Which Misperceptions Persist?

Drew Fudenberg and Giacomo Lanzani[*]

First posted version: October 12, 2020
This version: February 11, 2021

Abstract

We use an evolutionary model to determine what sorts of misperceptions are likely to persist. 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. Different agents may employ different subjective models, but subjective models that induce better performing actions increase their relative prevalence. Every steady state corresponds to a Berk-Nash equilibrium, but only some of them can resist mutations that lead a small fraction of agents to enlarge their subjective models. Mutations to a more permissive but still misspecified subjective model fail to spread if incorrect inference from equilibrium data in this larger set lead the agents to take worse actions. 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.

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: Agents may have a behavioral bias such as overconfidence and correlation neglect, or they may over-simplify a complex environment by omitting some relevant variables or interactions between them. For example, when agents misperceives a progressive tax schedule as linear they end up working too much, since they equate their marginal cost of effort to the average instead of marginal tax rate.

We propose an evolutionary criterion to evaluate the plausibility of misspecified Bayesian models and whether they are likely to persist. In our model, the agents are Bayesians who estimate the parameters that best fit their data given their possibly misspecified prior beliefs. 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 estimates the parameters of their subjective model that best fit the data generated by the actions and outcomes of the previous generation.[1] The agents then choose a best reply to a belief that is concentrated on these best-fitting parameters.

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. Steady states in which all agents have the same model coincide with the Berk-Nash equilibria (cf. Esponda and Pouzo, 2016): the action played is a best response to posterior beliefs that fit the data generated by playing that action as well as the agents' subjective model allows.

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 mutation will spread, or whether the existing paradigm is "resistant to mutations."[2] At a high level, our conceptual point is that whether a mutation spreads depends not only on whether it allows a better explanation of the existing data, but also on whether the behavior it induces generates higher payoffs for the agents.

Any self-confirming Berk-Nash equilibrium is automatically resistant to mutations: Be-

---

[1] "Best fit" here means minimizing the Kullback-Leibler divergence from the 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.

cause the current model perfectly fits the outcome distribution induced by the equilibrium action, no mutation can provide a better fit. Importantly, equilibria that are not self-confirming can be stable as well, because the population share of mutants only increases if they do better than agents using the prevailing paradigm. Mutations to more permissive but still incorrectly specified models do not guarantee that better actions are adopted. Whether this is the case depends on the data generated by the prevailing paradigm, and how the mutated agents interpret it.

We consider two particular ways that subjective models can expand: "local" expansions, and "one-hypothesis relaxations." Local expansions have a quantitative nature: The agents add nearby parameters to their subjective models when these better explain the evidence.[3] In contrast, one-hypothesis relaxations capture qualitative changes. Here the subjective model of the agent is composed of a finite set of assumptions about the data generating process, and the mutations lead offspring to challenge one of them while maintaining the others.

First, we show that uniformly strict Berk-Nash equilibria (Fudenberg, Lanzani, and Strack, 2021) 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, since the evidence points in the direction of reducing this bias. We show that in problems with a "directional" structure, only equilibria that are pure and/or self-confirming are stable under local mutations. This structure is present in a number of examples of interest, including when the agent incorrectly believes the tax schedule is linear.

However, these biases can persist when mutations are to a structurally different model, such as a one-hypothesis relaxations of the paradigm. One-hypothesis relaxations allow larger changes in beliefs than local ones do, but equilibria can be resistant to one-hypothesis 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. We characterize stability to one-hypothesis relaxations by considering the KL-

---

[3]This is similar to the neighborhoods of posited "structural" subjective model that are considered in the macroeconomic literature on robust control following Hansen and Sargent (2008).

minimizers in relaxed subjective models where one of the hypotheses is dropped. Specifically, by using the assumption that a finite number of hypotheses determines the subjective model, we show that an equilibrium is resistant if and only if the KL minimizers of the relaxed problem, with respect to the equilibrium action, induce an action that yields less than the equilibrium payoff.

For models of correlation neglect, we argue that misspsecified beliefs are less resistant to mutations in "noisy" environments, because the noise helps the agents correctly infer the correlation between the variables. Since the noise can include preference heterogeneity, this suggest that a homogeneous closed group of agents is more likely to maintain misspecified beliefs.

Although we think that the two forms of mutations we study 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.[4]

In Section 6 we provide an explicit learning foundation for the dynamic process we study. We show that the strategy distributions correspond to the limit of Bayesian updating of beliefs and best replies when the number of agents in each generation grows to infinity. Finally, Section 7 extends the analysis to settings with a continuum of actions, so small changes in the paradigm can induce different actions even when the original equilibrium action was the unique best reply. We show that in this setting the key concept for stability is the derivative of the indirect utility function of the agent, and we reconcile this finding with the previous analysis by showing how Berk-Nash equilibria with continuum of actions are approximated by mixed equilibria on a fine grid of finite actions.

---

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

1.1 Related work

The analysis of misspecified learning began with Berk (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 (1973) gives the first general framework for this problem, and Nyarko (1991) points out that the combination of misspecification and endogenous observations can lead to cycles. Esponda and Pouzo (2016) defines Berk–Nash equilibrium, which relaxes Nash equilibrium by replacing the requirement that players' beliefs are correct with the requirement that each player's belief minimizes the KL divergence from their observations on the support of their prior.[5]

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 (2020), Frick, Iijima, and Ishii (2020b), and Fudenberg, Lanzani, and Strack (2021). Esponda, Pouzo, and Yamamoto (2020) uses stochastic approximation to establish when the agent's *action frequency* converges in an environment with finitely many actions. Frick, Iijima, and Ishii (2020b) provides conditions for local and global convergence of the agent's beliefs without explicitly modelling the agent's actions. Fudenberg, Lanzani, and Strack (2021) 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.[6]

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

---

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

[6]Earlier papers such as Fudenberg, Romanyuk, and Strack (2017), Heidhues, Kőszegi, and Strack (2018), Heidhues, Koszegi, and Strack (2020), and Molavi (2019) analyze the dynamics of misspecified learning in specific applications or examples. Bohren (2016), Bohren and Hauser (2020), and Frick, Iijima, and Ishii (2020a) and He (2019) consider misspecified social learning where all agents have the same form of misspecification. In one of He (2019)'s models, each period a large generation of agents learns from the data generated by the previous generation.

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).[7]

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).

---

[7]Also, Blume and Easley (1992), Blume and Easley (2002), and Blume and Easley (2006) study models of the evolution of the wealth shares of agents with different utility functions, priors, or updating rules in rational expectations market equilibria. Farther afield, King and Masel (2007) and Fudenberg and Imhof (2012) study the evolutionary selection of mutation rates in stochastic environments, and Nolfi, Parisi, and Elman (1994) and Belew (1990) study the joint effect of evolution and learning on neural networks.

2 The single-agent problem

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

2.1 Static Model

Actions, utility, and data generating process An agent chooses an action a from the finite set A after observing a signal s from the finite set S. The agent then observes an outcome $y \in Y \subseteq \mathbb{R}^m$. 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 the pure strategies of the agents, i.e., the maps from signals to actions, by $\Pi = A^S$.[8] 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 $\pi : S \to A$:

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

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 θ, where each θ is associated with a family of probability measures $Q_\theta(\cdot|s, a)$, one for each signal-action pair (s, a).[9] The agent's initial uncertainty about the value of the parameter is described by a *belief* $\mu \in \Delta(\mathcal{H})$ with support Θ. A *subjective model* for an agent is the particular subset of parameters $\Theta \subseteq \mathcal{H}$ they consider possible.

We do not require that the agent is *correctly specified*, i.e. that there is a $\theta^* \in \Theta$ such that $Q^* = Q_{\theta*}$, or even that Q^* can be approximated by Q_θ for some $\theta \in \Theta$. We do allow

---

[8]For simplicity we do not allow individual agents to randomize. However, the effect of randomization can be captured by allowing different agents with the same belief to play different actions as long as all of those actions maximize their subjective payoff.

[9]If 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.

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 θ 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 utility function u and beliefs μ determine their subjective expected utility function

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

With a small abuse of notation, we let $U_\theta = U_{\delta_\theta}$ where δ_θ is the Dirac measure on θ. We let $BR(\mu) = \operatorname{argmax}_{\pi \in \Pi} U_\mu(\pi)$ denote the set of pure best replies to μ, and for every $C \subseteq \Delta(\Theta)$, we let $BR(C) = \bigcup_{\mu \in C} BR(\mu)$.

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 the joint distribution $Q_\psi \in \Delta(S \times A \times Y)$ induced by the 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

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

7

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. This is the force behind Berk (1966)'s result that as sample size grows, beliefs concentrate on the parameters that minimize the Kullback-Leibler divergence from the true distribution.[10]

The likelihood of an outcome under true distribution Q^* depends on both the action and the signal, so to apply Berk's result to our model, we weight the realizations by their objective probability:

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

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

$$\Theta(\psi) = \operatorname*{argmin}_{\theta \in \Theta} H_\psi(Q^*, Q_\theta).$$

We call these parameters the *KL minimizers* for ψ. We assume that the agent's posterior after observing Q_ψ is a probability distribution supported over $\Theta(\psi)$. Proposition 6 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_ψ.

2.2 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 ψ.

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

(i) *Pure* if $\psi = \pi$ 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 $\operatorname{supp} \psi = BR(\Delta(\Theta(\psi)))$.
(iv) *Uniformly strict* if $\psi = \pi = BR(\mu)$ for every $\mu \in \Delta(\Theta(\psi))$.

---

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

When an equilibrium is not pure we will call it "mixed," but it is important to recall that we do not allow agents to randomize— mixed equilibria here correspond to different agents playing different strategies. Esponda and Pouzo (2016) defines the unitary version of Berk-Nash equilibrium, which requires that each of the equilibrium strategies can be rationalized with the same belief. Note that unitary equilibria need not be pure, it is sufficient that all of the strategies played are best responses to the same belief μ. Non-unitary 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 (2021), such multiple minimizers can arise from symmetry constraints or the use of low-dimensional functional forms. 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. Uniformly strict equilibrium (Fudenberg, Lanzani, and Strack, 2021) requires the equilibrium strategy to be a strict best reply to all of the KL-minimizing parameters; such equilibria are clearly quasi-strict.

A *self-confirming equilibrium*[11] is a Berk-Nash equilibrium (Θ, ψ) such that there is a $\theta \in \Theta$ with $Q_\theta(\cdot|s, \pi(s)) = Q^*(\cdot|s, \pi(s))$ for all $\pi \in \operatorname{supp} \psi$ and $s \in S$. Self-confirming equilibrium requires that the subjective model of the agents contains at least one parameter that induces the same distribution over observables as the equilibrium does. These equilibria do not trigger a change of paradigm, as confirmed by Corollary 2 below.

2.3   Model expansions

We will consider two ways that mutated agents might expand the set of models they consider possible. The first is through a quantitative local expansion that allows for all the parameter values within ε of a parameter their parents thought was possible. To define this, we let $B_\varepsilon(\theta) = \{\theta \in \mathcal{H} : ||\theta - \theta'||_2 \leqslant \varepsilon\}$.

The second applies when the subjective model is described by a finite collection of hypotheses about the underlying parameter that are expressed in the form of quantitative statements. For example, agents might 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. The hypotheses can also take the form of joint restrictions on the parameters, as with an agent who believes that two variables are independent. These hypotheses describe the parts of the subjective model of an agent that can be separately relaxed

---

[11]C.f. Fudenberg and Levine (1993).

by a mutation. 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 $\Theta = \{\theta \in \mathcal{H} : f_i(\theta) \geqslant 0, \forall i \in \{1, ..., m\}\} := \Theta(\mathcal{F})$.

Definition. [Model expansions]
  (i) The subjective model Θ_ε is the ε *expansion of* Θ if $\Theta_\varepsilon = \bigcup_{\theta \in \Theta} B_\varepsilon(\theta)$.
  (ii) The subjective model Θ^l is a *one-hypothesis relaxation of* $\Theta(\mathcal{F})$ in hypothesis $l \in \{1, ..., m\}$ if $\Theta^l = \Theta(\mathcal{F}\backslash\{f_l\})$.

3 Illustrative Examples

Our first example considers two versions of the problem of a misspecified monopolist. The two versions of this example have the same payoff function for the monopolist and the same true demand function,they have specifications of the agent's subjective model. The first version shows that even a unique and isolated equilibrium may not be resistant to mutations. In the second 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 probability to the suboptimal action.

Example 1. *The monopolist chooses price $a \in \{2, 10\}$ 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 ω is a standard normal shock. The true demand function is given by $(l^*, i^*) = (4, 42)$.*

a. *Suppose that the subjectively possible parameter values are $[3/2, 5/2] \times [28, 32]$, as in the example of Nyarko (1991). The unique Berk-Nash equilibrium assigns probability $1/4$ to price 2, sustained by a Dirac belief on $\theta_\alpha = (5/2, 30)$. As shown in Figure 1, the binding constraint of the KL minimization problem is $l \leqslant 5/2$, so the KL-minimizer of the ε expansion of the subjective model mostly adjusts the slope upwards, which lead the mutants to choose action 2. Since action 2 performs better than mixing, this Berk-Nash equilibrium is not be resistant to local mutations.[12]*

b. *The second 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 $[3, 10/3] \times$*

---

[12]In Nyarko's version of the example, both the subjective model and the correct data generating process differ from those in Esponda and Pouzo. To emphasize the role of the subjective model in determining the stability of the equilibrium in an otherwise identical objective environment, we transposed Nyarko's example so both examples have the same data generating process.
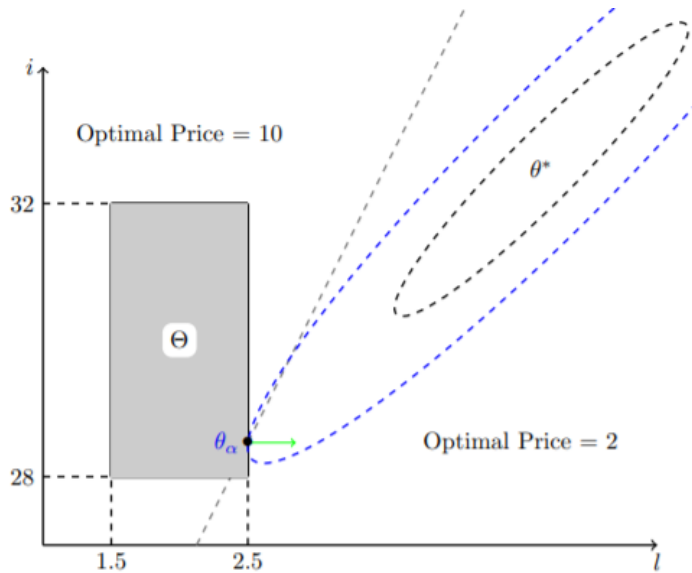
Figure 1: The ellipses are level curves for the KL divergence in the unique equilibrium of part *b*; the green arrow points in the direction of the greatest KL improvement.

[33, 40]. *There is a continuum of mixed Berk-Nash equilibria, indexed by the probability of price* 2 *in* $[7/8, 35/36]$, *sustained by a Dirac belief on the KL-minimizing parameter* $\theta_\alpha = (10/3, 40)$. *In all these equilibria, both the slope and intercept constraints are binding. Because the average likelihood of the realized quantities depends on the probabilities that each price is charged, so does the KL minimizer for a slightly enlarged subjective model. Specifically, when the low price is charged almost all the time, the main unexplained feature in the equilibrium data is high demand, so the KL minimizer for an* ε *expansion of the subjective model revises the intercept upward. This includes the original equilibrium in Esponda and Pouzo (2016) where the low price is charged with probability* 35/36, *illustrated in Figure 2. 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. When both actions are played with sufficiently high probability* $(\psi(2) < \frac{97}{100})$, *the main unexplained feature is high price sensitivity instead of generally high demand: for price* 2 *demand is higher than expected and for price* 10 *it is lower. Therefore, the mutants revise the slope upward. This evidence suggests that the agent should change their belief about the slope of the demand curve. This makes* 2 *the subjectively optimal action, which yields a larger payoff. As we will see, the first sort of equilibrium resists local mutations—that is, local expansions of* Θ— *and the second type does not.*
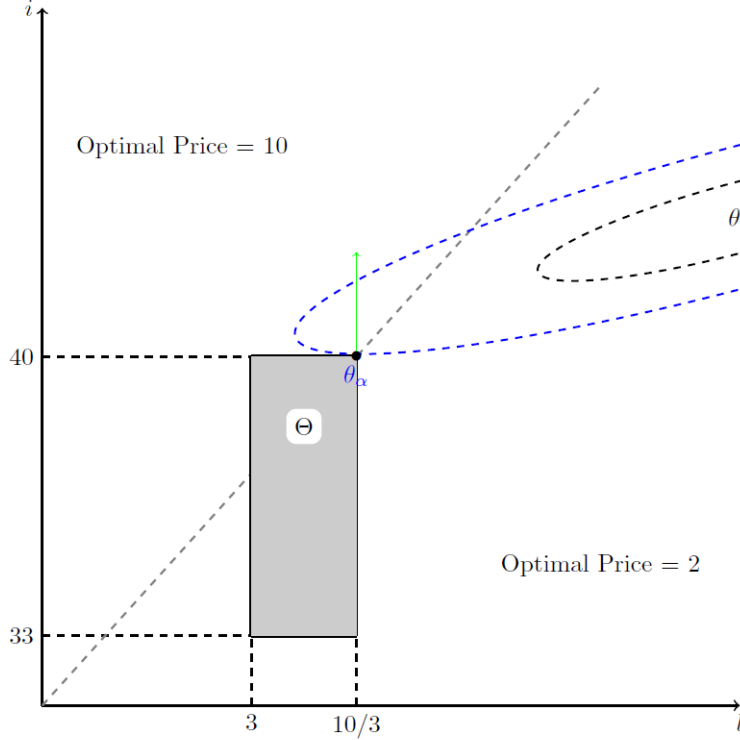
Figure 2: The ellipses are level curves for the KL divergence in the equilibrium of part a where $\psi(2) = 35/36$; the green arrow points in the direction of the greatest KL improvement.

The next example shows that a qualitative relaxation of the subjective model can lead to overadjustment of beliefs in the direction of the relaxed constraint. This shift may induce overshooting the optimal action, and lead to a lower payoff than before. In particular, we consider an equilibrium in which the agents misperceive a progressive tax schedule as linear, and end up exerting an excessive effort, as in an example of Esponda and Pouzo, 2016. Mutated agents who realize that the tax schedule might be progressive overestimate its convexity, because they assess it at the equilibrium action. This overestimate leads to very low effort, which yields less than the equilibrium payoff, so the equilibrium resists the mutation.

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

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

12

.

The agent obtains payoff $u(a, (z, x)) = z - x - c(a)$, so their optimal action is 4.

Their subjective model of the tax schedule is $\tau_\theta(z) = (\theta_1 + \varepsilon)z + (\theta_2 + \varepsilon)z^2$, where the first component of the parameter $(\theta_1, \theta_2) \in \mathbb{R}^2$ capture how steep is the tax schedule, and the second captures how progressive it is.

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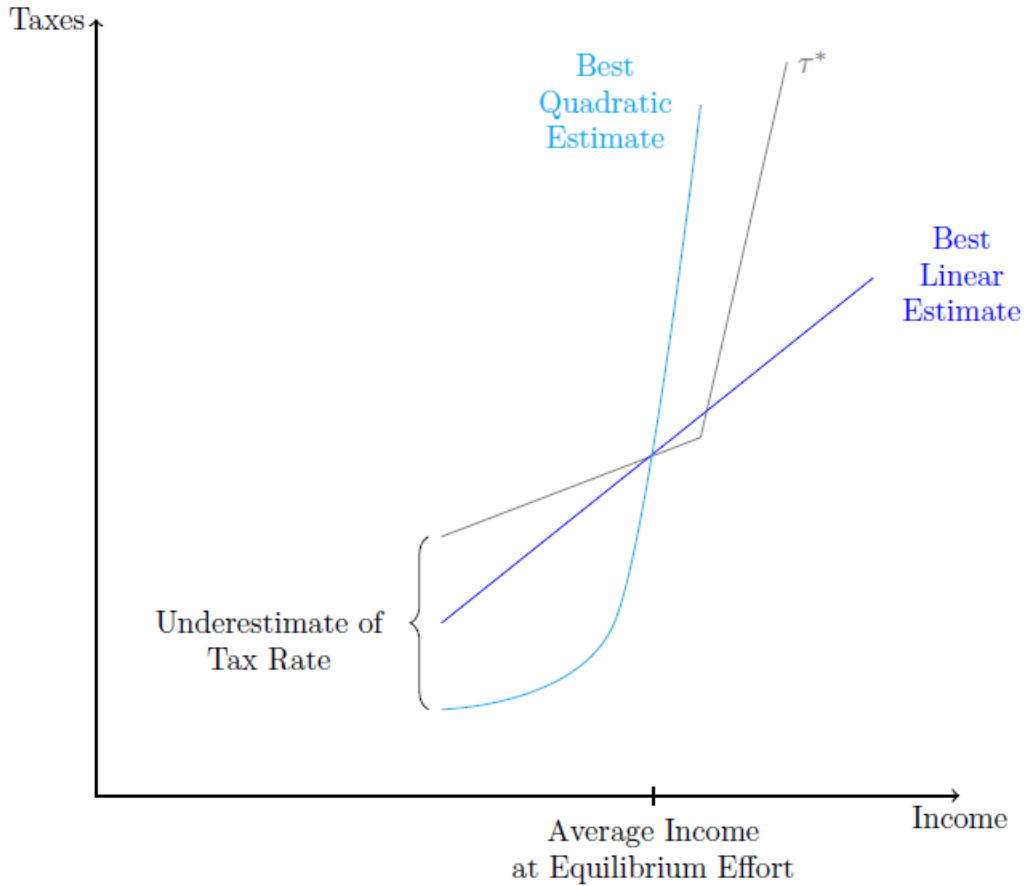

Figure 3: Misspecified Taxation Schedule

An agent who drops the linearity assumption and shifts to the subjective model $\Theta' = \mathbb{R} \times \mathbb{R}_+$ estimates a very high quadratic term: The equilibrium action makes the average income very close to the shift point between the two brackets, so that the agent observes high progressivity. Their quadratic subjective model extrapolates this progressivity as a global feature of the tax schedule (instead of a kink in the tax schedule) which leads them to choose the minimal action 3. The objectively optimal action 4 is lower than the equilibrium action 5, but the mutated agent overshoots the optimum. For this reason, the equilibrium is resistant to "one-hypothesis" mutations.

4 Evolutionary Dynamics and Steady States

4.1 States and State Evolution

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

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 Θ 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:

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

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

The reason that $p^{t+1}(\cdot|\Theta)$ takes values in $\Delta(BR(\Delta(\Theta(p_\Pi^t))))$ as opposed to the smaller set $BR(\Delta(\Theta(p_\Pi^t)))$ is that different agents with the same subjective model may play different best responses: They may have different beliefs over the KL minimizers when $\Theta(p_\Pi^t)$ is not

a singleton, and multiple strategies may be best replies to the same beliefs. We provide an explicit learning foundation for this process in Section 6 under the assumption that either there is a unique best reply to the KL minimizers (which covers the case of a uniformly strict Berk-Nash equilibrium) or that Θ is finite.

To handle situations where agents with the same Θ play different strategies, we extend U^* to mixed strategies in the standard way: Given a distribution of strategies $p(\cdot|\Theta)$ in the subpopulation with subjective model Θ, we have $U^*(p(\cdot|\Theta)) = \sum_{\pi \in \Pi} p(\pi|\Theta)U^*(\pi)$.

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

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

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

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

A sequence $(p^t)_{t \in \mathbb{N}_0} \in P^{\mathbb{N}_0}$ is a *solution* if there is an evolutionary map such that for all $t \in \mathbb{N}_0$ equations (1) and (2) hold. These dynamics can be interpreted as the result of biological reproduction or as the result of social learning and imitation.[13] 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. This is the reason why the evolutionary map depends on the current average performance of the different paradigms.

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

---

[13]A leading example is the discrete-time replicator dynamics (Hofbauer and Sigmund (1988), Hines (1980), Dekel and Scotchmer (1992), and Cabrales and Sobel (1992)).

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.[14]

4.2   Steady states

We focus on the rest points of the evolutionary dynamics.

Definition. A *steady state* is a $\hat{p} \in P$ such that $(\hat{p})_{t \in \mathbb{N}_0}$ is a solution and $\hat{p}_{\mathcal{K}} = \delta_{\Theta}$ for some $\Theta \in \mathcal{K}$. A steady state is *unitary* if $\hat{p}(\cdot|\Theta) \in \Delta(BR(\mu))$ for some $\mu \in \Delta(\Theta(\hat{p}_{\Pi}))$.

Note that in a steady state all agents have the same subjective model. Almost all of the literature on misspecified learning has focused on this case,[15] and Lemma 1 shows that they correspond to Berk-Nash equilibrium.

In a 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.

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

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.[16]

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

---

[14]If these features were inherited, some Berk-Nash equilibria would not be steady states.

[15]Bohren and Hauser (2020) is an exception.

[16]Esponda and Pouzo (2016) assumes Y is finite, but this is not needed for the proof of their Theorem 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 ε *mutation of a steady state* $\delta_\Theta \times \psi$ *to* Θ' if
  (i) $p_\mathcal{K} = (1 - \varepsilon)\delta_\Theta + \varepsilon \delta_{\Theta'}$ and
  (ii) $p(\cdot|\tilde{\Theta}) \in \Delta(BR(\Delta(\tilde{\Theta}(\psi)))) \qquad \forall \tilde{\Theta} \in \{\Theta, \Theta'\}.$

Note that both the mutated and unmutated agents choose their actions based on the same data, namely the distribution of play that prevailed before the mutation occurred.

Definition. A Berk-Nash equilibrium (Θ, ψ) *resists mutation to* Θ' if there is a collection of solutions $(p_\varepsilon^t)_{t \in \mathbb{N}_0, \varepsilon \in (0,1)}$, such that p_ε^0 is the ε mutation of $\delta_\Theta \times \psi$ to Θ', and

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

For every given ε, the inner limit gives the long-run strategy distribution following an ε mutation to Θ'; the outer limit sends the fraction of mutated agents to 0. The equilibrium (Θ, ψ) is resistant to this mutation if this iterated limit converges back to ψ for some solution that starts from the ε mutation to Θ'.[17]

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 induce a different best response to the evidence generated from equilibrium play, so the distribution over strategies does not change.

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

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

---

[17]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 self-confirming equilibria that are not Nash equilibria, even when the agent is correctly specified.

subjective model that generates a worse fit to the data can generate higher payoffs, as when a point mutation to a misspecified model leads the agent to play the optimal strategy. However, if Θ' does not drop the KL-minimizing parameters, it cannot destabilize the equilibrium.[18] For this reason, 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 (Θ, ψ) resists mutation to Θ'.*

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

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

This follows immediately from the definitions, as in a self-confirming equilibrium the subjective model perfectly matches the observed distribution, so the KL divergence between the agent's beliefs and observations is 0.

Conversely, if an explanation-improving mutation to Θ' makes the agent correctly specified, and the map from Θ' to the observations generated under the equilibrium strategy is injective, i.e., Θ' is "strongly identified" (Esponda and Pouzo (2016)) then the mutants will perform at least as well as their parents. In this case a strategy can only resist mutation to Θ' if 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.

---

[18]We illustrate these two facts in Appendix B.5. 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 an initial mutation can be the springboard for subsequent mutations.

5.1 Local mutations

In an ε local mutation, a fraction ε of the agents in the new generation reacts to unexplained evidence by considering a moderately more permissive paradigm.[19]

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

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

Proposition 2.
 (i) *Every uniformly strict Berk-Nash equilibrium is resistant to local mutations.*
 (ii) *Every Berk-Nash equilibrium $\Theta \times \psi$ in which $\Theta(\psi)$ is in the interior of $\mathcal{H}$ is resistant to local mutations.*

In a uniformly strict equilibrium, the equilibrium strategy π is a strict best response to every parameter that minimizes the KL divergence given that π 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 π. Therefore, both mutated and conformist agents continue to play π at every period, which yields Proposition 2(i). Part (ii) shows that if none of the restrictions imposed by the model is binding, the equilibrium is resistant to local mutations. It follows from the fact that when the KL minimizers are in the interior of Θ, they have a strictly lower divergence than any parameter on or near the boundary of Θ.

Proposition 2 reinforces the finding of Fudenberg, Lanzani, and Strack (2021) 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 π as the unique best reply. If the mutation leads to a KL minimizer that is near θ', the mutated agents may start to play a different strategy than π, 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 steady state $\delta_\Theta \times \psi$ and an $\varepsilon \in \mathbb{R}_{++}$ we define the *most improving parameters* as

---

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

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

These are the parameters that generate the largest decrease in H.[20] The identity of the most improving parameters depends on the initial paradigm and also on the equilibrium distribution of strategies: Ceteris paribus, the most improving parameters are more sensitive to the the consequences of strategies played by more of the of agents.

By an argument paralleling that of Berk, small mutations induce beliefs that are concentrated on the most improving parameters. We are going to show that 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 play converges back to the original equilibrium. Let

$$\Pi_{\mathcal{M}_{\Theta,\psi}} = \limsup_{\varepsilon \to 0} BR(\Delta(\mathcal{M}_{\Theta,\psi}(\varepsilon)))$$

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

Proposition 3. *Let (Θ, ψ) be a Berk-Nash equilibrium.*
(i) If for every $\pi \in \Pi_{\mathcal{M}_{\Theta,\psi}}$, $U^*(\pi) > U^*(\psi)$, then (Θ, ψ) is not resistant to local mutations.
(ii) If for some $\pi' \in \Pi_{\mathcal{M}_{\Theta,\psi}}$, $U^*(\pi') \leqslant U^*(\psi)$, and (Θ, ψ) is a quasi-strict equilibrium, (Θ, ψ) is resistant to local mutations.

To illustrate the role of the payoff comparisons in Proposition 3, we revisit the monopoly pricing problems of Example 1. In the equilibrium of Example 1a, only the slope constraint is binding. An upward revision of the slope makes the low price the unique optimal choice, and since the low price performs better than mixing, by Proposition 3(i) this Berk-Nash equilibrium is not resistant to local mutations. In Example ??b, the KL minimizing parameters underestimate both the intercept and the slope of the demand curve. Here a Berk-Nash equilibrium resists local mutations if and only if the (optimal) low price is played by a large fraction of agents, so that the average loss induced the misspecification is small. When almost all the agents choose a low price, mutants revise the intercept upward, which induces the high price. This action performs worse than the equilibrium, so by Proposition 3(ii) the equilibrium is resistant to local mutations. When both actions are played with suffi-

---

[20]In general $\mathcal{M}_\Theta(\varepsilon)$ need not be a singleton, but it is a singleton for sufficiently small ε in the examples we analyze. Moreover, as we show in Lemma 4, it is always a singleton for sufficiently small ε if $\Theta(\psi)$ is unique and Q_θ is linear in θ.

ciently high probability, the agents revisions induce the objectively optimal high price, and by Proposition 3(i) the equilibrium is not resistant to local mutations.

Example 10 shows that the quasi strictness assumption in part (ii) of the Proposition is needed. Without it, the feedback gathered from mutated agents playing a strategy that is not used in equilibrium may change the behavior of the old population, even if they were performing better than the mutants.

In Example 1, 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 3. [Monopoly pricing as a Bandit Problem] *A monopolist faces a single consumer in every period, and chooses 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: $\mathcal{H} = [0,1]^2$ with typical element $\theta = (\theta_{a_h}, \theta_{a_l})$, where θ_a is the probability that the consumer buys when the price is equal to a, and the true parameter is θ^*. As in Rothschild (1974), the monopolist's belief is a product measure with support Θ, so that the monopolist perceives their choice as a bandit problem.*

Let (Θ, ψ) 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 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 is a smooth function of the parameters, there is a convenient way to check the conditions on $\Pi_{\mathcal{M}_{\Theta,\psi}}$ in

Proposition 3. Given a strategy distribution ψ and $v \in \mathbb{R}^k$, let

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

be the ψ-weighted directional derivative of the relative entropy H in direction v at θ. As we show in Lemma 5 in the Appendix, it is enough to compute the direction (θ, v) in which $D_\psi(\theta, v)$ is maximal, and look at the best replies to the parameters along that direction.

We now introduce an environment where small enlargements of the subjective model never induce worse strategies. In these environments, 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.

Definition. The environment is *directional* if we can order the space of strategies $(\Pi, >)$ and parameters $(\mathcal{H} \cup \{\theta^*\}, \geqslant)$ so that:

(i) The order on the parameter space is continuous and for each $\theta \in \Theta$ and $\varepsilon \in \mathbb{R}_{++}$ there exist $\theta', \theta'' \in B_\varepsilon(\theta)$ such that $\theta' > \theta > \theta''$;

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

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

$$H(Q_\theta\left(\cdot|s, \pi(s)\right), Q_{\theta'}\left(\cdot|s, \pi(s)\right)) < H(Q_\theta\left(\cdot|s, \pi(s)\right), Q_{\theta''}\left(\cdot|s, \pi(s)\right))$$

and

$$H(Q_{\theta''}\left(\cdot|s, \pi(s)\right), Q_{\theta'}\left(\cdot|s, \pi(s)\right)) < H(Q_{\theta''}\left(\cdot|s, \pi(s)\right), Q_\theta\left(\cdot|s, \pi(s)\right));$$

(iv) $Q_{\theta*} = Q^*$, and either $\theta \leqslant \theta^*$ for all $\theta \in \Theta$ or $\theta \geqslant \theta^*$ for all $\theta \in \Theta$.

The simplest case of a directional environment is the one in which $\mathcal{H}$ is a subset of the real line endowed with the usual order. Condition (i) is a generalization that allows multidimensional parameters, where each dimension measures the performance of a different action.

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 $\underline{\theta} \in$ $\mathrm{argmin}\{\theta \in \Theta\}$. Then directionality of the environment implies that the most improving

parameter is slightly lower than $\underline{\theta}$, 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 θ^*, so the equilibrium is not resistant to local mutations.

Because we work with a finite set of actions, some pure equilibria with misspecified beliefs can resist local mutations in directional environments: Even though the mutated agents have more accurate beliefs, this may not be sufficient to induce them to change their actions. Section 7 shows that only self-confirming equilibria can resist local mutations in directional environments with a continuum of actions.

5.2   One-hypothesis Mutations

One-hypothesis mutations 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. [21]

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

We say that a Berk-Nash equilibrium $(\Theta(\mathcal{F}), \psi)$ is *resistant to one-hypothesis mutations* if it resists every one-hypothesis ε mutation for sufficiently small ε. Note that there may be multiple sets of hypotheses that determine the same set Θ. Which one the agent uses—that is, how they mentally encode the set of possible models— will determine which equilibria are resistant to one-hypothesis mutations.

Given a steady state $p = \delta_\Theta \times \psi$ the *l-agnostic KL minimizers* are $\mathcal{P}_l(p)\colon = \operatorname{argmin}_{\theta \in \Theta^l} H_\psi(Q^*, Q_\theta)$, and $\Pi_{p,l} = BR(\Delta(\mathcal{P}_l(p)))$ denotes the set of best replies after the relaxation of hypothesis l.

Proposition 5. *Let $(\Theta(\mathcal{F}), \psi)$ be a Berk-Nash equilibrium.*
  (i) *If for some $l \in \{1, ..., k\}$, $U^*(\pi) > U^*(\psi)$ for every $\pi \in \Pi_{p,l}$, then (Θ, ψ) is not resistant to one-hypothesis mutations.*
  (ii) *If for every $l \in \{1, ..., k\}$ there is a $\pi' \in \Pi_{p,l}$ such that $U^*(\pi') \leqslant U^*(\psi)$, and either $\pi' \in \operatorname{supp} \psi$ or ψ is a uniformly strict Berk-Nash equilibrium, then (Θ, ψ) is resistant to one-hypothesis mutations.*

In the monopoly pricing problem of Example 1a, the level curve for the KL divergence has a downward slope at equilibria with $\psi(2) < 35/36$, so a one-hypothesis mutation that

---

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

allows for larger intercepts leads agents to play the objectively optimal price 2. Thus by part (i) of Proposition 5 these equilibria are not stable to one-hypothesis mutations. Example 8 in the Online Appendix shows that an equilibrium can be stable to one-hypothesis mutations and not stable to local mutations, so that the two stability conditions are not nested.

The next example considers a buyer who has correlation neglect: they do not understand that the price charged by a seller is positively correlated with the value of the good. In the first version of the example, the preferences of the buyer and the seller for the good are the same (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 introduce a stochastic misalignment between the preferences by adding a noise term to the buyer's evaluation, 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 ω in a double auction with price at the buyer's bid, so the seller sets their bid x equal to their value. The value ω is good, $\omega = 3$, with probability 1/3, bad, $\omega = 2$, with probability 1/2 and very bad, $\omega = 1$, with probability 1/6. 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., $a \geqslant x$, and $\tilde{\omega} = \#$ if the transaction fails, i.e., $a < x$.*

Here a parameter θ is a vector $(p_1, p_2, p_3, F(1|1), F(2|1), F(1|2), F(2|2), F(1|3), F(2|3))$, where p_i is the probability that the seller's asking price asked is equal to i and $F(i|j)$ is the probability that the value is lower or equal than i given that the seller asked price j. So $\mathcal{H}$ is the subset of $\mathbb{R}^9$ such that $\sum_{i=1}^3 p_i = 1$, and $0 \leqslant F(1|i) \leqslant F(2|i) \leqslant 1$ for $i \in \{1, 2, 3\}$. The true parameter is

$$\theta^* = \begin{pmatrix} p_1^* = 1/6, p_2^* = 1/2, p_3^* = 1/3, \\ F(1|1) = 1, F(1|2) = F(1|3) = 0, \\ F(2|1) = F(2|2) = 1, F(2|3) = 0, \end{pmatrix}$$

so the objectively optimal strategy is to bid 3 for any $s \in \{-1, 0, 1\}$. However, the agent believes that price and value are independent.

Suppose first that the noise term is degenerate, with $s = 0$. Here the strategy $a = 2 + s = 2$ is a Berk-Nash equilibrium and the KL-minimizing parameter is an independent joint probability distribution that is correct about the distribution of seller bids. However, because value is 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 is $\{(1, \frac{1}{6}); (2, \frac{5}{6}); (3, 0)\}$.

This equilibrium is resistant to one-hypothesis mutations. Because 3 is not accepted, a one-dimensional mutation is explanation improving only if it involves the relation between the conditional value distributions for 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 in the equilibrium.

Now suppose that the signal s is a mean-zero shock uniformly distributed over $\{-1, +1\}$, independent of ω. 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 $\{(1, \frac{1}{5}); (2, \frac{3}{5}); (3, \frac{1}{5})\}$ is too pessimistic, leading to the (objectively suboptimal) bid of price 2 after signal $s = -1$.

This equilibrium is uniformly strict, so by Proposition 2 it is resistant to local mutations. 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:

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

Agents with this model have a posterior concentrated on

$$\hat{\theta} = \begin{pmatrix} p_1 = 1/6, p_2 = 1/2, p_3 = 1/3, \\ F(1|1) = F(1|2) = F(1|3) = 1/6, \\ F(2|1) = F(2|2) = 1, \\ F(2|3) = 2/3. \end{pmatrix}$$

Since $BR(\hat{\theta}) = \{3\}$, by Proposition 5 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)(GRS) is that an agent who is certain that the true parameter is in Θ will only pay attention to the frequency of events that help distinguish between parameters in Θ that induce different best replies.[22] They motivate this assumption with reference to several examples of persistent sub-optimization, such as Indonesian seaweed farmers who persistently 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, GRS uses the concept of "attention partitions," which are defined on a finite set Y of outcomes; we will likewise restrict to finite Y in this subsection.

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

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

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

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

Gagnon-Bartsch, Rabin, and Schwartzstein (2020) assumes that actions do not influence the distribution over outcomes, i.e., $Q_\theta(\cdot|s, a) = Q_\theta(\cdot|s, a')$ and $Q^*(\cdot|s, a) = Q^*(\cdot|s, a')$ for all $\theta \in \mathcal{H}, a, a' \in A$ and $s \in S$. Under this assumption, whether an attention-improving mutation is triggered is independent of the current equilibrium distribution, as is the inference made by the decision maker in the enlarged model Θ'. We will use our model to explore the impact of relaxing that assumption. GRS also restricts attention to unitary equilibria, and consider only the change from the initial paradigm Θ to $\Theta \cup \theta^*$, so that the "mutants" add the true model θ^* to the support of their prior. This leads to the following definition:

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

---

[22]The paper considers other ways of allocating attention, but focuses on this one.

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

If actions do not influence the distribution over outcomes, an objectively suboptimal equilibrium is resistant to a mutation to the correctly specified model if and only if it is attentionally stable, so there is no role for evolutionary forces. That an attentionally stable equilibrium is resistant is an immediate consequence of Proposition 1: If the alternative subjective paradigm does not improve the explanation it is not going to destabilize the equilibrium. Conversely, if the equilibrium is not attentionally stable, the correctly specified model provides a better explanation of the observed outcomes than the old subjective model. Thus since actions do not change the outcome distribution, a correctly specified agent is able to infer the objectively optimal action from the equilibrium observables, so if the equilibrium had a suboptimal payoff it is destabilized by a mutation that adds the correct model $\theta*$ to Θ .

If actions can change the outcome distribution, and we replace the GRS assumption of a point mutation to the true model with other sorts of mutations, 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 next example, an equilibrium that resists local attention-partition mutations does not resist more general local mutations. Example 9 in the Online Appendix shows that the attention partition can also have the opposite effect.

Example 5. [Channeled Attention and Overconfidence] *An agent can choose between a hard and easy task, $A = \{a_h, a_e\}$. The agent's utility is their performance, i.e., $u(a, y) = y$. They believe that their performance $\{y_L = 1, y_M = 2, y_H = 3\}$ in each task is a function of their skill κ and a relative difficulty parameter ρ. The initial prevailing paradigm is dogmatic about the difficulty level $1/5$, and posits that the skill is $\kappa = 1/3$ or $\kappa = 0$, so $\Theta = \{1/5\} \times \{0, 1/3\}$. Moreover, the agent believes that the frequency of high outcomes is a sufficient statistic for the skill, so one minimal attention partition for Θ is $\Theta_{\mathcal{Y}} = \{\{y_H\}, \{y_M, y_L\}\}$. Section B.3 shows that the unique Berk-Nash equilibrium has $\psi(h) \approx 0.56$, sustained by the belief $\mu((1/5, 0)) = 1/3, \mu((1/5, 1/3)) = 2/3$.*

By Lemma 5 and Proposition 3, the most improving parameter for the partition $\Theta_{\mathcal{Y}}$ is a slight variation of the low skill parameter $\mathcal{M}(\varepsilon) = (1/5, \varepsilon)$ and the mutated agents adopt the suboptimal action a_h, so that they will eventually die out. Intuitively, the minimal attention

partition $\Theta_{\mathcal{Y}}$ 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 is $\mathcal{M}(\varepsilon) = (1/5 + \varepsilon, 1/3)$. Since the best reply to these parameters is the optimal action a_e, the equilibrium would not resist mutations: When the agent tracks every outcome, they realize that failures come from the high difficulty of the hard task rather than low skill.

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.[23] Suppose that all agents with subjective model Θ have the same prior μ_{Θ}.[24] Each agent born in a given period observes an independent sample of n randomly drawn individual experiences from the previous generation, updates their prior using Bayes' rule, and chooses a best reply to their posterior.

Formally, let $p_{\Pi} \in \Delta(\Pi)$ be a distribution over strategies. Each agent $i \in [0, 1]$ is endowed with a measurable best response function $R_i : \Delta(\Theta) \to A$, with $R_i(\nu) \in BR(\nu)$ for all $\nu \in \Delta(\Theta)$.[25]

Each agent observes $n \in \mathbb{N}$ individual experiences drawn from the distribution over strategies p_{Π}, and computes posterior belief μ_n^i:

$$\mu_n^i(B \mid a^n, s^n, y^n) = \frac{\int_B \prod_{l=1}^n q_\theta(y_l|s_l, a_l) d\mu_\Theta(\theta)}{\int_\Theta \prod_{l=1}^n q_\theta(y_l|s_l, a_l) d\mu_\Theta(\theta)}.$$

The agent then best replies to the posterior using the best response function, so aggregate play is given by

$$\psi_n(\Theta, p_\Pi)(\pi) = \int 1_{R_i(\mu_n^i) = \pi} di.$$

We assume that the individual observations are drawn independently across agents, and that the map $i \mapsto R_i$ is measurable.

---

[23]This is only one way to provide a foundation for our model; we provide it to show the plausibility of the dynamics we study. We conjecture that the same steady states would be asymptotic limits if a single agent acted each period, as in He (2019) or Bohren and Hauser (2020).

[24]In Section B.6 of the Online Appendix, we show that in equilibria where the KL-minimizing parameters are identified, as in our examples, the limit belief will be independent of the prior.

[25]Our result covers the case where all agents use the same R but does not impose this restriction.

Let $\mathcal{E}(\theta, \psi)$ be the set of parameters that are indistinguishable from θ if the distribution over strategies is ψ, i.e., those parameters θ' such that $Q_\theta(\cdot|\cdot,\cdot) = Q_{\theta'}(\cdot|\cdot,\cdot)$ almost surely under Q_ψ.

Definition. We say that *strategies are distinguishable under p_Π* if $\forall \pi, \pi' \in \Pi, \pi \neq \pi', \exists \theta \in \Theta(p_\Pi)$ such that $U_{\mu_\Theta(\cdot|\mathcal{E}(\theta,p_\Pi))}(\pi) \neq U_{\mu_\Theta(\cdot|\mathcal{E}(\theta,p_\Pi))}(\pi')$.

Proposition 6. *If either*
  (i) $BR(\Delta(\Theta(p_\Pi)))$ is a singleton, or
  (ii) Θ is finite and strategies are distinguishable under p_Π,
then $\lim_{n\to\infty} \psi_n(\Theta, p_\Pi)$ exists, and is in $\Delta(BR(\Delta(\Theta(p_\Pi))))$.

The proposition shows that if agents observe a sufficiently large number of individual experiences from the previous generation, the aggregate distribution of strategies is an element of $\Delta(BR(\Delta(\Theta(p_\Pi))))$. 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, and it is not needed if the best reply function is continuous in beliefs, as in our extension of Section 7, since in that case convergence in distribution of the beliefs automatically translates in the convergence of the distribution of best replies.

To prove this result, we use an argument similar to those of Berk (1966) and Esponda and Pouzo (2016) to show that the probability assigned to the models that do not minimize the KL divergence goes to 0. We then prove that although beliefs may not converge, their distribution does. We prove this by showing that the vector of likelihood ratios between KL-minimizers is a random walk with positive definite covariance matrix, and applying the central limit theorem to obtain convergence.[26] 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.

---

[26]Fudenberg, Lanzani, and Strack (2021) also combines the properties of a random walk with the properties of the KL-divergence, but they consider a different random walk (the difference between the realized empirical distribution and the true distribution) and do not prove that the distribution of beliefs converges.

7 Infinitely Many Strategies

So far we have assumed there is a finite number of strategies. However, in some applications, 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. We show here how our analysis can be applied to continuum environments.

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

Assumption 3.

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

(ii) S is a subset of a Euclidean space, endowed with a full-support, objective Borel probability measure σ.

(iii) U^* is continuously differentiable in actions and signals.

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

The cardinality of the action space is irrelevant for the analysis of one-hypothesis mutations, and those results extend more or less immediately. For local mutations, the cardinality of the action space does matter: With any fixed set of actions a vanishingly small ε is eventually smaller than the "gap" between the actions, but this is not the case when the action space is a connected interval in $\mathbb{R}$. Instead, in any uniformly strict equilibrium there is a nearby action that performs almost as well, and arbitrarily small changes in beliefs generally induce a change in the best reply. As we show below, this allows local mutations to invade some uniformly strict equilibria in settings with a continuum of actions. We also show that any unstable uniformly strict equilibrium that is an attractor for the dynamic process corresponds to a limit of equilibria that are mixed and unstable along a sequence of increasingly fine finite action grids.[27]

To prove these results, we endow the set of strategies with the L_1 norm. 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 $V(\theta) = U^*(BR(\theta))$.

---

[27]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 $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is a singleton for sufficiently small ε, let $V'(\mathcal{M}_{\Theta,\psi}, \psi) = \liminf_{\varepsilon \to 0} \frac{V(\mathcal{M}_{\Theta,\psi}(\varepsilon)) - V(\Theta(\psi))}{\varepsilon}$ be the derivative of V in the direction of the most improving parameter.

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

The previous result shows that the derivative of the static indirect utility function can be used to check the resistance of an equilibrium to mutation. If the derivative in the direction of the most improving parameter is positive, as is always the case for directional environments, the equilibrium does not resist 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 η is a standard normal, θ_0 is the perceived correlation between performance in the two periods, and θ_a is the perceived influence of action a on performance. Consider a population of instructors that are certain that $\theta_0 = 1$, so that modulo their action, second period performance is centered around that of the first period, with $\Theta = \{1\} \times [-K, +K]^2$. Esponda and Pouzo (2016) shows that for sufficiently large K in the unique Berk-Nash equilibrium the instructor criticizes too often: there is a unique threshold T such that the instructor criticizes if and only if performance is below the threshold $T = (\theta_{a_c}(T) - \theta_{a_r}(T))/k > 0$ where $\theta_{a_c}(T, \theta_0) = E[y - \theta_0 s | s < T] > 0$ and $\theta_{a_r}(T, \theta_0) = E[y - \theta_0 s | s > T] < 0$. Since $\theta_{a_c}(T, \theta_0)$ and $\theta_{a_c}(T, \theta_0)$ are respectively increasing and decreasing in θ_0, the most improving parameter for a sufficiently small ε is $((1, \theta_{a_c}(T), \theta_{a_r}(T)) + (v_0, v_r, v_c))$ 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 $V'((1, \theta_{a_c}(T), \theta_{a_r}(T)) + (v_0, v_r, v_c)) = kT > 0$, by Proposition 7, the equilibrium is not resistant to local mutations.

Section B.3 shows there is a sequence of approximating discrete problems with unique Berk-Nash equilibria, where the instructor criticizes performances below T and praises performance above T, and each equilibrium also is not resistant to local mutations.

Definition. Let S be a singleton. We say that a pure Berk Nash equilibrium $(\Theta, \hat{a})$ is an *attractor* if $\hat{a} \in A^o$ and there is an $\varepsilon > 0$ such $||a - \hat{a}|| \leqslant \varepsilon$ 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.[28] 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 Θ 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 $||A_n - A|| \to 0$

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

8 Conclusion

We used the idea of resistance to mutations to help understand what sorts of misspecification are likely to persist for a long time. If after a mutation the aggregate play converges back to the original equilibrium, we say that the equilibrium is resistant to mutations. This can happen because even explanation improving mutations that lead to a better but imperfect fit in a statistical sense can lead to lower payoffs. 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

---

[28]The assumption of a singleton S guarantees that the space of strategies is unidimensional, so the direction of the deviation from equilibrium completely determines the direction of the best reply to the evidence generated by this deviation.

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 model, we expect that some broadenings of perspective will lead to lower payoffs and thus 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.

A Appendix

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

Conversely, if (Θ, ψ) is a Berk-Nash equilibrium, for every $\pi \in \operatorname{supp}\psi$ there exists $\mu_\pi \in \Delta(\Theta(\psi))$ such that $\pi \in BR(\mu_\pi)$, and so $\psi \in \Delta(BR(\Delta(\Theta(\psi))))$. Therefore, $p_t = \delta_\Theta \times \psi$ for every $t \in \mathbb{N}$ satisfies equations (1) and equation (2), so $\delta_\Theta \times \psi$ is a steady state. The equilibrium is unitary if and only if we can choose $\mu = \mu_\pi$ for all $\pi \in \operatorname{supp}\psi$, so 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 $\lim_{t\to\infty} p_\Pi^t = \psi$, where p_ε is the ε mutation of $\delta_\Theta \times \psi$ to Θ'.

Define p_t by $p_\mathcal{K}^t(\Theta) = 1 - \varepsilon, p_\mathcal{K}^t(\Theta') = \varepsilon$, and $p_\Pi^t(\cdot|\Theta) = \psi = p_\Pi^t(\cdot|\Theta')$ 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 generated by the two models are the same at every period. ∎

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

The proof of this lemma is in the Online Appendix.

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

$$G(\varepsilon) = \min_{\pi' \in \Pi \setminus \{\pi\}} \min_{\mu \in \Delta(\Theta(\pi, \varepsilon))} (U_\mu(\pi) - U_\mu(\pi')).$$

Because Π is finite, U is continuous in μ, and $\varepsilon \mapsto \Theta(\pi, \varepsilon)$ is a continuous and compact-valued correspondence, G is continuous by the Maximum Theorem. And since (Θ, π) is a uniformly strict Berk-Nash equilibrium, $G(0) > 0$, and there is an $\hat{\varepsilon}$ such that if $\varepsilon \leqslant \hat{\varepsilon}$, $G(\varepsilon) > 0$.

Now we show that there is an $\varepsilon' \in (0, \hat{\varepsilon})$ such that if Θ' is an $\varepsilon < \varepsilon'$ local expansion of Θ, then $\Theta'(\pi) \subseteq \Theta(\pi, \hat{\varepsilon})$. 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 Θ'_n is an ε_n local expansion of Θ with $\varepsilon_n \downarrow 0$, and $\theta_n \in \Theta'_n(\pi) \setminus \Theta(\pi, \hat{\varepsilon})$. Since $(\theta_n)_{n \in \mathbb{N}}$ is a sequence in the compact Θ'_1, it has an accumulation point $\theta \in \Theta$. If $\theta \in \Theta(\pi)$ this means that $(\theta_n)_{n \in \mathbb{N}}$ is eventually in $\Theta(\pi, \hat{\varepsilon})$, a contradiction. If $\theta \notin \Theta(\pi)$, then since H is lower semi-continuous in its second argument, (see, e.g., Chapter 1 of Liese and Vajda, 1987) eventually

$$\sum_{s \in S} \sigma(s) H \left(Q^*(\cdot|s, \pi(s)), Q_{\theta_n}(\cdot|s, \pi(s))\right) > \min_{\tilde{\theta} \in \Theta} \sum_{s \in S} \sigma(s) H \left(Q^*(\cdot|s, \pi(s)), Q_{\tilde{\theta}}(\cdot|s, \pi(s))\right)$$
$$\geqslant \min_{\tilde{\theta} \in \Theta'_n} \sum_{s \in S} \sigma(s) H \left(Q^*(\cdot|s, \pi(s)), Q_{\tilde{\theta}}(\cdot|s, \pi(s))\right),$$

a contradiction.

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

(ii) Let $\partial\Theta$ denote the boundary of Θ. By Assumption 1(i) $\partial\Theta$ is a compact set, so when the KL minimizers are in the interior of Θ, there is a $K \in \mathbb{R}_{++}$ and an $\hat{\varepsilon} \in \mathbb{R}_{++}$ such that if

$\theta' \in \partial\Theta$ and θ is in $B_{\hat{\varepsilon}}(\theta')$ around θ' then

$$H_\psi(Q^*, Q_\theta) - \operatorname*{argmin}_{\theta \in \Theta} H_\psi(Q^*, Q_\theta) > K.$$

This in turn implies that $\Theta_\varepsilon(\psi) = \Theta(\psi)$, when $\varepsilon < \hat{\varepsilon}$. Thus the sequence in which both the mutated and the conformist agents continue to play ψ at every period and their shares remain fixed is a solution: equation (1) is satisfied since $\Theta_\varepsilon(\psi) = \Theta(\psi)$, and equation (2) is trivially satisfied since both subpopulations have the same distribution over strategies. Therefore the equilibrium is resistant to local mutations. ∎

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

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

denote the set of states where the strategy used by agents with subjective model Θ' outperforms the strategy used by agents with subjective model Θ by at least γ, and both populations are larger than β.

The proofs of Propositions 3, 5, and 7 use the following lemma.

Lemma 3. *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.*

The proof of this lemma is in the Online Appendix.

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

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

$$||\psi' - \psi|| < \varepsilon^* \Rightarrow U^*(\psi') - \min_{\pi \in \Pi_{\mathcal{M}_{\Theta,\psi}}} U^*(\pi) < -\gamma. \tag{3}$$

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

$$||\psi' - \psi|| < \bar{\varepsilon} \Rightarrow BR(\Delta(\Theta_\varepsilon(\psi'))) \subseteq \Pi_{\mathcal{M}_{\Theta,\psi}}. \tag{4}$$

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

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

that is

$$U^* \left(p_\varepsilon^{t+1}(\cdot | \Theta_\varepsilon) \right) > U^* \left(p_\varepsilon^{t+1}(\cdot | \Theta) \right) + \frac{\gamma}{p_\varepsilon^{t+1}(\Theta)}. \tag{5}$$

But then, applying Lemma 3 with

$$\beta = \min \left\{ p_{\mathcal{K}}^1(\Theta_\varepsilon), 1 - \max_{\psi' : ||\psi - \psi'|| \leqslant \frac{\min\{\varepsilon', \varepsilon^*\}}{2}} \psi' \left(\Pi_{\mathcal{M}_{\Theta, \psi}} \right) \right\},$$

and $\lambda = \frac{\gamma}{p_\varepsilon^0(\Theta)}$ we obtain the contradiction that $\lim_{t \to \infty} p_\varepsilon^t(\Theta_\varepsilon) \to 1$.

(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 for equilibria that are not necessarily quasi-strict: If for some $\pi' \in \Pi_{\mathcal{M}_{\Theta, \psi}}$, $U^*(\pi') \leqslant U^*(\psi)$, and $\pi' \in \operatorname{supp} \psi$, then (Θ, ψ) is resistant to local mutations. Part (ii) of the Proposition follows from the fact that by the upper hemicontinuity of $\Theta(\cdot)$ (see Lemma 2) and $BR(\cdot)$, the requirement $\pi' \in \operatorname{supp} \psi$ is always satisfied in a quasi-strict equilibrium.

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

Let $\hat{\varepsilon} = \min_{\pi \in \operatorname{supp} \psi} \psi(\pi)$. 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 $\lim_{t \to \infty} p_\Pi^t = \psi$, where p_ε is the ε local mutation of p. Fix such an ε and let Θ_ε be the local ε expansion of Θ. Define the candidate solution by letting p_0 be given by $p_{\mathcal{K}}^0(\Theta) = 1 - \varepsilon, p_{\mathcal{K}}^0(\Theta_\varepsilon) = \varepsilon, p_\Pi^0(\cdot | \Theta_\varepsilon) = \pi'$ and

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

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

36

defined as $\frac{p_{\mathcal{K}}^{t+1}(\Theta)}{p_{\mathcal{K}}^{t+1}(\Theta_\varepsilon)} = \frac{T(p^t)(\Theta)}{T(p^t)(\Theta_\varepsilon)}$, $p_\Pi^{t+1}(\cdot|\Theta_\varepsilon) = \pi'$ and

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

By definition, $(p^t)_{t \in \mathbb{N}_0}$ satisfies equation (2). It only remains to show that for every $t \in \mathbb{N}_0$ $\psi(\pi') \geqslant p_{\mathcal{K}}^t(\Theta_\varepsilon)$ 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 Θ_ε is the ε local mutation of Θ, and $\varepsilon < \hat{\varepsilon} = \min_{\pi \in \operatorname{supp}\psi} \psi(\pi)$, $\psi(\pi') \geqslant p_{\mathcal{K}}^0(\Theta_\varepsilon) = \varepsilon$. 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 $\psi(\pi') \geqslant p_{\mathcal{K}}^t(\Theta_\varepsilon)$ by the inductive hypothesis, we have $p_{\mathcal{K}}^{t+1}(\Theta_\varepsilon) < p_{\mathcal{K}}^t(\Theta_\varepsilon) \leqslant \psi(\pi')$. And since $\varepsilon < \varepsilon'$, equation (1) is satisfied for $t + 1$ as well. ∎

Lemma 4. *Let (Θ, ψ) be a Berk-Nash equilibrium. If $\Theta(\psi)$ is a singleton and Q_θ is linear in θ then there is $\varepsilon' \in \mathbb{R}^{++}$ such that $\mathcal{M}_{\Theta,\psi}(\varepsilon)$ is a singleton for all $\varepsilon \leqslant \varepsilon'$.*

Proof of Lemma 4. Let $\{\hat{\theta}\} = \Theta(\psi)$. By Lemma 2 there is $\varepsilon' > 0$ such that for all $\varepsilon \leqslant \varepsilon'$, $\mathcal{M}_{\Theta,\psi}(\varepsilon) \subseteq B_\varepsilon(\hat{\theta})$. Assume by way of contradiction that $\theta', \theta' \in \mathcal{M}_{\Theta,\psi}(\varepsilon)$, $\varepsilon \leqslant \varepsilon'$ $\theta' \neq \theta'$. By the convexity of the norm, $\theta'/2 + \theta'/2 \in B_\varepsilon(\hat{\theta})$, and by the strict convexity of KL divergence in the second argument, $H_\psi\left(Q^*, Q_{\theta'/2+\theta'/2}\right) < H_\psi\left(Q^*, Q_\theta\right)$, a contradiction. ∎

Let $\mathcal{S}$ denote the sphere of radius 1 in $\mathbb{R}^k$ with respect to the $||\cdot||_2$ norm.

Lemma 5. *Assume that 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 $\theta \mapsto H(Q^*(\cdot|s,a), Q_\theta(\cdot|s,a))$ is twice continuously Gateaux differentiable at $\hat{\theta}$. Let (Θ, ψ) be a Berk-Nash equilibrium. If Θ is finite and $\operatorname{argmax}_{\theta \in \Theta(\psi), v \in \mathcal{S} \cup \{0\}} D_\psi(\theta, v) = (\hat{\theta}, \hat{v})$ is a singleton then $\Pi_{\mathcal{M}_{\Theta,\psi}} \subseteq \limsup_{\varepsilon \to 0} \operatorname{argmax}_{\pi \in \Pi} U_{\delta_{\hat{\theta}+\varepsilon\hat{v}}}(\pi)$.*

Proof. Since $H_\psi(Q^*, Q_\theta)$ is finite at the KL-minimizers, it is Gateaux continuously differentiable at $\hat{\theta}$, and D_{p_Π} is continuous at every $(\hat{\theta}, \hat{v})$. Then, by finiteness of the parameters, for all $\lambda > 0$ there exists an $\bar{\varepsilon} \in \mathbb{R}_{++}$ such that for all $\varepsilon < \bar{\varepsilon}$, $\Theta_\varepsilon(\psi) \subseteq B_\varepsilon(\hat{\theta})$ and for all $v \in \mathbb{R}_k$, $\left| D_\psi(\hat{\theta}, v) - \varepsilon[H_\psi\left(Q^*, Q_{\hat{\theta}+v*\varepsilon}\right) - H_\psi\left(Q^*, Q_{\hat{\theta}}\right)] \right| < \lambda$. Finally, by finiteness of the strategies, if $\hat{a}$ is in $\operatorname{argmax}_{\pi \in \Pi} U_{\delta_{\hat{\theta}+\varepsilon\hat{v}}}(\pi)$, $\hat{a} \in BR(\Delta(\mathcal{M}_{\Theta,\psi}(\varepsilon)))$. ∎

Proof of Proposition 4. Let (Θ, ψ) be a Berk-Nash equilibrium that is neither pure nor self-confirming, and denote its corresponding steady state $p = \delta_\Theta \times \psi$. Since (Θ, ψ) is not self-confirming, there is no θ such that for all $s \in S$, $Q_\theta(\cdot|s, \pi(s)) = Q^*(\cdot|s, \pi(s)) = Q_{\theta*}(\cdot|s, \pi(s))$,

and in particular, $\theta^* \notin \Theta(\psi)$. Since the equilibrium is mixed and the environment is directional, either $\Theta(\psi) \subseteq \{\underline{\theta} : \forall \theta \in \Theta, \underline{\theta} \leqslant \theta\}$, or $\Theta(\psi) \subseteq \{\bar{\theta} : \forall \theta \in \Theta, \bar{\theta} \geqslant \theta\}$. We treat the first case; the other is symmetric. By the directionality assumption, for sufficiently small ε every most improving parameter θ_ε is such that $\theta_\varepsilon < \theta$ for all $\theta \in \Theta$. Let $\underline{\pi}$ be the minimal element of $BR(\Delta(\{\underline{\theta} : \forall \theta \in \Theta, \underline{\theta} \leqslant \theta\}))$, and set $\pi \in \operatorname{supp} \psi \backslash \{\underline{\pi}\}$, $\pi \neq \underline{\pi}$. Since both π and $\underline{\pi}$ are played in equilibrium and $\Theta(\psi) \subseteq \{\underline{\theta} : \forall \theta \in \Theta, \underline{\theta} \leqslant \theta\}$, we have $U_{\hat{\theta}}(\pi) = U_{\hat{\theta}}(\underline{\pi})$ for all $\hat{\theta} \in \{\underline{\theta} : \forall \theta \in \Theta, \underline{\theta} \leqslant \theta\}$. Moreover, $\Pi_{\mathcal{M}_{\Theta,\psi}} = \{\underline{\pi}\}$, and since the environment is directional, $U^*(\pi) < U^*(\underline{\pi})$. Since π was arbitrarily chosen in $\operatorname{supp} \psi \backslash \{\underline{\pi}\} \neq \varnothing$, we also have $U^*(\psi) < U^*(\underline{\pi})$, and the result follows by Proposition 3. ∎

Proof of Proposition 5. By Lemma 2, $\mathcal{P}_l(p)$ is compact for each $l \in \{1, ..., k\}$. Because the action and signal spaces are finite and $\mathcal{P}_l(p)$ is compact, there is $\varepsilon > 0$ such that if $\nu(\{\theta : \exists \bar{\theta} \in \mathcal{P}_l(p), ||\theta - \bar{\theta}|| < \varepsilon\}) > 1 - \varepsilon$ and $\pi \in BR(\nu)$, then $\pi \in \Pi_{p,l}$. Because $\Theta(\cdot)$ is upper-hemicontinuous (see Lemma 2), it follows that there exists $\varepsilon' > 0$ such that if Θ' is a one-hypothesis ε' mutation of Θ,

$$||\psi' - \psi|| < \varepsilon' \Rightarrow BR(\Delta(\Theta'(\psi))) \subseteq \bigcup_{l \in \{1,...,k\}} \Pi_{p,l}. \tag{6}$$

(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 $l \in \{1, ..., k\}$, we have $U^*(\pi) > U^*(\psi)$ for every $\pi \in \Pi_{p,l}$. By continuity of U^* and the finiteness of Π there is $\varepsilon^* > 0$ such that

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

Let Θ_l be the one-hypothesis relaxation of Θ 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

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

Moreover, by Lemma 2 $\mathcal{P}_l$ is upper hemicontinuous.

Suppose by way of contradiction that for $\varepsilon < \min\{\varepsilon', \varepsilon^*\}$ there is an ε mutation to Θ^l, p_ε^0, such that $\lim_{t\to\infty}(p_\varepsilon^t)_\Pi = \psi$. This means that after some $\tau > 0$, for all $t > \tau$, $\|(p_\varepsilon^t)_\Pi - \psi\| < \varepsilon$. Since $\varepsilon < \varepsilon'$, by equation (8), $p_\varepsilon^{t+1}(\cdot|\Theta_l) \in \Delta(\Pi_{p,l})$. Since $\varepsilon < \varepsilon^*$, by equation (7), $U^*((p_\varepsilon^{t+1})_\Pi) + \gamma < U^*(p_\varepsilon^{t+1}(\cdot|\Theta_l))$. Therefore, $U^*(p_\varepsilon^{t+1}(\cdot|\Theta_l)) > U^*((p_\varepsilon^{t+1})_\Pi) + \gamma$ $= p_\varepsilon^{t+1}(\Theta)U^*(p_\varepsilon^{t+1}(\cdot|\Theta)) + (1 - p_\varepsilon^{t+1}(\Theta))U^*(p_\varepsilon^{t+1}(\cdot|\Theta_l)) + \gamma$, so

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

Then applying Lemma 3 with $\Theta' = \Theta_l$, $\lambda = \frac{\gamma}{p_\varepsilon^0(\Theta)}$, and $\beta = \min\{p_\varepsilon^1(\Theta_l), 1 - \max_{\psi':\|\psi-\psi'\|\leqslant\varepsilon}\psi'(\Pi_{p,l})\}$, yields $\lim_{t\to\infty} p_\varepsilon^t(\Theta_l) \to 1$, a contradiction.

(ii) Suppose first that $\pi' \in \text{supp}\,\psi$. Let $\hat{\varepsilon} = \min_{\pi\in\text{supp}\,\psi}\psi(\pi)$. We prove case (ii) by showing that for every $l \in \{1,...,k\}$ and $\varepsilon < \min\{\varepsilon', \hat{\varepsilon}\}$ there exists a solution $(p^t)_{t\in\mathbb{N}_0}$ where p^0 is a mutation of p to the one-hypothesis relaxation of Θ in hypothesis $l \in \{1,...,m\}$ and $\lim_{t\to\infty}(p_\varepsilon^t)_\Pi = \psi$. Fix such an ε and let Θ' be the one-hypothesis relaxation of Θ in hypothesis $l \in \{1,...,m\}$. Define the candidate solution by letting p_0 be given by $p_{\mathcal{K}}^0(\Theta) = 1 - \varepsilon$, $p_{\mathcal{K}}^0(\Theta') = \varepsilon$, $p_\Pi^0(\cdot|\Theta') = \pi'$ and

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

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

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

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

We prove this by induction. For the initial step, since Θ' is an ε mutation of Θ, and $\varepsilon < \hat{\varepsilon} = \min_{\pi\in\text{supp}\,\psi}\psi(\pi)$, $\psi(\pi') \geqslant p_{\mathcal{K}}^0(\Theta') = \varepsilon$. Moreover, by equation (6), equation (1) is satisfied for $t = 0$. For the inductive step, observe that since the operator T is payoff monotone, $U^*(\pi') \leqslant U^*(\psi)$, and $\psi(\pi') \geqslant p_{\mathcal{K}}^t(\Theta')$ by the inductive hypothesis, we have

$p_{\mathcal{K}}^{t+1}(\Theta') \leqslant p_{\mathcal{K}}^t(\Theta') \leqslant \psi(\pi')$. Moreover, since by the inductive step $(p^t)_\Pi = \psi$, by equation (6), equation (1) is satisfied for $t+1$ as well.

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

Define the candidate solution by letting p_0 be given by $p_{\mathcal{K}}^0(\Theta) = 1 - \varepsilon, p_{\mathcal{K}}^0(\Theta') = \varepsilon$, $p_\Pi^0(\cdot|\Theta') = \pi_0$ and $p_\Pi^0(\cdot|\Theta) = \hat{\pi}_0$, where π_0 is an arbitrary element of $\Pi_{p,l}$ and $\hat{\pi}_0$ is an arbitrary element of $BR(\Delta(\Theta(\psi)))$. Subsequent period states are recursively defined as $\frac{p_{\mathcal{K}}^{t+1}(\Theta)}{p_{\mathcal{K}}^{t+1}(\Theta')} = \frac{T(p^t)(\Theta)}{T(p^t)(\Theta')}$, $p_\Pi^{t+1}(\cdot|\Theta') = \pi_t$ and $p_\Pi^{t+1}(\pi|\Theta) = \hat{\pi}$, where π_0 is an arbitrary element of $BR(\Delta(\Theta'(p_\Pi^t)))$. By definition, $(p^t)_{t \in \mathbb{N}_0}$ satisfies equation (2). It only remains to show that for every $t \in \mathbb{N}_0$, $\hat{\varepsilon} \geqslant p_{\mathcal{K}}^t(\Theta')$ so that equation (1) is satisfied and $BR(\Delta(\Theta'(p_\Pi^t))) \subseteq \Pi_{p,l}$. We prove this by induction. The initial step, follows from the definition of $\hat{\varepsilon}$. For the inductive step, observe that since the operator T is payoff monotone, $U^*(\pi') \leqslant U^*(\hat{\pi})$, for every $\pi' \in \Pi_{p,l}$ and $\hat{\varepsilon} \geqslant p_{\mathcal{K}}^t(\Theta')$ by the inductive hypothesis, we have $p_{\mathcal{K}}^{t+1}(\Theta') \leqslant p_{\mathcal{K}}^t(\Theta') \leqslant \hat{\varepsilon}$. Finally, the proof is concluded by applying Lemma 3. ∎

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 ψ, and they do not assign the same probability to all events that ψ gives positive probability, then their strict convex combination has a strictly lower KL divergence given ψ.

Lemma 6. *Let $X \in \mathcal{B}(\mathbb{R}^m)$, and let $\Phi, \Phi^1, \Phi^2 \in \Delta(X)$ be Borel probability measure with densities $\phi, \phi^1, \phi^2 \in \Delta(X)$ such that*

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

and such that ϕ^1 is not Φ almost surely equal to ϕ^2. For every $v \in (0, 1)$

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

The proof of this lemma is in the Online Appendix.

Proof of Proposition 6. The proof for the case in which $\{\hat{\pi}\} = BR(\Delta(\Theta(p_\Pi)))$ is simple: the main theorem in Berk (1966) guarantees that almost surely $\lim_{n \to \infty} \mu_t(\Theta(p_\Pi)) = 1$, 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_Π. Claim 3 shows that beliefs that induce ties have Lebesgue measure zero in the space of likelihood ratios. Claim 2 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 (2021); we discuss some of the differences below.

Partition the elements of Θ in the equivalence classes $\left\{\tilde{\theta}^l\right\}_{l=1}^C$ such that

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

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

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

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

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

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

The proof of this claim is in the Online Appendix. It is similar to the proof of Lemma 2 of Esponda and Pouzo (2016), but also covers the case in which Y is infinite. In that case, we combine the LLN with the Monotone Convergence Theorem to obtain the convergence of the likelihood ratios between two parameters when the Radon-Nykodim derivative between their distributions is unbounded.

Fudenberg, Lanzani, and Strack (2021) shows that the probability assigned to a particular KL-minimizing parameter is large infinitely often. We sharpen their conclusion, and prove 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 defi-

nite, which implies that every set of Lebesgue measure zero has measure zero under the limit distribution. For every $\nu \in \Delta\left(\left\{\tilde{\theta}^1, ..., \tilde{\theta}^K\right\}\right)$, define $U(\nu, a) = \sum_{i=1}^K \nu(\tilde{\theta}^i) \sum_{\theta_j \in \tilde{\theta}^i} \frac{\mu_\theta(\theta_j)}{\mu_\theta(\tilde{\theta}^i)} U_{\theta_j}(a)$.

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

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

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

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:

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

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

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

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_{p_Π} 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_{p_Π} almost surely $\log q_{\theta^1}(y|s, a) \leq \log \sum_{i=2}^K v_i q_{\theta^i}(y|s, a)$, 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_ψ almost surely $q_{\theta^1}(y|s, a) = \sum_{i=2}^K v_i q_{\theta^i}(y|s, a)$. By Lemma 6, this contradicts $\theta^2 \in \Theta(p_\Pi)$. Thus $v^T \Sigma v > 0$, and since v was chosen arbitrarily Σ is positive definite, proving Claim 3. □

Claim 3. *The set of $\nu \in \Delta\left(\left\{\tilde{\theta}^1, ..., \tilde{\theta}^K\right\}\right)$ such that $U(\nu, a) = U(\nu, a')$ for some $a \neq a'$ has Lebesgue measure 0 in $\mathbb{R}^K$.*

The proof of this claim is also in the Online Appendix.

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

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

the distribution of μ_m converges to some $\nu \in \Delta\left(\Delta\left(\Theta\left(p_\Pi\right)\right)\right)$ and the distribution of beliefs in the overall population converges to ν as well. Because the covariance matrix Σ is positive definite by Claim 3, the induced distribution of strategies converges to an element of $\Delta\left(BR\left(\Delta\left(\Theta\left(p_\Pi\right)\right)\right)\right)$. This concludes the proof of part (ii). ∎

References

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

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

Belew, R. K. (1990). "Evolution, learning and culture: Computational metaphors for adaptive algorithms". *Complex Systems*, pp. 11–49.

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

Binmore, K. and L. Samuelson (1997). "Muddling through: Noisy equilibrium selection". *Journal of Economic Theory* 74, pp. 235–265.

— (1999). "Evolutionary drift and equilibrium selection". *The Review of Economic Studies* 66, pp. 363–393.

Björnerstedt, J. and J. W. Weibull (1995). *Nash equilibrium and evolution by imitation.* Ed. by E. K. A. et al. London, MacMillan.

Blume, L. E. and D. Easley (2002). "Optimality and natural selection in markets". *Journal of Economic Theory* 107, pp. 95–135.

Blume, L. and D. Easley (1992). "Evolution and market behavior". *Journal of Economic theory* 58, pp. 9–40.

— (2006). "If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets". *Econometrica* 74, pp. 929–966.

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

Bohren, J. A. and D. Hauser (2020). "Bounded Rationality and Learning: A Framework and a Robustness Result".

Cabrales, A. and J. Sobel (1992). "On the limit points of discrete selection dynamics". *Journal of Economic Theory* 57, pp. 407–419.

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

Dekel, E. and S. Scotchmer (1992). "On the evolution of optimizing behavior". *Journal of Economic Theory* 57, pp. 392–406.

— (1999). "On the evolution of attitudes towards risk in winner-take-all games". *Journal of Economic Theory* 87, pp. 125–143.

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

— (1990). "On the Uniform Consistency of Bayes Estimates for Multinomial Probabilities". *The Annals of Statistics* 18, pp. 1317–1327.

Esponda, I. and D. Pouzo (2016). "Berk–Nash equilibrium: A Framework for Modeling Agents with Misspecified Models". *Econometrica* 84, pp. 1093–1130.

Esponda, I., D. Pouzo, and Y. Yamamoto (2020). "Asymptotic Behavior of Bayesian Learners with Misspecified Models". *mimeo*. eprint: `1904.08551`.

Esponda, I., E. Vespa, and S. Yuksel (2020). "Mental Models and Learning: The Case of Base-Rate Neglect". *mimeo*.

Frick, M., R. Iijima, and Y. Ishii (2020a). "Misinterpreting Others and the Fragility of Social Learning". *Econometrica* 88, pp. 2281–2328.

— (2020b). "Stability and Robustness in Misspecified Learning Models". *mimeo*.

— (2020c). "Welfare Comparisons for Biased Learning". *In preparation*.

Fudenberg, D. and L. A. Imhof (2012). "Phenotype switching and mutations in random environments". *Bulletin of mathematical biology* 74, pp. 399–421.

Fudenberg, D. and D. M. Kreps (1994). "Learning in extensive form games, II: Experimentation and Nash equilibrium". *mimeo*.

Fudenberg, D., G. Lanzani, and P. Strack (2021). "Limits Points of Endogenous Misspecified Learning". *Forthcoming at Econometrica*.

Fudenberg, D. and D. K. Levine (1993). "Self-confirming equilibrium". *Econometrica: Journal of the Econometric Society*, pp. 523–545.

Fudenberg, D., G. Romanyuk, and P. Strack (2017). "Active Learning with a Misspecified Prior". *Theoretical Economics* 12, pp. 1155–1189.

Fudenberg, D. and J. Tirole (1991). *Game theory*. MIT press.

Gagnon-Bartsch, T., M. Rabin, and J. Schwartzstein (2020). *Channeled Attention and Stable Errors*. Harvard Business School.

Hansen, L. P. and T. J. Sargent (2008). *Robustness*. Princeton university press.

Harsanyi, J. C. (1973). "Oddness of the number of equilibrium points: a new proof". *International Journal of Game Theory* 2, pp. 235–250.

He, K. (2019). "Mislearning from Censored Data: The Gambler's Fallacy in Optimal-Stopping Problems". *mimeo*.

He, K. and J. Libgober (2020). "Evolutionarily Stable (Mis)specifications: Theory and Application". *In preparation*.

Heidhues, P., B. Koszegi, and P. Strack (2020). "Convergence in Misspecified Learning Models with Endogenous Actions". *Forthcoming in Theoretical Economics*.

Heidhues, P., B. Kőszegi, and P. Strack (2018). "Unrealistic Expectations and Misguided Learning". *Econometrica* 86, pp. 1159–1214.

Hines, W. G. (1980). "Three characterizations of population strategy stability". *Journal of Applied Probability* 17, pp. 333–340.

Hofbauer, J. and K. Sigmund (1988). *Dynamical systems and the theory of evolution.*

Hong, H., J. C. Stein, and J. Yu (2007). "Simple forecasts and paradigm shifts". *The Journal of Finance* 62, pp. 1207–1242.

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

Kahneman, D. and A. Tversky (1973). "On the psychology of prediction." *Psychological review* 80, p. 237.

Kimura, M. (1983). *The neutral theory and molecular evolution.* Springer.

King, O. D. and J. Masel (2007). "The evolution of bet-hedging adaptations to rare scenarios". *Theoretical population biology* 72, pp. 560–575.

Liese, F. and I. Vajda (1987). *Convex statistical distances.* Vol. 95. Teubner.

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

Netzer, N. (2009). "Evolution of time preferences and attitudes toward risk". *American Economic Review* 99, pp. 937–55.

Nolfi, S., D. Parisi, and J. L. Elman (1994). "Learning and evolution in neural networks". *Adaptive Behavior* 3, pp. 5–28.

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

Ortoleva, P. (2012). "Modeling the change of paradigm: Non-Bayesian reactions to unexpected news". *American Economic Review* 102, pp. 2410–36.

Robson, A. J. (1996a). "A biological basis for expected and non-expected utility". *Journal of economic theory* 68, pp. 397–424.

— (1996b). "The evolution of attitudes to risk: Lottery tickets and relative wealth". *Games and economic behavior* 14, pp. 190–207.

Robson, A. J. and L. Samuelson (2007). "The evolution of intertemporal preferences". *American Economic Review* 97, pp. 496–500.

— (2019). "Evolved attitudes to idiosyncratic and aggregate risk in age-structured populations". *Journal of Economic Theory* 181, pp. 44–81.

Rothschild, M. (1974). "A two-armed bandit theory of market pricing". *Journal of Economic Theory* 9, pp. 185–202.

Samuelson, L. and J. Zhang (1992). "Evolutionary stability in asymmetric games". *Journal of Economic Theory* 57, pp. 363–391.

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

Traulsen, A. and C. Hauert (2009). "Stochastic evolutionary game dynamics". *Reviews of nonlinear dynamics and complexity* 2, pp. 25–61.

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

By continuity of U^* in ψ, there is an $\varepsilon^* \in (0, \varepsilon')$ and $\gamma > 0$ such that for all $\psi' \in B_{\varepsilon*}(\psi)$ $\tilde{\psi} \in BR(\Delta(B_{\bar{\varepsilon}}(\mathcal{M}_{\Theta,\psi}(\varepsilon))))$,

$$U^*(\psi') - U^*(\tilde{\psi}) < -\gamma. \tag{9}$$

By the upper hemicontinuity (see Lemma 2) of $\Theta(\cdot)$ there is $\bar{\varepsilon}$ such that $||\psi' - \psi|| < \bar{\varepsilon}$ implies $\Theta_\varepsilon(\psi') \subseteq B_{\bar{\varepsilon}}(\mathcal{M}_{\Theta,\psi}(\varepsilon))$. Thus if there were a t such that $||p_\Pi^\tau - \psi|| < \min \varepsilon^*, \bar{\varepsilon}$ for all $\tau \geqslant t$, it would follow from equation (9) that

$$U^*(p^{\tau+1}(\cdot|\Theta)) + \frac{\gamma}{p_\varepsilon^0(\Theta)} < U^*(p^{\tau+1}(\cdot|\Theta'))). \tag{10}$$

By Lemma 3 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 $\{\hat{a} + \frac{1}{m}c\}_{c \in \mathbb{Z}\backslash\{0\}} \cap A$, 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 Esponda 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 (Θ, π_m) and a justifying belief $\mu_m \in \Delta(\Theta(\pi_m))$. Since A is compact, $\Delta(A)$ is also compact, and therefore (π_m, μ_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 $U_{\cdot}(\cdot)$ is jointly continuous in beliefs and actions

$$U_{\mu_n}(\pi_n) \geqslant U_{\mu_n}(a_n) \qquad \forall n \in \mathbb{N}$$

implies $U_{\hat{\mu}}(\hat{\pi}) \geqslant U(a)$. Moreover, by the upper hemicontinuity (see Lemma 2) of Θ,

$$\mu_n \in \Delta(\Theta(\pi_n)) \qquad \forall n \in \mathbb{N}$$

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 $n > N$, $\text{supp}\,\pi_n \subseteq B_\varepsilon(\hat{a})$. Therefore, since $(\hat{a}, \Theta)$ is an attractor, there is an $N' > N$ such that for $n \geqslant N'$, $\pi_n^*(\{a : a < \hat{a}\}) > 0$ and $\pi_n^*(\{a : a > \hat{a}\}) > 0$.

Since U^* is continuously differentiable and $\left(\delta_{\hat{a}}, \delta_{\Theta(\hat{a})}\right)$ 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 $V\left(\mathcal{M}_{\Theta,\psi}(\varepsilon)\right) > V\left(\mathcal{M}_{\Theta,\psi}(\hat{a})\right) = U^*(\hat{a})$ for all $\varepsilon < \hat{\varepsilon}$, there exists an $\bar{\varepsilon} < \hat{\varepsilon}$ such that for all $\varepsilon < \hat{\varepsilon}$, $BR\left(\mathcal{M}_{\Theta,\psi}(\varepsilon)\right) \in (\hat{a}, \hat{a} + K)$. Fix an $\varepsilon < \min\{\bar{\varepsilon}, \hat{\varepsilon}\}$. By the upper hemicontinuity of $\Theta_{\varepsilon}(\cdot)$, we have that $\Theta_{\varepsilon}(\pi_n)$ is converging to $\mathcal{M}_{\Theta,\psi}(\varepsilon) = \Theta_{\varepsilon}(\hat{a})$. Therefore, there exists $N'' > N'$ such that for all $n > N''$, $BR_n\left(\Theta_{\varepsilon}(\pi_n)\right) \in \left(\frac{\hat{a} + BR\left(\mathcal{M}_{\Theta,\psi}(\varepsilon)\right)}{2}, \hat{a} + K\right)$. 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 $n > N''$, $U^*\left(BR_n\left(\Theta_{\varepsilon}(\pi_n)\right)\right) > U(\pi_n)$, so that π_n is not resistant to an ε expansion of Θ. ∎

B.2 Proofs of results stated in the Appendix

B.2.1 Proof of Lemma 2

Proof. Since Θ 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 $\hat{\theta} \in \Theta$ and $K \in \mathbb{R}_{++}$ such that $H_\psi\left(Q^*, Q_{\hat{\theta}}\right) + K < H_\psi\left(Q^*, Q_\theta\right)$. Since $H_\psi\left(Q^*, \cdot\right)$ is jointly lower semicontinuous see, e.g., Chapter 1 of Liese and Vajda, 1987, there exists an $N \in \mathbb{N}$ such that $H_{\psi_n}\left(Q^*, Q_{\theta_n}\right) > H_{\psi_n}\left(Q^*, Q_{\hat{\theta}}\right) + K$ for all $n \geqslant N$. Moreover, since $H_\psi\left(Q^*, Q_{\hat{\theta}}\right) < \infty$, $H_\psi\left(Q^*, Q_{\hat{\theta}}\right)$ is continuous in the strategy distribution at ψ, and there exists $N' > N$ such that $H_{\psi_n}\left(Q^*, Q_{\hat{\theta}}\right) < H_{\psi_n}\left(Q^*, Q_{\hat{\theta}}\right) + \frac{K}{2}$, for all $n \geqslant N$. But then, we would have

$$H_{\psi_n}\left(Q^*, Q_{\hat{\theta}}\right) < H_{\psi_n}\left(Q^*, Q_{\theta_n}\right),$$

a contradiction. ∎

B.2.2 Proof of Lemma 3

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 $p_{\mathcal{K}}(\Theta') \geqslant \beta > 0$ on $P_{\lambda,\beta}(\Theta, \Theta')$, so $\frac{T(p)(\Theta')}{T(p)(\Theta)} \frac{p_{\mathcal{K}}(\Theta)}{p_{\mathcal{K}}(\Theta')}$ 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 Θ' strictly outperforms the strategy used by agents with subjective model Θ on $P_{\lambda,\beta}(\Theta, \Theta')$ implies that this minimum is strictly larger than 1. ∎

B.2.3 Proof of Lemma 6

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

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

the set B can be chosen such that

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

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

$$\rho = \min_{z \in [\underline{K}, \bar{K}], z' \in [z+K, \bar{K}]} \log(vz + (1-v)z') - v\log(z) - (1-v)\log(z') > 0 \qquad (11)$$

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

$$-\int_{Y \backslash B} \log\left(v\phi^1(x) + (1-v)\phi^2(x)\right) d\Phi(x) - \int_B \log\left(v\phi^1(x) + (1-v)\phi^2(x)\right) d\Phi(x)$$

$$\leqslant -\int_{Y \backslash B} \log\left(v\phi^1(x) + (1-v)\phi^2(x)\right) d\Phi(x) - \int_B \left(\log v\phi^1(x) + (1-v)\phi^2(x) + \rho\right) d\Phi(x)$$

$$= -v\int_Y \log \phi^1(x) d\Phi(x) + (1-v)\int_Y \log \phi^2(x) d\Phi(x) - \rho\Phi(B)$$

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

as desired. ∎

B.2.4 Proof of Claim 1

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

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

Since $\Theta \in \mathcal{K}$, $\mathbb{E}\left[(L_m^l)^+ \mid (Z_i^l)_{i=1}^{m-1}\right] < \infty$ and so by the Strong Law of Large Numbers and the Monotone Convergence Theorem, it follows that $\lim_{m \to \infty} e^{Z_m^l} = 0$ a.s. Therefore,

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

proving the claim. □

B.2.5   Proof of Claim 3

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 μ who solve the equations are a vector subspace of $\mathbb{R}^K$. However, assumption (1) of the proposition implies that there exists $\tilde{\theta}_l \in \Theta(p_\Pi)$ such that $U_{\mu(\cdot \mid \tilde{\theta}_l)}(a) \neq U_{\mu(\cdot \mid \tilde{\theta}_l)}(a')$, 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 $\nu \in \Delta\left(\left\{\tilde{\theta}^1, ..., \tilde{\theta}^K\right\}\right) \subseteq \mathbb{R}^K$ such that $U(\nu, a) = U(\nu, a')$ for some $a \neq a'$ has Lebesgue measure 0 as well. □

B.3   Examples

B.3.1   Example 1

a) Esponda and Pouzo (2016) shows that

$$\sum_{a \in \{1,2\}} \psi(a) H\left(Q^*(\cdot \mid a), Q_\theta(\cdot \mid s, a)\right) = \psi(2)(34 - i + 2l)^2 + \psi(10)(2 - i + 10l)^2. \tag{12}$$

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

$$-2\psi(2)(34 - 40 + 2 * \frac{10}{3}) - 2(1 - \psi(2))(2 - 40 + 10 * \frac{10}{3}) \leqslant 0 \tag{13}$$

$$4\psi(2)(34 - 40 + 2 * \frac{10}{3}) + 20(1 - \psi(2))(2 - 40 + 10 * \frac{10}{3}) \leqslant 0. \tag{14}$$

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

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

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

This expression is maximized at a $\tilde{v}$ with $\frac{\sqrt{(1-\tilde{v}^2)}}{\tilde{v}} > 1/12$ if and only if $\psi(2) > 13245113/13623544 - (21\sqrt{73/2})/1702943 \approx 0.972$. Finally, the claim that $A_{\mathcal{M}_{\Theta,\psi}} = \{10\}$ if and only if $\psi(2) > \approx 0.972$ follows from

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

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

b) When $\psi(2) = 1$, the parameter that minimizes equation (12) is $(3/2,32)$, and since $BR(3/2, 32) = \{10\}$, $(\Theta, 2)$ is not a Berk-Nash equilibrium. When $\psi(10) = 1$, the parameter that minimizes equation (12) is $(3/2, 28)$. Since $BR(3/2, 28) = \{2\}$, $(\Theta, 10)$ is not a Berk-Nash equilibrium. For every totally mixed ψ, the Hessian of equation (12) 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 (12) 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 ε expansion of the model is:

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

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

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: $\Theta = [\underline{\theta}_H, \bar{\theta}_H] \times [\underline{\theta}_L, \bar{\theta}_L]$. 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 $\underline{\theta}_i > 2/3 > \theta_i^*$. 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* $\sigma(s_i|y_i)$. *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:* $Q(y_i|a, s_i) = \underline{\theta}_i = \bar{\theta}_i = \sigma(s_i|y_i)$. *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 2

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

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

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

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

B.3.4   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 θ_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 Example 5

The subjective model of the agent is the following:

$$Q_{(\rho,\kappa)}(y_H|a_h) = 1/3 + \rho - \kappa/2, Q_{(\rho,\kappa)}(y_M|a_h) = 1/3 + \kappa/4 + \rho/2, Q_{(\rho,\kappa)}(y_L|a_h) = 1/3 + \kappa/4 - 3\rho/2$$

and

$$Q_{(\rho,\kappa)}(y_H|a_e) = 1/3 + \kappa, Q_{(\rho,\kappa)}(y_M|a_e) = 1/3 - \kappa/2, Q_{(\rho,\kappa)}(y_L|a_e) = 1/3 - \kappa/2.$$

Because the agent is sure that $\rho = 1/5$, $\Theta_{\mathcal{Y}} = \{\{y_H\}, \{y_M, y_L\}\}$ is a minimal attention partition for Θ that the agent thinks is sufficient to infer the value of κ. In reality, low outcomes are more unlikely, and the objectively optimal action is a_e.

$$Q^*(y_H|a_h) = 11/42, Q^*(y_L|a_h) = 1/100 \text{ and } Q^*(y_H|a_e) = 4/21, Q^*(y_L|a_e) = 51/700.$$

B.3.6 An equilibrium that is resistant to one-hypothesis mutations but is not resistant to local mutations.

Example 8. *An agent can choose between a hard, medium or easy task, $A = \{a_h, a_m, a_e\}$. They believe that the expected performance in each task is a function of their skill κ and a relative difficulty parameter ρ. In particular, $Q_{(\kappa,\rho)}(y|a_h)$, $Q_{(\kappa,\rho)}(y|a_m)$, and $Q_{(\kappa,\rho)}(y|a_l)$, are normal distributions with unit variance, centered at $2\kappa - 3\rho, \kappa - 2\rho$ and $-\frac{3}{2}\rho$ respectively. The agent's utility is their performance, i.e., $u(a, y) = y$. In reality $Q^*(y|a_h), Q^*(y|a_m)$, and $Q^*(y|a_l)$, are normal distributions with unit variance, centered at $3, 1$, and -1 respectively, so that the objectively optimal action is a_m. The agent is sure that their skill is between 1 and 10, and that the relative difficulty is between 0 and 2, so $\Theta = [1, 10] \times [0, 2]$. Here $(\psi(a_h) = 2/5$, $\psi(a_m) = 3/5)$ is a Berk-Nash equilibrium sustained by a Dirac belief on $(1, 1)$. The equilibrium is unstable to local mutations: Along the most informative ray $\mathcal{M}(\varepsilon) = (1 - \varepsilon, 1)$, the mutated agent lowers their self-assessment, which induces a_m as the unique best reply. And since a_m gives a higher payoff than mixing between a_h and a_m, Proposition 3 implies that the equilibrium is unstable to local mutations. However, it is resistant to one-hypothesis mutations: Relaxing the hypothesis on relative difficulty does not induce a change in behavior, since it is not binding in the equilibrium, and relaxing the assumption that skill is larger than 1 makes $(-0.81, 0)$ the KL-minimizing parameter, which induces a_e as the unique best reply. Since a_e gives a lower payoff than ψ, Proposition 5 implies that the equilibrium is stable to one-hypothesis mutations.*

B.3.7 An equilibrium that is not stable with attention partitions but is stable without them

Example 9. *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*

$$u\left(a, (y_1, y_2)\right) = \begin{cases} 1 & a = l = y_1 \\ 4 & a = r = y_1 \\ 0 & a \neq y_1 \end{cases}.$$

Here $\mathcal{H} = \left\{ (\theta_{al}, \theta_{au})_{a \in A} \in [0,1]^A \right\}$ where θ_{al} is the probability of $y_1 = l$ given a and θ_{au} is the probability of $y_2 = u$ given that a has been played. The agent believes that the action does not affect the outcome, i.e., $\Theta = \left\{ \theta \in \mathcal{H} : \theta_{ll} = \theta_{rl}, \theta_{lu} = \theta_{ru} = \frac{1}{2} \right\}$. In reality $\theta_{ll}^* = \frac{1}{4}$, $\theta_{rl}^* = \frac{2}{3}$, $\theta_{lu}^* = \frac{1}{10}$ 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, θ_{lr}. 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, θ_{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.3.8   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 $\left\{ T + \frac{m}{n} : m \in \mathbb{N} \right\}$ 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.[29] 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

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

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

---

[29]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.

threshold, $\pi(T)(r) = 1$. Then,

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

and

$$\theta_r(\pi) = -E_{\sigma_g}[s|s \geqslant T] < -E_\sigma[s|s \geqslant T] = \theta_r(T)$$

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 $((1, \theta_{a_c}(\hat{s}), \theta_{a_r}(\hat{s})) + (v_0, v_r, v_c))$ with $v_0 < 0, v_r \geqslant 0, v_c \leqslant 0$. The best reply to these parameters is same as the equilibrium strategy, except at the threshold performance it praises instead of mixing. Since $\hat{s} < 0$, this strategy outperforms the equilibrium, which is thus destabilized by the mutation.

B.4   The role of quasi-strictness

Example 10. *Suppose that there are three actions $A = \{a, b, c\}$ and that the outcomes have three components that are either 1 or 0, i.e., $Y = \{0,1\} \times \{0,1\} = \{0,1\}$. The utility of a and b depends only on the first component; a is better if the first component is likely to be 1, b if it is likely to be 0:*

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

The utility to c depends only on the third component:

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

The parameter space has two dimensions $\mathcal{H} = [0, 1] \times [0, 1]$, where for a $\theta = (\theta_1, \theta_2) \in \mathcal{H}$, θ_1 is both the probability that the first component is equal to 1 while playing a, b, or c, and the probability that the second component is equal to 1 while playing a, b, or c. Instead, θ_2 is the probability that the third component is equal to 1 while playing b or c. Finally, the agent

(correctly) believes that the third component is always equal to 0 if they play a. Formally

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

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

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

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

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

To see that the equilibrium does not resist local mutations, note that for small ε, the most improving parameter is $\mathcal{M}_{\Theta,\delta_a}\left(\varepsilon\right)$ decreases θ_1 to better match the frequency observed on the second component. This makes b strictly preferable to a. However, even if b performs less well than a, behavior does not converge back to the equilibrium: a small fraction of mutated agents playing b allows all the agents to learn that c is preferable to both alternatives, and therefore the entire population switches to c forever from the second period onward.

B.5 Mutations to smaller subjective models

Invasion by a non-improving mutation Consider the subjective model Θ and an arbitrary completely mixed equilibrium ψ of Example 1a. 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 Θ': 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 ψ.

Definition (Evidence-Based Simplifications)**.** The ε mutation of a steady state $\delta_\Theta \times \psi$ to Θ' is an *evidence-based simplification* if $\Theta' \subseteq \Theta$ and $\operatorname{argmin}_{\theta \in \Theta'} H_\psi\left(Q^*, Q_\theta\right) = \operatorname{argmin}_{\theta \in \Theta} H_\psi\left(Q^*, Q_\theta\right)$.

A mutation is an evidence-based simplification if the mutated agents reduce the size of their subjective model, eliminating some of the parameters that provide a suboptimal explanation of the equilibrium data.

Remark 1. If the mutation of $\delta_\Theta \times \psi$ to Θ' is an evidence-based simplifications, then (Θ, ψ) resists mutation to Θ'.

Since evidence-based simplifications do not drop any KL-minimizer, the equilibrium strategy remains a best reply to some belief over the KL minimizing parameters in the simplified paradigm.

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

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

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