Which misspecifications persist?

Drew Fudenberg and Giacomo Lanzani*

First posted version: October 12, 2020
This version: January 26, 2023

Abstract

We use an evolutionary model to determine which misperceptions can persist. Every period, a new generation of agents use their subjective models and the data generated by the previous generation to update their beliefs, and models that induce better actions become more prevalent. An equilibrium can resist mutations that lead agents to use a model that better fits the equilibrium data but induce the mutated agents to take an action with lower payoffs. We characterize which steady states resist mutations to a nearby model, and which resist mutations that drop a qualitative restriction such as independence.

Keywords: Misspecified learning, Berk-Nash equilibrium, evolution, payoff monotone dynamics

*Department of Economics, MIT. We thank Pierpaolo Battigalli, Renee Bowen, Daniel Clark, Roberto Corrao, Tristan Gagnon-Bartsch, Francesca Galbiati, Ying Gao, Kevin He, Elliot Lipnowski, Pooya Molavi, Stephen Morris, Frank Schilbach, Josh Schwartzstein, Philipp Strack, Takuo Sugaya, Dmitry Taubinsky, Jorgen Weibull, and Alex Wolitzky for helpful conversations, and NSF grants SES 1643517 and 1951056 and the Guido Cazzavillan Scholarship for financial support.

1 Introduction

Economic agents are often misspecified in the sense that their prior beliefs rule out the data generating process they actually face. This misspecification may have different roots: Agents may have a behavioral bias such as overconfidence and correlation neglect, or they may over-simplify a complex environment by omitting some relevant variables or interactions, or by positing an incorrect functional form. Many of these misspecifications have important consequences for behavior. For example, when agents misperceive a progressive tax schedule as linear they end up working too much, since they equate their marginal cost of effort to the average instead of marginal tax rate, and when buyers misperceive price and quality as independent they may bid prices that are too low.

We study the effect of mutations in an evolutionary model where models that induce higher-payoff actions become more prevalent. In our model, the agents face single-agent decision problems where their optimal action depends on some parameters of the outcome-generating function. Each generation, agents estimate the parameters of their subjective model (or "paradigm") that best fit the data generated by the actions and outcomes of the previous generation.[1] The agents then choose a best reply to a belief that is concentrated on these best-fitting parameters. Steady states in which all agents have the same model coincide with Berk-Nash equilibria (Esponda and Pouzo, 2016): The actions played are a best response to posterior beliefs that fit the equilibrium data as well as the model allows.

A purely Bayesian agent can never come to assign positive probability to a data generating process that lies outside the support of their subjective model, so the standard Bayesian model predicts that all misspecifications will persist forever. Instead, we consider what happens when a small fraction of agents adopts an expanded subjective model: Will the equilibrium resist the mutation, in the sense that the original behavior persists?

Our analysis uncovers a few general insights. The first is that there are two different pathways by which a mutation can destabilize an equilibrium misperception. Under the direct channel, a better explanation of the data may lead the mutants to adopt a better action, so their share of the population grows. However, whether the actions of the mutants are better is determined by the inference they draw from the equilibrium data. Several examples show that unless the new model is correctly specified, this need not be the case. Under the indirect channel, the mutants might obtain the same or lower payoff as agents using the prevailing model, but the information their actions generate helps agents with the

[1] "Best fit" here means maximizing the likelihood of the data.

old model realize they can increase their payoff with another action. We also found that the ability of a mixed equilibrium to resist mutations depends on the mixing probabilities, which may not not be pinned down by the equilibrium conditions. This effect creates an effective bound on the inefficiency induced by a bias to persist.

Now we give a more detailed summary of our analysis. A first observation is that every steady state resists any mutation that does not provide a better explanation of the equilibrium data. For that reason, every self-confirming equilibrium resists all mutations, even though the agents may misperceive the consequences of some non-equilibrium actions. In contrast, if an equilibrium relies on misspecified beliefs about the consequences of equilibrium actions, it can be overturned by mutations. As noted above, there are two ways this can occur, the direct channel and the indirect one. The effectiveness of these channels depends both on the nature of the equilibrium and that of the mutations. In a *uniformly strict* Berk-Nash equilibrium (Fudenberg, Lanzani, and Strack, 2021), the action played is the unique best reply to all parameters that minimize the Kullback-Leibler (henceforth "KL") divergence from the equilibrium data. These equilibria resist *local mutations*, where agents add nearby parameters to their subjective models when these better explain their data.[2] Whether a Berk-Nash equilibrium that is not uniformly strict resists local mutations depends on the actions the mutations induce and their associated payoffs. We show that the effectiveness of the direct channel depends on the payoff of the best responses that remain optimal against the nearby parameters that most improve the fit to the equilibrium data. If and only if all these "local responses" give an objectively higher payoff, the mutant will become more prevalent, and the equilibrium does not resist local mutations.

Even though all uniformly strict equilibria resist local mutations, some of them do not resist mutations to a paradigm with a more general structure. We model this with the idea of *one-hypothesis mutations.* Here the agent's paradigm consists of a finite set of assumptions about the data generating process, and mutations relax one assumption while maintaining the others. This large change in paradigm can lead agents to take new actions with higher payoffs. However, some equilibria resist one-hypothesis mutations and not local ones, because the one-hypothesis relaxation can lead to over-adjustment in the direction of the relaxed constraint, and thus to overshooting the optimal action. We characterize resistance to one-hypothesis mutations by considering the KL minimizers in relaxed subjective models where one of the hypotheses is dropped. Specifically, we show that a uniformly strict equilibrium

[2] This is similar to the neighborhoods of a subjective model considered in the macroeconomics literature on robust control following Hansen and Sargent (2008).

is resistant if and only if the KL minimizers of the relaxed problem induce an action that yields less than the equilibrium payoff.

In equilibria that are not uniformly strict, there may be an unused action that is a best response to the KL minimizers. Here the indirect channel can operate, because the action induced by the mutation can provide evidence that leads agents with the old paradigm to change to an action with higher payoffs. In this case the misperception does not persist even though the mutants may receive a lower payoff.

In models of correlation neglect, misspecified beliefs are less resistant to mutations in "noisy" environments, because the noise helps the agents correctly infer the correlation between the variables. Since in our setting the distribution of an initial signal is equivalent to a distribution over heterogeneous preferences, this suggests that a homogeneous closed group of agents is more likely to maintain misspecified beliefs: If the agents share a subjective model but have different preferences, in equilibrium they can play different best replies. Thus the KL minimizers for a new subjective model will reflect the consequences of multiple actions, which makes it less likely that the adjustment induced by the mutation is detrimental.

We show that the continuum-of-agents dynamic process we study corresponds to the limit of finite-population processes as the number of agents goes to infinity. Finally, we consider settings with a continuum of actions. We explain how the conditions for resisting mutations need to be modified to account for the fact that small changes in paradigm can induce different actions even when the original equilibrium action was the unique best reply, and how the modified definition relates to the limits of finer and finer action grids.

1.1 Related work

Berk (1966) shows that the misspecified beliefs asymptotically concentrate on the models that minimize the KL divergence from the objective data generating process when this process is exogenous. In many economic applications, actions and associated signal distributions aren't fixed, but depend on agents' actions, so misspecification has implications for what the agents observe and thus for their long-run beliefs. Arrow and Green (1973) gives the first general framework for this problem, and Nyarko (1991) points out that the combination of misspecification and endogenous data can lead to cycles. Esponda and Pouzo (2016) defines Berk-Nash equilibrium, which relaxes Nash equilibrium by replacing the requirement that player's beliefs are correct with the requirement that each player's belief minimizes the KL divergence of their observations from their subjective model.[3]

[3] Jehiel (2020) surveys various equilibrium concepts for misspecified agents.

Esponda, Pouzo, and Yamamoto (2021) uses stochastic approximation to establish when the agent's action frequency converges. Frick, Iijima, and Ishii (2021a) provides conditions for local and global convergence of the agent's beliefs without explicitly modeling the agent's actions. Fudenberg, Lanzani, and Strack (2021) introduces uniform Berk-Nash equilibria and uniformly strict Berk-Nash equilibria. It shows that uniform Berk-Nash equilibria are the only possible limit actions, and that uniformly strict Berk-Nash equilibria are the only stable equilibria. Bohren and Hauser (2021) characterize the long-run beliefs of a sequence of heterogeneous misspecified agents, and Bowen, Dmitriev, and Galperti (2022) studies polarization of misspecified agents on a network.[4]

Gagnon-Bartsch, Rabin, and Schwartzstein (2021) proposes that agents only pay attention to events they believe are payoff-relevant, and that an agent whose model is wrong about the probability of one of these events may switch to a model that includes the objective data generating process. It assumes actions do not influence the distribution over outcomes, so the issues that we address do not arise. He and Libgober (2021) studies competition between two models in a game setting, where even correctly specified models can be out-performed by some mutants. The inference in their model does not depend on data that was generated before the mutation. Massari and Newton (2022) justifies a generalized Bayes rule as the result of evolutionary competition between different updating rules. This extends to a single misspecified decision maker the failure of Bayesian updating that Blume and Easley (1992) obtained when correctly specified agents trade with "irrational" ones. Grant and Quiggin (2017) studies how the evolutionarily stable profiles of a two-player game change when some agents in one population become aware of additional strategies.[5]

Esponda, Vespa, and Yuksel (2022) shows that when given unexplained evidence, misspecified agents make small adjustments but do not typically include the correct model. It also shows that agents use the effects of one action to help predict the consequences of others. This extrapolation is at the core of why some misspecifications can persist: A model that better fits the equilibrium data may lead agents to switch to an action with lower payoffs. Our work is also related to models of agents who subject their models to tests for misspec-

[4]Fudenberg, Romanyuk, and Strack (2017), Heidhues, Kőszegi, and Strack (2018), and Molavi (2019) analyze misspecified learning in specific applications. Bohren (2016), Frick, Iijima, and Ishii (2020), and He (2022) consider misspecified social learning where all agents have the same misspecification. Like us, He (2022) considers a model where agents learn from the data generated by the previous generation.

[5]Also, Schwartzstein and Sunderam (2021) studies model selection in a setting without actions or payoffs, Montiel Olea, Ortoleva, Pai, and Prat (2021) studies how the bids of misspecified agents vary with their priors, and Frick, Iijima, and Ishii (2021b) characterizes the efficiency of updating with incorrect likelihood functions on exogenous data.

ification, as in Fudenberg and Kreps (1994), Hong, Stein, and Yu (2007), Cho and Kasa (2017), Ba (2022), and Lanzani (2022).

2 The single-agent problem

Before developing our large-population model we introduce the single-agent problem.

2.1 Static model

Actions, utility, and data generating process An agent chooses an action a from the finite set A after observing a signal s from the finite set S. The agent then observes an outcome $y \in Y$, which is a subset of $\mathbb{R}^m$ for some finite m. The objective data generating process is determined by a full-support probability distribution over signals $\sigma \in \Delta(S)$ and an action and signal contingent probability measure over outcomes $Q^*(\cdot|\cdot) \in \Delta(Y)^{S\times A}$.[6]

The *individual experience* of an agent consists of a (signal, action, outcome) triplet (s, a, y). The agent's realized flow utility depends on their individual experience through the utility function $u : S \times A \times Y \to \mathbb{R}$. We denote the pure strategies of the agent, i.e., the maps from signals to actions, by $\Pi = A^S$. For notational simplicity we do not allow individual agents to randomize.[7] The objective expected utility of strategy π is $U^*(\pi) = \sum_{s\in S} \sigma(s) \int_Y u(s, \pi(s), y) dQ^*(y|s, \pi(s))$, which we assume is well defined and finite for each $\pi \in \Pi$.

Subjective models The agent uses parametric models to describe the environment. Formally, there is a compact and convex subset $\mathcal{H}$ of a Euclidean space $\mathbb{R}^k$ whose elements θ are associated with a family of probability measures $Q_\theta(\cdot|s, a)$, one for each signal-action pair (s, a).[8] The agent's initial uncertainty about the value of the parameter is described by a *belief* $\mu \in \Delta(\mathcal{H})$, the agent's *subjective model* is the subset of parameters $\operatorname{supp} \mu = \Theta \subseteq \mathcal{H}$ the agent considers possible.

[6]For every subset X of a Euclidean space, we let $\mathcal{B}(X)$ denote its relative Borel sigma-algebra, and $\Delta(X)$ denote the set of Borel probability distributions on X endowed with the Levy-Prokhorov metric.

[7]This is without loss of generality since we will allow different agents with the same belief to play different actions as long as all of those actions maximize their subjective payoff, so that each individual randomization can be replicated at the aggregate level.

[8]Compactness guarantees that for every observed distribution of actions and outcomes there is at least one best explanation. Convexity only plays a role in our analysis of local mutations. See Diaconis and Freedman (1986) for reasons to restrict to a finite-dimensional parameter space.

Preferences and best replies The agent's utility function and beliefs determine their subjective expected utility as a function of their strategy:

$$U_\mu(\pi) = \int_\Theta \sum_{s\in S} \sigma(s) \int_Y u(s, \pi(s), y) dQ_\theta(y|s, \pi(s)) d\mu(\theta).$$

We let $U_\theta = U_{\delta_\theta}$ where δ_θ is the Dirac measure on θ, and assume that $U_\mu(\pi)$ is finite for all (π, μ) pairs. We let $BR(\mu) = \text{argmax}_{\pi\in\Pi} U_\mu(\pi)$ denote the set of pure best replies to μ, and for every $C \subseteq \Delta(\Theta)$, we let $BR(C) = \bigcup_{\mu\in C} BR(\mu)$.

Inference and Kullback-Leibler minimizers Given two distributions over outcomes $Q, Q' \in \Delta(Y)$ we define $H(Q, Q') = -\int_{y\in Y} \log q'(y) dQ(y)$.[9] Note that $-H(Q, Q')$ is the expected log-likelihood of an outcome under subjective distribution Q' when the objective distribution is Q, so Q' with smaller $H(Q, Q')$ better explains distribution Q. This is the force behind Berk (1966)'s result that as sample size grows, beliefs concentrate on the parameters that minimize the Kullback-Leibler divergence from the objective distribution.[10]

The likelihood of an outcome under the objective distribution Q^* depends on both the action and the signal. Given the signals, actions, and outcomes of a continuum population with strategy shares $\psi \in \Delta(\Pi)$, we define the weighted KL divergence

$$H_\psi(Q^*, Q_\theta) = \sum_{s\in S} \sigma(s) \sum_{\pi\in\Pi} \psi(\pi) H\left(Q^*(\cdot|s, \pi(s)), Q_\theta(\cdot|s, \pi(s))\right).$$

We let $\Theta(\psi)$ denote the parameters in Θ that minimize the weighted KL divergence from the observed distribution:

$$\Theta(\psi) = \underset{\theta\in\Theta}{\text{argmin}}\, H_\psi(Q^*, Q_\theta),$$

and call these the *KL minimizers*. Our evolutionary model will assume that the agent's posterior after observing the experience of a population of agents that used strategy distribution ψ is a probability distribution over $\Theta(\psi)$. Proposition 6 shows that this describes the limit as the agent observes a larger and larger number of individual experiences.

Regularity assumptions The agent is *correctly specified* if there is a $\theta^* \in \Theta$ such that $Q^* = Q_{\theta^*}$. We allow this case, but our focus is on the case where the agent is misspecified in

[9] We use the notation $q(y)$ for the probability of outcome y if Y is finite, and for the probability density function of Q wrt to the Lebesgue measure evaluated at y if Y is infinite.

[10] The Kullback-Leibler divergence between Q and Q' is given by $H(Q, Q') - H(Q, Q)$, so any Q' that minimizes $H(Q, Q')$ also minimizes the KL divergence between Q and Q'.

the sense their prior rules out the objective outcome distribution for at least some actions. Let $B_\varepsilon(\theta) = \{\theta' \in \mathcal{H} : ||\theta - \theta'||_2 \leqslant \varepsilon\}$ denote the ε ball around Θ.

Assumption 1. (i) $\theta \mapsto Q_\theta(\cdot|s, a)$ is continuous for all $s \in S$ and $a \in A$.

(ii) Either Y is finite, or for every $\theta \in \mathcal{H}$ and $(s, a) \in S \times A$, $Q^*(\cdot|s, a)$ and $Q_\theta(\cdot|s, a)$ admit probability density functions.

(iii) For every $\varepsilon > 0$, there is an $r \in \mathbb{R}_+$ such that

$$\min_{\theta \in B_\varepsilon(\hat{\theta})} H\left(Q^*(\cdot|s, a), Q_\theta(\cdot|s, a)\right) < r \qquad \forall \hat{\theta} \in \Theta, \forall s \in S, \forall a \in A.$$

Assumption 1(i) guarantees that the set of KL minimizers is non-empty and compact. Without it, the equilibrium notions we define can fail to exist. Assumption 1(ii) requires that either every parameter specifies a discrete outcome distribution for each action, or every parameter specifies a continuous density on outcomes for each action. The assumption is made for simplicity. It allows both the finite outcomes case mostly studied in the literature (see, e.g., Esponda and Pouzo (2016)) and examples with a Gaussian structure. Assumption 1(iii) is a mild boundedness condition that guarantees the upper hemicontinuity of $\Theta(\cdot)$.

2.2 Equilibrium concepts

Here we introduce the static equilibrium concepts that we will relate to the steady states of our evolutionary model. To do so, let $\mathcal{K}$ denote the collection of compact subsets of $\mathcal{H}$.

Definition. A *Berk-Nash equilibrium* is a $(\Theta, \psi) \in \mathcal{K} \times \Delta(\Pi)$ such that for every $\pi \in \operatorname{supp} \psi$ there exists a belief $\mu \in \Delta(\Theta(\psi))$ with $\pi \in BR(\mu)$. A Berk-Nash equilibrium (Θ, ψ) is:

(i) *Pure* if $\psi = \delta_\pi$ for some $\pi \in \Pi$; otherwise it is *mixed*.

(ii) *Unitary* if there exists a belief $\mu \in \Delta(\Theta(\psi))$ with $\psi \in \Delta(BR(\mu))$.

(iii) *Quasi-strict* if $\operatorname{supp} \psi = BR(\Delta(\Theta(\psi)))$.

(iv) *Uniformly strict* if $\psi = \delta_\pi$ and $\{\pi\} = BR(\mu)$ for every $\mu \in \Delta(\Theta(\psi))$.

Berk-Nash equilibrium requires beliefs to be supported on the parameters that best explain the equilibrium data. We do not allow agents to randomize; mixed equilibria here correspond to different agents playing different strategies. Esponda and Pouzo (2016) defines the unitary version of Berk-Nash equilibrium, which requires that all of the equilibrium strategies can be rationalized with the same belief. It shows Berk-Nash equilibria exist, and that if play converges, it converges to a unitary Berk-Nash equilibrium.[11] Section 6 extends this necessary condition to our large population setting, where non-unitary Berk-Nash

[11] As with Nash equilibria in games, pure strategy Berk-Nash equilibria need not exist.

equilibria can also arise in the limit. Note that unitary equilibria need not be pure; it is sufficient that all of the strategies played are best responses to the same belief μ. Non-unitary equilibria only arise if multiple parameters minimize the weighted KL divergence from the equilibrium outcome distribution.[12]

Quasi-strict equilibrium requires that all the strategies that are best replies to some belief over the KL minimizers are played with positive probability; this generalizes the strong equilibrium of Harsanyi (1973) (renamed quasi-strict by Fudenberg and Tirole (1991)) to allow for misspecified beliefs. We will see that the quasi-strictness property helps equilibria resist mutations. The more demanding concept of uniformly strict equilibrium (Fudenberg, Lanzani, and Strack, 2021) requires the equilibrium strategy to be a strict best reply to all of the KL-minimizing parameters. Uniformly strict equilibria are clearly quasi-strict, and because a strict best reply remains so after small changes in beliefs, uniformly strict equilibria resist all local mutations (see Proposition 2). But neither quasi-strict nor uniformly strict equilibria are guaranteed to exist.

A *self-confirming equilibrium* is a Berk-Nash equilibrium (Θ, ψ) such that there is a $\theta \in \Theta$ with $Q_\theta(\cdot|s, \pi(s)) = Q^*(\cdot|s, \pi(s))$ for all $\pi \in \operatorname{supp} \psi$ and $s \in S$. Self-confirming equilibrium requires that the subjective model of the agents contains at least one parameter that induces the same distribution over observables as the equilibrium does. These equilibria always resist mutations, as shown by Corollary 2 below.

3 Illustrative examples

To illustrate our main ideas and motivate our analysis, we present two examples of when mutations do and do not lead to a change in paradigm. The examples use several concepts that are not formally defined until Section 5, but we think that the intuition is still clear. Our first example considers "local" mutations in two versions of the problem of a misspecified seller facing an unknown demand function. The two versions have the same payoff function for the seller and the same objective demand function, but different specifications of the seller's subjective model. The first version shows that even a unique and isolated equilibrium may not resist mutations. In the second version there is a continuum of equilibria, and an equilibrium resists mutations if and only if it doesn't assign too much probability to the suboptimal action. In both versions, playing one action generates evidence that the other

[12]This can be the case due to e.g. symmetry constraints or the use of low-dimensional functional forms; see Fudenberg, Lanzani, and Strack (2021) for examples.

action would be better, so all the equilibria are mixed. The computations backing the claims in this and all subsequent examples are in Appendix A.2.

Example 1. The seller chooses price $a \in \{2, 10\}$ and receives payoff $u(a, y) = ay = a(i^* - \beta^* a + \omega)$, where i^* and β^* are the unknown intercept and slope of the demand function, and ω is a standard normal shock. The objective demand function is given by $(\beta^*, i^*) = (4, 42)$.

a. Suppose that the subjectively possible parameter values are $[3/2, 5/2] \times [28, 32]$, as in the example of Nyarko (1991).[13] The unique Berk-Nash equilibrium assigns probability 1/4 to price 2, sustained by a Dirac belief on $(5/2, 30)$. As shown in Figure 1, the binding constraint of the KL minimization problem is $\beta \leqslant 5/2$. Thus if some mutant agents think a slightly larger set of parameters is possible, their KL minimizer mostly adjusts the slope upwards, which leads the mutants to choose the optimal price of 2.[14] Hence this Berk-Nash equilibrium does not resist the mutation.

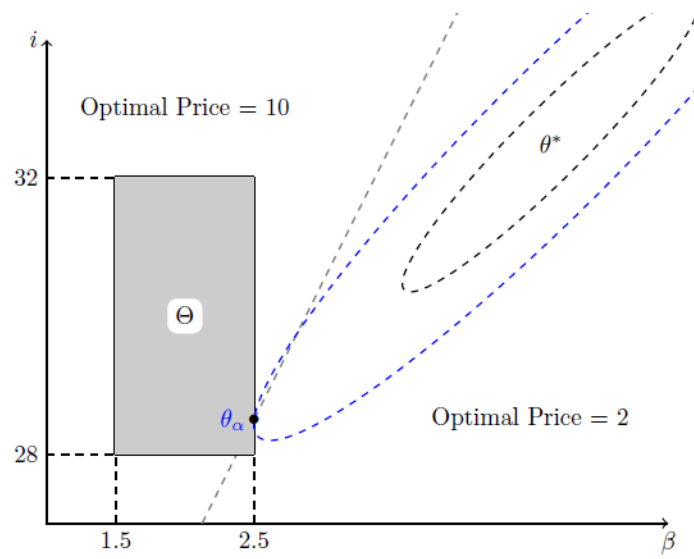

Figure 1: The ellipses are KL-level curves in the unique equilibrium of part a.

b. Here the seller thinks that the possible values of the slopes and intercepts are $[3, 10/3] \times$

[13]In Nyarko's version of the example, both the subjective model and the correct data generating process differ from those in Esponda and Pouzo. To emphasize the role of the subjective model in determining the stability of the equilibrium in an otherwise identical objective environment, we transposed Nyarko's example so both examples have the same data generating process.

[14]The fact that the KL divergence is minimized at the indifference line between actions is a consequence of the fixed point condition that characterizes Berk Nash equilibria: the probabilities of the two actions determine the curvature of the KL divergence, and, analogously to mixed equilibria in games, the Berk-Nash equilibria are the distributions over actions that make the KL minimizers lie on the indifference curve.

[33, 40]. There is a continuum of mixed Berk-Nash equilibria, indexed by the probability of price 2 in [7/8, 35/36], sustained by a Dirac belief on the KL-minimizing parameter (10/3, 40). In all these equilibria, both the slope and intercept constraints on the KL minimization bind, and because the average likelihood depends on the probabilities that each price is charged, so do the KL minimizers for slightly enlarged subjective models. Specifically, when the low price is charged almost all the time, the main unexplained feature in the equilibrium data is high demand, so the KL minimizer for a slightly larger subjective model revises the intercept upward. This is the case for the equilibrium in Esponda and Pouzo (2016) where the low price is charged with probability 35/36, illustrated in Figure 2. Since the new KL minimizer lies above the diagonal indifference curve, it induces 10 as the unique best reply, which yields a lower payoff than the equilibrium action, so this equilibrium resists local mutations. In contrast, when $\psi(2) < 97/100$, the main unexplained feature is high price sensitivity, so the mutants revise their belief about the slope upward. This makes the optimal price 2 subjectively optimal, so this equilibrium does not resist local mutations. ▲

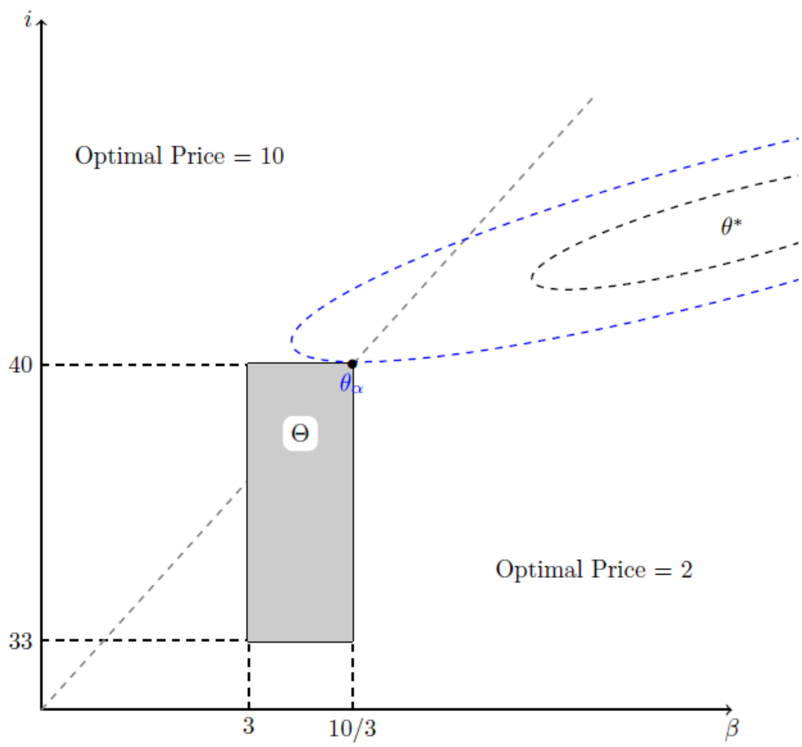

Figure 2: The ellipses are the KL-level curves in the equilibrium of part b where $\psi(2) = 35/36$.

The next example shows that a qualitative relaxation of the subjective model can lead to overadjustment in the direction of the relaxed constraint and a lower payoff than before. We consider an equilibrium in which the agents exert excessive effort because they mistakenly simplify a progressive tax schedule to a linear one, as in an example of Esponda and Pouzo (2016).[15] Mutated agents who realize that the tax schedule might be progressive overestimate its convexity, because they use the data generated by the equilibrium action. This overestimate leads to excessively low effort, which yields less than the equilibrium payoff, so the equilibrium resists the mutation.

Example 2. [Non-linear taxation] An agent chooses effort $a \in A = \{3, 4, 5\}$ at cost $c(3) = 0$, $c(4) = 0.5$, and $c(5) = 1.38$, and obtains income $z = a + \omega$, where $\omega \sim N(0, 1)$. The agent pays taxes $x = \tau^*(z)$, where τ^* has two income brackets, and the higher one is heavily taxed:

$$\tau^*(z) = \begin{cases} z/6, & \text{if } z \leqslant 16/3 \\ \frac{11}{12}z - 4, & \text{if } z > 16/3. \\ \end{cases}$$

The agent's payoff is $u(a, (z, x)) = z - x - c(a)$, so the objectively optimal action is 4.

The agent observes $y = (z, x)$ at the end of each period. The original paradigm is that the tax schedule is linear with random coefficients, as in Sobel (1984), i.e. $\tau_\theta(z) = (\theta + \eta)z + \eta z^2$, $\eta \sim N(0, 1)$, with $\Theta = \mathbb{R}$. Given any action a, the KL-minimizing parameter is given by $\Theta(a) = (\mathbb{E}[\tau^*(a + \omega)/(a + \omega)])$: The agent believes that the expected marginal rate is the actual average rate. Since the actual tax schedule is progressive, the agent exerts too much effort. The unique pure Berk-Nash equilibrium is uniformly strict and has $\pi = 5$, with a Dirac belief on 0.21.

An agent who relaxes linearity by shifting to a quadratic subjective model $\tau_\theta(z) = (\theta_1 + \eta)z + (\theta_2 + \eta)z^2$, $\Theta' = \mathbb{R} \times \mathbb{R}_+$, estimates a very high quadratic term: The equilibrium action makes average income very close to the shift point between the brackets, so the agent observes high progressivity. Their quadratic subjective model extrapolates this progressivity as a global feature of the tax schedule which leads them to choose the minimal action 3. The objectively optimal action 4 is lower than the equilibrium action 5, but the mutated agent overshoots the optimum and ends up using an action that performs even worse than ~~the equilibrium one. For this reason,~~ the equilibrium resists one-hypothesis mutations. ▲

[15] Liebman and Zeckhauser (2004) and Rees-Jones and Taubinsky (2020) provide evidence for this misperception. Our example could also describe agents with other sorts of misperception, for example the agent might be self-employed and untaxed but misperceive the productivity of effort.

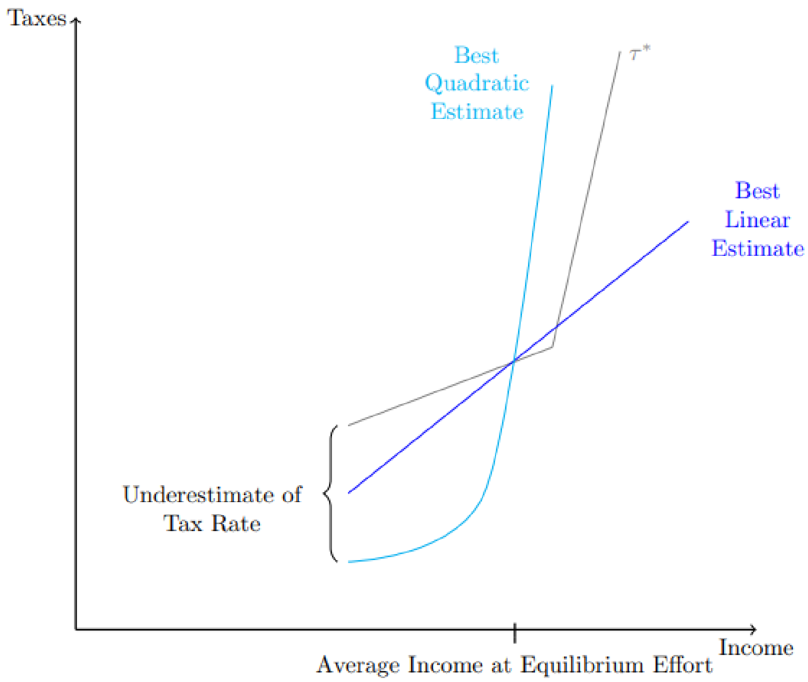

Figure 3: Misspecified Taxation

4 Evolutionary dynamics and steady states

We consider a model that combines individual Bayesian learning with evolutionary competition between the subjective models. There is a continuum of agents, all with the same utility function. The *state* of the system at every period $t \in \mathbb{N}$ is a joint distribution p in the space P of finite-support measures on $\mathcal{K} \times \Pi$ over the subjective models and strategies of the agents. We denote the marginal distributions of p as $p_{\mathcal{K}}$ and p_{Π}.[16]

Inference and actions Let $p^{t+1}(\cdot|\Theta)$ denote the distribution over strategies played at time $t+1$ by the agents with subjective model Θ when the previous state is p^t. We require that this distribution satisfies the following inclusion, which captures the effect of learning and optimization:

$$p^{t+1}(\cdot|\Theta) \in \Delta(BR(\Delta(\Theta(p^t_{\Pi})))). \tag{1}$$

This formula says that each agent plays a best response to a posterior belief that is supported on the KL-minimizing parameters in the agent's model given the data from the

[16]The assumption that at any point in time there is only a finite number of different subjective models in the population is made to guarantee that the process is well-defined.

previous period. The reason that $p^{t+1}(\cdot|\Theta)$ takes values in $\Delta(BR(\Delta(\Theta(p^t_\Pi))))$ as opposed to the smaller set $BR(\Delta(\Theta(p^t_\Pi)))$ is that different agents with the same subjective model may play different best responses: They may have different beliefs over the KL minimizers when $\Theta(p^t_\Pi)$ is not a singleton, and multiple strategies may be best replies to the same beliefs. We provide an explicit learning foundation for this process in Section 6 under the assumption that either there is a unique best reply to the KL minimizers (which covers the case of a uniformly strict Berk-Nash equilibrium) or that Θ is finite.

Evolutionary dynamics We assume that the share of agents with a particular subjective model evolves according to an exogenously fixed $T : P \to \Delta(\mathcal{K})$, so that

$$p^{t+1}_{\mathcal{K}} = T(p^t). \tag{2}$$

We say that T is an *evolutionary map* if it is continuous, with $\operatorname{supp} p_{\mathcal{K}} = \operatorname{supp} T(p)$ and *payoff monotone* (Samuelson and Zhang, 1992), meaning that

$$U^*(p(\cdot|\Theta)) > (=) U^*(p(\cdot|\Theta')) \implies \frac{T(p)(\Theta)}{T(p)(\Theta')} > (=) \frac{p_{\mathcal{K}}(\Theta)}{p_{\mathcal{K}}(\Theta')} \qquad \forall p \in P,^{17}$$

where $U^*(p(\cdot|\Theta))$ is the average payoff of $\sum_{\pi \in \Pi} p(\pi|\Theta) U^*(\pi)$ of agents with model Θ.

This simple model of paradigm change can be interpreted as the result of biological reproduction or as the result of social learning and imitation.[18] Under the biological perspective, payoffs correspond to fitness, and agents whose subjective model induces fitter actions have more offspring. Parents transmit their subjective model— i.e., the support of their prior— but not their beliefs, strategy, or data, and the offspring then perform Bayesian updating based on the actions and outcomes in the previous period. The biological interpretation of payoff monotonicity is better suited to misspecifications due to behavioral biases such as overconfidence or correlation neglect, and can help to explain why evolutionary forces may or may not be able to eradicate those biases. For example, an overconfident agent may also be overconfident about the skill of their offspring, and in turn this may induce the offspring to be more confident about themselves. Other economic examples, such as the misspecified beliefs of a seller about a demand function, are better interpreted as arising from imitation. Under this interpretation, agents in the new generation receive noisy signals about the per-

[17] Recall that we restrict attention to states with finitely many models and strategies.

[18] One example of a payoff monotone dynamic is the discrete-time replicator dynamics (Hines (1980), Hofbauer and Sigmund (1988), Dekel and Scotchmer (1992), and Cabrales and Sobel (1992)).

formance of the different subjective models as in Björnerstedt and Weibull (1995), Schlag (1998), and Binmore and Samuelson (1997), and use these signals to decide whether to stick with their parent's worldview or adopt a different one.

Payoff monotonicity requires that the average payoff of each subjective model is enough to rank them in terms of relative growth. This rules out dynamics that are based on non-linear functions of each agent's payoffs. However, whenever the ordinal ranking can be summarized in a single statistic (e.g., the median as in Ellison and Fudenberg (1993)) Propositions 3 and 4 continue to hold by replacing the average payoff in the statement with, say, the median.[19]

The combination of Bayesian inference within a model and payoff monotone evolution of the subjective model shares can be seen as a generalized cross-validation procedure. Under cross validation, a statistician has to decide which statistical model Θ to use to perform inference (typically in the form of a subset of parameters they try to estimate). To do so they rely on a past sample of observations, which they divide in two parts, with the first (the training sample) used to estimate the model parameters and the second (the validation sample) to see how well the estimated model performs. For an agent in generation $t+1$, generation $t-1$'s outcomes act as the training dataset, and the induced performance of the estimated parameter in period t acts as the validation sample. Indeed, generation $t-1$'s outcomes are used by generation t agents to make decision, and generation $t+1$ agents select the models depending on their performance at time t. As under cross validation, models that are more successful at the validation stage are more likely to be adopted/maintained. In the procedure, the utility function plays the role of a loss function: a model is viewed more favorably when its average utility on the validation sample is higher.

Notice that the offspring "inherit" their subjective model from their parents, but do not inherit their parent's parameter estimates. That is, the offspring inherit ways of interpreting data, but are left to make their own inferences based on the data that they observe. The offspring also do not inherit how to handle cases where there are multiple KL-minimizing beliefs or multiple best responses to the same belief. If these features were inherited, the only change would be that mixed Berk-Nash equilibria in which actions with different payoffs are played by a positive fraction of agents would not be steady states.

In models where evolutionary pressure acts directly on strategies (rather than subjective models), payoff monotonicity implies that for every solution the average payoff in the population increases over time. This is not the case in our model, as the same subjective model may induce different inferences on different data. As a consequence, the dynamics can cycle

[19]Proposition 7 holds as well provided that the statistic is Gateaux differentiable in the payoff distribution.

between states with different payoffs, as in Example 5 in the Appendix.

A sequence $(p^t)_{t \in \mathbb{N}_0} \in P^{\mathbb{N}_0}$ is a *solution* if there is an evolutionary map such that for all $t \in \mathbb{N}_0$ equations (1) and (2) hold. A *steady state* is a $\hat{p} \in P$ such that the sequence constant at $\hat{p}$ is a solution, and $\hat{p}_{\mathcal{K}} = \delta_\Theta$ for some $\Theta \in \mathcal{K}$, so the solution is a constant point mass on a single model.[20] In a steady state all agents have the same subjective model, but they can have different beliefs over the KL minimizers unless the minimizer is unique. A steady state is *unitary* if all the strategies are best replies to the same belief. This does not require that the distribution over strategies is a point mass, because different agents can break ties between best replies in different ways.

Lemma 1. *For all $\Theta \in \mathcal{K}$ and $\psi \in \Delta(\Pi)$, $\delta_\Theta \times \psi$ is a steady state if and only if (Θ, ψ) is a Berk-Nash equilibrium. Moreover, $\delta_\Theta \times \psi$ is unitary if and only if (Θ, ψ) is unitary.*

The proofs of this and all subsequent results are in the Appendix. Lemma 1 combined with Theorem 1 of Esponda and Pouzo (2016) guarantees the existence of a steady state.[21]

Corollary 1. *For every objective environment (S, A, Y, u, Q^*), every evolutionary map T, and every subset of subjective models $C \in \mathcal{K}$ there exists a steady state p with $p_{\mathcal{K}}(C) = 1$.*

5 Mutations

5.1 Explanation-Improving Mutations

We now consider mutations that lead agents to expand the subjective model they inherited from their parents. We suppose that mutant agents consider a larger set of possible parameter values, and use their data from the previous generation to estimate which parameters fit best.

Our first step is to define what we mean by an ε mutation.

Definition. p is an *ε mutation of a steady state $\delta_\Theta \times \psi$ to $\Theta' \supseteq \Theta$* if

(i) $p_{\mathcal{K}} = (1 - \varepsilon)\delta_\Theta + \varepsilon \delta_{\Theta'}$ and

(ii) $p(\cdot|\tilde{\Theta}) \in \Delta(BR(\Delta(\tilde{\Theta}(\psi)))) \qquad \forall \tilde{\Theta} \in \{\Theta, \Theta'\}$.

Note that both the mutated and unmutated agents choose their actions based on the same data, namely the distribution of play that prevailed before the mutation occurred. Note also that ε mutations must enlarge the set of subjective models. In Section 5.5 we explain that

[20] The misspecified learning literature has focused on this case; Bohren and Hauser (2021) is an exception.

[21] Esponda and Pouzo (2016) assumes Y is finite, but this is not needed for the proof of their Theorem 1.

as long as the induced mutations are minimally responsive to the evidence generated by the steady state this is without loss of generality.

Definition. A Berk-Nash equilibrium (Θ, ψ) *resists mutation to* Θ' if there is a sequence of solutions $(p^t_{\varepsilon_n})_{t,n\in\mathbb{N}_0}$ where $p^0_{\varepsilon_n}$ is an $\varepsilon_n \in (0,1)$ mutation of $\delta_\Theta \times \psi$ to Θ', $\lim_{n\to\infty} \varepsilon_n = 0$, and $\lim_{n\to\infty} \lim_{t\to\infty} (p^t_{\varepsilon_n})_\Pi = \psi$.

For every ε_n, the inner limit gives the long-run strategy distribution following an ε_n mutation to Θ'; the outer limit sends the fraction of mutated agents to 0. The equilibrium (Θ, ψ) resists this mutation if this iterated limit converges back to ψ for some solution that starts from an ε_n mutation to Θ'.

This definition is purposefully weaker than the requirement that every possible solution converges back to ψ. Using this definition lets us identify the unambiguous departures from what is predicted by the purely Bayesian benchmark; the alternative stronger definition would leave open whether an equilibrium that "failed to resist" could nevertheless persist.

Our definition of resisting mutations is also weaker than alternatives that require the equilibrium to persist even when agents are forced to experiment. Forced experimentation would force even correctly specified agents to move away from self-confirming equilibria that are not Nash equilibria as they learn the consequences of non-equilibrium play. When such experimentation-inducing mutations are present, we expect them to be rare and thus to operate on a much slower time scale than paradigm changes that are driven by inconsistencies with the equilibrium data.

A first reason why an equilibrium can resist a mutation is that the mutation may not induce a different best response. In particular this happens when the mutation does not lead to a better explanation of the equilibrium data.

Definition. The mutation of a steady state $\delta_\Theta \times \psi$ to Θ' is *explanation improving* if $\min_{\theta\in\Theta'} H_\psi(Q^*, Q_\theta) < \min_{\theta\in\Theta} H_\psi(Q^*, Q_\theta)$.

In an explanation-improving mutation, a fraction of agents realizes that they could better explain their data within a more permissive paradigm Θ'.

Proposition 1. *A Berk-Nash equilibrium resists every mutation that is not explanation improving.*

The proof of this is simple: Because the subjective model Θ' of the mutated agents contains Θ and Θ' is not explanation improving, the best explanations in Θ are also best explanations in Θ', and one possible continuation path is for the mutants and conformists to both play the same ψ as before the mutation.

Corollary 2. *A self-confirming equilibrium resists every mutation.*

This follows immediately from the definitions, as in a self-confirming equilibrium the subjective model perfectly matches the observed distribution, so the KL divergence between the agent's beliefs and observations is 0. Conversely, only equilibria where the strategy is objectively optimal or that are self-confirming resist a mutation that adds the correct data generating process to Θ.

Which equilibria resist mutations depends on the types of mutations that can occur. We consider two different sorts of enlargements that do not necessarily include the correct model, "local mutations" that make small enlargements of the current parameter space, and "one-hypothesis mutations" where the mutated agents drop one of the restrictions of their subjective models. Local mutations relax the quantitative specification of the model with the idea that a more robust approach may be beneficial. One-hypothesis mutations instead relax a qualitative restriction on the data generating process. For example, agents might restrict the set of possible values for one dimension of the parameter, as in the case of an overconfident agent who is sure that their skill is higher than some threshold. The hypotheses can also take the form of joint restrictions on the parameters, as with an agent who believes that two variables are independent.

5.2 Local mutations

Definition. Subjective model Θ_ε is the ε *expansion of* Θ if $\Theta_\varepsilon = \bigcup_{\theta \in \Theta} B_\varepsilon(\theta)$.

In an ε local mutation, a fraction ε of the agents in the new generation reacts to unexplained evidence by considering a moderately more permissive paradigm.[22]

Definition. p is an ε *local mutation* of a steady state $\delta_\Theta \times \psi$ if it is an ε mutation of $\delta_\Theta \times \psi$ to the ε expansion of Θ.

A Berk-Nash equilibrium (Θ, ψ) *resists local mutations* if it resists mutation to every sufficiently small ε local mutation. That is, an equilibrium resists local mutations if after small mutations aggregate behavior converges back to the equilibrium.

Proposition 2. *(i) Every uniformly strict Berk-Nash equilibrium resists local mutations.*
(ii) Every Berk-Nash equilibrium with $\Theta(\psi)$ in the interior of Θ resists local mutations.

[22]To lighten notation we use ε in two roles here, but nothing would change if we instead had share ε' of agents adopt an ε'' expansion.

Proposition 2(i) reinforces the finding of Fudenberg, Lanzani, and Strack (2021) that uniformly strict Berk-Nash equilibria have strong local stability properties. Although that result was obtained in a different single-agent setting and with a different proof technique, both conclusions follow from the fact that in a uniformly strict equilibrium, the equilibrium strategy π is a strict best response to all beliefs that assign a sufficiently high probability to the parameters that minimize the weighted KL divergence given that π is played. This implies that after small mutations π is still the unique best reply to the agents' beliefs. Part (ii) follows from the fact that when the KL minimizers are in the interior of Θ, they have a strictly lower divergence than any parameter on or near the boundary of Θ.

In other Berk-Nash equilibria there may be a parameter $\theta' \in \Theta(\psi)$ that does not induce a unique best reply. If the mutation leads to a KL minimizer that is near θ', the mutated agents may start to play a different strategy, inducing a departure from equilibrium play.

To evaluate the stability of Berk-Nash equilibria that are not uniformly strict, we will use a measure of how much enlarging the parameter space in a particular direction improves the explanation of the equilibrium outcome. Given a steady state $\delta_\Theta \times \psi$ and an $\varepsilon \in \mathbb{R}_{++}$ we define

$$\mathcal{M}_{\Theta,\psi}(\varepsilon) = \operatorname*{argmin}_{\theta \in \Theta_\varepsilon} H_\psi\left(Q^*, Q_\theta\right). \tag{3}$$

These parameters generate the largest decrease in H.[23]

By an argument paralleling that of Berk, small mutations induce beliefs that are concentrated on $\mathcal{M}_{\Theta,\psi}(\varepsilon)$. We will show that if the strategies induced by these beliefs perform better than the equilibrium strategy distribution, the mutation will not die out, permanently destabilizing the equilibrium. Conversely, if the strategies of the mutated agents lead to lower payoffs and the equilibrium is quasi-strict, the mutated agents will eventually disappear, and play converges back to the original equilibrium. To formalize this, let $\Pi_{\mathcal{M}_{\Theta,\psi}} = \limsup_{\varepsilon \to 0} BR(\Delta(\mathcal{M}_{\Theta,\psi}(\varepsilon)))$ denote the limits of the strategies that are optimal against distributions over $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ as ε goes to 0. We call these the *local responses* at (Θ, ψ).

Proposition 3. *Let (Θ, ψ) be a Berk-Nash equilibrium.*

(i) If for every local response π at (Θ, ψ), $U^*(\pi) > U^*(\psi)$, then (Θ, ψ) does not resist local mutations.

(ii) If for some local response π' at (Θ, ψ), $U^*(\pi') \leqslant U^*(\psi)$, and (Θ, ψ) is quasi-strict, it resists local mutations.

Part (ii) illustrates the similarity and differences between the stability notions when sub-

[23] $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ need not be a singleton, but it is a singleton for small ε in the examples we analyze. Moreover, as shown in Lemma 5, it is a singleton for sufficiently small ε if $\Theta(\psi)$ is unique and Q_θ is linear in θ.

jective models rather than actions are inherited. The sufficient condition for resistance has two requirements, mirroring the conditions defining an evolutionarily stable strategy. In both cases, the first requirement is that the mutants do not perform better *given the equilibrium strategy.* For evolutionary stability "given the equilibrium strategy" means *against the equilibrium strategy.* Here "given the equilibrium strategy" means *given the beliefs generated by the equilibrium.* The second requirement is that the equilibrium strategy does not perform poorly given the mutants strategy. For evolutionary stability, "given the mutants' strategy" means *against the mutants strategy.* Here "given the mutants' strategy" means *for some best reply to the evidence generated by the mutants' strategy.* Requiring that the equilibrium is quasi-strict guarantees this second requirement is satisfied: In a quasi-strict equilibrium all best replies are played with positive probability, and so by upper hemicontinuity of $\Theta(\cdot)$ and $BR(\cdot)$ all local responses are already played with positive probability, so no completely new evidence is generated after mutation.

To illustrate the role of payoff comparisons in Proposition 3, we revisit Example 1. In the equilibrium of Example 1a, only the slope constraint is binding. An upward revision of the slope makes the low price the unique optimal choice, and since the low price performs better than mixing, by Proposition 3(i) this Berk-Nash equilibrium does not resist local mutations. In Example 1b, the constraints on the intercept and the slope are both binding. Here a Berk-Nash equilibrium resists local mutations if and only if the low price is played by a large fraction of agents. When almost all the agents choose a low price, mutants revise the intercept upward, which induces the high price. This action performs worse than the equilibrium, so by Proposition 3(ii) the equilibrium resists local mutations. When both actions are played sufficiently often, the agents' revisions induce the objectively optimal low price, and by Proposition 3(i) the equilibrium does not resist local mutations.

Section 5.4 shows why the quasi-strictness assumption in part (ii) of the proposition is needed. Without it, the feedback gathered from mutated agents playing a strategy that is not used in equilibrium may change the behavior of the old population, even if they were performing better than the mutants. However, because $\mathcal{M}_{\Theta,\psi}(\cdot)$ and $BR(\cdot)$ have closed graphs, local mutations never introduce a completely novel strategy in a quasi-strict equilibrium.

When Θ is finite and H is continuously differentiable in θ, there is a convenient way to check Proposition 3's conditions on the local responses. Let $\mathcal{S}$ denote the sphere of radius 1 in $\mathbb{R}^k$ with respect to the $||\cdot||_2$ norm. Given a strategy distribution ψ and $v \in \mathcal{S}$, let

$$D_\psi(\theta, v) = \liminf_{h\to 0} \sum_{s\in S} \sigma(s) \sum_{\pi\in\Pi} \psi(\pi) \left(\int_Y \log \frac{q_{\theta+hv}(y|s,\pi(s))}{q_\theta(y|s,\pi(s))} dQ^*(y|s,\pi(s)) \right) /h$$

be the ψ-weighted directional derivative of $-H$ in direction v at θ. As we show in Lemma 6 in the Appendix, it is enough to compute the direction v in which $D_\psi(\theta, v)$ is maximal, and look at the best replies to the parameters along that direction.

5.3 One-hypothesis mutations

One-hypothesis mutations capture the idea that mutations that only change one dimension of the model are much more likely than adjustments that involve multiple aspects of the model at once. As an illustration, recall that the refined version of the Ptolemaic cosmological model as formalized by Tycho Brahe has two central tenets: that the Sun revolves around the Earth, and that stars and planets revolve around the Earth-Sun pair (see Kuhn, 1957). The fit of the Ptolemaic model is imperfect but very good, and relaxing either assumption to include the correct model separately does not improve performance.[24] This may help explain why the Ptolemaic system persisted for over a thousand years.

Formally, there is a finite collection of continuous functions $\mathcal{F} = \{f_l\}_{l=1}^m$, where each $f_l : \mathcal{H} \to \mathbb{R}$, such that $\Theta = \{\theta \in \mathcal{H} : f_l(\theta) \geqslant 0, \forall l \in \{1, ..., m\}\} := \Theta_{\mathcal{F}}$.

Definition. Subjective model Θ^l is a *one-hypothesis relaxation of* $\Theta_{\mathcal{F}}$ in hypothesis $l \in \{1, ..., m\}$ if $\Theta^l = \Theta_{\mathcal{F}\setminus\{f_l\}}$.

Observe that several collections of hypotheses can describe the same Θ, and they can have different sets of one-hypothesis relaxations.[25] This is natural, as the hypotheses are part of the agents' model of the world. These hypotheses describe the parts of the subjective model of an agent that can be separately relaxed by a mutation. As an illustration, consider Example 2. We model the constraint $\theta_2 = 0$ as a pair of inequality constraints, $\theta_2 \leqslant 0$ and $\theta_2 \geqslant 0$. Relaxing the first constraint and allowing negative values has no effect; relaxing the second constraint has the same effect as dropping both of them.[26]

Definition. $\bar{p}$ is a *one-hypothesis* ε *mutation* of a steady state $\delta_\Theta \times \psi$ if it is an ε mutation to some one-hypothesis relaxation of Θ.

[24] The unsuccessful "Andalusian revolt" in Islamic astronomy relaxed the latter assumption without abandoning geocentrism, see Sabra (1984), while the Copernican model (before the additions by Kepler) only relaxed the former assumption.

[25] As a trivial special case, if $\mathcal{F}$ contains two copies of each of the f's it includes, then one-hypothesis mutations have no effect on Θ.

[26] Formally, to make the example a special case of our model, we need $\mathcal{H} = [-K_1, K_1] \times [-K_2, K_2]$ for some large K_1, K_2 and $\mathcal{F} = \{-\theta_2, \theta_2\}$.

We say that a Berk-Nash equilibrium $(\Theta_{\mathcal{F}}, \psi)$ *resists one-hypothesis mutations* if it resists every one-hypothesis ε mutation for sufficiently small ε.[27] Given a steady state $p = \delta_\Theta \times \psi$ the *l-agnostic KL minimizers* are $\mathcal{P}_l(p) := \operatorname{argmin}_{\theta \in \Theta^l} H_\psi(Q^*, Q_\theta)$, and $\Pi_{p,l} = BR(\Delta(\mathcal{P}_l(p)))$ denotes the set of best replies when hypothesis l is dropped.

Proposition 4. *Let $(\Theta_{\mathcal{F}}, \psi)$ be a Berk-Nash equilibrium.*

(i) If for some $l \in \{1, ..., k\}$, $U^*(\pi) > U^*(\psi)$ for every $\pi \in \Pi_{p,l}$, then (Θ, ψ) does not resist one-hypothesis mutations.

(ii) If for every $l \in \{1, ..., k\}$ and $\pi' \in \Pi_{p,l}$, $U^*(\pi') \leqslant U^*(\psi)$, and either $\Pi_{p,l} \subseteq \operatorname{supp} \psi$ or ψ is a uniformly strict Berk-Nash equilibrium, then (Θ, ψ) resists one-hypothesis mutations.

The intuition for part (i) is that if the strategy distribution converges back to the equilibrium, then eventually the mutated agents will use strategies that are a best reply to the l-agnostic KL minimizers, while conformists would perform strictly worse. The intuition for part (ii) is that under the conditions considered, the feedback received from the play of the mutants would not move the conformist play too far from the equilibrium behavior. Since the equilibrium payoff is larger than the ones induced by the best-reply to the l-agnostic KL minimizers, the mutated agents will eventually die out.

In Example 1b, a one-hypothesis mutation that allows for larger slopes leads agents to play the objectively optimal price 2. Thus by Proposition 4(i) these equilibria do not resist one-hypothesis mutations.

The next example considers a buyer who has correlation neglect: they do not understand that the price charged by a seller is positively correlated with the value of the good. In the first version of the example, the difference between the buyer's and seller's values for the good is constant. Here the equilibrium resists one-hypothesis mutations, because the buyer never bids a high price, and so even after mutation does not learn that higher bids attract higher value sellers. This may help explain the apparent pervasiveness of correlation neglect. However, if there is a stochastic shock to the buyer's valuation, they make a wider range of bids, which allows one-hypothesis mutations to lead them to find better actions. This suggests that correlation neglect should be less frequent in settings where taste heterogeneity leads the agents to use a range of actions instead of always using the same one.

Example 3. [Additive Lemons and Cursed Equilibrium]

a. **Persistent Correlation Neglect** The agent, a buyer whose value for an object is $v = \omega + 3.1$, faces a seller who owns the object and values it at ω. They play a double auction

[27] Note that here, unlike with local mutations, ε plays a single role.

with price at the buyer's bid, so the seller sets their ask x equal to their value, and a sale occurs if buyer's bid a is at least x. The value ω is 3 with probability 1/3, 2 with probability 1/2, and 1 with probability 1/6. The value is observed only if a transaction occurs, so the outcome is the pair $y = (\tilde{\omega}, x) \in (\Omega \cup \{\#\}) \times X$, where $\tilde{\omega} = \omega$ if $a \geqslant x$, and $\tilde{\omega} = \#$ otherwise.

Here a parameter θ consists of probability distribution on seller ask prices (p_1, p_2, p_3) and a family of conditional probabilities $(F(1|1), F(2|1), F(1|2), F(2|2), F(1|3), F(2|3))$, where $F(i|j)$ is the probability that the value is less than or equal to i given that the seller asked price j. So $\mathcal{H}$ is the subset of $\mathbb{R}^9$ such that $\sum_{i=1}^{3} p_i = 1$, and $0 \leqslant F(1|i) \leqslant F(2|i) \leqslant 1$ for $i \in \{1, 2, 3\}$. The objective price distribution is $(1/6, 1/2, 1/3)$, with conditional probabilities $(1, 1, 0, 1, 0, 0)$, so the objectively optimal strategy is to bid 3.

Suppose that as in Esponda (2008) the agent believes that seller ask price and value are independent. Because the value is only observed when a transaction occurs, the buyer doesn't realize that a higher bid would increase average quality conditional on the seller accepting the offer, and as we show in Appendix A.2.3, $a = 2$ is a uniformly strict Berk-Nash equilibrium. The KL-minimizing parameter is an independent joint probability distribution that is correct about the distribution of seller asks and with value distribution $(1/4, 3/4, 0)$.

Because the buyer never offers 3, a one-hypothesis mutation can be explanation improving only if it better fits the conditional value distributions for asks 1 and 2. Such mutations do not induce a different strategy, so the mutated agents do not obtain a higher payoff, and by Propositions 1 and 4 this equilibrium resists one-hypothesis mutations.

b. **A non-resistant uniformly strict Berk Nash-equilibrium** Now suppose that the buyer's value is $v = \omega + 3.1 + s$, where s is either -1 or 1 with probability 1/2 each, independent of ω. The objectively optimal strategy is to bid 3 after both signals, but $\pi(-1) = 2$, $\pi(1) = 3$ is a Berk-Nash equilibrium. The KL-minimizing parameter is an independent joint probability distribution that is correct about the distribution of seller bids. However, because the values are only observed when the transaction is realized, and the buyer doesn't realize that a higher bid would increase average quality conditional on the sellers accepting the offer, the corresponding distribution over values $(1/5, 3/5, 1/5)$ is too pessimistic, leading to the (objectively suboptimal) bid of price 2 after signal $s = -1$.

This equilibrium is uniformly strict, so by Proposition 2 resists local mutations. However, the one-hypothesis relaxation that allows for the possibility that a high value is more likely to be observed after the seller has asked for a high price leads to the subjective model:

$$\Theta' = \left\{ \theta \in \mathbb{R}^9_+ : \begin{array}{c} p_1 + p_2 + p_3 = 1, \\ F(1|1) = F(1|2) = F(1|3), \\ F(2|1) = F(2|2) \geqslant F(2|3), \end{array} \right\}$$

which generates a posterior concentrated on $\hat{\theta} = ((1/6, 1/2, 1/3), (1/5, 1, 1/5, 1, 1/5, 1/5))$. Since $BR(\hat{\theta}) = \{3\}$, by Proposition 4 the equilibrium does not resist one-hypothesis mutations. ▲

The difference between the cases is that payoff shocks lead agents to use more actions, which makes it easier to spot errors in the subjective model and find better strategies. More generally, a Berk-Nash equilibrium cannot resist mutation to a correctly specified model if the payoff shocks lead the agent to assign positive probability to every action.

5.4 Misspecification driven innovation

Here we sharpen our previous sufficient conditions for an equilibrium not to persist. Those conditions considered a direct channel between paradigm change and destabilization, where mutants obtain a higher payoff. Now we focus on an indirect channel: The new strategies the mutation induces may have lower payoff, but provide information that lets the agents with the old subjective model realize that their previous play was suboptimal. This possibility is captured by the following definition.

Definition. A Berk-Nash equilibrium (Θ, ψ) is *innovation vulnerable* if there exist $\pi_I, \pi_U \in \Pi \backslash \operatorname{supp} \psi$ and $\bar{\varepsilon} > 0$ such that $\{\pi_U\} = \operatorname{argmax}_{\pi \in \Pi} U_\mu(\pi)$ for all $\mu \in \Delta(\Theta(\psi'))$ with $\psi' \in B_{\bar{\varepsilon}}(\psi)$, $\psi'(\pi_I) > 0$. When (Θ, ψ) is innovation vulnerable, we say that $\Theta' \supset \Theta$ is *innovation inducing* for (Θ, ψ) if $BR(\Delta(\Theta'(\psi))) = \{\pi_I\}$.

In words, an equilibrium is innovation vulnerable if there is an unplayed best response π_U to some belief over equilibrium KL minimizers, and an innovative strategy π_I that provides evidence in favor of π_U even if π_I has a low payoff. A trivial case of an equilibrium that is not vulnerable to innovation is when Θ is a singleton. More generally, quasi-strict (and henceforth uniformly strict) Berk-Nash equilibria are not innovation vulnerable, since they do not have any unplayed best responses to the belief over equilibrium KL minimizers. By restricting the second part of the statement to quasi-strict and uniformly strict Berk-Nash equilibria respectively, Propositions 3 and 4 rule out this indirect channel. Other sorts of equilibria can be innovation vulnerable, and innovation vulnerable equilibria do not persist.

Proposition 5. *An innovation vulnerable equilibrium does not resist a mutation to an innovation-inducing model.*

The intuition is that if the solution of the dynamic process returns to the old paradigm, the data provided by the mutated agents breaks ties among old paradigm's best-fitting models in a way that favors a non-equilibrium best reply.

The role of innovation vulnerability can be vividly illustrated with the case of thalidomide. In the original subjective model doctors believed that no treatment is viable for the nausea and "morning sickness" experienced in some pregnancies, so these symptoms were not treated. It was also known that blocking the growth of blood vessels slows down myeloma, but there was uncertainty about which substances do so without severe side effects. Finally, thalidomide had been used (and found effective) in treating insomnia.

The mutation came in the form of understanding the similarity of the histamine levels seen in patients with morning sickness and patients with insomnia. This evidence-driven shift led to the use of thalidomide as a cure for morning sickness, but that had a very low payoff: While effective against morning sickness, thalidomide has a dramatic effect on the fetus, which led to the "thalidomide tragedy." However, the data observed in the tragedy, i.e., the inhibition of the growth of blood vessels without side effects beyond those for the fetus, led to the very successful use of thalidomide as a treatment for myeloma (Franks, Macpherson, and Figg (2004)). Notice that adopting thalidomide for myeloma did not require a shift of paradigm, since it is consistent also with the original subjective model that believes no treatment is possible for morning sickness. Example 6 in the Appendix provides a fully detailed example of an innovation vulnerable equilibrium.

5.5 Mutations to smaller subjective models

Invasion by a non-improving mutation A mutation that is not explanation improving can only invade when agents discard a parameter that provides the best fit. Consider for example a completely mixed equilibrium ψ of Example 1b, and let Θ' be the single point $(3, 33)$, so that mutated agents drop the best-fitting parameter $(10/3, 40)$. This mutation is not explanation improving, but the equilibrium does not resist it: the mutated agents start to play 2 at every period, and either the conformists eventually switch to 2, or they eventually disappear. In either case, play does not converge back to ψ.

However, in the more plausible case where the smaller subjective model retains the best-fitting parameters, the smaller model cannot invade.

Definition. p is an *evidence-based simplification* of a steady state $\delta_\Theta \times \psi$ to Θ' if

(i) $p_{\mathcal{K}} = (1-\varepsilon)\delta_\Theta + \varepsilon\delta_{\Theta'}$,

(ii) $p(\cdot|\tilde{\Theta}) \in \Delta(BR(\Delta(\tilde{\Theta}(\psi)))) \qquad \forall\tilde{\Theta} \in \{\Theta, \Theta'\}$,

(iii) $\Theta' \subseteq \Theta$ and $\text{argmin}_{\theta\in\Theta'} H_\psi(Q^*, Q_\theta) = \text{argmin}_{\theta\in\Theta} H_\psi(Q^*, Q_\theta)$.

A mutation is an evidence-based simplification if the mutated agents reduce the size of their subjective model by only eliminating parameters that provide worse fits to the equilibrium data.

Since evidence-based simplifications do not drop any KL minimizer, the equilibrium strategy remains a best reply to some belief over the KL minimizing parameters. Thus every Berk-Nash equilibrium (Θ, ψ) resists every evidence-based simplification. This conclusion follows from our assumption that mutations that do not discard the best fitting parameters, as opposed to being "blind." Blindness may be natural under a purely biological interpretation, but in our case of the evolution of subjective models it seems less plausible that new agents reject the most successful explanations of the previous generation. The new agents may well drop some aspects of the model they inherit or imitate, but we see little reason for the dropped aspects to be those that actually work in explaining the observables. More generally, although mutations that are not evidence-based simplifications do not seem completely implausible, we expect them to be much less common and so to operate on a much slower time scale.

6 Large finite data sets

Our evolutionary model is deterministic because agents observe an infinite number of individual experiences. Here we show that this process emerges as the limit of observing large finite data sets.[28]

Suppose that all agents with subjective model Θ have the same prior μ_Θ.[29] Each agent $i \in [0,1]$ observes $n \in \mathbb{N}$ individual experiences drawn from a population playing $p_\Pi \in \Delta(\Pi)$, independently across agents, computes posterior belief μ_n^i using Bayes rule, and chooses a best reply according to a measurable best response function $R : \Delta(\Theta) \to \Pi$. So aggregate play is $\psi_n(\Theta, p_\Pi)(\pi) = \int_0^1 1_{R(\mu_n^i)=\pi} di$.

[28]This is only one way to provide a foundation for our model; we provide it to show the plausibility of the dynamics we study. We conjecture that the same steady states would be asymptotic limits if a single agent acted each period, as in He (2022) or Bohren and Hauser (2021).

[29]Section A.1 of the Appendix shows that the limit belief is independent of the prior.

Definition. We say that p_Π distinguishes parameters and strategies if:

(i) For all $\theta, \theta' \in \Theta(p_\Pi)$, there is $s \in S$ such that there is positive probability under $\sum_{\pi \in \Pi} p_\Pi(\pi) Q^*(\cdot|s, \pi(s))$ that $\sum_{\pi \in \Pi} p_\Pi(\pi) Q_\theta(\cdot|s, \pi(s)) \neq \sum_{\pi \in \Pi} p_\Pi(\pi) Q_{\theta'}(\cdot|s, \pi(s))$.

(ii) For all $\pi, \pi' \in \Pi, \pi \neq \pi'$, there is $\theta \in \Theta(p_\Pi)$ such that $U_\theta(\pi) \neq U_\theta(\pi')$.

In words, p_Π distinguishes parameters and strategies if every two KL-minimizing parameters disagree on the probability of some events, and if for every pair of strategies there is a KL-minimizing parameter under which they are not indifferent.

Proposition 6. *If either*

(i) $BR(\Delta(\Theta(p_\Pi))$ *is a singleton, or*

(ii) Θ *is finite and* p_Π *distinguishes parameters and strategies,*

then $\lim_{n \to \infty} \psi_n(\Theta, p_\Pi)$ *exists, and is in* $\Delta(BR(\Delta(\Theta(p_\Pi))))$.

This shows that if agents observe enough individual experiences from the previous generation, the aggregate distribution of strategies approaches $\Delta(BR(\Delta(\Theta(p_\Pi))))$. The case in which $BR(\Delta(\Theta(p_\Pi))$ is a singleton covers uniformly strict Berk-Nash equilibria, and provides a complete learning foundation for our results about them. To handle the case of multiple best replies to the KL minimizers, we add the assumption that every agent has a finite set of possible models and that π_Π distinguishes parameters and strategies.[30] We do not think that the finiteness assumption is necessary, but it simplifies the proof considerably. It is not needed if the best reply function is continuous in beliefs, as in Section 7, since then when the distribution of beliefs converges so does the distribution of best replies.

To prove this result, we use an argument similar to those of Berk (1966) and Esponda and Pouzo (2016) to show that the probability assigned to models that do not minimize the weighted KL divergence goes to 0. We then prove that although beliefs may not converge, their distribution does. We prove this by showing that the vector of likelihood ratios between KL minimizers is a random walk with positive definite covariance matrix, and applying the central limit theorem to obtain convergence.[31] The exact law of large numbers applied to the continuum of agents implies that the distribution of beliefs in the population converges as well. Finally, we show that the limit distribution assigns probability 0 to beliefs that induce ties between strategies, so the distribution of strategies converges.

[30] The Appendix proves this result under a more general condition that allows incomplete identification.

[31] Fudenberg, Lanzani, and Strack (2021) also combines the properties of a random walk with the properties of the KL divergence, but considers a different random walk (the difference between the realized empirical distribution and the objective distribution) and does not prove that the distribution of beliefs converges.

7 Infinitely many strategies

So far we have assumed there is a finite number of strategies. However, in some applications, there are many actions and/or signals, and it is more convenient to analyze the problem using a continuum approximation. We show here how our analysis can be applied to continuum environments. We assume that actions are real numbers and Θ is convex, as in many examples in the literature.

Assumption 2. (i) A is a compact subset of $\mathbb{R}$, with non-empty interior A^o.

(ii) S is a Borel subset of a Euclidean space, endowed with a full-support objective Borel probability measure σ. Π is the set of measurable functions from S to A.

(iii) u is continuously differentiable in a and s. The set $\{U^*(\pi) : \pi \in \Pi\}$ is bounded.

(iv) Θ is convex, and for all $\theta \in \Theta$, $BR(\theta)$ is a singleton and $Q_\theta(\cdot|a, s)$ is continuous in (a, s).

The results on one-hypothesis mutations extend immediately to real-valued actions. For local mutations, the cardinality of the action space does matter: With any finite set of actions a vanishingly small ε is eventually smaller than the "gap" between the actions, but this is not the case when the action space is an interval in $\mathbb{R}$. Instead, in any uniformly strict equilibrium there is a nearby action that performs almost as well, and arbitrarily small changes in beliefs generally induce a change in the best reply. As we show below, this allows local mutations to invade some uniformly strict equilibria in settings with a continuum of actions. We also show that any unstable uniformly strict equilibrium that is an attractor for the dynamic process corresponds to a limit of equilibria that are mixed and unstable along a sequence of increasingly fine finite action grids.[32]

The stability of an equilibrium in this setting depends on $\mathcal{M}_{\Theta,\psi}(\varepsilon)$, introduced in Section 5.2 and the (objective) indirect utility function of the agent, which is $V(\theta) = U^*(BR(\theta))$. We assume that V is continuously Gateaux differentiable. If $\Theta(\psi)$ is a singleton and $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is a singleton for sufficiently small ε, let $V'(\mathcal{M}_{\Theta,\psi}, \psi) = \liminf_{\varepsilon \to 0}(V(\mathcal{M}_{\Theta,\psi}(\varepsilon)) - V(\Theta(\psi))/\varepsilon)$ be the derivative of V in the direction $\mathcal{M}_{\Theta,\psi}(\varepsilon)$.

Proposition 7. *Let (Θ, ψ) be a Berk-Nash equilibrium such that $\Theta(\psi)$ is a singleton and $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is a singleton for sufficiently small ε. If $V'(\mathcal{M}_{\Theta,\psi}, \psi) > 0$ then (Θ, ψ) does not resist local mutations.*

[32] Convergence here means convergence with respect to the Hausdorff metric on the compact subsets of A. In some cases, there are ways of specifying the approximating action grid so that the unstable limit equilibrium is the limit of equilibria that are stable with finitely many actions, but these approximations rely on exactly including the equilibrium action of the continuum case as one of the elements of the grid.

This shows that if the derivative of the static indirect utility function in direction $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is positive the equilibrium does not resist local mutations.

Example 4. [Regression to the Mean] An instructor observes the initial performance $s \in \mathbb{R}$ of a student and decides whether to praise them, $a = a_r$, or criticize them, $a = a_c$. Then the student performs again, and the instructor observes their performance y. The instructor's utility is

$$u(s, a, y) = \begin{cases} y - k|s| & \text{if } s > 0 \text{ and } a = a_c, \text{ or } s < 0 \text{ and } a = a_r \\ y & \text{otherwise.} \end{cases}$$

The truth is that s and y are independent standard normals and the instructor cannot influence performance, so it is optimal to praise if $s > 0$.

The instructor believes that $y = \theta_0 s + \theta_a + \eta$, where η is a standard normal, θ_0 is the perceived correlation between performance in the two periods, and θ_a is the perceived effect of action a. Suppose the instructor is certain that $\theta_0 = 1$, with $\Theta = \{1\} \times [-K, +K]^2$. Esponda and Pouzo (2016) shows that for sufficiently large K the instructor criticizes too often in the unique Berk-Nash equilibrium: there is a threshold T such that the instructor criticizes if and only if performance is below $T = (\theta_{a_c}(T, \theta_0) - \theta_{a_r}(T, \theta_0))/k > 0$, where $\theta_{a_c}(T, \theta_0) = \mathbb{E}[y - \theta_0 s | s < T] > 0$ and $\theta_{a_r}(T, \theta_0) = \mathbb{E}[y - \theta_0 s | s > T] < 0$. Since $\theta_{a_r}(T, \theta_0)$ and $\theta_{a_c}(T, \theta_0)$ are respectively decreasing and increasing in θ_0, for sufficiently small ε, $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is $((1, \theta_{a_c}(T, \theta_0), \theta_{a_r}(T, \theta_0)) + (v_0(\varepsilon), v_r(\varepsilon), v_c(\varepsilon)))$ with $v_0(\varepsilon) < 0, v_c(\varepsilon) \leqslant 0, v_r(\varepsilon) \geqslant 0$. This corresponds to a lower correlation between the two periods, higher effectiveness of praise, and lower effectiveness of criticism. Since $V'((1, \theta_{a_c}(T, \theta_0), \theta_{a_r}(T, \theta_0)) + (v_0(\varepsilon), v_c(\varepsilon), v_r(\varepsilon))) = kT > 0$, by Proposition 7, the equilibrium does not resist local mutations.

Definition. Let S be a singleton. We say that a pure Berk Nash equilibrium $(\Theta, \hat{a})$ is an *attractor* if $\hat{a} \in A^o$ and there is an $\varepsilon > 0$ such that $||a - \hat{a}|| \leqslant \varepsilon$ implies $\Theta(a)$ is a singleton and $(BR(\Theta(a)) - \hat{a})(a - \hat{a}) < 0$.

A Berk-Nash equilibrium is an attractor if slightly changing the action in one direction induces a KL minimizer whose best reply is in the opposite direction.[33] In all continuum of actions versions of the examples of this paper, every interior Berk-Nash equilibrium is an attractor. We say that a sequence of finite action sets $(A_n)_{n \in \mathbb{N}}$ *approximates* A if for each $n \in \mathbb{N}$, A_n is a finite subset of A, and $||A_n - A|| \to 0$.

[33] When S is a singleton the space of strategies is unidimensional, so the direction of the deviation completely determines the direction of the best reply to the evidence it generates.

Proposition 8. *Suppose* $(\Theta, \hat{a})$ *is the unique Berk-Nash equilibrium, and is an attractor that satisfies the assumptions of Proposition 7. Then for every sufficiently small* $\varepsilon > 0$*, there is* $(A_n)_{n\in\mathbb{N}}$ *that approximates* A *and* $(\psi_n)_{n\in\mathbb{N}} \to \hat{a}$ *such that* (Θ, ψ_n) *is a Berk-Nash equilibrium of the environment with actions* A_n *that does not resist an* ε *expansion of* Θ.

8 Conclusion

We say that an equilibrium resists mutations if after a mutation the aggregate play converges back to the original equilibrium. This can happen because even explanation improving mutations that lead to a better but imperfect fit can lead to lower payoffs. We considered two sorts of mutations: local mutations that add all parameters close to the support of the original beliefs, and one-hypothesis mutations that completely abandon a particular constraint. These two forms of mutations have different implications for which equilibria resist mutations. Local mutations are effective at destabilizing mixed equilibria, but cannot destabilize Berk-Nash equilibria that are uniformly strict. One-hypothesis mutations can destabilize such equilibria, but they can fail to invade when they lead the agent to overshoot the optimal action, as in our income tax example.

The two forms of mutations we study are natural benchmarks, but other sorts of mutation may be worth exploring, such as mutations that weaken but do not entirely drop a single hypothesis. Another interesting case is mutations in which the new subjective model must include at least one parameter that perfectly fits the observed data, but may be incorrect about what is not observed in equilibrium. Here too we expect that some expansions of the parameter space may lead to lower payoffs and thus be abandoned, depending on how the expanded model interprets the equilibrium data.[34]

Our framework can be expanded to the misspecification-driven inattention proposed by Gagnon-Bartsch, Rabin, and Schwartzstein (2021), where agents only pay attention to the coarsest partition of outcomes that allows for all the inference they think is payoff relevant. Gagnon-Bartsch, Rabin, and Schwartzstein (2021) assumes that actions do not influence the distribution over outcomes, which implies that only equilibria that are self-confirming given the agents' "attention partition" resist mutations. It also assumes that all mutations include the objective model. An earlier version of this paper shows that when these assumptions are relaxed, the effect of attention partitions is ambiguous.

Our model's combination of Bayesian learning and evolutionary dynamics has a larger

[34]For an axiomatic approach to these "backup" models see Ortoleva (2012).

potential scope, as it can be used as a framework to study competition between paradigms without focusing, as we do, on steady states. For example, one might study cycles between subjective models in a setting without mutations, or the ergodic distribution of subjective models when mutations do sometimes occur but are rare. We could also extend the model to consider agents who have an intrinsic preference for particular beliefs, either innately or from peer effects.

A Appendix

Proof of Lemma 1. If $\delta_\Theta \times \psi$ is a steady state, then by equation (1), $\psi \in \Delta(BR(\Delta(\Theta(\psi))))$, so for every $\pi \in \operatorname{supp}\psi$ there exists $\mu_\pi \in \Delta(\Theta(\psi))$ such that $\pi \in BR(\mu_\pi)$, so (Θ, ψ) is a Berk-Nash equilibrium. The steady state is unitary if and only there is $\mu \in \Delta(\Theta)$ such that this μ_π can be chosen to be equal to μ for all $\pi \in \operatorname{supp}\psi$, so the equilibrium is unitary as well.

Conversely, if (Θ, ψ) is a Berk-Nash equilibrium, for every $\pi \in \operatorname{supp}\psi$ there exists $\mu_\pi \in \Delta(\Theta(\psi))$ such that $\pi \in BR(\mu_\pi)$, and so $\psi \in \Delta(BR(\Delta(\Theta(\psi))))$. Therefore, $p_t = \delta_\Theta \times \psi$ for all t satisfies equations (1) and (2), so $\delta_\Theta \times \psi$ is a steady state. The equilibrium is unitary if and only if we can choose $\mu = \mu_\pi$ for all $\pi \in \operatorname{supp}\psi$, so the steady state is unitary as well. ■

Proof of Proposition 1. Let (Θ, ψ) be a Berk-Nash equilibrium and suppose that the mutation of $\delta_\Theta \times \psi$ to Θ' is not explanation improving. We show that for every $\varepsilon \in (0, 1)$ the following constant path is a solution: $p^t_{\mathcal{K}}(\Theta) = 1 - \varepsilon, p^t_{\mathcal{K}}(\Theta') = \varepsilon$, and $p^t_\Pi(\cdot|\Theta) = \psi = p^t_\Pi(\cdot|\Theta')$ for all $t \in \mathbb{N}_0$. By Lemma 1, $\psi \in \Delta(BR(\Delta(\Theta(\psi))))$. Since $\Theta' \supseteq \Theta$ is not explanation improving with respect to $\delta_\Theta \times \psi$, $\Theta(\psi) \subseteq \Theta'(\psi)$, and $\psi \in \Delta(BR(\Delta(\Theta(\psi)))) \subseteq \Delta(BR(\Delta(\Theta'(\psi))))$, so equation (1) is satisfied, and equation (2) is satisfied because the distributions of strategies generated by the two models are the same at every period. ■

Lemma 2. $\Theta(\cdot)$ *is upper hemicontinuous, nonempty-valued, and compact-valued.*

The proof of this lemma is an immediate adaptation of Lemma 1 of Esponda and Pouzo (2016) to the case of infinitely many outcomes. Define $\Theta(\pi, \varepsilon) := \{\theta \in \mathcal{H} : \exists \theta' \in \Theta(\pi), ||\theta - \theta'||_2 \leqslant \varepsilon\}$.

Lemma 3. *For every* $\Theta \in \mathcal{K}$ *and* $\varepsilon' > 0$ *there is an* $\hat{\varepsilon} \in (0, \varepsilon')$ *such that if* Θ' *is an* $\varepsilon < \hat{\varepsilon}$ *local expansion of* Θ*, then* $\Theta'(\pi) \subseteq \Theta(\pi, \hat{\varepsilon})$.

Proof. Suppose not, and let $(\Theta'_n)_{n\in\mathbb{N}}$ be a sequence of ε_n local expansions of Θ with $\varepsilon_n \downarrow 0$ and $\theta_n \in \Theta'_n(\pi)\backslash\Theta(\pi,\hat{\varepsilon})$. Since Θ'_1 is compact, the sequence has an accumulation point $\theta \in \Theta$. If $\theta \in \Theta(\pi)$ then a subsequence of $(\theta_n)_{n\in\mathbb{N}}$ is eventually in $\Theta(\pi,\hat{\varepsilon})$, a contradiction. If $\theta \notin \Theta(\pi)$, then since H is lower semi-continuous in θ,[35] eventually $H_\pi(Q^*, Q_{\theta_n}) > \min_{\tilde{\theta}\in\Theta} H_\pi(Q^*, Q_{\tilde{\theta}}) \geqslant \min_{\tilde{\theta}\in\Theta'_n} H_\pi(Q^*, Q_{\tilde{\theta}})$, which contradicts with $\theta_n \in \Theta'_n(\pi)$. ■

Proof of Proposition 2. (i) Let (Θ,π) be a uniformly strict Berk-Nash equilibrium. By Lemma 2, $\Theta(\pi)$ is compact, and by the triangle inequality so is $\Theta(\pi,\varepsilon)$. The result is immediate if Π is a singleton. Otherwise, let $G(\varepsilon) = \min_{\pi'\in\Pi\backslash\{\pi\}} \min_{\mu\in\Delta(\Theta(\pi,\varepsilon))} (U_\mu(\pi) - U_\mu(\pi'))$. Because Π is finite and U is linear and bounded on $\Delta(\Theta)$, U is continuous by Lemma 5.64 in Aliprantis and Border (2013). Moreover, $\varepsilon \mapsto \Theta(\pi,\varepsilon)$ is a continuous and compact-valued correspondence, and so G is continuous by the Maximum Theorem. And since (Θ,π) is a uniformly strict Berk-Nash equilibrium, $G(0) > 0$, and there is an $\hat{\varepsilon}$ such that if $\varepsilon \leqslant \hat{\varepsilon}$, $G(\varepsilon) > 0$.

By Lemma 3 there is an $\varepsilon' \in (0,\hat{\varepsilon})$ such that if Θ' is an $\varepsilon < \varepsilon'$ local expansion of Θ, then $\Theta'(\pi) \subseteq \Theta(\pi,\hat{\varepsilon})$. Let p_ε be an ε local mutation of $\delta_\Theta \times \pi$ for $\varepsilon < \varepsilon'$ and let Θ' be the ε local expansion of Θ. We prove by induction that $p^t_\Pi = \pi$ for every solution $(p^t)_{t\in\mathbb{N}}$ with $p^0 = p_\varepsilon$, concluding the proof of the statement. For the initial step, note that since $\varepsilon < \varepsilon' \leqslant \hat{\varepsilon}$, $\Theta'(\pi) \subseteq \Theta(\pi,\hat{\varepsilon})$. But then $p_\varepsilon(\cdot|\Theta') \in \Delta(BR(\Delta(\Theta'(\pi))) \subseteq \Delta(BR(\Delta(\Theta(\pi,\hat{\varepsilon})))) = \{\pi\}$, where the last equality follows from $G(\hat{\varepsilon}) > 0$. Moreover, since (Θ,π) is a uniformly strict Berk-Nash equilibrium, $p_\varepsilon(\cdot|\Theta) = \{\pi\}$ as well, concluding the base step. Suppose the statement is true for some $t \in \mathbb{N}_0$. Since $\varepsilon < \varepsilon' \leqslant \hat{\varepsilon}$ we have $\Theta'(p^t_\Pi) = \Theta'(\pi) \subseteq \Theta(\pi,\hat{\varepsilon})$, and by definition $\Theta(\pi) \subseteq \Theta(\pi,\hat{\varepsilon})$. Since $G(\hat{\varepsilon}) > 0$, this implies $p^{t+1}_\Pi = \{\pi\}$. Since (Θ,π) is a uniformly strict Berk-Nash equilibrium, $p_\varepsilon(\cdot|\Theta) = \{\pi\}$, which completes the inductive step.

(ii) Let $\partial\Theta$ denote the boundary of Θ. Since $\partial\Theta$ is compact, when the KL minimizers are in the interior of Θ, there is a $K \in \mathbb{R}_{++}$ and an $\hat{\varepsilon} \in \mathbb{R}_{++}$ such that if $\theta' \in \partial\Theta$ and θ is in $B_{\hat{\varepsilon}}(\theta')$ then $H_\psi(Q^*, Q_\theta) - \operatorname{argmin}_{\tilde{\theta}\in\Theta} H_\psi(Q^*, Q_{\tilde{\theta}}) > K$. This in turn implies that $\Theta_\varepsilon(\psi) = \Theta(\psi)$, when $\varepsilon < \hat{\varepsilon}$. Thus the sequence in which both the mutated and the conformist agents play ψ every period and their shares remain fixed is a solution: equation (1) is satisfied since $\Theta_\varepsilon(\psi) = \Theta(\psi)$, and equation (2) is trivially satisfied since both subpopulations have the same distribution over strategies. Therefore the equilibrium resists local mutations. ■

For $\lambda \in \mathbb{R}_{++}, \alpha \in (0,1)$ and $\Theta, \Theta' \in \mathcal{K}$, let

[35] The lower semicontinuity of H in the probability measure follows from e.g., Theorem 1.47 of Liese and Vajda (1987), and then the lower semicontinuity in θ follows from the continuity in θ of the probability measure imposed by Assumption 1(i).

$$P_{\lambda,\alpha}(\Theta,\Theta') = \{p \in P : p_{\mathcal{K}}(\{\Theta,\Theta'\}) = 1, U^*(p(\cdot|\Theta')) - U^*(p(\cdot|\Theta)) \geqslant \lambda, \min\{p_{\mathcal{K}}(\Theta), p_{\mathcal{K}}(\Theta')\} \geqslant \alpha\}$$

denote the states where the strategy used by agents with model Θ' outperforms the strategy used by agents with model Θ by at least λ, and both population shares are larger than α.

The proofs of Propositions 3, 4, and 7 use the following lemma. It uses continuity and compactness arguments and the payoff monotonicity of T to show that the change in the relative prevalence of any two subjective models Θ, Θ' is bounded away from 1 on $P_{\lambda,\alpha}(\Theta,\Theta')$, regardless of their initial population shares.

Lemma 4. *For every $\lambda, \alpha \in (0,1)$ and $\Theta, \Theta' \in \mathcal{K}$, $\min_{p \in P_{\lambda,\alpha}(\Theta,\Theta')}[T(p)(\Theta')p_{\mathcal{K}}(\Theta)]/[T(p)(\Theta)p_{\mathcal{K}}(\Theta')]$ is well defined and strictly larger than 1. Thus, there is no solution that eventually stays in $P_{\lambda,\alpha}(\Theta,\Theta')$.*

Proof. Because U^* is continuous, $P_{\lambda,\alpha}(\Theta,\Theta')$ is a compact subset of P. Therefore, since $T(\cdot)(\Theta)$ is continuous and strictly positive on $P_{\lambda,\alpha}(\Theta,\Theta')$, it is bounded away from 0 on this set. Moreover, by definition $p_{\mathcal{K}}(\Theta') \geqslant \alpha > 0$ on $P_{\lambda,\alpha}(\Theta,\Theta')$, so $[T(p)(\Theta')p_{\mathcal{K}}(\Theta)]/[T(p)(\Theta)p_{\mathcal{K}}(\Theta')]$ is continuous on $P_{\lambda,\alpha}(\Theta,\Theta')$ and attains a minimum m. Because T is payoff monotone and the strategy used by agents with subjective model Θ' strictly outperforms the strategy used by agents with subjective model Θ on $P_{\lambda,\alpha}(\Theta,\Theta')$, the minimum satisfies $m > 1$.

To prove the last part of the lemma, note that when $p_t \in P_{\lambda,\alpha}(\Theta,\Theta')$ the ratio between the shares of models Θ and Θ' grows multiplicatively by at least m, so there is a $\tau \leqslant \log_m(1/\alpha^2)+t$ such that $p_\tau(\Theta) < \alpha$. Thus the solution leaves $P_{\lambda,\alpha}(\Theta,\Theta')$ before time τ. ∎

Proof of Proposition 3. Since the set of strategies is finite, there is $\varepsilon' \in \mathbb{R}_{++}$ such that for all $\varepsilon \in (0,\varepsilon')$, $BR(\Delta(\mathcal{M}_{\Theta,\psi}(\varepsilon))) \subseteq \Pi_{\mathcal{M}_{\Theta,\psi}}$. Fix such an ε' for the entire proof.

(i) Since U^* is continuous and $\Pi_{\mathcal{M}_{\Theta,\psi}}$ is finite, there are $\varepsilon^* \in \mathbb{R}_{++}$ and $\gamma \in \mathbb{R}_{++}$ such that

$$||\psi' - \psi|| < \varepsilon^* \Rightarrow U^*(\psi') - \min_{\pi \in \Pi_{\mathcal{M}_{\Theta,\psi}}} U^*(\pi) < -\gamma. \tag{4}$$

Let $(p^t)_{t\in\mathbb{N}_0}$ be a solution with $p^0 = p_\varepsilon$, where p_ε is an ε local mutation of $\delta_\Theta \times \psi$ and $\varepsilon < \min\left\{\varepsilon', \varepsilon^*, \left(1 - \psi\left(\Pi_{\mathcal{M}_{\Theta,\psi}}\right)\right)/2\right\}$. Because Π is finite and $\Theta_\varepsilon(\cdot)$ is upper-hemicontinuous by Lemma 2, there is $\bar{\varepsilon} \in (0,\varepsilon)$ such that

$$||\psi' - \psi|| < \bar{\varepsilon} \Rightarrow BR\left(\Delta\left(\Theta_\varepsilon\left(\psi'\right)\right)\right) \subseteq \Pi_{\mathcal{M}_{\Theta,\psi}}. \tag{5}$$

Suppose by way of contradiction that $\lim p_\Pi^t = \psi$ and so $||p_\Pi^t - \psi|| < \bar{\varepsilon}$ for all t larger than some $\tau > 0$. For such t, $BR\left(\Delta\left(\Theta_\varepsilon\left(p_\Pi^t\right)\right)\right) \subseteq \Pi_{\mathcal{M}_{\Theta,\psi}}$ from equation (5), and since $\bar{\varepsilon} < \varepsilon < \varepsilon^*$, equation (4) implies:

$$U^*\left(p_\varepsilon^{t+1}(\cdot|\Theta_\varepsilon)\right) > U^*\left((p_\varepsilon^{t+1})_\Pi\right) + \gamma = p_\varepsilon^{t+1}(\Theta)U^*\left(p_\varepsilon^{t+1}(\cdot|\Theta)\right) + (1-p_\varepsilon^{t+1}(\Theta))U^*\left(p_\varepsilon^{t+1}(\cdot|\Theta_\varepsilon)\right) + \gamma$$

$$\text{so } U^*\left(p_\varepsilon^{t+1}(\cdot|\Theta_\varepsilon)\right) > U^*\left(p_\varepsilon^{t+1}(\cdot|\Theta)\right) + \frac{\gamma}{p_\varepsilon^{t+1}(\Theta)}. \tag{6}$$

Moreover, by equation (5) for all $t > \tau$, the mutated agents only play strategies in $\Pi_{\mathcal{M}_{\Theta,\psi}}$, i.e., $\operatorname{supp} p_\varepsilon^{t+1}(\cdot|\Theta_\varepsilon) \subseteq BR\left(\Delta\left(\Theta_\varepsilon\left(p_\Pi^t\right)\right)\right) \subseteq \Pi_{\mathcal{M}_{\Theta,\psi}}$. This, together with $U^*(\pi) > U^*(\psi)$ for every $\pi \in \Pi_{\mathcal{M}_{\Theta,\psi}}$, implies that $p_\varepsilon^{t+1}(\Theta) > \left(1-\psi\left(\Pi_{\mathcal{M}_{\Theta,\psi}}\right)\right)/2 > 0$. But then $p^t \in P_{\lambda,\alpha}(\Theta,\Theta_\varepsilon)$ for all $t > \tau$ with $\alpha = \min\left\{p_\varepsilon^{\tau+1}(\Theta_\varepsilon), \left(1-\psi\left(\Pi_{\mathcal{M}_{\Theta,\psi}}\right)\right)/2\right\}$ and $\lambda = \gamma/p_\varepsilon^{\tau+1}(\Theta)$, a contradiction by Lemma 4.

(ii) We prove the following stronger result: If for some $\pi' \in \Pi_{\mathcal{M}_{\Theta,\psi}}$, $U^*(\pi') \leqslant U^*(\psi)$, and $\pi' \in \operatorname{supp}\psi$, then (Θ,ψ) resists local mutations. Part (ii) of the Proposition follows because the upper hemicontinuity of $\mathcal{M}_{\Theta,\psi}(\cdot)$ and $BR(\cdot)$ implies that $\pi' \in \operatorname{supp}\psi$ in any quasi-strict equilibrium.

Suppose that $U^*(\pi') \leqslant U^*(\psi)$, and $\pi' \in \operatorname{supp}\psi$ for some $\pi' \in \Pi_{\mathcal{M}_{\Theta,\psi}}$. We will show that there is a solution in which the mutated agents always play $\pi' \in \Pi_{\mathcal{M}_{\Theta,\psi}}$ and the conformists play a strategy distribution very close to the equilibrium in every period. Upper hemicontinuity of the best reply make the case in period 1 for a sufficiently small share of mutants. Payoff monotonicity and the fact that π' has lower payoff than the equilibrium distribution guarantee that the share of mutants does not increase.

Let $\hat{\varepsilon} = \min_{\pi\in\operatorname{supp}\psi}\psi(\pi)$, and let $(\varepsilon_n)_{n\in\mathbb{N}_0} \in (0,\min\{\varepsilon',\hat{\varepsilon}\})^{\mathbb{N}_0}$ be such that $\pi' \in BR(\Delta(\mathcal{M}_{\Theta,\psi}(\varepsilon_n)))$ for all $n \in \mathbb{N}_0$ and $(\varepsilon_n)_{n\in\mathbb{N}_0} \to 0$. We now show that for every ε_n there is a solution $(p^t)_{t\in\mathbb{N}_0}$ with $p^0 = p_{\varepsilon_n}$ and $\lim_{t\to\infty} p_\Pi^t = \psi$, where p_{ε_n} is an ε_n local mutation of $\delta_\Theta \times \psi$.

Set $\varepsilon = \varepsilon_n$ for some $n \in \mathbb{N}_0$ and let Θ_ε be the local ε expansion of Θ. To define the candidate solution, let $p_\mathcal{K}^0(\Theta) = 1-\varepsilon, p_\mathcal{K}^0(\Theta_\varepsilon) = \varepsilon$, $p_\Pi^0(\cdot|\Theta_\varepsilon) = \pi'$ and

$$p_\Pi^0(\pi|\Theta) = \begin{cases} \frac{\psi(\pi)}{1-\varepsilon} & \pi \neq \pi' \\ \frac{\psi(\pi)-\varepsilon}{1-\varepsilon} & \pi = \pi'. \end{cases} \tag{7}$$

Also, recursively define $p_\mathcal{K}^{t+1}(\Theta) = T(p^t)(\Theta)$, $p_\mathcal{K}^{t+1}(\Theta_\varepsilon) = T(p^t)(\Theta_\varepsilon)$

$$p_\Pi^{t+1}(\cdot|\Theta_\varepsilon) = \pi', \text{ and } p_\Pi^{t+1}(\pi|\Theta) = \begin{cases} \frac{\psi(\pi)}{p_\mathcal{K}^{t+1}(\Theta)} & \pi \neq \pi' \\ \frac{\psi(\pi)-p_\mathcal{K}^{t+1}(\Theta_\varepsilon)}{p_\mathcal{K}^{t+1}(\Theta)} & \pi = \pi'. \end{cases}$$

By definition, $(p^t)_{t\in\mathbb{N}_0}$ satisfies equation (2). It remains to show that $\psi(\pi') \geqslant p_\mathcal{K}^t(\Theta_\varepsilon)$ for all

t, so that $p^{t+1}(\cdot|\Theta)$ is a well defined element of $\Delta(\Pi)$, and that equation (1) is satisfied.

We prove this by induction. For the initial step, since Θ_ε is an ε local mutation of Θ, and $\varepsilon < \hat{\varepsilon} = \min_{\pi \in \operatorname{supp} \psi} \psi(\pi)$, $\psi(\pi') \geqslant p^0_{\mathcal{K}}(\Theta_\varepsilon) = \varepsilon$. Moreover, by definition of $\varepsilon = \varepsilon_n$ equation (1) is satisfied for $t = 0$. For the inductive step, since the operator T is payoff monotone, $U^*(\pi') \leqslant U^*(\psi)$, and $\psi(\pi') \geqslant p^t_{\mathcal{K}}(\Theta_\varepsilon)$ by the inductive hypothesis, we have $p^{t+1}_{\mathcal{K}}(\Theta_\varepsilon) \leqslant p^t_{\mathcal{K}}(\Theta_\varepsilon) \leqslant \psi(\pi')$, and because $\varepsilon = \varepsilon_n$ equation (1) is satisfied for $t + 1$. ■

Lemma 5. *Let (Θ, ψ) be a Berk-Nash equilibrium. If Θ is finite, $\Theta(\psi)$ is a singleton, Q_θ is linear in θ, and for every $\theta' \neq \theta''$ $Q_{\theta'}$ and $Q_{\theta''}$ are not Q^*-almost surely equal, then there is $\varepsilon' \in \mathbb{R}^{++}$ such that $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is a singleton for all $\varepsilon \leqslant \varepsilon'$.*

Proof. Let $\{\hat{\theta}\} = \Theta(\psi)$. By Lemma 3 there is $\varepsilon' > 0$ such that for all $\varepsilon \leqslant \varepsilon'$, $\mathcal{M}_{\Theta,\psi}(\varepsilon) \subseteq B_\varepsilon(\hat{\theta})$. Assume by way of contradiction that $\theta', \theta'' \in \mathcal{M}_{\Theta,\psi}(\varepsilon)$, $\varepsilon \leqslant \varepsilon'$, $\theta' \neq \theta''$. By the convexity of the norm, $\theta'/2 + \theta''/2 \in B_\varepsilon(\hat{\theta})$, and by the strict convexity of the weighted KL divergence in the second argument, $H_\psi\left(Q^*, Q_{\theta'/2+\theta''/2}\right) < H_\psi\left(Q^*, Q_{\theta'}\right)$, a contradiction. ■

Lemma 6. *Assume that for every $a \in A$, $s \in S$, and $\hat{\theta} \in \mathbb{R}^k$, if $H(Q^*(\cdot|s,a), Q_{\hat{\theta}}(\cdot|s,a))$ is finite then $\theta \mapsto H(Q^*(\cdot|s,a), Q_\theta(\cdot|s,a))$ is twice continuously Gateaux differentiable at $\hat{\theta}$. Let (Θ, ψ) be a Berk-Nash equilibrium. If Θ is finite and* $\operatorname{argmax}_{\theta \in \Theta(\psi), v \in \mathcal{S}} D_\psi(\theta, v) = \{(\hat{\theta}, \hat{v})\}$ *is a singleton then $\Pi_{\mathcal{M}_{\Theta,\psi}} \subseteq \limsup_{\varepsilon \to 0} \operatorname{argmax}_{\pi \in \Pi} U_{\delta_{\hat{\theta}+\varepsilon\hat{v}}}(\pi)$.*

Proof. By Assumption 1(iii) $H_\psi(Q^*, Q_\theta)$ is finite at the KL minimizers, so it is Gateaux continuously differentiable at $\hat{\theta}$, and $D_\psi\left(\hat{\theta}, \cdot\right)$ is continuous.

We now prove that for all $c \in \mathbb{R}_{++}$, there exists $\bar{\varepsilon} \in \mathbb{R}_{++}$ such that for all $\varepsilon < \bar{\varepsilon}$, and $\theta \in \mathcal{M}_{\Theta,\psi}(\varepsilon)$, $\left|\left|\frac{(\theta-\hat{\theta})}{\varepsilon} - \hat{v}\right|\right| < c$. Suppose by contradiction that there exist $(\varepsilon_n)_{n\in\mathbb{N}} \to 0$ and $(\theta_n)_{n\in\mathbb{N}}$ with $\theta_n \in \Theta_{\varepsilon_n}(\psi)$ and $\left|\left|\frac{(\theta_n-\hat{\theta})}{\varepsilon_n} - \hat{v}\right|\right| \geqslant c$. Since $\Theta_{\varepsilon_1}(\psi)$ is compact by Lemma 2, the $(\theta_n)_{n\in\mathbb{N}}$ can be taken to be convergent to some $\tilde{\theta}$. Let $v_n = (\theta_n - \tilde{\theta})/||\theta_n - \tilde{\theta}||_2$. Again, by possibly passing to a subsequence v_n can be taken to be convergent to some $\tilde{v}$. Observe that by construction we must have $(\hat{\theta}, \hat{v}) \neq (\tilde{\theta}, \tilde{v})$. By definition of θ_n we have

$$\sum_{s\in S} \sigma(s) \sum_{\pi\in\Pi} \psi(\pi) \left(\int_{y\in Y} \log \frac{q_{\hat{\theta}+\hat{v}\varepsilon_n}(y|s,\pi(s))}{q_{\theta_n}(y|s,\pi(s))} dQ^*(y|s,\pi(s)) \right) \leqslant 0$$

and since $H(Q^*(\cdot|s,a), Q_\theta(\cdot|s,a))$ is twice continuously differentiable at $\hat{\theta}$ and $\tilde{\theta}$ this implies that $D_\psi(\hat{\theta}, \hat{v}) \leqslant D_\psi(\tilde{\theta}, \tilde{v})$ a contradiction with $\operatorname{argmax}_{\theta\in\Theta(\psi), v\in\mathcal{S}\cup\{0\}} D_\psi(\theta, v) = \{(\hat{\theta}, \hat{v})\}$.

Then the result follows by the upperhemicontinuity of BR and $\mathcal{M}_{\Theta,\psi}$. ■

Proof of Proposition 4. Because $\Theta^l(\cdot)$ is upper-hemicontinuous (see Lemma 2), there is $\varepsilon' > 0$ such that

$$||\psi' - \psi|| < \varepsilon' \Rightarrow BR\left(\Delta\left(\Theta^l\left(\psi'\right)\right)\right) \subseteq \Pi_{p,l} \qquad \forall l \in \{1, ..., k\}. \tag{8}$$

(i) We prove this case by contradiction. Suppose that for some $l \in \{1, ..., k\}$, we have $U^*(\pi) > U^*(\psi)$ for every $\pi \in \Pi_{p,l}$. Because U^* is continuous and Π is finite, there are $\varepsilon^* > 0$ and $\gamma > 0$ such that

$$||\psi' - \psi|| < \varepsilon^* \Rightarrow U^*\left(\psi'\right) - U^*\left(\psi''\right) < -\gamma \qquad \forall \psi'' \in \Delta(\Pi_{p,l}). \tag{9}$$

Let Θ^l be the one-hypothesis relaxation of Θ in hypothesis l. By definition of ε' we have

$$||p^t_\Pi - \psi|| < \varepsilon' \Rightarrow p^{t+1}(\cdot|\Theta^l) \in \Delta(\Pi_{p,l}) \qquad \forall p^t \in P. \tag{10}$$

Suppose by way of contradiction that for $\varepsilon < \min\{\varepsilon', \varepsilon^*, (1 - \psi(\Pi_{p,l}))/2\}$ there is an ε mutation to Θ^l, p^0_ε, such that $\lim_{t\to\infty}(p^t_\varepsilon)_\Pi = \psi$. This means that after some $\tau > 0$, for all $t > \tau$, $||(p^t_\varepsilon)_\Pi - \psi|| < \varepsilon$. Since $\varepsilon < \varepsilon'$, by equation (10), $p^{t+1}_\varepsilon(\cdot|\Theta^l) \in \Delta(\Pi_{p,l})$. This, together with the assumption that for every $\pi \in \Pi_{p,l}$, $U^*(\pi) > U^*(\psi)$, implies that $p^{t+1}_\varepsilon(\Theta) > (1 - \psi(\Pi_{p,l}))/2 > 0$.

Since $\varepsilon < \varepsilon^*$, by equation (9), $U^*\left((p^{t+1}_\varepsilon)_\Pi\right) + \gamma < U^*\left(p^{t+1}_\varepsilon(\cdot|\Theta^l)\right)$. Therefore, $U^*\left(p^{t+1}_\varepsilon(\cdot|\Theta^l)\right) > p^{t+1}_\varepsilon(\Theta)U^*\left(p^{t+1}_\varepsilon(\cdot|\Theta)\right) + (1 - p^{t+1}_\varepsilon(\Theta))U^*\left(p^{t+1}_\varepsilon(\cdot|\Theta^l)\right) + \gamma$, so

$$U^*\left(p^{t+1}_\varepsilon(\cdot|\Theta^l)\right) > U^*\left(p^{t+1}_\varepsilon(\cdot|\Theta)\right) + \frac{\gamma}{p^{t+1}_\varepsilon(\Theta)}.$$

But then $p^t \in P_{\lambda,\alpha}(\Theta, \Theta^l)$ for all $t > \tau$ with $\alpha = \min\left\{p^{\tau+1}_\varepsilon(\Theta^l), (1 - \psi(\Pi_{p,l}))/2\right\}$ and $\lambda = \gamma/p^{\tau+1}_\varepsilon(\Theta)$, a contradiction by Lemma 4.

(ii.a) Let $\hat{\varepsilon} = \min_{\pi\in\operatorname{supp}\psi}\psi(\pi)$. We will show that for every $l \in \{1, ..., k\}$ and $\varepsilon < \min\{\varepsilon', \hat{\varepsilon}\}$ there exists a solution $(p^t)_{t\in\mathbb{N}_0}$ where p^0 is an ε mutation of $\delta_\Theta \times \psi$ to the one-hypothesis relaxation of Θ in hypothesis $l \in \{1, ..., m\}$ and $\lim_{t\to\infty}(p^t_\varepsilon)_\Pi = \psi$. Fix such an ε and let Θ' be the one-hypothesis relaxation of Θ in hypothesis $l \in \{1, ..., m\}$. Initialize the candidate solution by setting $p^0_\mathcal{K}(\Theta) = 1-\varepsilon$, $p^0_\mathcal{K}(\Theta') = \varepsilon$, $p^0_\Pi(\cdot|\Theta') = \pi' \in \Pi_{p,l}$ and $p^0_\Pi(\pi|\Theta)$ as in equation (7), and recursively define subsequent states by $p^{t+1}_\mathcal{K}(\Theta)/p^{t+1}_\mathcal{K}(\Theta') = T(p^t)(\Theta)/T(p^t)(\Theta')$, $p^{t+1}_\Pi(\cdot|\Theta') = \pi'$ and

$$p_\Pi^{t+1}(\pi|\Theta) = \begin{cases} \frac{\psi(\pi)}{p_{\mathcal{K}}^{t+1}(\Theta)} & \pi \neq \pi' \\ \frac{\psi(\pi) - p_{\mathcal{K}}^{t+1}(\Theta')}{p_{\mathcal{K}}^{t+1}(\Theta)} & \pi = \pi'. \end{cases}$$

By definition, $(p^t)_{t\in\mathbb{N}_0}$ satisfies equation (2). It only remains to show that $\psi(\pi') \geqslant p_{\mathcal{K}}^t(\Theta')$ for all t, so that $p^{t+1}(\cdot|\Theta)$ is a well defined element of $\Delta(\Pi)$, and that equation (1) is satisfied.

We prove this by induction. For the initial step, since $\varepsilon < \hat{\varepsilon} = \min_{\pi\in\operatorname{supp}\psi}\psi(\pi)$, $\psi(\pi') \geqslant p_{\mathcal{K}}^0(\Theta') = \varepsilon$. Moreover, since $\pi' \in \Pi_{p,l}$, equation (1) is satisfied for $t = 0$. For the inductive step, since T is payoff monotone, $U^*(\pi') \leqslant U^*(\psi)$, and $\psi(\pi') \geqslant p_{\mathcal{K}}^t(\Theta')$ by the inductive hypothesis, we have $p_{\mathcal{K}}^{t+1}(\Theta') \leqslant p_{\mathcal{K}}^t(\Theta') \leqslant \psi(\pi')$. Moreover, since by the inductive step $(p^t)_\Pi = \psi$ and $\pi' \in \Pi_{p,l}$, equation (1) is satisfied for $t+1$ as well.

(ii.b) Now suppose (Θ, ψ) is a uniformly strict Berk-Nash equilibrium. Let Θ' be the one-hypothesis relaxation of Θ in hypothesis $l \in \{1, ..., m\}$. Let $\hat{\varepsilon} > 0$ be small enough that $||\psi - \psi'|| \leqslant \hat{\varepsilon}$ implies $BR(\Delta(\Theta(\psi')) \subseteq BR(\Delta(\Theta(\psi)))$ and $BR(\Delta(\Theta'(\psi')) \subseteq \Pi_{p,l}$ for every $l \in \{1, ..., k\}$. (Such $\hat{\varepsilon}$ exists by Lemma 2.) We show that for every $\varepsilon < \min\{\varepsilon', \hat{\varepsilon}\}$ there is a solution $(p^t)_{t\in\mathbb{N}_0}$ where (i) p^0 is a mutation of $\delta_\Theta \times \psi$ to the one-hypothesis relaxation of Θ in hypothesis l and (ii) $\lim_{t\to\infty}(p_\varepsilon^t)_\Pi = \psi$. Fix such an ε.

Initialize the candidate solution by setting $p_{\mathcal{K}}^0(\Theta) = 1 - \varepsilon, p_{\mathcal{K}}^0(\Theta') = \varepsilon$, $p_\Pi^0(\cdot|\Theta') = \pi_0$ and $p_\Pi^0(\cdot|\Theta) = \psi$, where π_0 is an arbitrary element of $\Pi_{p,l}$, and recursively define subsequent states by $p_{\mathcal{K}}^{t+1}(\Theta) = T(p^t)(\Theta)$ and $p_{\mathcal{K}}^{t+1}(\Theta') = T(p^t)(\Theta')$, $p_\Pi^{t+1}(\cdot|\Theta') = \pi_t$ and $p_\Pi^{t+1}(\pi|\Theta) = \psi$, where π_t is an arbitrary element of $BR(\Delta(\Theta'(p_\Pi^t)) \cap \Pi_{p,l}$. By definition, $(p^t)_{t\in\mathbb{N}_0}$ satisfies equation (2). It only remains to show that for every $t \in \mathbb{N}_0$, $\hat{\varepsilon} \geqslant p_{\mathcal{K}}^t(\Theta')$ so that equation (1) is satisfied because $BR(\Delta(\Theta'(p_\Pi^t)) \cap \Pi_{p,l} \neq \varnothing$ and $\psi \in \Delta(BR(\Delta(\Theta(p_\Pi^t))))$. We prove this by induction. The initial step follows from the definition of $\hat{\varepsilon}$. For the inductive step, since the operator T is payoff monotone, $U^*(\pi') \leqslant U^*(\psi)$, for every $\pi' \in \Pi_{p,l}$, and $\hat{\varepsilon} \geqslant \varepsilon \geqslant p_{\mathcal{K}}^t(\Theta')$ by the inductive hypothesis, we have $p_{\mathcal{K}}^{t+1}(\Theta') \leqslant p_{\mathcal{K}}^t(\Theta') \leqslant \varepsilon \leqslant \hat{\varepsilon}$. The proof is concluded by observing that the previous inequality implies that as $\varepsilon \to 0$ we have $\lim_{t\to\infty} p_{\mathcal{K}}^t(\Theta') = 0$. ■

Proof of Proposition 5. Let Θ' be innovation-inducing for the innovation-vulnerable equilibrium (Θ, ψ). Since U^* is continuous and Θ' is upper hemicontinuous by Lemma 2, there is $\varepsilon^* > 0$ such that

$$||\psi' - \psi|| < \varepsilon^* \Rightarrow BR(\Delta(\Theta'(\psi'))) = \{\pi_I\}. \tag{11}$$

Let $(p^t)_{t\in\mathbb{N}_0}$ be a solution with $p^0 = p_\varepsilon$, where p_ε is an ε mutation of $\delta_\Theta \times \psi$ to Θ' and $\varepsilon < \varepsilon^*$.

By assumption, there is $\bar{\varepsilon} \in (0, \varepsilon)$ such that

$$||p_\Pi - \psi|| < \bar{\varepsilon} \text{ and } p(\cdot|\Theta') = \{\pi_I\} \Rightarrow BR\left(\Delta\left(\Theta\left(p_\Pi\right)\right)\right) = \{\pi_U\}. \tag{12}$$

Suppose by way of contradiction that after some $\tau > 0$, for all $t > \tau$, $||p_\Pi^t - \psi|| < \bar{\varepsilon}$. From equation (11) this implies that $BR\left(\Delta\left(\Theta'\left(p_\Pi^t\right)\right)\right) = \{\pi_I\}$. Since $\bar{\varepsilon} < \varepsilon < \varepsilon^*$, by equation (12), $BR\left(\Delta\left(\Theta\left(p_\Pi^t\right)\right)\right) = \{\pi_U\}$ a contradiction to $||p_\Pi^{t+1} - \psi|| < \bar{\varepsilon}$. ■

Proofs of Proposition 6 and 9. Let $\mathcal{E}\left(\theta, \psi\right)$ be the parameters that are indistinguishable from θ under strategy distribution ψ, i.e., the θ' such that for all $s \in S$ and for $\sum_{\pi \in \Pi} Q^*(\cdot|s, \pi(s))\psi(\pi)$-almost every y, $\sum_{\pi \in \Pi} Q_\theta(\cdot|s, \pi(s))\psi(\pi) = \sum_{\pi \in \Pi} Q_{\theta'}(\cdot|s, \pi(s))\psi(\pi)$. When aggregate play is ψ, the relative likelihood of elements of $\mathcal{E}\left(\theta, \psi\right)$ is determined by the prior. Let $U(\pi|\mathcal{E}\left(\theta, \psi\right))$ denote the subjective utility of strategy π under posterior $\mu_\Theta\left(\cdot|\mathcal{E}\left(\theta, p_\Pi\right)\right)$.

Definition. We say that *strategies are subjectively different under* p_Π if for all $\pi, \pi' \in \Pi, \pi \neq \pi'$, there is $\theta \in \Theta(p_\Pi)$ such that $U(\pi|\mathcal{E}\left(\theta, p_\Pi\right)) \neq U(\pi'|\mathcal{E}\left(\theta, p_\Pi\right))$.

In words, strategies are subjectively different under p_Π if for every two strategies there is a class of indistinguishable parameters $\mathcal{E}\left(\theta, p_\Pi\right)$ that minimize the weighted KL divergence given p_Π such that the utility of the strategies conditional to $\mathcal{E}\left(\theta, p_\Pi\right)$ is different. If p_Π distinguishes parameters and strategies, strategies are subjectively different under p_Π, as the former is the case when each $\mathcal{E}\left(\theta, p_\Pi\right)$ is a singleton. Thus Proposition 6 follows from Proposition 9 below.

The next lemma is used in the proof of Proposition 9. It generalizes Jensen's inequality by showing that if two parameters have the same weighted KL divergence given ψ, and they do not assign the same probability to all events that ψ gives positive probability, then their strict convex combination has a strictly lower weighted KL divergence given ψ.

Lemma 7. *Let* $X \in \mathcal{B}(\mathbb{R}^m)$*, and let* $\Phi, \Phi^1, \Phi^2 \in \Delta(X)$ *be Borel probability measures with densities* $\phi, \phi^1, \phi^2 \in \Delta(X)$ *such that* $-\int_X \log \phi^1(x) d\Phi(x) = -\int_X \log \phi^2(x) d\Phi(x)$ *and* ϕ^1 *is not* Φ*-almost surely equal to* ϕ^2*. For every* $v \in (0, 1)$

$$-\int_X \log[v\phi^1(x) + (1-v)\phi^2(x)] d\Phi(x) > -\int_X \log \phi^1(x) d\Phi(x).$$

Proof. Since ϕ^1 is not Φ-almost surely equal to ϕ^2 there exists $B \in \mathcal{B}(X)$ with $\Phi(B) > 0$, and $K \in \mathbb{R}^+$ such that $\phi^1(x) > \phi^2(x) + K$ for all $x \in B$. Moreover, since

$$-\int_{x\in X} \log \phi^1(x) d\Phi(x) = -\int_{x\in X} \log \phi^2(x) d\Phi(x)$$

the set B can be chosen such that

$$\underline{K} \leqslant \phi^1(x) \leqslant \bar{K} \text{ and } \underline{K} \leqslant \phi^2(x) \leqslant \bar{K} \text{ for all } x \in B$$

for some $\underline{K}, \bar{K}$. Let

$$\rho = \min_{z\in[\underline{K},\bar{K}],z'\in[z+K,\bar{K}]} \log(vz + (1-v)z') - v\log(z) - (1-v)\log(z') > 0 \tag{13}$$

where the strict inequality follows from Jensen's inequality, the strict concavity of log, and the compactness of the set over which the expression is minimized.

Notice that the formula for relative entropy can be expanded as

$$\begin{aligned}
&-\int_{Y\setminus B} \log\left(v\phi^1(x) + (1-v)\phi^2(x)\right) d\Phi(x) - \int_B \log\left(v\phi^1(x) + (1-v)\phi^2(x)\right) d\Phi(x) \\
&\leqslant -\int_{Y\setminus B} \log\left(v\phi^1(x) + (1-v)\phi^2(x)\right) d\Phi(x) - \int_B \left(v\log\phi^1(x) + (1-v)\log\phi^2(x) + \rho\right) d\Phi(x) \\
&\leqslant -v\int_Y \log\phi^1(x)d\Phi(x) + (1-v)\int_Y \log\phi^2(x)d\Phi(x) - \rho\Phi(B) \\
&= \int_{x\in X} \log\phi^1(x)d\Phi(x) - \rho\Phi(B)
\end{aligned}$$

as desired. ■

Proposition 9. *If either*
(i) $BR(\Delta(\Theta(p_\Pi))$ *is a singleton, or*
(ii) Θ *is finite and strategies are subjectively different under* p_Π,
then $\lim_{n\to\infty}\psi_n(\Theta, p_\Pi)$ *exists, and is in* $\Delta(BR(\Delta(\Theta(p_\Pi))))$.

Proof. If $\{\hat{\pi}\} = BR(\Delta(\Theta(p_\Pi))$, an argument analogous to the main theorem in Berk (1966) guarantees that almost surely $\lim_{n\to\infty}\mu_n(C) = 1$ for all open sets $C \supseteq \Theta(p_\Pi)$, and the upper-hemicontinuity of the best-reply correspondence implies that $\psi_n \to \hat{\pi}$.

The proof for part (ii) follows from three claims. Claim 1 shows that almost surely the posterior beliefs will assign probability 1 to the KL-minimizing parameters for p_Π. Claim

2 shows that the likelihood ratios between different minimizers is a non-degenerate random walk. We show this by adapting and extending an argument from Fudenberg, Lanzani, and Strack (2021) to allow for infinitely many outcomes and a related but different random walk. Claim 3 shows that beliefs that induce ties have Lebesgue measure zero in the space of likelihood ratios. The proposition then follows from the central limit theorem.

Define $Q_{p_\Pi} \in \Delta(S \times A \times Y)$ by $Q_{p_\Pi}(s, a, B) = \sigma(s)p_\Pi\{\pi : \pi(s) = a\}Q^*(B|s, a)$. Partition the elements of Θ in equivalence classes $\{\tilde{\theta}^l\}_{l=1}^C$ such that

$$\sum_{\pi\in\Pi} q_{\theta^1}(\cdot|s, \pi(s))p_\Pi(\pi) = \sum_{\pi\in\Pi} q_{\theta^2}(\cdot|s, \pi(s))p_\Pi(\pi) \quad \forall s \in S, \forall\theta^1, \theta^2 \in \tilde{\theta}^l$$

$\sum_{\pi\in\Pi} Q^*(\cdot|s, \pi(s))p_\Pi(\pi)$-almost surely, and for every $i \in \{1, ..., C\}$ let θ^i be an arbitrary element of $\tilde{\theta}^i$. Let $\tilde{\theta}^1, ..., \tilde{\theta}^K$ be the equivalence classes that contain the elements of $\Theta(p_\Pi)$, and let $\tilde{\theta}^1$ contain at least one element of $\text{argmax}_{\theta\in\Theta(p_\Pi)} \int_{S\times A\times Y} q_\theta(y|s, a)dQ_{p_\Pi}(s, a, y)$. For every $m \in \mathbb{N}$, let

$$\mu_m(\tilde{\theta}^l) = \mu_\Theta(\tilde{\theta}^l)\frac{\prod_{j=1}^m q_{\theta^l}(y_j|s_j, a_j)}{\sum_{i\in\{1,...,C\}} \mu_\Theta(\tilde{\theta}^i) \prod_{j=1}^m q_{\theta^i}(y_j|s_j, a_j)} \qquad \forall l \in \{1, ..., C\},$$

which is well defined Q_{p_Π}-almost surely. With this, for all $l \in \{1, ..., C\}$ define

$$Z_m^l = \log \frac{\mu_m(\tilde{\theta}^l)}{\mu_m(\tilde{\theta}^1)} \text{ and } L_m^l = \log \frac{q_{\theta^l}(y_m|s_m, a_m)}{q_{\theta^1}(y_m|s_m, a_m)}, \text{ so } Z_m^l = Z_0^l + \sum_{i=1}^m L_i^l.$$

Claim 1. *The probability assigned to the KL-minimizing parameters goes to* 1 Q_{p_Π}*-almost surely, i.e.* $\liminf_{m\to\infty} \mu_m (\Theta(p_\Pi)) = 1$.

The proof of this claim combines the SLLN with the Monotone Convergence Theorem to show the likelihood ratio between two parameters converges even when the ratio between their densities may be unbounded.[36]

Proof. If $\Theta = \Theta(p_\Pi)$ the result is immediate. Suppose $K < C$. For $l > K$, $\mathbb{E}\left[L_m^l | \left(Z_i^l\right)_{i=1}^{m-1}\right]$ is equal to

$$\sum_{s\in S} \sigma(s) \sum_{\pi\in\Pi} p_\Pi(\pi) \left[H\left(Q^*(\cdot|s, \pi(s)), Q_{\theta^1}(\cdot|s, \pi(s))\right) - H\left(Q^*\left(\cdot|s, \pi(s)\right), Q_{\theta^l}\left(\cdot|s, \pi(s)\right)\right)\right] < 0.$$

Since $\Theta \in \mathcal{K}$, $\mathbb{E}\left[(L_m^l)^+ | \left(Z_i^l\right)_{i=1}^{m-1}\right] < \infty$ and so by the Strong Law of Large Numbers and the

[36]Unlike the related Lemma 2 of Esponda and Pouzo (2016), this result allows Y to be infinite.

Monotone Convergence Theorem, it follows that $\lim_{m\to\infty} e^{Z_m^l} = 0$ a.s. Therefore,

$$\limsup_{m\to\infty} \log \frac{\mu_m(\Theta\backslash\Theta(p_\Pi))}{\mu_m(\Theta(p_\Pi))} \leqslant \limsup_{m\to\infty} \log \frac{\mu_m(\Theta\backslash\Theta(p_\Pi))}{\mu_m\left(\tilde{\theta}^1\right)} = \limsup_{m\to\infty} \log \sum_{l=K+1}^{\Theta} \exp Z_m^l \overset{a.s.}{=} -\infty,$$

proving the claim. ∎

Claim 2. *The process $\left(Z^l\right)_{l=2}^K$ is a multi-dimensional random walk in $\mathbb{R}^{K-1}$, and the covariance matrix of its increments is positive definite.*

Proof. For every $l \in \{2, ..., K\}$, $\mathbb{E}\left[L_m^l | \left(Z_i^l\right)_{i=1}^{m-1}\right] = H_{p_\Pi}(Q^*, Q_{\theta^1}) - H_{p_\Pi}(Q^*, Q_{\theta^l})=0$, so $\left(Z^l\right)_{l=2}^K$ is a multi-dimensional random walk. Because $Q^*(\cdot|s,a)$ is absolutely continuous with respect to $Q_{\theta^1}(\cdot|s,a)$ for all $s \in S$ and $a \in \operatorname{supp} p_\Pi(s)$, the increments L_t have covariance matrix Σ given by

$$\Sigma_{ij} = cov(L^i, L^j) = \mathbb{E}\left[L^i L^j\right] = \int_{S\times A\times Y} \log\left(\frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)}\right) \log\left(\frac{q_{\theta^j}(y|s,a)}{q_{\theta^1}(y|s,a)}\right) dQ_{p_\Pi}(s,a,y)\,.$$

To show this covariance matrix is positive definite, we will show that $v^T\Sigma v > 0$ for all $v \in \mathbb{R}_{++}^{K-1}$ with $||v||_1 = 1$. This is sufficient because these vectors include the canonical orthogonal basis of $\mathbb{R}^{K-1}$. The claim trivially holds if $K = 2$. Therefore, suppose $K > 2$, observe first that $v^T\Sigma v$ is non-negative:

$$\begin{aligned}
v^T\Sigma v &= \sum_{i=2}^K\sum_{j=2}^K v_i\Sigma_{ij}v_j = \sum_{i=2}^K\sum_{j=2}^K v_iv_j \int_{S\times A\times Y} \log\left(\frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)}\right) \log\left(\frac{q_{\theta^j}(y|s,a)}{q_{\theta^1}(y|s,a)}\right) dQ_{p_\Pi}(s,a,y)\\
&= \int_{S\times A\times Y} \sum_{i=2}^K\sum_{j=2}^K v_i \log\left(\frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)}\right) v_j \log\left(\frac{q_{\theta^j}(y|s,a)}{q_{\theta^1}(y|s,a)}\right) dQ_{p_\Pi}(s,a,y)\\
&= \int_{S\times A\times Y} \left(\sum_{i=2}^K v_i \log\left(\frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)}\right)\right)^2 dQ_{p_\Pi}(s,a,y) \geqslant 0.
\end{aligned}$$

Since the last expression is the integral of a weakly positive function, it equals zero if and only if the integrand is Q_{p_Π}-almost surely equal to 0. Moreover, we have:

$$0 = \sum_{i=2}^K v_i \log\left(\frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)}\right) \Rightarrow \log q_{\theta^1}(y|s,a) = \sum_{i=2}^K v_i \log q_{\theta^i}(y|s,a).$$

By Jensen's inequality this implies that $\log q_{\theta^1}(y|s,a) \leqslant \log \sum_{i=2}^K v_i q_{\theta^i}(y|s,a)$, Q_{p_Π}-almost surely, so $q_{\theta^1}(y|s,a) \leqslant \sum_{i=2}^K v_i q_{\theta^i}(y|s,a)$. And as θ^1 maximizes $\int_{S\times A\times Y} q_\theta(y|s,a) dQ_{p_\Pi}(s,a,y) = 1$ on $\Theta(p_\Pi)$ this implies that Q_{p_Π}-almost surely $q_{\theta^1}(y|s,a) = \sum_{i=2}^K v_i q_{\theta^i}(y|s,a)$. By Lemma 7, this contradicts $\theta^2 \in \Theta(p_\Pi)$. Thus $v^T \Sigma v > 0$, so Σ is positive definite, proving Claim 2. ■

Claim 3. *The set of* $\nu \in \Delta\left(\left\{\tilde{\theta}^1, ..., \tilde{\theta}^K\right\}\right)$ *such that*

$$\sum_{i=1}^{K} \nu(\tilde{\theta}^i) \sum_{\theta_j \in \tilde{\theta}^i} \frac{\mu_\theta(\theta_j)}{\mu_\theta(\tilde{\theta}^i)} U_{\theta_j}(\pi) = \sum_{i=1}^{K} \nu(\tilde{\theta}^i) \sum_{\theta_j \in \tilde{\theta}^i} \frac{\mu_\theta(\theta_j)}{\mu_\theta(\tilde{\theta}^i)} U_{\theta_j}(\pi'). \tag{14}$$

for some $\pi \neq \pi'$ *has Lebesgue measure* 0 *in* $\mathbb{R}^K$.

Proof. Fix $\pi \neq \pi'$. Equation (14) is a linear equation in the K unknowns $\nu(\tilde{\theta}^i)$, so its solutions are a vector subspace of $\mathbb{R}^K$. Since strategies are subjectively different under p_Π there exists $\tilde{\theta}_l \in \Theta(p_\Pi)$ such that $U_{\mu\left(\cdot|\tilde{\theta}_l\right)}(\pi) \neq U_{\mu\left(\cdot|\tilde{\theta}_l\right)}(\pi')$, so the set of beliefs under which $U_\mu(\pi) = U_\mu(\pi')$ has dimension at most $K-1$, and hence Lebesgue measure 0. Since the set of actions is finite and π, π' are chosen arbitrary, the set of beliefs $\nu \in \Delta\left(\left\{\tilde{\theta}^1, ..., \tilde{\theta}^K\right\}\right) \subseteq \mathbb{R}^K$ such that $U(\nu, \pi) = U(\nu, \pi')$ for some $\pi \neq \pi'$ has Lebesgue measure 0 as well. ■

Note that because $\left(Z^l\right)_{l=2}^K$ is a martingale with positive definite covariance matrix of the increments, the central limit theorem implies that $\left(Z_m^l/\sqrt{m}\right)_{l=2}^K$ converges in distribution to a $K-1$ dimensional normal distribution with mean $\vec{0}$ and covariance matrix Σ. Since

$$\mu_m(\tilde{\theta}_l) = \frac{\exp Z_m^l}{\sum_{i=2}^K \exp Z_m^i + 1} \qquad \forall l \in \{2, ..., K\},$$

the distribution on the indifference classes induced by μ_m converges to some $\nu \in \Delta\left(\Delta\left(\Theta\left(p_\Pi\right)\right)\right)$, as does the distribution of beliefs over the indifference classes in the overall population. Claim 2 shows that Σ is positive definite, so by Claim 3 beliefs that induce ties between the strategies' payoffs have 0 limit probability. Therefore, the induced distribution of strategies converges to an element of $\Delta\left(BR\left(\Delta\left(\Theta\left(p_\Pi\right)\right)\right)\right)$. This concludes the proof of part (ii). ■

Proof of Proposition 7. We endow the set of strategies with the L_1 norm. Let $\varepsilon' \in (0,1)$ be such that for all $\varepsilon < \varepsilon'$, $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is a singleton, and $V(\mathcal{M}_{\Theta,\psi}(\varepsilon)) > U^*(\psi)$. We now prove that (δ_Θ, ψ) does not resist mutation to Θ' if Θ' is an ε expansion of Θ for $\varepsilon < \varepsilon'$.

Assumption 2(iii) implies that U^* is continuous in ψ.[37] So there is an $\varepsilon^* \in (0, \varepsilon')$ and $\gamma > 0$ such that for all $\psi' \in B_{\varepsilon^*}(\psi)$, $\tilde{\psi} \in BR(\Delta(B_{\varepsilon^*}(\mathcal{M}_{\Theta,\psi}(\varepsilon)))$,

$$U^*\left(\psi'\right) - U^*(\tilde{\psi}) < -\gamma. \tag{15}$$

[37]See e.g. Lemma 5.24 of Aliprantis and Border (2013).

By the upper hemicontinuity (see Lemma 2) of $\Theta'(\cdot)$ there is $\bar{\varepsilon}$ such that $||\psi' - \psi|| < \bar{\varepsilon}$ implies $\Theta'(\psi') \subseteq B_{\varepsilon*}(\mathcal{M}_{\Theta,\psi}(\varepsilon))$. Thus if there were a t such that $||p^{\tau}_{\Pi} - \psi|| < \min\{\varepsilon^*, \bar{\varepsilon}\}$ for all $\tau \geqslant t$, it would follow from equation (15) that $U^*(p^{\tau+1}(\cdot|\Theta)) + \gamma/p^t_{\varepsilon}(\Theta) < U^*(p^{\tau+1}(\cdot|\Theta')))$. By Lemma 4 this concludes the proof. ∎

Proof of Proposition 8. Suppose that $\hat{a}$ is an attractor with associated unique KL minimizer $\hat{\theta}$. Let A_m be given by $\{\hat{a} + c/m\}_{c\in\mathbb{Z}\backslash\{0\}} \cap A$, and let $\hat{\varepsilon}$ be such that for all $\varepsilon < \hat{\varepsilon}$, $V(\mathcal{M}_{\Theta,\hat{a}}(\varepsilon)) > V(\hat{\theta})$. Let $BR_n(\theta) = \text{argmax}_{a\in A_n} U_\theta(a)$. Since $\hat{a}$ is in the interior of A, A_m is nonempty for sufficiently large m. By Theorem 1 in Esponda and Pouzo (2016), for every $m \in \mathbb{N}$, the environment with action set A_m admits at least one Berk-Nash equilibrium, so for all $m \in \mathbb{N}$ there is an equilibrium (Θ, ψ_m) with justifying belief $\mu_m \in \Delta(\Theta(\psi_m))$. Since A is compact, by Theorem 15.11 in Aliprantis and Border (2013) $\Delta(A)$ is also compact, and therefore (ψ_m, μ_m) has an accumulation point $\left(\hat{\psi}, \hat{\mu}\right)$.

We claim that this accumulation point must be $\left(\delta_{\hat{a}}, \delta_{\Theta(\hat{a})}\right)$. To see this, recall that there is a unique Berk-Nash equilibrium in the environment with a continuum of actions. Then note that for all $a \in A$, there is a sequence $(a_n)_{n\in\mathbb{N}}$ such that $a_n \to a$, $a_n \in A_n$, and since $U_{(\cdot)}(\cdot)$ is jointly continuous in beliefs and actions $U_{\mu_n}(\psi_n) \geqslant U_{\mu_n}(a_n)$ for all $n \in \mathbb{N}$ implies $U_{\hat{\mu}}\left(\hat{\psi}\right) \geqslant U_{\hat{\mu}}(a)$. Moreover, by the upper hemicontinuity of $\Theta(\cdot)$ (see Lemma 2), $\mu_n \in \Delta(\Theta(\psi_n))$ for all $n \in \mathbb{N}$ implies that $\hat{\mu} \in \Delta\left(\Theta(\hat{\psi})\right)$. That is, $\left(\hat{\psi}, \hat{\mu}\right)$ must be the unique Berk-Nash equilibrium of the environment with a continuum of actions.

Next, we will show that this subsequence $(A_n, \psi_n)_{n\in\mathbb{N}}$ satisfies the requirements in the statement of the proposition. Since A is compact, $||A_n - A|| \to 0$. Let $\hat{\varepsilon}$ be as the ε in the definition of an attractor for $(\hat{a}, \Theta)$. Since $\mu_n \to \delta_{\Theta(\hat{a})}$, for every $\tilde{\varepsilon}$ there exists $N > 0$ such that for all $n > N$, $\psi_n(B_{\hat{\varepsilon}}(\hat{a})) > 1 - \tilde{\varepsilon}$.

Fix an $\varepsilon < \hat{\varepsilon}$. Because $\Theta_\varepsilon(\cdot)$ is upper hemicontinuous (see Lemma 2), $\Theta_\varepsilon(\psi_n)$ converges to $\mathcal{M}_{\Theta,\psi}(\varepsilon) = \Theta_\varepsilon(\hat{a})$. Thus there exists $N'' > N'$ such that for all $n > N''$, $U^*(BR_n(\Theta_\varepsilon(\psi_n))) \subseteq (U^*([\hat{a}) + V((\mathcal{M}_{\Theta,\psi}(\varepsilon)))]/2, \infty)$. But since $U^*(\psi_n) \to U^*(\hat{a})$, there exists $N''' > N''$ such that for all $n > N''$, $U^*(BR_n(\Theta_\varepsilon(\psi_n))) > U(\psi_n)$, so that ψ_n does not resist an ε expansion of Θ. ∎

A.1 Prior-independent limit aggregate behavior

Here we show that the limit aggregate behavior identified in Proposition 6 does not depend on the prior of the agents, and that all the best replies to a KL minimizing parameter are played by a positive fraction of agents.

Proposition 10. *If the assumptions of Proposition 6 are satisfied then* $\lim_{n\to\infty}\psi_n(\Theta, p_\Pi)$ *is independent of the prior, and if* $\{\pi\} = BR(\delta_\theta)$ *for some* $\theta \in \Theta(p_\Pi)$*, then* $\lim_{n\to\infty}\psi_n(\Theta, p_\Pi)(\pi) > 0$.

Proof. In this proof, we continue to use the notation introduced in the proof of Proposition 6. That the limit beliefs do not depend on the prior follows from Proposition 6, which shows that the beliefs over equivalence classes converges to the limit distribution ν that is independent of the prior. Suppose $\{\pi\} = BR(\delta_\theta)$ for some $\theta \in \Theta(p_\Pi)$. Since θ^1 was chosen arbitrarily, suppose without loss of generality that $\theta = \theta^1$. Since π is the unique best reply to θ^1, and $Z_m^l \overset{a.s.}{\to} -\infty$ for all $l \in \{K+1, ..., C\}$ by Claim 1, there exists $c < 0$ such that if $(Z_m^l)_{l=2}^K$ is coordinate by coordinate less than c, the best reply to the corresponding belief is to play π. Consider the events E_m that $(Z_m^l)_{l=2}^K$ is coordinate-wise less than c: $E_m = \left\{Z_m^l \leqslant c, \forall l \in \{2, ..., K\}\right\}$. As $Z_m/\sqrt{m}$ converges to a normal random variable we have that

$$\lim_{m\to\infty} \mathbb{P}\left[E_m\right] = \lim_{m\to\infty} \mathbb{P}\left[\frac{Z_m}{\sqrt{m}} \leqslant \frac{c}{\sqrt{m}}\right] = \mathbb{P}\left[\tilde{Z} \leqslant \frac{c}{\sqrt{m}}\right],$$

where $\tilde{Z}$ is a random variable that is Normally distributed with mean $\vec{0}$ and covariance matrix Σ. As Σ is positive definite, this distribution admits a strictly positive density and hence $\mathbb{P}[\tilde{Z} \leqslant c/\sqrt{m}] > 0$. ∎

A.2 Examples

A.2.1 Example 1

Esponda and Pouzo (2016) shows that

$$\sum_{a\in\{10,2\}} \psi(a)H\left(Q^*(\cdot|a), Q_\theta(\cdot|s,a)\right) = \psi(2)(34 - i + 2\beta)^2 + \psi(10)(2 - i + 10\beta)^2. \tag{16}$$

a) When $\psi(2) = 1$, the parameter that minimizes equation (16) is (3/2,32), and since $BR(\delta_{(3/2,32)}) = \{10\}$, $(\Theta, 2)$ is not a Berk-Nash equilibrium. When $\psi(10) = 1$, the parameter that minimizes equation (16) is $(5/2, 28)$. Since $BR(\delta_{(5/2,28)}) = \{2\}$, $(\Theta, 10)$ is not a Berk-Nash equilibrium. For every totally mixed ψ, the Hessian of equation (16) as a function of β and i, $[200 - 192\psi(2), 16\psi(2) - 20; 16\psi(2) - 20, 2]$, is positive definite for every $(\beta, i) \in \Theta$ so there is a unique KL minimizer. Moreover, plugging $i = 12\beta$ in to equation (16) shows that the derivative in β of the resulting expression is strictly negative for every ψ. Therefore, the unique parameter on the line $i = 12\beta$ where the two actions are indifferent that minimizes equation (16) for some ψ is $\hat{\theta} = (5/2, 30)$ with $\psi = (1/4, 3/4)$, and so the latter is the unique Berk-Nash equilibrium.

The first order condition for the KL-minimizing intercept after an ε expansion of the model is:

$$-2\frac{1}{4}(34 - i + 2 * (2.5 + \varepsilon)) - 2(1 - \frac{1}{4})(2 - i + 10 * (2.5 + \varepsilon)) = 0$$

so that by equation (3), $\mathcal{M}_{\Theta,\psi}(\varepsilon) = (2.5 + \varepsilon, 30 + 8\varepsilon)$.

b) When $\psi(2) = 1$, the parameter that minimizes equation (16) is (3,40), and since $BR(\delta_{(3,40)}) = \{10\}$, $(\Theta, 2)$ is not a Berk-Nash equilibrium. When $\psi(10) = 1$, all the (β, i) with $i = 10\beta + 2$, $\beta \in (3, 10/3)$ minimize equation (16). Since $BR(\delta_{(\beta,i)}) = \{2\}$ for all such (β, i), $(\Theta, 10)$ is not a Berk-Nash equilibrium. The first order conditions for $(10/3, 40)$ to be the KL minimizer are

$$-2\psi(2)(34 - 40 + 2 * \frac{10}{3}) - 2(1 - \psi(2))(2 - 40 + 10 * \frac{10}{3}) \leqslant 0 \tag{17}$$

$$4\psi(2)(34 - 40 + 2 * \frac{10}{3}) + 20(1 - \psi(2))(2 - 40 + 10 * \frac{10}{3}) \leqslant 0. \tag{18}$$

The first inequality gives $\psi(2) \geqslant 7/8$, while the second gives $\psi(2) \leqslant 35/36$.

Each parameter $\tilde{v}$ on the unit circle $\mathcal{S}$ can be written as $\tilde{v} = (\sqrt{(1 - v^2)}, v)$ for some $v \in [0, 1]$. With this,

$$\begin{aligned} D_\psi\left(\hat{\theta}, (\sqrt{(1 - v^2)}, v)\right) = &-\sqrt{(1 - v^2)}[4\psi(2)(34 - 40 + 2 * \frac{10}{3}) + 20(1 - \psi(2))(2 - 40 + 10 * \frac{10}{3})] \\ &+ v * 2\psi(2)(34 - 40 + 2 * \frac{10}{3}) + 2(1 - \psi(2))(2 - 40 + 10 * \frac{10}{3}). \end{aligned}$$

This expression is maximized at a $\tilde{v}$ with $\sqrt{(1 - v^2)}/v > 1/12$ if and only if $\psi(2) > 427/438 \approx 0.97$.

A.2.2 Example 2

Each parameter θ generates distribution $q_\theta(z, x|a) = \phi_1(z - a)\phi_\theta(x|z)$ on (z, x), where ϕ_1 is the pdf of a standard normal distribution and $\phi_\theta(\cdot|z)$ is a normal density with mean $\theta_1 z + \theta_2 z^2$ and variance $z^2 + z^4$. Since ϕ_θ is a normal density, for the restricted linear model where $\theta_2 = 0$ we have

$$H\left(Q^*(\cdot|a), Q_\theta(\cdot|a)\right) \propto -\frac{1}{2}\int \left(\frac{\tau(a+\omega)}{(a+\omega)} - \theta_1\right)^2 d\phi(\omega).$$

An agent who drops the linearity assumption and shifts to the subjective model $\Theta^2 = \mathbb{R} \times \mathbb{R}_+$ finds that the KL-minimizing parameters solve:

$$\underset{(\theta_1,\theta_2)\in\mathbb{R}\times\mathbb{R}_+}{\operatorname{argmin}} \ \mathbb{E}\left[(\tau(5+\omega) - \theta_1(5+\omega) - \theta_2(5+\omega)^2)^2\right].$$

Numerical calculations in Mathematica (available here) then show that the subjectively optimal actions are 5 for the restricted linear model and 3 for the agent with model Θ^2.

A.2.3 Example 3

We add a constant signal $s = 0$ to the deterministic version of the example so we can state some conclusions that apply to both deterministic and stochastic version at the same time.

The buyer's payoff is

$$u(s, a, y) = \sum_{\omega=1}^{a} p_\omega(\omega + 3.1 + s - a)1_{a\geqslant\omega}.$$

To see that bidding 3 is objectively optimal after every signal, note that

$$\begin{aligned}\sum_{y\in Y} u(-1,3,y)p^*(y) &= \frac{1}{3}(3+3.1-1) + \frac{1}{2}(2+3.1-1) + \frac{1}{6}(1+3.1-1) - 3 \\ &> \frac{1}{2}(2+3.1-1-2) + \frac{1}{6}(1+3.1-1-2) \\ &= \sum_{y\in Y} u(-1,2,y)p^*(y) > \frac{1}{6}(1+3.1-1-1) = \sum_{y\in Y} u(-1,1,y)p^*(y).\end{aligned}$$

Since bidding 3 is optimal when $s = -1$ and the utility function is strictly supermodular in a and s, it is also optimal to bid 3 when s is 0 or 1.

Online Appendix B of Esponda and Pouzo (2016) shows that in a Berk-Nash equilibrium of this example, beliefs have correct marginals over prices asked and a marginal distribution over valuations equal to the one observed in equilibrium. Therefore, when s is identically zero, $\Theta(2) = \tilde{\theta} := ((\hat{p}_1, \hat{p}_2, \hat{p}_3), (\hat{F}(1|1), \hat{F}(2|1), \hat{F}(1|2), \hat{F}(2|2), \hat{F}(1|3), \hat{F}(2|3))) = ((1/6, 1/2, 1/3), (1/4, 1, 1/4, 1, 1/4, 1))$, where $\hat{p}_i$ is the conjectured probability that the seller asks price i and $\hat{F}(i|j)$ is the conjectured probability that the value is less than or equal to i given that the seller asked price j. Moreover, under $\tilde{\theta}$ bidding 2 is optimal:

$$U_{\tilde{\theta}}(1) = \frac{1}{6}(\frac{1}{4}(4.1{-}1){+}\frac{3}{4}(5.1{-}1)) < (\frac{1}{4}(4.1{-}3){+}\frac{3}{4}(5.1{-}3)) = U_{\tilde{\theta}}(3) < \frac{2}{3}(\frac{1}{4}(4.1{-}2){+}\frac{3}{4}(5.1{-}2)) = U_{\tilde{\theta}}(2).$$

In this equilibrium, for every $\omega \in \{1,2,3\}$, relaxing the hypotheses $F(\omega|3) \geqslant F(\omega|3)$ or $F(\omega|3) \leqslant F(\omega|3)$ is not explanation improving, as in equilibrium the agent never observes the value after an ask price equal to 3.

In the stochastic case, there are two signals, $s \in \{-1, 1\}$. With $\pi(-1) = 2$ and $\pi(1) = 3$, we have $\bar{\theta} :\ = \Theta(\pi) = ((1/6, 1/2, 1/3), (1/5, 4/5, 1/5, 4/5, 1/5, 4/5))$, and under $\bar{\theta}$ bidding 2 after signal $s = -1$ is optimal:

$$\begin{aligned}\sum_{y\in Y} u(-1,1,y)p_{\bar{\theta}}(y) &= \frac{1}{6}(\frac{1}{5}(3.1-1) + \frac{3}{5}(4.1-1) + \frac{1}{5}(5.1-1)) \\ &< (\frac{1}{5}(3.1-3) + \frac{3}{5}(4.1-3) + \frac{1}{5}(5.1-3)) = \sum_{y\in Y} u(-1,3,y)p_{\bar{\theta}}(y) \\ &< \frac{2}{3}(\frac{1}{5}(3.1-2) + \frac{3}{5}(4.1-2) + \frac{1}{5}(5.1-2)) = \sum_{y\in Y} u(-1,2,y)p_{\bar{\theta}}(y).\end{aligned}$$

Moreover, under $\bar{\theta}$ bidding 3 after signal $s = 1$ is optimal:

$$\begin{aligned}\sum_{y\in Y} u(1,1,y)p_{\bar{\theta}}(y) &= \frac{1}{6}(\frac{1}{5}(5.1-1) + \frac{3}{5}(6.1-1) + \frac{1}{5}(7.1-1)) \\ &< \frac{2}{3}(\frac{1}{5}(5.1-2) + \frac{3}{5}(6.1-2) + \frac{1}{5}(7.1-2)) = \sum_{y\in Y} u(1,2,y)p_{\bar{\theta}}(y) \\ &< (\frac{1}{5}(5.1-3) + \frac{3}{5}(6.1-3) + \frac{1}{5}(7.1-3)) = \sum_{y\in Y} u(1,3,y)p_{\bar{\theta}}(y).\end{aligned}$$

Finally, the minimizing parameter $\hat{\theta}$ after the one hypothesis relaxation to $F(2|3) \leqslant F(2|2) = F(2|1)$ is obtained as the unique element of

$$\operatorname*{argmin}_{F(1|1),F(2|1),F(2|3)} -[\log(F(1|1))/6 + \log(F(2|1) - F(1|1))/2] - \frac{1}{2}[\log(1 - F(2|3))/3].$$

A.2.4 An example of a cycle

Example 5. Let $A = \{a, b, c\}$, $Y = \{0, 1\}$, $u(a, y) = y$ and let $\theta = (\theta_a, \theta_b, \theta_c) = [0, 1]^3$ correspond to the probability of success ($y = 1$) under the three actions. The objective parameter is $(0.5, 0.01, 0.02)$ so that a is the optimal action. Suppose that $\Theta_1 = \{(0.5, 0.9, 0.02), (0.5, 0.3, 0.1)\}$, $\Theta_2 = \{(0.5, 0.01, 0.9), (0.5, 0.1, 0.3)\}$. Consider the two states

$$\hat{p} = 0.1\delta_{\Theta_1 \times a} + 0.9\delta_{\Theta_2 \times c} \text{ and } \bar{p} = 0.9\delta_{\Theta_1 \times b} + 0.1\delta_{\Theta_2 \times a}.$$

Let T be an arbitrary payoff monotone dynamic such that $T(\hat{p})(\Theta_1) = 0.9$ and $T(\bar{p})(\Theta_1) = 0.1$. Notice that payoff monotonicity is satisfied, as under $\hat{p}$ the performance of the agents with subjective model Θ_1 is higher (they play a) than that of those with subjective model Θ_2 (they play c). Moreover under $\bar{p}$ the performance of the agents with subjective model Θ_1 is lower (they play b) than that of those with subjective model Θ_2 (they play a). The unique solution with $p^0 = \hat{p}$ has $p^t = \hat{p}$ in all even periods and $p^t = \bar{p}$ in all odd periods and the system cycles forever. Moreover, the average payoff is $0.5 * 0.1 + 0.01 * 0.9$ in the odd periods and $0.5 * 0.1 + 0.02 * 0.9$ in the even periods, so that the average payoff sequence is not monotone. ▲

A.2.5 An innovation-vulnerable equilibrium

Example 6. Suppose that $A = \{a, b, c\}$ and that the outcomes have three components that are either 1 or 0, i.e., $Y = \{0,1\} \times \{0,1\} \times \{0,1\}$. The utility of a and b depends only on the first component; a is better if the first component is likely to be 1, b if it is likely to be 0:

$$u(a, (1, y_2, y_3)) = 1 = u(b, (0, y_2, y_3))$$
$$u(a, (0, y_2, y_3)) = 0 = u(b, (1, y_2, y_3)).$$

The utility to c depends only on the third outcome component:

$$u(c, (y_1, y_2, 1)) = 1;$$
$$u(c, (y_1, y_2, 0)) = 0.$$

The parameter space has two dimensions $\mathcal{H} = [0,1] \times [0,1]$, where $\theta = (\theta_1, \theta_2) \in \mathcal{H}$, θ_1 is both the probability that the first component is equal to 1 (regardless of the action) and the probability that the second component is equal to 1 (regardless of the action), and θ_2 is the probability that the third component is equal to 1 while playing b or c. The agent (correctly) believes that the third component is always equal to 0 if they play a, and they believe that the outcomes are independent. Formally

$$q_\theta(y|a) = \begin{cases} 0 & y_3 = 1 \\ (1-\theta_1)^2 & y_1 = y_2 = y_3 = 0 \\ \theta_1^2 & y_1 = y_2 = 1, y_3 = 0 \\ \theta_1(1-\theta_1) & y_1 \neq y_2, y_3 = 0 \end{cases}$$

$$q_\theta(y|b) = \begin{cases} (1-\theta_1)^2(1-\theta_2) & y_1 = y_2 = y_3 = 0 \\ \theta_1^2(1-\theta_2) & y_1 = y_2 = 1, y_3 = 0 \\ \theta_1(1-\theta_1)(1-\theta_2) & y_1 \neq y_2, y_3 = 0 \\ \theta_1^2\theta_2 & y_1 = y_2 = y_3 = 1 \\ (1-\theta_1)^2\theta_2 & y_1 = y_2 = 0, y_3 = 1 \\ \theta_1(1-\theta_1)\theta_2 & y_1 \neq y_2, y_3 = 1 \end{cases}$$

$$q_\theta(y|c) = \theta_1^{y_1+y_2}(1-\theta_1)^{2-y_1-y_2}\theta_2^{y_3}(1-\theta_2)^{y_3}.$$

In reality, the probability of having the first and second component equal to 1 are not equal, the former is equal to 2/3 and the latter is equal to 1/4 under every action. Moreover, the probability of $y_3 = 1$ given b or c is equal to 3/4.

The initial subjective model is $\Theta = \{1/2\} \times [0,1]$, and $p = (\delta_\Theta, a)$ is a steady state: every parameter induces the same weighted KL divergence and a is a best reply to any subjective model in which $\theta_1 = 1/2$ and $\theta_2 \leqslant 1/2$. The equilibrium is not quasi-strict: for every belief supported on Θ, a is a best reply if and only if b is.

A mutation to Θ_ε with $\varepsilon < 1/2$ induces b as the unique best reply, since the mutated agents decrease θ_1 to better match the observed frequency of the second component, which makes b strictly preferable to a. And even if b performs less well than a, behavior does not converge back to the evidence generated by a small fraction of mutated agents playing b allows all the agents to learn that c, the unused best reply to $\Delta(\Theta(a))$, is better than either alternatives. Thus the equilibrium is innovation-vulnerable. ▲

References

Aliprantis, Charambolos and Kim Border (2013). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Berlin. Springer-Verlag.

Arrow, KJ and JR Green (1973). "Notes on Expectations Equilibria in Bayesian Settings." Working Paper No. 33, Stanford University.

Ba, Cuimin (2022). "Robust Model Misspecification and Paradigm Shifts." *arXiv:2106.12727*.

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

Binmore, Ken and Larry Samuelson (1997). "Muddling Through: Noisy Equilibrium Selection." *Journal of Economic Theory*, 74, 235–265.

Björnerstedt, Jonas and Jörgen W Weibull (1995). *Nash Equilibrium and Evolution by Imitation.* Ed. by Ed. K. Arrow et al. London, MacMillan.

Blume, Lawrence and David Easley (1992). "Evolution and Market Behavior." *Journal of Economic theory*, 58, 9–40.

Bohren, J Aislinn (2016). "Informational Herding with Model Misspecification." *Journal of Economic Theory*, 163, 222–247.

Bohren, J Aislinn and Daniel Hauser (2021). "Learning with Model Misspecification: Characterization and Robustness." *Econometrica*, 89, 3025–3077.

Bowen, Renee, Danil Dmitriev, and Simone Galperti (2022). "Learning from Shared News: When Abundant Information Leads to Belief Polarization."

Cabrales, Antonio and Joel Sobel (1992). "On the Limit Points of Discrete Selection Dynamics." *Journal of Economic Theory*, 57, 407–419.

Cho, In-Koo and Kenneth Kasa (2017). "Gresham's Law of Model Averaging." *American Economic Review*, 107, 3589–3616.

Dekel, Eddie and Suzanne Scotchmer (1992). "On the Evolution of Optimizing Behavior." *Journal of Economic Theory*, 57, 392–406.

Diaconis, Persi and David Freedman (1986). "On the Consistency of Bayes Estimates." *The Annals of Statistics*, 14, 1–26.

Ellison, Glenn and Drew Fudenberg (1993). "Rules of Thumb for Social Learning." *Journal of Political Economy*, 101, 612–643.

Esponda, Ignacio (2008). "Information Feedback in First Price Auctions." *The RAND Journal of Economics*, 39, 491–508.

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

Esponda, Ignacio, Demian Pouzo, and Yuichi Yamamoto (2021). "Asymptotic Behavior of Bayesian Learners with Misspecified Models." *Journal of Economic Theory*, 195, 105260.

Esponda, Ignacio, Emanuel Vespa, and Sevgi Yuksel (2022). "Mental Models and Learning: The Case of Base-Rate Neglect." *mimeo.*

Franks, Michael E, Gordon R Macpherson, and William D Figg (2004). "Thalidomide." *The Lancet*, 363, 1802–1811.

Frick, Mira, Ryota Iijima, and Yuhta Ishii (2020). "Misinterpreting Others and the Fragility of Social Learning." *Econometrica*, 88, 2281–2328.

Frick, Mira, Ryota Iijima, and Yuhta Ishii (2021a). "Belief Convergence under Misspecified Learning: A Martingale Approach." *mimeo.*

— (2021b). "Welfare Comparisons for Biased Learning." *mimeo.*

Fudenberg, Drew and David M Kreps (1994). "Learning in Extensive Form Games, II: Experimentation and Nash Equilibrium." *mimeo.*

Fudenberg, Drew, Giacomo Lanzani, and Philipp Strack (2021). "Limits Points of Endogenous Misspecified Learning." *Econometrica*, 89, 1065–1098.

Fudenberg, Drew, Gleb Romanyuk, and Philipp Strack (2017). "Active Learning with a Misspecified Prior." *Theoretical Economics*, 12, 1155–1189.

Fudenberg, Drew and Jean Tirole (1991). *Game Theory*. MIT press.

Gagnon-Bartsch, Tristan, Matthew Rabin, and Joshua Schwartzstein (2021). "Channeled Attention and Stable Errors." *mimeo.*

Grant, Simon and John Quiggin (2017). "The evolution of awareness." *Journal of Economic Psychology*, 63, 86–92.

Hansen, Lars Peter and Thomas J Sargent (2008). *Robustness*. Princeton university press.

Harsanyi, John C (1973). "Oddness of the Number of Equilibrium Points: a New Proof." *International Journal of Game Theory*, 2, 235–250.

He, Kevin (2022). "Mislearning from Censored Data: The Gambler's Fallacy and Other Correlational Mistakes in Optimal-Stopping Problems." *Theoretical Economics*, 17, 1269–1312.

He, Kevin and Jonathan Libgober (2021). "Evolutionarily Stable (Mis)specifications: Theory and Application." *arXiv:2012.15007.*

Heidhues, Paul, Botond Kőszegi, and Philipp Strack (2018). "Unrealistic Expectations and Misguided Learning." *Econometrica*, 86, 1159–1214.

Hines, William GS (1980). "Three Characterizations of Population Strategy Stability." *Journal of Applied Probability*, 17, 333–340.

Hofbauer, Josef and Karl Sigmund (1988). *Dynamical Systems and the Theory of Evolution.*

Hong, Harrison, Jeremy C Stein, and Jialin Yu (2007). "Simple Forecasts and Paradigm Shifts." *The Journal of Finance*, 62, 1207–1242.

Jehiel, Philippe (2020). "Analogy-Based Expectation Equilibrium and Related Concepts: Theory, Applications, and Beyond." *mimeo.*

Kuhn, Thomas (1957). *The Copernican revolution: Planetary astronomy in the development of Western thought*. Vol. 16. Harvard University Press.

Lanzani, Giacomo (2022). "Dynamic Concern for Misspecification."

Liebman, Jeffrey B and Richard J Zeckhauser (2004). "Schmeduling."

Liese, Friedrich and Igor Vajda (1987). *Convex Statistical Distances*. Vol. 95. Teubner.

Massari, Filippo and Jonathan Newton (2022). "Learning and Equilibrium in Misspecified Models." *Available at SSRN 3473630*.

Molavi, Pooya (2019). "Macroeconomics with Learning and Misspecification: A General Theory and Applications." *mimeo*.

Montiel Olea, José Luis, Pietro Ortoleva, Mallesh Pai, and Andrea Prat (2021). "Competing models." *arXiv:1907.03809*.

Nyarko, Yaw (1991). "Learning in Mis-specified Models and the Possibility of Cycles." *Journal of Economic Theory*, 55, 416–427.

Ortoleva, Pietro (2012). "Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News." *American Economic Review*, 102, 2410–36.

Rees-Jones, Alex and Dmitry Taubinsky (2020). "Measuring "schmeduling"." *The Review of Economic Studies*, 87, 2399–2438.

Sabra, Abdelhamid I (1984). "The Andalusian Revolt Against Ptolemaic Astronomy: Averroes and Al-Bitruji." *Transformation and Tradition in the Sciences*, 133–53.

Samuelson, Larry and Jianbo Zhang (1992). "Evolutionary Stability in Asymmetric Games." *Journal of Economic Theory*, 57, 363–391.

Schlag, Karl H (1998). "Why Imitate, and If So, How?: Exploring a Model of Social Evolution." *Journal of Economic Theory*, 78, 130–156.

Schwartzstein, Joshua and Adi Sunderam (2021). "Shared Models in Networks, Organizations, and Groups." *mimeo*.

Sobel, Joel (1984). "Non-linear Prices and Price-Taking Behavior." *Journal of Economic Behavior & Organization*, 5, 387–396.