Which Misperceptions persist?

Drew Fudenberg and Giacomo Lanzani

Harvard University

February 2021
Introduction

• Economic agents are often misspecified: Their prior beliefs may rule out the true data generating process.
• The misspecification may be due to:
  – Behavioral biases such as overconfidence, correlation neglect, etc.
  – Oversimplifications of a complex environment that omit some relevant variables.
• Misspecified agents need not learn their optimal actions even when they have lots of data.
• But when will the data lead agents to realize they are misspecified?
• Bayesian updating can’t lead to a positive probability on a data generating process that lies outside the support of the prior.
• We propose an evolutionary criterion to evaluate the stability of misspecified Bayesian models.
• We allow offspring to become more open-minded than their parents in face of unexplained evidence.
• Our Bayesian agents have beliefs over “parameters”, where each parameter is associated with an action-contingent outcome distribution.
• We call the support of the prior beliefs their “subjective model.”
• Each agent observes the actions and outcomes of the previous generation, and uses them to update beliefs within their subjective model.
• The agents then choose an action that is a best reply to these beliefs.
• Different agents may employ different subjective models, and the relative frequency of the models that induce better actions increases over time.
• We show that steady states of this process coincide with Berk-Nash equilibria (Esponda and Pouzo, 2016, henceforth BN-E).
• “Mutations” lead some agents to adopt an expanded subjective model with a larger support.

• They use the expanded model to make their inferences and choose their actions.

• We then ask whether the use of the expanded model will spread, or whether the existing model “resists mutations.”

• Not all equilibria are unstable, because the share of mutants only increases if they do better than agents using the prevailing paradigm.

• We consider two ways that subjective models can expanded: “local” expansions to nearby subjective models, and “one-hypothesis relaxations” that drop one of the hypotheses that characterize the subjective model.

• We characterize stability in both cases.

• We apply the results to several common misspecifications.
Some Related Work

• Berk (1966) studies Bayesian updating with fixed i.i.d. data.

• Esponda and Pouzo (2016) relates Berk-Nash equilibrium to the long-run outcome of misspecified learning from endogenous data.

• Fudenberg, Romanyuk, Strack (2017), Heidheus, Koszegi, and Strack (2018), Bohren and Hauser (2018), and He (2019) characterize long-run beliefs in specific settings, e.g. binary states

• Esponda, Pouzo, and Yamamoto (2019), Frick, Iijima, and Ishii (2020), and Fudenberg, Lanzani, and Strack (2021) study the convergence of actions and/or beliefs with misspecified learners in more general settings.
• Esponda, Vespa and Yuksel (2020) experiment finds that misspecified agents make small partial adjustments in the face of large, unexplained evidence, as in our analysis of local mutations, and that agents use their inferences from a particular task or action to adjust the predicted consequences for other actions.

• He and Libgober (2020) and Murooka and Yamamoto (2021) both study misspecified learning in a game.

• HL looks at the competition between a misspecified model and the correctly specified model.

• MY shows that when all the agents share the same misspecification, deviations from the objectively optimal behavior are amplified by the strategic interaction.

• Closest: Gagnon-Bartsch, Rabin, and Schwartzstein (GRS) (2019); more on this paper later.
Single Agent Decision Problem

• An agent chooses an action \( a \in A \) after having observed a signal \( s \in S \).

• \( A \) and \( S \) are finite. Strategies are elements of \( \Pi = A^S \).

• Outcome \( y \in Y \subseteq \mathbb{R}^m \).

• Utility function \( u : S \times A \times Y \rightarrow \mathbb{R} \).

• Objective data generating process:
  - probability distribution over signals \( \sigma \in \Delta(S) \),
  - contingent outcome distributions \( q^*(\cdot|\cdot) \in \Delta(Y)^{S \times A} \).

• The objective expected utility of strategy \( \pi : S \rightarrow A \) is:

\[
U^*(\pi) = \sum_{s \in S} \sigma(s) \int_Y u(s, \pi(s), y) dq^*(y|s, \pi(s)).
\]
Subjective Models

• The agent uses parametric models to describe the environment.

• A subjective model for an agent is the set of parameters \( \Theta \subseteq \mathbb{R}^k \) they consider possible, where each \( \theta \in \Theta \) is associated with family of probability distributions \( q_\theta(\cdot|s,a) \).

• The agent’s uncertainty about the value of the parameter is described by a belief \( \mu \in \Delta(\Theta) \).

• The subjective expected utility of \( \pi \) is \( U_\mu(\pi) \).

• We let \( BR(\mu) = \arg\max_{\pi \in \Pi} U_\mu(\pi) \) denote the set of pure best replies to \( \mu \).
Inference and KL minimizers

• Given two $q, q' \in \Delta(Y)$ we define

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

• $q'$ with smaller $H(q, q')$ better explain the true distribution $q$.

• Given a distribution of strategies $\psi$ in the population, we let

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

• Let $\Theta(\psi)$ denote KL minimizers for $\psi$:

$$\Theta(\psi) : = \argmin_{\theta \in \Theta} H_\psi(Q^*, Q_\theta).$$

• When an agent with subjective model $\Theta$ observes random draws from a large population playing $\psi$, their beliefs concentrate on $\Theta(\psi)$. 
State of the System

- There is a continuum of agents.

- The state of the system at every period \( t \in \mathbb{N} \) is a finite-support joint distribution \( p \in \Delta(\mathcal{K} \times \Pi) =: P \) over the subjective models and strategies of the agents.

- We denote the marginal distributions of \( p \) as \( p_{\mathcal{K}} \) and \( p_{\Pi} \).

- Each agent’s posterior beliefs are supported on the KL-minimizing parameters in their \( \Theta \) with respect to the distribution over actions in the last period.

- The agent plays some best response to these beliefs.

- \( p^{t+1}(\cdot | \Theta) \in \Delta(\Pi) \) denotes the distribution over actions played at time \( t + 1 \) by the agents with subjective model \( \Theta \) when the previous state is \( p^t \).
Evolutionary Dynamics

• We assume that the share of agents with a particular subjective model evolves according to a payoff monotone dynamic $T : P \to \Delta(K)$.
• Biological interpretation: parents transmit their subjective model.
• In our model, the beliefs themselves are not inherited; what is inherited is how to use observables to reach a conclusion.
• The offspring then perform Bayesian updating based on the actions and outcomes in the previous period.
• Other economic examples, such as the misspecified beliefs of a seller about a demand function, are better interpreted as arising from imitation.
• Informally, a solution of the evolutionary system is a sequence of states where the shares of the subjective models evolve according to $T$, and each agent plays a best reply to the KL-minimizers given the previous period’s data. And a steady state is a fixed point of that process, i.e. a constant solution.
Steady States and Equilibria

Berk-Nash equilibria require that the agents’ beliefs are supported on the set of parameters that best explain the observables given $\Theta$.

**Definition**

A Berk-Nash equilibrium is a $(\Theta, \psi)$ such that for every $\pi \in \text{supp} \psi$ there exists a belief $\mu \in \Delta(\Theta(\psi))$ with $\pi \in BR(\mu)$.

The equilibrium is uniformly strict if $\psi = BR(\mu)$ for every $\mu \in \Delta(\Theta(\psi))$.

**Lemma**

For all $\Theta \in \mathcal{K}$ and $\psi \in \Delta(\Pi)$, $(\delta_\Theta \times \psi)$ is a steady state if and only if $(\Theta, \psi)$ is a Berk-Nash equilibrium.
Mutations

<table>
<thead>
<tr>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\overline{p}$ is an $\varepsilon$ mutation of a steady state $\delta_\Theta \times \psi$ to $\Theta'$ if</td>
</tr>
<tr>
<td>(a) $\overline{p} = (1 - \varepsilon)\delta_\Theta + \varepsilon\delta_{\Theta'}$ and</td>
</tr>
<tr>
<td>(b) $\overline{p}(\cdot</td>
</tr>
</tbody>
</table>

- A fraction $\varepsilon$ of the agents shift to $\Theta'$, and play a best reply to the KL-minimizing parameters in $\Theta'$. 
Resistance to Mutations

**Definition**
A Berk-Nash equilibrium \((\Theta, \psi)\) resists invasion by \(\Theta'\) if, after a sufficiently small mutation, the aggregate behavior of the population converges back to \(\psi\).

**Definition (Improving Mutations)**
The \(\varepsilon\) mutation of a steady state \(\delta_\Theta \times \psi\) to \(\Theta'\) is improving if \(\Theta'\) allows a lower KL divergence w.r.t. \(\psi\) than \(\Theta\) does.

**Proposition**
*If the mutation of \(\delta_\Theta \times \psi\) to \(\Theta' \supseteq \Theta\) is not explanation improving, then \((\Theta, \psi)\) resists invasion by \(\Theta'\): A more open-minded model can destabilize an equilibrium only if it better explains the equilibrium distribution.*
Local Mutations

Definition (Local Mutations)

Subjective model $\Theta'$ is the $\varepsilon$ expansion of $\Theta$ if

$$\Theta' = \{ \theta' \in \mathbb{R}^k : \exists \theta \in \Theta, \| \theta - \theta' \| \leq \varepsilon \}.$$  

Definition

A Berk-Nash equilibrium $(\Theta, \psi)$ resists local mutations if it resists invasion by every sufficiently small $\varepsilon$ expansion of $\Theta$. 
Uniformly Strict Berk-Nash Equilibria

- One reason an equilibrium can resist local mutations is if all small mutations may lead agents to stick with the same best responses (given the data generated in equilibrium).

**Proposition**

*Every uniformly strict Berk-Nash equilibrium resists local mutations.*

- In a uniformly strict equilibria, the equilibrium action $a$ is a strict best response to every KL-minimizing parameter.
- This implies the beliefs of the agents under small mutations are concentrated on a neighborhood where the unique best reply is still $a$. 
Most improving Parameters

- Uniformly strict Berk-Nash equilibria need not exist. At other Berk-Nash equilibria, multiple strategies are a best reply to some belief over KL-minimizers.
- And only some of these are best replies to the KL minimizers of the expanded model.
- The comparison between the performance of this subset of strategies and the equilibrium is what determines resistance to local mutations.
- To identify the relevant strategies we introduce the following notion.
- The most improving parameters at a steady state are
  \[ M_{\Theta, \psi}(\varepsilon) = \arg\min_{\theta \in \Theta_{\varepsilon}} H_{\psi}(Q^*, Q_{\theta}). \]
- Following an \( \varepsilon \) expansion, the mutants’ beliefs will be concentrated on these parameters, because they make the greatest local improvement in the explanation of equilibrium data.
Characterization

Proposition

• If all best replies to the most improving parameters at an equilibrium induce a higher payoff, the equilibrium does not resist local mutations.

• If the equilibrium is quasi-strict and some best reply to the the most improving parameters give a lower payoff, the equilibrium resists local mutations.

• An equilibrium is quasi-strict if all the best replies are played with positive probability.

• If the equilibrium is not quasi-strict, the feedback gathered from a tiny fraction of mutated agents playing a more revealing strategy that is not used in equilibrium may change the behavior of the old population, even if they were performing better than the mutants.
Monopoly Pricing and Linear Demand

- A monopolist faces demand function $y = i^* - l^* a + \omega$.

- $a$ is the price chosen by the monopolist and $\omega$ is a standard normal shock.

- The monopolist’s payoff is $u(a, y) = ay$.

- It is uncertain about the the intercept $i \in \mathbb{R}$ and slope $l \in \mathbb{R}$ of the demand function.

- The true values of the parameters are $(i^*, l^*) = (42, 4)$.

- The monopolist has two actions, $A = \{2, 10\}$.

- Objectively optimal to use action 2.
Nyarko (1994)

- The quasi-strict equilibrium $\psi(2) = \frac{1}{5}$ is unstable to local mutations because the most improving parameter relaxes the unique binding constraint, allowing for a larger slope, which induces the (optimal) low price.

(Green arrow points towards greatest KL improvement.)
Esponda and Pouzo (2016)

- The quasi-strict equilibrium $\psi(2) = \frac{35}{36}$ resists local mutations because the most improving parameter has a larger intercept, which induces the (suboptimal) high price.

(Green arrow points towards greatest KL improvement.)
One-hypothesis Mutations

- We also consider agents whose subjective model is described by a finite collection of hypotheses about the underlying parameter.

- Quantitative statements like:
  - Restrictions on the possible values of one of the dimensions of the parameter, e.g. an overconfident agent who is sure that their skill is higher than a threshold.
  - Joint restrictions on the parameters, as independence between two variables.

- The hypotheses describe the parts of the agent’s model that can be separately relaxed by a mutation.

- Formally, there is a finite collection of continuous functions $(f_i)_{i=1}^m$, $f_i : \mathbb{R}^k \to \mathbb{R}$ such that

$$
\Theta = \{\theta : f_i(\theta) \geq 0, \forall i \in \{1, \ldots, m\}\}.
$$
Definition

- The subjective model $\Theta^l$ is a one-hypothesis relaxation of $\Theta = \{ \theta \in \mathbb{R}^k : f_i(\theta) \geq 0, \forall i \in \{1, \ldots, m\} \}$ if $\Theta' = \{ \theta \in \mathbb{R}^k : f_i(\theta) \geq 0, \forall i \in \{1, \ldots, m\} \setminus \{l\} \}$.

- A Berk-Nash equilibrium $(\Theta, \psi)$ resists one-hypothesis mutations if it resists invasion from every one-hypothesis relaxation.

- Several collections of hypotheses can describe the same $\Theta$, and they can have different sets of one-hypothesis relaxations.

- This is natural, as the hypotheses are part of the agents’ model of the world.
Characterization

**Definition**

Given a steady state \( p \) we define the \( l \)-agnostic minimizers, \( \mathcal{P}_l(p) \), as the set of the KL-minimizers for the relaxed problem where hypothesis \( l \) is not imposed.

**Proposition**

Let \( (\Theta, \psi) \) be a uniformly strict Berk-Nash equilibrium.

- If for some \( l \), all the best replies to \( \mathcal{P}_l(p) \) induce a larger payoff than the equilibrium strategy distribution, then \( (\Theta, \psi) \) does not resist one-hypothesis mutations.

- If for every \( l \) some best replies to \( \mathcal{P}_l(p) \) induce a smaller payoff than the equilibrium strategy distribution, then \( (\Theta, \psi) \) resists one-hypothesis mutations.

**General Case**
Taxation and Overshooting

• 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 $\tau^*$ has two income brackets, and the higher one is heavily taxed:

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

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

• Their subjective model of the tax schedule is quadratic with random coefficients:

$$
\tau_\theta(z) = (\theta_1 + \eta)z + (\theta_2 + \eta)z^2,
$$

where $\eta$ is a standard normal.
• The agent observes $y = (z, x)$ at the end of each period.

• The original paradigm is that the tax schedule is linear.

• Given any action $a$, the KL-minimizing parameter treats the expected marginal rate as the actual average rate.

• The unique pure Berk-Nash equilibrium is uniformly strict and has too much effort.
• The agent observes \( y = (z, x) \) at the end of each period.

• The original paradigm is that the tax schedule is linear.

• Given any action \( a \), the KL-minimizing parameter treats the expected marginal rate as the actual average rate.

• The unique pure Berk-Nash equilibrium is uniformly strict and has too much effort.

• An agent who drops the linearity assumption estimates a very high quadratic term, because most realized income levels will be near the shift point between brackets.

• They extrapolate this progressivity as a global feature which leads them to choose the minimal action 3.

• The mutated agent overshoots the optimum, and the equilibrium resists to “one-hypothesis” mutations.
Figure: Misspecified Taxation Schedule
The agent is a buyer with valuation \( v = \omega + 5 + s \).

Seller who owns the object and values it at \( \omega \).

The signal \( s \) is a mean-zero shock independent of \( \omega \).

The mechanism used is double action with price at the buyer’s bid, so the seller sets their bid \( x \) equal to their value.

The value \( \omega \) is \( \omega = 3 \), with probability \( 1/2 \), \( \omega = 2 \) with probability \( 1/4 \) and \( \omega = 1 \) with probability \( 1/4 \).

The value is observed only if a sale occurs.
• A parameter is a probability distribution over prices, and a conditional distribution over values given the price.

• The true distribution of values conditional on an ask price depends on the price.

• However, the agent believes that the price and the value are independent.
Aligned Preferences

- Suppose the distribution of $s$ is a point mass on 0.
- Then the objectively optimal strategy is to bid 3.
- Bidding 2 is a Berk-Nash equilibrium.
- The KL-minimizing parameter is an independent joint probability distribution that is correct about the distribution of seller bids.
- Because 3 is not accepted, the mutated agents cannot infer that there is correlation at the high price level, and so they do not increase their bid.
- Note that the equilibrium does not resist a point mutation that adds the true data generating process: In that case, the only parameter that allows the agent to explain the observed correlation between values at low and medium prices leads them to the correct belief about the distribution of values when they bid 3.
• Suppose $s$ instead that is uniform over $\{-1, 1\}$.
• The optimal strategy is still to always bid 3 after every signal.
• The strategy $a = 2 + s$ is a Berk-Nash equilibrium— now the agent sometimes bids 3.
• The KL-minimizer is again an independent joint probability distribution that is correct about the distribution of seller bids.
• Since the value is observed only when a transaction occurs, the observed distribution over values is too pessimistic, which leads the agent to bids 1 after signal $-1$.
• The equilibrium does not resist one-hypothesis mutations: a mutation that drops the restriction that high values have the same probability after each price lets the agent realize the high price is correlated with high value.
• Realizing this leads the agent to always bid high ($a = 3$).
• This did not happen without the noise since in that case the agent never makes the high bid.
Channeled Attention

- (GRS) suggest that subjective models channel the attention to a subset of events.
- The attention partition $\mathcal{A}_\Theta$ includes only the events that are needed to distinguish between parameters in $\Theta$ that induce different best replies.
- If the agent only pays attention to this partition, their beliefs concentrate on parameters that minimize the KL divergence of the distribution on $\mathcal{A}_\Theta$.
- In an attention-improving mutation the subjective model of the mutants allows for lower KL divergence on the attention partition.
- Attention-improving mutations are a subset of improving mutations because they require the improvement occurs over a coarser partition of outcomes.
• GRS assumes that actions do not influence the distribution over outcomes.

• This implies that whether an attention-improving mutation is triggered is independent of the current equilibrium distribution.

• And so is the inference made by a decision maker with an enlarged model.

• Here an objectively suboptimal equilibrium resists a mutation to the correctly specified model if and only if it is attentionally stable—no role for evolutionary forces.

• We let actions change the outcome distribution, and replace the GRS assumption of a point mutation to the true model with other sorts of mutations.

• Now 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.
Example

In the next example, an equilibrium that resists local attention-partition mutations does not resist more general local mutations.

• An agent can choose between a hard and easy task.

• They believe that their performance $y \in \{y_L, y_M, y_H\}$ in each task is a function of their skill and a relative difficulty parameter.

• The agent’s utility is their performance, i.e., $u(a, y) = y$.

• The agent is overconfident about their skill, which makes them think that the frequency of high performance is sufficient to determine the optimal action.

• So they pool $y_L$ and $y_M$ in the attention partition.
• The resulting equilibrium resists local mutations: tracking only high performance the agent revises the difficulty of the tasks instead of their ability, which induces them to choose the suboptimal hard task.

• If the agents instead tracked the complete outcome distribution, mutations would lead them to update their belief about their skill, and switch to the objectively optimal task.

• We show by example that attention partitions can also have the opposite effect— an equilibrium might resist local mutations but not attention-partition mutations.
Each agent with the subjective model $\Theta$ has a prior $\mu_\Theta$ with support $\Theta$.

Let $p_\Pi \in \Delta(\Pi)$ be a distribution over strategies, and suppose that a continuum of agents each observes $n$ randomly drawn agents in the previous generation.

**Proposition**

Suppose that either $\Theta$ is finite or that there is a unique best reply to the KL-minimizers.

Then as the number of observed agents increases, the distribution of strategies in the population converges to a distribution over the best replies to the minimizers.

The convergence of beliefs follows from an argument similar to Berk [1966]; we use the fact that the likelihood ratio between KL minimizers is a random walk to show strategies converge.
Continuum of Actions

• The cardinality of the action space is irrelevant for one-hypothesis mutations.

• For local mutations, the cardinality of the action space does matter: With any fixed set of actions a vanishingly small $\varepsilon$ is eventually smaller than the “gap" between the actions, not so 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.

• So local mutations can invade some uniformly strict equilibria in settings with a continuum of actions.

• These unstable equilibria correspond to a limit of equilibria that are mixed and unstable along a sequence of increasingly fine finite action grids.
Conclusion

• We propose a model combining learning and evolutionary pressure among subjective models.

• We use mutations to help understand why some misspecifications are likely to persist for a long time.

• Mutations that lead to a better but imperfect fit in a statistical sense can lead to lower payoffs.

• We considered 2 sorts of mutations and characterize stability with respect to them.

• And we used our model to extend the analysis of attention-channeled mutations to allow actions to influence outcomes and to mutations other than to the true model.
Thanks!
Proof Sketch

- The proof has three steps:
- Since the observations are i.i.d., standard arguments show the posterior beliefs will assign probability one to the KL-minimizing parameters for $p_\Pi$.
- Then we show that the relative likelihood of two KL minimizers is a random walk, and use this to obtain convergence of beliefs. This part uses a technique similar to Fudenberg, Lanzani, and Strack (2020)– there we show that the probability of a specific KL-minimizing parameter is arbitrarily large i.o., while here we need that the entire distribution of beliefs converges.
- Finally, to show the limit distribution assigns probability zero to ties we show that the that the limit distribution of odds ratio is normal, and its covariance matrix is positive definite.

Back to slides
Subjective Expected Utility

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

• \( p_{t+1}^t(\Theta) = T(p^t)(\Theta) \), where \( T \) is continuous and such that the dynamic is payoff monotone, meaning that

\[
\frac{U^*(p_t(\cdot | \Theta))}{U^*(p_t(\cdot | \Theta'))} > (\geq) 1 \implies \frac{T(p)(\Theta)}{T(p)(\Theta')} = \frac{p_{t+1}^t(\Theta)}{p_{t+1}^t(\Theta')} > (\geq) \frac{p_t^t(\Theta)}{p_t^t(\Theta')}.
\]

(1)

Back to slides
Improving mutation

The \( \varepsilon \) mutation of a steady state \( \delta_\Theta \times \psi \) to \( \Theta' \) is improving if

\[
\min_{\theta \in \Theta'} \sum_{\pi \in \Pi} \psi(\pi) \sum_{s \in S} \sigma(s) H \left( q^* (\cdot | s, \pi(s)), q_\theta (\cdot | s, \pi(s)) \right) \\
< \min_{\theta \in \Theta} \sum_{\pi \in \Pi} \psi(\pi) \sum_{s \in S} \sigma(s) H \left( q^* (\cdot | s, \pi(s)), q_\theta (\cdot | s, \pi(s)) \right)
\]
Projection

Given a steady state $p$ we define the projection on hypothesis $l$ as

$$P_l(p): = \text{argmin}_{\theta \in \Theta^l} \sum_{s \in S} \sigma(s) \sum_{\pi \in \Pi} p_{\Pi}(\pi) H (q_{\theta}(|s, \pi(s)), q^*(|s, \pi(s))).$$

Back to slides
Solution

**Definition**

A sequence \((p^t)_{t \in \mathbb{N}_0} \in P^{\mathbb{N}_0}\) is a solution if satisfies equations

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

and

\[
\frac{U^*(p^t(\cdot|\Theta))}{U^*(p^t(\cdot|\Theta'))} > (=) 1 \quad \Rightarrow \quad \frac{p^{t+1}_K(\Theta)}{p^{t+1}_K(\Theta')} > (=) \frac{p^t_K(\Theta)}{p^t_K(\Theta')}.\]

for all \(t \in \mathbb{N}_0\).

Back to slides
A steady state is a $\hat{p} \in P$ such that $(\hat{p})_{t \in N_0}$ is a solution and $\hat{p}_K = \delta_{\Theta}$ for some $\Theta \in$. A steady state is unitary if $\hat{p}(\cdot|\Theta) \in \Delta(BR(\mu))$ for some $\mu \in \Delta(\Theta(\hat{p}_\Pi))$. 
A Berk-Nash equilibrium \((\Theta, \psi)\) resists mutation to \(\Theta'\) if there is a collection of solutions \((p_t^\varepsilon)_{t \in \mathbb{N}_0, \varepsilon \in (0,1)}\), such that \(p_0^\varepsilon\) is the \(\varepsilon\) mutation of \(\delta_\Theta \times \psi\) to \(\Theta'\), and

\[
\lim_{\varepsilon \to 0} \lim_{t \to \infty} (p_t^\varepsilon)_{\Pi} = \psi.
\]
Resistance to one-hypothesis mutations

**Proposition**

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

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

The additional restrictions of part 2 plays the role of quasi-strictness in the local mutation case, ruling out mutations that induce low performance but highly revealing strategies.

Back to slides