Which Misperceptions Persist?

Drew Fudenberg and Giacomo Lanzani

First posted version: October 12, 2020
This version: December 22, 2020

Abstract

We use an evolutionary model to determine what sorts of misperceptions are likely to persist. In our model, every period a new generation of agents use their subjective models and the outcome distribution induced by the previous generation to update their beliefs about some underlying parameters, and steady states correspond to Berk-Nash equilibria. Different agents may employ different subjective models, but subjective models that induce better performing actions increase their relative prevalence. To test the stability of steady states, we introduce the possibility that a small fraction of agents enlarge their subjective models in a way that lets them better fit their observations. We characterize which steady states are resistant to “local mutations” to a nearby model, and which are resistant to mutations where agents drop one of the hypotheses they had been imposing on their beliefs, e.g. relaxing independence to allow for positive correlation.

*Department of Economics, MIT. Drew Fudenberg (corresponding author): 77 Massachusetts Avenue, Building E52-418, Cambridge, MA 02139; email: drewf@mit.edu. We thank Pierpaolo Battigalli, Daniel Clark, Roberto Corrao, Kevin He, Pooya Molavi, Stephen Morris, Frank Schilbach, Josh Schwartzstein, Philipp Strack, Dmitry Taubinsky, 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 about the data generating process they face. That is, their prior beliefs may rule out the true state of the world. This misspecification may have different roots: The agent may have a behavioral bias such as overconfidence and correlation neglect, or they may over-simplify a complex environment by only considering simpler but misspecified models that omit some relevant variables or interactions between them.

We propose an evolutionary criterion to evaluate the plausibility and possible persistence of misspecified Bayesian models. In our model, the agents are Bayesians, but the misspecified models prevalent in the society may change in response to mutations and evolutionary pressure. Specifically, beliefs take the form of a distribution over parameters, where each parameter is associated with an action-contingent outcome distribution. We call the support of the prior beliefs the agent’s “subjective model” or “paradigm.” Each agent in a generation observes the actions and outcomes of the previous generation, and estimates the best fitting parameters within the support of their prior distribution.\(^1\) The agents then choose a best reply to a belief that is concentrated on these best-fitting parameters. Proposition 6 shows that this corresponds to the limit of the standard Bayesian updating of beliefs and best replies when the number of observed agents in the previous generation grows to infinity.

We allow different agents to employ different subjective models, but suppose that the distribution of the models used in the population evolves according to evolutionary pressure: The relative frequency of the models that induce better actions increases over time. The monomorphic steady states (i.e. the steady states in which all agents have the same model) of this process coincide with the Berk-Nash equilibria (cf. Esponda and Pouzo, 2016).

To evaluate the stability of these steady states, we introduce the possibility of mutations: From time to time, a small fraction of agents may adopt an expanded subjective model if this allows them to better explain their observations. We then ask whether the use of the expanded model will spread, or whether the existing model is “resistant to mutations.”\(^2\)

Any self-confirming equilibrium is automatically resistant to mutations, as in this case the existing model perfectly fits the observed data. Moreover, not all equilibria with incorrect on-path beliefs are unstable, because the population share of mutants only increases if they do better than agents using the prevailing paradigm. Not all mutations lead to better

\(^1\)Here “best fitting” means the parameters in the support of the prior for which the corresponding distribution minimizes the Kullback-Leibler divergence from the agent’s data.

\(^2\)This bypasses the fact that a Bayesian agent can never assign positive probability to a data generating process that lies outside the support of their prior.
performance; we characterize the ones that do.

We consider two ways that subjective models can expand: “local” expansions to nearby subjective models, and “one-hypothesis relaxations” where the agent drops one of the hypotheses that characterize their subjective model.

First, we show that uniformly strict Berk-Nash equilibria (Fudenberg, Lanzani, and Strack, 2020) are always stable with respect to local mutations. These are pure equilibria in which the action played is the unique best reply to all the parameters that minimize the Kullback-Leibler (henceforth “KL”) divergence from the true data generating process under the equilibrium action. Whether a Berk-Nash equilibrium that is not uniformly strict is stable with respect to local mutations hinges on the concept of the “most improving parameters,” which are the parameters near the KL minimizer that most improve the explanation of the equilibrium distribution. If a small mutation towards these parameters leads agents to change to an action that gives them a higher payoff, the equilibrium is unstable to local mutations. We show that in problems with a “directional” structure, a Berk-Nash equilibrium must be pure and/or self-confirming to be stable.

One-hypothesis relaxations allow larger changes in beliefs. Nevertheless, equilibria may be resistant to these mutations and not to local ones, because the one-hypothesis relaxation may lead to over-adjustment in the direction of the relaxed constraint and thus to overshooting the optimal action. More generally, we provide a static criterion to determine whether an equilibrium is resistant to one-hypothesis relaxations.

We apply these results to study the stability of the equilibria in common examples of misspecifications such as overconfidence, base-rate neglect, correlation neglect, and misperception of a demand curve or tax schedule.

Although we think that these two forms of mutations are natural benchmarks, our model’s combination of Bayesian learning and evolutionary dynamics has a larger 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 relaxation 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 stability with respect to arbitrary sorts
of mutations to alternative models.\textsuperscript{3}

1.1 Related work

The analysis of misspecified learning began with Berk \textsuperscript{(1966)}, which shows that the beliefs of a misspecified agent asymptotically concentrate on the set of models that minimize the KL divergence from the true data generating process when this process is exogenous. In many economic applications, actions and associated signal distributions aren’t fixed but rather change endogenously over time depending on the action taken by the agent, so the agent’s misspecification has implications for what they observe and thus for their long-run beliefs. Arrow and Green \textsuperscript{(1973)} gives the first general framework for this problem, and Nyarko \textsuperscript{(1991)} points out that the combination of misspecification and endogenous observations can lead to cycles. Esponda and Pouzo \textsuperscript{(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 from their observations on the support of their prior.\textsuperscript{4}

Much of the subsequent literature focuses on the learning process for a single misspecified agent. The problem of convergence and stability of equilibria is tackled on a fairly general level by Esponda, Pouzo, and Yamamoto \textsuperscript{(2020)}, Frick, Iijima, and Ishii \textsuperscript{(2020b)}, and Fudenberg, Lanzani, and Strack \textsuperscript{(2020)}. Esponda, Pouzo, and Yamamoto \textsuperscript{(2020)} uses stochastic approximation to establish when the agent’s \textit{action frequency} converges in an environment with finitely many actions. Frick, Iijima, and Ishii \textsuperscript{(2020b)} provides conditions for local and global convergence of the agent’s beliefs without explicitly modelling the agent’s actions. Fudenberg, Lanzani, and Strack \textsuperscript{(2020)} proposes two refinements of Berk-Nash equilibria, uniform Berk-Nash equilibria and uniformly strict Berk-Nash equilibria. It shows that uniform Berk-Nash equilibria are the only possible limit points of the DM’s actions, and that uniformly strict Berk-Nash equilibria are the unique stable equilibria.\textsuperscript{5}

Gagnon-Bartsch, Rabin, and Schwartzstein \textsuperscript{(2020)}, like this paper, formalizes a process that lets agents realize that their model is misspecified. It proposes that agents only pay

\textsuperscript{3}For an axiomatic approach to these “backup” models see Ortoleva \textsuperscript{(2012)}.

\textsuperscript{4}Jehiel \textsuperscript{(2020)} surveys various equilibrium concepts for misspecified agents.

\textsuperscript{5}Earlier papers such as Fudenberg, Romanyuk, and Strack \textsuperscript{(2017)}, Heidhues, Kőszegi, and Strack \textsuperscript{(2018)}, Heidhues, Koszegi, and Strack \textsuperscript{(2020)}, and Molavi \textsuperscript{(2019)} analyze the dynamics of misspecified learning in specific applications or examples. Bohren \textsuperscript{(2016)}, Bohren and Hauser \textsuperscript{(2020)}, and Frick, Iijima, and Ishii \textsuperscript{(2020a)} and He \textsuperscript{(2019)} consider misspecified social learning where all agents have the same form of misspecification. In one of He \textsuperscript{(2019)}’s models, each period a large generation of agents learns from the data generated by the previous generation.
attention to the frequency of the events that they believe have payoff-relevant consequences. If an agent’s misspecified model is wrong about the probability of one of these events, they can switch to a different model that is correct in that respect. Unlike in our model, this switch occurs whether or not it actually leads to higher payoffs. We say more about the difference this makes in Section 5.3. He and Libgober (2020) studies a 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. Frick, Iijima, and Ishii (2020c) compares the asymptotic efficiency of two biased updating procedures. In contrast to our model, their agents always assign a positive probability to the correct state, but fail to perform Bayesian updating upon observing a signal, and the data generating process is exogenous.

The experimental findings of Esponda, Vespa, and Yuksel (2020) support our analysis, even though the details of their design differ from our theoretical model in a number of respects. First, they show that agents who use a misspecified model make small partial adjustments in the face of large, unexplained evidence, as in our analysis of local mutations. Second, they show that agents use their inferences from a particular task or action to adjust the predicted consequences for other actions. This possibility is at the core of why some misspecifications are stable in our setting: A subjective model that is closer to the correct one with respect to the observed equilibrium distribution of actions may induce beliefs about non-equilibrium actions that in turn induce a behavior that performs even worse than the status quo.

At a high level of abstraction, our work is related to papers that study evolutionary selection of behavioral rules or utility functions that then determine actions, such as risk preferences (Robson, 1996b, Robson, 1996a, Dekel and Scotchmer, 1999, Robson and Samuelson, 2019) or time preferences (Robson and Samuelson, 2007, Netzer, 2009). Our paper is also related to the macroeconomics literature on non-Bayesian agents who test their models. For example, in Cho and Kasa (2017) the agent tests whether the model parameters are constant, and in Hong, Stein, and Yu (2007) the agents switch between two univariate models on the basis of a statistical test. The possibility that the agent may reject

---

or expand their subjective model in the face of overwhelming evidence against it resembles the statistical test of stationarity considered in Fudenberg and Kreps (1994).

2 Illustrative Example

Example 1. We consider two versions of the problem of a misspecified monopolist. The monopolist chooses price \( a \) and receives payoff \( u(a, y) = ay = a(i^* - l^*a + \omega) \), where \( i^* \) and \( l^* \) are the unknown intercept and slope of the demand function, and \( \omega \) is a standard normal shock. The monopolist has two actions, \( A = \{2, 10\} \).

The two versions of this example have the same payoff function for the monopolist and the same true demand function, given by \((l^*, i^*) = (4, 42)\); they have different assumptions on the agent's subjective model. In the first version there is a continuum of equilibria, and our criterion shows that an equilibrium is resistant to mutations if and only if it doesn't assign too much \( \epsilon \) probability to the suboptimal action 10. The second version shows that even a unique and isolated equilibrium may not be resistant to mutations.

a. The first version is a slight variation of Example 2.1 of Esponda and Pouzo (2016), where the monopolist 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 with price 2 in \([7/8, 35/36]\), sustained by a Dirac belief on the KL-minimizing parameter \((l, i) = (10/3, 40)\).

In the equilibria in which the optimal price of 2 is charged with probability higher than \( \approx \frac{97}{100} \), the KL minimizer for an \( \epsilon \)-enlarged subjective model is approximately \((10/3, 40 + \epsilon)\). This includes the original equilibrium in Esponda and Pouzo (2016) where the low price is charged with probability 35/36, illustrated in Figure 1. Since this parameter lies above the diagonal indifference curve, it induces 10 as the unique best reply, which yields a lower payoff than the equilibrium action. As we will see, this makes these equilibria resistant to local mutations. In the equilibria in which 2 is charged with probability lower than \( \frac{97}{100} \), the evidence points toward the region in which the subjectively optimal action is 2, which yields a larger payoff. As we will see, this makes these equilibria not resistant to local mutations.

b. Now suppose that the subjectively possible parameter values are \([3/2, 5/2] \times [28, 32]\), as in the example of Nyarko (1991). The unique Berk-Nash equilibrium is mixed, and assigns probability 1/4 to price 2, sustained by a Dirac belief on \((l, i) = (5/2, 30)\), illustrated in Figure
Figure 1: The ellipses are level curves for the KL divergence in the equilibrium of part a where \( \psi(2) = 35/36 \); the green arrow points in the direction of the greatest KL improvement.

Here the binding constraint is \( l \leq 5/2 \), so an agent with an \( \varepsilon \)-enlarged subjective model will mostly adjust the slope upwards, which leads them to choose action 2. Since action 2 performs better than mixing, this Berk-Nash equilibrium will not be resistant to local mutations.\(^7\)

3 The single-agent problem

We will consider learning in a large population of agents. To begin, we introduce the single-agent problem.

\(^7\)In Nyarko’s version of example b, 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.
Figure 2: The ellipses are level curves for the KL divergence in the unique equilibrium of part b; the green arrow points in the direction of the greatest KL improvement.

3.1 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 $y \in Y \subseteq \mathbb{R}^m$. Thus 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 $u : S \times A \times Y \to \mathbb{R}$.

For every subset $X$ of some Euclidean space we let $\mathcal{B}(X)$ denote its relative Borel sigma-algebra, and $\Delta(X)$ denote the set of Borel probability distributions on $X$. 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 $Q^*(\cdot | \cdot) \in \Delta(Y)^{S \times A}$. We denote strategies of the agents, i.e., the maps from signals to actions, by $\Pi = A^S$. The utility function $u : S \times A \times Y \to \mathbb{R}$ of the agent, paired with $Q^*(\cdot | \cdot)$, induces the objective expected utility of strategy $\pi : S \to A$:

$$U^*(\pi) = \sum_{s \in S} \sigma(s) \int_Y u(s, \pi(s), y) dQ^*(y|s, \pi(s)).$$

The expected utility function is extended to mixed strategies $\psi \in \Delta(\Pi)$ in the standard way.
**Subjective models**  As in Esponda and Pouzo (2016), the agent uses parametric models to describe the environment. Formally, there is a compact and convex subset $\mathcal{H}$ of a Euclidean space $\mathbb{R}^k$ with elements $\theta$, where each $\theta$ is 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 $\mu$ with support $\Theta$. A *subjective model* for an agent is the particular subset of parameters $\Theta \subseteq \mathcal{H}$ they consider possible.

We do not require 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 do allow these cases, but our focus is on the case where the agent is misspecified in the sense their prior rules out the true outcome distribution for at least some actions. We let $\mathcal{K}$ denote the collection of compact subsets of $\mathcal{H}$, and make the following technical assumptions:

**Assumption 1.**

(i) The subjective model of the agent belongs to $\mathcal{K}$, and $\theta \mapsto Q_{\theta} (\cdot | s, a)$ is continuous.

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

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. With a slight abuse of notation, 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.

**Preferences and best replies**  The agent’s subjective expected utility function given belief $\mu \in \Delta(\Theta)$ is

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

With a small abuse of notation, we let $U_\theta = U_{\delta_\theta}$ where $\delta_\theta$ is the Dirac measure on $\theta$. 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 $BR(C) = \bigcup_{\mu \in C} BR(\mu)$.

---

8If $Y$ is finite, $\mathcal{H}$ can be set to be the space of all signal and action contingent distributions, but in general it is a strict subset of this space. See Diaconis and Freedman (1986) and Diaconis and Freedman (1990) for arguments in favor of restricting to a finite-dimensional parameter space.
Assumption 2. For every $\pi \in \Pi$, $\Theta \in \mathcal{K}$, and $\mu \in \Delta(\Theta)$, $U^*(\pi)$ and $U_\mu(\pi)$ are finite.

Inference and Kullback-Leibler minimizers Before taking an action, the agent observes a joint distribution $Q_\psi \in \Delta(S \times A \times Y)$ induced by a distribution over strategies $\psi \in \Delta(\Pi)$ played in the previous period:

$$Q_\psi(s, a, B) = \sigma(s)\psi(\{\pi : \pi(s) = a\})Q^*(B|s, a) \quad \forall s \in S, a \in A, B \in \mathcal{B}(Y).$$

Given two distributions over outcomes $Q, Q' \in \Delta(Y)$ we define

$$H(Q, Q') = -\int_{y \in Y} \log q'(y)dQ(y).$$

Note that $-H(Q, Q')$ is the expected log-likelihood of an outcome under subjective distribution $Q'$ when the true distribution is $Q$, so $Q'$ with smaller $H(Q, Q')$ better explain distribution $Q$.\(^9\) The relative entropy between the objective distribution $Q^*$ and the distribution induced by $\theta$ depends both on the signal and strategy distribution, so let

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

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

$$\Theta(\psi) = \arg\min_{\theta \in \Theta} H_{\sigma, \psi}(Q^*, Q_\theta).$$

We call these parameters the KL minimizers for $\psi$. We assume that the posterior over $\Theta$ of the agent after having observed $Q_\psi$ is a probability distribution supported over $\Theta(\psi)$. Proposition \(^\Box\) shows that this describes the limit as the agent observes a larger and larger finite number of individual experiences from the previous period that are drawn from $Q_\psi$.

3.2 Equilibrium Concepts

Here we introduce the static equilibrium concepts that we will work with. As we will see, they correspond to the steady states of our model. Berk-Nash equilibria will be our primary focus. These equilibria require that beliefs are supported on the set of parameters that best explain the observables given the equilibrium distribution over strategies $\psi$.

\(^9\)The Kullback-Leibler (KL) 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 $Q'$.\(^9\)
Definition. A Berk-Nash equilibrium is a $(\Theta, \psi) \in K \times \Delta(\Pi)$ 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 $\psi = \pi$ for some $\pi \in \Pi$.

(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 $\psi = \pi = BR(\mu)$ for every $\mu \in \Delta(\Theta(\psi))$.

Esponda and Pouzo (2016) defines the unitary version of Berk-Nash equilibrium, which requires that each of the equilibrium strategies can be rationalized with the same belief. 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. Non-unitary steady states can only arise if multiple parameters minimize the KL divergence from the equilibrium distribution over outcomes. As noted by Fudenberg, Lanzani, and Strack (2020), such multiple minimizers can arise from symmetry constraints or the use of low-dimensional functional forms. Uniformly strict equilibrium (Fudenberg, Lanzani, and Strack, 2020) requires the equilibrium strategy to be a strict best reply to all of the KL-minimizing parameters; such equilibria are clearly quasi-strict.

A self-confirming equilibrium is a Berk-Nash equilibrium $(\Theta, \psi)$ such that there exists $\theta \in \Theta$ with $Q_{\theta}(|s, \pi(s)) = Q^*(|s, \pi(s))$ 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 true data generating process does, under the equilibrium strategy distribution. Therefore, these equilibria do not trigger a change of paradigm, as confirmed by Corollary 2 below.

4 Evolutionary Dynamics and Steady States

4.1 States and State Evolution

There is a continuum of agents, all with the same utility function $u$. The state of the system at every period $t \in \mathbb{N}$ is a finite-support joint distribution $p \in P := \Delta(K \times \Pi)$ over the subjective models and strategies of the agents. We denote the marginal distributions of $p$ as $p_K$ and $p_\Pi$.

We consider a dynamic process where the state evolves in a way that combines individual Bayesian learning and evolutionary competition between the various subjective models. Our
focus will be on the possible steady states of the system and their stability, but we first introduce our dynamic model.

**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 state $p^t$. 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 evolutionary map $T : P \rightarrow \Delta(K)$ that is continuous, with $	ext{supp } p_K = \text{supp } T(p)$ and payoff monotone (Samuelson and Zhang, 1992), meaning that

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

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

$$p^t_{K+1} = T(p^t)$$

(2)

hold.

These dynamics can be interpreted as the result of biological reproduction or as the result of social learning and imitation.\(^{10}\) Under the biological perspective, payoffs correspond to

---

fitness, and agents whose subjective model induce fitter actions have more offspring. Parents transmit their subjective model—i.e., the support of their prior—but neither their beliefs nor their 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 well 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 information about the performance of the different subjective models as in Björnerstedt and Weibull [1995], Schlag et al. [1998], and Binmore and Samuelson [1997]. Notice that the offspring inherit their subjective model from their parents, but do not inherit how to handle cases where there are multiple KL-minimizing beliefs or multiple best responses to the same belief.\(^{11}\)

4.2 Steady states

We study the joint dynamics of actions and subjective models that satisfy the best reply and evolutionary conditions we introduced above. In particular, we are interested in properties of the rest points of these dynamics.

**Definition.** A **steady state** is a constant solution, i.e. a \( \hat{p} \in P \) such that \( (\hat{p})_{t \in \mathbb{N}_0} \) is a solution. A steady state is **monomorphic** if \( \hat{p}_K = \delta_\Theta \) for some \( \Theta \in K \). It is **unitary** if \( \hat{p}(\cdot | \Theta) \in \Delta(BR(\mu)) \) for some \( \mu \in \Delta(\Theta(\hat{p}_\Pi)) \).

In a steady state, every subjective model that is used by a positive fraction of the agents induces a strategy distribution that performs equally well. The steady state is monomorphic if all agents have the same subjective model. In a monomorphic steady state, agents with the same subjective model can have different beliefs over the KL minimizers unless there is a unique minimizer given the data and their model. A steady state is unitary if all the strategies are best replies to the same belief about the KL-minimizing parameters. This does not require that the distribution over strategies is a point mass, because ties between best replies may be broken in different ways by different agents.

\(^{11}\)If these features were inherited, some Berk-Nash equilibria would not be steady states.
Lemma 1. For all $\Theta \in \mathcal{K}$ and $\psi \in \Delta(\Pi)$, $\delta_{\Theta} \times \psi$ is a monomorphic 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.

This lemma shows that Berk-Nash equilibria are the only rest points of evolutionary competition between subjective models when a unique paradigm has prevailed. Its proof follows readily from unpacking the definitions. 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.\(^\text{12}\)

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

5 Mutations

We now allow for mutations that lead offspring to expand the subjective model they inherited from their parents. The idea is that agents are induced to mutate if enlarging their model lets them find a better explanation for their observations. We suppose that mutant agents enlarge the parameter space of models that they think are possible, and then use their observations from the previous generation to estimate the best fitting parameters.

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

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

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

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

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 collection of solutions $(p_\varepsilon^t)_{t \in \mathbb{N}, \varepsilon \in (0,1)}$, such that $p_\varepsilon^0$ is the $\varepsilon$ mutation of $\delta_{\Theta} \times \psi$ to $\Theta'$, and

$$\lim_{\varepsilon \to 0} \lim_{t \to \infty} (p_\varepsilon^t)_\Pi = \psi.$$  

\(^{12}\)Esponda and Pouzo (2016) assumes a finite $Y$, but this is not needed for the proof of their Theorem 1.
For every given $\varepsilon$, the inner limit gives the long-run strategy distribution following an $\varepsilon$ mutation to $\Theta'$; the outer limit sends the fraction of mutated agents to 0. The equilibrium $(\Theta, \psi)$ is resistant to this mutation if this iterated limit converges back to $\psi$ for some solution that starting from the $\varepsilon$ mutation to $\Theta'$.

An equilibrium can resist a mutation for two distinct reasons. First, the mutated agents may play a strategy that performs worse than the equilibrium, leading to the asymptotic extinction of the mutated agents and a return to the equilibrium behavior. Second, the mutation may not be sufficient to alter the best response to the evidence generated from equilibrium play, so the distribution over strategies does not change.

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

In an explanation-improving $\varepsilon$ mutation, a fraction $\varepsilon$ of agents realize that they could better explain their data with a larger parameter set $\Theta'$. Mutations that generate worse fits to the data can still generate higher payoffs, as in [B.3] in the Appendix, where a point mutation to a misspecified model induces the agent play the optimal strategy. However, we will focus on mutations that strictly enlarge the set of subjective models. For such mutations, only explanation-improving mutations can be successful.

**Proposition 1.** *If the mutation of $\delta_\Theta \times \psi$ to $\Theta' \supseteq \Theta$ is not explanation improving, then $(\Theta, \psi)$ resists mutation to $\Theta'$.*

The proof of this is simple: Because $\Theta' \supseteq \Theta$ and $\Theta'$ is not explanation improving, the best explanations in $\Theta$ are also best explanations in $\Theta'$, and one possible continuation path is for the mutants and conformists to all continue to play the same $\psi$ as before the mutation.

**Corollary 2.** *If $(\Theta, \psi)$ is a self-confirming equilibrium and $\Theta' \supseteq \Theta$, then $(\Theta, \psi)$ resists mutation to $\Theta'$.*

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.

---

13 A more demanding definition would require that long-run play converges back to the equilibrium in every solution. This more demanding definition would rule out all the self-confirming equilibria that are not Nash equilibria, even when the agent is correctly specified.

14 Here we implicitly assume that mutations that do not lead to changes in play are abandoned, so that we do not need to worry about the cumulative effects of multiple mutations. This rules out the possibility of neutral drift, (Kimura, 1983; Binmore and Samuelson, 1999; Traulsen and Hauert, 2009), where the initial mutation keeps a constant $\varepsilon$ share of the population and can then be the springboard for subsequent mutations.
Conversely, if an explanation-improving mutation to $\Theta'$ makes the agent correctly specified, and the map from $\Theta'$ to the observations generated under the equilibrium strategy is injective, i.e., $\Theta'$ is “strongly identified” (Esponda and Pouzo (2016)) then the mutants will perform at least as well as their parents. In this case a strategy will not resist mutation to $\Theta'$ unless it is a Nash equilibrium.

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.

5.1 Local mutations

In an $\varepsilon$ local mutation, a fraction $\varepsilon$ of the agents in the new generation reacts to unexplained evidence by considering a moderately more permissive paradigm: they allow for all the parameter values within $\varepsilon$ of a parameter their parents thought was possible.\footnote{To lighten notation we use $\varepsilon$ in two roles here, but nothing would change if we instead had share $\varepsilon'$ of agents adopt an $\varepsilon''$ mutation.}

Definition (Local Mutations).

(i) The subjective model $\Theta_\varepsilon$ is the $\varepsilon$ expansion of $\Theta$ if $\Theta_\varepsilon = \{\theta' \in \mathcal{H} : \exists \theta \in \Theta, ||\theta - \theta'||_2 \leq \varepsilon\}$. 

(ii) $p$ is the $\varepsilon$ local mutation of a monomorphic steady state $\delta_\Theta \times \psi$ if it is an $\varepsilon$ mutation of $\delta_\Theta \times \psi$ to the $\varepsilon$ expansion of $\Theta$.

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

An equilibrium is resistant to local mutations if, after a mutation that is sufficiently small both in the extent of affected agents and the in the size of the adjustment, the aggregate behavior of the population converges back to the original equilibrium behavior. Local mutations are particularly powerful in selecting among non-pure Berk-Nash equilibria. By definition, in these equilibria multiple strategies are best replies to beliefs about the KL-minimizing parameters. Arbitrarily small enlargements of the subjective model can break this tie, with the mutated agents playing only one of the strategies in the equilibrium distribution.

Proposition 2.

(i) Every uniformly strict Berk-Nash equilibrium is resistant to local mutations.
(ii) Every Berk-Nash equilibrium $\delta_\Theta \times \psi$ in which $\Theta(\psi)$ is in the interior of $\Theta$ is resistant to local mutations.

In a uniformly strict equilibrium, the equilibrium strategy $\pi$ is a strict best response to every parameter that minimizes the KL divergence given that $\pi$ is played. We show that this implies the beliefs of the agents under small mutations are concentrated on a neighborhood where the unique best reply is still $\pi$. Therefore, both mutated and conformist agents continue to play $\pi$ at every period, 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 (2020) that uniformly strict Berk-Nash equilibria have desirable stability properties. In contrast, in a Berk-Nash equilibrium that is not uniformly strict, there may be one parameter $\theta \in \Theta(\pi)$ that does not induce $\pi$ as the unique best reply. If the most improving expansion of the subjective model (a concept we formalize below) is near $\theta$, the mutated agents may start to play a different strategy than $\pi$, which would induce a departure from equilibrium play.

To evaluate the stability of Berk-Nash equilibria that are not uniformly strict, we will use a concept that captures how much enlarging the parameter space in a particular direction improves the explanation of the equilibrium outcome.

Given a monomorphic steady state $\delta_\Theta \times \psi$ and an $\varepsilon$ we define the most improving parameters as

$$M_{\Theta}(\varepsilon) = \arg\min_{\theta \in \Theta_\varepsilon} H_{\sigma,\psi}(Q^*, Q_\theta).$$

These are the parameters that generate the largest decrease in $H$.\textsuperscript{16}

Small mutations induce beliefs that are concentrated on the most improving parameters. If the strategies that are 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 than the original population, the mutated agents will eventually disappear, and the play converges back to the original equilibrium. Let

$$\Pi_{M_{\Theta}} = \limsup_{\varepsilon \to 0} BR(\Delta(M_{\Theta}(\varepsilon)))$$

denote the strategies that are optimal against the the most improving parameters.

\textsuperscript{16}In principle $M_{\Theta}(\varepsilon)$ need not be a singleton, but in all the examples where we analyze local mutations it will be a singleton for sufficiently small $\varepsilon$. 
Proposition 3. Let \((\Theta, \psi)\) be a Berk-Nash equilibrium.

(i) If for every \(\pi \in \Pi_{\Theta \psi}, U^*(\pi) > U^*(\psi),\) then \((\Theta, \psi)\) is not resistant to local mutations.

(ii) If for every \(\pi' \in \Pi_{\Theta \psi}, U^*(\pi') \leq U^*(\psi),\) and \((\Theta, \psi)\) is a quasi-strict equilibrium, \((\Theta, \psi)\) is resistant to local mutations.

To illustrate Proposition 3, we revisit the monopoly pricing problems of Example 1. Some Berk-Nash equilibria in the version from Esponda and Pouzo (2016) are resistant to local mutations, while the equilibrium in the version based on Nyarko (1991) is not. The difference is in whether or not the equilibrium action is a best reply to the most improving parameters. In the Online Appendix, which contains the supporting computations for the examples, we show that in example 1a), in the original equilibrium of Esponda and Pouzo (2016), where \(\psi(2) = 35/36,\) the most improving parameters have an increased intercept, which makes 10 the unique optimal response. This action performs worse than the equilibrium, so by Proposition 3, the equilibrium is resistant to local mutations. In example 1b), though, the most improving parameter makes 2 the unique optimal choice. Action 2 performs better than mixing, so by Proposition 3, this Berk-Nash equilibrium is not resistant to local mutations.

In the previous example, some misspecified beliefs could be stable because the agent’s parametric model led them to incorrectly extrapolate their observations of demand at one price to what demand would be at another. In contrast, if the agent thinks that observations at one price have no information at all about demand at others, explanation-improving mutations do not induce different beliefs about non-equilibrium actions. As the next example illustrates, this can restrict the set of stable beliefs.

Example 2. [Monopoly pricing as a Bandit Problem] Suppose a monopolist faces a single consumer in every period and can choose between a high and a low price \(A = \{a_h, a_l\}.\) The monopolist is uncertain about the probability that the consumer buys the product at any given price: \(H = [0, 1]^2\) with typical element \(\theta = (\theta_{a_h}, \theta_{a_l}),\) where \(\theta_a\) is the probability that the consumer buys when the price is equal to \(a,\) and the true parameter is \(\theta^*.\) As in Rothschild (1974), suppose that the monopolist’s belief is a product measure with support \(\Theta,\) so that the monopolist perceives their choice as a bandit problem.

Let \((\Theta, \psi)\) be a Berk-Nash equilibrium. We say that the agent is overoptimistic (overpessimistic) about the price \(a\) if \(\theta_a > \theta^*_a\) (\(\theta_a < \theta^*_a\)) for all \(\theta \in \Theta(\psi).\) In any Berk-Nash equilibrium in which both prices are charged with positive probability, and is resistant to local mutations, the agent must be either overoptimistic about both prices or overpessimistic about both.

To see why, consider the case in which the objectively optimal action is \(a_l,\) and suppose that the there is a non-pure equilibrium where the monopolist is overpessimistic about only
one price. Subjective optimality of mixing then implies that the monopolist is overoptimistic about the demand at \( a_h \), and overpessimistic about the demand at \( a_l \). The most improving parameters must (weakly) reduce these two biases, and strictly reduce one of them. But then the unique best reply to the new beliefs is to play the optimal action \( a_l \), and, by Proposition 3, the equilibrium is not resistant to local mutations. In contrast, in Example 1a, the agent is overoptimistic about the demand at the high price 10, and overpessimistic at the low price 2, and the equilibrium is resistant to local mutations.

When the subjective model is finite and the perceived data generating process changes smoothly in the parameters there is a convenient way to check the conditions on \( \Pi_M \Theta \) from Proposition 3. For every \( a \in A \) and \( s \in S \), and \( \hat{\theta} \in \mathbb{R}^k \), if \( H(Q^*(\cdot|s,a), Q_{\hat{\theta}}(\cdot|s,a)) \) is finite then \( \theta \mapsto H(Q^*(\cdot|s,a), Q_{\hat{\theta}}(\cdot|s,a)) \) is continuously Gateaux differentiable at \( \hat{\theta} \). Given a strategy distribution \( \psi \) and \( v \in \mathbb{R}^k \), let

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

be the \( \psi \)-weighted directional derivative of the relative entropy \( H \) in direction \( v \) at \( \theta \). Let \( S \) denote the sphere of radius 1 in \( \mathbb{R}^k \) with respect to the \( \| \cdot \|_2 \) norm.

In the case of a finite \( \Theta \), the computations of the strategies played by the mutated agents are greatly simplified. In particular, it is enough to compute the direction in which there is a maximal decrease in the KL-divergence from one of the KL-minimizers, and look at the best replies to the parameters along that direction.

**Lemma 2.** Let \((\Theta, \psi)\) be a Berk-Nash equilibrium. If \( \Theta \) is finite and \( \arg \max_{\theta \in \Theta(s), v \in S \cup \{0\}} D_{\psi}(\theta, v) = (\hat{\theta}, \hat{v}) \) is a singleton then \( \Pi_M \Theta \subseteq \limsup_{\varepsilon \to 0} \arg \max_{\pi \in \Pi} U_{\delta_{\hat{\theta}+\varepsilon \hat{v}}} (\pi) \).

Next, we introduce an environment where small enlargements of the subjective model always induce better strategies.

**Definition.** The environment is directional if we can strictly order the space of actions \((\Pi, >)\) and:

(i) \( \Theta \subset \mathbb{R} \);

(ii) \( U_\theta(\pi) \) is strictly supermodular in \( \pi \) and \( \theta \);

(iii) if \( \theta > \theta' > \theta'' \), then for all \( s \in S \) and \( \pi \in \Pi \) we have

\[
H(Q_\theta(\cdot|s, \pi(s)), Q_{\theta'}(\cdot|s, \pi(s))) < H(Q_\theta(\cdot|s, \pi(s)), Q_{\theta''}(\cdot|s, \pi(s)))
\]
and
\[ H(Q_{\theta^*} (\cdot | s, \pi(s)), Q_{\theta^*} (\cdot | s, \pi(s))) < H(Q_{\theta^*} (\cdot | s, \pi(s)), Q_{\theta} (\cdot | s, \pi(s))); \]

(iv) There is \( \theta^* \in \mathbb{R} \) such that \( Q_{\theta^*} = Q^* \), and either \( \theta \leq \theta^* \) for all \( \theta \in \Theta \) or \( \theta \geq \theta^* \) for all \( \theta \in \Theta \).

In a directional environment the actions and parameters are ordered, and higher parameters make higher actions more preferred. Moreover, the misspecification of the agent is one-sided, in that the parameters they consider possible are either all lower or all higher than the truth.

**Proposition 4.** *In a directional environment, a Berk-Nash equilibrium that is resistant to local mutations is either pure or self-confirming (or both).*

To see why the proposition holds, suppose that the true parameter is lower than \( \theta = \min\{\theta \in \Theta\} \). Then directionality of the environment implies that the most improving parameter is \( \theta - \varepsilon \), since lower parameters can better match the true distribution. Strict supermodularity implies that the action selected will be lower as well, and, because we are looking at a local change, such a lower action performs strictly better under the true low parameter \( \theta^* \), so the equilibrium is not resistant to local mutations.

### 5.2 One-hypothesis Mutations

In this section, we consider agents whose subjective model is described by a finite collection of hypotheses about the underlying parameter. These hypotheses are expressed in the form of quantitative statements. For example, they may restrict the set of possible values for one of the dimensions of the parameter, as in the case of an overconfident agent who is sure that their skill is higher than some threshold. Alternatively, they can come in 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. We assume that a mutated agent does not completely discard the subjective model they inherited from their parents, but instead drops one of their inherited hypotheses in a way that lets them better explain their data.

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

**Definition** (One-Hypothesis Mutations).
(i) The subjective model $\Theta^l$ is a one-hypothesis relaxation of $\Theta(F)$ in hypothesis $l \in \{1, \ldots, m\}$ if $\Theta^l = \{ \theta \in H : f_i(\theta) \geq 0, \forall f_i \in F \setminus \{f_l\} \}$.

(ii) $\bar{p}$ is a one-hypothesis $\varepsilon$ mutation of a monomorphic steady state $\delta_\Theta \times \psi$ if it is an $\varepsilon$ mutation to some one-hypothesis relaxation of $\Theta$. 17

One-hypothesis relaxations capture the idea that paradigm changes that affect only one dimension of the model are much more likely than adjustments that involve multiple aspects of the model at once. We say that a Berk-Nash equilibrium $(\Theta(F), \psi)$ is resistant to one-hypothesis mutations if it resists every one-hypothesis $\varepsilon$ mutation for sufficiently small $\varepsilon$. 18

Given a monomorphic steady state $p = \delta_\Theta \times \psi$ the $l$-agnostic KL minimizers are $P_l(p) := \arg\min_{\theta \in \Theta^l} H_{\sigma,p}(Q^*, Q_\theta)$, and $\Pi_{p,l} = BR(\Delta(P_l(p)))$ denotes the set of best replies after the relaxation of hypothesis $l$.

Proposition 5. Let $(\Theta(F), \psi)$ be a Berk-Nash equilibrium.

(i) If for some $l \in \{1, \ldots, k\}$, $U^*(\pi) > U^*(\psi)$ for every $\pi \in \Pi_{p,l}$, then $(\Theta, \psi)$ is not resistant to one-hypothesis mutations.

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

In the monopoly pricing problem of Example 1(a) the level curve for the KL divergence from the true parameter has a downward slope at equilibria in which $\psi(2) < 35/36$. Therefore, a one-hypothesis mutation that allows for larger intercepts leads agents to play the objectively optimal price 2. Thus such equilibria are not stable to one-hypothesis mutations.

The next example shows how relaxing one hypothesis while maintaining the others can lead to beliefs that induce overshooting of the optimal action.

Example 3. [Non-linear taxation as in Esponda and Pouzo, 2016] An agent chooses effort $a \in A = \{3, 4, 5\}$ at cost $c(a) = 2a/3$ and obtains income $z = a + \omega$, where $p(\omega = 1) = p(\omega = 0) = p(\omega = -1) = \frac{1}{3}$ is a zero-mean shock. The agent pays taxes $x = \tau^*(z)$, where $\tau^*$ is the following nonlinear tax schedule:

$$
\tau^*(z) = \begin{cases} 
    z/6, & \text{if } z \leq 5 \\
    21/12, & \text{if } z = 6 
\end{cases}
$$

17 Note that here, unlike with local mutations, $\varepsilon$ plays a single role.

18 There may be multiple sets of hypotheses that determine the same set $\Theta$, and which one the agent uses will determine which equilibria are resistant to one-hypothesis mutations.
Thus the tax schedule is linear, except for the highest income bracket that is instead heavily taxed. The agent obtains payoff $u(a, (z, x)) = z - x - c(a)$, so the optimal action is 4.

The agent observes $y = (z, x)$, and the parameters are $(\theta_1, \theta_2) \in \mathbb{R}^2$, where $q_\theta(z, x|a) = p(z-a)\phi(x|z)$, and $\phi(x|z)$ is a normal density with mean $\theta_1 z + \theta_2 z^2$ and variance $z + z^2$.

The agent believes that the tax schedule is linear, i.e. $\Theta = \mathbb{R} \times \{0\}$. Given any action $a$, the KL-minimizing parameter is given by $\Theta(a) = \left(\mathbb{E}\left[\frac{\tau(a + \omega)}{a + \omega}\right], 0\right)$. Here the agent believes that the expected marginal rate is equal to the true average rate. Since the actual tax schedule is progressive, this leads the agent to exert too much effort: the unique pure Berk-Nash equilibrium has $\pi = 5$, with a Dirac belief on $(0.2, 0)$.

An agent with a one-hypothesis mutation adopts the subjective model $\Theta^2 = \mathbb{R} \times \mathbb{R}_+$, so the KL-minimizing parameters solve:

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

Since the actual tax schedule is progressive, the mutated agent chooses an action that is lower than $\pi$, action 3. Here, even though the objectively optimal action 4 is lower than the equilibrium action 5, the mutated agent overshoots the optimum. By assessing the convexity of the tax schedule at the equilibrium action $\pi$, the mutated agent concludes that $\theta_2$ is so high that a very low effort is called for, and this low effort induces a payoff lower than in equilibrium. So by Proposition 5 the equilibrium is resistant to one-hypothesis mutations.

In the next example we consider 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 preferences of the buyer and the seller for the good are always aligned (modulo a constant shift). In this case, the equilibrium is resistant to one-hypothesis mutations, because the buyer never bids a high price, and so cannot discover the correlation between high bids and high values. However, if we reduce the correlation between the preferences of the market participants by adding a noise term, one-hypothesis mutations can lead the buyer to better actions because they sometimes bid high.

Example 4. [Additive Lemons and Cursed Equilibrium] The agent, a buyer whose value for an object is $v = \omega + 3 + s$, faces a seller who owns the object and values it at $\omega$. The mechanism used is a double auction with price at the buyer’s bid, so the seller sets their bid $x$ equal to their value. The value $\omega$ is good, $\omega = 3$, with probability $1/3$, bad, $\omega = 2$, with probability $1/2$ and very bad, $\omega = 1$, with probability $1/6$, so the objectively optimal strategy is to bid 3 for any $s \in \{-1, 0, 1\}$. The value is observed only if the transaction is concluded,
so the outcome is the pair \( y = (\tilde{\omega}, x) \in (\Omega \cup \{\#\}) \times X \), where \( \tilde{\omega} = \omega \) if the transaction is concluded, i.e., \( a \geq x \), and \( \tilde{\omega} = \# \) if the transaction fails, i.e., \( a < x \).

Here \( \theta = (p_1, p_2, p_3, F(1|1), F(2|1), F(1|2), F(2|2), F(1|3), F(2|3)) \), where \( p_i \) is the probability that the seller’s asking price is equal to \( i \) and \( F(i|j) \) is the probability that the value is lower or equal than \( i \) given that the seller has asked price \( j \). So \( \mathcal{H} \) 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 all \( i \in \{1, 2, 3\} \). The true parameter is

\[
\theta^* = \begin{cases}
  p_1^* = 1/6, p_2^* = 1/2, p_3^* = 1/3, \\
  F(1|1) = 1, F(2|1) = 1, \\
  F(1|2) = 0, F(2|2) = 1, \\
  F(1|3) = 0, F(2|3) = 0.
\end{cases}
\]

However, the agent believes that the price and the value are independent:

\[
\Theta = \left\{ \theta \in \mathcal{H} : \begin{array}{c}
  F(1|1) = F(1|2), \\
  F(1|2) = F(1|3), \\
  F(2|1) = F(2|2), \\
  F(2|2) = F(2|3).
\end{array} \right\}
\]

Suppose first that the noise term is degenerate, with \( s = 0 \). The strategy \( a = 2 + s = 2 \) is a Berk-Nash equilibrium. Indeed, the KL-minimizing parameter is an independent joint probability distribution that is correct about the distribution of seller bids. However, given that 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 \( \{(1, \frac{1}{6}); (2, \frac{5}{6}); (3, 0)\} \) assigns probability 0 to the object being good.

This equilibrium is resistant to one-hypothesis mutations. An ask of 3 is never accepted in this equilibrium, so a one-dimensional mutation is explanation improving only if it involves the relation between the conditional distribution of the value given asks 1 and 2. Mutations of this form either do not induce a different strategy, or they induce the agent to bid 1 after some signal, so the mutated agents will obtain a payoff weakly lower than the equilibrium one.

Now suppose that the signal \( s \) is a mean-zero shock uniformly distributed over \( \{-1, +1\} \) and is independent of \( \omega \). The strategy \( a(-1) = 2 \) and \( a(1) = 3 \) 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, \frac{1}{2}); (2, \frac{2}{3}); (3, \frac{1}{2}) \} is too pessimistic. This pessimism justifies the (objectively
suboptimal) bid of price 2 after signal \( s = -1 \).

This equilibrium is uniformly strict, so, by Proposition 2, it is resistant to local muta-
tions. However, it is not resistant to one-hypothesis mutations. Consider 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, which leads to the subjective model:

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

Agents with this model have a posterior concentrated on

\[
\hat{\theta} = \left\{ \begin{array}{l}
p_1 = 1/6, p_2 = 1/2, p_3 = 1/3, \\
F(1|1) = F(1|2) = F(1|3) = 1/6, \\
F(2|1) = F(2|2) = 1, \\
F(2|3) = 2/3.
\end{array} \right\}
\]

Since \( BR(\hat{\theta}) = \{3\} \), by Proposition 3 the equilibrium is not resistant to one-hypothesis
mutations. The intuition is that noise that disentangles the action and outcome distributions
makes it easier to spot errors in the subjective model and find better strategies.

5.3 Attention-channeled mutations

The idea of the “minimal sufficient attentional strategy” in Gagnon-Bartsch, Rabin, and
Schwartzstein (2020) is that an agent who is certain that the true parameter is in \( \Theta \) will only
pay attention to the frequency of the events that help distinguish between parameters in \( \Theta \)
that induce different best replies.\(^{19}\) They motivate this assumption with reference to several
examples of persistent sub-optimization, such as Indonesian seaweed farmers who persistently

\(^{19}\)The paper considers other ways of allocating attention, but focuses on this one.
fail to optimize along a dimension (pod size) they wrongly ignore despite being exposed to rich data from which they could learn if they paid attention (Hanna, Mullainathan, and Schwartzstein 2014).

To formalize the idea of channeled attention, they use 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

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

**Definition** (Minimal Attention Partition). 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.\(^{20}\)

In the next example, we replicate the analysis of sections 5.1, replacing $\Theta(\psi)$ with $\Theta_\mathcal{Y}(\psi)$ where $\mathcal{Y}$ is a minimal attention partition.

**Example 5.** [Channeled Attention and Overconfidence] An agent can choose between a hard and easy task, $A = \{a_h, a_e\}$. 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 $\rho$.\(^{21}\) The agent is sure that the difficulty is 1/5, and believes that their skill is either high, $\kappa = 1/3$, or low $\kappa = 0$ and so $\Theta = \{1/5\} \times \{0, 1/3\}$.

The agent’s utility is their performance, i.e., $u(a, y) = y$. In reality, low outcomes are more unlikely, with $Q^*(y_H|a_h) = 11/42, Q^*(y_L|a_h) = 1/100$ and $Q^*(y_H|a_e) = 4/21, Q^*(y_L|a_e) = 51/700$, so the objectively optimal action is $a_e$.

A minimal attention partition for $\Theta$ is $\Theta_\mathcal{Y} = \{\{y_H\}, \{y_M, y_L\}\}$: The agent does not need to keep track of the relative frequencies of intermediate and low outcomes, since they believe

\(^{20}\)Gagnon-Bartsch, Rabin, and Schwartzstein (2020) imposes the additional assumption that the action does 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 $\Theta'$. The paper focuses on the case $\Theta' = \Theta \cup \{\theta\}$, where $Q^*(\cdot|\cdot) = Q_\theta(\cdot|\cdot)$, so that the new model adds the correct data generating process.

\(^{21}\)In particular, $Q_{(\rho, \kappa)}(y_H|a_h) = 1/3 + \rho - \kappa/2, Q_{(\rho, \kappa)}(y_M|a_h) = 1/3 + \kappa/4 + \rho/2, Q_{(\rho, \kappa)}(y_L|a_h) = 1/3 + \kappa/4 - 3\rho/2$, and $Q_{(\rho, \kappa)}(y_H|a_e) = 1/3 + \kappa, Q_{(\rho, \kappa)}(y_M|a_e) = 1/3 - \kappa/2, Q_{(\rho, \kappa)}(y_L|a_e) = 1/3 - \kappa/2$. 

24
the fraction of high outcomes identifies the true parameter. Here the unique Berk-Nash equilibrium is \( \psi(h) \approx 0.56 \), sustained by the belief \( \mu((1/5, 0)) = 1/3, \mu((1/5, 1/3)) = 2/3 \).

The most improving parameter for the partition \( \Theta_Y \) is \( M(\varepsilon) = (1/5, \varepsilon) \): The evidence suggests a revision of the parameter with a low skill towards a higher skill. Since the best reply to these parameters is the suboptimal action \( a_h \), the mutated agents will eventually die out. Intuitively, the minimal attention partition \( \Theta_Y \) lets the agent ignore the fact that most of the failures to achieve the high outcome are intermediate and not low.

However, the most improving parameter for a player who distinguishes between all the outcomes has a higher difficulty: \( M(\varepsilon) = (1/5 + \varepsilon, 1/3) \). Since the best reply to these parameters is the optimal action \( a_e \), the equilibrium would not resist mutations: When the agent tracks all of the outcomes, they realize that failures come from the high difficulty of the hard task rather than low skill.

This example shows how the attention partition matters for which equilibria can resist local mutations, because the most improving parameters for the events in the partition and the most improving parameters for the entire set of observables can differ. If the best replies to these parameters are different, an equilibrium that is stable in one case need not be stable in the other. In the example above, an equilibrium that resists local attention-partition mutations does not resist more general local mutations. Example 8 in the Online Appendix shows that the attention partition can also have the opposite effect.

6 Large Finite Data Sets

Our evolutionary model is deterministic because each generation of agents observes the deterministic outcome of an infinite number of individual experiences. Here we show that this process emerges as the limit of observing large finite data sets.\(^{22}\) Suppose that each agent with the subjective model \( \Theta \) has a prior \( \mu_\Theta \) with support \( \Theta \). Let \( \Pi_1 \in \Delta(\Pi) \) be the distribution over strategies at time \( t \), and suppose that an agent born in period \( t + 1 \) observes the individual experiences of \( n \) randomly drawn agents from the previous generation, with draws independent across the period \( t + 1 \) agents. Then each agent updates their prior using Bayes’ rule and chooses a best reply to their posterior. Denote the distribution over strategies induced by this procedure by \( \psi_n(\Theta, \Pi_1) \).

\(^{22}\)This is only one way to provide a foundation for our model. We conjecture that the same steady states would be asymptotic limits if a single agent acted each period, as in He (2019) or Bohren and Hauser (2020).
Let $\mathcal{E}(\theta, \psi)$ be the set of parameters that are indistinguishable from $\theta$ if the distribution over strategies is $\psi$, i.e., those parameters $\theta'$ such that $Q_\theta(\cdot | \cdot, \cdot) = Q_{\theta'}(\cdot | \cdot, \cdot)$ almost surely under $Q_\psi$.

**Proposition 6.** If either $BR(\Delta(\Theta(p_\Pi)))$ is a singleton, or $\Theta$ is finite and
\[
\forall \pi, \pi' \in \Pi, \pi \neq \pi', \exists \theta \in \Theta(p_\Pi), U_{\mu_\theta}(\cdot | \mathcal{E}(\theta, p_\Pi))(\pi) \neq U_{\mu_{\theta'}}(\cdot | \mathcal{E}(\theta, p_\Pi))(\pi'),
\]
then $\lim_{n \to \infty} \psi_n(\Theta, p_\Pi)$ exists, and is in $\Delta(BR(\Delta(\Theta(p_\Pi))))$.

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. We do not think this restriction is necessary, but it simplifies the proof considerably. It is not needed if the best reply function is continuous in beliefs, as in our extension of Section 7. With that proviso, the result confirms the informal motivation for condition (1) that we proposed above.

The proposition shows that if newborn agents observe a sufficiently large number of individual experiences from the previous generation, the aggregate distribution of strategies is an element of $\Delta(BR(\Delta(\Theta(p_\Pi))))$. Proposition holds under the assumption that for every two strategies there is a class of parameters that are indistinguishable under the equilibrium distribution and make one of the strategies strictly preferred to the other.

To prove this result, we use an argument similar that of Berk (1966) to show that the probability assigned to the models that do not minimize the KL divergence goes to 0. We then prove that although beliefs may not converge, their distribution does, and 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 as well.

---

23The assumption is needed to guarantee that the state converges because the best-reply correspondence is not continuous. We could drop it by considering smooth best replies instead of exact utility maximization. It is clear that if two KL-minimizing parameters are indistinguishable given the strategy distribution, $\lim_{n \to \infty} \psi_n(\Theta, p_\Pi)$ may depend on the prior, because even with a very large data set the agents will be not able to tell the two parameters apart, so their relative probability under the limit belief will be completely determined by the prior. In Section B.3.7 of the Online Appendix, we show that if in equilibrium the KL-minimizing parameters are identified, as in our examples, the limit belief will be independent of the prior.
7 Infinitely Many Strategies

So far we have assumed there is a finite number of strategies. However, in some applications, either the grid of available actions is very fine, or the set of signals is very fine, and it is more convenient to analyze the problem using a continuum approximation. As we now explain, our analysis can be usefully applied to the continuum environment as well, and it has essentially the same conclusions there.

For simplicity, we state this extension for the case of real-valued actions and a convex $\Theta$, which covers many of the applications and examples in the mispecified learning literature.

Assumption 3. (i) The set of actions $A$ is a compact subset of $\mathbb{R}$, with interior $A^\circ$ (ii) The set of signals $S$ is a subset of a Euclidean space, endowed with a full-support, objective Borel probability measure $\sigma$. (iii) The utility function $U^*$ is continuously differentiable in actions and signals. (iv) $\Theta$ is convex, $BR(\theta)$ is a singleton for every $\theta \in \Theta$, and $Q_\theta$ is continuous in $(a, s)$ for all $\theta \in \Theta$.

The cardinality of the action space makes no real difference for the analysis of one-hypothesis mutations, and those results extend more or less immediately. With local mutations, though, there is a key difference: with any fixed set of actions a vanishingly small $\varepsilon$ is eventually smaller than the “gap” between the actions, and this is not the case when the action space is a connected interval in $\mathbb{R}$. With a continuum of actions, 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, which is not the case with a finite action set. As we show below, this allows local mutations to invade some uniformly strict equilibria in settings with a continuum of actions. However, 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.$^{24}$

To prove these claims, we endow the set of strategies with the $L_1$ norm, i.e., we say that $\pi_n \to \pi$ if and only if $\int_S |\pi_n(s) - \pi(s)|\sigma(ds) \to 0$. The stability of an equilibrium in this setting will rely on two key objects: the most improving parameters introduced in Section 5.1 and the (objective) indirect utility function of the agent, which is $V(\theta) = U^*(BR(\theta))$.  

$^{24}$Convergence here means convergence in the Hausdorff metric with respect to the usual metric on $A \subseteq \mathbb{R}$. 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.
We assume that \( V \) is continuously Gateaux differentiable. If \( \Theta(\psi) \) is a singleton and \( M_{\Theta}(\varepsilon) \) is a singleton for sufficiently small \( \varepsilon \), let \( V'(M_{\Theta}(\varepsilon), \psi) = \liminf_{\varepsilon \to 0} \frac{V(M_{\Theta}(\varepsilon)) - V(\Theta(\psi))}{\varepsilon} \) be the derivative of \( V \) in the direction of the most improving parameter.

**Proposition 7.** Let \((\Theta, \psi)\) be a Berk-Nash equilibrium such that \( \Theta(\psi) \) is a singleton and \( M_{\Theta}(\varepsilon) \) is a singleton for sufficiently small \( \varepsilon \). If \( V'(M_{\Theta}, \psi) > 0 \) then \((\Theta, \psi)\) is not resistant to local mutations.

**Example 6.** [Regression to the Mean] An instructor observes the initial performance \( s \in \mathbb{R} \) of a student and decides whether to praise them, \( a = a_r \), or criticize them, \( a = a_c \). Then the student performs again, and the instructor observes their performance \( y \). The instructor’s utility is

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

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

The instructor believes that \( y = \theta_0 s + \theta_a + \eta \), where \( \eta \) is a standard normal, \( \theta_0 \) is the perceived correlation between performance in the two periods, and \( \theta_a \) is the perceived influence of action \( a \) on performance. Suppose that we start with a population of instructors that are certain that \( \theta_0 = 1 \), so that modulo their action, second period performance is centered around the first period one. More precisely, suppose that \( \Theta = \{1\} \times [-K, +K]^2 \) where \( K \) is a large positive number. Esponda and Pouzo (2016) shows that in the unique Berk-Nash equilibrium the instructor overcriticizes: there is a unique threshold \( T \) such that the instructor criticizes if and only if performance is below the threshold \( T = (\theta_{a_r}(T) - \theta_{a_c}(T))/k > 0 \) where \( \theta_{a_r}(T, \theta_0) = E[y - \theta_0 s|s < T] > 0 \) and \( \theta_{a_c}(T, \theta_0) = E[y - \theta_0 s|s > T] < 0 \). Since \( \theta_{a_r}(T, \theta_0) \) and \( \theta_{a_c}(T, \theta_0) \) are respectively increasing and decreasing in \( \theta_0 \), the most improving parameter for a sufficiently small \( \varepsilon \) is \((1, \theta_{a_r}(T), \theta_{a_c}(T)) + (v_0, v_r, v_c)\) with \( v_0 < 0, v_r \geq 0, v_c \leq 0 \). This encodes a lower correlation between the two period outcomes, a higher effectiveness of praise, and a lower effectiveness of criticism. Since \( V'(1, \theta_{a_r}(T), \theta_{a_c}(T)) + (v_0, v_r, v_c)) = kT > 0 \), by Proposition 7, the equilibrium is not resistant to local mutations.

In the Online Appendix, we show there is a sequence of approximating discrete-signal problems with unique Berk-Nash equilibria, where the instructor criticizes performances below \( T \) and praises performance above \( T \), and each equilibrium is also not resistant to local mutations.
Definition. Let $S$ be a singleton. We say that a pure Berk Nash equilibrium $(\hat{a}, \Theta)$ is an attractor if $\hat{a} \in A^o$ and there is an $\varepsilon > 0$ such $||a - \hat{a}|| \leq \varepsilon$ implies $(BR(\Theta(a)) - \hat{a})(a - \hat{a}) < 0$.

A Berk-Nash equilibrium is an attractor if slightly changing the action in one direction induces a KL-minimizer that induces a best reply in the opposite direction. In all the continuum of actions versions of the examples of this paper with an interior Berk-Nash equilibrium, this equilibrium is an attractor. The condition can fail for finite $\Theta$ and if multiple actions induce the same payoff for all the parameters. 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 $||A_n - A|| \rightarrow 0$

Proposition 8. Suppose that $\hat{a}$ is the unique Berk-Nash equilibrium. Suppose further that it is an attractor that satisfies the assumption of Proposition 7. Then for every sufficiently small $\varepsilon$, there is a sequence $(A_n)_{n \in \mathbb{N}}$ that approximates $A$, and a sequence $(\pi_n)_{n \in \mathbb{N}}$ such that $\pi_n \rightarrow \delta_{\hat{a}}$, and $\pi_n$ is a Berk Nash equilibrium of the environment with finite actions $A_n$ that is not resistant to an $\varepsilon$ expansion of $\Theta$.

8 Conclusion

We have used the idea of resistance to mutations to help understand what sorts of misspecification are likely to persist for a long time. Mutations that lead to a better but imperfect fit in a statistical sense can lead to lower payoffs. If this is the case for all of the plausible mutations at a given equilibrium, we say that the equilibrium is resistant to mutations.

We considered two sorts of mutations: local mutations that consider all parameters that are 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 are resistant to mutations. Local mutations are most 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.

Our model provides a way to obtain precise results about the stability of some canonical examples of misspecification. More generally, even outside of this precise model, we expect that some broadenings of perspective will lead to lower payoffs and thus will be abandoned. Our approach highlights that whether this happens rests on the way the broadened model interprets the data that is generated in equilibrium. Since this data often does not fully
reveal the consequences of non-equilibrium actions, the inferences drawn with the more general model may lead to worse actions. As we point out in our lemons example, noise that disentangles the action and outcome distributions makes it easier for agents to discover models that lead them to take better actions.

Other plausible forms of mutation may have yet other implications. More work is needed to determine what sorts of innovations or mutations seem most empirically relevant.

A Appendix

Proof of Lemma 1. If $\delta \Theta \hat{\psi}$ is a steady state, then by equation (1), $\psi \in \Delta(BR(\Delta(\Theta(\psi))))$, so that for every $\pi \in \text{supp} \psi$ 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 if we can choose $\mu = \mu_\pi$ for all $\pi \in \text{supp} \psi$, so that the equilibrium is unitary as well.

Conversely, if $(\Theta, \psi)$ is a Berk-Nash equilibrium, for every $\pi \in \text{supp} \psi$ 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 every $t \in \mathbb{N}$ satisfies equations (1) and equation (2), so $\delta \Theta \hat{\psi}$ is a steady state. The equilibrium is unitary if and only if we can choose $\mu = \mu_\pi$ for all $\pi \in \text{supp} \psi$, so that the steady-state is unitary as well.

Proof of Proposition 1. We show that for every $\varepsilon \in (0, 1)$ there exists a solution $(p_t)_{t \in \mathbb{N}_0}$ with $p^0 = p_\varepsilon$ and $\lim_{t \to \infty} p^t_\Pi = \psi$, where $p_\varepsilon$ is the $\varepsilon$ mutation of $\delta \Theta \times \psi$ to $\Theta'$.

Define $p_t$ by $p^t_\Pi(\Theta) = 1 - \varepsilon, p^t_\Pi(\Theta') = \varepsilon$, and $p^t_\Pi(\cdot | \Theta) = \psi = p^t_\Pi(\cdot | \Theta')$ for all $t \in \mathbb{N}_0$. By Lemma 1, $\psi \in \Delta(BR(\Delta(\Theta(\psi))))$. Since $\Theta' \supseteq \Theta$ is not explanation improving with respect to $\delta \Theta \times \psi$, $\Theta(\psi) \subseteq \Theta'(\psi)$, and $\psi \in \Delta(BR(\Delta(\Theta(\psi)))) \subseteq \Delta(BR(\Delta(\Theta'(\psi))))$, so that equation (1) is satisfied.

Moreover, equation (2) is satisfied at every period and for every evolutionary map because the distributions of strategies contingent to the two models are the same at every period.

Lemma 3. $\Theta(\cdot)$ is a nonempty, compact-valued upper hemicontinuous correspondence.

The proof of this lemma is in the Online Appendix.

Proof of Proposition 2. (i) Let $(\Theta, \delta_\pi)$ be a uniformly strict Berk-Nash equilibrium, and define $\Theta(\pi, \varepsilon) := \{\theta \in \mathcal{H} : \exists \theta' \in \Theta(\pi), ||\theta - \theta'||_2 \leq \varepsilon\}$. By Lemma 3, $\Theta(\pi)$ is compact, and by the triangle inequality so is $\Theta(\pi, \varepsilon)$. Let

$$G(\varepsilon) = \min_{\pi' \in \Pi, \{\pi\}} \min_{\mu \in \Delta(\Theta(\pi, \varepsilon))} (U_\mu(\pi) - U_\mu(\pi'))$$.
Because $\Pi$ is finite, $U$ is continuous in $\mu$, and $\varepsilon \mapsto \Theta(\pi, \varepsilon)$ is a continuous and compact-valued correspondence, $G$ is continuous by the Maximum Theorem. And since $\pi$ is a uniformly strict Berk-Nash equilibrium, $G(0) > 0$, and there is an $\hat{\varepsilon}$ such that if $\varepsilon < \hat{\varepsilon}$, $G(\varepsilon) > 0$.

Now we show that there is an $\varepsilon' \in (0, \hat{\varepsilon})$ such that if $\Theta'$ is an $\varepsilon < \varepsilon'$ local expansion of $\Theta$, then $\Theta'(\pi) \subseteq \Theta(\pi, \varepsilon)$. This is clearly true if there exists $\theta^* \in \Theta'(\pi)$ with $Q_{\theta^*}(\cdot | s, \pi(s)) = Q^*(\cdot | s, \pi(s))$ for all $s \in S$. By way of contradiction, let $(\Theta'_{n})_{n \in \mathbb{N}}$ be a sequence such that $\Theta'_{n}$ is an $\varepsilon_{n}$ local expansion of $\Theta$ with $\varepsilon_{n} \downarrow 0$, and $\theta_{n} \in \Theta'_{n}(\pi) \setminus \Theta(\pi, \varepsilon)$. Since $(\theta_{n})_{n \in \mathbb{N}}$ is a sequence in the compact $\Theta'_{1}$, it has an accumulation point $\theta \in \Theta$. If $\theta \notin \Theta(\pi)$ this means that $(\theta_{n})_{n \in \mathbb{N}}$ is eventually in $\Theta(\pi, \varepsilon)$, a contradiction. If $\theta \notin \Theta(\pi)$, then since $H$ is lower semi-continuous in its second argument, (see, e.g., Chapter 1 of Liese and Vajda, 1987) eventually

$$
\sum_{s \in S} \sigma(s) H (Q^*(\cdot | s, \pi(s)), Q_{\theta_{n}}(\cdot | s, \pi(s))) > \min_{\theta \in \Theta} \sum_{s \in S} \sigma(s) H (Q^*(\cdot | s, \pi(s)), Q_{\theta}(\cdot | s, \pi(s))) \\
\geq \min_{\theta \in \Theta'} \sum_{s \in S} \sigma(s) H (Q^*(\cdot | s, \pi(s)), Q_{\theta}(\cdot | s, \pi(s)))
$$

a contradiction.

Finally, let $p_{\varepsilon}$ be an explanation-improving $\varepsilon$ 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 for every solution $(p')_{t}$ with $p^{0} = p_{\varepsilon}, p^{t}_{\Pi} = \pi$, concluding the proof of the statement. For the initial step, note that since $\varepsilon < \varepsilon' \leq \hat{\varepsilon}$, $\Theta'(\pi) \subseteq \Theta(\pi, \varepsilon)$. But then $p(\cdot | \Theta') \in \Delta(BR(\Delta(\Theta'(\pi))) \subseteq \Delta(BR(\Delta(\Theta(\pi, \varepsilon)))) = \{ \pi \}$, where the last equality follows from $G(\hat{\varepsilon}) > 0$. Moreover, since $(\Theta, \delta_{\pi})$ is a uniformly strict Berk-Nash equilibrium, $p_{\varepsilon}(\cdot | \Theta') = \{ \pi \}$ as well, concluding the base step. Suppose the statement is true for some $t \in \mathbb{N}_{0}$. Since $\varepsilon < \varepsilon' \leq \hat{\varepsilon}$ we have $\Theta'(p^{t}_{\Pi}) = \Theta'(\pi) \subseteq \Theta(\pi, \varepsilon)$, and by definition $\Theta(\pi) \subseteq \Theta(\pi, \varepsilon)$. Since $G(\hat{\varepsilon}) > 0$, this implies that $p^{t+1}_{\Pi} \in BR(\Delta(\Theta(\pi, \varepsilon))) = \{ \pi \}$, as desired.

(ii) Let $\partial \Theta$ denote the boundary of $\Theta$. By Assumption $\Pi(1)$ $\partial \Theta$ is a compact set, so 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 the $\hat{\varepsilon}$ ball $B_{\hat{\varepsilon}}(\theta')$ around $\theta'$ then

$$
H_{\sigma, \psi}(Q^{*}, Q_{\theta}) - \arg\min_{\hat{\theta} \in \Theta} H_{\sigma, \psi}(Q^{*}, Q_{\hat{\theta}}) > K.
$$

This in turn implies that $\Theta_{\varepsilon}(\psi) = \Theta(\psi)$, when $\varepsilon < \hat{\varepsilon}$. Therefore, the equilibrium is resistant to every $\varepsilon$ local mutation. To see this notice that the sequence in which both the mutated and the conformist agents continue to play $\psi$ at every period, and the share of mutated and conformist agents remain fixed is a solution: equation $\Pi$ is satisfied since $\Theta_{\varepsilon}(\psi) = \Theta(\psi)$,
and equation (2) is trivially satisfied since both subpopulations have the same distribution over strategies.

For every $\gamma, \beta \in (0, 1)$ and $\Theta, \Theta' \in \mathcal{K}$, let

$$P_{\gamma, \beta}(\Theta, \Theta') = \{p \in P : U^*(p|\Theta')) - U^*(p|\Theta)) \geq \gamma, \min\{p_{\mathcal{K}}(\Theta), p_{\mathcal{K}}(\Theta') \geq \beta\}$$

denote the set of states where the strategy used by agents with subjective model $\Theta'$ outperforms the strategy used by agents with subjective model $\Theta$ by at least $\gamma$, and both populations are larger than $\beta$.

**Lemma 4.** For every $\lambda, \beta \in (0, 1)$ and $\Theta, \Theta' \in \mathcal{K}$, $\min_{p \in P_{\lambda, \beta}(\Theta, \Theta')} \frac{T(p|\Theta')}{T(p|\Theta)} \frac{p_{\mathcal{K}}(\Theta)}{p_{\mathcal{K}}(\Theta')}$ is well defined and strictly larger than 1.

**Proof of Proposition 3.** Since the set of strategies is finite, there is $\varepsilon_1$ such that for all $\varepsilon < \varepsilon_1$, $BR(\Delta(M_\Theta(\varepsilon))) \subseteq \Pi_{M_\Theta}$. Fix such an $\varepsilon_1$ for the entire proof.

(i) Since $U^*$ is continuous by Assumption 2 and Lemma 6.4 in Aliprantis and Border (2013) and $\Pi_{M_\Theta}$ is finite, there are $\varepsilon_1 > 0$ and $\gamma > 0$ such that

$$\|\psi' - \psi\| < \varepsilon \Rightarrow U^*(\psi') - \min_{\pi \in \Pi_{M_\Theta}} U^*(\pi) < -\gamma. \quad (3)$$

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 $\varepsilon < \min \{\varepsilon', \varepsilon^*\}$. Because $\Pi$ is finite and $\Theta_\varepsilon(\cdot)$ is upper-hemicontinuous by Lemma 3, there is $\varepsilon \in (0, \varepsilon)$ such that

$$\|\psi' - \psi\| < \varepsilon \Rightarrow BR(\Delta(\Theta_\varepsilon(\psi'))) \subseteq \Pi_{M_\Theta}. \quad (4)$$

Suppose by way of contradiction that after some $\tau > 0$, for all $t > \tau$, $\|p^t_{\Pi} - \psi\| < \varepsilon$. From equation (4) this implies that $BR(\Delta(\Theta_\varepsilon(p^0_{\Pi}))) \subseteq \Pi_{M_\Theta}$. Since $\varepsilon < \varepsilon < \varepsilon^*$, by equation (3) we get:

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

that is

$$U^*(p^t_{\varepsilon} + 1(\cdot | \Theta_\varepsilon)) > U^*(p^t_{\varepsilon} + 1(\cdot | \Theta_\varepsilon)) + \frac{\gamma}{p^t_{\varepsilon} + 1(\Theta)}. \quad (5)$$

But then, applying Lemma 4 with

$$\beta = \min \left\{ \frac{1}{p_{\mathcal{K}}(\Theta_\varepsilon)}, 1 - \max_{\psi' : \|\psi' - \psi\| \leq \min\{\varepsilon', \varepsilon^*\}} \psi'(P_{M_\Theta}) \right\}$$

32
and $\lambda = \frac{\gamma}{p^1(\Theta)}$, we obtain the contradiction that $\lim_{t \to \infty} p^t(\Theta_\varepsilon) \to 1$.

(ii) We prove this case by constructing a solution in which the play converges back to the equilibrium distribution for a sufficiently small mutation. In particular, we will prove the following stronger result. If for every $\pi' \in \Pi_{\mathcal{M}_\Theta}$, $U^*(\pi') \leq U^*(\psi)$, and $\pi' \in \supp \psi$, then $(\Theta, \psi)$ is resistant to local mutations. Then, observe that the statement in the text follows from the fact that both $BR$ and $\mathcal{M}_\Theta$ are upper hemicontinuous correspondences, so if $\psi$ is a quasi-strict Berk-Nash equilibrium, $\Pi_{\mathcal{M}_\Theta} \cap \supp \psi \subseteq BR(\Delta(\Theta(\psi))) \subseteq \supp \psi$. To do so, we show that there is a solution in which the mutated agents always play $\pi' \in \Pi_{\mathcal{M}_\Theta}$ and the conformists play a strategy distribution very close to the equilibrium in every period. Upper hemicontinuity of the best reply make this possible for period 1 and for a sufficiently small share of mutants. For subsequent periods, this is possible as long as the share of mutants does not increase above the initial step. Payoff-monotonicity and the fact that $\pi'$ has lower payoff than the equilibrium distribution guarantee that the share of mutants does not increase.

Let $\hat{\varepsilon} = \min_{\pi \in \supp \psi(\pi)}\psi(\pi)$. We show that for every $\varepsilon < \min\{\varepsilon', \hat{\varepsilon}\}$ there exists a solution $(p^t)_{t \in \mathbb{N}_0}$ with $p^0 = p_\varepsilon$ and $\lim_{t \to \infty} p^t = \psi$, where $p_\varepsilon$ is the $\varepsilon$ local mutation of $p$. Fix such an $\varepsilon$ and let $\Theta_\varepsilon$ be the local $\varepsilon$ expansion of $\Theta$. Define the candidate solution by letting $p_0$ be given by $p^0_\Pi(\Theta) = 1 - \varepsilon, p^0_\Theta(\Theta_\varepsilon) = \varepsilon, p^0_\Pi(\cdot|\Theta_\varepsilon) = \pi'$ and

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

where $\pi' \in \Pi_{\mathcal{M}_\Theta} \cap \supp \psi \cap BR(\Delta(\mathcal{M}_\Theta(\varepsilon)))$. Subsequent period states are recursively defined as

$$p^{t+1}_\Pi(\cdot|\Theta_\varepsilon) = \frac{T(p^t(\Theta_\varepsilon))}{T(p^t(\Theta_\varepsilon))}, p^{t+1}_\Pi(\cdot|\Theta_\varepsilon) = \pi'$$

and

$$p^{t+1}_\Pi(\pi|\Theta) = \begin{cases} \frac{\psi(\pi)}{p^{t+1}_\Pi(\Theta_\varepsilon)} & \pi \neq \pi' \\ \frac{\psi(\pi)-p^{t+1}_\Pi(\Theta_\varepsilon)}{p^{t+1}_\Pi(\Theta_\varepsilon)} & \pi = \pi' \end{cases}$$

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

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

**Proof of Lemma 2** First notice that since \( H_{(\sigma, \psi)}(Q^*, Q_\theta) \) is finite at the KL-minimizers, so by Assumption (i) it follows that it is Gateaux continuously differentiable at \( \hat{\theta} \), and \( D_{\rho_\pi} \) is continuous at every \((\hat{\theta}, \hat{\nu})\). Then, by finiteness of the strategies and parameters, if \( \hat{\nu} \) is in \( \arg\max_{\pi \in P} U_{\hat{\theta} + \varepsilon_0}(\pi) \), \( \hat{\nu} \in BR(\Delta(\mathcal{M}_\Theta(\varepsilon))) \).

**Proof of Proposition 4.** Let \((\Theta, \psi)\) be a Berk-Nash equilibrium that is neither pure nor self-confirming, and denote its corresponding steady state \( p = \delta_\Theta \times \psi \). Since \((\Theta, \psi)\) is not self-confirming, there is no \( \theta \) such that for all \( s \in S, Q_\theta(s, \pi(s)) = Q^*(s, \pi(s)) = Q_{\theta^*}(s, \pi(s)) \), and in particular, \( \theta^* \neq \Theta(\psi) \). Since the equilibrium is mixed and the environment is directional, either \( \Theta(\psi) = \{\hat{\theta}\}, \hat{\theta} > \theta^* \), or \( \Theta(\psi) = \{\hat{\theta}\}, \hat{\theta} < \theta^* \). We treat the case \( \hat{\theta} > \theta^* \); the other is symmetric. Here by the directionality assumption, for sufficiently small \( \varepsilon \) the most improving parameter is given by \( \mathcal{M}_\Theta(\varepsilon) = \hat{\theta} - \varepsilon \). Let \( \pi \) be the minimal element of \( BR(\delta_{\hat{\theta}}) \), and \( \pi \in \text{supp } \psi \setminus \{\pi\}, \pi \neq \pi \). Since both \( \pi \) and \( \pi \) are played in equilibrium and \( \Theta(\psi) = \{\hat{\theta}\} \), we have \( U_{\hat{\theta}}(\pi) = U_{\hat{\theta}}(\pi) \). Moreover, \( \Pi_{\mathcal{M}_\Theta} = \{\pi\} \), and since the environment is directional, \( U^*(\pi) < U^*(\pi) \). Since \( \pi \) was arbitrarily chosen in \( \text{supp } \psi \setminus \{\pi\} \neq \emptyset \), we also have \( U^*(\psi) < U^*(\pi) \), and the result follows by Proposition 3.

**Proof of Proposition 5.** First notice that for every \( l \in \{1, ..., k\} \), \( \mathcal{P}_l(p) \) is compact. Given the finiteness of the action and signal spaces and compactness of \( \mathcal{P}_l(p) \), there is \( \varepsilon > 0 \) such that if \( \nu \{\theta : \exists \hat{\theta} \in \mathcal{P}_l(p), ||\theta - \hat{\theta}|| < \varepsilon \} > 1 - \varepsilon \) and \( \pi \in BR(\nu) \), then \( \pi \in \Pi_{p,l} \). Because \( \Theta(\cdot) \) is upper-hemicontinuous (see Lemma 3), it follows that there exists \( \varepsilon' > 0 \) such that if \( \Theta' \) is a one-hypothesis \( \varepsilon' \) mutation of \( \Theta \),

\[
||\psi' - \psi|| < \varepsilon' \Rightarrow BR(\Delta(\Theta'(\psi))) \subseteq \bigcup_{l \in \{1, ..., k\}} \Pi_{p,l}.
\]

(i) We prove this case by contradiction. If the mutated agents play strategies that are best replies to the \( l \)-agnostic KL minimizers, and if their share of the population increases to 1, then the aggregate play does not converge back to the equilibrium. We now show that if the strategy distribution converges back to the equilibrium, then after some finite time the mutated agents use strategies that are a best reply to the \( l \)-agnostic KL minimizers, while conformists perform strictly worse. But then from payoff monotonicity, the share of mutants would converge to 1, a contradiction.
Suppose that for some \( l \in \{1, \ldots, k\} \), we have \( U^*(\pi) > U^*(\psi) \) for every \( \pi \in \Pi_{p,l} \). By continuity of \( U^* \) and the finiteness of \( \Pi \) there is \( \varepsilon^* > 0 \) such that

\[
||\psi' - \psi|| < \varepsilon^* \Rightarrow U^*(\psi') - U^*(\psi'') < -\gamma \quad \forall \psi'' \in \Delta(\Pi_{p,l}). \tag{7}
\]

Let \( \Theta_l \) be the one-hypothesis relaxation of \( \Theta \) in hypothesis \( l \). Observe that \( \mathcal{P}_l(p) \) is compact. Given the finiteness of the action and signal spaces and compactness of \( \mathcal{P}_l(p) \), there is \( \varepsilon' > 0 \) such that

\[
||p_{\Pi}^l - \psi|| < \varepsilon' \Rightarrow p^{t+1}(\cdot|\Theta_l) \in \Delta(\Pi_{p,l}) \quad \forall p^t \in P. \tag{8}
\]

Moreover, by Lemma \( \boxplus \) \( \mathcal{P}_l \) is upper hemicontinuous.

Suppose by way of contradiction that for \( \varepsilon < \min\{\varepsilon', \varepsilon^*\} \) there is an \( \varepsilon \) mutation to \( \Theta' \), \( p^0_\varepsilon \), such that \( \lim_{t \to \infty} (p^t_\varepsilon)_\Pi = \psi \). This means that after some \( \tau > 0 \), for all \( t > \tau \), \( ||(p^t_\varepsilon)_\Pi - \psi|| < \varepsilon \). Since \( \varepsilon < \varepsilon' \), by equation \( \square \), \( p^{t+1}_\varepsilon(\cdot|\Theta_l) \in \Delta(\Pi_{p,l}) \). Since \( \varepsilon < \varepsilon^* \), by equation \( \Box \), \( U^*((p^{t+1}_\varepsilon)_\Pi) + \gamma < U^*(p^{t+1}_\varepsilon(\cdot|\Theta_l)) \). Therefore, \( U^*(p^{t+1}_\varepsilon(\cdot|\Theta_l)) > U^*((p^{t+1}_\varepsilon)_\Pi) + \gamma 

\[
U^* (p^{t+1}_\varepsilon(\cdot|\Theta_l)) > U^* (p^{t+1}_\varepsilon(\cdot|\Theta)) + \frac{\gamma}{p^{t+1}_\varepsilon(\cdot|\Theta)}. \tag{9}
\]

Then applying Lemma \( \Box \) with \( \Theta' = \Theta_l, \lambda = \frac{\gamma}{p^{t+1}_\varepsilon(\cdot|\Theta)}, \) and \( \beta = \min \{p^{t+1}_\varepsilon(\cdot|\Theta_l), 1 - \max_{\psi':||\psi - \psi'||<\varepsilon} \psi' (\Pi_{p,l})\} \), yields \( \lim_{t \to \infty} p^t_\varepsilon(\Theta_l) \to 1 \), a contradiction.

(ii) Suppose first that \( \pi' \in \text{supp} \psi \). Let \( \bar{\varepsilon} = \min_{\pi \in \text{supp} \psi} \psi(\pi) \). We prove case (ii) by showing that for every \( l \in \{1, \ldots, k\} \) and \( \varepsilon < \min\{\varepsilon', \bar{\varepsilon}\} \) there exists a solution \( (p^t)_{t \in \mathbb{N}_0} \) where \( p^0 \) is a 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 \( l \in \{1, \ldots, m\} \). Define the candidate solution by letting \( p_0 \) be given by \( p^0_\varepsilon(\cdot|\Theta) = 1 - \varepsilon, p^0_\varepsilon(\cdot|\Theta') = \varepsilon, p^{0}_\Pi(\cdot|\Theta) = \pi' \) and

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

\( \pi' \) is an arbitrary element of \( \Pi_{p,l} \). Subsequent period states are recursively defined as
\[ \frac{p_{K}^{t+1}(\Theta)}{p_{K}^{t+1}(\Theta')} = \frac{T(p^t)(\Theta)}{T(p^t)(\Theta')}, \quad p_{\Pi}^{t+1}(\cdot | \Theta') = \pi \text{ and} \]

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

By definition, \((p^t)_{t\in\mathbb{N}_0}\) satisfies equation \([2]\). It only remains to show that for every \(t \in \mathbb{N}_0\), \(\psi(\pi') \geq p_{K}^{t}(\Theta')\) 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 \(\Theta'\) is an \(\epsilon\) mutation of \(\Theta\), and \(\epsilon < \hat{\epsilon} = \min_{\pi \in \text{supp} \psi} \psi(\pi), \psi(\pi') \geq p_{K}^{0}(\Theta') = \epsilon\). Moreover, by equation \([6]\), 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 \(\psi(\pi') \geq p_{K}^{t}(\Theta')\) by the inductive hypothesis, we have \(p_{K}^{t+1}(\Theta') \leq p_{K}^{t+1}(\Theta') \leq \psi(\pi')\). Moreover, since by the inductive step \((p^t)_{\Pi} = \psi\), by equation \([6]\), equation \([1]\) is satisfied for \(t + 1\) as well.

We prove the case in which \((\Theta, \psi)\) is a uniformly strict Berk-Nash equilibrium by letting the \(\hat{\epsilon}\) used in the definition of \(\epsilon\) be such that \(|a - \psi'| \leq \hat{\epsilon}\) implies \(BR(\Delta(\Theta(\psi'))) \subseteq BR(\Delta(\Theta(\psi)))\) and \(BR(\Delta(\Theta'(\psi'))) \subseteq \Pi_{p,t}\). We show that for every \(l \in \{1, \ldots, k\}\) and \(\epsilon < \min\{\epsilon', \hat{\epsilon}\}\) there exists a solution \((p^t)_{l \in \mathbb{N}_0}\) where \(p^0\) is a 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 \(\epsilon\) and let \(\Theta'\) be the one-hypothesis relaxation of \(\Theta\) in hypothesis \(l \in \{1, \ldots, m\}\). Define the candidate solution by letting \(p_0\) be given by \(p_{K}^{0}(\Theta) = 1 - \epsilon, p_{K}^{0}(\Theta') = \epsilon, p_{\Pi}^{0}(\cdot | \Theta') = \pi_0\) and \(p_{\Pi}^{0}(\cdot | \Theta) = \hat{\pi}_0\), where \(\pi_0\) is an arbitrary element of \(\Pi_{p,t}\) and \(\hat{\pi}_0\) is an arbitrary element of \(BR(\Delta(\Theta(\psi)))\). Subsequent period states are recursively defined as \(\frac{p_{K}^{t+1}(\Theta)}{p_{K}^{t+1}(\Theta')} = \frac{T(p^t)(\Theta)}{T(p^t)(\Theta')}, p_{\Pi}^{t+1}(\cdot | \Theta') = \pi_t\) and \(p_{\Pi}^{t+1}(\pi | \Theta) = \hat{\pi}\), where \(\pi_0\) is an arbitrary element of \(BR(\Delta(\Theta(p^t)_{\Pi}))\). By definition, \((p^t)_{t \in \mathbb{N}_0}\) satisfies equation \([2]\). It only remains to show that for every \(t \in \mathbb{N}_0\), \(\hat{\epsilon} \geq p_{K}^{t}(\Theta')\) so that equation \([1]\) is satisfied and \(BR(\Delta(\Theta'(p^t)_{\Pi})) \subseteq \Pi_{p,t}\). 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, \(U^*(\pi') \leq U^*(\hat{\pi})\), for every \(\pi' \in \Pi_{p,t}\) and \(\hat{\epsilon} \geq p_{K}^{t}(\Theta')\) by the inductive hypothesis, we have \(p_{K}^{t+1}(\Theta') \leq p_{K}^{t}(\Theta') \leq \hat{\epsilon}\). Finally, the proof is concluded by applying Lemma \([4]\).

The next lemma is used in the proof of Proposition \([6]\). It generalizes Jensen’s inequality by showing that if two parameters have the same KL divergence given \(\psi\), and they do not assign the same probability to all events that have positive probability under \(\psi\), then their strict convex combination has a strictly lower KL divergence given \(\psi\).

36
Lemma 5. For $\psi \in \Delta(\Pi)$ let

$$Q_\psi(s, a, B) = \sigma(s) \psi(\{\pi : \pi(s) = a\})Q^*(B|s, a) \quad \forall s \in S, a \in A, B \in B(Y).$$

Suppose that $\theta^1, \theta^2, \theta^3 \in \mathbb{R}^k$ are such that $q_{\theta^1}(|\cdot|, \cdot)$ is not $Q_\psi$ almost surely equal to $q_{\theta^3}(|\cdot|, \cdot)$ and $q_{\theta^1}(y|s, a) = vq_{\theta^2}(y|s, a) + (1 - v) q_{\theta^3}(y|s, a)$, $Q_\psi$ a.s. for some $v \in (0, 1)$. If

$$\sum_{s \in S} \sum_{\pi \in A^S} \sigma(s) \psi(\pi)H(Q^*(|s, \pi(s)), Q_{\theta^2}(|s, \pi(s))), Q_{\theta^3}(|s, \pi(s))) = \sum_{s \in S} \sum_{\pi \in A^S} \sigma(s) \psi(\pi)H(Q^*(|s, \pi(s)), Q_{\theta^3}(|s, \pi(s))) \neq \infty$$

then

$$\sum_{s \in S} \sum_{\pi \in A^S} \sigma(s) \psi(\pi)H(Q^*(|s, \pi(s)), Q_{\theta^1}(|s, \pi(s))), Q_{\theta^2}(|s, \pi(s))) < \sum_{s \in S} \sum_{\pi \in A^S} \sigma(s) \psi(\pi)H(Q^*(|s, \pi(s)), Q_{\theta^3}(|s, \pi(s))) .$$

The proof of this lemma is in the Online Appendix.

Proof of Proposition 6. The proof for the case in which $\{\hat{\pi}\} = \text{BR}(\Delta(\Theta(p_{\Pi}))$ is simple: the main theorem in Berk (1966) guarantees that almost surely $\lim_{n \to \infty} \mu_{i}(\Theta(p_{\Pi})) = 1$, and the upper-hemicontinuity of the best-reply correspondence guarantees that $\psi_{n} \to \delta_{\hat{\pi}}$.

The proof of the case with multiple best replies to the KL-minimizing parameters follows from three claims. Claim 1 shows that almost surely the posterior beliefs will assign probability one to the KL-minimizing parameters for $p_\Pi$. This is similar to Lemma 2 of Esponda and Pouzo (2016), but it also covers the case in which $Y$ is infinite. Claim 2 shows that beliefs that induce ties have Lebesgue measure zero in the space of likelihood ratios. Claim 3 shows that the likelihood ratios between different minimizers is a random walk and that it converges. We show this with an argument similar to one in Fudenberg, Lanzani, and Strack (2020); we discuss some of the differences below.

Partition the elements of $\Theta$ in the equivalence classes $\{\tilde{\theta}^i\}_{i=1}^C$ such that

$$\theta^1, \theta^2 \in \tilde{\theta}^i \implies q_{\theta^1}(|\cdot|, \cdot) = q_{\theta^2}(|\cdot|, \cdot) \quad Q_{p_{\Pi}} \text{ almost surely.}$$

Without loss of generality, let $\tilde{\theta}^1, ..., \tilde{\theta}^K$ be the equivalence classes that contain the elements of $\Theta$ in $\Theta(p_{\Pi})$. In what follows for every $i \in \{1, ..., C\}$ we denote by $\theta^i$ an arbitrary element of $\tilde{\theta}^i$. For every $m \in \mathbb{N}$, let

$$\mu_m(\tilde{\theta}^i) = \mu_{\Theta}(\tilde{\theta}^i) \frac{q_{\theta^i}(y_m|s_m, a_m)}{\sum_{i \in \{1, ..., C\}} q_{\theta^i}(y_m|s_m, a_m)} \quad \forall i \in \{1, ..., C\} .$$

37
which is well defined $Q_{p_l}$ almost surely. With this, for all $l \in \{1, \ldots, C\}$ define

$$Z^l_m = \log \frac{\mu_m(\tilde{\theta}^l)}{\hat{\mu}_m(\tilde{\theta}^1)} \text{ and } L^l_m = \log \frac{q_{\theta_l}(y_m|s_m, a_m)}{q_{\theta_l}(y_m|s_m, a_m)}; \text{ so } Z^l_m = Z_0^l + \sum_{i=1}^{m} L^l_i.$$

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

For every $\nu \in \Delta \left(\left\{\tilde{\theta}^1, \ldots, \tilde{\theta}^K\right\}\right)$, define $U(\nu, a) = \sum_{\theta_j \notin \theta} \nu(\tilde{\theta}^i) \sum_{\theta_j \notin \theta} \frac{\mu_\theta(\theta)^{\tilde{\theta}}}{\mu_{\theta}(\theta)} U_{\theta_j}(a)$.

**Claim 2.** The set of $\nu \in \Delta \left(\left\{\tilde{\theta}^1, \ldots, \tilde{\theta}^K\right\}\right)$ such that $U(\nu, a) = U(\nu, a')$ for some $a \neq a'$ has Lebesgue measure 0 in $\mathbb{R}^K$.

Fudenberg, Lanzani, and Strack (2020) shows that the probability assigned to a particular KL-minimizing parameter is large infinitely often; here we need to show that the entire distribution of beliefs converges. To do that we first show that the limit distribution assigns probability zero to ties. This is done by working with the distribution of odds ratios, showing that its limit is normal, and that its covariance matrix is positive definite, which implies that every set of Lebesgue measure zero has measure zero under the limit distribution.

**Claim 3.** The process $\left(Z^l_{i}\right)_{i=2}^K$ is a multi-dimensional random walk in $\mathbb{R}^{K-1}$, and the covariance matrix of its increments is positive definite.

**Proof.** For every $l \in \{2, \ldots, K\}, E \left[ L^{l}_m | \left(Z^l_{1}\right)_{i=1}^{m-1} \right] = H_{\sigma, p_l}(Q^*, Q_{\theta^l}) - H_{\sigma, p_l}(Q^*, Q_{\theta})$, so $\left(Z^l\right)_{i=2}$ is a multi-dimensional random walk. Because $Q_{\theta^l}(\cdot|s, a)$ is absolutely continuous with respect to $Q^*(\cdot|s, a)$ for all $s \in S$ and $a \in A$, the increments $L_l$ have covariance matrix $\Sigma$ given by

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

We want to show that this covariance matrix is positive definite. This is equivalent to $v^T \Sigma v > 0$ for all $v \in \mathbb{R}^{K-1}_+$ with $||v||_1 = 1$. Take an arbitrary vector $v \in \mathbb{R}^{K-1}_+$ with $||v||_1 = 1.$
We first observe that $v^T \Sigma v$ is non-negative:

$$
v^T \Sigma v = \sum_{i=2}^{K} \sum_{j=2}^{K} v_i \Sigma_{ij} v_j = \sum_{i=2}^{K} \sum_{j=2}^{K} v_i v_j \int_{S \times A \times Y} \log \left( \frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)} \right) \log \left( \frac{q_{\theta^j}(y|s,a)}{q_{\theta^1}(y|s,a)} \right) dQ_{\pi}(s,a,y)
$$

$$
= \int_{S \times A \times Y} \sum_{i=2}^{K} \sum_{j=2}^{K} v_i \log \left( \frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)} \right) v_j \log \left( \frac{q_{\theta^j}(y|s,a)}{q_{\theta^1}(y|s,a)} \right) dQ_{\pi}(s,a,y)
$$

$$
= \int_{S \times A \times Y} \left( \sum_{i=2}^{K} v_i \log \left( \frac{q_{\theta^i}(y|s,a)}{q_{\theta^1}(y|s,a)} \right) \right)^2 dQ_{\pi}(s,a,y) \geq 0.
$$

Moreover, since the last expression in the display above is the integral of a weakly positive function, it equals zero if and only if the integrand is $Q_{\pi}$ almost surely equal to 0:

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

By Jensen’s inequality the previous expression implies that $Q_{\pi}$ almost surely $\log q_{\theta^i}(y|s,a) \leq \log \sum_{i=2}^{K} v_i q_{\theta^i}(y|s,a)$, so $q_{\theta^1}(y|s,a) \leq \sum_{i=2}^{K} v_i q_{\theta^i}(y|s,a)$. And as $\int_{S \times A \times Y} q_{\theta^i}(y|s,a)dQ_{\psi}(s,a,y) = 1$ for all $i \in \{1, ..., K\}$ this implies that $Q_{\psi}$ almost surely $q_{\theta^1}(y|s,a) = \sum_{i=2}^{K} v_i q_{\theta^i}(y|s,a)$. By Lemma 5 this contradicts that $\theta^2 \in \Theta(\pi_1)$. Thus, $v^T \Sigma v > 0$ and since $v$ was chosen arbitrarily, $\Sigma$ is positive definite, proving Claim 3.

Because $(Z_i^l)_{l=2}^K$ is a martingale with positive definite covariance matrix of the increments, the central limit theorem implies that $(\frac{Z_i^l}{\sqrt{m}})_{l=2}^K$ converges in distribution to a $(K-1)$ dimensional Normal distribution with mean 0 and covariance matrix $\Sigma$. Since

$$
\mu_m(\hat{\theta}_l) = \frac{\exp Z_i^l}{\sum_{i=2}^{C} \exp Z_i^l + 1}, \quad \forall l \in \{2, ..., C\},
$$

the distribution of $\mu_m$ is converging to some $\nu \in \Delta(\Delta(\Theta(\pi_1)))$ and the distribution of beliefs at time $m$ in the overall population converges to $\nu$ as well. Because the covariance matrix $\Sigma$ is positive definite by Claim 2, the induced distribution of strategies converges to an element of $\Delta(BR(\Delta(\Theta(\pi_1))))$. This concludes the proof of part (i).
References


Eprint: [1904.08551](https://arxiv.org/abs/1904.08551)

41


B Online Appendix

B.1 Proof of the results of Section 7

Proof of Proposition 7 Let \( \varepsilon' \in (0, 1) \) be such that for all \( \varepsilon < \varepsilon' \), \( \mathcal{M}_\Theta(\varepsilon) \) is a singleton, and \( V(\mathcal{M}_\Theta(\varepsilon)) > U^*(\psi) \). 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 a \( \varepsilon^* \in (0, \varepsilon') \) and \( \gamma > 0 \) such that for all \( \psi \in BR(\Delta(B_\varepsilon(\mathcal{M}_\Theta(\varepsilon))) \cap B_\varepsilon(\psi) \),

\[
U^*(\psi') - U^*(\psi) < -\gamma.
\]

(10)

By the upper hemicontinuity (see Lemma 3) of \( \Theta(\cdot) \) there is \( \varepsilon \) such that \( ||\psi' - \psi|| < \varepsilon \) implies \( \Theta_\varepsilon(\psi') \subseteq B_\varepsilon(\mathcal{M}_\Theta(\varepsilon)) \). Thus if there were a \( t \) such that \( ||p_{\tilde{n}} - \psi|| < \min \varepsilon^* , \varepsilon \) for all \( \tau > t \), it would follow from equation (10) that

\[
U^*(p_{\tau+1}(\cdot|\Theta)) + \frac{\gamma}{p_{\tau}^0(\Theta)} < U^*(p_{\tau+1}(\cdot|\Theta')).
\]

(11)

By Lemma 4 this concludes the proof.

Proof of Proposition 8 Suppose that \( \hat{a} \) is an attractor with associated KL minimizer \( \hat{\theta} = \Theta(\hat{a}) \). Let \( A_m \) be given by \( \{\hat{a} + \frac{i}{m} c\} \cap A \), and let \( \hat{\varepsilon} \) be such that for all \( \varepsilon < \hat{\varepsilon} \), \( V(\mathcal{M}_{\hat{\theta}}(\varepsilon)) > V(\hat{\theta}) \). Let \( BR_n(\theta) = \arg \max_{a \in A_n} U_\theta(a) \). By Theorem 1 in Espinosa 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 \( \pi_m \) and a justifying belief \( \mu_m \in \Delta(\Theta(\pi_m)) \). Since \( A \) is compact, \( \Delta(A) \) is also compact, and therefore \( (\pi_m, \mu_m) \) admits a converging subsequence to a limit \( (\hat{\pi}, \hat{\mu}) \).

We claim that the limit of this subsequence is \( (\delta_\hat{a}, \delta_{\Theta(\hat{a})}) \). 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 \( a_n \to a \), \( a_n \in A_n \), and since \( U(\cdot) \) is jointly continuous in beliefs and actions

\[
U_{\mu_n}(\pi_n) \geq U_{\mu_n}(a_n) \quad \forall n \in \mathbb{N}
\]

implies \( U_{\hat{\mu}}(\hat{\pi}) \geq U(a) \). Moreover, by the upper hemicontinuity (see Lemma 3) of \( \Theta \),

\[
\mu_n \in \Delta(\Theta(\pi_n)) \quad \forall n \in \mathbb{N}
\]

implies that \( \hat{\mu} \in \Delta(\Theta(\hat{\pi})) \). That is, \( (\hat{\pi}, \hat{\mu}) \) must be the unique Berk-Nash equilibrium of the environment with a continuum of actions.

Next, we will show that this subsequence \( (A_n, \pi_n)_{n \in \mathbb{N}} \) satisfies the requirement in the statement of the proposition. It is clear that \( ||A_n - A|| \to 0 \). Since \( \mu_n \to \delta_{\Theta(\hat{a})} \), there exists a \( N > 0 \) such that for all \( n > N \), \( \sup \pi_n \subseteq B_\varepsilon(\hat{a}) \). Therefore, since \( (\hat{a}, \Theta) \) is an attractor,
there is an $N' > N$ such that for $n \geq N'$, $\pi^*_n (\{a : a < \hat{a}\}) > 0$ and $\pi^*_n (\{a : a > \hat{a}\}) > 0$.

Since $U^*$ is continuously differentiable and $(\delta_{\hat{a}}, \delta_{\Theta(\hat{a})})$ is not resistant to 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 $[\hat{a} - K, \hat{a}]$. We prove the result in the first case, the other case being analogous. By $V (\mathcal{M}_\Theta (\epsilon)) > V (\mathcal{M}_\Theta (\hat{a})) = U^* (\hat{a})$ for all $\epsilon < \hat{\epsilon}$, there exists an $\epsilon' < \hat{\epsilon}$ such that for all $\epsilon < \epsilon'$, $BR (\mathcal{M}_\Theta (\epsilon)) \in (\hat{a}, \hat{a} + K)$. Fix an $\epsilon < \min \{\epsilon', \hat{\epsilon}\}$. By the upper hemicontinuity of $\Theta (\cdot)$, we have that $\Theta (\pi_n)$ is converging to $M_\Theta (\epsilon) = \theta (\hat{a})$. Therefore, there exists $N'' > N'$ such that for all $n > N''$, $BR_n (\Theta (\pi_n)) \in \left(\frac{\hat{a} + BR (M_\Theta (\epsilon))}{2}, \hat{a} + K\right)$. But since $U^*$ is strictly increasing on $[\hat{a}, \hat{a} + K]$, $U^* (\pi_n) \to U^* (\hat{a})$, and $\pi^*_n (\{a : a < \hat{a}\}) > 0$, there exists a $N'' > N''$ such that for all $n > N''$, $U^* (BR_n (\Theta (\pi_n))) > U (\pi_n)$, so that $\pi_n$ is not resistant to an $\epsilon$ expansion of $\Theta$.

\section*{B.2 Proofs of results stated in the Appendix}

\subsection*{B.2.1 Proof of Lemma 3}

\textbf{Proof.} Since $\Theta$ is compact, it is enough to show that $\Theta (\cdot)$ has a closed graph (see, e.g., Theorem 17.11 in Aliprantis and Border (2013)). Fix any $(\psi_n)_{n \in \mathbb{N}}$ and $(\theta_n)_{n \in \mathbb{N}}$ such that $\psi_n \to \psi$, $\theta_n \to \theta$, and $\theta_n \in \Theta (\psi_n)$ for all $n \in \mathbb{N}$. Suppose by way of contradiction that $\theta \notin \Theta (\psi)$. Then there exists $\delta \in \Theta$ and $K \in \mathbb{R}^+$ such that $H_{\sigma, \psi} (Q^*, Q_\theta) + K < H_{\sigma, \psi} (Q^*, Q_\theta)$. Since $H_{\sigma, \psi} (Q^*, \cdot)$ is jointly lower semicontinuous see, e.g., Chapter 1 of Liese and Vajda, 1987, there exists an $N \in \mathbb{N}$ such that $H_{\sigma, \psi_n} (Q^*, Q_\theta_n) > H_{\sigma, \psi_n} (Q^*, Q_\theta) + K$ for all $n > N$. Moreover, since $H_{\sigma, \psi} (Q^*, Q_\theta) < \infty$, $H_{\sigma, \psi} (Q^*, Q_\theta)$ is continuous in the strategy distribution at $\psi$, and there exists $N' > N$ such that $H_{\sigma, \psi_n} (Q^*, Q_\theta) < H_{\sigma, \psi_n} (Q^*, Q_\theta) + \frac{K}{2}$, for all $n > N$.

But then, we would have

$$H_{\sigma, \psi_n} (Q^*, Q_\theta) < H_{\sigma, \psi_n} (Q^*, Q_\theta),$$

a contradiction.

\subsection*{B.2.2 Proof of Lemma 4}

\textbf{Proof.} Because $U^*$ is continuous, $P_{\lambda, \beta} (\Theta, \Theta')$ is a compact subset of $P$. Therefore, since $T (\cdot) (\Theta)$ is continuous and strictly positive on $P_{\lambda, \beta} (\Theta, \Theta')$, it is bounded away from 0 on this set. Moreover, by definition $p_K (\Theta') \geq \beta > 0$ on $P_{\lambda, \beta} (\Theta, \Theta')$, so $\frac{T (p (\Theta') p_K (\Theta))}{T (p (\Theta)) p_K (\Theta')}$ is continuous on $P_{\lambda, \beta} (\Theta, \Theta')$, so it attains a minimum. The payoff monotonicity of $T$ and the fact that the strategy used by agents with subjective model $\Theta'$ strictly outperforms the strategy used by agents with subjective model $\Theta$ on $P_{\lambda, \beta} (\Theta, \Theta')$ implies that this minimum is strictly larger than 1.
Proof. Since \( q_{\theta^{2}}(\cdot, \cdot) \) is not \( Q_{\psi} \) almost surely equal to \( q_{\theta^{3}}(\cdot, \cdot) \) there exist \( s' \in S, a' \in A, \pi' \in \text{supp } \psi, B \in \mathcal{B}(Y) \) with \( Q^{*}(B|s', a') > 0 \), and \( K \in \mathbb{R}^{+} \) such that \( a = \pi(s) \) and \( q_{\theta^{2}}(y|a, s) > q_{\theta^{3}}(y|a, s) + K \) for all \( y \in B \). Moreover, since

\[
\sum_{s \in S} \sigma(s) \sum_{\pi \in \text{supp } A} \psi(\pi) H \left( Q^{*}(\cdot|s, \pi(s)), Q_{\theta^{2}}(\cdot|s, \pi(s)) \right) \\
= \sum_{s \in S} \sigma(s) \sum_{\pi \in \text{supp } A} \psi(\pi) H \left( Q^{*}(\cdot|s, \pi(s)), Q_{\theta^{3}}(\cdot|s, \pi(s)) \right) \neq \infty
\]

the set \( B \) can be chosen such that

\[
K \leq q_{\theta^{2}}(y|a, s) \leq \bar{K} \quad \text{and} \quad \underline{K} \leq q_{\theta^{3}}(y|a, s) \leq \bar{K} \quad \text{for all } y \in B
\]

for some \( \underline{K}, \bar{K} \). Let

\[
\rho = \min_{x \in [\underline{K}, \bar{K}], x' \in [1, 1+\bar{K}]} \log (vx + (1 - v) x') - v \log (x) - (1 - v) \log (x') > 0 \tag{12}
\]

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

Notice that the formula for relative entropy can be expanded as

\[
\sum_{s \in S} \sigma(s) \sum_{\pi \in \text{supp } A} \psi(\pi) H \left( Q^{*}(\cdot|s, \pi(s)), Q_{\theta^{1}}(\cdot|s, \pi(s)) \right) \\
= - \sum_{s \in S \setminus \{s'\}} \sigma(s) \sum_{\pi \in \text{supp } A} \psi(\pi) \int_{Y} \log \left( v q_{\theta^{2}}(y|s, \pi(s)) + (1 - v) q_{\theta^{3}}(y|s, \pi(s)) \right) dQ^{*}(y|s, \pi(s)) \\
- \sigma(s') \sum_{\pi \in \text{supp } A \setminus \{\pi'\}} \psi(\pi) \int_{Y} \log \left( v q_{\theta^{2}}(y|s', \pi(s)) + (1 - v) q_{\theta^{3}}(y|s, \pi(s)) \right) dQ^{*}(y|s', \pi(s)) \\
- \sigma(s') \psi(\pi') \int_{Y \setminus B} \log \left( v q_{\theta^{2}}(y|s', a') + (1 - v) q_{\theta^{3}}(y|s', a') \right) dQ^{*}(y|s', a') \\
- \sigma(s') \psi(\pi') \int_{B} \log \left( v q_{\theta^{2}}(y|s', a') + (1 - v) q_{\theta^{3}}(y|s', a') \right) dQ^{*}(y|s', a').
\]
By equation (12) and Jensen’s inequality, the last expression is weakly lower than
\[- \sum_{s \in S \setminus \{s'\}} \sigma(s) \sum_{\pi \in A^S} \psi(\pi) \int_Y (v \log (q_{\theta^2}(y|s, a)) + (1 - v) \log(q_{\theta^3}(y|s, a))) \, dQ^* (y|s, \pi(s)) \]
\[- \sigma(s') \sum_{\pi \in A^S \setminus \{\pi'\}} \psi(\pi) \int_Y (v \log q_{\theta^2}(y|s', \pi(s)) + (1 - v) \log q_{\theta^3}(y|s, \pi(s))) \, dQ^*(y|s', \pi(s')) \]
\[- \sigma(s') \psi(\pi') \int_{Y \setminus B} (v \log q_{\theta^2}(y|s', a') + (1 - v) \log q_{\theta^3}(y|s', a') + \rho) \, dQ^*(y|s, \pi(s')) \]
\[= \sum_{s \in S} \sigma(s) \sum_{\pi \in A^S} \psi(\pi) H(Q^*(\cdot|s, \pi(s)), Q_{\theta^1}(\cdot|s, \pi(s))) - \rho Q^*(B|s', a') \psi(\pi') \sigma(s') \]
which is strictly less than
\[
\sum_{s \in S} \sigma(s) \sum_{\pi \in A^S} \psi(\pi) H(Q^*(\cdot|s, \pi(s)), Q_{\theta^2}(\cdot|s, \pi(s))) .
\]
Thus
\[
\sum_{s \in S} \sigma(s) \sum_{\pi \in A^S} \psi(\pi) H(Q^*(\cdot|s, \pi(s)), Q_{\theta^1}(\cdot|s, \pi(s))) < \sum_{s \in S} \sigma(s) \sum_{\pi \in A^S} \psi(\pi) H(Q^*(\cdot|s, \pi(s)), Q_{\theta^2}(\cdot|s, \pi(s))) ,
\]
as desired.

B.2.4 Proof of Claim

Proof. If \( \Theta = \Theta(p_{\Pi}) \) the result immediately follows. Suppose \( K < C \). For \( l > K \),
\[ E \left[ L^l_m \left| (Z^l_i)_{i=1}^{m-1} \right. \right] \]is equal to
\[
\sum_{s \in S} \sigma(s) \sum_{\pi \in A^S} p_{\Pi}(\pi) \left[ H \left( Q^*(\cdot|s, \pi(s)), Q_{\theta^1}(\cdot|s, \pi(s)) \right) - H \left( Q^*(\cdot|s, \pi(s)), Q_{\theta^2}(\cdot|s, \pi(s)) \right) \right] < 0.
\]
Since \( \Theta \in \mathcal{K} \), \[ E \left[ (L^l_m)^+ \left| (Z^l_i)_{i=1}^{m-1} \right. \right] < \infty \] 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,
\[
\limsup_{m \to \infty} \frac{\mu_m((\Theta|\Theta(p_{\Pi}))}{\mu_m((\Theta(p_{\Pi}))} \leq \limsup_{m \to \infty} \frac{\mu_m((\Theta|\Theta(p_{\Pi}))}{\mu_m((\Theta(p_{\Pi})}} = \limsup_{m \to \infty} \frac{\Theta}{\sum_{l=K+1}^{\infty} \exp Z^l_m \text{ a.s.}} = -\infty,
\]
proving the claim.
B.2.5 Proof of Claim \[2\]

Proof Fix \(a \neq a'\). The condition \(U(\nu, a) = U(\nu, a')\) by definition is equal to

\[
U(\nu, a) = \sum_{i=1}^{K} \nu(\tilde{\theta}^{i}) \sum_{\theta_j \in \tilde{\theta}^{i}} \frac{\mu_{\theta}(\theta_j)}{\mu_{\theta}(\tilde{\theta}^{i})} U_{\theta_j}(a) = U(\nu, a') = \sum_{i=1}^{K} \nu(\tilde{\theta}^{i}) \sum_{\theta_j \in \tilde{\theta}^{i}} \frac{\mu_{\theta}(\theta_j)}{\mu_{\theta}(\tilde{\theta}^{i})} U_{\theta_j}(a').
\]

Since this is a linear equation in the \(K\) unknowns \(\nu(\tilde{\theta}^{i})\), the \(\mu\) who solve the equations are a vector subspace of \(\mathbb{R}^{K}\). However, assumption (1) of the proposition implies that there exists \(\tilde{\theta}_i \in \Theta(p_{\Pi})\) such that \(U_{\mu(\tilde{\theta}_i)}(a) \neq U_{\mu(\tilde{\theta}_i)}(a')\), so the set of beliefs under which \(U_{\mu}(a) = U_{\mu}(a')\) has dimension at most \(K - 1\), and hence Lebesgue measure 0. Since the set of actions is finite and \(a, a'\) are chosen arbitrary, the set of beliefs \(\nu \in \Delta\left(\{\tilde{\theta}^1, ..., \tilde{\theta}^K\}\right) \subseteq \mathbb{R}^{K}\) such that \(U(\nu, a) = U(\nu, a')\) for some \(a \neq a'\) has Lebesgue measure 0 as well. \(\square\)

B.3 Examples

B.3.1 Example 1

a) Esponda and Pouzo (2016) shows that

\[
\sum_{a \in \{1,2\}} \psi(a) H(Q^{*}(\cdot|a), Q_{\theta}(\cdot|s, a)) = \psi(2)(34 - i + 2l)^2 + \psi(10)(2 - i + 10l)^2. \tag{13}
\]

The first order conditions for \((10/3, 40)\) to be the KL-minimizer are

\[
\begin{align*}
-2\psi(2)(34 - 40 + 2 \ast \frac{10}{3}) - 2(1 - \psi(2))(2 - 40 + 10 \ast \frac{10}{3}) &\leq 0 \tag{14} \\
4\psi(2)(34 - 40 + 2 \ast \frac{10}{3}) + 20(1 - \psi(2))(2 - 40 + 10 \ast \frac{10}{3}) &\leq 0. \tag{15}
\end{align*}
\]

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 \(\tilde{v} = (\sqrt{1 - \tilde{v}^2}, \nu)\) for some \(\nu \in [0, 1]\). With this,

\[
D_{\psi(2), \psi(10)}(\tilde{\theta}, \tilde{v}) = \sqrt{(1 - \tilde{v}^2)}[4(10 + (10/3)(50 - 48\psi(2)) + 24\psi(2) + 40(-5 + 4\psi(2)))]
+ \tilde{v} \ast 2(-2 + 40 - 32\psi(2) + 2(10/3)(-5 + 4 \ast 40)).
\]

This expression is maximized at a \(\tilde{v}\) with \(\sqrt{(1 - \tilde{v}^2)} > 1/12\) if and only if \(\psi(2) > 13245113/13623544 - (21\sqrt(\pi/2))/1702943 \approx 0.972\). Finally, the claim that \(A_{M_{\Theta}} = \{10\}\) if and only if \(\psi(2) \approx 0.972\) follows from

\[
\lim_{h \to 0} \frac{U_{(10/3, 40) + h(\sqrt{(1 - \tilde{v}^2)}, \tilde{v})}(10) - U_{(10/3, 40) + h(\sqrt{(1 - \tilde{v}^2)}, \tilde{v})}(2)}{h} = (-100 - (-4))\sqrt{(1 - \tilde{v}^2) + (10 - 2)\tilde{v}}
\]

OA-5
and $U_{(10/3,40)}(10) = U_{(10/3,40)}(2)$.

b) When $\psi(2) = 1$, the parameter that minimizes equation (13) is $(3/2,32)$, and since $BR(3/2,32) = \{10\}$, 2 is not a Berk-Nash equilibrium. When $\psi(10) = 1$, the parameter that minimizes equation (13) is $(3/2,28)$. Since $BR(3/2,28) = \{2\}$, 10 is not a Berk-Nash equilibrium. For every totally mixed $\psi$, the Hessian of equation (13) as a function of $l$ and $i$, $[2,-4\psi(2) - 20\psi(10)l; -4\psi(2) - 20\psi(10)l, 200 - 196\psi(2)]$, is positive definite for every $(l,i) \in \Theta$ so there is a unique KL-minimizer. For $\psi = (1/4,3/4)$, the unique KL-minimizing parameter $\hat{\theta} = (5/2,30)$ obtained from equation (13) lies on the line $i = 12l$ where the two actions indifferent, and so it is a Berk-Nash equilibrium.

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

$$-2\frac{1}{4}(34 - 30 + 2 \cdot (2.5 + \varepsilon)) - 2(1 - \frac{1}{4})(2 - 30 + 10 \cdot (2.5 + \varepsilon)) = 0$$

so that $M_\Theta(\varepsilon) = (2.5 + \varepsilon, 30 + 8\varepsilon)$.

**Invasion by a non-improving mutation** Consider the subjective model $\Theta$ and the equilibrium $\psi$ of part a). Let $\Theta' = \{(3,33)\}$. Since $\Theta' \subset \Theta$, and $\Theta(\psi) = (10/3,40)$, the mutation of $\delta_\Theta \times \psi$ to $\Theta' = \{(3,33)\}$ is not improving. However, the equilibrium is not resistant to 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, the play does not converge back to $\psi$.

**B.3.2 Example 7**

**Example 7.** [Overconfidence and base-rate neglect] Every period, the agent either buys an asset, $a_B$, sells it, $a_S$, or maintains a neutral position, $a_N$. The asset has either high performance, $y_H$, or low, $y_L$. Before taking an action, the agent receive a binary signal $s \in \{s_L,s_H\}$. The true model is that for all $a \in A$ and $i \in \{H,L\}$, $Q(y_i|a,s_i) = \theta^*_i$, where $\theta^*_i \in (1/2,2/3)$ is the precision of the signal. The agent is uncertain about the precision of their signal: $\Theta = [\theta_H,\bar{\theta}_H] \times [\theta_L,\bar{\theta}_L]$. Their payoff function is

$$u(s,a,y) = \begin{cases} 
1 & (a,y) \in \{(a_B,y_H),(a_S,y_L)\} \\
0 & a = a_N \\
-2 & (a,y) \in \{(a_B,y_L),(a_S,y_H)\}
\end{cases}$$

The objectively optimal action is to play $a_N$ after every signal.

Overconfidence. The agent is overconfident when $\theta^*_i > 2/3 > \theta^*_i$. In the unique Berk-Nash equilibrium the misspecified agent plays $a_B$ after $s_H$ and $a_S$ after $s_L$. Since the equilibrium is uniformly strict, it is resistant to local mutations.

Base-rate neglect. Suppose the agents are told that the unconditional probability of high performance is $p$, and that conditional on outcome $y_i$ they will receive signal $s_i$ with probability
Agents with base-rate neglect (Kahneman and Tversky (1973)) are certain that the precision of a high (low) signal is equal to the probability of the high (low) signal in the (high) state, neglecting the impact of the base rate: \(Q(y_i|a,s_i) = \theta_i = \theta_1 = \sigma(s_i|y_i)\). If \(\sigma(s_i|y_i) > 2/3\), again the unique Berk-Nash equilibrium is that the misspecified agent plays \(a_B\) after \(s_H\) and \(a_S\) after \(s_L\). This equilibrium is uniformly strict, hence resistant to local mutations.

**B.3.3 Example 3**

Since \(\phi_\theta\) is a normal density, we have

\[
\psi(a)H(Q^*(\cdot|a), Q_\theta(\cdot|s,a)) = -\frac{11}{26} \left( \frac{\tau(x+1)}{(x+1)} - \theta_1 \right)^2 + \left( \frac{\tau(x)}{(x)} - \theta_1 \right)^2 + \left( \frac{\tau(x-1)}{(x-1)} - \theta_1 \right)^2.
\]

**B.3.4 Example 4**

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

\[
\sum_{y \in Y} u(-1,3,y) = \frac{1}{2}(3+3-1) + \frac{1}{4}(2+3-1) + \frac{1}{4}(1+3-1) - 3 > \frac{1}{4}(2+3-1-2) + \frac{1}{4}(1+3-1-2)
\]

\[
= \sum_{y \in Y} u(-1,2,y) > \frac{1}{4}(1+3-1-1) = \sum_{y \in Y} u(-1,1,y).
\]

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

The fact that in a Berk-Nash equilibrium the beliefs have correct marginals over prices asked and valuations is proved in the Online Appendix B of Esponda and Pouzo (2016).

Therefore, when \(s\) is identically zero, \(\theta_0 = \Theta(2) = (\frac{1}{6}, \frac{1}{2}, \frac{1}{3}, \frac{1}{6}, 1, \frac{1}{6}, 1, \frac{1}{6}, 1)\), and under \(\theta_0\) bidding 2 is optimal:

\[
U_{\theta_0}(1) = \frac{1}{6}(\frac{1}{6}(4-1)+\frac{5}{6}(5-1)) < (\frac{1}{6}(4-3)+\frac{5}{6}(5-3)) = U_{\theta_0}(3) < \frac{2}{3}(\frac{1}{6}(4-2)+\frac{5}{6}(5-2)) = U_{\theta_0}(2).
\]

**B.3.5 An equilibrium that is not stable with attention partitions but is stable without them**

**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

\[
u(a,(y_1,y_2)) = \begin{cases} 
1 & a = l = y_1 \\
4 & a = l = y_1 \\
0 & a \neq y_1
\end{cases}.
\]
Here $\mathcal{H} = \left\{ (\theta_{al}, \theta_{au})_{a \in A} \in [0, 1]^A \right\}$ where $\theta_{al}$ is the probability of $y_1 = l$ given $a$ and $\theta_{au}$ is the probability of $y_2 = u$ given $a$ has been played. The agent believes that the action does not affect the outcome, i.e., $\Theta = \left\{ \theta \in \mathcal{H} : \theta_{ll} = \theta_{rl}, \theta_{lu} = \theta_{ru} = \frac{1}{2} \right\}$. In reality $\theta_{rl}^* = \frac{1}{4}$, $\theta_{ru}^* = \frac{2}{3}$, $\theta_{lu}^* = \frac{1}{10}$ and $\theta_{ru}^* = \frac{9}{10}$. The unique Berk-Nash equilibrium has $\pi(l) = \frac{2}{5}$ and $\pi(r) = \frac{3}{5}$ supported by the Dirac measure on $\left( \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right)$. Since the agent (correctly) believes that the second outcome component is payoff irrelevant, the minimal attention partition pools outcomes with the same second component. The most improving parameter for the partition information increases the probability of $l$ after $r$ is played, $\theta_{rl}$. Since the best reply to these parameters is the optimal action $r$, the equilibrium would not resist mutations. However, the most improving parameters for a player who distinguishes between all the outcomes increase the (payoff irrelevant) probability of $u$ after $r$, $\theta_{ru}^*$ and the 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.

### B.3.6 Grid approximation of Example 6

Recall that in the model with a continuum of signals the instructor criticizes performance below $T = (-E[-s|s > T] - E[-s|s < T]) / k$ and praises performance above it. Fix an integer $n$, and suppose that the signals take values on the grid $\{ T + \frac{m}{n} : m \in \mathbb{N} \}$ so that as $n$ increases the grid becomes finer, 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. 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

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

Denote the new probability distribution over states as $\sigma_\delta$. 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). Suppose by way of contradiction that the instructor always praises at the threshold, $\pi(T)(r) = 1$. Then,

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

The key is that the distance between the threshold and its closest element is $O(n)$. If the grid is coarser than this, there can also be uniformly strict equilibria.
and
\[ \theta_r(\pi) = -E_{\sigma} [s|s \geq T] < -E_{\sigma} [s|s \geq T] = \theta_r(T) \]
and therefore
\[ \frac{\theta_c(\pi) - \theta_r(\pi)}{k} > \frac{\theta_c(T) - \theta_r(T)}{k} = T. \]

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

This equilibrium is again not resistant to local mutations: the most improving parameter is still a \( ((1, \theta_{ac}(s), \theta_{ar}(s)) + (v_0, v_r, v_c)) \) with \( v_0 < 0, v_r \geq 0, v_c \leq 0 \). The best reply to these parameters is 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 destabilized by the mutation.

### B.3.7 Prior independent limit aggregate behavior

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

**Proposition 9.** If the assumptions of Proposition 6 are satisfied and for every \( \theta \in \Theta(p_\Pi) \), \( \mathcal{E}(\theta, p_\psi) \) is a singleton, 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. The fact that the limit beliefs do not depend on the prior follows from the fact that in Proposition 6 we have shown that the beliefs over equivalence classes is converging 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 \( Z_m \Rightarrow a.s. \to -\infty \) for all \( l \in \{K + 1, \ldots, C\} \) by Claim 1, there exists \( c < 0 \) such that if \( (Z_m^l)_{l=2}^K \) is coordinate by coordinate less than \( c \), the best reply to the corresponding belief is to play \( a \). Consider the events \( E_m \) that \( (Z_m^l)_{l=2}^K \) is coordinate-wise less than \( c \): \( E_m = \{Z_m^l \leq c, \forall l \in \{2, \ldots, K\}\} \). As \( \frac{Z_m}{\sqrt{m}} \) converges to Normal random variable we have that
\[
\lim_{m \to \infty} \mathbb{P}[E_m] = \lim_{m \to \infty} \mathbb{P}\left[ \frac{Z_m}{\sqrt{m}} \leq \frac{c}{\sqrt{m}} \right] = \mathbb{P}\left[ \tilde{Z} \leq 0 \right],
\]
where \( \tilde{Z} \) is a random variable that is Normally distributed with mean \( \bar{0} \) and covariance matrix \( \Sigma \). As \( \Sigma \) is positive definite, this distribution admits a strictly positive density and hence \( \mathbb{P}[\tilde{Z} \leq 0] > 0 \).

\[ \text{OA-9} \]