the model Θ^{An} adherents. All agents, regardless of their model, actually have the strategy space $\mathbb{A}$.

5.3 Results

The next proposition provides a justification for why we might expect agents with coarse analogy classes given by $\mathbb{A}^{An}$ to persist in the society.

Proposition 4. *Suppose $K \geq 4$ and $g > \frac{2}{K-2}\ell$. The correctly specified model $\Theta^\bullet$ is evolutionarily stable against itself, but it is not evolutionarily stable against the misspecified model Θ^{An}. Also, Θ^{An} is not evolutionarily stable against $\Theta^\bullet$.*

Thus, the correctly specified model is not evolutionarily stable against a coarse reasoner. Here, the conditional fitness of Θ^{An} against both $\Theta^\bullet$ and Θ^{An} can strictly improve on the correctly specified residents' equilibrium fitness. This is because the matches between two adherents of $\Theta^\bullet$ must result in Dropping at the first move in equilibrium, while matches where at least one player is an adherent of Θ^{An} either lead to the same outcome or lead to a Pareto dominating payoff profile as the misspecified agent misperceives the opponent's continuation probability and thus chooses Across at almost all of the decision nodes.

However, Θ^{An} is not evolutionarily stable against $\Theta^\bullet$ either. The correctly specified agents can exploit the analogy reasoners' mistake and receive higher payoffs in matches against them than the misspecified agents receive in matches against each other. Hence, no homogeneous population can be stable, as the resident model would have lower fitness than the entrant model in equilibrium. Thus we determine stable shares as defined in Section 5.1, focusing on the EZs where Across is played as often as possible.

Suppose $K \geq 4$ and $g > \frac{2}{K-2}\ell$. Consider the *maximal continuation EZ*: (1) misspecified agents always play Across except at node K where they choose Drop, and (2) correctly specified agents (i) when matched with misspecified agents, play Drop at nodes $K-1$ and K and Across otherwise, and (ii) when matched with correctly specified agents, always play Drop. We verify this indeed forms an EZ.

Proposition 5. Suppose $K \geq 4$ and $g > \frac{2}{K-2}\ell$. The only stable population share (p_A^*, p_B^*) supported by the maximal continuation EZ described above is $p_B^* = 1 - \frac{\ell}{g(K-2)}$. We have p_B^* is strictly increasing in g and K, and strictly decreasing in ℓ.

Intuitively, p_B^* reflects the fraction of society expected to be analogy reasoners if long-run population changes are determined by fitness. Under the maintained assumption $g > \frac{2}{K-2}\ell$, the stable population share of misspecified agents is strictly more than 50%, and the share grows with more periods and a larger increase in payoffs from continuation. The main intuition is that the misspecified model has a higher conditional fitness than the rational model against rational opponents. The former leads to many periods of continuation and a high payoff for the biased agent when the rational agent eventually drops, but the latter leads to 0 payoff from immediate dropping. On the other hand, the misspecified model has a lower conditional fitness than the rational model against misspecified opponents. For the two groups to have the same expected fitness, there must be fewer rational opponents (i.e., a smaller stable population share p_A^*) when g and K are higher.

Note that, when payoffs are specified as above, two successive periods of continuation lead to a strict Pareto improvement in payoffs. Consider instead the so-called "dollar game" Reny (1993) in Figure 2, a variant with a more "competitive" payoff structure, where an agent always gets zero when the opponent plays Drop, at all parts of the game tree. Assume total payoff increases by 1 in each round. If the first player stops immediately, payoffs are (1, 0). If the second player continues at the final node n^K, payoffs are $(K+2, 0)$.

Proposition 6. For every population size $(p, 1-p)$ with $p \in [0,1]$, the maximal continuation EZ is an EZ where the fitness of $\Theta^{\bullet}$ is strictly higher than that of Θ^{An}.

While maximal continuation remains an EZ, the rational model strictly outperforms the misspecified model for all population shares. Provided the maximal continuation EZ remains

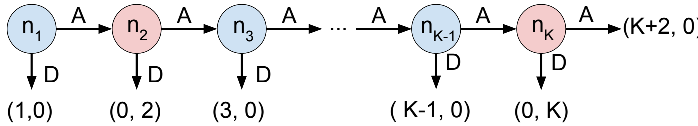

Figure 2: The dollar game. Players 1 (blue) and 2 (red) alternate in choosing Across (A) or Drop (D). Payoff profiles are shown at the terminal nodes.

focal, we would expect no analogy reasoners in the long run with this stage game. Intuitively, the payoffs imply one player can only do better *at the expense* of the opponent. This implies the less cooperative strategy will be selected.

In a recent survey, Jehiel (2020) points out that the misspecified Bayesian learning approach to analogy classes should aim for "a better understanding of how the subjective theories considered by the players may be shaped by the objective characteristics of the environment."[19] Taken together, our analysis in this section provides predictions regarding when coarse reasoning should be more prevalent, specifically when the payoff structure is "less competitive." When this is indeed the case, the bias becomes more prevalent with a longer horizon and with faster payoff growth.

6 Concluding Discussion

We have introduced an evolutionary approach to predict the persistence of misspecified models under Bayesian learning. We have emphasized the implications and significance of belief endogeneity for evolutionary stability and the viability of models. Our contributions are twofold. First, we show that belief endogeneity may confer strategic benefits in cases where dogmatic beliefs do not. This is because endogenous beliefs enable flexible commitments that are tailored to the realized situation. Second, we show that the endogeneity of misspecified beliefs makes it difficult to extrapolate the performance of a fixed bias across environments. More broadly, we hope to have shown that incorporating inference enables the evolutionary approach to speak to new applications and patterns.

We acknowledge that our framework does not account for which errors appear in the first place. It is plausible that some first-stage filter prevents certain obvious misspecifications from ever reaching the stage that we study in the evolutionary framework. For this reason,

[19] Jehiel (2020) interprets ABEEs as players adopting the "simplest" explanations of observed aggregate statistics of play with coarse feedback. An objectively coarse feedback structure can lead agents to adopt the subjective belief that others behave in the same way in all contingencies in the same coarse analogy class.

the applications we focused on reflected misspecifications that seem psychologically plausible.

We have used an otherwise off-the-shelf framework to describe the selection of specifications. The goal of this paper is not to identify suitable definitions of fitness to justify particular errors (which is the focus for many of the papers that Robson and Samuelson (2011) survey). Rather, our goal has been to determine what evolutionary forces would suggest about the persistence of misspecified models. We have therefore focused more on the implications of belief endogeneity in an otherwise standard evolutionary setup.

References

Alger, I. and J. Weibull (2019): "Evolutionary models of preference formation," *Annual Review of Economics*, 11, 329–354.

Aliprantis, C. and K. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer Science & Business Media.

Ba, C. (2025): "Robust misspecified models," *Working Paper*.

Ba, C. and A. Gindin (2023): "A multi-agent model of misspecified learning with overconfidence," *Games and Economic Behavior*, 142, 315–338.

Berk, R. H. (1966): "Limiting behavior of posterior distributions when the model is incorrect," *The Annals of Mathematical Statistics*, 37, 51–58.

Berman, R. and Y. Heller (2024): "Naive analytics: The strategic advantage of algorithmic heuristics," *Working Paper*.

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

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

Cho, I.-K. and K. Kasa (2015): "Learning and model validation," *Review of Economic Studies*, 82, 45–82.

——— (2017): "Gresham's law of model averaging," *American Economic Review*, 107, 3589–3616.

Cho, I.-K. and J. Libgober (2025): "Learning underspecified models," *Journal of Economic Theory*, 226, 106015.

Dasaratha, K. and K. He (2020): "Network structure and naive sequential learning," *Theoretical Economics*, 15, 415–444.

Dekel, E., J. Ely, and O. Yilankaya (2007): "Evolution of preferences," *Review of Economic Studies*, 74, 685–704.

Esponda, I. and D. Pouzo (2016): "Berk–Nash equilibrium: A framework for modeling agents with misspecified models," *Econometrica*, 84, 1093–1130.

Esponda, I., D. Pouzo, and Y. Yamamoto (2021): "Asymptotic behavior of Bayesian learners with misspecified models," *Journal of Economic Theory*, 195, 105260.

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

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

——— (2024): "Welfare comparisons for biased learning," *American Economic Review*, 114, 1612–1649.

Fudenberg, D. and G. Lanzani (2023): "Which misspecifications persist?" *Theoretical Economics*, 18, 1271–1315.

Fudenberg, D., G. Lanzani, and P. Strack (2021): "Limit points of endogenous misspecified learning," *Econometrica*, 89, 1065–1098.

——— (2024): "Selective-Memory Equilibrium," *Journal of Political Economy*, 132, 3978–4020.

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

Gagnon-Bartsch, T., M. Rabin, and J. Schwartzstein (2021): "Channeled attention and stable errors," *Working Paper*.

Güth, W. and S. Napel (2006): "Inequality aversion in a variety of games - An indirect evolutionary analysis," *Economic Journal*, 116, 1037–1056.

He, K. (2022): "Mislearning from censored data: The gambler's fallacy in optimal-stopping problems," *Theoretical Economics*, 17, 1269–1312.

He, K. and J. Libgober (2025): "Higher-Order Beliefs and (Mis)learning from Prices," *Working Paper*.

Heidhues, P., B. Koszegi, and P. Strack (2018): "Unrealistic expectations and misguided learning," *Econometrica*, 86, 1159–1214.

Heller, Y. (2015): "Three steps ahead," *Theoretical Economics*, 10, 203–241.

Heller, Y. and E. Winter (2016): "Rule rationality," *International Economic Review*, 57, 997–1026.

——— (2020): "Biased-belief equilibrium," *American Economic Journal: Microeconomics*, 12, 1–40.

Jehiel, P. (1995): "Limited horizon forecast in repeated alternate games," *Journal of Economic Theory*, 67, 497–519.

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

——— (2018): "Investment strategy and selection bias: An equilibrium perspective on overoptimism," *American Economic Review*, 108, 1582–97.

——— (2020): "Analogy-based expectation equilibrium and related concepts: Theory, applications, and beyond," *Working Paper*.

Jehiel, P. and E. Mohlin (2023): "Categorization in games: A bias-variance perspective," Working Paper.

Jehiel, P. and G. Weber (2025): "Endogenous clustering and analogy-based expectation equilibrium," *Review of Economic Studies*.

Lanzani, G. (2025): "Dynamic Concern for Misspecification," *Econometrica*.

Levy, G., R. Razin, and A. Young (2022): "Misspecified politics and the recurrence of populism," *American Economic Review*, 112, 928–962.

Massari, F. (2020): "Under-reaction: Irrational behavior or robust response to model misspecification?" *Working Paper*.

Massari, F. and J. Newton (2020): "Learning and equilibrium in misspecified models," Working Paper.

——— (2024): "Rational beliefs when the truth is not an option," Working Paper.

Molavi, P. (2019): "Macroeconomics with learning and misspecification: A general theory and applications," *Working Paper*.

Montiel Olea, J. L., P. Ortoleva, M. M. Pai, and A. Prat (2022): "Competing models," *Quarterly Journal of Economics*, 137, 2419–2457.

Murooka, T. and Y. Yamamoto (2023): "Higher-order misspecification and equilibrium stability," Working Paper.

Nyarko, Y. (1991): "Learning in mis-specified models and the possibility of cycles," *Journal of Economic Theory*, 55, 416–427.

Reny, P. J. (1993): "Common belief and the theory of games with perfect information," *Journal of Economic Theory*, 59, 257–274.

Robson, A. J. (1996): "A biological basis for expected and non-expected utility," *Journal of Economic Theory*, 68, 397–424.

Robson, A. J. and L. Samuelson (2011): "The evolutionary foundations of preferences," in *Handbook of Social Economics*, Elsevier, vol. 1, 221–310.

Sandholm, W. (2001): "Preference evolution, two-speed dynamics, and rapid social change," *Review of Economic Dynamics*, 4, 637–679.

Sandroni, A. (2000): "Do markets favor agents able to make accurate predictions?" *Econometrica*, 68, 1303–1341.

Schwartzstein, J. and A. Sunderam (2021): "Using models to persuade," *American Economic Review*, 111, 276–323.

Appendix

A Omitted Proofs from the Main Text

A.1 Proof of Proposition 1

Proof. We first note that for an agent i who believes in the parameters $r > 0$ and β, the subjective expected utility from the strategy profile (a_i, a_{-i}) is $a_i \cdot (\beta - r(a_i + a_{-i}) - c)$ (since ε is mean zero). The second derivative in a_i is $-2r < 0$, so the maximizer is characterized by the first-order condition. Taking FOC, we get the subjective best response $a_i = \frac{\beta - c - ra_{-i}}{2r}$.

We also note that the correctly specified residents must have an equilibrium belief that assigns probability 1 to $\beta = \beta^\bullet$ in every EZ. This is because in an EZ where the residents play the strategy a_{AA}, $\beta = \beta^\bullet$ has zero KL divergence whereas any other value of β has strictly positive KL divergence.

Now we prove the three parts of the proposition.

Part 1: The only Nash equilibrium of the game is for both players to choose $\frac{\beta^\bullet - c}{3r^\bullet}$ (since the rational best response function is linear). So, the fitness of the residents is given by $\frac{\beta^\bullet - c}{3r^\bullet} \cdot (\beta^\bullet - r^\bullet \cdot 2 \cdot \frac{\beta^\bullet - c}{3r^\bullet} - c)$, which simplifies to $\frac{(\beta^\bullet - c)^2}{9r^\bullet}$.

Part 2: In an equilibrium zeitgeist where the entrants use the strategy a_{BA}, the correctly specified residents (who have correct beliefs about all the parameters in equilibrium) best respond with $a_{AB} = \frac{\beta^\bullet - c - r^\bullet a_{BA}}{2r^\bullet}$. So, the fitness of the entrant is given by $a_{BA} \cdot (\beta^\bullet - r^\bullet(a_{BA} + \frac{\beta^\bullet - c - r^\bullet a_{BA}}{2r^\bullet}) - c)$. Simplifying we get $\frac{1}{2}[(a_{BA}) \cdot (\beta^\bullet - c) - (a_{BA})^2 \cdot (r^\bullet)]$. This expression is quadratic in a_{BA} and it must be maximized at the Stackelberg strategy of the game, hence the fitness of the entrants is strictly decreasing in the distance between a_{BA} and the Stackelberg strategy. The Stackelberg strategy is found by taking the first-order condition of the expression $\frac{1}{2}[(a_{BA}) \cdot (\beta^\bullet - c) - (a_{BA})^2 \cdot (r^\bullet)]$, which gives $a_{\text{stack}} = \frac{\beta^\bullet - c}{2r^\bullet}$.

Part 3: In an EZ where entrants with slope misperception $\hat{r}$ play a_i against residents who play a_{-i}, the distribution of consequences has zero KL divergence only for the parameter $\hat{\beta}$ that solves $\beta^\bullet - r^\bullet A = \hat{\beta} - \hat{r}A$, so entrants must infer $\hat{\beta} = \beta^\bullet + (a_i + a_{-i})(\hat{r} - r^\bullet)$. But if the entrants play a_i, in equilibrium the residents must play the rational best response against it, which is $a_{-i} = \frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet}$. So the entrants infer $\hat{\beta} = \beta^\bullet + (a_i + \frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet})(\hat{r} - r^\bullet)$ in an EZ where they choose $a_{BA} = a_i$. Under the beliefs $(\hat{\beta}, \hat{r})$, the entrants' subjective best response to $\frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet}$ is $\frac{\hat{\beta} - \hat{r}\frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet} - c}{2\hat{r}}$, and so we must have $a_i = \frac{\hat{\beta} - \hat{r}\frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet} - c}{2\hat{r}}$. Making the substitution that $\hat{\beta} = \beta^\bullet + (a_i + \frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet})(\hat{r} - r^\bullet)$ on the right-hand side, we get

$a_i = \frac{\beta^\bullet + (a_i + \frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet}) \cdot (\hat{r} - r^\bullet) - \hat{r}\frac{\beta^\bullet - c - r^\bullet a_i}{2r^\bullet} - c}{2\hat{r}}$. Simplifying this linear equation in a_i gives us the unique solution $a_i = \frac{\beta^\bullet - c}{2\hat{r} + r^\bullet}$, which is the unique EZ value of a_{BA} when entrants have the slope misperception $\hat{r}$. □

A.2 Proof of Theorem 1

Part 1: Suppose the hypotheses hold and let us construct the misspecified model $\hat{\Theta} = \{F_G : G \in \mathcal{G}\}$. Towards defining the parameter F_G for each situation G, first consider $\tilde{F}_G$ where $\tilde{F}_G(a_i, a_{-i}) := F^\bullet(a_i, \underline{\text{BR}}(a_i, G), G)$ for every $a_{-i} \in \mathbb{A}$. Now for each $(a_i, a_{-i}, G) \in \mathbb{A} \times \mathbb{A} \times \mathcal{G}$, define the full-support distribution $F_G(a_i, a_{-i}) \in \Delta(\mathbb{Y})$ as a sufficiently small perturbation of $\tilde{F}_G(a_i, a_{-i})$, such that for every $a_i, a_{-i} \in \mathbb{A}$ and every $G \in \mathcal{G}$, $\min_{\hat{G} \in \mathcal{G}} KL(F^\bullet(a_i, a_{-i}, G) \parallel F_{\hat{G}}(a_i, a_{-i}))$ has a unique solution. This can be done because there are finitely many strategies and situations.

Consider any EZ $\mathfrak{Z}$ where the resident is the minimal correctly specified model and the entrant is $\hat{\Theta}$. By situation identifiability and because we are in an environment of strategic certainty, in $\mathfrak{Z}$ the correctly specified residents must believe in the true $F^\bullet(\cdot, \cdot, G)$ in every situation G. When the fraction of entrants $\epsilon > 0$ is sufficiently small, the entrants cannot hold a mixed belief in any situation G, by the construction of the parameters in $\hat{\Theta}$ to rule out ties in KL divergence if entrants only use the consequences in their matches against the residents to make inferences. We show further that entrants must believe in F_G in situation G for ϵ small enough. This is because if they instead believed in $F_{G'}$ for some $G' \neq G$, then they must play $\bar{a}_{G'}$ as the Stackelberg strategy is assumed to be unique. Let a_{-i} be the rational best response to $\bar{a}_{G'}$ in situation G and a'_{-i} be the rational best response to $\bar{a}_{G'}$ in situation G', both unique by assumption. In their matches against the residents, the entrants' expected distribution of consequences $F_{G'}(\bar{a}_{G'}, a_{-i})$ is a perturbed version of $F^\bullet(\bar{a}_{G'}, a'_{-i}, G')$, while the true distribution of consequences $F^\bullet(\bar{a}_{G'}, a_{-i}, G)$ is a perturbed version of $F_G(\bar{a}_{G'}, a_{-i})$. We have $F^\bullet(\bar{a}_{G'}, a'_{-i}, G') \neq F^\bullet(\bar{a}_{G'}, a_{-i}, G)$ by Stackelberg identifiability, so $KL(F^\bullet(\bar{a}_{G'}, a_{-i}, G) \parallel F_G(\bar{a}_{G'}, a_{-i})) < KL(F^\bullet(\bar{a}_{G'}, a_{-i}, G) \parallel F_{G'}(\bar{a}_{G'}, a_{-i}))$ when the perturbations are sufficiently small. When $\epsilon > 0$ is small enough, this contradicts the entrants believing in $F_{G'}$ in situation G as the parameter F_G generates smaller weighted KL divergence across all of the entrant's data (since data from matches against entrants get weighted by ϵ and the full-support nature of all processes in the model implies that KL divergence of the data from such matches is bounded). So the entrants get the Stackelberg payoff in each situation when playing the resident, which means they have higher fitness than the residents in every EZ for ϵ small enough, since $\bar{v}_G > v_G^{\mathrm{NE}}$ for at least one situation and q has full support. Finally, there exists at least one EZ: for $\epsilon > 0$ small enough, it is an EZ for the residents to believe in $F^\bullet(\cdot, \cdot, G)$ in every situation G, to play the symmetric Nash profile that results in v_G^{NE} when matched with other residents (this profile exists by hypothesis of the theorem), and for the entrants to believe in F_G and play $(\bar{a}_G, \underline{\mathrm{BR}}(\bar{a}_G, G))$ in matches against residents in situation G.

Part 2: Let $\mathcal{V}$ be the convex hull of $\{(v_G^b)_{G\in\mathcal{G}} \mid b : \mathbb{A} \rightrightarrows \mathbb{A}\}$, and let $\mathcal{U} = \{(u_G)_{G\in\mathcal{G}} : u_G \leq v_G \text{ for all } G \text{ for some } v \in \mathcal{V}\}$. Note $\mathcal{U}$ is closed and convex (since $\mathcal{V}$ is convex). By hypothesis, v^{NE} is not in the interior or on the boundary of $\mathcal{U}$. So by the separating hyperplane theorem, there exists a real number c and a vector $q \in \mathbb{R}^{|\mathcal{G}|}$ with $q_G \neq 0$ for every G, so that $q \cdot v^{\mathrm{NE}} > c > q \cdot u$ for every $u \in \mathcal{U}$. Furthermore, $q_G \geq 0$ for every G. This is because if $q_{G'} < 0$ for some G', then since $\mathcal{U}$ contains vectors with arbitrarily negative values in the G' dimension, we cannot have $q \cdot v^{\mathrm{NE}} \geq q \cdot u$ for every $u \in \mathcal{U}$. We may then without loss view q as a distribution on $\mathcal{G}$. In fact, we can take q to be full support. To see this, note that since $|\mathcal{G}| < \infty$ and $\mathcal{U}$ is convex, we have

$$\lim_{\varepsilon\to 0} \max_{v\in\mathcal{U}} \left[(1-\varepsilon)q + \frac{\varepsilon}{|\mathcal{G}|}(1, 1, \ldots, 1)\right] \cdot v = \max_{v\in\mathcal{U}} q \cdot v,$$

by continuity of the support function of convex sets in $\mathbb{R}^n$ (given that the support function on $\mathcal{U}$ is bounded for all $q \geq 0$, since v_G^b is bounded above for every b and every G). Thus, setting $\tilde{q}(\varepsilon) = (1-\varepsilon)q + \frac{\varepsilon}{|\mathcal{G}|}(1, 1, \ldots, 1)$, we have $\tilde{q}(\varepsilon)$ is a full support distribution with $\tilde{q}(\varepsilon) \cdot v^{NE} > c > \tilde{q}(\varepsilon) \cdot u$ whenever ε is sufficiently small, since we have that these inequalities hold in the limit.

Now consider any singleton model $\mathcal{F} = \{F\}$, and let $b : \mathbb{A} \rightrightarrows \mathbb{A}$ be the subjective best-response correspondence that F induces. If $v_G^b \neq -\infty$ for every G, then, for each G we can find a strategy profile (a_i^G, a_{-i}^G) where $a_i^G \in b(a_{-i}^G)$, a_{-i}^G is a rational best response to a_i^G in situation G, and the strategy pair gives payoff v_G^b to the first player. For any population shares of the two models, there is an EZ where the resident correctly specified agents get v_G^{NE} in situation G when playing against each other, and the entrants with model Θ play (a_i, a_{-i}) in matches against the residents and get utility v_G^b in the same situation. Under the distribution of situations q, as the fraction of the entrants approaches 0, the residents' fitness approaches $q \cdot v^{\mathrm{NE}}$ while that of the entrants approaches $q \cdot v^b$, and the former is strictly larger by construction of q since $v^b \in \mathcal{U}$. This EZ shows the correctly specified model is not evolutionarily fragile against $\{F\}$. Otherwise, if we have that $v_G^b = -\infty$ for some G, then there are no EZs, so the correctly specified model is not evolutionarily fragile against $\{F\}$ by the emptiness of the set of EZs.

A.3 Proof of Proposition 2

Proof. Let two singleton models Θ_A, Θ_B be given. By contradiction, suppose they exhibit stability reversal. Let $\mathfrak{Z} = (\mu_A, \mu_B, p = (0,1), (a))$ be any EZ where Θ_B is resident. By the definition of EZ, $\mathfrak{Z}' = (\mu_A, \mu_B, p = (1,0), (a))$ is also an EZ where Θ_A is resident. Let $u_{g,g'}$ be model Θ_g's conditional fitness against group g' in the EZ $\mathfrak{Z}'$. Part (i) of the definition of stability reversal requires that $u_{AA} > u_{BA}$ and $u_{AB} > u_{BB}$. These conditional fitness levels remain the same in $\mathfrak{Z}$. This means the fitness of Θ_A is strictly higher than that of Θ_B in $\mathfrak{Z}$, a contradiction. □

A.4 Proof of Proposition 3

Proof. To show the first claim, suppose $\mathfrak{Z} = (\mu_A, \mu_B, p = (1,0), (a_{AA}, a_{AB}, a_{BA}, a_{BB}))$ is an EZ, and $\tilde{\mathfrak{Z}} = (\mu_A, \mu_B, p = (0,1), (\tilde{a}_{AA}, \tilde{a}_{AB}, \tilde{a}_{BA}, \tilde{a}_{BB}))$ is another EZ where the adherents of Θ_B hold the same belief μ_B (group A's belief cannot change as Θ_A is the correctly specified singleton model). By the optimality of behavior in $\mathfrak{Z}$, a_{BA} best responds to a_{AB} under the belief μ_B, and a_{AB} best responds to a_{BA} under the belief μ_A, therefore $\tilde{\mathfrak{Z}}' = (\mu_A, \mu_B, p = (0,1), (\tilde{a}_{AA}, a_{AB}, a_{BA}, \tilde{a}_{BB}))$ is another EZ. This holds because the distributions of observations for the adherents of Θ_B are identical in $\tilde{\mathfrak{Z}}$ and $\tilde{\mathfrak{Z}}'$, since they only face data generated from the profile $(\tilde{a}_{BB}, \tilde{a}_{BB})$. At the same time, since $\tilde{a}_{BB}$ best responds to itself under the belief μ_B, we have that $\mathfrak{Z}' = (\mu_A, \mu_B, p = (1,0), (a_{AA}, a_{AB}, a_{BA}, \tilde{a}_{BB}))$ is an EZ. Part (i) of the definition of stability reversal applied to $\mathfrak{Z}'$ requires that $U^\bullet(a_{AB}, a_{BA}) > U^\bullet(\tilde{a}_{BB}, \tilde{a}_{BB})$ (where $U^\bullet$ is the objective expected payoffs), but part (ii) of the same definition applied to $\tilde{\mathfrak{Z}}'$ requires $U^\bullet(\tilde{a}_{BB}, \tilde{a}_{BB}) \geq U^\bullet(a_{AB}, a_{BA})$, a contradiction.

To show the second claim, by way of contradiction suppose Θ_B is strategically independent and $\mathfrak{Z} = (\mu_A, \mu_B, p = (0,1), (a_{AA}, a_{AB}, a_{BA}, a_{BB}))$ is an EZ. By strategic independence, the adherents of Θ_B find it optimal to play a_{BB} against any opponent strategy under the belief μ_B. So, there exists another EZ of the form $\mathfrak{Z}' = (\mu_A, \mu_B, p = (0,1), (a_{AA}, a'_{AB}, a_{BB}, a_{BB}))$, where a'_{AB} is an objective best response to a_{BB}. The belief μ_B is sustained because in both $\mathfrak{Z}$

and $\mathfrak{Z}'$, the adherents of Θ_B have the same data: from the strategy profile (a_{BB}, a_{BB}). In $\mathfrak{Z}'$, Θ_A's fitness is $U^\bullet(a'_{AB}, a_{BB})$ and Θ_B's fitness is $U^\bullet(a_{BB}, a_{BB})$. We have $U^\bullet(a'_{AB}, a_{BB}) \geq U^\bullet(a_{BB}, a_{BB})$ since a'_{AB} is an objective best response to a_{BB}, contradicting the definition of stability reversal. $\square$

A.5 Details Behind Example 2

Let $b^*(a_i, a_{-i})$ solve

$$\min_{b\in\mathbb{R}} D_{KL}(F^\bullet(a_i, a_{-i}) \parallel \hat{F}(a_i, a_{-i}; b, m))),$$

where $F^\bullet(a_i, a_{-i})$ is the objective distribution over consequences under the investment profile (a_i, a_{-i}), and $\hat{F}(a_i, a_{-i}; b, m)$ is the distribution under the same investment profile if productivity is given by $P = b(x_i + x_{-i}) - m + \epsilon$. We find that $b^*(a_i, a_{-i}) = b^\bullet + \frac{m}{a_i + a_{-i}}$. It is clear that $D_{KL}(F^\bullet(a_i, a_{-i}) \parallel \hat{F}(a_i, a_{-i}; b^*(a_i, a_{-i}), m))) = 0$, while this KL divergence is strictly positive for any other choice of b.

Now we show that Example 2 exhibits stability reversal. In every EZ with $p = (1, 0)$, we must have $a_{AA} = a_{AB} = 1$. If $a_{BA} = 2$, then the adherents of Θ_B infer $b^*(1, 2) = b^\bullet + \frac{m}{3}$. With this inference, the biased agents expect $1 \cdot (2(b^\bullet + \frac{m}{3}) - m) = 2b^\bullet - \frac{m}{3}$ from playing 1 against rival investment 1, and expect $2 \cdot (3(b^\bullet + \frac{m}{3}) - m) - c = 6b^\bullet - c$ from playing 2 against rival investment 1. Since $4b^\bullet + \frac{m}{3} - c > 0$ from Condition 2, there is an EZ with $a_{BA} = 2$ and μ_B puts probability 1 on $b^\bullet + \frac{m}{3}$. It is impossible to have $a_{BA} = 1$ in EZ. This is because $b^*(1, 1) > b^*(1, 2)$, and under the inference $b^*(1, 2)$ we already have that the best response to 1 is 2, so the same also holds under any higher belief about complementarity. Also, we have $a_{BB} = 2$, since 2 must best respond to both 1 and 2. So in every such EZ, Θ_A's conditional fitness against group A is $2b^\bullet$ and Θ_B's conditional fitness against group A is $6b^\bullet - c$, with $2b^\bullet > 6b^\bullet - c$ by Condition 1. Also, Θ_A's conditional fitness against group B is $3b^\bullet$, while Θ_B's conditional fitness against group B is $8b^\bullet - c$. Again, $3b^\bullet > 8b^\bullet - c$ by Condition 1.

Next, we show Θ_B has strictly higher fitness than Θ_A in every EZ with $p_B = 1$. There is no EZ with $a_{BB} = 1$. This is because $b^*(1, 1) = b^\bullet + \frac{m}{2}$. As discussed before, under this inference the best response to 1 is 2, not 1. Now suppose $a_{BB} = 2$. Then μ_B puts probability 1 on $b^*(2, 2) = b^\bullet + \frac{m}{4}$. With this inference, the biased agents expect $1 \cdot (3(b^\bullet + \frac{m}{4}) - m) = 3b^\bullet - \frac{m}{4}$ from playing 1 against rival investment 2, and expect $2 \cdot (4(b^\bullet + \frac{m}{4}) - m) - c = 8b^\bullet - c$ from playing 2 against rival investment 2. We have $5b^\bullet + \frac{m}{4} - c > 0$ from Condition 2, so 2 best responds to 2. We must have $a_{AA} = a_{AB} = 1$. We conclude the unique EZ behavior is

$(a_{AA}, a_{AB}, a_{BA}, a_{BB}) = (1, 1, 1, 2)$, since the biased agents expect $1 \cdot (2(b^\bullet + \frac{m}{4}) - m) = 2b^\bullet - \frac{m}{2}$ from playing 1 against rival investment 1, and expect $2 \cdot (3(b^\bullet + \frac{m}{4}) - m) - c = 6b^\bullet - \frac{m}{2} - c$ from playing 2 against rival investment 1. We have $4b^\bullet - c < 0$ from Condition 1, so 1 best responds to 1. In the unique EZ with $p = (0, 1)$, the fitness of Θ_A is $2b^\bullet$ and the fitness of Θ_B is $8b^\bullet - c$, where $8b^\bullet - c > 2b^\bullet$ by Condition 1.

A.6 Proof of Proposition 4

Proof. When $\Theta_A = \Theta_B = \Theta^\bullet$, with any (p_A, p_B), we show adherents of both models have 0 fitness in every EZ. Suppose instead that the match between groups g and g' reach a terminal node other than z_1 with positive probability. Let n_L be the last non-terminal node reached with positive probability, so we must have $L \geq 2$, and also that nodes $n_1, ..., n_{L-1}$ are also reached with positive probability. So Drop must be played with probability 1 at n_L. Since n_L is reached with positive probability, correctly specified agents hold correct beliefs about opponent's play at n_L, which means at n_{L-1} it cannot be optimal to play Across with positive probability since this results in a loss of ℓ compared to playing Drop, a contradiction.

Now let $\Theta_A = \Theta^\bullet$, $\Theta_B = \Theta^{An}$ and let $p_B \in (0, 1)$. We claim there is an EZ where $d^k_{AA} = 1$ for every k, $d^k_{AB} = 0$ for every even k with $k < K$, $d^k_{AB} = 1$ for every other k, $d^k_{BA} = 0$ for every odd k and $d^k_{BA} = 1$ for every even k, and $d^k_{BB} = 0$ for every k with $k < K$, $d^K_{BB} = 1$. It is easy to see that the behavior (d_{AA}) is optimal under correct belief about opponent's play. In the Θ_A vs. Θ_B matches, the conjecture about A's play $\hat{d}^k_{AB} = 2/K$ for k even, $\hat{d}^k_{AB} = 1$ for k odd minimizes KL divergence among all strategies in $\mathbb{A}^{An}$, given B's play. To see this, note that when B has the role of P2, opponent Drops immediately. When B has the role of P1, the outcome is always z_K. So a conjecture with $\hat{d}^k_{AB} = x$ for every even k has the conditional KL divergence of:

$$\sum_{k \leq K-1 \text{ odd}} \underbrace{0 \cdot \ln\left(\frac{0}{0}\right)}_{(1,z_k) \text{ for } k \leq K-1 \text{ odd}} + \sum_{k \leq K-1 \text{ even}} \underbrace{0 \cdot \ln\left(\frac{0}{(1/2) \cdot (1-x)^{(k/2)-1} \cdot x}\right)}_{(1,z_k) \text{ for } k \leq K-1 \text{ even}}$$
$$+ \underbrace{\frac{1}{2} \ln\left(\frac{1/2}{(1/2) \cdot (1-x)^{(K/2)-1} \cdot x}\right)}_{(1,z_K)} + \underbrace{0 \cdot \ln\left(\frac{0}{(1-x)^{(K/2)}}\right)}_{(1,z_{end})}$$

when matched with an opponent from Θ_A. Using $0 \cdot \ln(0) = 0$, the expression simplifies to $\frac{1}{2}\ln\left(\frac{1}{(1-x)^{(K/2)-1}\cdot x}\right)$, which is minimized among $x \in [0,1]$ by $x = 2/K$. Against this conjecture, the difference in expected payoff at node n_{K-1} from Across versus Drop is $(1-2/K)(g)+(2/K)(-\ell)$. This is strictly positive when $g > \frac{2}{K-2}\ell$. This means the continuation value at n_{K-1} is at least g larger than the payoff of Dropping at n_{K-3}, so again Across has strictly higher expected payoff than Drop. Inductively, (d^k_{BA}) is optimal given the belief $(\hat{d}^k_{AB})$. Also, (d^k_{AB}) is optimal as it results in the highest possible payoff. We can similarly show that the conjecture $\hat{d}^k_{BB}$ with $\hat{d}^k_{BB} = 2/K$ for k even, $\hat{d}^k_{BB} = 0$ for k odd minimizes KL divergence conditional on Θ_B opponent, and (d^k_{BB}) is optimal given this conjecture.

As $p_B \to 0$, we find an EZ where adherents of A have fitness approaching 0, whereas the adherents of B have fitness approaching at least $\frac{1}{2}(((K/2)-1)g - \ell) > 0$ since $g > \frac{2}{K-2}\ell$. This shows Θ_A is not evolutionarily stable against Θ_B.

But consider the same (d_{AA}, d_{AB}, d_{BA}) and suppose $d^k_{BB} = 1$ for every k. Taking $p_B \to 1$, we find an EZ where adherents of B have fitness 0, adherents of A have fitness $\frac{1}{2}\cdot((K/2)g+\ell) > 0$. This shows Θ_B is not evolutionarily stable against Θ_A. □

A.7 Proof of Proposition 5

Proof. Take $g > \frac{2}{K-2}\ell$ in the centipede game. The misspecified agent thinks a group B agent in the role of P2 and a group A agent in either role has a probability $2/K$ of stopping at every node. Under this belief, choosing to continue instead of drop means there is a $(K-2)/K$ chance of gaining g, but a $2/K$ chance of losing ℓ. Since we assume $g > \frac{2}{K-2}\ell$, it is strictly better to continue. When p fraction of the agents are correctly specified, the fitness of $\Theta^\bullet$ is $p \cdot 0 + (1-p) \cdot (\frac{1}{2}\frac{g(K-2)}{2} + \frac{1}{2}(\frac{gK}{2} + \ell))$, while the fitness of Θ^{An} is $p \cdot [\frac{1}{2}(\frac{g(K-2)}{2} - \ell) + \frac{1}{2}\frac{g(K-2)}{2}] + (1-p)[\frac{1}{2}(\frac{g(K-2)}{2} - \ell) + \frac{1}{2}(\frac{gK}{2} + \ell)]$. The difference in fitness is $-p[\frac{1}{2}(\frac{g(K-2)}{2} - \ell) + \frac{1}{2}\frac{g(K-2)}{2}] + (1-p)\frac{1}{2}\ell$. Simplifying, this is $\frac{1}{2}\ell - p \cdot \frac{g(K-2)}{2}$, a strictly decreasing function in p. When $p = \frac{\ell}{g(K-2)}$, which is a number strictly between 0 and 1/2 from the assumption $g > \frac{2}{K-2}\ell$ in the centipede game, the two models have the same fitness. Furthermore, since the payoff difference is linear in p with a negative slope, the difference in fitness is negative when $p > \frac{\ell}{g(K-2)}$—so that Θ^{An} outperforms $\Theta^\bullet$ under these population shares—and conversely, the difference in fitness is positive when $p < \frac{\ell}{g(K-2)}$. Thus, we have this fraction of the population being correctly specified forms a stable population share. □

A.8 Proof of Proposition 6

Proof. In the Θ^{An} vs. Θ^{An} match, the adherents of Θ^{An} hold the belief that $\hat{d}^k_{BB} = 2/K$ for every even k. In the role of P1, at node k for $k \leq K-3$, stopping gives them k but continuing gives them a $(K-2)/K$ chance to get at least $k+2$, and we have $k \leq \frac{K-2}{K}(k+2) \iff 2k \leq 2K-4 \iff k \leq K-2$. At node $K-1$, the agent gets $K-1$ from dropping but expects $(K+2) \cdot \frac{K-2}{K}$ from continuing, and $(K+2) \cdot \frac{K-2}{K} - (K-1) = \frac{K^2-4-K^2+K}{K} = \frac{K-4}{K} > 0$ since $K \geq 6$.

In the $\Theta^\bullet$ vs. Θ^{An} match, the adherents of Θ^{An} hold the belief that $\hat{d}^k_{AB} = 2/K$ for every k. By the same arguments as before, the behavior of the adherents of Θ^{An} are optimal given these beliefs. Also, the adherents of $\Theta^\bullet$ have no profitable deviations since they are best responding both as P1 and P2.

When p fraction of the agents are correctly specified, in the dollar game the fitness of $\Theta^\bullet$ is $p \cdot 0.5 + (1-p) \cdot (\frac{1}{2}(K-1) + \frac{1}{2}K)$, while the fitness of Θ^{An} is $p \cdot 0 + (1-p) \cdot (\frac{1}{2} \cdot 0 + \frac{1}{2}K)$. For any p, the fitness of $\Theta^\bullet$ is strictly higher than that of Θ^{An}. □

B Existence and Continuity of Equilibrium Zeitgeists

We provide a few technical results about the existence of EZ and the upper-hemicontinuity of the set of EZs with respect to population share. We suppose that $|\mathcal{G}| = 1$ for simplicity, but analogous results would hold for environments with multiple situations. Note that the same belief endogeneity that generates new stability phenomena in Section 4 also leads to some difficulty in establishing existence and continuity results, as agents draw different inferences with different rates of interactions with the various groups.

We provide two sets of results. The first concerns environments where the expected KL divergence of any parameter in the model is finite under any strategy profile (for example, when every parameter conjectures a full-support distribution over consequences in $\mathbb{Y}$ under every strategy profile, and the support of the monitoring signal does not vary with opponent's strategy). The second focuses on environments with strategic certainty, so monitoring signals do not have full support and instead perfectly reveal opponent's strategy. But, we impose the same finite KL divergence requirement on the consequences.

For each $g, g' \in \{A, B\}$, define $K_{g,g'} : \mathbb{A}^2 \times \mathcal{G} \times \Theta_g \to \mathbb{R}$ by $K_{g,g'}(a_i, a_{-i}, G; (a_A, a_B, F)) = D_{KL}(F^\bullet(a_i, a_{-i}, G) \times \varphi^\bullet(a_{-i}) \parallel F(a_i, a_{g'}) \times \varphi^\bullet(a_{g'}))$. This is the KL divergence of the parameter

$(a_A, a_B, F) \in \Theta_g$ in situation G based on the data generated from the strategy profile (a_i, a_{-i}).

B.1 Environments with Full-Support Monitoring Signals

Let two models, Θ_A, Θ_B be fixed. Also fix population shares p. For $g, g' \in \{A, B\}$, define $V_{g,g'} : \mathbb{A}^2 \times \Theta_g \to \mathbb{R}$ to be $V_{g,g'}(a_i, a_{-i}, (\hat{a}_A, \hat{a}_B, F)) := \mathbb{E}_{y \sim F(a_i, \hat{a}_{g'})}(\pi(y))$, the expected payoff from choosing strategy a_i under the parameter $(\hat{a}_A, \hat{a}_B, F)$ when matched with an opponent from group g'. Extend the domain of the third argument of $V_{g,g'}$ from Θ_g to $\Delta(\Theta_g)$ by linearity. The $K_{g,g'}$ function defined above specializes in the case of of $|\mathcal{G}| = 1$ to be $K_{g,g'}(a_i, a_{-i}; (\hat{a}_A, \hat{a}_B, F)) = D_{KL}(F^\bullet(a_i, a_{-i}) \times \varphi^\bullet(a_{-i}) \parallel F(a_i, \hat{a}_{g'}) \times \varphi^\bullet(\hat{a}_{g'}))$.

Assumption A.1. $\mathbb{A}, \Theta_A, \Theta_B$ *are compact metrizable spaces.*

Assumption A.2. *For every* $g, g' \in \{A, B\}$, $V_{g,g'}$ *is continuous.*

Assumption A.3. *For every* $g, g' \in \{A, B\}$, $K_{g,g'}$ *is well-defined and finite on its domain* $\mathbb{A}^2 \times \Theta_g$.

Assumption A.4. *For every* $g, g' \in \{A, B\}$, $K_{g,g'}$ *is continuous.*

Assumption A.5. $\mathbb{A}$ *is convex and, for* $g, g' \in \{A, B\}$, *all* $a_{-i} \in \mathbb{A}$ *and all* $\mu_g \in \Delta(\Theta_g)$, $a_i \mapsto V_{g,g'}(a_i, a_{-i}; \mu_g)$ *is quasiconcave.*

We show the existence of equilibrium zeitgeists using the Kakutani-Fan-Glicksberg fixed point theorem, applied to the correspondence which maps strategy profiles and beliefs over parameters into best replies and beliefs over KL-divergence minimizing parameter. We start with a lemma.

Lemma A.1. *For* $g \in \{A, B\}$, $a = (a_{AA}, a_{AB}, a_{BA}, a_{BB}) \in \mathbb{A}^4$, *and* $0 \leq p_g \leq 1$, *let*

$$\Theta_g^*(a, p_g) := \underset{\hat{\theta} \in \Theta_g}{\arg\min} \left\{ \, p_g \cdot K_{g,g}(a_{g,g}, a_{g,g}; \hat{\theta}) + (1 - p_g) \cdot K_{g,-g}(a_{g,-g}, a_{-g,g}; \hat{\theta}) \, \right\}.$$

Then, Θ_g^* *is upper hemicontinuous in its arguments.*

This lemma says the set of KL-minimizing parameters is upper hemicontinuous in strategy profile and population share. This leads to the existence result.

Proposition A.1. *Under Assumptions A.1, A.2, A.3, A.4, and A.5, an equilibrium zeitgeist exists.*

Next, upper hemicontinuity in p_g in Lemma A.1 allows us to deduce the upper hemicontinuity of the EZ correspondence in population shares.

Proposition A.2. *Fix two models Θ_A, Θ_B. The set of equilibrium zeitgeists is an upper hemicontinuous correspondence in p_B under Assumptions A.1, A.2, A.3, and A.4.*

B.2 Proofs of Results in Appendix B.1

B.2.1 Proof of Lemma A.1

Proof. Write the minimization objective as

$$W(a, p_g, \hat{\theta}) := p_g \cdot K_{g,g}(a_{g,g}, a_{g,g}; \hat{\theta}) + (1 - p_g) \cdot K_{g,-g}(a_{g,-g}, a_{-g,g}; \hat{\theta}),$$

a continuous function of $(a, p_g, \hat{\theta})$ by Assumption A.4. Suppose we have a sequence $(a^{(n)}, p_g^{(n)}) \to (a^*, p_g^*) \in \mathbb{A}^4 \times [0,1]$ and let $\hat{\theta}^{(n)} \in \Theta_g^*(a^{(n)}, p_g^{(n)})$ for each n, with $\theta^{(n)} \to \theta^* \in \Theta_g$. For any other $\theta' \in \Theta_g$, note that $W(a^*, p_g^*, \theta') = \lim_{n\to\infty} W(a^{(n)}, p_g^{(n)}, \theta')$ by continuity. But also by continuity, $W(a^*, p_g^*, \theta^*) = \lim_{n\to\infty} W(a^{(n)}, p_g^{(n)}, \hat{\theta}^{(n)})$ and $W(a^{(n)}, p_g^{(n)}, \hat{\theta}^{(n)}) \leq W(a^{(n)}, p_g^{(n)}, \theta')$ for every n. It therefore follows $W(a^*, p_g^*, \theta^*) \leq W(a^*, p_g^*, \theta')$. □

B.2.2 Proof of Proposition A.1

Proof. Consider the correspondence $\Gamma : \mathbb{A}^4 \times \Delta(\Theta_A) \times \Delta(\Theta_B) \rightrightarrows \mathbb{A}^4 \times \Delta(\Theta_A) \times \Delta(\Theta_B)$,

$$\Gamma(a_{AA}, a_{AB}, a_{BA}, a_{BB}, \mu_A, \mu_B) :=$$
$$(\mathrm{BR}_{AA}(\mu_A), \mathrm{BR}_{AB}(\mu_A), \mathrm{BR}_{BA}(\mu_B), \mathrm{BR}_{BB}(\mu_B), \Delta(\Theta_A^*(a)), \Delta(\Theta_B^*(a))),$$

where $\mathrm{BR}_{gg'}(\mu_g) := \arg\max_{a_i \in \mathbb{A}} V_{g,g'}(a_i, a_{-i}; \mu_g)$ (this is well-defined because $V_{g,g'}$ does not depend on its second argument) and, for each $g \in \{A, B\}$, we have omitted the dependence of the correspondence Θ_g^* on p_g. It is clear that fixed points of Γ are equilibrium zeitgeists.

We apply the Kakutani-Fan-Glicksberg theorem (see, e.g, Corollary 17.55 in Aliprantis and Border (2006)). By Assumptions A.1 and A.5, $\mathbb{A}$ is a compact and convex metric space, and each Θ_g is a compact metric space, so it follows the domain of Γ is a nonempty, compact and convex metric space. We need only verify that Γ has closed graph, non-empty values, and convex values.

To see that Γ has closed graph, the previous lemma shows the upper hemicontinuity of $\Theta_A^*(a)$ and $\Theta_B^*(a)$ in a, and Theorem 17.13 of Aliprantis and Border (2006) then implies $\Delta(\Theta_A^*(a))$ and $\Delta(\Theta_B^*(a))$ are also upper hemicontinuous in a. It is a standard argument that since Assumption A.2 supposes $V_{AA}, V_{AB}, V_{BA,}V_{BB}$ are continuous, it implies the best-response correspondences $\text{BR}_{AA}(\mu_A)$, $\text{BR}_{AB}(\mu_A)$, $\text{BR}_{BA}(\mu_B)$, $\text{BR}_{BB}(\mu_B)$ have closed graphs.

To see that Γ is non-empty, recall that each $a_i \mapsto V_{g,g'}(a_i, a_{-i}; \mu_g)$ is a continuous function on a compact domain, so it must attain a maximum on $\mathbb{A}$. Similarly, the minimization problem that defines each $\Theta_g^*(a)$ is a continuous function of the parameter over a compact domain of possible parameters, so it attains a minimum. Thus each $\Delta(\Theta_g^*(a))$ is the set of distributions over a non-empty set.

To see that Γ is convex valued, clearly $\Delta(\Theta_A^*(a))$ and $\Delta(\Theta_B^*(a))$ are convex valued by definition. Also, $a_i \mapsto V_{AA}(a_i, a_{-i}; \mu_A)$ is quasiconcave by Assumption A.5. That means if $a_i^{'}, a_i^{''} \in \text{BR}_{AA}(\mu_A)$, then for any convex combination $\tilde{a}_i$ of $a_i^{'}, a_i^{''}$, we have $V_{AA}(\tilde{a}_i, a_{-i}; \mu_A) \geq \min(V_{AA}(a_i^{'}, a_{-i}; \mu_A),\, V_{AA}(a_i^{''}, a_{-i}; \mu_A)) = \max_{a_i \in \mathbb{A}} V_{AA}(a_i, a_{-i}; \mu_A)$. Therefore, $\text{BR}_{AA}(\mu_A)$ is convex. For similar reasons, $\text{BR}_{AB}(\mu_A)$, $\text{BR}_{BA}(\mu_B)$, $\text{BR}_{BB}(\mu_B)$ are convex. □

B.2.3 Proof of Proposition A.2

Proof. Since $\mathbb{A}^4 \times \Delta(\Theta_A) \times \Delta(\Theta_B)$ is compact by Assumption A.1, we need only show that for every sequence $(p_B^{(k)})_{k\geq 1}$ and $(a^{(k)}, \mu^{(k)})_{k\geq 1} = (a_{AA}^{(k)}, a_{AB}^{(k)}, a_{BA}^{(k)}, a_{BB}^{(k)}, \mu_A^{(k)}, \mu_B^{(k)})_{k\geq 1}$ such that for every k, $(a^{(k)}, \mu^{(k)})$ is an EZ with $p = (1 - p_B^{(k)}, p_B^{(k)})$, $p_B^{(k)} \to p_B^*$, and $(a^{(k)}, \mu^{(k)}) \to (a^*, \mu^*)$, then (a^*, μ^*) is an EZ with $p = (1 - p_B^*, p_B^*)$.

We first show for all $g, g^{'} \in \{A, B\}$, $a_{g,g'}^*$ is optimal under the belief μ_g^*. By Assumption A.2, $V_{g,g'}(a_i, a_{-i}; \mu_g)$ is continuous, so by the property of convergence in distribution, $V_{g,g'}(a_{g,g'}^{(k)}, a_{g',g}^{(k)}; \mu_g^{(k)}) \to V_{g,g'}(a_{g,g'}^*, a_{g',g}^*; \mu_g^*)$. For any other $a_i' \in \mathbb{A}$, $V_{g,g'}(a_i', a_{g',g}^{(k)}; \mu_g^{(k)}) \to V_{g,g'}(a_i', a_{g',g}^*; \mu_g^*)$ and for every k, $V_{g,g'}(a_{g,g'}^{(k)}, a_{g',g}^{(k)}; \mu_g^{(k)}) \geq V_{g,g'}(a_i', a_{g',g}^{(k)}; \mu_g^{(k)})$. Therefore $a_{g,g'}^*$ best responds to the belief μ_g^*.

Next, we show parameters in the support of μ_g^* minimize weighted KL divergence for group g. Since $\Theta_g^*(a, p_g)$ represents the minimizers of a continuous function on a compact domain (by Lemma A.1), it is non-empty and closed. By Theorem 17.13 of Aliprantis and Border (2006), the correspondence $\tilde{H} : \mathbb{A}^4 \times [0, 1] \rightrightarrows \Delta(\Theta_g)$ defined so that $\tilde{H}(a, p_g) := \Delta(\Theta_g^*(a, p_g))$ is also upper hemicontinuous. For every k, $\mu_g^{(k)} \in \tilde{H}(a^{(k)}, p_g^{(k)})$, and $\mu_g^{(k)} \to \mu_g^*$, $a^{(k)} \to a^*, p_g^{(k)} \to p_g^*$. Therefore, $\mu_g^* \in \tilde{H}(a^*, p_g^*)$, that is to say μ_g^* is supported on the minimizers of weighted KL divergence. □

B.3 Environments with Strategic Certainty

Let two models, Θ_A, Θ_B be fixed. Suppose we are in an environment with strategic certainty, so each Θ_g has the form $\mathbb{A}^2 \times \mathcal{F}_g$, $\mathbb{M} = \mathbb{A}$, and for every $a_{-i} \in \mathbb{A}$, $\varphi^\bullet(a_{-i})$ puts probability 1 on a_{-i}. Fix population shares p. As discussed in Section 2.4, we omit the part of the parameter that corresponds to conjectures about others' strategies and simply view beliefs as elements in $\Delta(\mathcal{F}_g)$.

For $g \in \{A, B\}$, let $U_g : \mathbb{A}^2 \times \mathcal{F}_g \to \mathbb{R}$ be defined by $U_g(a_i, a_{-i}; F) = U_i(a_i, a_{-i}; \delta_F)$ with $U_i : \mathbb{A}^2 \times (\Delta(\mathcal{F}_A) \cup \Delta(\mathcal{F}_B))$ as defined before. Extend the domain of the third argument of U_g from $\mathcal{F}_g$ to $\Delta(\mathcal{F}_g)$ by linearity. Also, the $K(a_i, a_{-i}; F)$ function specializes in the case of $|\mathcal{G}| = 1$ to be $K(a_i, a_{-i}; F) = D_{KL}(F^\bullet(a_i, a_{-i}) \parallel F(a_i, a_{-i}))$.

Assumption A.6. $\mathbb{A}, \mathcal{F}_A, \mathcal{F}_B$ *are compact metrizable spaces.*

Assumption A.7. *For each* $g \in \{A, B\}$, U_g *is continuous.*

Assumption A.8. K *is well-defined and finite on its domain* $\mathbb{A}^2 \times (\mathcal{F}_A \cup \mathcal{F}_B)$.

Under Assumption A.8, we can define $K_A : \mathbb{A}^2 \times \mathcal{F}_A \to \mathbb{R}$ and $K_B : \mathbb{A}^2 \times \mathcal{F}_B \to \mathbb{R}$, with $K_g(a_i, a_{-i}; F) = K(a_i, a_{-i}; F)$ for each $g \in \{A, B\}$, $a_i, a_{-i} \in \mathbb{A}$, and $F \in \mathcal{F}_g$.

Assumption A.9. *For each* $g \in \{A, B\}$, K_g *is continuous.*

Assumption A.10. $\mathbb{A}$ *is convex and, for* $g \in \{A, B\}$, *all* $a_{-i} \in \mathbb{A}$ *and all* $\mu_g \in \Delta(\mathcal{F}_g)$, $a_i \mapsto U_g(a_i, a_{-i}; \mu_g)$ *is quasiconcave.*

We show the existence of equilibrium zeitgeists using the Kakutani-Fan-Glicksberg fixed point theorem, applied to the correspondence which maps strategy profiles and beliefs over parameters into best replies and beliefs over KL-divergence minimizing parameter. We start with a lemma.

Lemma A.2. *For* $g \in \{A, B\}$, $a = (a_{AA}, a_{AB}, a_{BA}, a_{BB}) \in \mathbb{A}^4$, *and* $0 \le p_g \le 1$, *let*

$$\Theta_g^*(a, p_g) := \underset{\hat{F} \in \mathcal{F}_g}{\arg\min} \left\{ \; p_g \cdot K_g(a_{g,g}, a_{g,g}; \hat{F}) + (1 - p_g) \cdot K_{g,}(a_{g,-g}, a_{-g,g}; \hat{F}) \; \right\}.$$

Then, Θ_g^* is upper hemicontinuous in its arguments.

This lemma says the set of KL-minimizing parameters is upper hemicontinuous in strategy profile and population share. This leads to the existence result.

Proposition A.3. *Under Assumptions A.6, A.7, A.8, A.9, and A.10, an equilibrium zeitgeist exists.*

Next, upper hemicontinuity in p_g in Lemma A.2 allows us to deduce the upper hemicontinuity of the EZ correspondence in population shares.

Proposition A.4. *Fix two models Θ_A, Θ_B. The set of equilibrium zeitgeists is an upper hemicontinuous correspondence in p_B under Assumptions A.6, A.7, A.8, and A.9.*

B.4 Proofs of Results in Appendix B.3

B.4.1 Proof of Lemma A.2

Proof. Write the minimization objective as

$$W(a, p_g, \hat{F}) := p_g \cdot K_g(a_{g,g}, a_{g,g}; \hat{F}) + (1 - p_g) \cdot K_{g,}(a_{g,-g}, a_{-g,g}; \hat{F}),$$

a continuous function of $(a, p_g, \hat{F})$ by Assumption A.9. Suppose we have a sequence $(a^{(n)}, p_g^{(n)}) \to (a^*, p_g^*) \in \mathbb{A}^4 \times [0,1]$ and let $F^{(n)} \in \Theta_g^*(a^{(n)}, p_g^{(n)})$ for each n, with $F^{(n)} \to F^* \in \mathcal{F}_g$. For any other $F' \in \mathcal{F}_g$, note that $W(a^*, p_g^*, F') = \lim_{n\to\infty} W(a^{(n)}, p_g^{(n)}, F')$ by continuity. But also by continuity, $W(a^*, p_g^*, F^*) = \lim_{n\to\infty} W(a^{(n)}, p_g^{(n)}, F^{(n)})$ and $W(a^{(n)}, p_g^{(n)}, F^{(n)}) \leq W(a^{(n)}, p_g^{(n)}, F')$ for every n. It therefore follows $W(a^*, p_g^*, F^*) \leq W(a^*, p_g^*, F')$. □

B.4.2 Proof of Proposition A.3

Proof. Consider the correspondence $\Gamma : \mathbb{A}^4 \times \Delta(\mathcal{F}_A) \times \Delta(\mathcal{F}_B) \rightrightarrows \mathbb{A}^4 \times \Delta(\mathcal{F}_A) \times \Delta(\mathcal{F}_B)$,

$$\Gamma(a_{AA}, a_{AB}, a_{BA}, a_{BB}, \mu_A, \mu_B) :=$$
$$(\text{BR}(a_{AA}, \mu_A), \text{BR}(a_{BA}, \mu_A), \text{BR}(a_{AB}, \mu_B), \text{BR}(a_{BB}, \mu_B), \Delta(\Theta_A^*(a)), \Delta(\Theta_B^*(a))),$$

where $\text{BR}(a_{-i}, \mu_g) := \underset{\hat{a}_i \in \mathbb{A}}{\arg\max} U_g(\hat{a}_i, a_{-i}; \mu_g)$ and, for each $g \in \{A, B\}$, we have omitted the dependence of the correspondence Θ_g^* on p_g. It is clear that fixed points of Γ are equilibrium zeitgeists.

We apply the Kakutani-Fan-Glicksberg theorem (see, e.g., Corollary 17.55 in Aliprantis and Border (2006)). By Assumptions A.6 and A.10, $\mathbb{A}$ is a compact and convex metric space, and each $\mathcal{F}_g$ is a compact metric space, so it follows the domain of Γ is a nonempty, compact and convex metric space. We need only verify that Γ has closed graph, non-empty values, and convex values.

To see that Γ has closed graph, the previous lemma shows the upper hemicontinuity of $\Theta_A^*(a)$ and $\Theta_B^*(a)$ in a, and Theorem 17.13 of Aliprantis and Border (2006) then implies $\Delta(\Theta_A^*(a))$ and $\Delta(\Theta_B^*(a))$ are also upper hemicontinuous in a. It is a standard argument that since Assumption A.7 supposes U_A, U_B are continuous, it implies the best-response correspondence BR has closed graph.

To see that Γ is non-empty, recall that each $a_i \mapsto U_g(a_i, a_{-i}; \mu_g)$ is a continuous function on a compact domain, so it must attain a maximum on $\mathbb{A}$. Similarly, the minimization problem that defines each $\Theta_g^*(a)$ is a continuous function of the parameter over a compact domain of possible parameters, so it attains a minimum. Thus each $\Delta(\Theta_g^*(a))$ is the set of distributions over a non-empty set.

To see that Γ is convex valued, clearly $\Delta(\Theta_A^*(a))$ and $\Delta(\Theta_B^*(a))$ are convex valued by definition. Also, $a_i \mapsto U_g(a_i, a_{-i}; \mu_g)$ is quasiconcave by Assumption A.10. That means if $a_i^{'}, a_i^{''} \in \text{BR}(a_{-i}, \mu_g)$, then for any convex combination $\tilde{a}_i$ of $a_i^{'}, a_i^{''}$, we have $U_g(\tilde{a}_i, a_{-i}; \mu_g) \geq \min(U_g(a_i^{'}, a_{-i}; \mu_g),\ U_g(a_i^{''}, a_{-i}; \mu_g)) = \max_{a_i \in \mathbb{A}} U_g(a_i, a_{-i}; \mu_g)$. Therefore, $\text{BR}(a_{-i}, \mu_g)$ is convex. □

B.4.3 Proof of Proposition A.4

Proof. Since $\mathbb{A}^4 \times \Delta(\mathcal{F}_A) \times \Delta(\mathcal{F}_B)$ is compact by Assumption A.6, we need only show that for every sequence $(p_B^{(k)})_{k\geq 1}$ and $(a^{(k)}, \mu^{(k)})_{k\geq 1} = (a_{AA}^{(k)}, a_{AB}^{(k)}, a_{BA}^{(k)}, a_{BB}^{(k)}, \mu_A^{(k)}, \mu_B^{(k)})_{k\geq 1}$ such that for every k, $(a^{(k)}, \mu^{(k)})$ is an EZ with $p = (1 - p_B^{(k)}, p_B^{(k)})$, $p_B^{(k)} \to p_B^*$, and $(a^{(k)}, \mu^{(k)}) \to (a^*, \mu^*)$, then (a^*, μ^*) is an EZ with $p = (1 - p_B^*, p_B^*)$.

We first show for all $g, g^{'} \in \{A, B\}$, $a_{g,g^{'}}^*$ is optimal under the belief μ_g^*. By Assumption A.7, $U_g(a_i, a_{-i}; \mu_g)$ is continuous, so by property of convergence in distribution, $U_g(a_{g,g^{'}}^{(k)}, a_{g^{'},g}^{(k)}; \mu_g^{(k)}) \to U_g(a_{g,g^{'}}^*, a_{g^{'},g}^*; \mu_g^*)$. For any other $a_i' \in \mathbb{A}$, $U_g(a_i', a_{g^{'},g}^{(k)}; \mu_g^{(k)}) \to U_g(a_i', a_{g^{'},g}^*; \mu_g^*)$ and for every k, $U_g(a_{g,g^{'}}^{(k)}, a_{g^{'},g}^{(k)}; \mu_g^{(k)}) \geq U_g(a_i', a_{g^{'},g}^{(k)}; \mu_g^{(k)})$. Therefore $a_{g,g^{'}}^*$ best responds to the belief μ_g^*.

Next, we show parameters in the support of μ_g^* minimize weighted KL divergence for group g. Since $\Theta_g^*(a, p_g)$ represents the minimizers of a continuous function on a compact domain (by

Lemma A.2), it is non-empty and closed. By Theorem 17.13 of Aliprantis and Border (2006), the correspondence $\tilde{H} : \mathbb{A}^4 \times [0,1] \rightrightarrows \Delta(\mathcal{F}_g)$ defined so that $\tilde{H}(a, p_g) := \Delta(\Theta_g^*(a, p_g))$ is also upper hemicontinuous. For every k, $\mu_g^{(k)} \in \tilde{H}(a^{(k)}, p_g^{(k)})$, and $\mu_g^{(k)} \to \mu_g^*$, $a^{(k)} \to a^*, p_g^{(k)} \to p_g^*$. Therefore, $\mu_g^* \in \tilde{H}(a^*, p_g^*)$, that is to say μ_g^* is supported on the minimizers of weighted KL divergence. □

C Learning Foundation of Equilibrium Zeitgeists

We provide a foundation for equilibrium zeitgeists as the steady state of a learning system. This foundation considers a world where agents start with prior beliefs over parameters in a model. As in our framework from Section 2, these parameters correspond to conjectures about the stage game and about others' strategies.

At the end of every match, each agent observes their consequence and a monitoring signal. We show that under any asymptotically myopic policy, if behavior and beliefs converge, then the limit steady state must be an EZ. If the models allow agents to make rich enough inferences about opponents' strategies, then sufficiently accurate monitoring signals about opponent's play imply that agents must hold correct beliefs about others' strategies in the limit steady state. In particular, in environments that approach strategic certainty (that is, the monitoring signals are full support but they almost perfectly reveal opponent's strategy), limit steady state beliefs about others' strategies must be correct. Finally, if the true situation is redrawn every T periods and the agents reset their beliefs over parameters to their prior belief when the situation is redrawn, then their average payoffs approach their fitness in the EZ when T is large.

C.1 Regularity Assumptions

We make some regularity assumptions on the objective environments and on the two models Θ_A, Θ_B. These are similar to the regularity assumptions from Appendix B.

Suppose the strategy set $\mathbb{A}$ and the space of monitoring signals $\mathbb{M}$ are finite. Suppose the marginals of the models Θ_A, Θ_B on the dimension of fundamental uncertainty, denoted as $\mathcal{F}_A, \mathcal{F}_B$, are compact and metrizable spaces. Endow Θ_A and Θ_B with the product metric. Suppose that every $(a_A, a_B, F) \in \Theta_A \cup \Theta_B$ is such that for every $(a_i, a_{-i}) \in \mathbb{A}^2$ and every situation G, whenever $f^\bullet(a_i, a_{-i}, G)(y) > 0$, we also get $f(a_i, a_A)(y) > 0$ and $f(a_i, a_B)(y) > 0$, where f is the density or probability mass function for F. Suppose the monitoring signal has full support on $\mathbb{M}$ for every $a_{-i} \in \mathbb{A}$.

For each $g, g^{'} \in \{A, B\}$, recall that we defined $K_{g,g'} : \mathbb{A}^2 \times \mathcal{G} \times \Theta_g \to \mathbb{R}$ in Appendix B by $K_{g,g'}(a_i, a_{-i}, G; (a_A, a_B, F)) = D_{KL}(F^\bullet(a_i, a_{-i}, G) \times \varphi^\bullet(a_{-i}) \parallel F(a_i, a_{g'}) \times \varphi^\bullet(a_{g'}))$. Suppose each $K_{g,g'}$ is well defined and a continuous function of the parameter (a_A, a_B, F).

For $g \in \{A, B\}$, $F \in \mathcal{F}_g$, let $U_g(a_i, a_{-i}; F)$ be the expected payoffs of the strategy profile (a_i, a_{-i}) for i when consequences are drawn according to F. Assume U_A, U_B are continuous.

Suppose for every model Θ_g and every $(a_A, a_B, F) \in \Theta_g$ and $\epsilon > 0$, there exists an open neighborhood $V \subseteq \Theta_g$ of (a_A, a_B, F), so that for every $(\hat{a}_A, \hat{a}_B, \hat{F}) \in V$, $1 - \epsilon \leq [f(a_i, a_A)(y) \cdot \varphi^\bullet(a_A)(m)]/[\hat{f}(a_i, \hat{a}_A)(y) \cdot \varphi^\bullet(\hat{a}_A)(m)] \leq 1 + \epsilon$ and $1 - \epsilon \leq [f(a_i, a_B)(y) \cdot \varphi^\bullet(a_B)(m)]/[\hat{f}(a_i, \hat{a}_B)(y) \cdot \varphi^\bullet(\hat{a}_B)(m)] \leq 1 + \epsilon$ for all $a_i \in \mathbb{A}, y \in \mathbb{Y}, m \in \mathbb{M}$. Also suppose there is some $C > 0$ so that $\ln(f(a_i, a_A)(y) \cdot \varphi^\bullet(a_A)(m))$ and $\ln(f(a_i, a_B)(y) \cdot \varphi^\bullet(a_B)(m))$ are bounded in $[-C, C]$ for all $(a_A, a_B, F) \in \Theta_g$, $a_i, a_{-i} \in \mathbb{A}, y \in \mathbb{Y}, m \in \mathbb{M}$.

C.2 Learning Environment

We first consider an environment with only one true situation, $|\mathcal{G}| = 1$. Time is discrete and infinite, $t = 0, 1, 2, ...$ A unit mass of agents, $i \in [0, 1]$, enter the society at time 0. A $p_A \in (0, 1)$ measure of them are assigned to model A and the rest are assigned to model B. Each agent born into model g starts with the same full support prior over the model, $\mu_g^{(0)} \in \Delta(\Theta_g)$, and believes there is some $(a_A, a_B, F) \in \Theta_g$ so that every group g opponent always plays a_g and the consequences are always generated by F.

In each period t, agents are matched up uniformly at random to play the stage game. Each person in group g has p_g chance of matching with someone from group g, and matches with someone from group $-g$ with the complementary chance. Each agent i observes their opponent's group membership and chooses a strategy $a_i^{(t)} \in \mathbb{A}$. At the end of the match, the agent observes own consequence $y_i^{(t)}$ and a monitoring signal $m_i^{(t)} \in \mathbb{M}$ about the opponent's play, where $m_i^{(t)}$ is drawn from the distribution $\varphi^\bullet(a_{-i})$ if their opponent uses strategy a_{-i}. One example of this would be $\mathbb{M} = \mathbb{A}$ and $m_i^{(t)}$ is equal to the opponent's strategy with probability $\tau \in [0, 1)$ and is uniformly random on $\mathbb{M}$ with the complementary probability. Our results for the case when τ is close enough to 1 and each model has the form $\Theta_g = \mathbb{A}^2 \times \mathcal{F}_g$ provide a foundation for EZs in environments with strategic certainty.

The space of histories from one period is $\{A, B\} \times \mathbb{A} \times \mathbb{Y} \times \mathbb{M}$, with typical element

$(g_i^{(t)}, a_i^{(t)}, y_i^{(t)}, m_i^{(t)})$. It records the group membership of i's opponent $g_i^{(t)}$, i's strategy $a_i^{(t)}$,, i's consequence $y_i^{(t)}$, and i's monitoring signal about the matched opponent's strategy, $m_i^{(t)}$. Let $\mathbb{H}$ denote the space of all finite-length histories.

Given the assumption on the two models, there is a well-defined Bayesian belief operator for each model g, $\mu_g : \mathbb{H} \to \Delta(\Theta_g)$, mapping every finite-length history into a belief over parameters in Θ_g, starting with the prior $\mu_g^{(0)}$.

We also take as exogenously given policy functions for choosing strategies after each history. That is, $\mathfrak{a}_{g,g'} : \mathbb{H} \to \mathbb{A}$ for every $g, g' \in \{A, B\}$ gives the strategy that a group g agent uses against a group g' opponent after every history. Assume these policy functions are asymptotically myopic.

Assumption A.11. *For every $\epsilon > 0$, there exists N so that for any history h containing at least N matches against opponents of each group, $\mathfrak{a}_{g,g'}(h)$ is an ϵ-best response to the Bayesian belief $\mu_g(h)$.*

From the perspective of each agent i in group g, i's play against groups A and B, as well as i's belief over Θ_g, is a stochastic process $(\tilde{a}_{iA}^{(t)}, \tilde{a}_{iB}^{(t)}, \tilde{\mu}_i^{(t)})_{t \geq 0}$ valued in $\mathbb{A} \times \mathbb{A} \times \Delta(\Theta_g)$. The randomness is over the groups of opponents matched with in different periods, the strategies they play, and the random consequences and monitoring signals drawn at the end of the matches. Since there is a continuum of agents, the distribution over histories within each population in each period is deterministic. As such, there is a deterministic sequence $(\alpha_{AA}^{(t)}, \alpha_{AB}^{(t)}, \alpha_{BA}^{(t)}, \alpha_{BA}^{(t)}, \nu_A^{(t)}, \nu_B^{(t)}) \in \Delta(\mathbb{A})^4 \times \Delta(\Delta(\Theta_A)) \times \Delta(\Delta(\Theta_B))$ that describes the distributions of play and beliefs that prevail in the two sub-populations in every period t.

C.3 Steady State Limits are Equilibrium Zeitgeists

We state and prove the learning foundation of EZs. For $(\alpha^{(t)})_t$ a sequence valued in $\Delta(\mathbb{A})$ and $a^* \in \mathbb{A}$, $\alpha^{(t)} \to a^*$ means $\mathbb{E}_{\hat{a} \sim \alpha^{(t)}} \parallel \hat{a} - a^* \parallel \to 0$ as $t \to \infty$. For $(\nu^{(t)})_t$ a sequence valued in $\Delta(\Delta(\Theta_g))$ and $\mu^* \in \Delta(\Theta_g)$, $\nu^{(t)} \to \mu^*$ means $\mathbb{E}_{\hat{\mu} \sim \nu^{(t)}} \parallel \hat{\mu} - \mu^* \parallel \to 0$ as $t \to \infty$.

Proposition A.5. *Suppose the regularity assumptions in Appendix C.1 hold, and suppose Assumption A.11 holds. Suppose there exists $(a_{AA}^*, a_{AB}^*, a_{BA}^*, a_{BB}^*, \mu_A^*, \mu_B^*) \in \mathbb{A}^4 \times \Delta(\Theta_A) \times \Delta(\Theta_B)$ so that $(\alpha_{AA}^{(t)}, \alpha_{AB}^{(t)}, \alpha_{BA}^{(t)}, \alpha_{BA}^{(t)}, \nu_A^{(t)}, \nu_B^{(t)}) \to (a_{AA}^*, a_{AB}^*, a_{BA}^*, a_{BB}^*, \mu_A^*, \mu_B^*)$ and for each agent i in group g, almost surely $(\tilde{a}_{iA}^{(t)}, \tilde{a}_{iB}^{(t)}, \tilde{\mu}_i^{(t)}) \to (a_{gA}^*, a_{gB}^*, \mu_g^*)$. Then, $(a_{AA}^*, a_{AB}^*, a_{BA}^*, a_{BB}^*, \mu_A^*, \mu_B^*)$ is an equilibrium zeitgeist.*

Suppose further that for each g, the model Θ_g has the form $\mathbb{A}^2 \times \mathcal{F}_g$. There exists some $\underline{\tau} < 1$ so that for every $\tau \in (\underline{\tau}, 1)$ and $(a^*_{AA}, a^*_{AB}, a^*_{BA}, a^*_{BB}, \mu^*_A, \mu^*_B)$ satisfying the above conditions, we have that μ^*_A puts probability 1 on (a^*_{AA}, a^*_{AB}), μ^*_B puts probability 1 on (a^*_{BA}, a^*_{BB}).

Proof. For μ a belief and $g \in \{A, B\}$, let $u^\mu(a_i; g)$ represent subjective expected payoff from playing a_i against group g. Suppose $a^*_{AA} \notin \text{argmax}_{\hat{a} \in \mathbb{A}} u^{\mu^*_A}(\hat{a}; A)$ (the other cases are analogous). By the continuity assumptions on U_A (which is also bounded because $\mathcal{F}_A$ is bounded), there are some $\epsilon_1, \epsilon_2 > 0$ so that whenever $\mu_i \in \Delta(\Theta_A)$ with $\| \mu_i - \mu^*_A \| < \epsilon_1$, we also have $u^{\mu_i}(a^*_{AA}; A) < \max_{\hat{a} \in \mathbb{A}} u^{\mu_i}(\hat{a}; A) - \epsilon_2$. By the definition of asymptotically empirical best responses, find N so that $\mathfrak{a}_{A,A}(h)$ must be a myopic ϵ_2-best response when there are at least N periods of matches against A and B. Agent i has a strictly positive chance to match with groups A and B in every period. So, at all except a null set of points in the probability space, i's history eventually records at least N periods of play by groups A and B. Also, by assumption, almost surely $\tilde{\mu}_i^{(t)} \to \mu^*_A$. This shows that by asymptotically myopic best responses, almost surely $\tilde{a}_{iA}^{(k)} \not\to a^*_{AA}$, a contradiction.

Now suppose some $\theta^*_A = (a^*_A, a^*_B, f^*)$ in the support of μ^*_A does not minimize the weighted KL divergence in the definition of EZ (the case of a parameter θ^*_B in the support of μ^*_B not minimizing is similar). Then we have

$$\theta^*_A \notin \underset{\hat{\theta} \in \Theta_A}{\text{argmin}} \left\{ \begin{array}{c} (p_A) \cdot D_{KL}(F^\bullet(a^*_{AA}, a^*_{AA}) \times \varphi^\bullet(a^*_{AA}) \parallel \hat{F}(a^*_{AA}, \hat{a}_A) \times \varphi^\bullet(\hat{a}_A)) \\ +(1 - p_A) \cdot D_{KL}(F^\bullet(a^*_{AB}, a^*_{BA}) \times \varphi^\bullet(a^*_{BA}) \parallel \hat{F}(a^*_{AB}, \hat{a}_B) \times \varphi^\bullet(\hat{a}_B)) \end{array} \right\}$$

where $\hat{\theta} = (\hat{a}_A, \hat{a}_B, \hat{F})$.

This is equivalent to:

$$\theta^*_A \notin \underset{\hat{\theta} \in \Theta_A}{\text{argmin}} \left[\begin{array}{c} (p_A) \cdot \mathbb{E}_{(y,m) \sim F^\bullet(a^*_{AA}, a^*_{AA}) \times \varphi^\bullet(a^*_{AA})} \ln(\hat{f}(a^*_{AA}, \hat{a}_A)(y) \cdot \varphi^\bullet(\hat{a}_A)(m)) \\ +(1 - p_A) \cdot \mathbb{E}_{(y,m) \sim F^\bullet(a^*_{AB}, a^*_{BA}) \times \varphi^\bullet(a^*_{BA})} \ln(\hat{f}(a^*_{AB}, \hat{a}_B)(y) \cdot \varphi^\bullet(\hat{a}_B)(m)) \end{array} \right]$$

Let this objective, as a function of $\hat{\theta}$, be denoted $WL(\hat{\theta})$. There exists $\theta^{opt}_A = (a^{opt}_A, a^{opt}_B, f^{opt}) \in \Theta_A$ and $\delta, \epsilon > 0$ so that $(1 - \delta) WL(\theta^{opt}_A) - 2\delta C - 3\epsilon > (1 - \delta) WL(\theta^*_A)$. By assumption on the primitives, find open neighborhoods V^{opt} and V^* of $\theta^{opt}_A, \theta^*_A$ respectively, so that for all $a_i \in \mathbb{A}$, $g \in \{A, B\}$, $y \in \mathbb{Y}$, $m \in \mathbb{M}$, $1 - \epsilon \leq [f^{opt}(a_i, a^{opt}_g)(y) \cdot \varphi^\bullet(a^{opt}_g)(m)] / [\hat{f}(a_i, \hat{a}_g)(y) \cdot \varphi^\bullet(\hat{a}_g)(m)] \leq 1 + \epsilon$, for all $\hat{\theta} = (\hat{a}_A, \hat{a}_B, \hat{f}) \in V^{opt}$, and also $1 - \epsilon \leq [f^*(a_i, a^*_g)(y) \cdot \varphi^\bullet(a^*_g)(m)] / [\hat{f}(a_i, \hat{a}_g)(y) \cdot$

$\varphi^\bullet(\hat{a}_g)(m)] \leq 1+\epsilon$ for all $\hat{\theta} = (\hat{a}_A, \hat{a}_B, \hat{f}) \in V^*$. Also, by convergence of play in the populations, find T_1 so that in all periods $t \geq T_1$, $\alpha^{(t)}_{AA}(a^*_{AA}) \geq 1-\delta$ and $\alpha^{(t)}_{BA}(a^*_{BA}) \geq 1-\delta$.

Consider a probability space defined by $\Omega := (\{A,B\} \times \mathbb{A}^2 \times (\mathbb{Y})^{\mathbb{A}^2} \times \mathbb{M}^{\mathbb{A}})^\infty$ that describes the randomness in an agent's learning process. For a point $\omega \in \Omega$ and each period $t \geq 1$, $\omega_t = (g, a_{-i,A}, a_{-i,B}, (y_{a_i,a_{-i}})_{(a_i,a_{-i})\in\mathbb{A}^2}, (m_{a_{-i}})_{a_{-i}\in\mathbb{A}})$ specifies the group g of the matched opponent, the play $a_{-i,A}, a_{-i,B}$ of hypothetical opponents from groups A and B, the hypothetical consequence $y_{a_i,a_{-i}}$ that would be generated for every pair of strategies (a_i, a_{-i}) played, and the hypothetical monitoring signal $m_{a_{-i}}$ that would be generated for every opponent strategy a_{-i}. As notation, let $o(\omega,t)$, $a_{-i,A}(\omega,t)$, $a_{-i,B}(\omega,t)$, $y_{a_i,a_{-i}}(\omega,t)$, $m_{a_{-i}}(\omega,t)$ denote the corresponding components of ω_t. For $T_2 \geq T_1$, define $\mathbb{P}_{T_2}$ over $\Omega^\infty_{T_2} := \times_{t=T_2}^\infty(\{A,B\} \times \mathbb{A}^2 \times (\mathbb{Y})^{\mathbb{A}^2} \times \mathbb{M}^{\mathbb{A}})$ in the natural way. That is, it is independent across periods, and within each period, the density (or probability mass function if $\mathbb{Y}$ is finite) of $\omega_t = (g, a_{-i,A}, a_{-i,B}, (y_{a_i,a_{-i}})_{(a_i,a_{-i})\in\mathbb{A}^2}, (m_{a_{-i}})_{a_{-i}\in\mathbb{A}})$ is

$$p_g \cdot \alpha^{(t)}_{AA}(a_{-i,A})\alpha^{(t)}_{BA}(a_{-i,B}) \cdot \prod_{(a_i,a_{-i})\in\mathbb{A}^2} f^\bullet(a_i,a_{-i})(y_{a_i,a_{-i}}) \cdot \prod_{a_{-i}\in\mathbb{A}} \varphi^\bullet(a_{-i})(m_{a_{-i}}).$$

For $\theta = (a^\theta_A, a^\theta_B, F^\theta) \in \Theta_A$ with f^θ the density of F^θ, $\omega \in \Omega^\infty_{T_2}$, consider the process in $s = 1,2,3,...$

$$\ell_s(\theta,\omega) := \frac{1}{s}\sum_{t=T_2+1}^{T_2+s} \ln[f^\theta(a^*_{A,o(\omega,t)}, a^\theta_{o(\omega,t)})(y_{a^*_{A,o(\omega,t)},a_{-i,o(\omega,t)}(\omega,t)}(\omega,t)) \cdot \varphi^\bullet(a_{-i,o(\omega,t)}(\omega,t))(m_{a_{-i,o(\omega,t)}(\omega,t)}(\omega,t))].$$

By choice of the neighborhood V^*, for every s,

$$\begin{aligned}\sup_{\theta_A\in V^*} \ell_s(\theta_A,\omega) &\leq \epsilon + \frac{1}{s}\sum_{t=T_2+1}^{T_2+s} \ln[f^\theta(a^*_{A,o(\omega,t)}, a^*_{o(\omega,t)})(y_{a^*_{A,o(\omega,t)},a_{-i,o(\omega,t)}(\omega,t)}(\omega,t)) \\ &\qquad \cdot \varphi^\bullet(a_{-i,o(\omega,t)}(\omega,t))(m_{a_{-i,o(\omega,t)}(\omega,t)}(\omega,t))] \\ &\leq \epsilon + \frac{1}{s}\sum_{t=T_2+1}^{T_2+s} 1_{\{a_{-i,o(\omega,t)}(\omega,t)=a^*_{o(\omega,t),A}\}} \cdot \ln[f^\theta(a^*_{A,o(\omega,t)}, a^*_{o(\omega,t)})(y_{a^*_{A,o(\omega,t)},a^*_{o(\omega,t),A}}(\omega,t)) \\ &\qquad \cdot \varphi^\bullet(a^*_{o(\omega,t),A})(m_{a^*_{o(\omega,t),A}}(\omega,t))] \\ &\qquad + (1 - 1_{\{a_{-i,o(\omega,t)}(\omega,t)=a^*_{o(\omega,t),A}\}}) \cdot C.\end{aligned}$$

Since $T_2 \geq T_1$, in every period t, $\mathbb{P}_{T_2}(a_{-i,o(\omega,t)}(\omega,t) = a^*_{o(\omega,t),A}) \geq 1-\delta$. Let $(\xi_k)_{k\geq 1}$ a related stochastic process: it is i.i.d. such that each ξ_k has δ chance to be equal to C, $(1-\delta)p_A$ chance to be distributed according to $\ln(f^*(a^*_{AA}, a^*_A)(y)\cdot\varphi^\bullet(a^*_A)(m))$ where $y \sim f^\bullet(a^*_{AA}, a^*_{AA})$ and $m \sim \varphi^\bullet(a^*_{AA})$, and $(1-\delta)p_B$ chance to be distributed according to $\ln(f^*(a^*_{AB}, a^*_B)(y)\cdot\varphi^\bullet(a^*_B)(m))$ where $y \sim f^\bullet(a^*_{AB}, a^*_{BA})$ and $m \sim \varphi^\bullet(a^*_{BA})$. By law of large numbers, $\frac{1}{s}\sum_{k=1}^s \xi_k$ converges almost surely to $\delta C + (1-\delta)WL(\theta^*_A)$. By this comparison, $\limsup_s \sup_{\theta_A\in V^*} \ell_s(\theta_A,\omega) \leq \epsilon + \delta C + (1-\delta)WL(\theta^*_A)$ $\mathbb{P}_{T_2}$-almost surely. By a similar argument, $\liminf_s \inf_{\theta_A\in V^{opt}} \ell_s(\theta_A,\omega) \geq -\epsilon - \delta C + (1-\delta)WL(\theta^{opt}_A)$ $\mathbb{P}_{T_2}$-almost surely.

Along any ω where we have both $\limsup_s \sup_{\theta_A\in V^*} \ell_s(\theta_A,\omega) \leq \epsilon + \delta C + (1-\delta)WL(\theta^*_A)$ and $\liminf_s \inf_{\theta_A\in V^{opt}} \ell_s(\theta_A,\omega) \geq -\epsilon - \delta C + (1-\delta)WL(\theta^{opt}_A)$, if ω also leads to i always playing a^*_{AA} against group A and a^*_{AB} against group B in all periods starting with T_2+1, then the posterior belief assigns to V^* must tend to 0, hence $\tilde{\mu}^{(t)}_i \not\to \mu^*_A$. Starting from any length T_2 history h, there exists a subset $\hat{\Omega}_h \subseteq \Omega^\infty_{T_2}$ that leads to i not playing the EZ strategy in at least one period starting with T_2+1. So conditional on h, the probability of $\tilde{\mu}^{(t)}_i \to \mu^*_A$ is no larger than $1-\mathbb{P}_{T_2}(\hat{\Omega}_h)$. The unconditional probability is therefore no larger than $\mathbb{E}_h[1-\mathbb{P}_{T_2}(\hat{\Omega}_h)]$, where $\mathbb{E}_h$ is taken with respect to the distribution of period T_2 histories for i. But this term is also the probability of i playing non-EZ action at least once starting with period T_2. Since there are finitely many actions and $(\tilde{a}^{(t)}_{iA}, \tilde{a}^{(t)}_{iB}) \to (a^*_{AA}, a^*_{AB})$ almost surely, $\mathbb{E}_h[1-\mathbb{P}_{T_2}(\hat{\Omega}_h)]$ tends to 0 as $T_2\to\infty$. We have a contradiction as this shows $\tilde{\mu}^{(t)}_i \not\to \mu^*_A$ with probability 1.

Now we prove the second part of this proposition. Let $\bar{K} < \infty$ be an upper bound on $D_{KL}(F^\bullet(a_i,a_{-i}) \parallel \hat{F}(a_i,\hat{a}_{-i}))$ across all $a_i, a_{-i} \in \mathbb{A}$, $(\hat{a}_A, \hat{a}_B, \hat{F}) \in \Theta_A \cup \Theta_B$. Here $\bar{K}$ is finite because $\mathbb{A}$ is finite and $K_{g,g'}$ is continuous in the parameter, which is from a compact domain. Let $\varphi^\bullet_\tau(a_{-i})$ be the distribution over $\mathbb{M}$ when opponent plays a_{-i} and the monitoring structure is such that the monitoring signal matches the opponent's strategy with probability τ and is uniformly random on $\mathbb{M}$ with the complementary probability. It is clear that there exists some $\underline{\tau} < 1$ so that for any $a_{-i} \neq a'_{-i}$, $\tau \in (\underline{\tau}, 1)$, we get $\min(p_A,p_B)\cdot D_{KL}(\varphi^\bullet_\tau(a_{-i}) \parallel \varphi^\bullet_\tau(a'_{-i})) > \bar{K}$. Therefore, given any $(a^*_{AA}, a^*_{AB}, a^*_{BA}) \in \mathbb{A}^3$, the solution to

$$\min_{\hat{\theta}\in\Theta_A} \begin{bmatrix} (p_A)\cdot D_{KL}(F^\bullet(a^*_{AA}, a^*_{AA}) \times \varphi^\bullet(a^*_{AA}) \parallel \hat{F}(a^*_{AA}, \hat{a}_A) \times \varphi^\bullet(\hat{a}_A)) \\ +(1-p_A)\cdot D_{KL}(F^\bullet(a^*_{AB}, a^*_{BA}) \times \varphi^\bullet(a^*_{BA}) \parallel \hat{F}(a^*_{AB}, \hat{a}_B) \times \varphi^\bullet(\hat{a}_B)) \end{bmatrix}$$

must satisfy $\hat{a}_A = a^*_{AA}$, $\hat{a}_B = a^*_{BA}$, because (a^*_{AA}, a^*_{BA}, F) for any $F \in \Theta_A$ has a KL divergence no larger than $\bar{K}$. On the other hand, any $(\hat{a}_A, \hat{a}_B, \hat{F})$ with either $\hat{a}_A \neq a^*_{AA}$ or $\hat{a}_B \neq a^*_{BA}$ has KL divergence strictly larger than $\bar{K}$ by the choice of τ. $\square$

C.4 Multiple Situations

Now suppose there are multiple situations $G \in \mathcal{G}$ and a distribution $q \in \Delta(\mathcal{G})$, with $\mathcal{G}$ finite. At the start of period $t = 1$, Nature draws a situation $G^{(1)}$ from $\mathcal{G}$ according to q, and consequences are generated according to $F^\bullet(\cdot, \cdot, G^{(1)})$ until period $t = T+1$. In period $T+1$, Nature again draws a situation $G^{(2)}$ from $\mathcal{G}$ according to q, and consequences are generated according to $F^\bullet(\cdot, \cdot, G^{(2)})$ until period $t = 2T+1$, and so forth. Agents start with a prior over parameters in their group's model, $\mu_g^{(0)} \in \Delta(\Theta_g)$. In periods $T+1, 2T+1, ...$ agents reset their belief to $\mu_g^{(0)}$, and their belief in each period over the parameters in their model only use histories since the last reset. This belief corresponds to agents thinking that the data-generating process is redrawn according to $\mu_g^{(0)}$ every T periods.

Suppose for every $G \in \mathcal{G}$, the hypotheses of Proposition A.5 hold in a society where G is the only true situation. Denote $(a^*_{AA}(G), a^*_{AB}(G), a^*_{BA}(G), a^*_{BB}(G), \mu^*_A(G), \mu^*_B(G))$ as the limit of the agents' behavior and beliefs with situation G. Then it is straightforward to see that in a society with the situation redrawn every T periods, the expected undiscounted average payoff of an agent in group g approaches the fitness of g in the EZ characterized by the behavior and beliefs $(a^*_{AA}(G), a^*_{AB}(G), a^*_{BA}(G), a^*_{BB}(G), \mu^*_A(G), \mu^*_B(G))_{G \in \mathcal{G}}$ with the distribution q over situations, as $T \to \infty$. This provides a foundation for fitness in EZ as the agents' objective payoffs when the true situation changes sufficiently slowly.