Which Misspecifications Persist?

Drew Fudenberg and Giacomo Lanzani

First posted version: October 12, 2020
This version: January 11, 2022

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. Mutations that lead agents to use a model that better fits the equilibrium data can fail to spread if the better-fitting model leads 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, Jörgen 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 that best fit the data generated by the actions and outcomes of the previous generation. 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 baseline Bayesian model predicts that all misspecifications will persist forever. Thus, we suppose that a small fraction of agents may adopt an expanded subjective model. We then ask whether the equilibrium will resist the mutation, in the sense that the original behavior persists. 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. There are two ways this can occur. One way is direct: the better explanation of the data may lead the mutants to adopt a better action, so their share of the population grows. The other channel is indirect: The mutants might obtain the same or lower payoff as agents using the prevailing paradigm.

1“Best fit” here means maximizing the likelihood of the data.
but the information generated by the mutants’ actions can help agents with the old paradigm realize they can increase their payoff with another action.

The effectiveness of these channels depends both on the nature of the equilibrium and that of the mutations. In a uniformly strict Berk-Nash equilibria (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. 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 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 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 payoff. 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

---

2This is similar to the neighborhoods of a subjective model considered in the macroeconomics literature on robust control following Hansen and Sargent (2008).
between the variables. Since in our large population 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 extend the analysis to settings with a continuum of actions, so small changes in paradigm can induce different actions even when the original equilibrium action was the unique best reply. In this setting the key concept for stability is the derivative of the indirect utility function of the agent, and we reconcile this finding with the previous analysis by showing how Berk-Nash equilibria with a continuum of actions are approximated by mixed equilibria on a fine finite action grid.

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 players’ beliefs are correct with the requirement that each player’s belief minimizes the KL divergence of their observations from their subjective model.3

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 modelling 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

3Jehiel (2020) surveys various equilibrium concepts for misspecified agents.
equilibria. Bohren and Hauser (2021) characterize the long-run belief of a sequence of myopic misspecified agents.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. Unlike us, they do not specify an explicit evolutionary dynamic, and inference in their model does not depend on data that was generated before the mutation. Massari and Newton (2020) justifies a generalized Bayes rule as the result of evolutionary competition between different models, and 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

Our work is related to the evolutionary selection of risk preferences (Robson, 1996a,b, Dekel and Scotchmer, 1999, Robson and Samuelson, 2019) or time preferences (Robson and Samuelson, 2007, Netzer, 2009).6 It is also related to models of agents who subject their models to statistical tests, as in Fudenberg and Kreps (1994), Hong, Stein, and Yu (2007), Cho and Kasa (2017), and Ba (2021).

Esponda, Vespa, and Yuksel (2020) 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.

2 The single-agent problem

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

4Fudenberg, Romanyuk, and Strack (2017), Heidhues, Köszegi, and Strack (2018), and Molavi (2019) analyze misspecified learning in specific applications. Bohren (2016), Frick, Iijima, and Ishii (2020), and He (2021) consider misspecified social learning where all agents have the same misspecification. Like us, He (2021) considers a model where agents learn from the data generated by the previous generation.

5Also, Schwartzstein and Sunderam (2021) studies model selection in a setting without actions or payoffs, Montiel Olea, Ortöleva, 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.

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 $Q^*(\cdot|\cdot) \in \Delta(Y)^{S \times A}$.

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 $\Pi^A_S$.

The objective expected utility of strategy $\pi$ 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 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 $\theta$ are associated with a family of probability measures $Q_\theta(\cdot|s, a)$, one for each signal-action pair $(s, a)$. 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 $\text{supp} \mu = \Theta \subseteq \mathcal{H}$ the agent considers possible.

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

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 Lévy–Prokhorov metric.

For simplicity we do not allow individual agents to randomize. Our model captures the effect of randomization by allowing different agents with the same belief to play different actions as long as all of those actions maximize their subjective payoff.

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.
Inference and Kullback-Leibler minimizers  Given two distributions over outcomes \(Q, Q' \in \Delta(Y)\) we define \(H(Q, Q') = -\int_{y \in Y} \log q'(y) dQ(y)\).\(^{10}\) 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 explain 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.\(^{11}\)

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

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

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

\[
\Theta(\psi) = \arg\min_{\theta \in \Theta} H_\psi(Q^*, Q_\theta),
\]

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 \(\psi\) is a probability distribution over \(\Theta(\psi)\). Proposition\([\text{?}]\) shows that this describes the limit as the agent observes a larger and larger number of individual experiences.

Regularity assumptions  We do not require that the agent is correctly specified, i.e. that there is a \(\theta^* \in \Theta\) such that \(Q^* = Q_{\theta^*}\), or even that \(Q^*\) can be approximated by \(Q_\theta\) for some \(\theta \in \Theta\). We allow these cases, but our focus is on the case where the agent is misspecified in the sense their prior rules out the objective outcome distribution for at least some actions. Let \(B_\varepsilon(\Theta) = \{\theta' \in \mathcal{H} : ||\theta - \theta'||_2 \leq \varepsilon\}\) denote the \(\varepsilon\) ball around \(\Theta\).

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

\(^{10}\)We use the notation \(q(y)\) for the probability of outcome \(y\) if \(Y\) is finite, and for the probability density function of \(Q\) evaluated at \(y\) if \(Y\) is infinite.

\(^{11}\)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 \(Q'\).
\[
\min_{\theta \in B_r(\hat{\theta})} H(Q^*(\cdot|s,a), Q_\theta(\cdot|s,a)) < r \quad \forall \hat{\theta} \in \Theta, \forall s \in S, \forall a \in A.
\]

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 \( \pi \in \text{supp} \psi \) there exists a belief \( \mu \in \Delta(\Theta(\psi)) \) with \( \pi \in BR(\mu) \). A Berk-Nash equilibrium \((\Theta, \psi)\) 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 \( \text{supp} \psi = BR(\Delta(\Theta(\psi))) \).

(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.\(^{12}\) Section 6 extends this necessary condition to our large population setting, where non-unitary Berk-Nash 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 \( \mu \). Non-unitary equilibria only arise if multiple parameters minimize the weighted KL divergence from the equilibrium outcome distribution.\(^{13}\)

---

\(^{12}\)As with Nash equilibria in games, pure strategy Berk-Nash equilibria need not exist.

\(^{13}\)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.
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 \((\Theta, \psi)\) such that there is a \(\theta \in \Theta\) with \(Q_\theta(\cdot|s, \pi(s)) = Q^*(\cdot|s, \pi(s))\) for all \(\pi \in \text{supp} \psi\) 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.

2.3 Model expansions

We now introduce the idea that mutations might lead agents to expand the set of models they consider possible. We focus on two sorts of these expansions. The first is a quantitative local expansion that allows for all the parameter values within \(\varepsilon\) of a parameter the previous generation thought was possible.

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

The second sort of expansion is relevant when the subjective model is described by a finite collection of hypotheses about the underlying parameter that are expressed in the form of qualitative statements. 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. These hypotheses describe the parts of the subjective model of an agent that can be separately relaxed by a mutation. Formally, there is a finite collection of continuous

---

\(^{14}\)We follow Esponda and Pouzo (2016) in defining equilibrium as pairs of action and belief, so a pair \((\Theta, \psi)\) where the agent plays the objectively optimal action but has a dogmatic wrong belief is not a self-confirming equilibrium, in contrast to the original definition in Fudenberg and Levine (1993).
functions $\mathcal{F} = \{ f_l \}_{l=1}^m$, where each $f_l : \mathcal{H} \to \mathbb{R}$, such that $\Theta = \{ \theta \in \mathcal{H} : f_l(\theta) \geq 0, \forall l \in \{1, \ldots, m\} \} := \Theta_\mathcal{F}$.

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

### 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 the concept of “resistance to mutations,” which is not defined until Section 5, but we think that the intuition should still be 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 of this example 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 our criterion shows that 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 action would be better, so all the equilibria are mixed. The computations backing the claims in this and all subsequent examples are in Online Appendix B.3.

**Example 1.** The seller chooses price $p \in \{2, 10\}$ and receives payoff $u(a, y) = ay = a(i^* - \beta^*a + \omega)$, where $i^*$ and $\beta^*$ are the unknown intercept and slope of the demand function, and $\omega$ 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).\(^{15}\) 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 $\beta \leq 5/2$, so the KL-minimizer of the $\epsilon$ expansion of the subjective model mostly adjusts the slope upwards, which lead the mutants to choose the

15In 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.
optimal price of 2.\textsuperscript{16} Hence this Berk-Nash equilibrium does not resist local mutations.

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 $[3, 10/3] \times [33, 40]$. 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 bind. Because the average likelihood of the realized quantities 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 an $\varepsilon$ expansion revises the intercept upward. This is the case for the equilibrium in Esponda and Pouzo (2016\textsuperscript{16}), 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 $\psi(2) < \frac{97}{100}$, the main unexplained feature is high price sensitivity, so the mutants revise their

\textsuperscript{16}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.
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 misperceive a progressive tax schedule as linear, as in an example of Esponda and Pouzo (2016). 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 $c(a) = a/2$ and obtains income $z = a + \omega$, where $\omega \sim N(0, 1)$. The agent pays taxes $x = \tau^*(z)$,
where $\tau^*$ has two income brackets, and the higher one is heavily taxed:

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

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. $\tau_\theta(z) = (\theta + \eta)z$ and $\Theta = \mathbb{R}$. Given any action $a$, the KL-minimizing parameter is given by $\Theta(a) = \left( \mathbb{E} \left[ \frac{\tau^*(a + \omega)}{a + \omega} \right] \right)$.

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.

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 \( p_{\Pi} \).\(^{17}\)

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 \( \Theta \) when the previous state is \( p^t \). We require that this distribution satisfies the following inclusion, which captures the effect of learning and optimization:

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

(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 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 \( \Theta \) 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

\[
 p^{t+1}_{\mathcal{K}} = T(p^t) .
\]

(2)

\(^{17}\)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.
We say that $T$ is an evolutionary map if it is continuous, with $\text{supp } p_K = \text{supp } T(p)$ and payoff monotone (Samuelson and Zhang, 1992), meaning that

$$U^*(p(\cdot|\Theta)) > (=)U^*(p(\cdot|\Theta')) \implies \frac{T(p)(\Theta)}{T(p)(\Theta')} > (=)\frac{p_K(\Theta)}{p_K(\Theta')} \quad \forall p \in P,$$

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

This simple model of paradigm change can be interpreted as the result of biological reproduction or as the result of social learning and imitation. 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 performance 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.

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 $\Theta$ 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

---

18Recall that we restrict attention to states with finitely many models and strategies.

19One 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)).

20Payoff monotonicity assumes that what matters for change in population shares is the average payoff. This is not the case for all social learning procedures, see Ellison and Fudenberg (1993).
parameter in period $t$ acts as the validation sample. 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.\footnote{If these features were inherited, 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 between states with different payoffs, as in Example 5 in the Online 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 $\hat{p}_\Theta = \delta_\Theta$ for some $\Theta \in \mathcal{K}$, so the solution is a constant point mass on a single model.\footnote{The misspecified learning literature has focused on this case; Bohren and Hauser (2021) is an exception.} 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 $(\Theta, \psi)$ is a Berk-Nash equilibrium. Moreover, $\delta_\Theta \times \psi$ is unitary if and only if $(\Theta, \psi)$ 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.\footnote{Esponda and Pouzo (2016) assumes $Y$ is finite, but this is not needed for the proof of their Theorem 1.}

**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 $p_{\mathcal{K}}(C) = 1$. 

\[\text{15}\]
5 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 $\varepsilon$ mutation.\footnote{While we allow the consequences of a mutation to depend on the payoff it generates, unlike Bergin and Lipman (1996) we do not allow payoffs to influence which mutations are more likely.}

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

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

(ii) $p(\cdot|\tilde{\Theta}) \in \Delta(BR(\Delta(\tilde{\Theta}(\psi))))$ \quad \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.

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

For every $\varepsilon_n$, the inner limit gives the long-run strategy distribution following an $\varepsilon_n$ mutation to $\Theta'$; the outer limit sends the fraction of mutated agents to 0. The equilibrium $(\Theta, \psi)$ resists this mutation if this iterated limit converges back to $\psi$ for some solution that starts from the $\varepsilon_n$ mutation to $\Theta'$. We do not impose the stronger requirement that every possible solution converges back to $\psi$, which would rule out all self-confirming equilibria that are not Nash equilibria, even when the agent is correctly specified.

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 $\Theta'$ 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 realize that they could better explain their data within a more permissive paradigm $\Theta'$. A mutation to a smaller subjective model that generates a worse fit to the data can generate higher payoffs, as when a point mutation to a misspecified model leads the agent to play the optimal strategy. However, if $\Theta'$
does not drop the KL-minimizing parameters, it cannot destabilize the equilibrium. This is why our definition of ε mutations requires them to enlarge the set of subjective models. For such mutations, only explanation-improving mutations can be successful.

**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.

### 5.1 Local mutations

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.

**Definition.** p is the ε local mutation of a steady state δΘ × ψ if it is an ε mutation of δΘ × ψ to the ε expansion of Θ.

---

25 Appendix B.5 illustrates this. Our setup implies that mutations that do not lead to changes in play are abandoned, which rules out neutral drift, (Kimura, 1983; Binmore and Samuelson, 1999) where an initial mutation can be the springboard for subsequent mutations. We expect such drift to operate at a slower time scale than changes induced by mutations that allow better fit.

26 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.
A Berk-Nash equilibrium \((\Theta, \psi)\) resists local mutations if it resists mutation to every sufficiently small \(\varepsilon\) expansion of \(\Theta\). 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 \(\Theta\) resists local mutations.

In a uniformly strict equilibrium, the equilibrium strategy \(\pi\) is a strict best response to every parameter that minimizes the weighted KL divergence given that \(\pi\) is played. We show that this implies the beliefs of the agents after small mutations are concentrated on a neighborhood where the unique best reply is still \(\pi\), which yields Proposition 2 (i). Part (ii) follows from the fact that when the KL minimizers are in the interior of \(\Theta\), they have a strictly lower divergence than any parameter on or near the boundary of \(\Theta\).

Proposition 2 reinforces the finding of Fudenberg, Lanzani, and Strack (2021) that uniformly strict Berk-Nash equilibria have strong stability properties. In other Berk-Nash equilibria there may be a parameter \(\theta' \in \Theta(\pi)\) that does not induce \(\pi\) as the unique best reply. If the mutation leads to a KL minimizer that is near \(\theta'\), 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

\[
\mathcal{M}_{\Theta,\psi}(\varepsilon) = \arg\min_{\theta \in \Theta_\varepsilon} H_{\psi}(Q^*, Q_\theta)
\]

These parameters generate the largest decrease in \(H\).

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 \(\Pi_{\mathcal{M}_{\Theta,\psi}} = \lim\sup_{\varepsilon \to 0} BR(\Delta(\mathcal{M}_{\Theta,\psi}(\varepsilon)))\) denote the limits of the strategies that are optimal against distributions over \(\mathcal{M}_{\Theta,\psi}(\varepsilon)\) as \(\varepsilon\) goes to 0. We call these the local responses at \((\Theta, \psi)\).

---

\(^{27}\)This relies on the finiteness of the strategy set; Section 7 considers the case of infinitely many strategies.

\(^{28}\)\(\mathcal{M}_{\Theta,\psi}(\varepsilon)\) need not be a singleton, but it is a singleton for small \(\varepsilon\) in the examples we analyze. Moreover, as shown in Lemma 5, it is a singleton for sufficiently small \(\varepsilon\) if \(\Theta(\psi)\) is unique and \(Q_\theta\) is linear in \(\theta\).
Proposition 3. Let \((\Theta, \psi)\) be a Berk-Nash equilibrium.

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

(ii) If for some local response \(\pi'\) at \((\Theta, \psi)\), \(U^*(\pi') \leq U^*(\psi)\), and \((\Theta, \psi)\) is quasi-strict, it resists local mutations.

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 5.3 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.

When \(\Theta\) is finite and \(q_\theta(y|s, a)\) is continuously differentiable in \(\theta\) for all \(s \in S\) and \(a \in A\), there is a convenient way to check the conditions on \(\Pi_{\mathcal{M}_\Theta, \psi}\) in Proposition 3. Let \(S\) denote the sphere of radius 1 in \(\mathbb{R}^k\) with respect to the \(\| \cdot \|_2\) norm. Given a strategy distribution \(\psi\) and \(v \in S\), let

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

be the \(\psi\)-weighted directional derivative of \(-H\) in direction \(v\) at \(\theta\). As we show in Lemma 6 in the Appendix, it is enough to compute the direction \((\theta, v)\) in which \(D_\psi(\theta, v)\) is maximal, and look at the best replies to the parameters along that direction.
5.2 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 Ptolemaic cosmological model has two central tenets: that the Sun revolves around the Earth, and that stars revolve around the Earth-Sun pair. 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. This may help explain why the Ptolemaic system persisted for over a thousand years.

Definition.

$p$ is a one-hypothesis $\epsilon$ mutation of a steady state $\delta \times \psi$ if it is an $\epsilon$ mutation to some one-hypothesis relaxation of $\Theta$.

We say that a Berk-Nash equilibrium $(\Theta_F, \psi)$ resists one-hypothesis mutations if it resists every one-hypothesis $\epsilon$ mutation for sufficiently small $\epsilon$. There may be multiple sets of hypotheses that determine the same set $\Theta$; the way the agent mentally encodes the possible models is a key determinant of which equilibria resist one-hypothesis mutations.

Given a steady state $p = \delta \times \psi$ the $l$-agnostic $KL$ minimizers are $\mathcal{P}_l(p) = \arg\min_{\theta \in \Theta_l} H_\psi (Q^*, Q_\theta)$, and $\Pi_{p,l} = BR(\Delta(\mathcal{P}_l(p)))$ denotes the set of best replies when hypothesis $l$ is dropped.

Proposition 4. Let $(\Theta_F, \psi)$ 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 $(\Theta, \psi)$ does not resist one-hypothesis mutations.

(ii) If for every $l \in \{1, ..., k\}$ and $\pi' \in \Pi_{p,l}$, $U^*(\pi') \leq U^*(\psi)$, and either $\Pi_{p,l} \subseteq \text{supp} \psi$ or $\psi$ is a uniformly strict Berk-Nash equilibrium, then $(\Theta, \psi)$ 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.

29The unsuccessful "Andalusian revolt" in Islamic astronomy relaxed the latter assumption without abandoning geocentrism, see Sabra (1984).

30Note that here, unlike with local mutations, $\epsilon$ plays a single role.
In Example 1b, a one-hypothesis mutation that allows for larger slopes leads agents to play the objectively optimal price 2. Thus by Proposition 3(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 lead 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 $\omega$. They play a double auction 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 $\omega$ 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 $a \geq x$, and $\tilde{\omega} = \#$ otherwise.

Here a parameter $\theta$ is a pair of a 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 $H$ is the subset of $\mathbb{R}^9$ such that $\sum_{i=1}^3 p_i = 1$, and $0 \leq F(1|i) \leq F(2|i) \leq 1$ 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 B.3.3, $a = 2$ 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 \(\omega\). 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:

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

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.

5.3 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 better 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 \((\Theta, \psi)\) is **innovation vulnerable** if there exists \(\pi_I, \pi_U \in \Pi \setminus \text{supp} \psi\) and \(\epsilon > 0\) such that \(\{\pi_U\} = \arg\max_{\pi \in \Pi} U_\mu(\pi)\) for all \(\mu \in \Delta(\Theta(\psi'))\) with \(\psi' \in B_\epsilon(\psi)\), \(\psi'(\pi_I) > 0\). When \((\Theta, \psi)\) is innovation vulnerable, we say that \(\Theta' \supset \Theta\) is **innovation inducing** for \((\Theta, \psi)\) if \(BR(\Delta(\Theta'(\psi))) = \{\pi_I\}\).

In words, an equilibrium is innovation vulnerable if there is an unplayed best response \(\pi_U\) to some belief over equilibrium KL minimizers, and an innovative strategy \(\pi_I\) that provides evidence in favor of \(\pi_U\) even if \(\pi_I\) has a low payoff. A trivial case of an equilibrium that is not vulnerable to innovation is when \(\Theta\) 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. 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, 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 Online Appendix provides a fully detailed example of an innovation vulnerable equilibrium.

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.31

Suppose that all agents with subjective model Θ have the same prior \( \mu_\Theta \). Each agent born in a given period observes an independent sample of \( n \) randomly drawn individual experiences from the previous generation, updates their prior, and chooses a best reply to their posterior. Each agent \( i \in [0, 1] \) is endowed with a measurable best response function \( R : \Delta(\Theta) \to \Pi \), with \( R(\nu) \in BR(\nu) \) for all \( \nu \in \Delta(\Theta) \). Each agent observes \( n \in \mathbb{N} \) individual experiences drawn from \( p_\Pi \in \Delta(\Pi) \), independently across agents, computes posterior belief \( \mu_i \) using Bayes rule, and chooses a best reply, so aggregate play is \( \psi_n(\Theta, p_\Pi)(\pi) = \int 1_{R(\mu_i) = \pi} di \).

**Definition.** We say that \( p_\Pi \) distinguishes parameters and strategies if:

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

2. 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, \( p_\Pi \) 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) \( \Theta \) is finite and \( p_\Pi \) distinguishes parameters and strategies,

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

31This 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 (2021) or Bohren and Hauser (2021).

32Section B.6 of the Online Appendix shows that the limit belief is independent of the prior.
This shows that if agents observe enough individual experiences from the previous generation, the aggregate distribution of strategies is an element of $\Delta(BR(\Delta(\Theta(p_\Pi))))$. 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 $\pi_\Pi$ distinguishes parameters and strategies.\footnote{The Appendix proves this result under a more general condition that allows incomplete identification.} 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.\footnote{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.} 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.

## 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 $\Theta$ is convex, as in many examples in the literature.

**Assumption 2.**

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

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

(iii) $u$ is continuously differentiable in $a$ and $s$. 

\begin{enumerate}
\item The Appendix proves this result under a more general condition that allows incomplete identification.
\item 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.
\end{enumerate}
(iv) Θ is convex, and for all θ ∈ Θ, BR(θ) is a singleton and Qθ(·|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 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.\textsuperscript{35}

The stability of an equilibrium in this setting depends on Mθ,ψ(ε), introduced in Section 3.1 and the (objective) indirect utility function of the agent, which is V(θ) = U*(BR(θ)). We assume that V is continuously Gateaux differentiable. If Θ(ψ) is a singleton and Mθ,ψ(ε) is a singleton for sufficiently small ε, let V′(Mθ,ψ, ψ) = \lim inf\_{ε→0} \frac{V(Mθ,ψ(ε))-V(Θ(ψ))}{ε} be the derivative of V in the direction Mθ,ψ(ε).

**Proposition 7.** Let (Θ, ψ) be a Berk-Nash equilibrium such that Θ(ψ) is a singleton and Mθ,ψ(ε) is a singleton for sufficiently small ε. If V′(Mθ,ψ, ψ) > 0 then (Θ, ψ) does not resist local mutations.

This shows that if the derivative of the static indirect utility function in direction Mθ,ψ(ε) is positive the equilibrium does not resist local mutations.

**Example 4.** [Regression to the Mean] An instructor observes the initial performance s ∈ 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

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

\textsuperscript{35}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.
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 \( \eta \) is a standard normal, \( \theta_0 \) is the perceived correlation between performance in the two periods, and \( \theta_a \) 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 \( T \). Suppose \( \epsilon \) satisfies the assumption of Proposition 7. Then for every sufficiently small \( \epsilon \), there exists a threshold \( T \) such that \( \epsilon < 0 \). Since \( \epsilon \) is a Berk-Nash equilibrium, \( \epsilon \) is a singleton, none of which resist local mutations, where the instructor criticizes performances below \( T \) and praises performance above \( T \).

**Definition.** Let \( S \) be a singleton. We say that a pure Berk Nash equilibrium \((\Theta, \hat{a})\) is an attractor if \( \hat{a} \in A^0 \) and there is an \( \epsilon > 0 \) such that \(|a - \hat{a}| \leq \epsilon \) 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. 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 \(|A_n - A| \to 0\).

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

---

36 Section 13.3.4 shows there is a sequence of approximating discrete problems with unique Berk-Nash equilibria, none of which resist local mutations, where the instructor criticizes performances below \( T \) and praises performance above \( T \).

37 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.

38 The condition can fail if multiple actions induce the same payoff for all the parameters.
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 consider 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.

Our framework can be expanded to misspecification-driven inattention, 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. Online Appendix B.7 shows that when these assumptions are relaxed, the effect of attention partitions is ambiguous.

Finally, our model’s combination of Bayesian learning and evolutionary dynamics has a larger potential scope. First, the dynamic process we introduce 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. Second, our definition of resistance to mutations does not rely on the local or one-hypothesis structure. It only requires that if a new paradigm explains the observed data better than the old one does, it leads to a lower payoff for the agents who adopt it. This
more general criterion can be used to test resistance to other sorts of mutations.  

Last, we have focused on agents whose value for information is purely instrumental. By directly including beliefs in the utility function, our framework could be extended to agents who have an intrinsic preference for particular beliefs, either innately or from peer effects.

A Appendix

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

Conversely, if $(\Theta, \psi)$ 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 $(\Theta, \psi)$ be a Berk-Nash equilibrium and suppose that the mutation of $\delta_\Theta \times \psi$ to $\Theta'$ is not explanation improving. We show that for every $\varepsilon \in (0, 1)$ the following constant path is a solution: $p^t_K(\Theta) = 1 - \varepsilon; p^t_K(\Theta') = \varepsilon$, and $p^t_H(\cdot | \Theta) = \psi = p^t_H(\cdot | \Theta')$ 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 $\Theta(\pi, \varepsilon) := \{ \theta \in \mathcal{H} : \exists \theta' \in \Theta(\pi), ||\theta - \theta'||_2 \leq \varepsilon \}$. The next lemma is proved in the Online Appendix.

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

Proof of Proposition 2 (i) Let $(\Theta, \delta_\pi)$ 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 $\Pi$ is a singleton. Otherwise, let $G(\varepsilon) = \min_{\pi' \in \Pi \setminus \{ \pi \}} \min_{\mu \in \Delta(\Theta(\pi, \varepsilon))} \left( U_\mu(\pi) - U_\mu(\pi') \right)$.  

39For an axiomatic approach to these “backup” models see Ortoleva (2012).
Because $\Pi$ 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 $(\Theta, \pi)$ is a uniformly strict Berk-Nash equilibrium, $G(0) > 0$, and there is an $\hat{\varepsilon}$ such that if $\varepsilon \leq \hat{\varepsilon}$, $G(\varepsilon) > 0$.

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

(ii) Let $\partial \Theta$ denote the boundary of $\Theta$. Since $\partial \Theta$ is compact, when the KL minimizers are in the interior of $\Theta$, there is a $K \in \mathbb{R}^{++}$ and an $\hat{\varepsilon} \in \mathbb{R}^{++}$ such that if $\theta' \in \partial \Theta$ and $\theta$ is in $B_{\varepsilon}(\theta')$ then $H_{\psi}(Q^\ast, Q_{\theta}) - \arg\min_{\theta \in \Theta} H_{\psi}(Q^\ast, Q_{\theta}) > K$. 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 $\psi$ 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

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

denote the states where the strategy used by agents with model $\Theta'$ outperforms the strategy used by agents with model $\Theta$ by at least $\lambda$, and both population shares are larger than $\alpha$.

The proofs of Propositions 3, 4, and 7 use the following lemma, whose proof is in the Online Appendix. 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 $\Theta, \Theta'$ 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 $\Theta, \Theta' \in \mathcal{K}$, $\min_{p \in P_{\lambda, \alpha}(\Theta, \Theta')} \frac{T(p)(\Theta') p_{K}(\Theta)}{T(p)(\Theta) p_{K}(\Theta')}$ is well defined and strictly larger than 1. Thus, there is no solution that eventually stays in $P_{\lambda, \alpha}(\Theta, \Theta')$.

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

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

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

(3)

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

$$||\psi' - \psi|| < \varepsilon \Rightarrow BR(\Delta(\Theta_\varepsilon(\psi'))) \subseteq \Pi_{M_{\Theta,\psi}}.$$  

(4)

Suppose by way of contradiction that $\lim_{t \to \infty} p^t_\Pi = \psi$ and so $||p^t_\Pi - \psi|| < \varepsilon$ for all $t$ larger than some $\tau > 0$. For such $t$, $BR(\Delta(\Theta_\varepsilon(p^t_\Pi))) \subseteq \Pi_{M_{\Theta,\psi}}$ from equation (4), and since $\varepsilon < \varepsilon < \varepsilon^*$, equation (3) implies:

$$U^*(p^t_{\varepsilon} + (\cdot |_{\Theta_\varepsilon})) > U^*(p^t_{\varepsilon} + (\cdot |_{\Theta_\varepsilon})) + \gamma = p^t_{\varepsilon} + (\cdot |_{\Theta_\varepsilon}) + (1 - p^t_{\varepsilon})(\cdot |_{\Theta_\varepsilon}) + \gamma$$

so

$$U^*(p^t_{\varepsilon} + (\cdot |_{\Theta_\varepsilon})) > U^*(p^t_{\varepsilon} + (\cdot |_{\Theta_\varepsilon})) + \frac{\gamma}{1 - p^t_{\varepsilon}}.$$  

(5)

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

(ii) We will prove the following stronger result: If for some $\pi' \in \Pi_{M_{\Theta,\psi}}$, $U^*(\pi') \leq U^*(\psi)$, and $\pi' \in \operatorname{supp} \psi$, then $(\Theta, \psi)$ resists local mutations. Part (ii) of the Proposition follows from the fact that by the upper hemicontinuity of $M_{\Theta,\psi}(\cdot)$ and $BR(\cdot)$, the requirement $\pi' \in \operatorname{supp} \psi$ is always satisfied in a quasi-strict equilibrium.

Suppose that $U^*(\pi') \leq U^*(\psi)$, and $\pi' \in \operatorname{supp} \psi$ for some $\pi' \in \Pi_{M_{\Theta,\psi}}$. We will show that there is a solution in which the mutated agents always play $\pi' \in \Pi_{M_{\Theta,\psi}}$ 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 $\pi'$ has lower payoff than the equilibrium distribution guarantee that the share of mutants does not increase.

Let $\varepsilon = \min_{\pi \in \Pi_{\Theta,\psi}} \psi(\pi)$, and let $(\varepsilon_n)_{n \in \mathbb{N}_0} \in (0, \min\{\varepsilon', \varepsilon\})^{\mathbb{N}_0}$ be such that $\pi' \in BR(\Delta(M_{\Theta,\psi}(\varepsilon_n)))$
for all \( n \in \mathbb{N}_0 \) and \((\varepsilon_n)_{n \in \mathbb{N}_0} \to 0\). We will show that for every \( \varepsilon_n \) there is a solution \((p^t)_{t \in \mathbb{N}_0}\) with \( p^0 = p_{\varepsilon_n} \) and \( \lim_{t \to \infty} p^t_{II} = \psi \), where \( p_{\varepsilon_n} \) is the \( \varepsilon_n \) local mutation of \( p \).

Set \( \varepsilon = \varepsilon_n \) for some \( n \in \mathbb{N}_0 \) and let \( \Theta_\varepsilon \) be the local \( \varepsilon \) expansion of \( \Theta \). To define the candidate solution, let \( p^0_K(\Theta) = 1 - \varepsilon, p^0_K(\Theta_\varepsilon) = \varepsilon, p^0_{II}(\cdot|\Theta_\varepsilon) = \pi' \) and

\[
p^0_{II}(\pi|\Theta) = \begin{cases}
\frac{\psi(\pi)}{1-\varepsilon} & \pi \neq \pi' \\
\frac{\psi(\pi) - \varepsilon}{1-\varepsilon} & \pi = \pi'.
\end{cases}
\]

Also, recursively define \( p^{t+1}_K(\Theta) = T(p^t)(\Theta), p^{t+1}_K(\Theta_\varepsilon) = T(p^t)(\Theta_\varepsilon) \)

\[
p^{t+1}_{II}(\cdot|\Theta_\varepsilon) = \pi', \quad \text{and} \quad p^{t+1}_{II}(\pi|\Theta) = \frac{\psi(\pi)}{p^{t+1}_K(\Theta)} - \pi' \quad \text{if} \quad \pi \neq \pi'.
\]

By definition, \((p^t)_{t \in \mathbb{N}_0}\) satisfies equation (2). It remains to show that \( \psi(\pi') \geq p^t_K(\Theta_\varepsilon) \) for all \( t \), so that \( p^{t+1}_{II}(\cdot|\Theta) \) is a well defined element of \( \Delta(II) \), and that equation (II) is satisfied.

We prove this by induction. For the initial step, since \( \Theta_\varepsilon \) is the \( \varepsilon \) local mutation of \( \Theta \), and \( \varepsilon < \varepsilon' = \min_{\pi \in \text{supp } \psi} \psi(\pi) \), \( \psi(\pi') \geq p^0_K(\Theta_\varepsilon) = \varepsilon \). Moreover, by definition of \( \varepsilon = \varepsilon_n \) equation (1) is satisfied for \( t = 0 \). For the inductive step, observe that since the operator \( T \) is payoff monotone, \( U^*(\pi') \leq U^*(\psi) \), and \( \psi(\pi') \geq p^t_K(\Theta_\varepsilon) \) by the inductive hypothesis, we have \( p^{t+1}_K(\Theta_\varepsilon) \leq p^t_K(\Theta_\varepsilon) \leq \psi(\pi') \), and because \( \varepsilon = \varepsilon_n \) equation (I) is satisfied for \( t + 1 \).

Lemma 5. Let \((\Theta, \psi)\) be a Berk-Nash equilibrium. If \( \Theta \) is finite, \( \Theta(\psi) \) is a singleton, \( Q_\theta \) is linear in \( \theta \), 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 \( \varepsilon \leq \varepsilon' \).

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

The proofs of these two lemmas are in the Online Appendix.

Proof of Proposition 4. By Lemma 2 \( \mathcal{P}_l(p) \) is compact for each \( l \in \{1, \ldots, k\} \), and so is \( \{ \theta \in \mathcal{H} : \exists \bar{\theta} \in \mathcal{P}_l(p), ||\theta - \bar{\theta}|| \leq \varepsilon \} \). Therefore, because the action and signal spaces are finite, there is \( \varepsilon > 0 \) such that if \( \nu \left( \{ \theta \in \mathcal{H} : \exists \bar{\theta} \in \mathcal{P}_l(p), ||\theta - \bar{\theta}|| \leq \varepsilon \} \right) > 1 - \varepsilon \) and \( \pi \in BR(\nu) \), then \( \pi \in \Pi_{p,l} \). Because \( \Theta^I(\cdot) \) is upper-hemicontinuous (see Lemma 2), there are \( \varepsilon' > 0 \) and \( \gamma > 0 \) such that

\[
||\psi' - \psi|| < \varepsilon' \Rightarrow BR \left( \Delta \left( \Theta^I(\psi') \right) \right) \subseteq \Pi_{p,l} \quad \forall l \in \{1, \ldots, k\}.
\]
(i) We prove this case by contradiction. Suppose that for some \( l \in \{1, \ldots, k\} \), we have \( U^*(\pi) > U^*(\psi) \) for every \( \pi \in \Pi_{\ell,l} \). Because \( U^* \) is continuous and \( \Pi \) is finite, there is \( \varepsilon^* > 0 \) such that

\[
||\psi' - \psi|| < \varepsilon^* \Rightarrow U^*(\psi') - U^*(\psi'') < -\gamma \quad \forall \psi'' \in \Delta(\Pi_{\ell,l}).
\]

Let \( \Theta' \) be the one-hypothesis relaxation of \( \Theta \) in hypothesis \( l \). By definition of \( \varepsilon' \) we have

\[
||p_{\Pi}^t - \psi|| < \varepsilon' \Rightarrow p_{\Pi}^{t+1}(\cdot|\Theta') \in \Delta(\Pi_{\ell,l}) \quad \forall p^t \in P.
\]

Suppose by way of contradiction that for \( \varepsilon < \min\{\varepsilon', \varepsilon^*, (1 - \psi(\Pi_{\ell,l}))/2\} \) there is an \( \varepsilon \) mutation to \( \Theta' \), \( p_\varepsilon^0 \), such that \( \lim_{t \to \infty}(p_\varepsilon^t)_{\Pi} = \psi \). This means that after some \( \tau > 0 \), for all \( t > \tau \), \( ||(p_\varepsilon^t)_{\Pi} - \psi|| < \varepsilon \). Since \( \varepsilon < \varepsilon' \), by equation (9), \( p_{\varepsilon}^{t+1}(\cdot|\Theta') \in \Delta(\Pi_{\ell,l}) \). This, together with the assumption that for every \( \pi \in \Pi_{\ell,l} \), \( U^*(\pi) > U^*(\psi) \), implies that \( p_{\varepsilon}^{t+1}(\Theta) > (1 - \psi(\Pi_{\ell,l}))/2 > 0 \).

Since \( \varepsilon < \varepsilon^* \), by equation (8), \( U^*((p_{\varepsilon}^{t+1})_{\Pi}) + \gamma < U^*(p_{\varepsilon}^{t+1}(\cdot|\Theta')) \). Therefore, \( U^*(p_{\varepsilon}^{t+1}(\cdot|\Theta')) > p_{\varepsilon}^{t+1}(\Theta)U^*(p_{\varepsilon}^{t+1}(\cdot|\Theta)) + (1 - p_{\varepsilon}^{t+1}(\Theta))U^*(p_{\varepsilon}^{t+1}(\cdot|\Theta')) + \gamma \), so

\[
U^*(p_{\varepsilon}^{t+1}(\cdot|\Theta')) > U^*(p_{\varepsilon}^{t+1}(\cdot|\Theta)) + \frac{\gamma}{p_{\varepsilon}^{t+1}(\Theta)}.
\]

But then \( p_{\varepsilon}^t \in P_{\lambda,\alpha}(\Theta, \Theta') \) for all \( t > \tau \) with \( \alpha = \min\{p_{\varepsilon}^{t+1}(\Theta'), (1 - \psi(\Pi_{\ell,l}))/2\} \) and \( \lambda = \frac{\gamma}{p_{\varepsilon}^{t+1}(\Theta')} \), contradicting by Lemma 4.

(ii.a) Let \( \hat{\varepsilon} = \min_{\pi \in \text{supp} \psi}(\psi(\pi)) \). We will show that for every \( l \in \{1, \ldots, k\} \) and \( \varepsilon < \min\{\varepsilon', \hat{\varepsilon}\} \) there exists a solution \( (p^t)_{t \in \mathbb{N}_0} \) where \( p^0 \) is the \( \varepsilon \) mutation of \( p \) to the one-hypothesis relaxation of \( \Theta \) in hypothesis \( l \in \{1, \ldots, m\} \) and \( \lim_{t \to \infty}(p^t)_{\Pi} = \psi \). Fix such an \( \varepsilon \) and let \( \Theta' \) be the one-hypothesis relaxation of \( \Theta \) in hypothesis \( l \in \{1, \ldots, m\} \). Initialize the candidate solution by setting \( p_\varepsilon^0(\Theta) = 1 - \varepsilon, p_\varepsilon^0(\Theta') = \varepsilon, p_{\Pi}^0(\pi|\Theta) = \pi^\prime \in \Pi_{\ell,l} \) and \( p_\Pi^0(\pi|\Theta) \) as in equation (6), and recursively define subsequent states by \( \frac{p_{\Pi}^{t+1}(\pi|\Theta)}{p_{\Pi}^{t+1}(\Theta')} = \frac{\psi(\pi)}{p_{\Pi}^t(\Theta')} \), \( p_{\Pi}^{t+1}(\cdot|\Theta') = \pi' \) and

\[
\psi^t(\pi) = \begin{cases} 
\frac{\psi(\pi)}{p_{\Pi}^{t+1}(\Theta')} & \pi \neq \pi' \\
\frac{\psi(\pi)}{p_{\Pi}^{t+1}(\Theta')} & \pi = \pi'.
\end{cases}
\]

By definition, \( (p^t)_{t \in \mathbb{N}_0} \) satisfies equation (6). It only remains to show that \( \psi(\pi') \geq p_{\Pi}^0(\Theta') \) for all \( t \), so that \( p_{\Pi}^{t+1}(\cdot|\Theta) \) is a well defined element of \( \Delta(\Pi) \), and that equation (1) is satisfied.

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

(ii.b) Now suppose $(\Theta, \psi)$ is a uniformly strict Berk-Nash equilibrium. Let $\hat{\epsilon} > 0$ be small enough that $||\psi - \psi'|| < \hat{\epsilon}$ 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 $\hat{\epsilon}$ exists by Lemma \([2]\).) We show that for every $l \in \{1, ..., k\}$ and $\epsilon < \min\{\epsilon', \hat{\epsilon}\}$ there is a solution $(p')_{\epsilon \in \mathbb{N}_{0}}$ such that (i) $p^{0}$ is a mutation of $p$ to the one-hypothesis relaxation of $\Theta$ in hypothesis $l$ and (ii) $\lim_{t \to \infty} (p_{l}^{t})_{\Pi} = \psi$. Fix such an $\epsilon$ and let $\Theta'$ be the one-hypothesis relaxation of $\Theta$ in hypothesis $l \in \{1, ..., m\}$.

Initialize the candidate solution by setting $p_{K}^{0}(\Theta) = 1 - \epsilon, p_{K}^{0}(\Theta') = \epsilon, p_{H}^{0}(\cdot|\Theta') = \pi_{0}$ and $p_{H}^{0}(\cdot|\Theta) = \psi$, where $\pi_{0}$ is an arbitrary element of $\Pi_{p,l}$, and recursively define subsequent states by $p_{K}^{t+1}(\Theta) = T(p_{t}(\Theta))$ and $p_{K}^{t+1}(\Theta') = T(p_{t}(\Theta'), p_{H}^{t+1}(\cdot|\Theta') = \pi_{t}$ and $p_{H}^{t+1}(\pi|\Theta) = \psi$, where $\pi_{t}$ is a uniformly strict Berk-Nash equilibrium. Let $\Theta_{0}$ be a solution with $p_{0}^{l}(p_{l}^{0}) \in \Pi_{p,l}$ be a solution with $p_{l}^{0}$ and $\pi_{l}$ for every $l \in \{1, ..., m\}$. By definition, $(p_{l}^{t})_{\epsilon \in \mathbb{N}_{0}}$ satisfies equation \([2]\). It only remains to show that for every $t \in \mathbb{N}_{0}$, $\hat{\epsilon} \geq p_{l}^{t}(\Theta')$ so that equation \([1]\) is satisfied because $BR(\Delta(\Theta'(p_{l}^{t}))) \cap \Pi_{p,l} \neq \emptyset$ and $\psi \in \Delta(BR(\Delta(\Theta(p_{l}^{t}))))$. We prove this by induction. The initial step follows from the definition of $\hat{\epsilon}$. For the inductive step, observe that since the operator $T$ is payoff monotone, $U^{*}(\pi') \leq U^{*}(\psi)$, for every $\pi' \in \Pi_{p,l}$ and $\hat{\epsilon} \geq \epsilon \geq p_{l}^{t}(\Theta')$ by the inductive hypothesis, we have $p_{l}^{t+1}(\Theta') \leq p_{l}^{t}(\Theta') \leq \epsilon \leq \hat{\epsilon}$. The proof is concluded by observing that the previous inequality implies that as we let $\epsilon$ go to 0 we have $\lim_{t \to \infty} p_{l}^{t}(\Theta') = 0$.

Proof of Proposition \([5]\). Let $\Theta'$ be innovation-inducing for the innovation-vulnerable equilibrium $(\Theta, \psi)$. Since $U^{*}$ is continuous and $\Theta'$ is upper hemicontinuous by Lemma \([2]\) there is $\epsilon^{*} > 0$ such that

$$||\psi' - \psi|| < \epsilon^{*} \Rightarrow BR(\Delta(\Theta'(\psi'))) = \{\pi_{I}\}. \quad (10)$$

Let $(p_{l}^{t})_{\epsilon \in \mathbb{N}_{0}}$ be a solution with $p^{0} = p_{\epsilon}$, where $p_{\epsilon}$ is the $\epsilon$ mutation of $p$ to $\Theta'$ and $\epsilon < \epsilon^{*}$. By assumption, there is $\tilde{\epsilon} \in (0, \epsilon)$ such that

$$||p_{\Pi} - \psi|| < \tilde{\epsilon} \text{ and } p_{\cdot|\Theta'} = \{\pi_{I}\} \Rightarrow BR(\Delta(\Theta(p_{\Pi}))) = \{\pi_{U}\}. \quad (11)$$

Suppose by way of contradiction that after some $r > 0$, for all $t > r$, $||p_{\Pi}^{t} - \psi|| < \tilde{\epsilon}$. From equation \([10]\) this implies that $BR(\Delta(\Theta'(p_{\Pi}^{t}))) = \{\pi_{I}\}$. Since $\tilde{\epsilon} < \epsilon < \epsilon^{*}$, by equation \([11]\), $BR(\Delta(\Theta(p_{\Pi}))) = \{\pi_{U}\}$ a contradiction to $||p_{\Pi}^{t} - \psi|| < \tilde{\epsilon}$.

Proofs of Proposition \([6]\) and \([9]\). Let $E(\theta, \psi)$ be the parameters that are indistinguishable from $\theta$ under strategy distribution $\psi$, i.e., the $\theta'$ such that $\sum_{s \in \Pi} Q_{\theta}(s|s, \pi(s))\psi(\pi) =$
\[ \sum_{\pi \in \Pi} Q_{\psi}(\cdot | s, \pi(s)) \psi(\pi) \text{ for all } s \in S \text{ and for } \sum_{\pi \in \Pi} Q^\ast(\cdot | s, \pi(s)) \psi(\pi)\text{-almost every } y. \]

When aggregate play is \( \psi \), the relative likelihood of elements of \( \mathcal{E}(\theta, \psi) \) is determined by the prior. Let \( U(\pi | \mathcal{E}(\theta, \psi)) \) denote the subjective utility of strategy \( \pi \) under posterior \( \mu_\theta(\cdot | \mathcal{E}(\theta, p_\Pi)) \).

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

In words, strategies are subjectively different under \( p_\Pi \) if for every two strategies there is a class of indistinguishable parameters \( \mathcal{E}(\theta, p_\Pi) \) that minimize the weighted KL divergence given \( p_\Pi \) such that the utility of the strategies conditional to \( \mathcal{E}(\theta, p_\Pi) \) is different. If \( p_\Pi \) distinguishes parameters and strategies, strategies are subjectively different under \( p_\Pi \), as the former is the case when each \( \mathcal{E}(\theta, p_\Pi) \) 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 \( \psi \), and they do not assign the same probability to all events that \( \psi \) gives positive probability, then their strict convex combination has a strictly lower weighted KL divergence given \( \psi \).

**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

\[ -\int_{x \in X} \log \phi^1(x)d\Phi(x) = -\int_{x \in X} \log \phi^2(x)d\Phi(x) \text{ and } \phi^1 \text{ is not } \Phi\text{-almost surely equal to } \phi^2. \]

For every \( v \in (0, 1) \)

\[ -\int_{x \in X} \log[v\phi^1(x) + (1-v)\phi^2(x)]d\Phi(x) > -\int_{x \in X} \log \phi^1(x)d\Phi(x). \]

The proof of this lemma is in the Online Appendix.

**Proposition 9.** If either

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

(ii) \( \Theta \) is finite and strategies are subjectively different under \( p_\Pi \),

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 \( \lim_{n \to \infty} \mu_n(\Theta(p_\Pi)) = 1 \), and the upper-hemicontinuity of the best-reply correspondence implies that \( \psi_n \to \delta_{\hat{\pi}} \).

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 \( p_\Pi \). 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 \cite{strack2021} 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 $Q_{p \Pi}(s,a,B) = \sigma(s)\psi(s)(a)Q(B|s,a)$. Partition the elements of $\Theta$ in equivalence classes $\{\hat{\theta}^i\}_{i=1}^C$ such that

$$\sum_{\pi \in \Pi} q_{\theta^i}(|s,\pi(s))\psi(\pi) = \sum_{\pi \in \Pi} q_{\theta^2}(|s,\pi(s))\psi(\pi) \quad \forall s \in S, \forall \theta^1, \theta^2 \in \hat{\theta}^i$$

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

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

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

$$Z^l_m = \log \frac{\mu_m(\hat{\theta}^l)}{\mu_m(\hat{\theta}^1)} \quad \text{and} \quad L^l_m = \log \frac{q_{\theta^l}(y_m|s_m,a_m)}{q_{\theta^1}(y_m|s_m,a_m)},$$

so $Z^l_m = Z^l_0 + \sum_{i=1}^m L^l_i$.

**Claim 1.** The probability assigned to the KL-minimizing parameters goes to 1 $Q_{p \Pi}$-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.\textsuperscript{40}

**Proof.** If $\Theta = \Theta(p \Pi)$ the result is immediate. Suppose $K < C$. For $l > K$, $\mathbb{E}\left[\left|L^l_m \right|^m \right]$ is equal to

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

Since $\Theta \in \mathcal{K}$, $\mathbb{E}\left[\left|L^l_m \right|^m \right] < \infty$ and so by the Strong Law of Large Numbers and the Monotone Convergence Theorem, it follows that $\lim_{m \to \infty} e^{Z^l_m} = 0$ a.s. Therefore,

\textsuperscript{40}Unlike the related Lemma 2 of Esponda and Pouzo \cite{esponda2016}, this result allows $Y$ to be infinite.
\[
\limsup_{m \to \infty} \log \frac{\mu_m(\Theta \setminus \Theta(p_\Pi))}{\mu_m(\Theta(p_\Pi))} \leq \limsup_{m \to \infty} \log \frac{\mu_m(\Theta \setminus \Theta(p_\Pi))}{\mu_m(\tilde{\Theta})} = \limsup_{m \to \infty} \sum_{l=K+1}^\varphi \exp Z_m \stackrel{a.s.}{=} -\infty,
\]
proving the claim.

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

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

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

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 \(\|v\|_1 = 1\). We first observe that \(v^T \Sigma v\) is non-negative:

\[
v^T \Sigma v = \sum_{i=2}^K \sum_{j=2}^K v_i \Sigma_{ij} v_j = \sum_{i=2}^K \sum_{j=2}^K v_i v_j \int_{S \times A \times Y} \log \frac{q_{\vartheta^l}(y|s, a)}{q_{\vartheta^1}(y|s, a)} \log \frac{q_{\vartheta^l}(y|s, a)}{q_{\vartheta^1}(y|s, a)} \, dQ_{p_\Pi}(s, a, y)
\]
\[
= \int_{S \times A \times Y} \sum_{i=2}^K v_i \log \left( \frac{q_{\vartheta^l}(y|s, a)}{q_{\vartheta^1}(y|s, a)} \right) \sum_{j=2}^K v_j \log \left( \frac{q_{\vartheta^l}(y|s, a)}{q_{\vartheta^1}(y|s, a)} \right) \, dQ_{p_\Pi}(s, a, y)
\]
\[
= \int_{S \times A \times Y} \left( \sum_{i=2}^K v_i \log \left( \frac{q_{\vartheta^l}(y|s, a)}{q_{\vartheta^1}(y|s, a)} \right) \right)^2 \, dQ_{p_\Pi}(s, a, y) \geq 0.
\]

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

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

By Jensen’s inequality this implies that \(\log q_{\vartheta^l}(y|s, a) \leq \log \sum_{i=2}^K v_i q_{\vartheta^l}(y|s, a), Q_{p_\Pi}\)-almost surely, so \(q_{\vartheta^l}(y|s, a) \leq \sum_{i=2}^K v_i q_{\vartheta^l}(y|s, a)\). And as \(\theta^1\) 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 \(Q_{p_\Pi}\)-almost surely \(q_{\theta^l}(y|s, a) = \sum_{i=2}^K v_i q_{\theta^l}(y|s, a)\). By Lemma 7 this contradicts \(\theta^2 \in \Theta(p_\Pi)\). Thus \(v^T \Sigma v > 0\), so \(\Sigma\) is positive definite, proving Claim 2. \(\blacksquare\)
Claim 3. The set of \( \nu \in \Delta \left( \left\{ \hat{\theta}^1, \ldots, \hat{\theta}^K \right\} \right) \) such that

\[
\sum_{i=1}^{K} \nu(\hat{\theta}^i) \sum_{\theta_j \in \hat{\theta}^i} \frac{\mu_\theta(\theta_j)}{\mu_\theta(\hat{\theta}^i)} U_{\theta_j}(\pi) = \sum_{i=1}^{K} \nu(\hat{\theta}^i) \sum_{\theta_j \in \hat{\theta}^i} \frac{\mu_\theta(\theta_j)}{\mu_\theta(\hat{\theta}^i)} U_{\theta_j}(\pi').
\]

(12)

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

The proof of this claim is in the Online Appendix.

Note that because \( \left( Z_l^t \right)_{l=2}^K \) is a martingale with positive definite covariance matrix of the increments, the central limit theorem implies that \( \left( \frac{Z_m^t}{\sqrt{\pi}} \right)_{l=2}^K \) converges in distribution to a \( K - 1 \) dimensional normal distribution with mean \( \hat{0} \) and covariance matrix \( \Sigma \). Since

\[
\mu_m(\hat{\theta}_l) = \frac{\exp Z_m^l}{\sum_{i=2}^C \exp Z_m^i + 1}, \quad \forall l \in \{2, \ldots, C\},
\]

the distribution on the indifference classes induced by \( \mu_m \) converges to some \( \nu \in \Delta \left( \Delta \left( \Theta (p_{\Pi}) \right) \right) \), as does the distribution of beliefs over the indifference classes in the overall population. Claim 2 shows that \( \Sigma \) 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 \( \Delta \left( BR \left( \Delta \left( \Theta (p_{\Pi}) \right) \right) \right) \). This concludes the proof of part (ii).

References


B For Online Publication

B.1 Proof of the results of Section 7

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 $\varepsilon < \varepsilon'$, $M_{\Theta, \psi}(\varepsilon)$ is a singleton, and $V(M_{\Theta, \psi}(\varepsilon)) > U^*(\psi)$. We now prove that $(\delta_{\Theta}, \psi)$ does not resist mutation to $\Theta'$ if $\Theta'$ is an $\varepsilon$ expansion of $\Theta$ for $\varepsilon < \varepsilon'$.

By continuity of $U^*$ in $\psi$, there is an $\varepsilon^* \in (0, \varepsilon')$ and $\gamma > 0$ such that for all $\psi' \in B_{\varepsilon^*}(\psi)$ $\tilde{\psi} \in BR(\Delta(B_{\varepsilon}(M_{\Theta, \psi}(\varepsilon)))$,
\[
U^*(\psi') - U^*(\tilde{\psi}) < -\gamma. \tag{13}
\]

By the upper hemicontinuity (see Lemma 2) of $\Theta'()$ there is $\bar{\varepsilon}$ such that $||\psi' - \psi|| < \bar{\varepsilon}$ implies $\Theta'(\psi') \subseteq B_{\varepsilon}(M_{\Theta, \psi}(\varepsilon))$. Thus if there were a $t$ such that $||p_{\Pi}^t - \psi|| < \min\{\varepsilon^*, \bar{\varepsilon}\}$ for all $t \geq t$, it would follow from equation (13) that
\[
U^*(p_{\Pi}^{t+1}(\cdot|\Theta)) + \frac{\gamma}{p_{\psi}^\varepsilon(\Theta)} < U^*(p_{\Pi}^{t+1}(\cdot|\Theta')). \tag{14}
\]

By Lemma 4 this concludes the proof.

Proof of Proposition 8. Suppose that $\hat{a}$ is an attractor with associated unique KL minimizer $\hat{\theta} \in \Theta(\hat{a})$. Let $A_m$ be given by $\{\hat{a} + \frac{1}{m}c\}_{c \in \mathbb{Z} \setminus \{0\}} \cap A$, and let $\hat{\varepsilon}$ be such that for all $\varepsilon < \hat{\varepsilon}$, $V(M_{\delta_{\Theta}, \hat{a}}(\varepsilon)) > V(\hat{\theta})$. Let $BR_n(\theta) = \arg\max_{a \in A_n} U_\theta(a)$. Since $\hat{a}$ is in the interior of $A$, $A_m$ is nonempty for sufficiently large $m$. By Theorem 1 in Espanda and Pouzo (2016), for every $m \in \mathbb{N}$, the environment with the discrete action set $A_m$ admits at least one Berk-Nash equilibrium. So, for all $m \in \mathbb{N}$ we can pick an equilibrium $(\Theta, \psi_m)$ and a justifying belief $\mu_m \in \Delta(\Theta(\psi_m))$. Since $A$ is compact, $\Delta(A)$ is also compact, and therefore $(\psi_m, \mu_m)$ admits a converging subsequence to a limit $\left(\hat{\psi}, \hat{\mu}\right)$.

We claim that the limit of this subsequence is $(\delta_{\hat{\theta}}, \delta_{\Theta(\hat{a})})$. 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 $a_n \rightarrow a$, $a_n \in A_n$, and since $U(\cdot)$ is jointly continuous in beliefs and actions $U_{\mu_n}(\psi_n) \geq U_{\mu_n}(a_n)$ for all $n \in \mathbb{N}$ implies $U_{\hat{\mu}}(\hat{\psi}) \geq U(a)$. Moreover, by the upper hemicontinuity (see Lemma 2) of $\Theta$, $\mu_n \in \Delta(\Theta(\psi_n))$ for all $n \in \mathbb{N}$ implies that $\hat{\mu} \in \Delta(\Theta(\hat{\psi}))$. 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 requirement in the statement of the proposition. It is clear that $||A_n \rightarrow A|| \rightarrow 0$. Since $\mu_n \rightarrow \delta_{\Theta(\hat{a})}$, there exists a $N > 0$ such that for all $n > N$, $\supp \psi_n \subseteq B_{\varepsilon}(\hat{a})$. Therefore, since $(\hat{a}, \Theta)$ is an attractor, there is an $N^{'} > N$ such that for $n \geq N^{'}$, $\psi_n \{a : a < \hat{a}\} > 0$ and $\psi_n \{a : a > \hat{a}\} > 0$.

Since $U^*$ is continuously differentiable and $(\delta_{\hat{a}}, \delta_{\Theta(\hat{a})})$ does not resist local mutations, either $U^*$ is strictly increasing in $a$ on $[\hat{a}, \hat{a} + K]$ for some $K > 0$, or $U^*$ is strictly decreasing on $[\hat{a} - K', \hat{a}]$ for some $K' > 0$. We prove the result in the first case, the other case is
analogous. Because \( V(M_{\Theta,\psi}(\varepsilon)) = V(M_{\Theta,\psi}(\hat{a})) = U^*(\hat{a}) \) for all \( \varepsilon < \hat{\varepsilon} \), there exists an \( \bar{\varepsilon} < \hat{\varepsilon} \) such that for all \( \varepsilon < \bar{\varepsilon} \), \( BR(M_{\Theta,\psi}(\varepsilon)) \in (\hat{a}, \hat{a} + K) \). Fix an \( \varepsilon < \bar{\varepsilon} \). Because \( \Theta_\varepsilon(\cdot) \) is upper hemicontinuous (see Lemma 2), \( \Theta_\varepsilon(\psi_n) \) converges to \( M_{\Theta,\psi}(\varepsilon) = \Theta_\varepsilon(\hat{a}) \). Thus there exists \( N'' > N' \) such that for all \( n > N'' \), \( BR_n(\Theta_\varepsilon(\psi_n)) \in \left(\frac{\hat{a} + BR(M_{\Theta,\psi}(\varepsilon))}{2}, \hat{a} + K\right) \). But since \( U^* \) is strictly increasing on \([\hat{a}, \hat{a} + K]\), \( U^*(\psi_n) \to U^*(\hat{a}) \), and \( \psi_n(\{a : a < \hat{a}\}) > 0 \), there exists a \( N''' > N'' \) such that for all \( n > N''' \), \( U^*(BR_n(\Theta_\varepsilon(\psi_n))) > U(\psi_n) \), so that \( \psi_n \) does not resist an \( \varepsilon \) expansion of \( \Theta \).

\[ \]  

B.2 Proofs of results stated in the Appendix

B.2.1 Proof of Lemma 3

**Proof.** Suppose not, and let \( (\Theta'_n)_{n \in \mathbb{N}} \) be a sequence of \( \varepsilon_n \) local expansions of \( \Theta \) with \( \varepsilon_n \downarrow 0 \) and \( \theta_n \in \Theta'_n(\pi) \setminus \Theta(\pi, \varepsilon) \). Since \( \Theta'_1 \) is compact, the sequence has an accumulation point \( \theta \in \Theta \). If \( \theta \in \Theta(\pi) \) then \( (\theta_n)_{n \in \mathbb{N}} \) is eventually in \( \Theta(\pi, \varepsilon) \), a contradiction. If \( \theta \notin \Theta(\pi) \), then since \( H \) is lower semi-continuous in \( \theta \), \( H_\pi(Q^*, Q_{\theta_n}) \geq \min_{\theta \in \Theta} H_\pi(Q^*, Q_\theta) \), which contradicts with \( \theta_n \in \Theta'_n(\pi) \).

B.2.2 Proof of Lemma 4

**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 \( p_K(\Theta') \geq \alpha > 0 \) on \( P_{\lambda,\alpha}(\Theta, \Theta') \), so \( \frac{T(p(\Theta'))}{T(p(\Theta)} \) 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 \( \Theta' \) strictly outperforms the strategy used by agents with subjective model \( \Theta \) on \( P_{\lambda,\alpha}(\Theta, \Theta') \), the minimum satisfies \( m \geq 1 \).

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 \( \Theta \) and \( \Theta' \) grows multiplicatively by at least \( m \), so there is a \( \tau \leq \log_m(1/\alpha^2) + t \) such that \( p_T(\Theta) < \alpha \). Thus the solution leaves \( P_{\lambda,\alpha}(\Theta, \Theta') \) before time \( \tau \).

B.2.3 Proof of Lemma 5

**Proof.** Let \( \{\hat{\theta}\} = \Theta(\psi) \). By Lemma 3, there is \( \varepsilon' > 0 \) such that for all \( \varepsilon \leq \varepsilon' \), \( M_{\Theta,\psi}(\varepsilon) \subseteq B_{\varepsilon}(\hat{\theta}) \). Assume by way of contradiction that \( \theta', \theta'' \in M_{\Theta,\psi}(\varepsilon), \varepsilon \leq \varepsilon', \theta' \neq \theta'' \). 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, \( H_\psi(Q^*, Q_{\theta'/2 + \theta''/2}) < H_\psi(Q^*, Q_{\theta'}) \), a contradiction.  

\(^{41}\)The lower semicontinuity of the \( H \) in the probability measure follows from e.g., Theorem 1.47 of Liese and Vajda (1987), and then the lower semicontinuity in \( \theta \) follows from the continuity in \( \theta \) of the probability measure imposed by Assumption \( \alpha(i) \).
B.2.4 Proof of Lemma \[6\]

**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 \(D_{\Pi}^\epsilon\) is continuous at every \((\hat{\theta}, \hat{v})\).

We now prove that for all \(\lambda > 0\) there exists an \(\epsilon \in \mathbb{R}_{++}\) such that for all \(\epsilon < \epsilon\), \(\mathcal{M}_{\theta, \psi}(\epsilon) \subseteq B_{\epsilon}(\hat{\theta})\), and for all \(v \in \mathbb{R}_k\), \(|D_\psi(\hat{\theta}, v) - \epsilon[H_\psi(Q^*, Q_{\theta + v, \epsilon}) - H_\psi(Q^*, Q_\theta)]| < \lambda\).

Suppose by contradiction that there exist \((\epsilon_n)_{n \in \mathbb{N}} \to 0\) and \((\theta_n)_{n \in \mathbb{N}}\) with \(\theta_n \in \Theta_{\epsilon_n}(\psi) \setminus B_{\epsilon_n}(\hat{\theta})\). Since \(\Theta_{\epsilon_1}(\psi)\) is compact by Lemma \[2\], the subsequence can be taken to be convergent to some \(\tilde{\theta}'\) and with \(\theta_n \in B_{\epsilon_n}(\tilde{\theta})\) for some \(\tilde{\theta} \in \Theta(\psi)\). 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 \(\tilde{v}\).

By definition of \(\theta_n\) we have

\[
\sum_{s \in S} \sigma(s) \sum_{\pi \in \Pi} \psi(\pi) \left( \int_{y \in Y} \log \frac{q_{\theta + v_n}(y|s, \pi(s))}{q_{\theta_n}(y|s, \pi(s))} dQ^*(y|s, \pi(s)) \right) \leq 0
\]

and since \(H(Q^*|s, a), Q_{\theta}(|s, a)\) is twice continuously differentiable at \(\hat{\theta}\) and \(\tilde{\theta}'\) this implies that \(D_\psi(\hat{\theta}, \tilde{v}) \leq D_\psi(\tilde{\theta}', \tilde{v})\) a contradiction with \(\arg\max_{\theta \in \Theta(\psi), v \in \mathbb{S} \cup \emptyset} D_\psi(\theta, v) = \{ (\hat{\theta}, \hat{v}) \}\).

Finally, because there are a finite number of strategies, if \(\hat{\pi}\) is in \(\text{argmax}_{\pi \in \Pi} U_{\delta + \epsilon}(\pi)\), \(\hat{\pi} \in BR(\Delta(\mathcal{M}_{\theta, \psi}(\epsilon)))\).

B.2.5 Proof of Lemma \[7\]

**Proof.** Since \(\phi^1\) is not \(\Phi\)-almost surely equal to \(\phi^2\) 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

\[
\overline{K} \leq \phi^1(x) \leq \overline{K} \text{ and } \overline{K} \leq \phi^2(x) \leq \overline{K} \text{ for all } x \in B
\]

for some \(\overline{K}, \overline{K}\). Let

\[
\rho = \min_{z \in [\overline{K}, \overline{K}], z' \in [z + \overline{K}, \overline{K}]} \log (vz + (1 - v) z') - v \log (z) - (1 - v) \log (z') > 0 \quad (15)
\]

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.

OA-3
Notice that the formula for relative entropy can be expanded as
\[-\int_{Y\setminus B} \log \left( v \phi^1(x) + (1 - v) \phi^2(x) \right) d\Phi(x) - \int_Y \log \left( v \phi^1(x) + (1 - v) \phi^2(x) \right) d\Phi(x)\]
\[\leq - \int_{Y\setminus B} \log \left( v \phi^1(x) + (1 - v) \phi^2(x) \right) d\Phi(x) - \int_B \left( v \log \phi^1(x) + (1 - v) \log \phi^2(x) + \rho \right) d\Phi(x)\]
\[= -v \int_Y \log \phi^1(x) d\Phi(x) + (1 - v) \int_Y \log \phi^2(x) d\Phi(x) - \rho \Phi(B)\]
\[= \int_{x \in X} \log \phi^1(x) d\Phi(x) - \rho \Phi(B)\]
as desired.

\section*{B.2.6 Proof of Claim 3}

\textbf{Proof} Fix $\pi \neq \pi'$. Equation (12) is a linear equation in the $K$ unknowns $\nu(\hat{\theta}^i)$, so its solutions are a vector subspace of $\mathbb{R}^K$. Since strategies are subjectively different under $p_H$ there exists $\hat{\theta}_i \in \Theta(p_H)$ such that $U_{\mu(\cdot|\hat{\theta}_i)}(\pi) \neq U_{\mu(\cdot|\hat{\theta}_i)}(\pi')$, 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 $\pi, \pi'$ are chosen arbitrary, the set of beliefs $\nu \in \Delta \left( \{\hat{\theta}^1, \ldots, \hat{\theta}^K\} \right) \subseteq \mathbb{R}^K$ such that $U(\nu, \pi) = U(\nu, \pi')$ for some $\pi \neq \pi'$ has Lebesgue measure 0 as well.

\section*{B.3 Examples}

\subsection*{B.3.1 Example 1}

Esponda and Pouzo (2016) shows that
\[
\sum_{a \in \{1, 2\}} \psi(a) H(Q^s(\cdot|a), Q^0(\cdot|s, a)) = \psi(2)(34 - i + 2\beta)^2 + \psi(10)(2 - i + 10\beta)^2. \tag{16}
\]

a) When $\psi(2) = 1$, the parameter that minimizes equation (16) is $(3/2, 32)$, and since $BR(\delta_{(3/2, 32)}) = \{10\}$, $(\Theta, \delta_2)$ is not a Berk-Nash equilibrium. When $\psi(10) = 1$, the parameter that minimizes equation (16) is $(5/2, 28)$. Since $BR(\delta_{(5/2, 28)}) = \{2\}$, $(\Theta, \delta_{10})$ is not a Berk-Nash equilibrium. For every totally mixed $\psi$, the Hessian of equation (16) as a function of $\beta$ 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 $\beta$ of the resulting expression is strictly negative for every $\psi$. Therefore, the unique parameter on the line $i = 12\beta$ where the two actions are indifferent that minimizes equation (16) for some $\psi$ 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 \( \varepsilon \) expansion of the model is:

\[
-2\frac{1}{4}(34 - i + 2 \ast (2.5 + \varepsilon)) - 2(1 - \frac{1}{4})(2 - i + 10 \ast (2.5 + \varepsilon)) = 0
\]

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

b) When \( \psi(2) = 1 \), the parameter that minimizes equation (16) is \( (3, 40) \), and since \( BR(\delta_{(3, 40)}) = \{10\} \), \( (\Theta, \delta_2) \) is not a Berk-Nash equilibrium. When \( \psi(10) = 1 \), all the \( (\beta, i) \) with \( i = 10\beta + 2 \), \( \beta \in (3, 10/3) \) minimize equation (16). Since \( BR(\delta_{(\beta, i)}) = \{2\} \) for all such \( (\beta, i) \), \( (\Theta, \delta_{10}) \) is not a Berk-Nash equilibrium. The first order conditions for \( (10/3, 40) \) to be the KL-minimizer are

\[
-2\psi(2)(34 - 40 + 2 \ast \frac{10}{3}) - 2(1 - \psi(2))(2 - 40 + 10 \ast \frac{10}{3}) \leq 0 \tag{17}
\]

\[
4\psi(2)(34 - 40 + 2 \ast \frac{10}{3}) + 20(1 - \psi(2))(2 - 40 + 10 \ast \frac{10}{3}) \leq 0. \tag{18}
\]

The first inequality gives \( \psi(2) \geq \frac{7}{8} \), while the second gives \( \psi(2) \leq \frac{35}{36} \).

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

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

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

### B.3.2 Example 2

Each parameter \( \theta \) generates distribution \( q_\theta(z, x | a) = \phi_1(z - a)\phi_\theta(x | z) \) on \( (z, x) \), where \( \phi_1 \) 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 \( \phi_\theta \) is a normal density, for the restricted linear model where \( \theta_2 = 0 \) we have

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

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:

\[
\arg\min_{(\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}_+} \mathbb{E} \left[ (\tau(5 + \omega) - \theta_1 (5 + \omega) - \theta_2 (5 + \omega)^2)^2 \right].
\]
B.3.3 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

\[
u(s, a, y) = \sum_{\omega=1}^{\alpha} p_\omega (\omega + 3.1 + s - a) 1_{a \geq \omega}.
\]

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

\[
\sum_{y \in Y} u(-1, 3, y)p^*(y) = \frac{1}{3}(3 + 3.1 - 1) + \frac{1}{2}(2 + 3.1 - 1) + \frac{1}{6}(1 + 3.1 - 1) - 3
\]

\[
> \frac{1}{2}(2 + 3.1 - 1 - 2) + \frac{1}{6}(1 + 3.1 - 1 - 2)
\]

\[
= \sum_{y \in Y} u(-1, 2, y)p^*(y) > \frac{1}{6}(1 + 3.1 - 1 - 1) = \sum_{y \in Y} u(-1, 1, y)p^*(y).
\]

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 \( s = 0 \) 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, \( \hat{\theta} := \Theta(2) = (\frac{1}{6}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, 1, \frac{1}{4}, 1, \frac{1}{4}, 1) \), and under \( \hat{\theta} \) bidding 2 is optimal:

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

In this equilibrium, for every \( \omega \in \{1, 2, 3\} \), relaxing the hypotheses \( F(\omega | 3) \geq F(\omega | 3) \) or \( F(\omega | 3) \leq F(\omega | 3) \) 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 \( \hat{\theta} := \Theta(\pi) = (\frac{1}{6}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{4}{5}, \frac{4}{5}, \frac{1}{4}, \frac{1}{4}) \), and under \( \hat{\theta} \) bidding 2 after signal \( s = -1 \) is optimal:

\[
\sum_{y \in Y} u(-1, 1, y)p_{\hat{\theta}}(y) = \frac{1}{6}(3.1 - 1) + \frac{1}{5}(4.1 - 1) + \frac{1}{5}(5.1 - 1)
\]

\[
< \frac{1}{5}(3.1 - 3) + \frac{3}{5}(4.1 - 3) + \frac{1}{5}(5.1 - 3) = \sum_{y \in Y} u(-1, 3, y)p_{\hat{\theta}}(y)
\]

\[
< \frac{2}{3}(3.1 - 2) + \frac{3}{5}(4.1 - 2) + \frac{1}{5}(5.1 - 2) = \sum_{y \in Y} u(-1, 2, y)p_{\hat{\theta}}(y).
\]
Moreover, under $\tilde{\theta}$ bidding 3 after signal $s = 1$ is optimal:

$$\sum_{y \in Y} u(1, 1, y) p_\theta(y) = \frac{1}{6} (\frac{1}{5} (5.1 - 1) + \frac{3}{5} (6.1 - 1) + \frac{1}{5} (7.1 - 1))$$

$$< \frac{2}{3} (\frac{1}{5} (5.1 - 2) + \frac{3}{5} (6.1 - 2) + \frac{1}{5} (7.1 - 2)) = \sum_{y \in Y} u(1, 2, y) p_\theta(y)$$

$$< (\frac{1}{5} (5.1 - 3) + \frac{3}{5} (6.1 - 3) + \frac{1}{5} (7.1 - 3)) = \sum_{y \in Y} u(1, 3, y) p_\theta(y).$$

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

$$\arg\min_{F(1|1), F(2|1), F(2|3)} -[\log(F(1|1))/6 + \log(F(2|1) - F(1|1))/2] - \frac{1}{2}[\log(1 - F(2|3))/3].$$

B.3.4 Grid approximation of Example 4

Recall that in the model with a continuum of signals the instructor criticizes performance below $T = (-E_s[-s|s > T] - E_s[-s|s < T])/k$ and praises performance above it. Fix an integer $n$, and suppose that the signals take values on the grid $\{T + \frac{m}{n} : m \in \mathbb{N}\} \cap [-n^2, n^2]$ so that as $n$ increases the grid becomes finer and with a wider range, and adjust the objective distribution over states by mapping the probability mass in any interval to one of the extremes so to maintain the same expected value.\footnote{The key is that the distance between the threshold and its closest element is $o(1/n)$. If the grid is coarser than this, there can also be uniformly strict equilibria.} Formally, let $T + \frac{m}{n}$ and $T + \frac{m+1}{n}$ be two elements of the grid. The fraction $\lambda_m \in (0, 1)$ of the probability of the interval $T + \frac{m}{n}$ and $T + \frac{m+1}{n}$ that goes to $m$ is such that

$$\lambda_m \left( T + \frac{m}{n} \right) + (1 - \lambda_m) \left( T + \frac{m+1}{n} \right) = E_s \left[ s | s \in \left( T + \frac{m}{n}, T + \frac{m+1}{n} \right) \right].$$

Denote the new probability distribution over states as $\sigma_g$. In this case, the unique equilibrium is such that $\pi(s) = \delta_c$ for $s < T$, $\pi(s) = \delta_r$ for $s > T$, and $\pi(T)(c) \in (0, 1)$. The behavior strictly below and above the threshold follows directly from the computations in Esponda and Pouzo (2016), so it only remains to show that $\pi(T)(c) \in (0, 1)$. Suppose by way of contradiction that the instructor always praises at the threshold, $\pi(T)(c) = 0$. Then,

$$\theta_c(\pi) = -E_{\sigma_g} \left[ s | s \leq T - \frac{1}{n} \right] > -E_{\sigma_g} \left[ s | s \leq T \right] = \theta_c(T)$$

and

$$\theta_r(\pi) = -E_{\sigma_g} \left[ s | s \geq T \right] < -E_{\sigma_g} \left[ s | s \geq T \right] = \theta_r(T).$$
and therefore
\[ \frac{\theta_c(\pi) - \theta_r(\pi)}{k} > \frac{\theta_c(T) - \theta_r(T)}{k} = T. \]

But then the threshold to praise is strictly larger than T, a contradiction. A symmetric argument shows why the instructor cannot always criticize at the threshold.

To see that this equilibrium does not resist local mutations, note that $M_{\Theta, \psi}(\varepsilon)$ is still $((1, \theta_a(s), \theta_b(s)), (v_0(\varepsilon), v_(\varepsilon), v_3(\varepsilon))$ with $v_0(\varepsilon) < 0, v_3(\varepsilon) \geq 0, v_3(\varepsilon) \leq 0$. The best reply to these parameters is the same as the equilibrium strategy, except at the threshold performance it praises instead of mixing. Since $s < 0$, this strategy outperforms the equilibrium, which is thus not resistant to the mutation.

### B.3.5 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 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

\[ \hat{p} = 0.1\delta_{\Theta_1 \times \delta_a} + 0.9\delta_{\Theta_2 \times \delta_c} \]

and $\bar{p} = 0.9\delta_{\Theta_1 \times \delta_a} + 0.1\delta_{\Theta_2 \times \delta_c}$.

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 $\Theta_1$ is higher (they play $a$) than that of those with subjective model $\Theta_2$ (they play $c$). Moreover under $\bar{p}$ the performance of the agents with subjective model $\Theta_1$ is lower (they play $b$) than that of those with subjective model $\Theta_2$ (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 $0.5 \cdot 0.1 + 0.01 \cdot 0.9$ in the even periods and $0.5 \cdot 0.1 + 0.02 \cdot 0.9$ in the odd periods, so that the average payoff sequence is not monotone. ▲

### B.4 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:

\[
\begin{align*}
  u(a, (1, y_2, y_3)) &= 1 = u(b, (0, y_2, y_3)) \\
  u(a, (0, y_2, y_3)) &= 0 = u(b, (1, y_2, y_3)) .
\end{align*}
\]

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

\[
\begin{align*}
  u(c, (y_1, y_2, 1)) &= 1; \\
  u(c, (y_1, y_2, 0)) &= 0.
\end{align*}
\]
The parameter space has two dimensions $\mathcal{H} = [0, 1] \times [0, 1]$, where $\theta = (\theta_1, \theta_2) \in \mathcal{H}$, $\theta_1$ 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 $\theta_2$ 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

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

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

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

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 $p = (\delta_0, \delta_a)$ is a steady state: every parameter induces the same weighted KL divergence and $\phi$ is a best reply to any subjective model in which $\theta_1 = 1/2$ and $\theta_2 \leq 1/2$. The equilibrium is not quasi-strict: for every belief supported on $\Theta$, $\phi$ is a best reply if and only if $b$ is.

A mutation to $\hat{\Theta}_\varepsilon$ with $\varepsilon < 1/2$ induces $b$ as the unique best reply, since the mutated agents decrease $\theta_1$ 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 $\Delta(\Theta(\psi))$, is better than either alternatives. Thus the entire population switches to $c$ forever from the second period onward, so the equilibrium is innovation-vulnerable. ▲

### B.5 Mutations to smaller subjective models

**Invasion by a non-improving mutation** Consider the subjective model $\Theta$ and any completely mixed equilibrium $\psi$ of Example [1]. Let $\Theta' = \{(3,33)\}$. Since $\Theta' \subset \Theta$, and $\Theta(\psi) = \{(10/3, 40)\}$, the mutation of $\delta_0 \times \psi$ to $\Theta' = \{(3,33)\}$ is not explanation improving. However, the equilibrium does not resist a mutation to $\Theta'$: 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 $\psi$. 

OA-9
A mutation that is not explanation improving can only invade when agents discard a parameter that provides the best explanation (here $(10/3, 40)$), which is implausible. 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 $\Theta'$ if

(i) $p_k = (1-\varepsilon)\delta_\Theta + \varepsilon\delta_{\Theta'}$,
(ii) $p(\cdot|\tilde{\Theta}) = \Delta(BR(\Delta(\tilde{\Theta}(\psi))))$ $\forall \tilde{\Theta} \in \{\Theta, \Theta'\}$,
(iii) $\Theta' \subseteq \Theta$ and $\arg\min_{\theta \in \Theta'} H_\psi(Q^*, Q_\theta) = \arg\min_{\theta \in \Theta} H_\psi(Q^*, Q_\theta)$.

A mutation is an evidence-based simplification if the mutated agents reduce the size of their subjective model, eliminating some of the parameters that provide a suboptimal explanation of 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 $(\Theta, \psi)$ resists every evidence-based simplification.

**B.6 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 $\pi$ is a best reply to 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 $\nu$ that is independent of the prior. Suppose $\{a\} = BR(\delta_\theta)$ for some $\theta \in \Theta(p_\Pi)$. Since $\theta^1$ was chosen arbitrarily, suppose without loss of generality that $\theta = \theta^1$. Since $a$ is the unique best reply to $\theta^1$, and $Z^l_m \xrightarrow{a.s.} -\infty$ for all $l \in \{K+1, \ldots, C\}$ by Claim 4, there exists $c < 0$ such that if $(Z^l_m)_{l=2}^K$ is coordinate by coordinate less than $c$, the best reply to the corresponding belief is to play $a$. Consider the events $E_m$ that $(Z^l_m)_{l=2}^K$ is coordinate-wise less than $c$: $E_m = \{Z^l_m \leq c, \forall l \in \{2, \ldots, K\}\}$. As $\frac{Z_m}{\sqrt{m}}$ converges to a normal random variable we have that

$$\lim_{m \to \infty} P[E_m] = \lim_{m \to \infty} P\left[\frac{Z_m}{\sqrt{m}} \leq \frac{c}{\sqrt{m}}\right] = P\left[\tilde{Z} \leq 0\right],$$

where $\tilde{Z}$ is a random variable that is Normally distributed with mean $\tilde{0}$ and covariance matrix $\Sigma$. As $\Sigma$ is positive definite, this distribution admits a strictly positive density and hence $P[\tilde{Z} \leq 0] > 0$. ■
B.7 Channeled attention

To formalize the idea of channeled attention, Gagnon-Bartsch, Rabin, and Schwartzstein (2021) uses the concept of “attention partitions,” which are defined on a finite set $Y$ of outcomes; we will likewise restrict to finite $Y$ in this subsection.

For every partition $\mathcal{Y}$ of $Y$, let

$$
\Theta_\mathcal{Y}(\psi) = \arg\min_{\theta \in \Theta} \sum_{s \in S} \sigma(s) \sum_{\pi \in \Pi} \psi(\pi) \sum_{B \in \mathcal{Y}} Q^*(B|s, \pi(s)) \log(Q_\theta(B|s, \pi(s)))
$$

denote the parameters that minimize the weighted KL divergence from the objective data generating process for the events in $\mathcal{Y}$.

**Definition.** Partition $\mathcal{Y}$ is an *minimal attention partition* induced by the subjective model $\Theta$ if for all $\pi \in \Pi$, $BR(\Delta(\Theta_\mathcal{Y}(\pi))) = BR(\Delta(\Theta(\pi)))$, and there is no coarser partition with the same property.

Note that a minimal attention partition need not be unique, because an outcome that is irrelevant for inference can be pooled with many different outcomes, and the resulting partitions are not ordered by coarseness.

Gagnon-Bartsch, Rabin, and Schwartzstein (GRS) assumes that actions do not influence the distribution over outcomes, i.e., $Q_\theta(\cdot|s, a) = Q_\theta(\cdot|s, a')$ and $Q^*(\cdot|s, a) = Q^*(\cdot|s, a')$ for all $\theta \in \mathcal{H}, a, a' \in A$ and $s \in S$. Under this assumption, whether an attention-improving mutation is triggered is independent of the current equilibrium distribution, as is the inference made by the decision maker in the enlarged model $\Theta'$. We will use our model to explore the impact of relaxing that assumption. GRS also restricts attention to unitary equilibria, and consider only the change from the initial paradigm $\Theta$ to $\Theta \cup \theta^*$, so that the “mutants” add the objective data generating process $Q^*$ to their subjective model. This leads to the following definition:

**Definition.** Unitary equilibrium $(\Theta, \psi)$ is *attentionally stable* if there is $\theta \in \Theta$ such that $Q_\theta(B|s, \pi(s)) = Q^*(B|s, \pi(s))$ for all $s \in S$ and $B \in \mathcal{Y}$.

That is, the equilibrium is attentionally stable if there is a parameter in $\Theta$ that is correct about the distribution over the attention partition.

If actions do not influence the distribution over outcomes, an objectively suboptimal equilibrium resists a mutation to a correctly specified model if and only if it is attentionally stable, so there is no role for evolutionary forces. That an attentionally stable equilibrium is resistant is an immediate consequence of Proposition [1]. If the alternative subjective paradigm does not improve the explanation it is not going to destabilize the equilibrium. Conversely, if the equilibrium is not attentionally stable, the correctly specified model provides a better explanation of the observed outcomes than the old subjective model. Thus since actions do not change the outcome distribution, a correctly specified agent is able to infer the objectively optimal action from the equilibrium observables, so if the equilibrium had a suboptimal payoff it is destabilized by a mutation that adds the correct model to $\Theta$.
When the data generating process is endogenous, only equilibria that are self-confirming resist mutations that include the correct model. With other sorts of mutations, the attention partition matters for which equilibria can resist mutations. For example, the local responses under the attention partition can be different than those when the agent tracks the entire set of observables, so an equilibrium that resists local mutations in one case may not resist in the other. In the next example, an equilibrium that resists local mutations when the agents use an attention partition does not resist local mutations with full attention.

Example 7. [Channeled Attention and Underconfidence] An agent can choose between a skill intensive, \( a_i \), and an unskilled, \( a_u \), task. The agent’s utility is their performance, i.e., \( u(a, y) = y \). They believe that their performance \( \{y_L = 1, y_M = 2, y_H = 3\} \) in each task is a function of their skill \( \kappa \) and a relative difficulty parameter \( \rho \). The prevailing paradigm is certain that the difficulty level is 8/15, and posits that skill is \( \kappa = 1/3 \) or \( \kappa = 0 \), so \( \Theta = \{8/15\} \times \{0, 1/3\} \). Formally:

\[
Q_{(\rho, \kappa)}(y_H|a_u) = \rho - \kappa/2, \quad Q_{(\rho, \kappa)}(y_M|a_u) = 1/6 + \kappa/4 + \rho/2, \quad Q_{(\rho, \kappa)}(y_L|a_u) = 1/6 + \kappa/4 - 3\rho/2
\]

and

\[
Q_{(\rho, \kappa)}(y_H|a_i) = \frac{5}{8}\rho + \kappa, \quad Q_{(\rho, \kappa)}(y_M|a_i) = \frac{5}{8}\rho - \kappa/2, \quad Q_{(\rho, \kappa)}(y_L|a_i) = \frac{5}{8}\rho - \kappa/2.
\]

Because the agent is sure that \( \rho = 8/15 \), \( \Theta_y = \{\{y_H\}, \{y_M, y_L\}\} \) is a minimal attention partition for \( \Theta \), as the agent thinks this is sufficient to infer the value of \( \kappa \), so one minimal attention partition for \( \Theta \) is \( \Theta_y = \{\{y_H\}, \{y_M, y_L\}\} \).

In the objective model, low outcomes are very unlikely with both tasks, and high outcomes are much more likely when performing the skill intensive task, so the objectively optimal action is to choose the skill intensive task:

\[
Q^*(y_H|a_u) = 11/42, \quad Q^*(y_L|a_u) = 1/100 \quad \text{and} \quad Q^*(y_H|a_i) = 4/21, \quad Q^*(y_L|a_i) = 51/700.
\]

The unique Berk-Nash equilibrium is mixed, because \( \Theta(a_i) = \{(8/15, 0)\} \) and \( BR(8/15, 0) = a_u \), while \( \Theta(a_u) = \{(8/15, 1/3)\} \) and \( BR(8/15, 1/3) = a_i \). In equilibrium about half of the population does the unskilled task, sustained by the belief that the probability of high skill is 5/8. We have

\[
U(a_u)(8/15, 1/3) = 2(8/15 - 1/6) + 1/6 + 1/12 + 8/30 = 1.2
\]

\[
U(a_i)(8/15, 1/3) = 2(1/3 + 1/3) + 1/3 - 1/6 = 1.5
\]

\[
U(a_u)(8/15, 0) = 2(8/15) + 1/6 + 8/30 = 1.5
\]

\[
U(a_i)(8/15, 0) = 2(1/3) + 1/3 = 1
\]

so that the agent is indifferent when

\[
1.2\mu((8/15, 1/3)) + 1.5(1 - \mu((8/15, 1/3))) = 1.5\mu((8/15, 1/3)) + (1 - \mu((8/15, 1/3)))
\]
i.e., \( \mu((8/15, 1/3)) = 5/8 \).

By Lemma 6 and Proposition 3, \( \mathcal{M}(\varepsilon) \) for the partition \( \Theta_Y \) is a slight variation of the low skill parameter \( (8/15, \varepsilon) \), so the mutated agents adopt the suboptimal unskilled task and thus eventually die out. Intuitively, the minimal attention partition \( \Theta_Y \) lets the agent ignore the fact that most failures to achieve the high outcome are intermediate and not low. However, for a player who distinguishes between all the outcomes, \( \mathcal{M}(\varepsilon) = (8/15 + \varepsilon, 1/3) \). Since the best reply to these parameters is the optimal action, the equilibrium would not resist mutations: When the agent tracks every outcome, they realize that failures come from high difficulty rather than low skill. ▲

The next example shows that the attention partition can also have the opposite effect.

**Example 8.** An agent can choose between left and right, \( A = \{l, r\} \). The set of outcomes is \( Y = \{l, r\} \times \{u, d\} \). The agent only cares about the first component of the outcome

\[
 u(a, (y_1, y_2)) = \begin{cases} 
 1 & a = l = y_1 \\
 3/2 & a = r = y_1 \\
 0 & a \neq y_1.
\end{cases}
\]

Here \( \mathcal{H} = \{(\theta_{al}, \theta_{au})_{a \in A} \in ([0, 1] \times [0, 1])^4\} \) where \( \theta_{al} \) is the probability of \( y_1 = l \) and \( \theta_{au} \) is the probability of \( y_2 = u \) given that \( a \) has been played. The agent believes that the action does not affect the outcome, i.e., \( \Theta = \{\theta \in \mathcal{H} : \theta_{ul} = \theta_{rl}, \theta_{lu} = \theta_{ru} = \frac{1}{2}\} \). In reality \( \theta^*_{ul} = \frac{3}{7}, \theta^*_{rl} = \frac{4}{7}, \theta^*_{lu} = \frac{1}{10} \) and \( \theta^*_{ru} = \frac{9}{10} \). The unique Berk-Nash equilibrium has \( \pi(l) = \frac{4}{25} \) and \( \pi(r) = \frac{21}{25} \) supported by the Dirac measure on \( (\frac{3}{7}, \frac{4}{7}, \frac{1}{2}, \frac{1}{2}) \). Since the agent (correctly) believes that the second outcome component is payoff irrelevant, the minimal attention partition pools outcomes with the same second component. With this attention partition, local mutations lead the agent to decrease the probability of \( l \) after \( l \) is played. Since the best reply to the updated beliefs is the optimal action \( r \), the equilibrium would not resist mutations. However, for a player who distinguishes between all the outcomes \( \mathcal{M}_{\Theta, \psi}(\varepsilon) \) increases the (payoff irrelevant) probability of \( u \) after \( r \), so mutated agents will eventually die out.

Intuitively, without attention partitions an explanation-improving local mutation does not lead to a change in action, and only mutations that do lead to changes can destabilize an equilibrium. Moreover, when the agent does use a minimal attention partition, they ignore the second component of the outcome, so explanation-improving local mutations lead to a different (and here, better) action. ▲