Evolutionarily Stable (Mis)specifications: Theory and Applications*

Kevin He† Jonathan Libgober‡

First version: December 20, 2020
This version: April 27, 2024

Abstract

Toward explaining the persistence of biased inferences, we propose a framework to evaluate competing (mis)specifications in strategic settings. Agents with heterogeneous (mis)specifications coexist and draw Bayesian inferences about their environment through repeated play. The relative stability of (mis)specifications depends on their adherents' equilibrium payoffs. A key mechanism is the *learning channel*: the endogeneity of perceived best replies due to inference. We characterize when a rational society is only vulnerable to invasion by some misspecification through the learning channel. The learning channel leads to new stability phenomena, and can confer an evolutionary advantage to otherwise detrimental biases in economically relevant applications.

*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, 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 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: while they learn from data, they exclude the true data-generating process from consideration. For instance, past work has documented a number of prevalent statistical biases. Reasoning about economic fundamentals under the spell of these biases constitutes misspecified learning. Economists have become increasingly interested in the implications of Bayesian learning under particular misspecifications, for the most part taking them to be exogenously imposed.

Compared with other errors and mistakes, a distinctive component of misspecified *learning* is using data to form beliefs about the world. This raises a natural question: how does the ability to draw inferences affect the viability of such mistakes? We introduce an *evolutionary approach* to answer this question in strategic settings. Specifically, we associate the viability of a particular (mis)specification with the objective payoffs of individuals who adopt it. In contrast to contemporary papers that use the same criterion in single-agent decision problems (Fudenberg and Lanzani, 2022; Frick, Iijima, and Ishii, 2021), our key innovation is to focus on games, where this objective performance depends on strategic behavior in equilibrium.

Our main message is that the *learning channel* — i.e., the ability for agents to learn and draw (possibly wrong) inferences from data — adds new ways for biased individuals to develop strategically beneficial commitments, as their perceived best replies become endogenously determined by feedback. In particular, our approach lets us distinguish agents with dogmatic beliefs (which are exogenous and do not depend on observed data) from those with flexible beliefs (which are endogenously determined in equilibrium).

Our contribution is to emphasize two implications of the learning channel:

1. Due to its greater flexibility, misinference can confer strategic benefits in cases where dogmatic beliefs do not.

2. Misspecified learners are *polymorphic*: agents with a fixed bias may be weak in one environment but become stronger in another due to (endogenous) changes in beliefs.

Our main results fall under one of these two themes. On the former, we find general conditions under which no dogmatically wrong belief can persist in a rational society, but some misspecified agents can nevertheless do strictly better than rational incumbents through the learning channel. We also study applications where the invading misspecification is encoded in economically meaningful and natural biases. In particular, we illustrate how the persistence of biases can depend on the possibility of learning. Predictions about the viability of a mistake can be reversed depending on whether the biased agents hold dogmatic beliefs or undertake misinference.

On the latter, polymorphism makes it harder to predict the viability of a given error across different economic environments. Without the learning channel, we find some sufficient conditions that let us use the welfare of a given error in one society to extrapolate that it will not persist in another society. But these conclusions no longer hold for biased agents who may develop different beliefs in different environments.

1.1 Inference and Selecting Misspecified Beliefs about Correlation

To articulate some intuition for why inference can affect the selection of biases, we informally describe the main application of our framework in this paper. Consider a linear-quadratic-normal (LQN) game of incomplete information as the stage game, interpreted as an incomplete-information version of Cournot duopoly. A population of players (firms) match in pairs every period to play the stage game. The intercept of the demand curve is drawn i.i.d. across games, and within every pair players receive correlated information about this intercept in their game. After observing this signal, players choose a production quantity. The market price depends on the intercept of demand, the quantity choices of the firms, and a price elasticity parameter (which is fixed across matches).

We suppose that a small fraction of firms hold a dogmatically wrong belief about the signal correlation and invade a society that has correct beliefs about all game parameters. An important property of this game is that players gain from strategic commitments, and which commitments are valuable depends on assortativity. That is, if entrants are only paired with each other (perfectly assortative matching), then they can improve payoffs by committing to more cooperative strategies. If entrants are paired with the rational incumbents (uniform matching), then committing to more aggressive strategies can help them obtain more favorable outcomes compared to when incumbents play each other. Our contribution is to show that whether a certain biased belief about signal correlation leads to more cooperative or more aggressive play, and hence whether it will be selected for a given matching assortativity, depends on whether the learning channel is present.

When learning is absent, an increase in the subjective perception of correlation makes a player choose *less* aggressive strategies. Intuitively, because production quantities are strategic substitutes, a player who believes signals are excessively correlated will not increase their production by a large amount following an optimistic signal about demand, expecting the opponent to also produce more. But when inference is present, an exaggerated perception of correlation also leads the player to believe that market price is less elastic relative to the truth. This is because the agent overestimates the opponent's production level and is thus surprised by how little the price adjusts. Inferring a more inelastic price makes the player choose *more* aggressive strategies. While these forces move in opposite directions, the second effect dominates. Thus, the presence of inference can reverse the conclusion of which misperception outperforms rationality.

Now suppose the underlying elasticity parameter can take on multiple possible values, and different types of agents are evaluated based on their average performance across these parameters. In this scenario, a fixed belief about elasticity can be beneficial for some realizations of the true elasticity parameter but harmful for others. We use this idea to show that generally, there is some amount of variation in elasticity under which no entrant with a fixed misperception about correlation and elasticity can invade a rational society, but some entrant who misperceives correlation and makes flexible inferences about elasticity can do strictly better than the incumbents. This is because misinference leads the agent to make different commitments tailored to different realizations of elasticity, illustrating our first contribution mentioned above.

1.2 A Framework of Competing Specifications

To proceed more formally, our general framework encodes specifications in *models* that delineate feasible beliefs about the stage game. These models serve as the basic unit of cultural transmission. The model's adherents think that one of the model parameters describes the true stage game. They estimate the best-fitting parameter which determines their subjective preference. Models rise and fall in prominence based on the objective welfare of adherents, as higher payoffs confer greater evolutionary success.

When we allow for inference in the example above, the incumbents and the entrants differ in their perceptions of the signal correlation structure in the stage game. Every firm learns about an aspect of the environment (price elasticity) through the lens of its model. Firms that believe in different correlations interpret the same observation differently when inferring price elasticity, as they make different estimates about rival firms' production based on their own demand signal.

Society consists of the adherents of multiple competing models who match up to play the stage game every period. We introduce the concept of a *zeitgeist* to capture the social interaction structure — the sizes of the subpopulations with different models and the matchmaking technology that pairs up opponents to play the game. Agents can identify which subpopulation their opponent is from, and (correctly) know that the game they play is orthogonal to the type of opponent.[1] Our framework assumes that the agents might face one of several possible games and therefore richer models can in principle help as they allow agents to adapt their behavior more. Conditional on the stage game, in equilibrium each agent forms a Bayesian belief about the game using data from all of her interactions, playing a subjective best response against every type of opponent given this belief.

We define the *evolutionary stability* of model A against model B based on whether model A has a weakly higher average equilibrium payoff than model B when the population share of model A is close to 1, with the average taken over the different stage games. This criterion is familiar from past work following what is known as the *indirect evolutionary approach*. Under this approach, evolution acts on some trait that determines best responses, as opposed to actions. We emphasize that our stability concepts reduce to standard notions under this approach when inference is absent. Rather, our contribution is to apply it to the selection of *models* that contain multiple feasible beliefs about the environment.

1.3 Implications of the Learning Channel: Tailored Commitments and Polymorphism

The ability to draw inferences within a model (as opposed to committing to a fixed belief) is necessary for misspecifications to defeat rationality in some contexts. In Section 4.1, we characterize environments where the correctly specified model is only evolutionarily fragile against invading models that allow for inferences. Our argument constructs an optimal misspecified model for invading a rational society. This misspecification resembles an "illusion of control" bias, where agents think the outcomes they get in a game only depend on their own strategy and not on the opponent's strategy. The model has the property that its

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

adherents end up adopting the optimal commitment against a correctly specified opponent game-by-game. Misinference thus becomes a channel to tailor commitments to the true game. The correctly specified model is evolutionarily fragile against this misspecified model with uniform matching unless the former already gets the Stackelberg payoff in every game.

More generally, we show that misspecified models exhibit different stability properties than wrong dogmatic beliefs in our framework. Our next two results highlight the aforementioned polymorphism of misspecified models—that they can appear weak against rational incumbents in one environment and yet grow stronger and successfully invade the rational society in another environment. The reason is that due to the learning channel, an adherent of a misspecified model may come to hold different beliefs about parameters of the underlying stage game, and thus (endogenously) adopt different best-reply functions when facing outcomes generated from different strategy profiles. Thus, changes in the population structure and matching process influence the perceived best replies for adherents of misspecified models.

Polymorphism enables a new stability phenomenon that we call *stability reversals.* Two models exhibit stability reversal if:

1. 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

2. Whenever model B is dominant, its adherents strictly outperform model A's adherents on average

In the absence of inference, condition (1) would imply that A outperforms B regardless of the two subpopulations' sizes. But this no longer holds when inference is possible. The reason is that the adherents of model B might make an evolutionarily advantageous inference only when they are matched up with each other sufficiently often. Thus, even if condition (1) held, model B might still drive out model A if model B adherents reach some critical mass.

Polymorphism also manifests in a non-monotonicity of stability with respect to matching assortativity. As discussed in Alger and Weibull (2013), the assortativity parameter can represent the degree of homophily in a society or the frequency of interaction with kin. Various versions of the idea that high assortativity selects for cooperative agents and low assortativity selects for competitive ones date back to at least Hamilton (1964a,b). But this simple dichotomous perspective becomes complicated with misspecifications. The reason is that in our framework, the preferences agents seek to maximize are *endogenously determined in equilibrium*, due to the learning channel. Because the adherents of a misspecified model can draw different misinferences about a fixed game's parameters when facing data generated by different opponent actions, one model may be favored over another only at *intermediate* levels of assortativities, but not favored at either very low or very high levels. Thus, a particular bias might only survive in moderately homophilous societies — a novel empirical implication of misspecified inference.

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. We consider a separate notion, *equilibrium zeitgeist with strategic uncertainty*, in Appendix C, when we allow agents to draw inferences about others' strategies in addition to learning about the fundamentals. Online Appendix OA 4 provides a combined learning foundation for both equilibrium concepts, but in the main text we primarily focus on the steady-state characterization. Section 2.6.1 sketches the framework without inference, which has been studied in past work.

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 i's consequence y_i determines her utility, according to a common utility function $\pi : \mathbb{Y} \to \mathbb{R}$. 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 $|\mathcal{G}| = 1$.

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 Appendix C for one application where agents play an asymmetric game).

2.2 Models and Parameters

Throughout this paper, we will take the strategy space $\mathbb{A}$, the set of consequences $\mathbb{Y}$, and the utility function over consequences π 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 $(a_i, a_{-i}) \mapsto F^\bullet(a_i, a_{-i}, G)$.

We focus on the case where society consists of two observably distinguishable groups of agents, A and B, who may behave differently in the stage game due to different beliefs about how y is generated. The two groups of agents entertain different *models* of the world that help resolve their fundamental uncertainty. A model Θ is a collection of data-generating processes $F : \mathbb{A}^2 \to \Delta(\mathbb{Y})$ about how strategy profiles translate into consequences for the agent, with different processes corresponding to different *parameters* of the model. Each F has associated with it a density or probability mass function $f(a_i, a_{-i}) : \mathbb{Y} \to \mathbb{R}_+$ for every $(a_i, a_{-i}) \in \mathbb{A}^2$. We thus view each model as a subset of $(\Delta(\mathbb{Y}))^{\mathbb{A}^2}$ and we assume it is metrizable.

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 as the *adherents* of that model. Each agent dogmatically believes that in every situation $G \in \mathcal{G}$, one of the parameters of her model accurately represents the stage game. We call $\Theta = \{F^\bullet(\cdot, \cdot, G) : G \in \mathcal{G}\}$ 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. In this case, the model is *misspecified*.

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 random opponent in every period and uses her personal experience in these matches to calibrate the most accurate parameter within her model. A zeitgeist describes the corresponding landscape.

Definition 1. Fix models Θ_A and Θ_B. A *zeitgeist* $\mathfrak{Z} = (\mu_A(G), \mu_B(G), p, \lambda, a(G))_{G \in \mathcal{G}}$ 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) a matching assortativity parameter $\lambda \in [0, 1]$; (4)

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 beliefs of each group. Parts (2) and (3) determine social composition and social interaction—the relative prominence of each model and the probability of interacting with one's own group versus with the overall population. In each period, λ is the probability an agent's opponent is from her own group, and $1 - \lambda$ is the probability the opponent is drawn uniformly from the population. Therefore, an agent from group g has probability $\lambda + (1 - \lambda)p_g$ of being matched with an opponent from her own group, and a complementary chance of being matched with an opponent from the other group. Part (4) 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.

2.4 Equilibrium Zeitgeists

A model's fitness corresponds to the equilibrium payoffs of its adherents. An equilibrium zeitgeist (EZ) imposes optimality conditions on inference and behavior in a zeitgeist. Optimality of behavior requires each player to best respond given her beliefs, and optimality of inference requires that the support of each player's belief only contains the "best-fitting" parameter from her model in the sense of minimizing Kullback-Leibler (KL) divergence.

We now formalize this criterion. For two distributions over consequences, $\Phi, \Psi \in \Delta(\mathbb{Y})$ with density or probability mass functions ψ, ϕ, define the KL divergence from Ψ to Φ as $D_{KL}(\Phi \parallel \Psi) := \int \phi(y) \ln\left(\frac{\phi(y)}{\psi(y)}\right) dy$. Recall that every data-generating process F, like the true fundamental $F^\bullet(\cdot, \cdot, G)$, outputs a distribution over consequences for every profile of own play and opponent's play, $(a_i, a_{-i}) \in \mathbb{A}^2$. For data-generating process F, let $K(F; a_i, a_{-i}, G) := D_{KL}(F^\bullet(a_i, a_{-i}, G) \parallel F(a_i, a_{-i}))$ be the KL divergence from the expected distribution $F(a_i, a_{-i})$ to the objective distribution $F^\bullet(a_i, a_{-i}, G)$ under the play (a_i, a_{-i}) and situation G. For a distribution μ over parameters, let $U_i(a_i, a_{-i}; \mu)$ represent i's subjective expected utility under the belief that the true parameter is drawn according to μ. That is, $U_i(a_i, a_{-i}; \mu) := \mathbb{E}_{F\sim\mu}(\mathbb{E}_{y\sim F(a_i,a_{-i})}[\pi(y)])$.

Definition 2. A zeitgeist $\mathfrak{Z} = (\mu_A(G), \mu_B(G), p, \lambda, a(G))_{G \in \mathcal{G}}$ is an *equilibrium zeitgeist (EZ)* if, for every $G \in \mathcal{G}$ and $g, g' \in \{A, B\}$, $a_{g,g'}(G) \in \underset{a_i \in \mathbb{A}}{\arg\max}\ U_i(a_i, a_{g',g}(G); \mu_g(G))$ and, for every $g \in \{A, B\}$, belief $\mu_g(G)$ is supported on

$$\underset{F \in \Theta_g}{\arg\min} \{(\lambda + (1-\lambda)p_g) \cdot K(F; a_{g,g}(G), a_{g,g}(G), G) + (1-\lambda)(1-p_g) \cdot K(F; a_{g,-g}(G), a_{-g,g}(G), G)\}$$

where $-g$ means the group other than g.

This definition requires agents from group g to choose a subjective best response against their opponents, given the belief μ_g about the fundamental uncertainty. No matter which group the agent is matched against, these choices are always made to selfishly maximize her (individual) subjective utility function. Each agent's belief μ_g is supported on the parameters in her 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, as in Esponda and Pouzo (2016). We note that here we assume inference occurs separately across situations. This reflects situation persistence, with agents having enough data to establish new beliefs and behavior if the situation were to change. Our learning foundation in Online Appendix OA 4 justifies this situation-by-situation updating, but we omit the details here as it otherwise plays no role in our results.

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, so that one model is more successful if the objective expected payoffs are higher. Given this criterion, 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 ϵ mass of agents who adhere to a *mutant model* Θ_B?

Evolutionary stability depends on the fitness of models Θ_A, Θ_B in EZs with $p_A = 1, p_B = 0$. But it is motivated by the invasion of a small but strictly positive population of model Θ_B adherents into an otherwise homogeneous society of model Θ_A adherents. Below, we directly analyze EZs with $p = (1, 0)$, but note that these EZs can be written as the limit of EZs where the population share of Θ_B is positive but approaching 0. Online Appendix OA 3 provides conditions for the existence of an EZ with $p = (1, 0)$ and to ensure that any limit of EZs with a positive but diminishing fraction of Θ_B remains an EZ with $p = (1, 0)$.

Definition 3. Say Θ_A is *evolutionarily stable [fragile]* against Θ_B under λ-matching if there exists at least one EZ with models Θ_A, Θ_B, $p = (1, 0)$, matching assortativity λ 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.[2] These two cases give sharp predictions about whether a small share of mutant-model invaders might grow in size, across all equilibrium selections. 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 mutant model may or may not grow in the society, depending on the equilibrium selection.

2.6 Discussion

Before using this framework to illustrate our main contributions—on tailored commitments and polymorphism, mentioned in the introduction—we clarify some important aspects of it.

2.6.1 Comparison to Other Evolutionary Frameworks

Our model applies the "indirect evolutionary approach" framework (see Robson and Samuelson (2011)) to settings where agents can draw inferences (especially misspecified inferences). We seek to build on this line of work. To appreciate the innovations in our framework, it may be instructive to consider the case without inference—i.e., where $|\Theta| = 1$ and $|\mathcal{G}| = 1$. Then:

- $F^\bullet$ and π define objective expected utilities following any action profile;
- Adherents are characterized by a subjective preference, which may differ from those implied by objective utilities;
- The environment pairs the population up, in a way that depends on an assortativity parameter and the population composition;

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

- Equilibrium requires players to maximize their subjective preferences against their opponent, given the opponent's strategy;
- Stability of preference A against preference B requires that A adherents obtain higher expected objective utility than B adherents, when dominant, in some equilibrium.

This setup is familiar from the literature on the selection or preferences—for instance, it encapsulates the main framework articulated in Alger and Weibull (2019).

Models are more general than preferences in that agents may adapt their beliefs (which determine their subjective preferences) endogenously. The reason we introduce *zeitgeists* is, relative to other evolutionary frameworks, ours requires beliefs about the data-generating process, μ, to be incorporated. The notion of an Equilibrium Zeitgeist tells us how to define the fitness of a specification when the resulting "preference" is part of the solution concept.

Allowing for multiple situations is the most direct way for inference itself to be beneficial. With only a single situation, any steady state outcome that emerges for some Θ can also emerge when $|\Theta| = 1$; the same is not true with multiple situations, as we will see. That said, one could also study settings with multiple situations without inference (see Güth and Napel (2006) for an example of such an exercise).

2.6.2 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.[3] If adherents were able to believe the fundamentals changed 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.

[3]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'.

We comment on some other modeling assumptions. First, 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. Observability assumptions are common in the literature on the indirect evolutionary approach; see Alger and Weibull (2019) and Dekel et al. (2007) for discussions. While there are a number of ways it can be relaxed, we expect the main insights to carry through given sufficient observability. In our context, one key assumption that makes our approach tractable is that players do not change their inferences in response to seeing their opponents' actions. In other words, players do not necessarily try to "read into" what others do when learning. This particular assumption seems plausible in many cases, as the inference problem on its own may be rather complex even before considering such higher-order inferences. Consider hedge funds that regularly trade against each other in a variety of settings. Funds hold differing philosophies, with some focusing on fundamental analysis and others on technical analysis.[4] But, simply observing another fund's actions would not lead a technical analyst to embrace efficient markets, or vice versa. Both fundamental analysis and technical analysis are complex forecasting systems that involve calibrating sophisticated models and take many years of training and experience to master. In settings such as these, agents need not know how others' models work even after identifying who they are.

Second, EZs as presented abstract away from the issues surrounding learning others' strategies. However, we study an extension in Appendix C allowing agents to be misspecified about others' strategies and hold wrong beliefs about these strategies in equilibrium.

Lastly, even as agents adjust their beliefs and behavior to achieve optimality, 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, Ely, and Yilankaya (2007)).

[4]In practice, each fund's model about the financial market is well known to other market participants, as it is always prominently marketed to their clients.

3 Higher-Order Misspecifications in LQN Games

We start by illustrating the relevance of the learning channel applied to a specific economically significant bias. This bias relates to how players perceive the correlation in private information in a strategic setting. We work with a class of linear quadratic normal (LQN) games. While prior work has exploited the tractability of this classic framework to derive comparative statics with respect to information (e.g., Bergemann and Morris (2013)), we innovate by accommodating both misspecifications and inference. In the main text we focus on a Cournot duopoly application, and extend the insights to general LQN games in Online Appendix OA 1.

We use this example to illustrate our first main contribution mentioned in the introduction: misinference can confer strategic benefits in cases where dogmatic beliefs do not. We show that a bias detrimental to welfare in the absence of learning may be beneficial in its presence, and that the welfare-improving bias depends on the social interaction structure. We also show that if there is uncertainty over which situation may emerge, the learning channel may be necessary to defeat rationality. We build on these results to deliver more general insights about the implications of inference on evolutionary stability in Section 4.

3.1 Stage Game and Misperceptions of Information Structure

We first describe the stage game. There is a demand state $\omega \sim \mathcal{N}(0, \sigma_\omega^2)$, where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2. Each player is a firm, with firm i receiving a private signal $s_i = \omega + \epsilon_i$, and then choosing $q_i \in \mathbb{R}$ (i.e., a quantity). The resulting market price is $P = \omega - r^\bullet \cdot \frac{1}{2}(q_1 + q_2) + \zeta$, where $\zeta \sim \mathcal{N}(0, (\sigma_\zeta^\bullet)^2)$ is an idiosyncratic independent price shock. Firm i's profit is $q_i P - \frac{1}{2}q_i^2$.

The stage game is parametrized by $\sigma_\omega^2, r^\bullet, (\sigma_\zeta^\bullet)^2 > 0$—i.e., variance in market demand, the elasticity of market price with respect to average quantity supplied, and the variance of price shocks, respectively. These parameters remain constant (so $|\mathcal{G}| = 1$). However, demand state ω, signals (s_i), and price shock ζ are redrawn independently across matches.

Note that market prices and quantity choices may be positive or negative. To interpret, when $P > 0$, the market pays for each unit of good supplied, and the market price decreases in total supply. When $P < 0$, the market pays for the disposal of the good. The cost $\frac{1}{2}q_i^2$ represents either a convex production cost or a convex disposal cost, depending on the sign of q_i.

We take the signals within a match to possibly be correlated conditional on ω, and study the perception of this correlation. Recalling that $s_i = \omega + \epsilon_i$, we assume in particular that $\epsilon_i = \frac{\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}} z + \frac{1-\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}} \eta_i$, where $\eta_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is the idiosyncratic component generated i.i.d. across players and $z \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is the common component. Higher κ leads to an information structure with higher conditional correlation. When $\kappa = 0$, s_i and s_{-i} are conditionally uncorrelated given ω. When $\kappa = 1$, we always have $s_i = s_{-i}$. This functional form for ϵ_i ensures $\text{Var}(s_i)$ is constant in κ, which facilitates tractability.

While objectively, $\kappa = \kappa^\bullet$, our interest will be in studying misspecifications in κ. Indeed, this particular bias is common in experiments, many of which show subjects often do not form accurate beliefs about the beliefs of others. We draw a connection between the misperception we study and such statistical biases:

Definition 4. Let $\tilde{\kappa}$ be a player's perceived κ. A player suffers from *correlation neglect* if $\tilde{\kappa} < \kappa^\bullet$. A player suffers from *projection bias* if $\tilde{\kappa} > \kappa^\bullet$.

Correlation neglect agents believe signals are less correlated relative to the truth, whereas projection bias agents "project" their own information onto others (exaggerating the similarity between others' signals and their own). We are agnostic about the origin of these misspecifications, e.g., cognitive biases or more complex mechanisms,[5] instead asking whether such misspecifications would persist under selection pressures were they to appear.

3.2 Formalizing Strategies and Models

This environment fits into the formalism from Section 2 as follows. A strategy is a function $Q_i : \mathbb{R} \to \mathbb{R}$ that assigns a quantity $Q_i(s_i)$ to every signal s_i, and a strategy is *linear* if $Q_i(s_i) = \alpha_i s_i$ for every $s_i \in \mathbb{R}$ and some $\alpha_i \geq 0$. Since the best response to any linear strategy is linear, regardless of the agent's belief about the correlation parameter and market price elasticity (Lemma 2 in Appendix A.1), we restrict attention to linear strategies and let $\mathbb{A} = [0, \bar{M}_\alpha]$ for $\bar{M}_\alpha < \infty$, with $\alpha_i \in \mathbb{A}$ referring to strategy $Q_i(s_i) = \alpha_i s_i$.

The stage game is common knowledge except for $r^\bullet, \kappa^\bullet$, and $\sigma_\zeta^\bullet$. Models are dogmatic and possibly wrong about κ, but allow inferences about r and σ_ζ. We set the consequence space

[5]For example, Hansen, Misra, and Pai (2021) show that multiple agents simultaneously conducting algorithmic price experiments in the same market may generate correlated information misinterpreted as independent information, a form of correlation neglect for firms. Goldfarb and Xiao (2019) structurally estimate a model of thinking cost and find that bar owners over-extrapolate the effect of today's weather shock on future profitability.

for agent i to be $\mathbb{Y} = \mathbb{R}^3$, where $y = (s_i, q_i, P) \in \mathbb{Y}$. Consequence y delivers utility $\pi(y) := q_i P - \frac{1}{2}q_i^2$. Since κ indexes models, we write $\Theta(\kappa) := \{F_{r,\kappa,\sigma_\zeta} : r \in [0, \bar{M}_r], \sigma_\zeta \in [0, \bar{M}_{\sigma_\zeta}]\}$ for some $\bar{M}_r, \bar{M}_{\sigma_\zeta} < \infty$. So, $\Theta(\kappa)$ is a set of parameters that reflect a dogmatic belief in the correlation parameter κ. Each $F_{r,\kappa,\sigma_\zeta} : \mathbb{A} \times \mathbb{A} \to \Delta(\mathbb{Y})$ is such that $F_{r,\kappa,\sigma_\zeta}(\alpha_i, \alpha_{-i})$ gives the distribution over i's consequences in a stage game with parameters $(r, \kappa, \sigma_\zeta)$, when i uses the linear strategy α_i against an opponent using linear strategy α_{-i}. While agents learn about both r and σ_ζ, (mis)inferences about r drives the main results.[6]

We assumed that the space of feasible linear strategies $\alpha_i \in [0, \bar{M}_\alpha]$ and the domain of inference over game parameters $r \in [0, \bar{M}_r], \sigma_\zeta \in [0, \bar{M}_{\sigma_\zeta}]$ are compact, to guarantee EZ existence. For some of our results, we utilize the following shorthand:

Notation 1. A result is said to hold "*with high enough price volatility and large enough strategy space and inference space*" if, whenever the strategy space $[0, \bar{M}_\alpha]$ has $\bar{M}_\alpha \geq \frac{1/\sigma_\epsilon^2}{1/\sigma_\epsilon^2 + 1/\sigma_\omega^2}$, there exist $0 < L_1, L_2, L_3 < \infty$ so that for any objective game $F^\bullet$ with $(\sigma_\zeta^\bullet)^2 \geq L_1$ and with models where $r \in [0, \bar{M}_r]$, $\sigma_\zeta \in [0, \bar{M}_{\sigma_\zeta}]$ are such that $\bar{M}_{\sigma_\zeta}^2 \geq (\sigma_\zeta^\bullet)^2 + L_2$ and $\bar{M}_r \geq L_3$, the result is true.

When imposed, these assumptions will ensure behavior and beliefs are interior. Our analysis relies on a number of technical lemmas, which we defer to Appendix A. We show, for example, that the set of EZs is non-empty and it is upper hemicontinuous in population sizes. We also derive closed-form expressions for the best-fitting inference and optimal behavior of misspecified agents.

3.3 The Impact of Misspecification: Some Intuition

Before presenting our results on the fragility of correct specifications, we briefly describe what happens when players entertain a dogmatically misspecified view of κ.

Most importantly, an agent's inference about r is strictly decreasing in her belief about the correlation parameter κ. To understand why, assume player i uses the linear strategy α_i and player $-i$ uses the linear strategy α_{-i}. After receiving a private signal s_i, player i expects to face a price distribution with a mean that is linearly increasing in $\mathbb{E}[s_{-i} \mid s_i]$, which in turn

[6]Since each firm's profit is linear in the market price, belief about the variance of the idiosyncratic price shock does not change her expected payoffs or behavior. The parameter σ_ζ absorbs changes in the variance of market price, creating significant tractability. To infer r, it is only necessary to consider the mean of the market price in the data, not its variance.

is linearly increasing in s_i (see Appendix A.1 for more details).[7] Now, under projection bias $\kappa > \kappa^\bullet$, $\mathbb{E}_\kappa[s_{-i} \mid s_i]$ is excessively steep in s_i, since the correlation is higher. For example, following a large and positive s_i, the agent overestimates the similarity of $-i$'s signal and wrongly predicts that $-i$ must also choose a very high quantity, and thus becomes surprised when the market price remains high. As a result, the agent then wrongly infers that the market price elasticity must be low. Therefore, in order to rationalize the average market price conditional on own signal, an agent with projection bias must infer $r < r^\bullet$. For similar reasons, an agent with correlation neglect infers $r > r^\bullet$.

The fact that projection-biased players believe the elasticity of demand is lower than the truth suggests that they will behave more aggressively than correctly specified players—with the converse holding for correlation neglect players. Intuitively, if price reacts less to quantity, then players should price more aggressively. While this does turn out to be true—and drives many of the results below—things are more subtle because increasing κ has an *a priori* ambiguous impact on the agent's equilibrium aggressiveness. In fact, in our characterization, we show that increasing κ but holding fixed the player's belief about price elasticity has the direct effect of *lowering* aggression. The results below show that the indirect effect through the learning channel dominates, and the evolutionary stability of correlational errors is driven by this channel. We show in Section 3.5 that the conclusions are reversed when we shut down the learning channel.

3.4 Selecting Biases under Uniform and Assortative Matching

We now consider the evolutionary instability of correctly specified beliefs about the information structure. We first take $\lambda = 0$; we note that this case requires some technical innovation to characterize the *asymmetric* equilibrium strategy profile in matches between the correctly specified residents and the projection-biased mutants.

Proposition 1 (Uniform Matching Selects Projection Bias)**.** *Let $r^\bullet > 0$, $\kappa^\bullet \in [0,1]$ be given. With high enough price volatility and large enough strategy space and inference space, there exist $\underline{\kappa} < \kappa^\bullet < \bar{\kappa}$ so that taking $(\Theta_A, \Theta_B) = (\Theta(\kappa^\bullet), \Theta(\kappa))$ for $\kappa \in [\underline{\kappa}, \bar{\kappa}]$, there is a unique EZ with uniform matching ($\lambda = 0$) and $(p_A, p_B) = (1, 0)$. The equilibrium fitness of $\Theta(\kappa)$ is strictly higher than that of $\Theta(\kappa^\bullet)$ if $\kappa > \kappa^\bullet$, and strictly lower if $\kappa < \kappa^\bullet$.*

[7] Specifically, Lemma 1 in Appendix A.1 shows there exists a strictly increasing and strictly positive function $\psi(\kappa)$ so that $\mathbb{E}_\kappa[s_{-i} \mid s_i] = \psi(\kappa) \cdot s_i$ for all $s_i \in \mathbb{R}, \kappa \in [0,1]$.

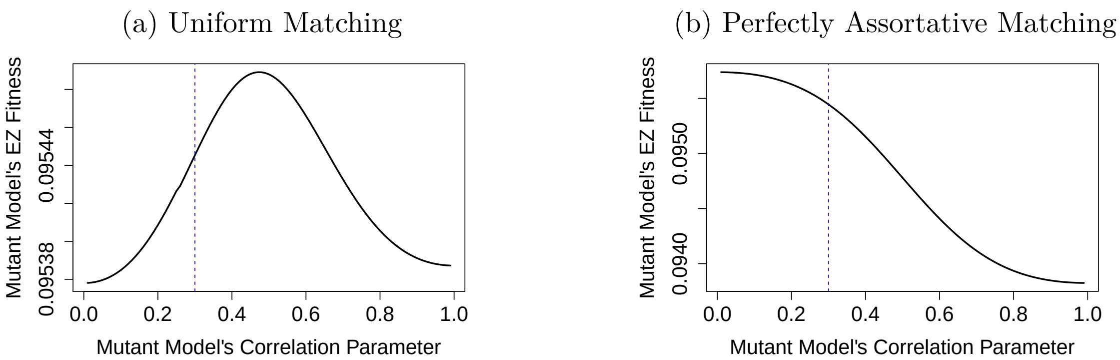

Figure 1: Fitness of mutant model against a correctly specified resident, as a function of κ.
Notes: The left panel assumes uniform matching (i.e., $\lambda = 0$) and the right panel assumes assortative matching (i.e., $\lambda = 1$). Both examples take $\kappa^\bullet = 0.3$, $r^\bullet = 1$, $\sigma_\omega^2 = \sigma_\epsilon^2 = 1$.

Figure 1a illustrates how, around $\kappa^\bullet$, mutant payoffs increase in κ. But misperception only helps to a point—the correct specification becomes evolutionarily stable for large enough κ.

The intuition for this result follows from the intuition outlined in Section 3.3— projection bias generates a commitment to aggression as it leads the biased agents to under-infer market price elasticity. It is well known that in Cournot oligopoly games, such commitment can be beneficial. For instance, if quantities are chosen sequentially, the first mover obtains a higher payoff compared to the case where quantities are chosen simultaneously. A similar force is at work here, but the source of the commitment is different. Misspecification about signal correlation leads to misinference about $r^\bullet$, which causes the mutants to credibly respond to their opponents' play in an overly aggressive manner. The rational residents, who can identify the mutants in the population, back down and yield a larger share of the surplus. While projection bias is beneficial in small measure, it is also intuitive that excessive aggression would be detrimental as well, as overproduction can be individually suboptimal.

By contrast, perfectly assortative matching favors biases that lead to more *cooperative* behavior, and thus the commitment to aggression is detrimental to fitness. Correspondingly, we obtain the opposite result: evolutionary stability selects correlation neglect.

Proposition 2 (Perfectly Assortative Matching Selects Correlation Neglect)**.** *Let $r^\bullet > 0$, $\kappa^\bullet \in [0,1]$ be given. With high enough price volatility and large enough strategy space and inference space, taking $(\Theta_A, \Theta_B) = (\Theta(\kappa_A), \Theta(\kappa_B))$ where $\kappa_A \leq \kappa_B$, the fitness of Θ_A is weakly higher than that of Θ_B in every EZ with any population proportion p and perfectly assortative matching ($\lambda = 1$).*

Correlation neglect leads agents to over-infer market price elasticity, enabling commitment to more cooperative behavior (i.e., linear strategies with a smaller coefficient α_i). Rational opponents would take advantage of such agents, but biased agents never match up against rational opponents in a society with perfectly assortative matching. The contrast with uniform matching is illustrated in Figure 1b—when $\lambda = 1$, the misspecified agents' payoffs are *decreasing* in κ around the true $\kappa^\bullet$.

In fact, the fragility of the correct specification is even starker when $\lambda = 1$ compared to $\lambda = 0$. Proposition 2 implies that mutant fitness is not only locally decreasing in κ around $\kappa^\bullet$, but monotonic *for all* κ (whereas Figure 1a illustrated the possibility of non-monotonicity of fitness in κ). Indeed, letting α^{TEAM} denote the symmetric linear strategy profile that maximizes the sum of the two firms' expected objective payoffs, we show that among symmetric strategy profiles, players' payoffs strictly decrease in their aggressiveness in the region $\alpha > \alpha^{TEAM}$. We also show that with $\lambda = 1$ and any $\kappa \in [0, 1]$, the equilibrium play among two adherents of $\Theta(\kappa)$ strictly increases in aggression as κ grows, always being strictly more aggressive than α^{TEAM}. Lowering the perception of κ confers an evolutionary advantage by bringing play monotonically closer to α^{TEAM} in equilibrium.

3.5 The Necessity of the (Mis)Learning Channel

In the previous sections, the misinference over r allows agents to commit to behavior that increases their equilibrium payoffs against their typical opponents. We establish two results to emphasize that the statistical biases may not be beneficial on their own, but only become beneficial due to the learning channel. First, assuming a single situation (as we have been working with so far), we show that if players were instead dogmatically correct about $r = r^\bullet$, then the predictions in Propositions 1 and 2 can be reversed:

Proposition 3. *Let $r^\bullet > 0$, $\kappa^\bullet \in [0, 1]$ be given. With high enough price volatility and large enough strategy space and inference space, there exists $\epsilon > 0$ so that for any $\kappa_l, \kappa_h \in [0, 1]$, $\kappa_l < \kappa^\bullet < \kappa_h \leq \kappa^\bullet + \epsilon$, the correctly specified model $\Theta(\kappa^\bullet)$ is evolutionarily stable against the singleton model $\{F_{r^\bullet, \kappa_h, \sigma_\zeta^\bullet}\}$ under uniform matching ($\lambda = 0$), and evolutionarily stable against the singleton model $\{F_{r^\bullet, \kappa_l, \sigma_\zeta^\bullet}\}$ under perfectly assortative matching ($\lambda = 1$).*

Using dogmatic beliefs over r to shut down the learning channel, misperceptions about κ that used to confer an evolutionary advantage for a $\lambda \in \{0, 1\}$ can no longer invade a society of correctly specified residents. Intuitively, this is because *an error about κ has the direct effect of lowering welfare*, but also causes mislearning about r and hence a stronger, indirect effect of increasing welfare. In the case of uniform matching, for instance, the direct effect of an increase in the perceived correlation κ is for players to use less aggressive strategies, anticipating that any favorable signal about market demand is also shared by the opponent.

For our second result on the necessity of mislearning, suppose that the environment features multiple situations given by multiple feasible values of $r^\bullet$. Mistaken agents who do not learn have a fixed belief about r that cannot be beneficial in all situations (i.e., for all values of $r^\bullet$), and so they do not end up with higher fitness than rational agents. But, misspecified agents who can make different inferences about price elasticity in different situations can invade a rational society. They outperform the correctly specified model in every situation, which is impossible for any fully dogmatic model.

Proposition 4. *For every $\bar{r} \geq 3$, there exists a $q \in \Delta([0, \bar{r}])$ such that the correctly specified model is evolutionarily stable against any singleton model with a fixed (r, κ) when $r^\bullet \sim q$. On the other hand, for every $\bar{r} > 0$, there exists a projection bias model with $\kappa > \kappa^\bullet$ so that the corrected specified model is evolutionarily fragile against it for any $\rho \in \Delta([0, \bar{r}])$.*

4 Learning Channel and New Stability Phenomena

We now illustrate some stability phenomena that distinguish misspecified learning from dogmatic beliefs in our framework. These phenomena underscore our two main contributions mentioned in the introduction. The main novelty of our framework relative to past work on the indirect evolutionary approach is that agents maximize endogenously determined subjective preferences, not exogenously fixed ones. The *learning channel* refers to this endogenous preference formation, and we showcase some of its unique implications in this section, toward making the aforementioned contributions.

The learning channel adds new ways for biased individuals to develop strategic commitments in games. First, unlike agents with fixed subjective preferences, misspecified learners can develop situation-specific commitments that are better tailored to the stage game. We show this mechanism expands the scope of invading rational societies. Second, misspecified learners can exhibit polymorphism as they form different beliefs in different environments. 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.

The idea that agents' personal experiences (and more broadly, the environments that generate these experiences) shape their preferences *beyond* their individual characteristics is empirically well documented. For instance, recent work studying attitudes toward immigrants (Bursztyn et al. (2022)) or attitudes among immigrants (Bolotnyy et al. (2022)) find that variation in a person's environment—plausibly independent from individual characteristics—can considerably influence their political behavior and preferences. In an experiment with Indian men, Lowe (2021) finds that favoritism for one's own caste changes in response to cross-caste contacts, in a way that depends on whether interactions are competitive or cooperative. Our framework derives the implications of these kinds of preference-formation mechanisms on the stability of misspecified models.

4.1 When does a Distribution over Situations Exist such that Learning is Necessary to Defeat Rationality?

Our first result characterizes when a misspecified model can *only* invade a rational society when inference is possible, using tailored commitments. The following example illustrates:

Example 1. Suppose there are two situations, G_A and G_B, which are equally likely, and consequences $\mathbb{Y} = \{g, b\}$, with $u(g) = 1$ and $u(b) = 0$. Suppose that the probability a given player obtains g given an action 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

Taking $\lambda = 0$, we show the correctly specified model is not evolutionarily fragile against any singleton mutant model $\Theta = \{F\}$. Indeed, the minimal correctly specified model obtains objective fitness 0.35 if (a_2, a_2) in situation G_A and (a_3, a_3) in situation G_B are played, as these are Nash equilibria. But under the singleton model $\{F\}$, one of the three must hold:

- If a_3 is a best response to a_3 under F, there is an EZ where (a_3, a_3) is always the outcome, and the expected fitness is $0.3 < 0.35$

- If a_2 is a best response to a_3 under F, there is an EZ where (a_2, a_3) is played by the mutant and resident in G_B, so the mutant's payoff is at most $\frac{1}{2} \cdot 0.3 + \frac{1}{2} \cdot 0.14 < 0.35$
- If a_1 is a best response to a_3 under F, then there is an EZ where (a_1, a_3) is played by the mutant and resident in G_A, so the mutant's payoff is at most $\frac{1}{2} \cdot 0.1 + \frac{1}{2} \cdot 0.55 < 0.35$.

Thus, the minimal correctly specified model is not evolutionarily fragile against any singleton. However, consider the misspecified model $\Theta = \{F_A, F_B\}$, where both F_A and F_B depend only on one's own strategies and not the opponent's. Under F_A, a_1, a_2, and a_3 lead to consequence g with probabilities 0.1, 0.3, and 0.2 respectively. 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.

The resident minimal correctly specified model is evolutionarily fragile against this misspecified model. Note that the mutants never choose a_3, since this is dominated under both F_A and F_B. Next, note that mutants would play a_2 when believing F_A and a_1 when believing F_B. We show these mutants play a_2 in G_A and a_1 in G_B against the resident. Indeed, if mutants were to play a_1 in situation G_A, the correctly specified residents would best respond with a_3 in G_A. The mutants then learn F_A in G_A, and would then deviate to a_2. If mutants play a_2 in situation G_B, once again the residents best respond with a_3 in G_B, and the mutants learn F_B. But under F_B, the mutants believe they should deviate to a_1. These arguments rule out all other EZ behavior, so the mutants must play a_2 in G_A and a_1 in G_B. In this EZ, mutant fitness is $(1/2) \cdot 0.3 + (1/2) \cdot 0.5 = 0.4 > 0.35$, higher than the resident's fitness.

The previous example features two notable features: (1) A misspecification resembling an "illusion of control" whereby individuals believe consequences only depend on their own actions, and (2) Inferences leading to a belief that a desirable action is dominant, in each situation. Models of this form allow us to determine when the ability to draw misinferences strictly expands the scope for invasion against rationality. Intuitively, if mutants can adopt the optimal commitment situation-by-situation, then the learning channel allows the mutants to tailor their commitment. But a mutant with only one model (i.e., an exogenous subjective preference) lacks the flexibility to play differently in different situations.

Some notation is needed to state the general result. Consider an arbitrary situation G. We let $v_G^{\mathrm{NE}} \in \mathbb{R}$ be the highest symmetric Nash equilibrium payoff in G, when agents choose strategies from $\mathbb{A}$. For each $a_i \in \mathbb{A}$, we let $\underline{\mathrm{BR}}(a_i, G)$ 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.,

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

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, let v_G^b denote the worst equilibrium payoff of an agent with the subjective best-response correspondence b when she plays against a rational opponent in situation G.[8]

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 $\bar{a}_G$ in situation G leads to different consequences than playing the same strategy in situation $G' \neq G$, provided the opponent chooses the rational best response to the strategy. We can now state our result.

Theorem 1. *Suppose $\lambda = 0$, there are finitely many situations, and there is a symmetric Nash equilibrium in $\mathbb{A} \times \mathbb{A}$ for every situation G.*

1. *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 correctly specified model is not evolutionarily fragile against any singleton model.*

2. *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 correctly specified model is evolutionarily fragile against $\hat{\Theta}$ under any full-support distribution $q \in \Delta(\mathcal{G})$.*

[8]More formally, given correspondence $b : \mathbb{A} \rightrightarrows \mathbb{A}$, let $v_G^b \in \mathbb{R}$ be defined as 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. If no such profile exists, let $v_G^b = -\infty$.

The core of the proof uses a separating hyperplanes argument to determine a distribution q under which the rational model cannot be invaded. One can check that indeed Example 1 satisfies both conditions of Theorem 1. Whenever the conditions are satisfied, there is some distribution over situations so that the minimal correctly specified model is evolutionarily fragile against some mutant model, but not evolutionarily fragile against any *singleton* mutant model. In these environments, the ability to adapt preferences endogenously to the relevant situation (i.e., the learning channel) is a necessary condition for an invading mutant to displace the rational incumbent. Hence, this result shows that mutants with misspecified models cannot in general be represented simply as mutants with fixed subjective best-response correspondences.

4.2 Stability Reversals

We now illustrate polymorphism and highlight one consequence of it: 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 $\lambda = 0$ and $(p_A, p_B) = (1, 0)$, 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 $\lambda = 0$ and $(p_A, p_B) = (0, 1)$, 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 is still stable against 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 the learning channel:

Proposition 5. *Suppose $|\mathcal{G}| = 1$. Two singleton models (i.e., two subjective preferences in the stage game) cannot exhibit stability reversal.*

Additionally, stability reversals cannot emerge in decision problems. We show this by introducing a class of games where 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 6. *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 6 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.

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 $a_i \cdot P - 1_{\{a_i=2\}} \cdot c$. 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). 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 adherents of Θ_B are polymorphic. 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 Online Appendix OA 2.1.) Due to Propositions 5 and 6, we conclude that this example is possible due to the non-trivial strategic interactions and Θ_B's inference about b (polymorphism through the learning channel).

Stability reversals provide a clear demonstration of polymorphism in models that permit inference. A mutant model may appear very weak when present in small proportions, doing worse than the incumbent model conditional on every type of opponent. Yet, if the population share of the mutant 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 incumbent model.

4.3 Non-Monotonic Stability in Matching Assortativity

Our last general result is also a consequence of the polymorphism of misspecified learners: a mutant model might successfully invade *only* when matching assortativity in the society is intermediate. This non-monotonicity arises because a misspecified agent can draw different inferences about the game's fundamentals depending on the relative frequency of in-group and out-group interactions, as these two groups of opponents choose different actions. The idea that social interaction structure shapes people's beliefs about the world has been empirically documented,[9] and our framework accommodates this mechanism and shows how it affects the stability of misspecified models.

We again assume there is only one situation, for simplicity. Note that without inference (i.e., in the setting of preference evolution), the fitness of a group is linear in matching

[9] For example, Bazzi et al. (2019) document how ethnic attachment in response to a resettlement policy in Indonesia has varying effects depending on whether a community is "fractionalized" (so that most interactions are not with one's own group members, i.e., λ is small) versus polarized (so that most interactions are with one's own group, i.e., λ is large).

assortativity. Thus, for singleton models, Θ_A being evolutionarily stable against Θ_B both when $\lambda = 0$ and when $\lambda = 1$ implies the same holds for all $\lambda \in (0, 1)$.

Proposition 7. *Suppose Θ_A, Θ_B are singleton models (i.e., subjective preferences in the stage game) and Θ_A is evolutionarily stable against Θ_B with λ-matching for both $\lambda = 0$ and $\lambda = 1$. Then, Θ_A is also evolutionarily stable against Θ_B with λ-matching for any $\lambda \in [0, 1]$.*

Crucially, inference leads to cases where the relevant "preference" changes depending on how frequently a model interacts with different types of opponents. This kind of polymorphism means a model's fitness may be *non-linear* in the matching probabilities. This phenomenon is a distinguishing feature of our framework and we show that the conclusion of Proposition 7 need not hold for models that allow for parameter inferences.

Consider a stage game where each player chooses an action from $\{a_1, a_2, a_3\}$. Every player then receives a random prize, $y \in \{g, b\}$, with utility values $\pi(g) = 1$ and $\pi(b) = 0$. The payoff matrix below displays the objective expected utilities associated with different action profiles, which also correspond to the probabilities that the row and column players receive the good prize g.

	a_1	a_2	a_3
a_1	0.25, 0.25	0.50, 0.20	0.70, 0.15
a_2	0.20, 0.50	0.40, 0.40	0.40, 0.20
a_3	0.15, 0.70	0.20, 0.40	0.20, 0.20

Let Θ_A be the correctly specified singleton model. The action a_1 is strictly dominant under the objective payoffs, so an adherent of Θ_A always plays a_1 in all matches. Let Θ_B be a misspecified model $\Theta_B = \{F_H, F_L\}$. Each model F_H, F_L stipulates that the prize g is generated according to the probabilities in the following table, where b and c are parameters that depend on the model. The model F_H has $(b, c) = (0.8, 0.2)$ and F_L has $(b, c) = (0.1, 0.4)$.

	a_1	a_2	a_3
a_1	0.10, 0.10	0.10, c	0.10, 0.15
a_2	c, 0.10	b, b	b, 0.20
a_3	0.15, 0.10	0.20, b	0.20, 0.20

The learning channel for the biased mutants leads the correctly specified model to have non-monotonic evolutionary stability in terms of matching assortativity.

Example 3. In this stage game, Θ_A is evolutionarily stable against Θ_B under λ-matching when $\lambda = 0$ and $\lambda = 1$, but it is also evolutionarily fragile under λ-matching when $\lambda \in (\lambda_l, \lambda_h)$, where $0 < \lambda_l < \lambda_h < 1$ are $\lambda_l = 0.25$, $\lambda_h \approx 0.56$.

Consider the match between two adherents of Θ_B. If they believe in F_H, they will play the action profile (a_2, a_2) and payoff profile $(0.4, 0.4)$, a Pareto improvement compared to the correctly specified outcome (a_1, a_1). The problem is that the data from play of (a_2, a_2) fit F_L better than F_H, since the objective 40% probability of getting prize g is closer to F_L's conjecture (10%) than F_H's conjecture (80%). A belief in F_H — and hence the profile (a_2, a_2) — cannot be sustained if the mutants only play each other. On the other hand, when an adherent of Θ_B plays a correctly specified Θ_A adherent, both F_H and F_L prescribe a best response of a_2 against the Θ_A adherent's play a_1. The data generated from the (a_2, a_1) profile lead biased agents to the parameter F_H that enables cooperative behavior within the mutant community. But, these matches against correctly specified opponents harm the mutant's welfare, as they only get an objective payoff of 0.2.

Therefore, the most advantageous interaction structure for the mutants is one where they can infer F_H using the data from matches against correctly specified opponents, then extrapolate this optimistic belief about b to coordinate on (a_2, a_2) in matches against fellow mutants. This requires the mutants to match with intermediate assortativity. Figure 2 depicts the equilibrium fitness of the mutant model Θ_B as a function of assortativity. While payoffs of Θ_B adherents increase in λ at first, eventually they drop when mutant-vs-mutant matches become sufficiently frequent that a belief in F_H can no longer be sustained. Note that a similar conclusion holds with fixed λ and varying population sizes: what ultimately matters is the probability with which Θ_B interacts with each model. Non-linearity of fitness in the population shares can emerge here as well, also a unique possibility due to inference.[10]

5 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 performance. Our two main contributions are to highlight that misinference in such a framework allows for (1) tailored commitments and (2) polymorphism. Most prior work

[10]See Appendix C.2 for a discussion of stability with intermediate population shares.

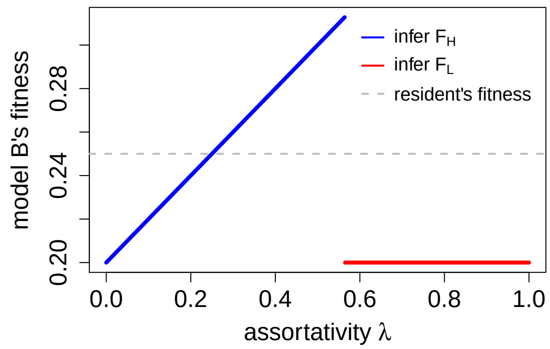

Figure 2: The EZ fitness of Θ_B for different values of λ when $p_B = 0$. (The EZ fitness of the resident model Θ_A is always 0.25.) In the blue region, adherents of Θ_B infer F_H and receive linearly increasing average payoffs across all matches as λ increases. In the red region, adherents of Θ_B infer F_L and receive a payoff of 0.2 in all matches, regardless of λ.

on misspecified Bayesian learning takes the misspecification as exogenous, studying the subsequent implications in both single-agent decision problems[11] and multi-agent games.[12] A number of papers establish general convergence properties of misspecified learning.[13] Our approach to endogenizing misspecified inference contrasts with those involving subjective expectations of payoffs[14] or goodness-of-fit tests.[15] To our knowledge, past work that *has* used objective payoffs to endogenize misspecified inference has restricted attention to financial markets (Sandroni, 2000; Massari, 2020).

This paper is closest to two independent and contemporaneous papers, Fudenberg and Lanzani (2022) and Frick, Iijima, and Ishii (2021), who consider welfare-based criteria for selecting among misspecifications in single-agent decision problems.[16] We differ in highlighting that the learning channel can *strictly* expand the possibility for misspecifications

[11]See Nyarko (1991); Fudenberg, Romanyuk, and Strack (2017); Heidhues, Koszegi, and Strack (2018); He (2022).

[12]See Bohren (2016); Bohren and Hauser (2021); Jehiel (2018); Molavi (2019); Dasaratha and He (2020); Ba and Gindin (2022); Frick, Iijima, and Ishii (2020); Murooka and Yamamoto (2021).

[13]See Esponda and Pouzo (2016); Esponda, Pouzo, and Yamamoto (2021); Frick, Iijima, and Ishii (2022); Fudenberg, Lanzani, and Strack (2021).

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

[15]See Cho and Kasa (2015, 2017); Ba (2022); Schwartzstein and Sunderam (2021); Lanzani (2022).

[16]Fudenberg and Lanzani (2022) study a framework where a continuum of agents with heterogeneous misspecifications arrive each period and learn from their predecessors' data. Frick, Iijima, and Ishii (2021) assign a *learning efficiency index* to every misspecified signal structure and conduct a robust comparison of welfare under different misspecifications.

to invade rational societies in strategic settings (relative to biased invaders who do not draw inferences), and we show that misspecifications can lead to different best responses in different environments and thus induce new stability phenomena.

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.

When agents entertain fundamental uncertainty about payoff parameters, our framework applies evolutionary forces to *sets of* preferences (i.e., models with multiple possible parameter values). This allows us to ask our central question: When does the ability to draw inference expand the scope for errors to invade rational societies? Developing a framework that accommodates inference is necessary to answer this question, providing the main point of departure from the literature on the indirect evolutionary approach. Our emphasis on *Bayesian* learning also distinguishes our work from papers that study the evolution of different belief-formation processes (Heller and Winter, 2020; Berman and Heller, 2022), who take a reduced-form (and possibly non-Bayesian) approach and consider arbitrary inference rules.

6 Concluding Discussion

We have introduced an evolutionary approach to predict the persistence of misspecified Bayesian learning. We have emphasized the implications and significance of the learning channel for evolutionary stability and the viability of biases. Our contributions are twofold. First, we show that the learning channel may confer strategic benefits in cases where dogmatic beliefs do not. This is because the learning channel enables flexible commitments that are tailored to the realized situation. Also, biases that are harmful under dogmatic beliefs may become beneficial when allowing for inferences. As highlighted in Section 3, the kinds of commitments generated under inference may even have the opposite direction relative to the commitments generated without it. Second, we show that misspecified agents are polymorphic. For this reason, the performance of a fixed bias may be difficult to extrapolate 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, 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 emergence of misspecified learning, and implications thereof. We have therefore focused more on the implications of the learning channel in an otherwise standard evolutionary setup.

References

Alger, I. and J. Weibull (2013): "Homo Moralis-Preference Evolution Under Incomplete Information and Assortative Matching," *Econometrica*, 81, 2269–2302.

——— (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.

Angeletos, G.-M. and A. Pavan (2007): "Efficient use of information and social value of information," *Econometrica*, 75, 1103–1142.

Ba, C. (2022): "Robust model misspecification and paradigm shift," *Working Paper*.

Ba, C. and A. Gindin (2022): "A multi-agent model of misspecified learning with overconfidence," *Working Paper*.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2019): "Unity in Diversity? How Intergroup Contact Can Foster Nation Building," *American Economic Review*, 109, 3978–4025.

Bergemann, D. and S. Morris (2013): "Robust predictions in games with incomplete information," *Econometrica*, 81, 1251–1308.

Berman, R. and Y. Heller (2022): "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.

Bolotnyy, V., M. Komisarchik, and B. Libgober (2022): "How Does Childhood Environment Shape Political Participation? Evidence from Refugees," *Working Paper*.

Bursztyn, L., T. Chaney, T. Hassan, and A. Rao (2022): "The Immigrant Next Door: Long-Term Contact, Generosity, and Prejudice," *Working Paper*.

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.

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.

——— (2021): "Welfare comparisons for biased learning," *Working Paper*.

——— (2022): "Belief convergence under misspecified learning: a martingale approach," *Review of Economic Studies, forthcoming*.

Fudenberg, D. and G. Lanzani (2022): "Which misperceptions persist?" *Theoretical Economics, forthcoming*.

Fudenberg, D., G. Lanzani, and P. Strack (2021): "Limit Points of Endogenous Misspecified Learning," *Econometrica*, 89, 1065–1098.

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*.

Goldfarb, A. and M. Xiao (2019): “Transitory shocks, limited attention, and a firm’s decision to exit,” *Working Paper.*

Güth, W. and S. Napel (2006): “Inequality Aversion in a Variety of Games - An Indirect Evolutionary Analysis,” *Economic Journal*, 116, 1037–1056.

Hamilton, W. D. (1964a): “The Genetical Evolution of Social Behaviour. 2,” *Journal of Theoretical Biology*, 7, 17–52.

——— (1964b): “The Genetical Evolution of Social Behaviour. I,” *Journal of Theoretical Biology*, 7, 1–16.

Hansen, K., K. Misra, and M. Pai (2021): “Frontiers: Algorithmic collusion: Supra-competitive prices via independent algorithms,” *Marketing Science*, 40, 1–12.

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

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.*

Lanzani, G. (2022): “Dynamic Concern for Misspecification,” *Working Paper.*

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

Lowe, M. (2021): “Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration,” *American Economic Review*, 111, 1807–44.

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

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

Murooka, T. and Y. Yamamoto (2021): "Multi-Player Bayesian Learning with Misspecified Models," *Working Paper*.

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

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

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

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 Additional Results for Section 3

A.1 Subjective Best Response and Misspecified Inference

In order to determine which models (i.e., perceptions of κ) are stable against rival models, we must characterize the relevant equilibrium zeitgeists. This section develops a number of preliminary results that relate beliefs about the game parameters to best responses, and conversely strategy profiles to the KL-divergence minimizing inferences. The proofs of these results appear in the Online Appendix OA 2.

We begin by proving the result alluded to in Section 3: every agent's inferences about the state and about opponent's signal are linear functions of her own signal. The linear coefficient on the latter increases with the correlation parameter κ.

Lemma 1. *There exists a strictly increasing function $\psi(\kappa)$, with $\psi(0) > 0$ and $\psi(1) = 1$, so that $\mathbb{E}_\kappa[s_{-i} \mid s_i] = \psi(\kappa) \cdot s_i$ for all $s_i \in \mathbb{R}$, $\kappa \in [0,1]$. Also, there exists a strictly positive $\gamma \in \mathbb{R}$ so that $\mathbb{E}_\kappa[\omega \mid s_i] = \gamma \cdot s_i$ for all $s_i \in \mathbb{R}$, $\kappa \in [0,1]$.*

Linearity of $\mathbb{E}[\omega \mid s_i]$ and $\mathbb{E}[s_{-i} \mid s_i]$ in s_i allows us explicitly characterize the corresponding linear best responses, given beliefs about κ and elasticity r. For Q_i, Q_{-i} (not necessarily linear) strategies in the stage game and $\mu \in \Delta(\Theta(\kappa))$, let $U_i(Q_i, Q_{-i}; \mu)$ be i's subjective expected utility from playing Q_i against Q_{-i}, under the belief μ.

Lemma 2. *For α_{-i} a linear strategy, $U_i(\alpha_i, \alpha_{-i}; \mu) = \mathbb{E}[s_i^2] \cdot \left(\alpha_i \gamma - \frac{1}{2}\hat{r}\alpha_i^2 - \frac{1}{2}\hat{r}\psi(\kappa)\alpha_i\alpha_{-i} - \frac{1}{2}\alpha_i^2\right)$ for every linear strategy α_i, where $\hat{r} = \int r \, d\mu(r, \kappa, \sigma_\zeta)$ is the mean of μ's marginal on elasticity. For $\kappa \in [0,1]$ and $r > 0$, $\alpha_i^{BR}(\alpha_{-i}; \kappa, r) := \frac{\gamma - \frac{1}{2}r\psi(\kappa)\alpha_{-i}}{1+r}$ best responds to α_{-i} among all (possibly non-linear) strategies $Q_i : \mathbb{R} \to \mathbb{R}$ for all $\sigma_\zeta > 0$.*

Lemma 2 shows that $\alpha_i^{BR}(\alpha_{-i}; \kappa, r)$ is not only the best-responding linear strategy when opponent plays α_{-i} and i believes in correlation parameter κ and elasticity r, it is also optimal among the class of all strategies $Q_i(s_i)$ against the same opponent play and under the same beliefs.

Call a linear strategy more *aggressive* if its coefficient $\alpha_i \geq 0$ is larger. One implication of Lemma 2 is that agent i's subjective best response function becomes more aggressive when i believes in lower κ or lower r. We have $\frac{\partial \alpha_i^{BR}}{\partial \kappa} < 0$ because the agent can better capitalize on her private information about market demand when her rival does not share the same information. We have $\frac{\partial \alpha_i^{BR}}{\partial r} < 0$ because the agent can be more aggressive when facing an inelastic market price.

We now turn to equilibrium inference about the market price elasticity $r^\bullet$. The following lemma shows that any linear strategy profile generates data whose KL-divergence can be minimized to 0 by a unique value of r. We also characterize how this inference about elasticity depends on the strategy profile and the agent's belief about the correlation parameter κ. As mentioned earlier, we focus on the case where the bounds on the inferences $r \in [0, \bar{M}_r]$, $\sigma_\zeta \in [0, \bar{M}_{\sigma_\zeta}]$ are sufficiently large to ensure that the KL-divergence minimization problem is well-behaved.

Lemma 3. *With high enough price volatility and large enough strategy space and inference space, for every $\alpha_i, \alpha_{-i} \in [0, \bar{M}_\alpha]$, we have $D_{KL}(F_{r^\bullet, \kappa^\bullet, \sigma_\zeta^\bullet}(\alpha_i, \alpha_{-i}) \parallel F_{\hat{r}, \kappa, \hat{\sigma}_\zeta}(\alpha_i, \alpha_{-i})) = 0$ for exactly one pair $\hat{r} \in [0, \bar{M}_r], \hat{\sigma}_\zeta \in [0, \bar{M}_{\sigma_\zeta}]$. This $\hat{r}$ is given by $r_i^{INF}(\alpha_i, \alpha_{-i}, ; \kappa^\bullet, \kappa, r^\bullet) := r^\bullet \frac{\alpha_i + \alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i + \alpha_{-i}\psi(\kappa)}$.*

Lemma 3 implies that an agent's inference about r is strictly decreasing in her belief about the correlation parameter κ. To understand why, assume player i uses the linear strategy α_i and player $-i$ uses the linear strategy α_{-i}. After receiving a private signal s_i, player i expects to face a price distribution with a mean of $\gamma s_i - r(\frac{1}{2}\alpha_i s_i + \frac{1}{2}\alpha_{-i}\mathbb{E}_\kappa[s_{-i} \mid s_i])$. Under projection bias $\kappa > \kappa^\bullet$, $\mathbb{E}_\kappa[s_{-i} \mid s_i]$ is excessively steep in s_i. For example, following a large and positive s_i, the agent overestimates the similarity of $-i$'s signal and wrongly predicts that $-i$ must also choose a very high quantity, and thus becomes surprised when market price remains high. The agent then wrongly infers that the market price elasticity must be low. Therefore, to rationalize the average market price conditional on its own signal, an agent with projection bias must infer $r < r^\bullet$. For similar reasons, an agent with correlation neglect infers $r > r^\bullet$.

Combining Lemma 2 and Lemma 3, we find that increasing κ has an *a priori* ambiguous impact on the agent's equilibrium aggressiveness. Increasing κ has the direct effect of lowering aggression (by Lemma 2), but it also causes the indirect effect of lowering inference about r (by Lemma 3) and therefore increases aggression (by Lemma 2).

Lemma 3 considers the problem of KL-divergence minimization when all of the data are generated from a single strategy profile, $(\alpha_{-i}, \alpha_{-i})$. It implies that if $\lambda \in \{0, 1\}$ and $(p_A, p_B) = (1, 0)$, that is matching is either perfectly uniform or perfectly assortative in a homogeneous society, then every agent can find a parameter to exactly fit her equilibrium data. This is because agents only match with opponents from one group in the EZ. The self-confirming property lends a great deal of tractability and allows us to provide sharp comparative statics and assess the stability of models.

With interior population shares, agents can observe consequences from matches against the adherents of both Θ_A and Θ_B. Thus, they must find a single set of parameters for the stage game that best fits all of their data, and even this best-fitting parameter will have positive KL divergence in equilibrium. The next lemma shows the LQN game satisfies the sufficient conditions from Online Appendix OA 3 (Assumptions OA2 through OA6) for the existence and upper hemicontinuity of EZs. So, the tractable analysis in homogeneous societies remains robust to the introduction of a small but non-zero share of a mutant model.

Lemma 4. *For every $r^\bullet, \sigma_\zeta^\bullet \geq 0$, $\lambda \in [0, 1]$, $\kappa^\bullet, \kappa \in [0, 1]$, $\bar{M}_\alpha, \bar{M}_{\sigma_\zeta}, \bar{M}_r < \infty$, the LQN with objective parameters $(r^\bullet, \kappa^\bullet, \sigma_\zeta^\bullet)$, strategy space $\mathbb{A} = [0, \bar{M}_\alpha]$ and models $\Theta(\kappa^\bullet), \Theta(\kappa)$ with parameter spaces $[0, \bar{M}_r]$, $[0, \bar{M}_{\sigma_\zeta}]$ satisfy Assumptions OA2, OA3, OA4, OA5, and OA6. Therefore, EZs in LQN are upper hemicontinuous in population sizes.*

B Proofs of Key Results from the Main Text

B.1 Proof of Theorem 1

Part 1: Let $\mathcal{V}$ be the convex hull of $\{(v_G^b)_{G\in\mathcal{G}} \mid b : \mathbb{A} \rightrightarrows \mathbb{A}\}$, and let $\mathcal{U} = \{(u_G)_{G\in\mathcal{G}} : u_G \leq v_G$ for all G 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 vector $q \in \mathbb{R}^{|\mathcal{G}|}$ with $q_G \neq 0$ for every G, so that $q \cdot v^{\text{NE}} > q \cdot u$ 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 $q \cdot v^{\text{NE}} \geq q \cdot u$ 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 $\tilde{q}(\varepsilon) \cdot v^{NE} > \tilde{q}(\varepsilon) \cdot u$ whenever ε is sufficiently small, since we have that this inequality holds in the limit.

Now consider any singleton model $\Theta = \{F\}$, 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 utility v_G^b to the first player. There is an EZ where the resident correctly specified agents get v_G^{NE} in situation G, and the mutants 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, the residents' fitness is $q \cdot v^{\text{NE}}$ while that of the mutants is $q \cdot v^b$, and the former is weakly 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.

Part 2: Suppose the hypotheses hold and let us construct the misspecified model $\hat{\Theta} = \{F_G : G \in \mathcal{G}\}$. To define the parameters F_G, first consider $\tilde{F}_G$ where $\tilde{F}_G(a_i, a_{-i}) :=$

$F^\bullet(a_i, \underline{\mathrm{BR}}(a_i, G), G)$ 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 distribution $F_G(a_i, a_{-i}) \in \Delta(\mathbb{Y})$ as a sufficiently small perturbation of the $\tilde{F}_G(a_i, a_{-i})$, 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 $\mathfrak{Z}$ with the correctly specified resident, $\hat{\Theta}$ as the mutant, $\lambda = 0$. By situation identifiability, in $\mathfrak{Z}$ the correctly specified residents must believe in the true $F^\bullet(\cdot, \cdot, G)$ in every situation G. The mutants 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. We show further that mutants must believe in F_G in situation G. 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. The mutants' 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. This contradicts the mutants believing in $F_{G'}$ in situation G as the parameter F_G generates smaller KL divergence. So the mutants get the Stackelberg payoff in each situation, which means they have higher fitness than the residents in every EZ 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: 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 mutants to believe in F_G and play $(\bar{a}_G, \underline{\mathrm{BR}}(\bar{a}_G, G))$ in matches against residents in situation G.

B.2 Proof of Proposition 2

Proof. We will show that in every EZ: (i) for each $g \in \{A, B\}$, μ_g puts probability 1 on $\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa_g)} r^\bullet$; (ii) for each $g \in \{A, B\}$, $\alpha_{gg} = \frac{\gamma}{1 + \frac{r^\bullet}{2}(1+\psi(\kappa^\bullet)) + \frac{r^\bullet}{2}(\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa_g)})}$; (iii) the equilibrium fitness of group A is weakly higher than that of group B if and only if $\kappa_A \leq \kappa_B$.

Choose L_1, L_2, L_3 as in Lemma 3, given $r^\bullet$ and $\bar{M}_\alpha$. In any EZ with behavior $(\alpha_{AA}, \alpha_{AB}, \alpha_{BA}, \alpha_{BB})$, since the adherents of each model matches with their own group with probability 1 under perfectly assortatively matching, we conclude that each of μ_g for $g \in \{A, B\}$ must put full weight on $r_i^{INF}(\alpha_{gg}, \alpha_{gg}; \kappa^\bullet, \kappa_g, r^\bullet) = \frac{\alpha_{gg}+\alpha_{gg}\psi(\kappa^\bullet)}{\alpha_{gg}+\alpha_{gg}\psi(\kappa_g)} r^\bullet = \frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa_g)} r^\bullet$, proving (i).

Given this belief, we must have $\alpha_{gg} = \frac{\gamma - \frac{1}{2}\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa_g)} r^\bullet \psi(\kappa_g)\alpha_{gg}}{1+\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa_g)} r^\bullet}$ by Lemma 2. Rearranging yields $\alpha_{gg} = \frac{\gamma}{1+\frac{r^\bullet}{2}(1+\psi(\kappa^\bullet))+\frac{r^\bullet}{2}(\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa)})}$, proving (ii).

From Lemma 2, the objective expected utility of each player when both play the strategy profile α_{symm} is $\mathbb{E}[s_i^2] \cdot \left(\alpha_{symm}\gamma - \frac{1}{2} r^\bullet \alpha_{symm}^2 - \frac{1}{2} r^\bullet \psi(\kappa^\bullet)\alpha_{symm}^2 - \frac{1}{2}\alpha_{symm}^2\right)$. This is a strictly concave quadratic function in α_{symm} that is 0 at $\alpha_{symm} = 0$. Therefore, it is strictly decreasing in α_{symm} for α_{symm} larger than the team solution α_{TEAM} that maximizes this expression, given by the first-order condition

$$\gamma - r^\bullet \alpha_{TEAM} - r^\bullet \psi(\kappa^\bullet)\alpha_{TEAM} - \alpha_{TEAM} = 0 \Rightarrow \alpha_{TEAM} = \frac{\gamma}{1 + r^\bullet + r^\bullet \psi(\kappa^\bullet)}.$$

For any value of $\kappa \in [0, 1]$, using the fact that $\psi(0) > 0$ and ψ is strictly increasing,

$$\frac{\gamma}{1 + \frac{r^\bullet}{2}(1 + \psi(\kappa^\bullet)) + \frac{r^\bullet}{2}(\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa)})} > \frac{\gamma}{1 + \frac{r^\bullet}{2}(1 + \psi(\kappa^\bullet)) + \frac{r^\bullet}{2}(1 + \psi(\kappa^\bullet))} = \alpha_{TEAM}.$$

Also, $\frac{\gamma}{1+\frac{r^\bullet}{2}(1+\psi(\kappa^\bullet))+\frac{r^\bullet}{2}(\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa)})}$ is a strictly increasing function in κ, since ψ is strictly increasing. We therefore conclude that each player's utility when they play $\frac{\gamma}{1+\frac{r^\bullet}{2}(1+\psi(\kappa^\bullet))+\frac{r^\bullet}{2}(\frac{1+\psi(\kappa^\bullet)}{1+\psi(\kappa)})}$ against each other is strictly decreasing in κ, proving (iii). □

B.3 Proof of Proposition 5

Proof. Let two singleton models Θ_A, Θ_B be given. By contradiction, suppose they exhibit stability reversal. Let $\mathfrak{Z} = (\mu_A, \mu_B, p = (0, 1), \lambda = 0, (a))$ be any EZ where Θ_B is resident. By the definition of EZ, $\mathfrak{Z}' = (\mu_A, \mu_B, p = (1, 0), \lambda = 0, (a))$ 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. □

B.4 Proof of Proposition 6

Proof. To show the first claim, by way of contradiction, suppose $\mathfrak{Z} = (\mu_A, \mu_B, p = (1, 0), \lambda = 0, (a_{AA}, a_{AB}, a_{BA}, a_{BB}))$ is an EZ, and $\tilde{\mathfrak{Z}} = (\mu_A, \mu_B, p = (0, 1), \lambda = 0, (\tilde{a}_{AA}, \tilde{a}_{AB}, \tilde{a}_{BA}, \tilde{a}_{BB}))$

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 $\tilde{\mathfrak{Z}}' = (\mu_A, \mu_B, p = (0,1), \lambda = 0, (\tilde{a}_{AA}, a_{AB}, a_{BA}, \tilde{a}_{BB}))$ 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 $\mathfrak{Z}' = (\mu_A, \mu_B, p = (1,0), \lambda = 0, (a_{AA}, a_{AB}, a_{BA}, \tilde{a}_{BB}))$ 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 $\mathfrak{Z} = (\mu_A, \mu_B, p = (0,1), \lambda = 0, (a_{AA}, a_{AB}, a_{BA}, a_{BB}))$ 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 $\mathfrak{Z}' = (\mu_A^{'}, \mu_B, p = (0,1), \lambda = 0, (a_{AA}, a_{AB}^{'}, a_{BB}, a_{BB}))$, 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. □

B.5 Proof of Proposition 7

Proof. Let $\lambda \in [0,1]$ be given and let $\mathfrak{Z} = (\mu_A, \mu_B, p = (1,0), \lambda, (a))$ be an EZ. Since Θ_A, Θ_B are singleton models, $\mathfrak{Z}_0 = (\mu_A, \mu_B, p = (1,0), \lambda = 0, (a))$ and $\mathfrak{Z}_1 = (\mu_A, \mu_B, p = (1,0), \lambda = 1, (a))$ are also EZs. Let $u_{g,g'}$ represent model Θ_g's conditional fitness against group g' in each of these three EZs. From the hypothesis of the proposition, $u_{A,A} \geq u_{B,A}$ and $u_{A,A} \geq u_{B,B}$. This means the fitness of Θ_A in $\mathfrak{Z}$, which is $u_{A,A}$, is weakly larger than the fitness of Θ_B in $\mathfrak{Z}$, which is $\lambda u_{B,B} + (1-\lambda) u_{B,A}$. This shows Θ_A has weakly higher fitness than Θ_B in every EZ with λ and $p = (1,0)$. Also, at least one such EZ exists with assortativity λ, for at least one EZ exists when $\lambda = 0$, and the same equilibrium belief and behavior also constitutes an EZ for any other assortativity. □

C Evolutionary Stability of Analogy Classes

Here we apply our framework 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 use our framework to endogenize them.

C.1 Relaxing the Observability of Strategies

To study analogy-based reasoning, we relax the assumption that people correctly know others' strategies in equilibrium. We introduce the concepts of extended parameters and extended models:

Definition A.1. An *extended parameter* is a triplet (a_A, a_B, F) with $a_A, a_B \in \mathbb{A}$ and $F : \mathbb{A}^2 \to \Delta(\mathbb{Y})$. An *extended model* $\overline{\Theta}$ is a collection of extended parameters: i.e., a subset of $\mathbb{A}^2 \times (\Delta(\mathbb{Y}))^{\mathbb{A}^2}$.

In addition to a conjecture F about how strategy profiles translate into consequences for the agent, extended models also contain conjectures about how group A and group B opponents will act. We assume the marginal of the extended model on $(\Delta(\mathbb{Y}))^{\mathbb{A}^2}$ is metrizable. As before, we also assume each F is given by a density or probability mass function $f(a_i, a_{-i}) : \mathbb{Y} \to \mathbb{R}_+$ for every $(a_i, a_{-i}) \in \mathbb{A}^2$. We say that an extended model $\overline{\Theta}$ is *correctly specified* if $\overline{\Theta} = \mathbb{A}^2 \times \{F^\bullet(\cdot, \cdot, G)\}$, so the agent can make unrestricted inferences about others' play and does not rule out the correct data-generating process $F^\bullet(\cdot, \cdot, G)$ for any situation G.

Defining zeitgeists for extended models is immediate, as we can simply replace "model" with "extended model" in Definition 1. The equilibrium notion, however, is subtly different:

[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.

Definition A.2. A zeitgeist with strategic uncertainty $\overline{\mathfrak{Z}} = (\overline{\Theta}_A, \overline{\Theta}_B, \mu_A(G), \mu_B(G), p, \lambda, a(G))_{G\in\mathcal{G}}$ is an *equilibrium zeitgeist with strategic uncertainty (EZ-SU)* if for every $G \in \mathcal{G}$ and $g, g^{'} \in \{A, B\}$, $a_{g,g'}(G) \in \arg\max_{\hat{a}\in\mathbb{A}} \mathbb{E}_{(a_A,a_B,F)\sim\mu_g(G)} \left[\mathbb{E}_{y\sim F(\hat{a},a_{g'})}(\pi(y))\right]$ and, for every $g \in \{A, B\}$, the belief $\mu_g(G)$ is supported on

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

where $-g$ means the group other than g.

The only difference with Definition 2 is that the KL divergence is now taken with respect to the *conjectured opponent's strategy*, part of the extended model. Conjectures now include others' play, in addition to stage game parameters.

C.2 Defining Stable Population Shares

In this Section, we will also be interested in stable population shares in a society that contains both rational and misspecified players. We briefly introduce the following solution concept.

Definition A.3. Given population share $p \in (0, 1)$ and an EZ (or EZ-SU), p is said to be a *stable population share* given the EZ (or EZ-SU) if both models have the same fitness.

Since EZ(-SU)s are defined with interior population shares, we can calculate the fitness of a model in terms of its adherents' objective expected payoff. Whereas Definition 3's stability notion reflects performance with $(p_A, p_B) = (1, 0)$, stability with interior population shares as in Definition A.3 correspond to both models being co-existing with equal fitness.

C.3 Centipede Games and Analogy-Based Reasoning

We now analyze analogy-based reasoning in the centipede game in Figure 3 (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, $1 \le k \le K$ 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 $g > 0$, but 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 $\frac{k}{2}g+\ell$.

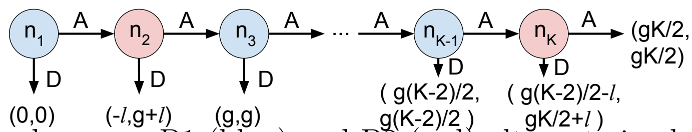

Figure 3: The centipede game. P1 (blue) and P2 (red) alternate in choosing Across (A) or Drop (D). Payoff profiles are shown at the terminal nodes.

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 $\mathbb{Y}=\{1,2\}\times(\{z_k:1\le k\le K\}\cup\{z_{end}\})$, 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 extended model is $\overline{\Theta}^\bullet:=\mathbb{A}\times\mathbb{A}\times\{F^\bullet\}$. The misspecified model of interest is $\overline{\Theta}^{An}:=\mathbb{A}^{An}\times\mathbb{A}^{An}\times\{F^\bullet\}$, 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 $\overline{\Theta}^{An}$ adherents. All agents, regardless of their model, actually have the strategy space $\mathbb{A}$.

C.4 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 A.1. *Suppose $K\ge4$ and $g>\frac{2}{K-2}\ell$. For any matching assortativity $\lambda\in[0,1]$, the correctly specified extended model $\overline{\Theta}^\bullet$ is evolutionarily stable with strategic uncertainty against itself, but it is not evolutionarily stable with strategic uncertainty against the misspecified extended model $\overline{\Theta}^{An}$. Also, $\overline{\Theta}^{An}$ is not evolutionarily stable against $\overline{\Theta}^\bullet$, unless $\lambda=1$.*

In contrast to the results from Section 3, whereby a misspecified inference over r was harmful for $\lambda = 1$ if and only if such an inference were helpful for $\lambda = 0$, in this environment the correctly specified extended model is not evolutionarily stable against a coarse reasoner for *any* level of assortativity. Here, the conditional fitness of $\overline{\Theta}^{An}$ against both $\overline{\Theta}^{\bullet}$ and $\overline{\Theta}^{An}$ can strictly improve on the correctly specified residents' equilibrium fitness. This is because the matches between two adherents of $\overline{\Theta}^{\bullet}$ must result in Dropping at the first move in equilibrium, while matches where at least one player is an adherent of $\overline{\Theta}^{An}$ 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, $\overline{\Theta}^{An}$ is not evolutionarily stable against $\overline{\Theta}^{\bullet}$ 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 mutant model in equilibrium. Thus we determine stable shares as defined in Section C.2, focusing on the EZ-SU where Across is played as often as possible.

We take $\lambda = 0$ throughout the remainder of this section. Suppose $K \geq 4$ and $g > \frac{2}{K-2}\ell$. Consider the *maximal continuation EZ-SU*: (1) misspecified agents always play Across except at node K where they choose Drop, and (2) correctly specified agents (i) matched with misspecified agents play Drop at nodes $K-1$ and K and Across otherwise, and (ii) matched with correctly specified agents always play Drop. We verify this indeed forms an EZ-SU.

Proposition A.2. Suppose $\lambda = 0$, $K \geq 4$ and $g > \frac{2}{K-2}\ell$. The two models have the same fitness in the maximal continuation EZ-SU of the centipede game if and only if $p_B^* = 1 - \frac{\ell}{g(K-2)}$, and thus 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 contintuation. 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 dollar game (Reny, 1993) in Figure 4, 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), and if the second player continues at the final node n^K, payoffs are $(K + 2, 0)$.

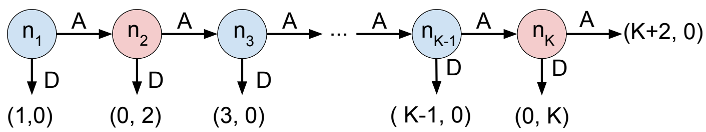

Figure 4: The dollar game. Players 1 (blue) and 2 (red) alternate in choosing Across (A) or Drop (D). Payoff profiles are shown at the terminal nodes.

Proposition A.3. For $\lambda = 0$ and every population size $(p, 1-p)$ with $p \in [0, 1]$, the maximal continuation EZ-SU is an EZ-SU where the fitness of $\overline{\Theta}^{\bullet}$ is strictly higher than that of $\overline{\Theta}^{An}$.

While maximal continuation remains an EZ-SU, the rational model strictly outperforms the misspecified model for all population shares. Provided the maximal continuation EZ-SU remains focal, we should thus expect no analogy reasoners in the long run with this stage game. Intuitively, the change in the payoffs means one player can only do better *at the expense* of the opponent. Since $\lambda = 0$, this implies the less cooperative strategy will be selected. But unlike Section 3, it is the correctly specified model that cannot be exploited.

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."[18] 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 become more prevalent with a longer horizon and with faster payoff growth.

[18] 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.

Online Appendix for "Evolutionarily Stable (Mis)specifications: Theory and Applications"

Kevin He and Jonathan Libgober

OA 1 More General LQN Games

We turn to general incomplete-information games and provide a condition for a model to be evolutionarily fragile against a "nearby" misspecified model. This condition shows how assortativity and the learning channel shape the evolutionary selection of models for a broader class of stage games and biases. We also relate the condition to the specific results studied so far in this application.

Consider a stage game where a state of the world ω is realized at the start of the game. Players 1 and 2 observe private signals $s_1, s_2 \in S \subseteq \mathbb{R}$, possibly correlated given ω. The objective distribution of (ω, s_1, s_2) is $\mathbb{P}^\bullet$. Based on their signals, players choose actions $q_1, q_2 \in \mathbb{R}$ and receive random consequences $y_1, y_2 \in \mathbb{Y}$. The distribution over consequences as a function of $(\omega, s_1, s_2, q_1, q_2)$ and the utility over consequences $\pi : \mathbb{Y} \to \mathbb{R}$ are such that each player i's objective expected utility from taking action q_i against opponent action q_{-i} in state ω is given by $u_i^\bullet(q_i, q_{-i}; \omega)$, differentiable in its first two arguments.

For an interval of real numbers $[\underline{\kappa}, \bar{\kappa}]$ with $\underline{\kappa} < \bar{\kappa}$ and $\kappa^\bullet \in (\underline{\kappa}, \bar{\kappa})$, suppose there is a family of models $(\Theta(\kappa))_{\kappa \in [\underline{\kappa}, \bar{\kappa}]}$. Fix $\lambda \in [0, 1]$ and a strategy space $\mathbb{A} \subseteq \mathbb{R}^S$, representing the feasible signal-contingent strategies. Suppose the two models in the society are $\Theta_A = \Theta(\kappa^\bullet)$ and $\Theta_B = \Theta(\kappa)$ for some $\kappa \in [\underline{\kappa}, \bar{\kappa}]$. The next assumption requires there to be a unique EZ with $(p_A, p_B) = (1, 0)$ in such societies with any $\kappa \in [\underline{\kappa}, \bar{\kappa}]$, and further requires the EZ to feature linear equilibria. Linear equilibria exist and are unique in a large class of games outside of the duopoly framework, and in particular in LQN games under some conditions on the payoff functions (see, e.g., Angeletos and Pavan (2007)).

Assumption OA1. *Suppose there is a unique EZ under λ-matching and population proportions $(p_A, p_B) = (1, 0)$ with $\Theta_A = \Theta(\kappa^\bullet)$, $\Theta_B = \Theta(\kappa)$ for every $\kappa \in [\underline{\kappa}, \bar{\kappa}]$. Suppose the κ-indexed EZ strategy profiles $(\sigma(\kappa)) = (\sigma_{AA}(\kappa), \sigma_{AB}(\kappa), \sigma_{BA}(\kappa), \sigma_{BB}(\kappa))$ are linear, i.e., $\sigma_{gg'}(\kappa)(s_i) = \alpha_{gg'}(\kappa) \cdot s_i$ with $\alpha_{gg'}(\kappa)$ differentiable in κ. Suppose that in the EZ with*

$\kappa = \kappa^\bullet$, $\alpha_{AA}(\kappa^\bullet)$ *is objectively interim-optimal against itself.*[19] *Finally, assume for every* κ*, Assumptions OA2, OA3, OA4, OA5, and OA6 are satisfied.*

Proposition OA1. *Let* $\alpha^\bullet := \alpha_{AA}(\kappa^\bullet)$*. Then, under Assumption OA1, if*

$$\mathbb{E}^\bullet\left[\mathbb{E}^\bullet\left[\frac{\partial u_1^\bullet}{\partial q_2}(\alpha^\bullet s_1, \alpha^\bullet s_2, \omega)\cdot[(1-\lambda)\alpha_{AB}^{'}(\kappa^\bullet)+\lambda\alpha_{BB}^{'}(\kappa^\bullet)]\cdot s_2 \mid s_1\right]\right] > 0,$$

then there exists some $\epsilon > 0$ *so that* $\Theta(\kappa^\bullet)$ *is evolutionarily fragile against models* $\Theta(\kappa)$ *with* $\kappa \in (\kappa^\bullet, \kappa^\bullet + \epsilon] \cap [\underline{\kappa}, \bar{\kappa}]$*. Also, if*

$$\mathbb{E}^\bullet\left[\mathbb{E}^\bullet\left[\frac{\partial u_1^\bullet}{\partial q_2}(\alpha^\bullet s_1, \alpha^\bullet s_2, \omega)\cdot[(1-\lambda)\alpha_{AB}^{'}(\kappa^\bullet)+\lambda\alpha_{BB}^{'}(\kappa^\bullet)]\cdot s_2 \mid s_1\right]\right] < 0,$$

then there exists some $\epsilon > 0$ *so that* $\Theta(\kappa^\bullet)$ *is evolutionarily fragile against models* $\Theta(\kappa)$ *with* $\kappa \in [\kappa^\bullet - \epsilon, \kappa^\bullet) \cap [\underline{\kappa}, \bar{\kappa}]$*. Here* $\mathbb{E}^\bullet$ *is the expectation with respect to the objective distribution of* (ω, s_1, s_2) *under* $\mathbb{P}^\bullet$*.*

Proposition OA1 describes a general condition to determine whether a correctly specified model is evolutionarily fragile against a nearby misspecified mutant model. The condition asks if a slight change in the mutant model's κ leads mutants' opponents to change their equilibrium actions such that the mutants become better off on average. These opponents are the residents under uniform matching $\lambda = 0$, so $\alpha_{AB}^{'}(\kappa^\bullet)$ is relevant. These opponents are other mutants under perfectly assortative matching $\lambda = 1$, so $\alpha_{BB}^{'}(\kappa^\bullet)$ is relevant.

Proposition OA1 implies that one should only expect the correctly specified model to be stable against all nearby models in "special" cases — that is, when the expectation in the statement of Proposition OA1 is exactly equal to 0. One such special case is when the agents face a decision problem where 2's action does not affect 1's payoffs, that is $\frac{\partial u_1^\bullet}{\partial q_2} = 0$. This sets the expectation to zero, so the result never implies that the correctly specified model is evolutionarily fragile against a misspecified model in such decision problems.

In the duopoly game analyzed previously, we have $\frac{\partial u_1^\bullet}{\partial q_2}(q_1, q_2, \omega) = -\frac{1}{2}r^\bullet q_1$. Player 1 is harmed by player 2 producing more if $q_1 > 0$, and helped if $q_1 < 0$. From straightforward algebra, the expectation in Proposition OA1 simplifies to

$$\mathbb{E}^\bullet[s_1^2]\cdot(-\frac{1}{2}\psi(\kappa^\bullet)r^\bullet\alpha^\bullet)\cdot[(1-\lambda)\alpha_{AB}^{'}(\kappa^\bullet)+\lambda\alpha_{BB}^{'}(\kappa^\bullet)].$$

[19] More precisely, for every $s_i \in S$, $\alpha_{AA}(\kappa^\bullet)\cdot s_i$ maximizes the agent's objective expected utility across all of $\mathbb{R}$ when $-i$ uses the same linear strategy $\alpha_{AA}(\kappa^\bullet)$.

The proof of Proposition 1 shows that when $\lambda = 0$, $\alpha'_{AB}(\kappa^\bullet) < 0$. The proof of Proposition 2 shows that when $\lambda = 1$, $\alpha'_{BB}(\kappa^\bullet) > 0$. The uniqueness of EZ also follow from these results, for an open interval of κ containing $\kappa^\bullet$. We restrict $\mathbb{A}$ to the set of linear strategies, and Lemma 2 implies linear strategies played by two correctly specified firms against each other are interim optimal. Finally, Lemma 4 verifies that Assumptions OA2 through OA6 are satisfied. So, the conditions of Proposition OA1 hold for $\lambda \in \{0, 1\}$, and we deduce the correctly specified model is evolutionarily fragile against slightly higher κ (for $\lambda = 0$) and slightly lower κ (for $\lambda = 1$).

OA 2 Proofs Omitted from the Appendix

OA 2.1 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 observations 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 in the model where productivity is given by $P = b(x_i + x_{-i}) - m + \epsilon$. We find that $b^*(a_i, a_{-i}) = b^\bullet + \frac{m}{a_i + a_{-i}}$. That is, adherents of Θ_B end up with different beliefs about the game parameter b depending on the behavior of their typical opponents, which in turn affects how they respond to different rival investment levels. Stability reversal happens because when Θ_A is resident and the adherents of Θ_B always meet opponents who play $a_i = 1$, they end up with a more distorted belief about the fundamental than when Θ_B is resident.

OA 2.2 Proof of Proposition 1

Proof. We can take L_1, L_2, L_3 as given by Lemma 3. Suppose there is an EZ with behavior $\alpha = (\alpha_{AA}, \alpha_{AB}, \alpha_{BA}, \alpha_{BB})$ and beliefs over parameters $\mu_A \in \Delta(\Theta(\kappa^\bullet))$, $\mu_B \in \Delta(\Theta(\kappa))$. By Lemma 3, both μ_A and μ_B must be degenerate beliefs that induce zero KL divergence, since both groups match up with group A with probability 1. Furthermore, since Θ_A is correctly specified, it is easy to see that the parameter $F_{r^\bullet, \kappa^\bullet, \sigma_\zeta^\bullet}$ generates 0 KL divergence, hence the belief of the adherents of Θ_A must be degenerate on this correct parameter.

In terms of behavior, from Lemma 2, $\alpha_i^{BR}(\alpha_{-i}; \kappa, r) \leq \gamma$ for all $\alpha_{-i} \geq 0, \kappa \in [0, 1], r \geq 0$. Since the upper bound $\bar{M}_\alpha \geq \gamma$, the adherents of each model must be best responding (across all linear strategies in $[0,\infty)$) in all matches, given their beliefs about the environment.

Using the equilibrium belief of group A, we must have $\alpha_{AA} = \alpha_i^{BR}(\alpha_{AA};\kappa^\bullet, r^\bullet)$, so $\alpha_{AA} = \frac{\gamma-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)\alpha_{AA}}{1+r^\bullet}$. We find the unique solution $\alpha_{AA} = \frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$. Next we turn to α_{AB}, α_{BA}, and μ_B. We know μ_B puts probability 1 on some r_B. For adherents of groups A and B to best respond to each others' play and for group B's inference to have 0 KL divergence (when paired with an appropriate choice of σ_ζ), we must have $\alpha_{AB} = \frac{\gamma-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)\alpha_{BA}}{1+r^\bullet}$, $\alpha_{BA} = \frac{\gamma-\frac{1}{2}r_B\psi(\kappa)\alpha_{AB}}{1+r_B}$, and $r_B = r^\bullet\frac{\alpha_{BA}+\alpha_{AB}\psi(\kappa^\bullet)}{\alpha_{BA}+\alpha_{AB}\psi(\kappa)}$ from Lemma 3. We may rearrange the expression for α_{BA} to say $\alpha_{BA} = \gamma - r_B\alpha_{BA} - \frac{1}{2}r_B\psi(\kappa)\alpha_{AB}$. Substituting the expression of r_B into this expression of α_{BA}, we get

$$\begin{aligned}\alpha_{BA} &= \gamma - r_B\cdot(\alpha_{BA}+\alpha_{AB}\psi(\kappa) - \frac{1}{2}\alpha_{AB}\psi(\kappa))\\ &= \gamma - \frac{r^\bullet\alpha_{BA}+r^\bullet\alpha_{AB}\psi(\kappa^\bullet)}{\alpha_{BA}+\alpha_{AB}\psi(\kappa)}\cdot(\alpha_{BA}+\alpha_{AB}\psi(\kappa)-\frac{1}{2}\alpha_{AB}\psi(\kappa))\\ &= \gamma - r^\bullet\alpha_{BA} - r^\bullet\alpha_{AB}\psi(\kappa^\bullet) + \frac{1}{2}\psi(\kappa)\alpha_{AB}\frac{r^\bullet\alpha_{BA}+r^\bullet\alpha_{AB}\psi(\kappa^\bullet)}{\alpha_{BA}+\alpha_{AB}\psi(\kappa)}\end{aligned}$$

Multiply by $\alpha_{BA}+\alpha_{AB}\psi(\kappa)$ on both sides and collect terms by powers of α,

$$(\alpha_{BA})^2\cdot[-1-r^\bullet]+(\alpha_{BA}\alpha_{AB})\cdot[-\psi(\kappa)-\frac{1}{2}r^\bullet\psi(\kappa)-r^\bullet\psi(\kappa^\bullet)]-(\alpha_{AB})^2\cdot[\frac{1}{2}r^\bullet\psi(\kappa^\bullet)\psi(\kappa)]+\gamma[\alpha_{BA}+\alpha_{AB}\psi(\kappa)]=0.$$

Consider the following quadratic function in x,

$$H(x) := x^2\left[-1-r^\bullet\right]+(x\cdot\ell(x))\cdot[-\psi(\kappa)-\frac{1}{2}r^\bullet\psi(\kappa)-r^\bullet\psi(\kappa^\bullet)]-(\ell(x))^2\cdot[\frac{1}{2}r^\bullet\psi(\kappa^\bullet)\psi(\kappa)]+\gamma\left[x+\ell(x)\psi(\kappa)\right]=0, \tag{2}$$

where $\ell(x) := \frac{\gamma-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)x}{1+r^\bullet}$ is a linear function in x. In an EZ, α_{BA} is a root of $H(x)$ in $[0,\frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}]$. To see why, if we were to have $\alpha_{BA} > \frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$, then $\alpha_{AB}=0$. In that case, $r_B = r^\bullet$ and so $\alpha_{BA} = \alpha_i^{BR}(0;\kappa^\bullet,r^\bullet) = \frac{\gamma}{1+r^\bullet}$. Yet $\frac{\gamma}{1+r^\bullet} < \frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$, contradiction. Conversely, for any root x^* of $H(x)$ in $[0,\frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}]$, there is an EZ where $\alpha_{BA}=x^*$, $\alpha_{AB}=\ell(x^*)\in[0,\gamma]$, and $r_B = r^\bullet\frac{\alpha_{BA}+\alpha_{AB}\psi(\kappa^\bullet)}{\alpha_{BA}+\alpha_{AB}\psi(\kappa)}$.

Claim A.1. There exist some $\underline{\kappa}_1 < \kappa^\bullet < \bar{\kappa}_1$ so that H has a unique root in $[0,\frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}]$ for all $\kappa\in[\underline{\kappa}_1,\bar{\kappa}_1]\cap[0,1]$.

By Claim A.1 (proved in the Online Appendix), for $\kappa\in[\underline{\kappa}_1,\bar{\kappa}_1]\cap[0,1]$, group B has only one possible belief about elasticity (denoted by $r_B(\kappa)$) in EZ), since there is only one possible outcome in the match between group A and group B. This means α_{BB} is also pinned down, since there is only one solution to $\alpha_{BB} = \alpha_i^{BR}(\alpha_{BB};\kappa,r_B(\kappa))$. So for every

$\kappa \in [\underline{\kappa}_1, \bar{\kappa}_1] \cap [0,1]$, there is a unique EZ, where equilibrium behavior is given as a function of κ by $\alpha(\kappa) = (\alpha_{AA}(\kappa), \alpha_{AB}(\kappa), \alpha_{BA}(\kappa), \alpha_{BB}(\kappa))$.

Recall from Lemma 2 that the objective expected utility from playing α_i against an opponent who plays α_{-i} is $U_i^\bullet(\alpha_i, \alpha_{-i}) = \mathbb{E}[s_i^2] \cdot (\alpha_i \gamma - \frac{1}{2} r^\bullet \alpha_i^2 - \frac{1}{2} r^\bullet \psi(\kappa^\bullet) \alpha_i \alpha_{-i} - \frac{1}{2}\alpha_i^2)$. If $-i$ plays the rational best response, then the objective expected utility of choosing α_i is $\bar{U}_i(\alpha_i) := \mathbb{E}[s_i^2] \cdot (\alpha_i \gamma - \frac{1}{2} r^\bullet \alpha_i^2 - \frac{1}{2} r^\bullet \psi(\kappa^\bullet) \alpha_i \frac{\gamma - \frac{1}{2} r^\bullet \psi(\kappa^\bullet)\alpha_i}{1+r^\bullet} - \frac{1}{2}\alpha_i^2)$. The derivative in α_i is $\bar{U}_i'(\alpha_i) = \gamma - r^\bullet \alpha_i - \frac{1}{2}\frac{r^\bullet}{1+r^\bullet}\gamma\psi(\kappa^\bullet) + \frac{1}{2}\frac{(r^\bullet)^2\psi(\kappa^\bullet)^2}{1+r^\bullet}\alpha_i - \alpha_i$. We also know that $\alpha_{AA} = \frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$ satisfies the first-order condition that $\gamma - r^\bullet \alpha_{AA} - \frac{1}{2} r^\bullet \psi(\kappa^\bullet)\alpha_{AA} - \alpha_{AA} = 0$, therefore

$$\begin{aligned}\bar{U}_i'(\alpha_{AA}) &= -\frac{1}{2}\frac{r^\bullet}{1+r^\bullet}\gamma\psi(\kappa^\bullet) + \frac{1}{2}\frac{(r^\bullet)^2\psi(\kappa^\bullet)^2}{1+r^\bullet}\alpha_{AA} + \frac{1}{2}r^\bullet\psi(\kappa^\bullet)\alpha_{AA} \\ &= \left[\frac{r^\bullet\psi(\kappa^\bullet)}{2}\right]\left(\frac{-\gamma}{1+r^\bullet} + \frac{\alpha_{AA}\psi(\kappa^\bullet)r^\bullet}{1+r^\bullet} + \alpha_{AA}\right).\end{aligned}$$

Making the substitution $\alpha_{AA} = \frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$,

$$\begin{aligned}\frac{-\gamma}{1+r^\bullet} + \frac{\alpha_{AA}\psi(\kappa^\bullet)r^\bullet}{1+r^\bullet} + \alpha_{AA} &= \frac{-\gamma(1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet) + \gamma\psi(\kappa^\bullet)r^\bullet + \gamma(1+r^\bullet)}{(1+r^\bullet)(1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet)} \\ &= \frac{\frac{1}{2}\gamma\psi(\kappa^\bullet)r^\bullet}{(1+r^\bullet)(1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet)} > 0.\end{aligned}$$

Therefore, if we can show that $\alpha_{BA}'(\kappa^\bullet) > 0$, then there exists some $\underline{\kappa}_1 \leq \underline{\kappa} < \kappa^\bullet < \bar{\kappa} \leq \bar{\kappa}_1$ so that for every $\kappa \in [\underline{\kappa}, \bar{\kappa}] \cap [0,1]$, $\kappa \neq \kappa^\bullet$ adherents of Θ_B have strictly higher or strictly lower equilibrium fitness in the unique EZ than adherents of Θ_A, depending on the sign of $\kappa - \kappa^\bullet$. Consider again the quadratic function $H(x)$ in Equation (2) and implicitly characterize the unique root x in $[0, \frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}]$ as a function of κ in a neighborhood around $\kappa^\bullet$. Denote this root by α^M, let $D := \frac{d\alpha^M}{d\psi(\kappa)}$ and also note $\frac{d\ell(\alpha^M)}{d\psi(\kappa)} = \frac{-r^\bullet}{2(1+r^\bullet)}\psi(k^\bullet) \cdot D$. We have

$$\begin{aligned}&(-1-r^\bullet)\cdot(2\alpha^M)\cdot D + (\alpha^M\ell(\alpha^M))(-1-\frac{1}{2}r^\bullet) \\ &+ (\ell(\alpha^M)D + \alpha^M\frac{-r^\bullet}{2(1+r^\bullet)}\psi(\kappa^\bullet)D)\cdot(-\psi(\kappa) - \frac{1}{2}r^\bullet\psi(\kappa) - r^\bullet\psi(\kappa^\bullet)) + (\ell(\alpha^M))^2\cdot(-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)) \\ &+ (2\ell(\alpha^M)\frac{-r^\bullet}{2(1+r^\bullet)}\psi(\kappa^\bullet)D)\cdot(-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)\psi(\kappa)) + \gamma(D + \ell(\alpha^M) + \psi(\kappa)\frac{-r^\bullet}{2(1+r^\bullet)}\psi(\kappa^\bullet)D) = 0\end{aligned}$$

Evaluate at $\kappa = \kappa^\bullet$, noting that $\alpha^M(\kappa^\bullet) = \ell(\alpha^M(\kappa^\bullet)) = x^* := \frac{\gamma}{1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet}$. The terms without D are:

$$(x^*)^2(-1-\frac{1}{2}r^\bullet)+(x^*)^2(\frac{1}{2}r^\bullet\psi(\kappa^\bullet))+\gamma x^* = x^*\cdot\left[-x^*\cdot\left(1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet-\frac{1}{2}r^\bullet\right)+\gamma\right]$$
$$= x^*\cdot\left[-\gamma+\frac{1}{2}x^*r^\bullet+\gamma\right]=\frac{1}{2}r^\bullet(x^*)^2>0.$$

The coefficient in front of D is:

$$(-1-r^\bullet)(2x^*)+(x^*+x^*\frac{-r^\bullet}{2(1+r^\bullet)}\psi(\kappa^\bullet))\cdot(-\psi(\kappa^\bullet)-\frac{3}{2}r^\bullet\psi(\kappa^\bullet))+\frac{1}{2}x^*\frac{(r^\bullet)^2}{(1+r^\bullet)}\psi(\kappa^\bullet)^3+\gamma+\gamma\psi(\kappa^\bullet)^2\cdot\frac{-r^\bullet}{2(1+r^\bullet)}.$$

Make the substitution $\gamma = x^*\cdot\left(1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet\right)$,

$$x^*\cdot\left\{-2-2r^\bullet+\left(1-\frac{r^\bullet}{2(1+r^\bullet)}\psi(\kappa^\bullet)\right)\cdot\psi(\kappa^\bullet)(-\frac{3}{2}r^\bullet-1)+\frac{(r^\bullet)^2}{2(1+r^\bullet)}\psi(\kappa^\bullet)^3\right\}$$
$$+x^*\cdot\left\{\left(1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet\right)\cdot(1-\psi(\kappa^\bullet)^2\frac{r^\bullet}{2(1+r^\bullet)})\right\}.$$

Collect terms inside the parenthesis based on powers of $\psi(\kappa^\bullet)$, we get

$$x^*\cdot\left\{\psi(\kappa^\bullet)^3\frac{(r^\bullet)^2}{2(1+r^\bullet)}-\frac{\psi(\kappa^\bullet)^2r^\bullet}{2(1+r^\bullet)}(-\frac{3}{2}r^\bullet-1)+\psi(\kappa^\bullet)(-\frac{3}{2}r^\bullet-1)-2r^\bullet-2\right\}$$
$$+x^*\cdot\left\{-\psi(\kappa^\bullet)^3\frac{(r^\bullet)^2}{4(1+r^\bullet)}-\frac{\psi(\kappa^\bullet)^2r^\bullet}{2(1+r^\bullet)}\cdot(1+r^\bullet)+1+r^\bullet+\frac{1}{2}\psi(\kappa^\bullet)r^\bullet\right\}.$$

Combine to get: $x^*\cdot\left[\psi(\kappa^\bullet)^3\frac{(r^\bullet)^2}{4(1+r^\bullet)}+\frac{\psi(\kappa^\bullet)^2(r^\bullet)^2}{4(1+r^\bullet)}-\psi(\kappa^\bullet)r^\bullet-\psi(\kappa^\bullet)-r^\bullet-1\right]$. Here $\psi(\kappa^\bullet)^3\frac{(r^\bullet)^2}{4(1+r^\bullet)}$ and $\frac{\psi(\kappa^\bullet)^2(r^\bullet)^2}{4(1+r^\bullet)}$ are positive terms with $\psi(\kappa^\bullet)^3\frac{(r^\bullet)^2}{4(1+r^\bullet)}+\frac{\psi(\kappa^\bullet)^2(r^\bullet)^2}{4(1+r^\bullet)}\le\frac{(r^\bullet)^2}{4(1+r^\bullet)}+\frac{(r^\bullet)^2}{4(1+r^\bullet)}\le\frac{1}{2}\cdot r^\bullet\cdot\frac{r^\bullet}{1+r^\bullet}\le\frac{1}{2}r^\bullet$. Now $-r^\bullet+\frac{1}{2}\cdot r^\bullet<0$, and also $-\psi(\kappa^\bullet)r^\bullet-\psi(\kappa^\bullet)-1<0$. Thus the coefficient in front of D is strictly negative. This shows $D(\kappa^\bullet)>0$. Finally, $\frac{d\alpha^M}{d\psi(\kappa)}$ has the same sign as $\frac{d\alpha^M}{d\kappa}$ since $\psi(\kappa)$ is strictly increasing in κ. □

OA 2.3 Proof of Proposition 4

We consider a distribution q over two situations that have different true values of $r^\bullet$, where $q(r^\bullet=0)=1-\varepsilon$ and $q(r^\bullet=\bar{r})=\varepsilon$, for some $\bar{r}\ge 3$. Suppose $p=(1,0)$ with the rational model as the resident. We claim that there are some $0<r_0<r_1$ such that the following three conditions hold.

- For $0 \leq r < r_0$, in every EZ, every singleton model (r, κ) obtains negative payoff when $r^\bullet = \overline{r}$, and no more than the rational model's payoff when $r^\bullet = 0$.

- For $r_0 \leq r \leq r_1$, in every EZ, every singleton model (r, κ) obtains strictly less than the rational payoff when $r^\bullet = 0$, and no more than the Stackelberg payoff against a rational opponent when $r^\bullet = \overline{r}$. Furthermore, the singleton model's highest EZ payoff when $r^\bullet = 0$ is given by a continuous function $\xi(r)$.

- For $r > r_1$, in every EZ, every singleton model (r, κ) obtains payoff less than half that of the rational payoff when $r^\bullet = 0$, and no more than the Stackelberg payoff against a rational opponent when $r^\bullet = \overline{r}$.

We show that if these conditions hold, then the correctly specified model is evolutionarily stable against any singleton model when ε is sufficiently small. Let $c_0 > 0$ be the rational model's payoff when $r^\bullet = 0$, let $c_{\overline{r}} > 0$ be the rational model's payoff when $r^\bullet = \overline{r}$, and let $c_s > 0$ be the Stackelberg payoff against the rational model when $r^\bullet = \overline{r}$. For every $r \in [r_0, r_1]$, there exists some $\epsilon_r > 0$ so that $\epsilon_r \cdot c_s + (1 - \epsilon_r) \cdot \xi(r) = \epsilon_r \cdot c_{\overline{r}} + (1 - \epsilon_r) \cdot c_0$. Since $\xi(r) < c_0$ for every r, we get that if $\varepsilon < \epsilon_r$, then the rational model is evolutionarily stable against the singleton model with r. We have that $\min_{r \in [r_0, r_1]} \epsilon_r > 0$ since $\xi(r)$ is continuous. Finally, there is some $\epsilon' > 0$ so that $\epsilon' \cdot c_s + (1 - \epsilon') \cdot (c_0/2) < \epsilon' \cdot c_{\overline{r}} + (1 - \epsilon') \cdot c_0$. Whenever $\varepsilon < \min\{\min_{r \in [r_0, r_1]} \epsilon_r, \epsilon'\}$, the rational model is evolutionarily stable against the singleton model with any $r \geq 0$.

As $\varepsilon \to 0$, by linearity of expectations the expected payoff converges to the payoff when $r^\bullet = 0$ with probability 1; for any $\hat{r} > 0$, a mutant who believes $r = \hat{r}$ obtains less than the correctly specified resident when $r^\bullet = 0$. Thus, a mutant with $r \in (0, \tilde{r}]$ does worse than the correctly specified resident. On the other hand, while the misspecified resident may do better than the correctly specified resident when $r > \tilde{r}$, they do significantly worse when $r = 0$, and uniformly so over all such r; as $\varepsilon \to 0$, the benefit vanishes uniformly and we have that again that the the rational model is stable against all such r.

Recall that Lemma 2 says the best replies are $\alpha_i^{BR}(\alpha_{-i}; \kappa, r) := \frac{\gamma - \frac{1}{2} r \psi(\kappa) \alpha_{-i}}{1+r}$. Suppose $r^\bullet = 0$. In this case, the rational player chooses $q(s_i) = \gamma s_i$, and therefore any other $\hat{r}$ chooses $q(s_i) = \gamma \left(\frac{1 - \frac{1}{2} \hat{r} \psi(\kappa)}{1+\hat{r}} \right) s_i$. The rational player's expected payoff is $\mathbb{E}[\mathbb{E}[\omega q_i(s_i) - \frac{1}{2} q_i(s_i)^2 \mid s_i]] = \mathbb{E}[s_i^2] \cdot \left(\frac{\gamma^2}{2} \right)$; the mutant playing strategy $q(s_i) = \alpha_i s_i$ obtains $\mathbb{E}[s_i^2](\gamma \alpha_i - \frac{1}{2} \alpha_i^2)$, which is quadratic in α_i and maximized at $\alpha_i = \gamma$. Therefore, the correctly specified resident obtains the highest payoff.

If $r^\bullet = \overline{r}$, then a mutant who believes $\hat{r} = 0$ uses strategy with slope $\alpha_i = \gamma$; the mutant obtains $\mathbb{E}[\mathbb{E}[\omega\alpha_i s_i - \overline{r}\left(\frac{1}{2}(\alpha_i + \alpha_{-i})\right)\alpha_i s_i^2 - \frac{\alpha_i^2 s_i^2}{2}] = \mathbb{E}[s_i^2]\left(\frac{\gamma^2}{2} - \overline{r}\gamma^2\left(\frac{1}{2}(1 + \frac{1-\frac{1}{2}\overline{r}\psi(\kappa)}{1+\overline{r}})\right)\right)$. Note that since $\frac{1}{2}\psi(\kappa)$ is bounded away from 1, $\frac{1-\frac{1}{2}\overline{r}\psi(\kappa)}{1+\overline{r}}$ is bounded away from 0. Therefore, as long as $\overline{r} \geq 1$, we have that the mutant's payoff will be negative. Since payoffs are continuous, taking $r_0 \to 0$, we can find some sufficiently small r_0 such that any mutant with $r < r_0$ obtains a negative payoff when $r^\bullet = \overline{r}$.

From Lemma 2, we know that the rational resident always chooses the linear strategy with $\alpha_{-i} = \gamma$ when $r^\bullet = 0$. Thus, an adherent of the singleton model with $r_0 \leq r \leq r_1$ chooses the linear coefficient $\frac{\gamma - \frac{1}{2}r\psi(\kappa)\gamma}{1+r} < \frac{\gamma}{1+r} < \gamma$ in every EZ when $r^\bullet = 0$. But the game with $r^\bullet = 0$ has $\alpha_i = \gamma$ as the strictly dominant strategy, so the mutant gets strictly lower payoff than the resident. The mutant's EZ strategy is a continuous function of r, so their payoff as a function of r must also be continuous. When $r^\bullet = \overline{r}$, because the resident must best respond to the mutant's strategy in an EZ, the mutant cannot get more than the Stackelberg payoff.

Find a small enough $x > 0$ so that $x\gamma - \frac{1}{2}x^2 < \frac{1}{4}\gamma^2$. By the same argument as before, an adherent of the singleton model with r chooses the linear coefficient $\frac{\gamma - \frac{1}{2}r\psi(\kappa)\gamma}{1+r}$. Set r_1 so that $\frac{\gamma}{1+r_1} = x$. For any $r \geq r_1$, we get the mutant's EZ strategy has a linear coefficient of $\frac{\gamma - \frac{1}{2}r\psi(\kappa)\gamma}{1+r} \leq \frac{\gamma}{1+r} \leq \frac{\gamma}{1+r_1} = x$, so their payoff is no larger than $x\gamma - \frac{1}{2}x^2 < \frac{1}{4}\gamma^4$. This is less than half of the payoff of the rational residents, who choose the linear coefficient γ and get $\frac{1}{2}\gamma^2$.

OA 2.4 Proof of Example 2

Proof. Define $b^*(a_i, a_{-i}) := b^\bullet + \frac{m}{a_i + a_{-i}}$. 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.

In every EZ with $\lambda = 0$ and $p = (1, 0)$, 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 $\lambda = 0, p_B = 1$. 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 $\lambda = 0$ and $p = (0,1)$, 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. □

OA 2.5 Proof of Example 3

Proof. Let $KL_{4,1} := 0.4 \cdot \ln \frac{0.4}{0.1} + 0.6 \cdot \ln \frac{0.6}{0.9} \approx 0.3112$, $KL_{4,8} := 0.4 \cdot \ln \frac{0.4}{0.8} + 0.6 \cdot \ln \frac{0.6}{0.2} \approx 0.3819$, and $KL_{2,4} := 0.2 \cdot \ln \frac{0.2}{0.4} + 0.8 \cdot \ln \frac{0.8}{0.6} \approx 0.0915$. Let λ_h be the unique solution to $(1-\lambda)KL_{2,4} - \lambda(KL_{4,8} - KL_{4,1}) = 0$, so $\lambda_h \approx 0.564$.

We show for any $\lambda \in [0, \lambda_h)$, there exists a unique EZ $\mathfrak{Z} = (\Theta_A, \Theta_B, \mu_A, \mu_B, p = (1,0), \lambda, (a))$, and that this EZ has μ_B putting probability 1 on F_H, $a_{AA} = a_1$, $a_{AB} = a_1$, $a_{BA} = a_2$, $a_{BB} = a_2$. First, we may verify that under F_H, a_2 best responds to both a_1 and a_2. Also, the KL divergence of F_H is $\lambda \cdot KL_{4,8}$ while that of F_L is $\lambda \cdot KL_{4,1} + (1-\lambda) \cdot KL_{2,4}$. Since $\lambda < \lambda_h$, we see that F_H has strictly lower KL divergence. Finally, to check that there are no other EZs, note we must have $a_{AA} = a_1$, $a_{AB} = a_1$, $a_{BA} = a_2$ in every EZ. In an EZ where a_{BB} puts probability $q \in [0,1]$ on a_2, the KL divergence of F_H is $\lambda p \cdot KL_{4,8}$and the KL divergence of F_L is $\lambda p \cdot KL_{4,1} + (1-\lambda) \cdot KL_{2,4}$. We have

$$\lambda q \cdot KL_{4,1} + (1-\lambda) \cdot KL_{2,4} - \lambda q \cdot KL_{4,8} = \lambda q \cdot (KL_{4,1} - KL_{4.8}) + (1-\lambda) KL_{2,4} \geq (1-\lambda) KL_{2,4} - \lambda(KL_{4,8} - KL_{4,1}).$$

Since $\lambda < \lambda_h$, this is strictly positive. Therefore we must have μ_B put probability 1 on F_H, which in turn implies $q = 1$.

When Θ_A is dominant, the equilibrium fitness of Θ_A is always 0.25 for every λ. The equilibrium fitness of Θ_B, as a function of λ, is $0.4\lambda + 0.2(1-\lambda)$. Let λ_l solve $0.25 = 0.4\lambda + 0.2(1-\lambda)$, that is $\lambda_l = 0.25$. This shows Θ_A is evolutionarily fragile against Θ_B for $\lambda \in (\lambda_l, \lambda_h)$, and it is evolutionarily stable against Θ_B for $\lambda = 0$.

Now suppose $\lambda = 1$. If there is an EZ with $p_A = 1$ where a_{BB} plays a_2 with positive probability, then μ_B must put probability 1 on F_L, since $KL_{4,1} < KL_{4,8}$. This is a contradiction, since a_2 does not best respond to itself under F_L. So the unique EZ involves $a_{AA} = a_1$, $a_{AB} = a_1$, $a_{BA} = a_2$, $a_{BB} = a_3$. In the EZ, the fitness of Θ_A is 0.25, and the fitness of Θ_B is 0.2. This shows Θ_A is evolutionarily stable against Θ_B for $\lambda = 1$. □

OA 2.6 Proof of Claim A.1

Proof. We show that $H(x)$ (i) has a unique root in $[0, \frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}]$ when $\kappa = \kappa^\bullet$; (ii) does not have a root at $x = 0$ or $x = \frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$, and (iii) the root in the interval is not a double root. By these three statements, since $H(x)$ is a continuous function of κ, there must exist some $\underline{\kappa}_1 < \kappa^\bullet < \bar{\kappa}_1$ so that it continues to have a unique root in $[0, \frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}]$ for all $\kappa \in [\underline{\kappa}_1, \bar{\kappa}_1] \cap [0, 1]$.

Statement (i) has to do with the fact that if $\kappa = \kappa^\bullet$, then we need $\alpha_{AB} = \frac{\gamma - \frac{1}{2}r^\bullet\psi(\kappa^\bullet)\alpha_{BA}}{1+r^\bullet}$ and $\alpha_{BA} = \frac{\gamma - \frac{1}{2}r^\bullet\psi(\kappa^\bullet)\alpha_{AB}}{1+r^\bullet}$. These are linear best response functions with a slope of $-\frac{1}{2}\frac{r^\bullet}{1+r^\bullet}\psi(\kappa^\bullet)$, which falls in $(-\frac{1}{2}, 0)$. So there can only be one solution to H in that region (even when we allow $\alpha_{AB} \neq \alpha_{BA}$), which is the symmetric equilibrium found before $\alpha_{AB} = \alpha_{BA} = \frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$.

For Statement (ii), we evaluate $H(0) = -(\frac{\gamma}{1+r^\bullet})^2\frac{1}{2}r^\bullet\psi(\kappa^\bullet)^2 + \frac{\gamma^2\psi(\kappa^\bullet)}{1+r^\bullet} = \frac{\psi(\kappa^\bullet)\gamma^2}{1+r^\bullet}(1 - \frac{(1/2)r^\bullet\psi(\kappa^\bullet)}{1+r^\bullet}) \neq 0$ because $1+r^\bullet > (1/2)r^\bullet\psi(\kappa^\bullet)$. Finally, we evaluate $H(\frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}) = (\frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)})^2(-1-r^\bullet)+\gamma\frac{\gamma}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)} = \frac{\gamma^2}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}(1-\frac{1+r^\bullet}{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)})$. This is once again not 0 because $1+r^\bullet > (1/2)r^\bullet\psi(\kappa^\bullet)$.

For Statement (iii), we show that $H'(x^*) < 0$ where $x^* = \frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$. We find that

$$\begin{aligned} H'(x) =& 2x(-1-r^\bullet) + \left(\frac{\gamma - r^\bullet\psi(\kappa^\bullet)x}{1+r^\bullet}\right)(-\psi(\kappa^\bullet) - \frac{1}{2}r^\bullet\psi(\kappa^\bullet) - r^\bullet\psi(\kappa^\bullet)) \\ &- 2\left(\frac{\gamma - \frac{1}{2}r^\bullet\psi(\kappa^\bullet)x}{1+r^\bullet}\right)\left(\frac{-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}{1+r^\bullet}\right)\left(\frac{1}{2}r^\bullet\psi(\kappa^\bullet)^2\right) + \gamma - \frac{\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}{1+r^\bullet}\gamma\psi(\kappa^\bullet). \end{aligned}$$

Collecting terms, the coefficient on x is

$$-2-2r^\bullet+\frac{\psi(\kappa^\bullet)^2r^\bullet}{1+r^\bullet}\left(\frac{3}{2}r^\bullet+1-\frac{1}{4}(\frac{(r^\bullet)^2\psi(\kappa^\bullet)^2}{1+r^\bullet})\right),$$

while the coefficient on the constant is

$$\frac{\gamma\psi(\kappa^\bullet)}{1+r^\bullet}\left(-\frac{3}{2}r^\bullet-1+\frac{1}{2}\frac{(r^\bullet)^2\psi(\kappa^\bullet)^2}{1+r^\bullet}-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)\right)+\gamma.$$

Therefore, we may calculate $H^{'}(x^*)\cdot\frac{1}{x^*}(1+r^\bullet)^2$, which has the same sign as $H^{'}(x^*)$, to be:

$$\begin{aligned}&-(1+r^\bullet)^2(2+2r^\bullet)+\psi(\kappa^\bullet)^2r^\bullet((1+r^\bullet)(\frac{3}{2}r^\bullet+1)-\frac{1}{4}(r^\bullet)^2\psi(\kappa^\bullet)^2)\\&+(1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet))\left[\psi(\kappa^\bullet)((1+r^\bullet)[-\frac{3}{2}r^\bullet-1-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)]+\frac{1}{2}(r^\bullet)^2\psi(\kappa^\bullet)^2)+(1+r^\bullet)^2\right].\end{aligned}$$

We have

$$-(1+r^\bullet)^2(2+2r^\bullet)+(1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet))(1+r^\bullet)^2\le(1+r^\bullet)^2(-1-\frac{1}{2}r^\bullet)<0,$$

since $0\le\psi(\kappa^\bullet)\le 1$. Also, for the same reason,

$$(1+r^\bullet)[-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)]+\frac{1}{2}(r^\bullet)^2\psi(\kappa^\bullet)^2\le-\frac{1}{2}(r^\bullet)^2\psi(\kappa^\bullet)+\frac{1}{2}(r^\bullet)^2\psi(\kappa^\bullet)^2\le 0.$$

Finally, $\psi(\kappa^\bullet)^2r^\bullet(1+r^\bullet)(\frac{3}{2}r^\bullet+1)+(1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet))\psi(\kappa^\bullet)(1+r^\bullet)(-\frac{3}{2}r^\bullet-1)$ is no larger than

$$\begin{aligned}&\psi(\kappa^\bullet)^2r^\bullet(\frac{3}{2}(r^\bullet)^2+\frac{5}{2}r^\bullet+1)+[r^\bullet\psi(\kappa^\bullet)r^\bullet(-(3/2)r^\bullet)]\\&+[r^\bullet\psi(\kappa^\bullet)r^\bullet(-1)+1\cdot\psi(\kappa^\bullet)r^\bullet(-(3/2)r^\bullet)]+[r^\bullet\psi(\kappa^\bullet)\cdot 1\cdot(-1)]\end{aligned}$$

where the negative terms in the first, second, and third square brackets are respectively larger in absolute value than the first, second and third parts in the expansion of the first summand. Therefore, we conclude $H^{'}(x^*)<0$. $\square$

OA 2.7 Proof of Lemma 1

Proof. For $i \neq j$, rewrite $s_i = \left(\omega + \frac{\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}} z\right) + \frac{1-\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}}\eta_i$ and $s_j = \left(\omega + \frac{\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}} z\right) + \frac{1-\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}}\eta_j$. Note that $\omega + \frac{\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}} z$ has a normal distribution with mean 0 and variance $\sigma_\omega^2 + \frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2$. The posterior distribution of $\left(\omega + \frac{\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}} z\right)$ given s_i is therefore normal with a mean of $\frac{1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}{1/(\sigma_\omega^2+\frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)+1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)} s_i$ and a variance of $\frac{1}{1/(\sigma_\omega^2+\frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)+1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}$.

Since η_j is mean-zero and independent of i's signal, the posterior distribution of $s_j \mid s_i$ under the correlation parameter κ is normal with a mean of

$$\frac{1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}{1/(\sigma_\omega^2 + \frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2) + 1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)} s_i$$

and a variance of $\frac{1}{1/(\sigma_\omega^2+\frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)+1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)} + \frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2$. We thus define $\psi(\kappa) := \frac{1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}{1/(\sigma_\omega^2+\frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)+1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}$ for $\kappa \in [0,1)$, and $\psi(1) := 1$. To see that $\psi(\kappa)$ is strictly increasing in κ, we have

$$\begin{aligned} 1/\psi(\kappa) &= 1 + \frac{\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2}{\sigma_\omega^2 + \frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2} \\ &= 1 + \frac{(1-\kappa)^2\sigma_\epsilon^2}{(\kappa^2+(1-\kappa)^2)\sigma_\omega^2 + \kappa^2\sigma_\epsilon^2} \end{aligned}$$

and then we can verify that the second term is decreasing in κ.

As $\kappa \to 1$, the term $1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)$ tends to ∞, so $\frac{1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}{1/(\sigma_\omega^2+\frac{\kappa^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)+1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}$ approaches $\frac{1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)}{1/(\frac{(1-\kappa)^2}{\kappa^2+(1-\kappa)^2}\sigma_\epsilon^2)} = 1$. We also verify that $\psi(0) = \frac{1/\sigma_\epsilon^2}{(1/\sigma_\omega^2)+(1/\sigma_\epsilon^2)} > 0$.

Finally, for any $\kappa \in [0,1]$, $\frac{\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}} z + \frac{1-\kappa}{\sqrt{\kappa^2+(1-\kappa)^2}}\eta_i$ has variance σ_ϵ^2 and mean 0, so $\mathbb{E}_\kappa[\omega \mid s_i] = \frac{1/\sigma_\epsilon^2}{1/\sigma_\epsilon^2+1/\sigma_\omega^2} s_i$. We then define γ as the strictly positive constant $\frac{1/\sigma_\epsilon^2}{1/\sigma_\epsilon^2+1/\sigma_\omega^2}$. $\square$

OA 2.8 Proof of Lemma 2

Proof. Player i's conditional expected utility given signal s_i is

$$\alpha_i s_i \cdot \mathbb{E}_\kappa[\mathbb{E}_{r\sim \text{marg}_r(\mu)}[\omega - \frac{1}{2}r\alpha_i s_i - \frac{1}{2}r\alpha_{-i}s_{-i} + \zeta] \mid s_i] - \frac{1}{2}(\alpha_i s_i)^2$$

by linearity, expectation over r is equivalent to evaluating the inner expectation with $r = \hat{r}$, which gives

$$\begin{aligned}
&\alpha_i s_i \cdot \mathbb{E}_\kappa[\omega - \frac{1}{2}\hat{r}\alpha_i s_i - \frac{1}{2}\hat{r}\alpha_{-i}s_{-i} + \zeta | s_i] - \frac{1}{2}(\alpha_i s_i)^2 \\
=&\alpha_i s_i \cdot (\gamma s_i - \frac{1}{2}\hat{r}\alpha_i s_i - \frac{1}{2}\hat{r}\psi(\kappa)s_i\alpha_{-i}) - \frac{1}{2}(\alpha_i s_i)^2 \\
=&s_i^2 \cdot (\alpha_i\gamma - \frac{1}{2}\hat{r}\alpha_i^2 - \frac{1}{2}\hat{r}\psi(\kappa)\alpha_i\alpha_{-i} - \frac{1}{2}\alpha_i^2).
\end{aligned}$$

The term in parenthesis does not depend on s_i, and the second moment of s_i is the same for all values of κ. Therefore this expectation is $\mathbb{E}[s_i^2] \cdot \left(\alpha_i\gamma - \frac{1}{2}\hat{r}\alpha_i^2 - \frac{1}{2}\hat{r}\psi(\kappa)\alpha_i\alpha_{-i} - \frac{1}{2}\alpha_i^2\right)$. The expression for $\alpha_i^{BR}(\alpha_{-i};\kappa,r)$ follows from simple algebra, noting that $\mathbb{E}[s_i^2] > 0$ while the second derivative with respect to α_i for the term in the parenthesis is $-\frac{1}{2}\hat{r} - \frac{1}{2} < 0$.

To see that the said linear strategy is optimal among all strategies, suppose i instead chooses any q_i after s_i. By above arguments, the objective to maximize is

$$q_i \cdot (\gamma s_i - \frac{1}{2}\hat{r}q_i - \frac{1}{2}\hat{r}\psi(\kappa)s_i\alpha_{-i}) - \frac{1}{2}q_i^2.$$

This objective is a strictly concave function in q_i, as $-\frac{1}{2}\hat{r} - \frac{1}{2} < 0$. First-order condition finds the maximizer $q_i^* = \alpha_i^{BR}(\alpha_{-i};\kappa,\hat{r})$. Therefore, the linear strategy also maximizes interim expected utility after every signal s_i, and so it cannot be improved on by any other strategy. □

OA 2.9 Proof of Lemma 3

Proof. Note that $\frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)} \geq 0$ and $\frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)} = 1 + \frac{\alpha_{-i}(\psi(\kappa^\bullet)-\psi(\kappa))}{\alpha_i+\alpha_{-i}\psi(\kappa)} \leq 1 + \frac{1}{\psi(0)}$ (recalling $\psi(0) > 0$). Hence let $L_3 = r^\bullet \cdot (1 + \frac{1}{\psi(0)})$. When $\bar{M}_r \geq L_3$, we always have $r_i^{INF}(\alpha_i, \alpha_{-i},; \kappa^\bullet, \kappa, r^\bullet) \leq \bar{M}_r$ for all $\alpha_i, \alpha_{-i} \geq 0$ and $\kappa^\bullet, \kappa \in [0,1]$.

Conditional on the signal s_i, the distribution of market price under the model $F_{\hat{r},\kappa,\hat{\sigma}_\zeta}$ is normal with a mean of

$$\mathbb{E}[\omega \mid s_i] - \frac{1}{2}\hat{r}\alpha_i s_i - \frac{1}{2}\hat{r}\alpha_{-i} \cdot \mathbb{E}_\kappa[s_{-i} \mid s_i] = \gamma s_i - \frac{1}{2}\hat{r}\alpha_i s_i - \frac{1}{2}\hat{r}\alpha_{-i}\psi(\kappa)s_i,$$

while the distribution of market price under the parameter $F_{r^\bullet,\kappa^\bullet,\sigma_\zeta^\bullet}$ is normal with a mean of

$$\mathbb{E}[\omega \mid s_i] - \frac{1}{2}r^\bullet\alpha_i s_i - \frac{1}{2}r^\bullet\alpha_{-i} \cdot \mathbb{E}_{\kappa^\bullet}[s_{-i} \mid s_i] = \gamma s_i - \frac{1}{2}r^\bullet\alpha_i s_i - \frac{1}{2}r^\bullet\alpha_{-i}\psi(\kappa^\bullet)s_i.$$

Matching coefficients on s_i, we find that if $\hat{r} = r^\bullet \frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)}$, then these means match after every s_i. On the other hand, for any other value of $\hat{r}$, these means will not match for any s_i and thus $D_{KL}(F_{r^\bullet,\kappa^\bullet,\sigma_\zeta^\bullet}(\alpha_i,\alpha_{-i}) \parallel F_{\hat{r},\kappa,\hat{\sigma}_\zeta}(\alpha_i,\alpha_{-i})) > 0$ for any $\hat{r} \neq r^\bullet \frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)}$.

Let $L_1 = \max_{\kappa\in[0,1]}\left\{\mathrm{Var}_\kappa[\omega \mid s_i] + \mathrm{Var}_\kappa\left[\frac{1}{2}r^\bullet \cdot (1+\frac{1}{\psi(0)})B_\alpha \cdot s_{-i} \mid s_i\right]\right\}$. This maximum exists and is finite, since the expression is a continuous function of κ on the compact domain $[0,1]$. Also, let $L_2 = \max_{\kappa\in[0,1]}\left\{\mathrm{Var}_\kappa[\omega \mid s_i] + \mathrm{Var}_\kappa\left[\frac{1}{2}r^\bullet B_\alpha \cdot s_{-i} \mid s_i\right]\right\}$,where the maximum exists for the same reason. Conditional on the signal s_i, the variance of market price under the parameter $F_{r^\bullet \frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)},\kappa,\hat{\sigma}_\zeta}$ is

$$\mathrm{Var}_\kappa\left[\omega - \frac{1}{2}r^\bullet\frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)}\alpha_{-i}s_{-i} \mid s_i\right] + \hat{\sigma}_\zeta^2.$$

Since ω and s_{-i} are positively correlated given s_i, and using the fact $r^\bullet \frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)} \leq r^\bullet\cdot(1+\frac{1}{\psi(0)})$ and $\alpha_{-i} \leq B_\alpha$, this variance is no larger than

$$\mathrm{Var}_\kappa\left[\omega \mid s_i\right] + \mathrm{Var}_\kappa\left[\frac{1}{2}r^\bullet \cdot (1+\frac{1}{\psi(0)})B_\alpha \cdot s_{-i} \mid s_i\right] + \hat{\sigma}_\zeta^2 = L_1 + \hat{\sigma}_\zeta^2.$$

On the other hand, the variance of market price under the parameter $F_{r^\bullet,\kappa^\bullet,\sigma_\zeta^\bullet}$ is

$$\mathrm{Var}_{\kappa^\bullet}\left[\omega - \frac{1}{2}r^\bullet\alpha_{-i}s_{-i} \mid s_i\right]+(\sigma_\zeta^\bullet)^2 \leq \mathrm{Var}_{\kappa^\bullet}[\omega \mid s_i]+\mathrm{Var}_{\kappa^\bullet}\left[\frac{1}{2}r^\bullet B_\alpha \cdot s_{-i} \mid s_i\right]+(\sigma_\zeta^\bullet)^2 \leq L_2+(\sigma_\zeta^\bullet)^2.$$

At the same time, since $(\sigma_\zeta^\bullet)^2 \geq L_1$, this conditional variance is at least L_1. Among values of $\hat{\sigma}_\zeta^2 \in [0,\bar{M}_{\sigma_\zeta}^2]$, there exists exactly one such that the conditional variance under $F_{r^\bullet \frac{\alpha_i+\alpha_{-i}\psi(\kappa^\bullet)}{\alpha_i+\alpha_{-i}\psi(\kappa)},\kappa,\hat{\sigma}_\zeta}$ is the same as that under $F_{r^\bullet,\kappa^\bullet,\sigma_\zeta^\bullet}$, since we have let $\bar{M}_{\sigma_\zeta}^2 \geq (\sigma_\zeta^\bullet)^2 + L_2$. Thus there is one choice of $\hat{\sigma}_\zeta \in [0,\bar{M}_{\sigma_\zeta}]$ with such that $D_{KL}(F_{r^\bullet,\kappa^\bullet,\sigma_\zeta^\bullet}(\alpha_i,\alpha_{-i}) \parallel$

$F_{r^{\bullet}\frac{\alpha_i+\alpha_{-i}\psi(\kappa^{\bullet})}{\alpha_i+\alpha_{-i}\psi(\kappa)},\kappa,\tilde{\sigma}_{\zeta}}(\alpha_i,\alpha_{-i})) = 0$. For any other choice of $\tilde{\sigma}_{\zeta}$, we conclude that $D_{KL}(F_{r^{\bullet},\kappa^{\bullet},\sigma_{\zeta}^{\bullet}}(\alpha_i,\alpha_{-i}) \parallel F_{r^{\bullet}\frac{\alpha_i+\alpha_{-i}\psi(\kappa^{\bullet})}{\alpha_i+\alpha_{-i}\psi(\kappa)},\kappa,\tilde{\sigma}_{\zeta}}(\alpha_i,\alpha_{-i})) > 0$. $\square$

OA 2.10 Proof of Lemma 4

Proof. Assumption OA2 holds as $\mathbb{A}, \Theta_A, \Theta_B$ are compact due to the finite bounds $\bar{M}_\alpha, \bar{M}_r, \bar{M}_{\sigma_\zeta}$. Also, from Lemma 2, the expected utility from playing α_i against α_{-i} in a model with parameters $(\hat{r},\kappa,\sigma_\zeta)$ is $\mathbb{E}[s_i^2]\cdot\left(\alpha_i\gamma - \frac{1}{2}\hat{r}\alpha_i^2 - \frac{1}{2}\hat{r}\psi(\kappa)\alpha_i\alpha_{-i} - \frac{1}{2}\alpha_i^2\right)$. This is a continuous function in $(\alpha_i,\alpha_{-i},\hat{r})$ and strictly concave in α_i. Therefore Assumptions OA3 and OA6 are satisfied.

To see the finiteness and continuity of the K functions, first recall that the KL divergence from a true distribution $\mathcal{N}(\mu_1,\sigma_1^2)$ to a different distribution $\mathcal{N}(\mu_2,\sigma_2^2)$ is given by $\ln(\sigma_2/\sigma_1) + \frac{\sigma_1^2+(\mu_1-\mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$. Under own play α_i, opponent play α_{-i}, correlation parameter κ, elasticity $\hat{r}$ and price idiosyncratic variance σ_ζ^2, the expected distribution of price after signal s_i is

$$-\frac{1}{2}\hat{r}\alpha_i s_i + (\omega - \frac{1}{2}\hat{r}\alpha_{-i}s_{-i} \mid s_i,\kappa) + \hat{\zeta}$$

where the first term is not random, the middle term is the conditional distribution of $\omega - \frac{1}{2}\hat{r}\alpha_{-i}s_{-i}$ given s_i, based on the joint distribution of (ω,s_i,s_{-i}) with correlation parameter κ. The final term is an independent random variable with mean 0, variance σ_ζ^2. The analogous true distribution of price is

$$-\frac{1}{2}r^{\bullet}\alpha_i s_i + (\omega - \frac{1}{2}r^{\bullet}\alpha_{-i}s_{-i} \mid s_i,\kappa^{\bullet}) + \zeta^{\bullet}$$

where $\zeta^{\bullet}$ is an independent random variable with mean 0, variance $(\sigma_\zeta^{\bullet})^2$. For a fixed κ, we may find $0 < \underline{\sigma}^2 < \bar{\sigma}^2 < \infty$ so that the variances of both distributions lie in $[\underline{\sigma}^2, \bar{\sigma}^2]$ for all $s_i \in \mathbb{R}$, $\alpha_i, \alpha_{-i} \in [0,\bar{M}_\alpha]$, $\hat{r} \in [0,\bar{M}_r]$. First note that as a consequence of the multivariate normality, the variances of these two expressions do not change with the realization of s_i. The lower bound comes from the fact that $\mathrm{Var}_\kappa(\omega - \frac{1}{2}\hat{r}\alpha_{-i}s_{-i} \mid s_i)$ is nonzero for all $\alpha_{-i},\hat{r}$ in the compact domains and it is a continuous function of these two arguments, so it must have some positive lower bound $\underline{\sigma}^2 > 0$. For a similar reason, the variance of the middle term has a upper bound for choices of the parameters $\alpha_{-i},\hat{r}$ in the compact domains, and the inference about σ_ζ^2 is also bounded.

The difference in the means of the two distributions is no larger than $s_i \cdot [\frac{1}{2}(\bar{M}_r + r^{\bullet})\cdot 1 +$

$\frac{1}{2}(\bar{M}_r + r^\bullet) \cdot 1 \cdot (\psi(\kappa) + \psi(\kappa^\bullet))]$. Thus consider the function

$$h(s_i) := \ln(\bar{\sigma}/\underline{\sigma}) + \frac{1}{2}(\bar{\sigma}^2/\underline{\sigma}^2) + \frac{[\frac{1}{2}(\bar{M}_r + r^\bullet) \cdot 1 + \frac{1}{2}(\bar{M}_r + r^\bullet) \cdot 1 \cdot (\psi(\kappa) + \psi(\kappa^\bullet))]^2}{2\underline{\sigma}^2} s_i^2 - \frac{1}{2}.$$

That is $h(s_i)$ has the form $h(s_i) = C_1 + C_2 s_i^2$ for constants C_1, C_2. It is absolutely integrable against the distribution of s_i, and it dominates the KL divergence between the true and expected price distributions at every s_i and for any choices of $\alpha_i, \alpha_{-i} \in [0, \bar{M}_\alpha], \hat{r} \in [0, \bar{M}_r], \sigma_\zeta^2 \in [0, \bar{M}_\zeta]$. This shows K_A, K_B are finite, so Assumption OA4 holds. Further, since the KL divergence is a continuous function of the means and variances of the price distributions, and since these mean and variance parameters are continuous functions of $\alpha_i, \alpha_{-i}, \hat{r}, \sigma_\zeta^2$, the existence of the absolutely integrable dominating function h also proves K_A, K_B (as integrals of KL divergences across different s_i) are continuous, so Assumption OA5 holds. □

OA 2.11 Proof of Proposition 3

Proof. Find L_1, L_2, L_3 as given by Lemma 3. Suppose $\Theta_A = \Theta(\kappa^\bullet)$, $\Theta_B = \{F_{r^\bullet,\kappa,\sigma_\zeta^\bullet}\}$ for any $\kappa \in [0, 1]$, $(p_A, p_B) = (1, 0)$, and $\lambda \in [0, 1]$, then arguments similar to those in the proof of Lemma 3 imply there exists exactly one EZ, and it involves the adherents of Θ_A holding correct beliefs and playing $\frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$ against each other.

We now analyze $\alpha_{BA}(\kappa)$ in such EZ. In the proof of Proposition 1, we defined $\bar{U}_i(\alpha_i)$ as i's objective expected utility of choosing α_i when $-i$ plays the rational best response. We showed that $\bar{U}_i'(\frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}) > 0$. In an EZ where i believes in the parameter $F_{r^\bullet,\kappa,\sigma_\zeta^\bullet}$ and $-i$ believes in the parameter $F_{r^\bullet,\kappa^\bullet,\sigma_\zeta^\bullet}$, using the expression for α_i^{BR} from Lemma 2, the play of i solves $x = \frac{\gamma - \frac{1}{2}r^\bullet\psi(\kappa)\left(\frac{\gamma-\frac{1}{2}r^\bullet\psi(\kappa^\bullet)x}{1+r^\bullet}\right)}{1+r^\bullet}$, which implies $\alpha_{BA}(\kappa) = \frac{\gamma(1+r^\bullet-\frac{1}{2}\psi(\kappa)r^\bullet)}{1+2r^\bullet+(r^\bullet)^2-\frac{1}{4}\psi(\kappa)\psi(\kappa^\bullet)(r^\bullet)^2}$. Taking the derivative and evaluating at $\kappa = \kappa^\bullet$, we find an expression with the same sign as $\frac{1}{4}\psi'(\kappa^\bullet)r^\bullet(1+r^\bullet)\gamma(-2(1+r^\bullet)+\psi(\kappa^\bullet)r^\bullet)$, which is strictly negative because $\psi'(\kappa^\bullet) > 0$, $r^\bullet > 0$, $\gamma > 0$, and $\psi(\kappa^\bullet) \le 1$. This shows there exists $\epsilon > 0$ so that for every $\kappa_h \in (\kappa^\bullet, \kappa^\bullet + \epsilon]$, we have $\bar{U}_i(\alpha_{BA}(\kappa_h)) < \bar{U}_i(\frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)})$, that is the adherents of $\{F_{r^\bullet,\kappa_h,\sigma_\zeta^\bullet}\}$ have strictly lower fitness than the adherents of $\Theta(\kappa^\bullet)$ with $\lambda = 0$ in the unique EZ. Finally, existence and upper-hemicontinuity of EZ in population proportion in such societies can be established using arguments similar to the proof of Propositions OA2 and OA3. This establishes the first claim to be proved.

Next, we turn to $\alpha_{BB}(\kappa)$. Using the expressing for α_i^{BR} in Lemma 2, we find that $\alpha_{BB}(\kappa) = \frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa)}$. Since $\psi' > 0$, we have $\alpha_{BB}(\kappa)$ is strictly larger than $\alpha_{AA} = \frac{\gamma}{1+r^\bullet+\frac{1}{2}r^\bullet\psi(\kappa^\bullet)}$ when $\kappa < \kappa^\bullet$. From the proof of Proposition 2, we know that objective payoffs in the stage game is strictly decreasing in linear strategies larger than the team solution $\alpha_{TEAM} = \frac{\gamma}{1+r^\bullet+r^\bullet\psi(\kappa^\bullet)}$. Since $\alpha_{BB}(\kappa) > \alpha_{AA} > \alpha_{TEAM}$, we conclude the adherents of $\{F_{r^\bullet,\kappa_l,\sigma_\zeta^\bullet}\}$ have strictly lower fitness than the adherents of $\Theta(\kappa^\bullet)$ with $\lambda = 1$ in the unique EZ, for any $\kappa_l < \kappa^\bullet$. Again , existence and upper-hemicontinuity of EZ in population proportion in such societies can be established using arguments similar to the proof of Propositions OA2 and OA3. This establishes the second claim to be proved. □

OA 2.12 Proof of Proposition OA1

Proof. Consider the society where $\Theta_A = \Theta_B = \Theta(\kappa^\bullet)$, $(p_A, p_B) = (1, 0)$. For any EZ with behavior $(\sigma_{AA}, \sigma_{AB}, \sigma_{BA}, \sigma_{BB})$ and beliefs (μ_A, μ_B), there exists another EZ $(\sigma'_{AA}, \sigma'_{AB}, \sigma'_{BA}, \sigma'_{BB})$ where $\sigma'_{g,g'} = \sigma_{AA}$ for all $g, g' \in \{A, B\}$ and all agents hold the belief μ_A. The uniqueness of EZ from Assumption OA1 implies $\alpha_{AB}(\kappa^\bullet) = \alpha_{BA}(\kappa^\bullet) = \alpha_{BB}(\kappa^\bullet) = \alpha^\bullet$.

Now consider the society where $\Theta_B = \Theta(\kappa)$, $(p_A, p_B) = (1, 0)$. By the same arguments as the existence arguments in Proposition OA2, there exists an EZ where $\alpha_{AA}(\kappa) = \alpha_{AA}(\kappa^\bullet)$. By the uniqueness of EZ from Assumption OA1, we must in fact have $\alpha_{AA}(\kappa) = \alpha_{AA}(\kappa^\bullet)$ for all κ, so the fitness of model $\Theta(\kappa^\bullet)$ in the unique EZ is

$$\mathbb{E}^\bullet\left[\mathbb{E}^\bullet\left[u_1^\bullet(\alpha^\bullet s_1, \alpha^\bullet s_2, \omega) \mid s_1\right]\right].$$

Under λ matching with mutant model $\Theta(\kappa)$, the mutant's fitness in the unique EZ is

$$\mathbb{E}^\bullet\left[\mathbb{E}^\bullet\left[(1-\lambda)u_1^\bullet(\alpha_{BA}(\kappa)s_1, \alpha_{AB}(\kappa)s_2, \omega) + (\lambda)u_1^\bullet(\alpha_{BB}(\kappa)s_1, \alpha_{BB}(\kappa)s_2, \omega) \mid s_1\right]\right].$$

Differentiate and evaluate at $\kappa = \kappa^\bullet$. At $\kappa = \kappa^\bullet$, adherents of Θ_A and Θ_B have the same fitness since they play the same strategies. So, a non-zero sign on the derivative would give the desired evolutionary fragility against either models with slightly higher or slightly lower κ. This derivative is:

$$\mathbb{E}^\bullet\left[\mathbb{E}^\bullet\left[\begin{array}{c} \frac{\partial u_1^\bullet}{\partial q_1}(\alpha^\bullet s_1, \alpha^\bullet s_2, \omega) \cdot [(1-\lambda)\alpha'_{BA}(\kappa^\bullet) + \lambda\alpha'_{BB}(\kappa^\bullet)] \cdot s_1 \\ +\frac{\partial u_1^\bullet}{\partial q_2}(\alpha^\bullet s_1, \alpha^\bullet s_2, \omega) \cdot [(1-\lambda)\alpha'_{AB}(\kappa^\bullet) + \lambda\alpha'_{BB}(\kappa^\bullet)] \cdot s_2 \end{array} \middle| s_1\right]\right].$$

Using the interim optimality part of Assumption OA1, $\mathbb{E}^{\bullet}\left[\frac{\partial u_1^{\bullet}}{\partial q_1}(\alpha^{\bullet}s_1, \alpha^{\bullet}s_2, \omega) \mid s_1\right] = 0$ for every $s_1 \in S$, using the necessity of the first-order condition. The derivative thus simplifies as claimed. □

OA 2.13 Proof of Proposition A.1

Proof. When $\Theta_A = \Theta_B = \Theta^{\bullet}$, for any matching assortativity λ and with $(p_A, p_B) = (1, 0)$, 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 $\Theta_A = \Theta^{\bullet}$, $\Theta_B = \Theta^{An}$. Suppose $\lambda \in [0, 1]$ and let $p_B \in (0, 1)$. We claim there is an EZ where $d_{AA}^k = 1$ for every k, $d_{AB}^k = 0$ for every even k with $k < K$, $d_{AB}^k = 1$ for every other k, $d_{BA}^k = 0$ for every odd k and $d_{BA}^k = 1$ for every even k, and $d_{BB}^k = 0$ for every k with $k < K$, $d_{BB}^K = 1$. 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 $\hat{d}_{AB}^k = 2/K$ for k even, $\hat{d}_{AB}^k = 1$ 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 $\hat{d}_{AB}^k = x$ 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 0, whereas the adherents of B have fitness 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 $p_B \to 1$, with $\lambda < 1$, we find an EZ where adherents of B have fitness 0, adherents of A have fitness $(1-\lambda)\cdot\frac{1}{2}\cdot((K/2)g+\ell) > 0$. This shows Θ_B is not evolutionarily stable against Θ_A. □

OA 2.14 Proof of Proposition A.2

Proof. In the centipede game, suppose $g > \frac{2}{K-2}\ell$. 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. □

OA 2.15 Proof of Proposition A.3

Proof. In the $\overline{\Theta}^{An}$ vs. $\overline{\Theta}^{An}$ match, the adherents of $\overline{\Theta}^{An}$ hold the belief that $\hat{d}^k_{BB} = 2/K$ for every even k. In the role of P1, at node k for $k \le K-3$, stopping gives them k but continuing gives them a $(K-2)/K$ chance to get at least $k+2$, and we have $k \le \frac{K-2}{K}(k+2) \iff 2k \le 2K-4 \iff k \le K-2$. 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 $K \ge 6$.

In the $\overline{\Theta}^{\bullet}$ vs. $\overline{\Theta}^{An}$ 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 $\overline{\Theta}^{\bullet}$ is $p\cdot 0.5+(1-p)\cdot(\frac{1}{2}(K-1)+\frac{1}{2}K)$, while the fitness of $\overline{\Theta}^{An}$ is $p\cdot 0+(1-p)\cdot(\frac{1}{2}\cdot 0+\frac{1}{2}K)$. For any p, the fitness of $\overline{\Theta}^{\bullet}$ is strictly higher than that of $\overline{\Theta}^{An}$. $\square$

OA 3 Existence and Continuity of EZ

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 learning channel 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 interaction structures.

Let two models, Θ_A, Θ_B be fixed. Also fix population shares p and matching assortativity λ. Let $U_A : \mathbb{A}^2 \times \Theta_A \to \mathbb{R}$ be such that $U_A(a_i, a_{-i}; F) = U_i(a_i, a_{-i}; \delta_F)$ and let $U_B : \mathbb{A}^2 \times \Theta_B \to \mathbb{R}$ be such that $U_B(a_i, a_{-i}; F) = U_i(a_i, a_{-i}; \delta_F)$.

Assumption OA2. *$\mathbb{A}, \Theta_A, \Theta_B$ are compact metrizable spaces.*

Assumption OA3. *U_A, U_B are continuous.*

Assumption OA4. *For every $F \in \Theta_A \cup \Theta_B$ and $a_i, a_{-i} \in \mathbb{A}$, $K(F; a_i, a_{-i})$ is well-defined and finite.*

Under Assumption OA4, we have the well-defined functions $K_A : \Theta_A \times \mathbb{A}^2 \to \mathbb{R}_+$ and $K_B : \Theta_B \times \mathbb{A}^2 \to \mathbb{R}_+$, where $K_g(F; a_i, a_{-i}) := D_{KL}(F^{\bullet}(a_i, a_{-i}) \parallel F(a_i, a_{-i}))$.

Assumption OA5. *K_A and K_B are continuous.*

Assumption OA6. $\mathbb{A}$ *is convex and, for all* $a_{-i} \in \mathbb{A}$ *and* $\mu \in \Delta(\Theta_A) \cup \Delta(\Theta_B)$, $a_i \mapsto U_i(a_i, a_{-i}; \mu)$ *is quasiconcave.*

We show existence of EZ 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 OA1. *For* $g \in \{A, B\}$, $a = (a_{AA}, a_{AB}, a_{BA}, a_{BB}) \in \mathbb{A}^4$, *and* $0 \leq m_g \leq 1$, *let*

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

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 matching assortativity. This leads to the existence result.

Proposition OA2. *Under Assumptions OA2, OA3, OA4, OA5, and OA6, an EZ exists.*

Next, upper hemicontinuity in m_g in Lemma OA1 allows us to deduce the upper hemicontinuity of the EZ correspondence in population shares.

Proposition OA3. *Fix two models* Θ_A, Θ_B. *Also fix matching assortativity* $\lambda \in [0, 1]$. *The set of EZ is an upper hemicontinuous correspondence in* p_B *under Assumptions OA2, OA3, OA4, and OA5.*

OA 3.1 Proofs of Results in Appendix OA 3

OA 3.1.1 Proof of Lemma OA1

Proof. Write the minimization objective as

$$W(a, F, m_g) := m_g K_g(F; a_{g,g}, a_{g,g}) + (1 - m_g) K_g(F; a_{g,-g}, a_{-g,g}),$$

a continuous function of (a, F, m_g) by Assumption OA5. Suppose we have a sequence $(a^{(n)}, m_g^{(n)}) \to (a^*, m_g^*) \in \mathbb{A}^4 \times [0, 1]$ and let $F^{(n)} \in \Theta_g^*(a^{(n)}, m_g^{(n)})$ for each n, with $F^{(n)} \to F^* \in \Theta_g$. For any other $\hat{F} \in \Theta_g$, note that $W(a^*, m_g^*, \hat{F}) = \lim_{n \to \infty} W(a^{(n)}, m_g^{(n)}, \hat{F})$ by continuity. But also by continuity, $W(a^*, m_g^*, F^*) = \lim_{n \to \infty} W(a^{(n)}, m_g^{(n)}, F^{(n)})$ and $W(a^{(n)}, m_g^{(n)}, F^{(n)}) \leq W(a^{(n)}, m_g^{(n)}, \hat{F})$ for every n. It therefore follows $W(a^*, m_g^*, F^*) \leq W(a^*, m_g^*, \hat{F})$. □

OA 3.1.2 Proof of Proposition OA2

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

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

where $\mathrm{BR}(a_{-i}, \mu_g) := \underset{\hat{a}_i \in \mathbb{A}}{\arg\max} U_g(\hat{a}_i, a_{-i}; \mu_g)$ and, for each $g \in \{A, B\}$, the correspondence Θ_g^* is defined with $m_g = \lambda + (1-\lambda)p_g$, $m_{-g} = 1 - m_g$. It is clear that fixed points of Γ are EZ.

We apply the Kakutani-Fan-Glicksberg theorem (see, e.g, Corollary 17.55 in Aliprantis and Border (2006)). By Assumptions OA2 and OA6, $\mathbb{A}$ is acompact 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 OA3 supposes U_A, U_B are continuous, it implies the best-response correspondences $\mathrm{BR}(a_{AA}, \mu_A)$, $\mathrm{BR}(a_{BA}, \mu_A)$, $\mathrm{BR}(a_{AB}, \mu_B)$, $\mathrm{BR}(a_{BB}, \mu_B)$ have closed graphs.

To see that Γ is non-empty, recall that each $\hat{a}_i \mapsto U_g(\hat{a}_i, a_{-i}; \mu_g)$ 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 F over a compact domain of possible F's, 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, $\hat{a}_i \mapsto U_A(\hat{a}_i, a_{AA}; \mu_A)$ is quasiconcave by Assumption OA6. That means if $a_i^{'}, a_i^{''} \in \mathrm{BR}(a_{AA}, \mu_A)$, then for any convex combination $\tilde{a}_i$ of $a_i^{'}, a_i^{''}$, we have $U_A(\tilde{a}_i, a_{AA}; \mu_A) \geq \min(U_A(a_i^{'}, a_{AA}; \mu_A),\, U_A(a_i^{''}, a_{AA}; \mu_A)) = \max_{\hat{a}_i \in \mathbb{A}} U_A(\hat{a}_i, a_{AA}; \mu_A)$. Therefore, $\mathrm{BR}(a_{AA}, \mu_A)$ is convex. For similar reasons, $\mathrm{BR}(a_{BA}, \mu_A)$, $\mathrm{BR}(a_{AB}, \mu_B)$, $\mathrm{BR}(a_{BB}, \mu_B)$ are convex. $\square$

OA 3.1.3 Proof of Proposition OA3

Proof. Since $\mathbb{A}^4 \times \Delta(\Theta_A) \times \Delta(\Theta_B)$ is compact by Assumption OA2, 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 against $a^*_{g',g}$ under the belief μ_g^*. Assortativity does not matter here, since optimality applies within all type match-ups. By Assumption OA3, $U_g(a_i, a_{-i}; F)$ is continuous, so by property of convergence in distribution, $U_g(a^{(k)}_{g,g'}, a^{(k)}_{g',g}; \mu_g^{(k)}) \to U_g(a^*_{g,g'}, a^*_{g',g}; \mu_g^*)$. For any other $\hat{a}_i \in \mathbb{A}$, $U_g(\hat{a}_i, a^{(k)}_{g',g}; \mu_g^{(k)}) \to U_g(\hat{a}_i, a^*_{g',g}; \mu_g^*)$ and for every k, $U_g(a^{(k)}_{g,g'}, a^{(k)}_{g',g}; \mu_g^{(k)}) \geq U_g(\hat{a}_i, a^{(k)}_{g',g}; \mu_g^{(k)})$. Therefore $a^*_{g,g'}$ best responds to $a^*_{g',g}$ under belief μ_g^*.

Next, we show parameters in the support of μ_g^* minimize weighted KL divergence for group g. First consider the correspondence $H : \mathbb{A}^4 \times [0,1] \rightrightarrows \Theta_g$ where $H(a, p_g) := \Theta_g^*(a, \lambda + (1-\lambda)(p_g))$. Then H is upper hemicontinuous by Lemma OA1. Since $H(a, p_g)$ represents the minimizers of a continuous function on a compact domain, 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 $\tilde{H}(a, p_g) := \Delta(H(a, p_g))$ 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. □

OA 4 Learning Foundation of EZ and EZ-SU

We provide a unified foundation for EZ and EZ-SU as the steady state of a learning system. This foundation considers a world where agents have prior beliefs over extended parameters in an extended models, as in Appendix C. At the end of every match, each agent observes her consequence and a noisy signal about the matched opponent's strategy. We show that under any asymptotically myopic policy, if behavior and beliefs converge, then the limit steady state must be an EZ-SU when the noisy signals about opponent's strategy are uninformative. Sufficiently accurate signals about opponent's play cause the steady states to be EZs, if the extended models allow agents to make rich enough inferences about opponents' strategies. Finally, if the true situation is redrawn every T periods and the agents reset their beliefs over extended parameters to their prior belief when the situation is redrawn, then their average payoffs approach their fitness in the EZ or EZ-SU when T is large.

OA 4.1 Regularity Assumptions

We make some regularity assumptions on the objective environments and on the extended models $\overline{\Theta}_A, \overline{\Theta}_B$. These are similar to the regularity assumptions from Online Appendix OA 3.

Suppose the strategy set $\mathbb{A}$ is finite. Suppose the marginals of the extended models $\overline{\Theta}_A, \overline{\Theta}_B$ on the dimension of fundamental uncertainty, denoted as Θ_A, Θ_B, are compact and metrizable spaces. Endow $\overline{\Theta}_A$ and $\overline{\Theta}_B$ with the product metric. Suppose that every $(a_A, a_B, F) \in \overline{\Theta}_A \cup \overline{\Theta}_B$ is so 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.

For each $g, g^{'} \in \{A, B\}$, define $K_{g,g'} : \mathbb{A}^2 \times \mathcal{G} \times \overline{\Theta}_g \to \mathbb{R}$ by $K_{g,g'}(a_i, a_{-i}, G; (a_A, a_B, F)) = D_{KL}(F^\bullet(a_i, a_{-i}, G) \parallel F(a_i, a_{g'}))$. This is the KL divergence of the parameter $(a_A, a_B, F) \in \overline{\Theta}_g$ in situation G based on the data generated from the strategy profile (a_i, a_{-i}). Suppose each $K_{g,g'}$ is well defined and a continuous function of the extended parameter (a_A, a_B, F).

For $g \in \{A, B\}$, $F \in \Theta_g$, 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 extended model $\overline{\Theta}_g$ and every $(a_A, a_B, F) \in \overline{\Theta}_g$ and $\epsilon > 0$, there exists an open neighborhood $V \subseteq \overline{\Theta}_g$ of (a_A, a_B, F), so that for every $(\hat{a}_A, \hat{a}_B, \hat{F}) \in V$, $1 - \epsilon \leq f(a_i, a_A)(y)/\hat{f}(a_i, \hat{a}_A)(y) \leq 1 + \epsilon$ and $1 - \epsilon \leq f(a_i, a_B)(y)/\hat{f}(a_i, \hat{a}_B)(y) \leq 1 + \epsilon$ for all $a_i \in \mathbb{A}, y \in \mathbb{Y}$. Also suppose there is some $M > 0$ so that $\ln(f(a_i, a_A)(y))$ and $\ln(f(a_i, a_B)(y))$ are bounded in $[-M, M]$ for all $(a_A, a_B, F) \in \overline{\Theta}_g$, $a_i, a_{-i} \in \mathbb{A}, y \in \mathbb{Y}$.

OA 4.2 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 extended model, $\mu_g^{(0)} \in \Delta(\overline{\Theta}_g)$, and believes there is some $(a_A, a_B, F) \in \overline{\Theta}_g$ 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 partially assortatively to play the stage game. Assortativity is $\lambda \in (0, 1)$. Each person in group g has $\lambda + (1 - \lambda)p_g$ 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 signal $x_i^{(t)} \in \mathbb{A}$ about the opponent's play, where $x_i^{(t)}$ equals the matched opponent's strategy a_{-i} with probability $\tau \in [0, 1)$, and it is uniformly random on $\mathbb{A}$ with the complementary probability. To give a foundation for a EZ-SU, we consider $\tau = 0$, so the signal x_i is uninformative. To give a foundation for EZ, we consider τ close to 1.

Thus, the space of histories from one period is $\{A, B\} \times \mathbb{A} \times \mathbb{Y} \times \mathbb{A}$, with typical element $(g_i^{(t)}, a_i^{(t)}, y_i^{(t)}, x_i^{(t)})$. 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 ex-post signal about the matched opponent's play, $x_i^{(t)}$. 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, $\mu_g : \mathbb{H} \to \Delta(\overline{\Theta}_g)$, mapping every finite-length history into a belief over extended parameters in $\overline{\Theta}_g$, 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 OA7. *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 $\overline{\Theta}_g$, is a stochastic process $(\tilde{a}_{iA}^{(t)}, \tilde{a}_{iB}^{(t)}, \tilde{\mu}_i^{(t)})_{t \geq 0}$ valued in $\mathbb{A} \times \mathbb{A} \times \Delta(\overline{\Theta}_g)$. The randomness is over the groups of opponents matched with in different periods, the strategies they play, and the random consequences and ex-post signals drawn at the end of the matches. At the same time, 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 $(\alpha_{AA}^{(t)}, \alpha_{AB}^{(t)}, \alpha_{BA}^{(t)}, \alpha_{BA}^{(t)}, \nu_A^{(t)}, \nu_B^{(t)}) \in \Delta(\mathbb{A})^4 \times \Delta(\Delta(\overline{\Theta}_A)) \times \Delta(\Delta(\overline{\Theta}_B))$ that describes the distributions of play and beliefs that prevail in the two sub-populations in every period t.

OA 4.3 Steady State Limits are EZ-SUs and EZs

We state and prove the learning foundation of EZ-SU and EZ. 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 $\Delta(\Delta(\overline{\Theta}_g))$ and $\mu^* \in \Delta(\overline{\Theta}_g)$, $\nu^{(t)} \to \mu^*$ means $\mathbb{E}_{\hat{\mu}\sim\nu^{(t)}} \parallel \hat{\mu} - \mu^* \parallel \to 0$ as $t \to \infty$.

Proposition OA4. *Suppose the regularity assumptions in Online Appendix OA 4.1 hold, and suppose Assumption OA7 holds.*

Suppose $\tau = 0$. Suppose there exists $(a^*_{AA}, a^*_{AB}, a^*_{BA}, a^*_{BB}, \mu^*_A, \mu^*_B) \in \mathbb{A}^4 \times \Delta(\overline{\Theta}_A) \times \Delta(\overline{\Theta}_B)$ so that $(\alpha^{(t)}_{AA}, \alpha^{(t)}_{AB}, \alpha^{(t)}_{BA}, \alpha^{(t)}_{BA}, \nu^{(t)}_A, \nu^{(t)}_B) \to (a^*_{AA}, a^*_{AB}, a^*_{BA}, a^*_{BB}, \mu^*_A, \mu^*_B)$ and for each agent i in group g, almost surely $(\tilde{a}^{(t)}_{iA}, \tilde{a}^{(t)}_{iB}, \tilde{\mu}^{(t)}_i) \to (a^*_{gA}, a^*_{gB}, \mu^*_g)$. Then, $(a^*_{AA}, a^*_{AB}, a^*_{BA}, a^*_{BB}, \mu^*_A, \mu^*_B)$ is an EZ-SU.

Suppose for each g, the extended model $\overline{\Theta}_g = \mathbb{A}^2 \times \Theta_g$ for some model Θ_g – that is, each group can make any inference about opponents' strategies. 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}), and $(a^*_{AA}, a^*_{AB}, a^*_{BA}, a^*_{BB}, \mu^*_A|_{\Theta_A}, \mu^*_B|_{\Theta_B})$ is an EZ, where $\mu^*_g|_{\Theta_g}$ is the marginal of the belief μ^*_g on the model Θ_g.

Proof. We first consider the case of $\tau = 0$, so the uninformative ex-post signals may be ignored.

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 Θ_A is bounded), there are some $\epsilon_1, \epsilon_2 > 0$ so that whenever $\mu_i \in \Delta(\overline{\Theta}_A)$ with $\parallel \mu_i - \mu^*_A \parallel < \epsilon_1$, 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 $\tilde{\mu}^{(t)}_i \to \mu^*_A$. This shows that by asymptotically myopic best responses, almost surely $\tilde{a}^{(k)}_{iA} \not\to a^*_{AA}$, 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-SU (the case of a parameter θ^*_B in the support of μ^*_B not minimizing is similar). Then we have

$$\theta^*_A \notin \underset{\hat{\theta}\in\overline{\Theta}_A}{\text{argmin}} \begin{bmatrix} (\lambda + (1-\lambda)p_A) \cdot D_{KL}(F^\bullet(a^*_{AA}, a^*_{AA}) \parallel \hat{F}(a^*_{AA}, \hat{a}_A)) \\ +(1-\lambda)(1-p_A) \cdot D_{KL}(F^\bullet(a^*_{AB}, a^*_{BA}) \parallel \hat{F}(a^*_{AB}, \hat{a}_B)) \end{bmatrix}$$

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

This is equivalent to:

$$\theta_A^* \notin \underset{\hat{\theta} \in \overline{\Theta}_A}{\operatorname{argmax}} \left[\begin{array}{c} (\lambda + (1-\lambda)p_A) \cdot \mathbb{E}_{y \sim F^\bullet(a_{AA}^*, a_{AA}^*)} \ln(\hat{f}(a_{AA}^*, \hat{a}_A)(y)) \\ +(1-\lambda)(1-p_A) \cdot \mathbb{E}_{y \sim F^\bullet(a_{AB}^*, a_{BA}^*)} \ln(\hat{f}(a_{AB}^*, \hat{a}_B)(y)) \end{array} \right]$$

Let this objective, as a function of $\hat{\theta}$, be denoted $WL(\hat{\theta})$. There exists $\theta_A^{opt} = (a_A^{opt}, a_B^{opt}, f^{opt}) \in \overline{\Theta}_A$ and $\delta, \epsilon > 0$ so that $(1-\delta)WL(\theta_A^{opt}) - 2\delta M - 3\epsilon > (1-\delta)WL(\theta_A^*)$. By assumption on the primitives, find open neighborhoods V^{opt} and V^* of $\theta_A^{opt}, \theta_A^*$ respectively, so that for all $a_i \in \mathbb{A}$, $g \in \{A, B\}$, $y \in \mathbb{Y}$, $1 - \epsilon \leq f^{opt}(a_i, a_g^{opt})(y)/\hat{f}(a_i, \hat{a}_g)(y) \leq 1 + \epsilon$, for all $\hat{\theta} = (\hat{a}_A, \hat{a}_B, \hat{f}) \in V^{opt}$, and also $1 - \epsilon \leq f^*(a_i, a_g^*)(y)/\hat{f}(a_i, \hat{a}_g)(y) \leq 1 + \epsilon$ 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 $t \geq T_1$, $\alpha_{AA}^{(t)}(a_{AA}^*) \geq 1 - \delta$ and $\alpha_{BA}^{(t)}(a_{BA}^*) \geq 1 - \delta$.

For $T_2 \geq T_1$, consider a probability space defined by $\Omega := (\{A, B\} \times \mathbb{A}^2 \times (\mathbb{Y})^{\mathbb{A}^2})^\infty$ that describes the randomness in an agent's learning process starting with period $T_2 + 1$. For a point $\omega \in \Omega$ and each period $T_2 + s$, $s \geq 1$, $\omega_s = (g, a_{-i,A}, a_{-i,B}, (y_{a_i,a_{-i}})_{(a_i,a_{-i}) \in \mathbb{A}^2})$ specifies the group g of the matched opponent, the play $a_{-i,A}, a_{-i,B}$ of hypothetical opponents from groups A and B, and the hypothetical consequence $y_{a_i,a_{-i}}$ that would be generated for every pair of strategies (a_i, a_{-i}) played. As notation, let $opp(\omega, s)$, $a_{-i,A}(\omega, s)$, $a_{-i,B}(\omega, s)$, and $y_{a_i,a_{-i}}(\omega, s)$ denote the corresponding components of ω_s. Define $\mathbb{P}_{T_2}$ over this space 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 $\omega_s = (g, a_{-i,A}, a_{-i,B}, (y_{a_i,a_{-i}})_{(a_i,a_{-i}) \in \mathbb{A}^2})$ is

$$m_g \cdot \alpha_{AA}^{(T_2+s)}(a_{-i,A}) \alpha_{BA}^{(T_2+s)}(a_{-i,B}) \cdot \prod_{(a_i,a_{-i}) \in \mathbb{A}^2} f^\bullet(a_i, a_{-i})(y_{a_i,a_{-i}}),$$

where m_g is the probability of i from group A being matched up against an opponent of group g, that is $m_A = (\lambda + (1-\lambda)p_A)$, $m_B = (1-\lambda)(1-p_A)$.

For $\theta = (a_A^\theta, a_B^\theta, F^\theta) \in \overline{\Theta}_A$ with f^θ the density of F^θ, $\omega \in \Omega$, consider the stochastic process

$$\ell_s(\theta, \omega) := \frac{1}{s} \sum_{t=T_2+1}^{T_2+s} \ln(f^\theta(a_{AA}^*, a_{opp(\omega,t)}^\theta)(y_{a_{AA}^*, a_{-i,opp(\omega,t)}(\omega,t)}(\omega, t)).$$

By choice of the neighborhood V^*,

$$\limsup_s \sup_{\theta_A \in V^*} \ell_s(\theta_A,\omega) \leq \epsilon + \frac{1}{s}\sum_{t=T_2+1}^{T_2+s} \ln(f^*(a^*_{AA}, a^*_{opp(\omega,t)})(y_{a^*_{AA},a_{-i,opp(\omega,t)}(\omega,t)}(\omega,t))$$
$$\leq \epsilon + \frac{1}{s}\sum_{t=T_2+1}^{T_2+s} \begin{array}{c} 1_{\{a_{-i,opp(\omega,t)}(\omega,t)=a^*_{opp(\omega,t),A}\}} \cdot \ln(f^*(a^*_{AA}, a^*_{opp(\omega,t)})(y_{a^*_{AA},a^*_{opp(\omega,t),A}}(\omega,t)) \\ (1 - 1_{\{a_{-i,opp(\omega,t)}(\omega,t)=a^*_{opp(\omega,t),A}\}}) \cdot M. \end{array}$$

Since $T_2 \geq T_1$, in every period t, $\mathbb{P}_{T_2}(a_{-i,opp(\omega,t)}(\omega,t) = a^*_{opp(\omega,t),A}) \geq 1-\delta$. 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 M, $(1-\delta)m_A$ chance to be distributed according to $\ln(f^*(a^*_{AA}, a^*_A)(y))$ where $y \sim f^\bullet(a^*_{AA}, a^*_{AA})$, and $(1-\delta)m_B$ chance to be distributed according to $\ln(f^*(a^*_{AB}, a^*_B)(y))$ where $y \sim f^\bullet(a^*_{AB}, a^*_{BA})$. By law of large numbers, $\frac{1}{s}\sum_{k=1}^{s} \xi_k$ converges almost surely to $\delta M + (1-\delta)WL(\theta^*_A)$. By this comparison, $\limsup_s \sup_{\theta_A \in V^*} \ell_s(\theta_A,\omega) \leq \epsilon + \delta M + (1-\delta)WL(\theta^*_A)$ $\mathbb{P}_{T_2}$-almost surely. By a similar argument, $\liminf_s \inf_{\theta_A \in V^{opt}} \ell_s(\theta_A,\omega) \geq -\epsilon - \delta M + (1-\delta)WL(\theta^{opt}_A)$ $\mathbb{P}_{T_2}$-almost surely.

Along any ω where we have both $\limsup_s \sup_{\theta_A \in V^*} \ell_s(\theta_A,\omega) \leq \epsilon + \delta M + (1-\delta)WL(\theta^*_A)$ and $\liminf_s \inf_{\theta_A \in V^{opt}} \ell_s(\theta_A,\omega) \geq -\epsilon - \delta M + (1-\delta)WL(\theta^{opt}_A)$, 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 $\hat{\Omega}_h \subseteq \Omega$ that leads to i not playing the EZ-SU 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-SU 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 consider the foundation for EZs. Suppose Let $\bar{K} < \infty$ be an upper bound on $K_{g,g'}(a_i, a_{-i}; (a_A, a_B, F))$ across all $g, g' \in \{A, B\}$, $a_i, a_{-i} \in \mathbb{A}$, $(a_A, a_B, F) \in \overline{\Theta}_g$. Here $\bar{K}$ is finite because $\mathbb{A}$ is finite and $K_{g,g'}$ is continuous in the extended parameter, which is from a compact domain. Let $F^X_\tau(a_{-i}) \in \Delta(\mathbb{A})$ represent the distribution of ex-post signals given precision τ, when opponent plays $a_{-i} \in \mathbb{A}$. 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 $\min(m_A, m_B) \cdot D_{KL}(F^X_\tau(a_{-i}) \parallel F^X_\tau(a'_{-i})) > \bar{K}$. Therefore, given any $(a^*_{AA}, a^*_{AB}, a^*_{BA}) \in \mathbb{A}^3$, the solution to

$$\min_{\hat{\theta} \in \overline{\Theta}_A} \left[\begin{array}{c} (\lambda + (1-\lambda)p_A) \cdot [D_{KL}(F^{\bullet}(a^*_{AA}, a^*_{AA}) \parallel \hat{F}(a^*_{AA}, \hat{a}_A)) + D_{KL}(F^X_\tau(a^*_{AA}) \parallel F^X_\tau(\hat{a}_A))] \\ +(1-\lambda)(1-p_A) \cdot [D_{KL}(F^{\bullet}(a^*_{AB}, a^*_{BA}) \parallel \hat{F}(a^*_{AB}, \hat{a}_B)) + D_{KL}(F^X_\tau(a^*_{BA}) \parallel F^X_\tau(\hat{a}_B)] \end{array} \right]$$

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. The rest of the argument is similar to the case of EZ-SU. □

OA 4.4 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 their group's extended model, $\mu^{(0)}_g \in \Delta(\overline{\Theta}_g)$. In periods $T + 1, 2T + 1, ...$ agents reset their belief to $\mu^{(0)}_g$, and their belief in each period over the extended parameters in their extended model only use histories since the last reset. This belief corresponds to agents thinking that the data-generating process is redrawn according to $\mu^{(0)}_g$ every T periods.

Suppose $\tau = 0$ and suppose for every $G \in \mathcal{G}$, the hypotheses of Proposition OA4 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-SU 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-SU as the agents' objective payoffs when the true situation changes sufficiently slowly (a similar foundation applies for the fitness in EZ.)

OA 5 The Single-Agent Case

This section records an observation related to our stability concepts when applied to the single-agent case. Specifically, situation G is a *decision problem* if $(a_i, a_{-i}) \mapsto F^\bullet(a_i, a_{-i}, G)$ only depends on a_i. If every situation is a decision problem, then the correctly specified model is evolutionarily stable against *any* other model, except when there are identification issues. We adapt the notion of strong identification from Esponda and Pouzo (2016).

Definition OA1. Model Θ_A is *strongly identified* in EZ $\mathfrak{Z} = (\mu_A(G), \mu_B(G), p, \lambda, a(G))_{G\in\mathcal{G}}$ if in every situation G, whenever $F', F'' \in \Theta_A$ both solve

$$\min_{F\in\Theta_A} \left\{(\lambda + (1-\lambda)p_A) \cdot K(F; a_{AA}, a_{AA}, G) + (1-\lambda)(1-p_A) \cdot K(F; a_{AB}, a_{BA}, G)\right\},$$

we have $F^{'}(a_i, a_{AA}) = F^{''}(a_i, a_{AA})$ and $F^{'}(a_i, a_{BA}) = F^{''}(a_i, a_{BA})$ for all $a_i \in \mathbb{A}$.

Proposition OA5. *Suppose every situation is a decision problem. Let λ and two models Θ_A, Θ_B be given, where Θ_A is correctly specified. Suppose there exists at least one EZ with $p_A = 1$, and Θ_A is strongly identified in all such equilibria. Then Θ_A evolutionarily stable under λ-matching against Θ_B.*

Proof. In any EZ, let $F \in \operatorname{supp}(\mu_A(G))$ and note that $F^\bullet(\cdot,\cdot,G) \in \Theta_A$ since Θ_A is correctly specified. Both F and $F^\bullet(\cdot,\cdot,G)$ solve the weighted minimization problem, the former because it is in the support of μ_A, the latter because it attains the lowest minimization objective of 0. By strong identification, the set of best responses to $a_{AA}(G)$ and $a_{BA}(G)$ under the belief μ_A is the same as set of actions that maximize payoffs in the decision problem given by $F^\bullet(\cdot,\cdot,G)$. Therefore, adherents of Θ_A obtain the highest possible objective payoffs in the stage game in situation G. This applies to every situation, so Θ_A has weakly higher fitness than Θ_B in the EZ. □

The result that a resident correct specification is immune to invasions from misspecifications echoes related results in Fudenberg and Lanzani (2022) and Frick, Iijima, and Ishii (2021). We primarily focus on stage games where multiple agents' actions jointly determine their payoffs and characterize which misspecifications can invade a rational society in which environments.