Esponda and Pouzo [2016]: Berk-Nash equilibrium

- Misspecified learning in static Bayesian games
- Each player privately observes a signal, which is their “type, ” and players then simultaneously choose actions.
- Agents have subjective beliefs about the map from strategy profiles to distributions of consequences.
- Berk-Nash equilibrium: each agent’s strategy (map from type to action) is optimal given beliefs, and beliefs minimize the KL divergence from what the agent sees.
- Applications include “cursed” behavior when people ignore the implications of the fact that they won an auction (Example 2.5) and ignoring regression to the mean (Example 2.3).
- Paper also provides a learning foundation for Berk-Nash equilibrium.

Model and Notation

- Players $i \in I$, states ω in a compact metric space Ω, finite signal space $\mathbb{S} = \times_{i \in I} \mathbb{S}^i$, where each i privately observes s_i.
- Objective probability distribution $p \in \Delta(\Omega \times \mathbb{S})$.
- Each i observes s_i, then chooses x_i from the finite set $\mathbb{X}^i$; $\mathbb{X} = \times_{i \in I} \mathbb{X}^i$
- Set of consequences $\mathbb{Y} = \times_{i \in I} \mathbb{Y}^i$.
- Main text of the paper restricts to finite Ω and Y, Online Appendix relaxes this (which is needed for some of the examples.) Stick with finite Ω and Y in these slides.
- Each player i observes the output of a consequence function $f^i : \Omega \times \mathbb{X} \to \mathbb{Y}^i$.
- Payoffs $\pi^i : \mathbb{X}^i \times \mathbb{Y}^i \to \mathbb{R}$ don't depend on the state ω or on the consequences y_j of the other agents.

- Strategies $\sigma^i : \mathbb{S}^i \to \Delta(\mathbb{X}^i)$
- Every strategy profile σ generates maps $Q^i_\sigma : \mathbb{S}^i \times \mathbb{X}^i \to \Delta(\mathbb{Y}^i)$ from signals and actions to distributions over each players' consequences; this is the objective model.
- Each agent i has a set of subjective models $\left(Q^i_{\theta^i}\right)_{\theta^i \in \Theta^i}$, where each $Q^i_{\theta^i} : \mathbb{S}^i \times \mathbb{X}^i \to \Delta(\mathbb{Y}^i)$. There are the models the agent believes are possible. (Implicitly this set is the support of the agent's prior, but the prior is only introduced when they get to the learning foundations.)
- Assume each Θ^i is a compact subset of a Euclidean space and that the Q^i are continuous in θ^i.
- Beliefs about the opponents' strategies are left implicit; agents may or may not be thinking about them.

- The KL divergence between a model's predictions and the true distribution is

$$K^i(\sigma, \theta^i) = \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_\sigma} \left[\ln \left(\frac{Q^i_\sigma(Y^i \mid s^i, x^i)}{Q^i_{\theta^i}(Y^i \mid s^i, x^i)} \right) \right] \sigma_i(x^i \mid s^i) p(s^i)$$

- The set of closest parameter values of player i given strategy σ is

$$\Theta^i(\sigma) = \operatorname*{argmin}_{\theta_i \in \Theta_i} K^i(\sigma, \theta^i).$$

- The agent's expected distribution over feedback given their beliefs is

$$\bar{Q}^i_{\mu^i}(y^i \mid s^i, x^i) \equiv \int_{\Theta^i} Q^i_{\theta^i}(y^i \mid s^i, x^i) \mu^i(d\theta^i).$$

When are there Multiple KL Miminimizers?

- In space of *all* probability distributions there is generically a unique KL minimizer.
- But frameworks with symmetry or parametric restrictions are not generic, and there multiple KL minimizers can arise naturally.
- Example: suppose that y is the color of the ball drawn from an urn which is known to contain 6 balls, with three possible colors, white, red, blue.
- The agent correctly believes their action doesn't affect y.
- Outcome distributions Q_θ correspond to the urn composition.
- The agent is certain that at most half of the balls have the same color, i.e., that $p(y) \leq 1/2$ for every y.
- In reality the urn has 4 white balls, 1 red, and 1 blue.
- So the two KL minimizers are (3 white, 2 blue, 1 red) and (3 white, 1 blue, 2 red).

Berk-Nash Equilibrium

Definition

A strategy profile σ is a Berk-Nash equilibrium (BNE) if for all $i \in I$, there is $\mu_i \in \Delta\left(\Theta^i\right)$ such that:

- $\operatorname{supp} \mu_i \subseteq \Theta^i(\sigma)$,
- σ^i is optimal given μ_i, i.e. $\sigma^i(x^i \mid s^i) > 0$ implies

$$x^i \in \operatorname{argmax}_{\bar{x}^i \in X^i} \sum_{y^i \in Y^i} \pi^i(\bar{x}^i, y^i) \bar{Q}^i_{\mu^i}(y^i \mid s^i, \bar{x}^i).$$

- If all agents consider the true model possible, then every Berk-Nash equilibrium is self-confirming: only the models for which $Q_{\theta^i} = Q^i_\sigma$ minimize the KL-divergence.

Theorem

A Berk-Nash equilibrium exists if either:

1. *For each $\theta \in \Theta$ and $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, $Q^i_{\theta^i}(\cdot|s^i, x^i)$ and $Q_{\delta_x}(\cdot|s^i, x^i)$ are mutually absolutely continuous*
2. *Every Q_{θ^i} can be approximated by a model where every feasible observation has full support.*

- The conditions aren't nested, but both are satisfied by the examples in the paper.
- We prove the theorem under the first assumption; the paper has a longer proof under the second.
- As with NE, pure-strategy best responses aren't convex valued, and existence of equilibrium can require mixed strategies.

Proof

- See the original game as $|2I|$ players game, $\{1, 1', 2, 2', ..., I, I'\}$.
- $i =$the original agent, $i' = i$'s "adversary."
- Action sets: $A_i = \mathbb{X}_i$, $A_{i'} = \Theta_i$.
- Utility functions:

$$U_i\left(a, \theta\right) = \int_Y u_i\left(a, y\right) dQ_{\theta'}\left(y|a\right),$$

$$U_{i'}\left(a, \theta\right) = \int_Y \log Q_{\theta_i}\left(y|a\right) dQ_{\delta_a}\left(y|a\right)$$

- i's adversary wants to choose a distribution on consequences to maximize the negative of KL divergence between i's beliefs and the true distribution.

- Denote mixed strategies by ψ_i and $\mu_{i'}$.
- Nash equilibrium requires both players only assign positive probability to their best responses.
- Player i''s pure best responses to ψ_i are

$$\begin{aligned} & \arg\max_{\theta_i \in \Theta} \sum_{a \in A} \prod_{i \in I} \psi_i (a_i) U_{i'} (a, \theta_i) \\ = & \arg\max_{\theta \in \Theta} \sum_{a \in A} \prod_{i \in I} \psi_i (a_i) \int_Y \log Q_{\theta_i} (y) \, dQ_{\delta_a} (y) \\ = & \Theta_i (\psi) . \end{aligned}$$

  where the first line follows from the fact that $U_{i'}$ doesn't depend on θ_j, $j \neq i$.
- Player i's best pure responses to $\mu_{i'}$ are $\arg\max_{a_i \in A_i} \mathbb{E}_{Q_{\mu_{i'}}} [u_i (a_i, y)]$.

- So a Nash equilibrium of the induced game is a pair (ψ^*, μ^*) where $\mu^* \in \Delta(\Theta(\psi^*))$ and $\psi^* \in \Delta\left(\arg\max_{a_i \in A_i} \mathbb{E}_{Q_{\mu^*}}[u_i(a_i, y)]\right)$.
- This is exactly a Berk-Nash equilibrium of the original game.
- When each Θ_i is finite, existence of Berk-Nash is an application of the Nash's result for games with finite set of strategies.
- When Θ_i is compact, the result follows from Glicksberg [1952] existence theorem for Nash equilibria of games with compact space of actions. (See e.g. Fudenberg and Tirole [1991] section 1.3.3.)
- To apply the theorem observe that Esponda and Pouzo [2016] assume A_i finite, Θ_i compact, and U_i continuous. The absolute continuity assumption guarantees that $U_{i'}$ is also continuous.

(fully) Cursed equilibrium (Eyster and Rabin [2005])

- Each agent i believes that conditional on their own signal s^i the state ω is independent of opponent's actions x^{-i}. (*ER also define "partially cursed" equilibrium...*)

- Motivated by evidence of the "winner's curse" in common-value auctions, where bidders fail to realize that it's bad news when others have low signals.

- EP show that with this sort of misspecification, if the state space is finite and each player gets perfect feedback (*see next slide*), a strategy profile is a cursed equilibrium iff it is Berk-Nash.

Example: Additive Lemons Problem

- Seller (player 1) owns an object and values it at $s^1 = \omega$.
- Buyer value is $v = \omega + 2.5$
- Double auction: Seller submits bid x^1, buyer submits bid $p = x^2$; trade at p iff $x^1 \leq x^2$.
- So seller sets $x^1 = \omega$.
- Distribution of ω is uniform on $\{1, 2, 3\}$.
- Objectively, payoffs are $(1/3)(1 - 1 + 2.5) = 5/6$ for $p = 1$, $(2/3)(-.5 + 2.5) = 4/3$ for $p = 2$ and $-1 + 2.5 = 1.5$ for $p = 3$.
- So NE price is 3.

Cursed equilibrium in the lemons example

- Buyer doesn't have a signal.
- Perfect feedback: $f^2(x, \omega) = (x^1, \omega)$.
- Buyer's models are all the product measures on $\mathbb{X}^1 \times \Omega$: buyer thinks seller's bid is independent of their price.
- BN-E: buyer learns that the distribution of seller bids and the distribution of ω are both uniform on $\{1, 2, 3\}$
- The buyer can't learn anything about the correlation between seller's bid and the value because they are convinced these are independent.
- Subjective payoff to $p = 1$ is $(1/3)(1 + 5/2) = 7/6$; subjective payoff to 2 is $(2/3)(5/2) = 5/3$ and subjective payoff to 3 is $-1 + 5/2 = 3/2$.
- Thus the Berk-Nash price is 2.

- Berk-Nash price is lower than in NE because buyer doesn't internalize the reduction in average quality conditional on the sellers accepting the offer.

- Buyer expects to get payoff $5/2$ conditional on purchase and actually gets 2, should we expect them to notice? If so how fast?

- See Gagnon-Bartsch, Rabin, and Schwartzstein [2023] and Fudenberg and Lanzani [2023].

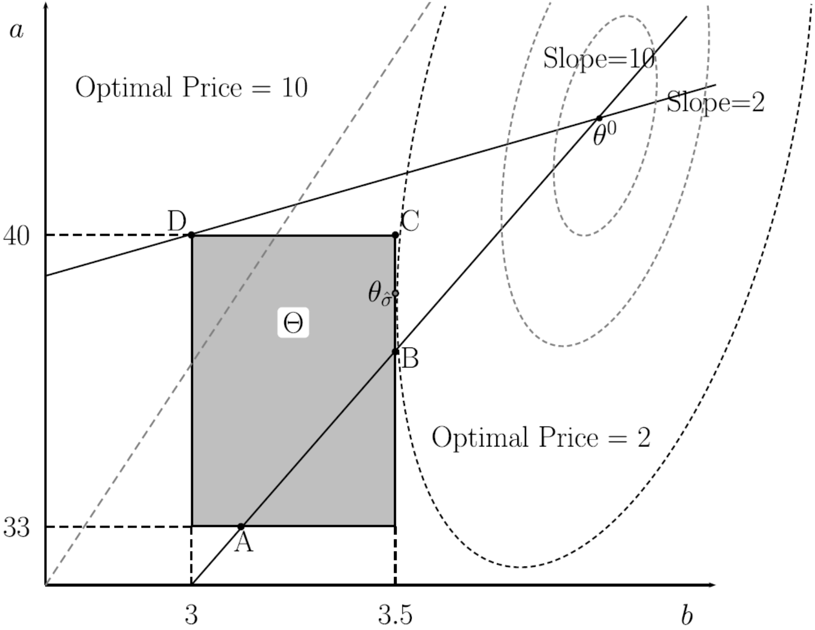
Monopolist with unknown demand

Nyarko [1991]

- Monopolist chooses price $x \in \{2, 10\}$
- Payoff: $\pi(x, y) = xy$
- Demand Function: $y = a^0 - b^0 x + \omega$.
- Noise term $\omega \sim N(0, 1)$, so $y \sim N(a^0 - b^0 x, 1)$.
- Monopolist believes $y \sim N(a - bx, 1)$ where $\theta = \{a, b\}$ is uniformly distributed on a square that doesn't contain (a^0, b^0).

Evolution of Beliefs and Actions

- Seller is myopic: sets prices each period to maximize that period's expected payoff.
- Nyarko shows that the price doesn't have a deterministic limit: If it converged to 2, the firm would come to believe it should charge 10, and if price converged to 10 firm eventually wants to charge 2.
- In contrast when the seller is correctly specified their beliefs converge even when the data is endogenous.
- Simpler observation: If price is fixed at 10 the KL minimizers are on the segment AB on the next figure, and if price is fixed at 2 the KL minimizers are on CD.


a
Optimal Price = 10
Slope=10
Slope=2
θ^0
D
C
40
$\theta_{\hat{\sigma}}$
Θ
B
Optimal Price = 2
33
A
3
3.5
b

Definition

A game is correctly specified given $\boldsymbol{\sigma}$ if for all $i \in I$ there is $\theta^i \in \Theta^i$ s.t. $Q^i_{\theta^i}(y^i \mid s^i, x^i) = Q^i_{\sigma}\left(y^i \mid s^i, x^i\right)$ for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$ and $y^i \in \mathbb{Y}^i$. The game is correctly specified if it is correctly specified for all σ; otherwise it is misspecified.

Definition

A game is strongly identified given $\boldsymbol{\sigma}$ if for all $i \in I$, if $\theta^i_1, \theta^i_2 \in \Theta^i(\sigma)$, then $Q^i_{\theta^i_1}(\cdot \mid s^i, x^i) = Q^i_{\theta^i_2}(\cdot \mid s^i, x^i)$ for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$.

Proposition

Berk-Nash equilibrium is equivalent to Nash equilibrium when the game is correctly specified and strongly identified for all σ.

Definition

Strategy profile σ is a unitary self-confirming equilibrium (SCE) if for each player i there is a conjecture $\mu^i \in \Delta(\Theta_i)$ such that for each s^i and x^i with $\sigma_i(s^i)(x^i) > 0$

a. x^i is a best response to μ^i given signal s^i, and
b. $\bar{Q}^i_{\mu^i}(\cdot \mid s^i, x^i) = Q^i_\sigma(\cdot \mid s^i, x^i)$.

Proposition

Berk-Nash equilibrium is equivalent to unitary self-confirming equilibrium when the game is correctly specified for all σ.

Berk-Nash in a bandit problem

- Out gives 0, this is outcome y_0.
- In leads to a move by Nature with outcomes y_1 and y_{-1}, payoffs 1 and -1 respectively.
- No signals.
- $\Omega = \{\omega_{0.1}, \omega_{0.6}\} = Q^i$; these are the probabilities Nature plays y_1.
- True distribution corresponds to $\omega_{0.6}$ so agent is correctly specified.
- Feedback function is that agent sees the realized outcome.
- When agent plays Out, both models give the same distribution on feedback.
- When agent plays In, the unique KL minimizer is the true model.

- "Out" together with $\mu(\omega_{0.1}) = .9$ is Berk Nash because with this belief the agent's expected payoff to In is $.9(-.8) + .1(.2) < 0$.

- Not Nash but a unitary SCE.

- *Proposition*: If a game is correctly specified at σ and σ is a Berk-Nash equilibrium then it is a SCE.

- If game isn't correctly specified, Berk-Nash isn't necessarily SCE: Misspecified agents needn't learn the path of play.

Molavi [2019]

- Extends EP 2016 to recursive dynamic general equilibrium macro settings with a continuum of agents and continuous actions, observables, and state variables.
- Prices and choices simultaneously determined through price-taking behavior and market clearing conditions.
- The economy has state variables and stochastic fluctuations so replace fixed "rest points" with ergodic distributions over aggregates.
- Uses results on continuous Markov chains over general state spaces to establish the convergence of empirical distributions and concentration of beliefs.
- Beliefs now minimize "weighted KL divergence" where the endogenous weights depend on the ergodic distribution of the limiting "temporary equilibrium." (cf Esponda and Pouzo, 2021)

Discussion

- Like EP, Molavi defends models of misspecified learning as a way to relax full rationality while letting behavior be endogenous.
- There is a long tradition of macro models with misspecified or boundedly rational learning—Bray (1982), Sargent (1993, 1999, ...), Marcet-Sargent (1989), Cho-Williams-Sargent (2002), Marcet-Nicolini (2003), Preston (2005), Adam-Marcet (2011), Evans-Honkapohja (2012), Malmendier-Nagel (2016), Eusepi-Preston (2018), etc.
- Molavi generalizes some of these (e.g. Bray [1982]) and is the special case of Adam-Marcet where beliefs come from Bayes rule. Molavi also gives a non-Bayesian MLE foundation that generalizes e.g. Evans-Honkapohja and others on OLS learning.

Fudenberg, Lanzani, and Strack [2021]

- Sharper necessary condition for an action to be a limit point of the learning process.
- A characterization of the actions that are limit points for all "nearby" beliefs.
- Sufficient conditions for an action to have positive probability of being the limit outcome from any initial beliefs.
- Main differences with previous work:
    1. Don't require random shocks to the payoff functions;
    2. Don't impose functional-form restrictions on the objective and subjective data generating processes;
    3. Don't assume myopia.

- Relate limit outcomes to two refinements of Berk-Nash equilibrium.
- A uniform Berk-Nash equilibrium is a best reply to *any* mixture over KL minimizers.
- A uniformly strict Berk-Nash equilibrium is an action that is a *strict* best reply to *every* mixture over KL minimizers.
- Any limit point must be a uniform Berk-Nash equilibrium.
- Uniformly strict Berk-Nash equilibria are uniformly stable: behavior converges to them with high probability from all nearby beliefs.
- Conversely, uniformly stable B-NE must be uniformly strict.
- Unif. Strict B-NE $=$ Unif. Stable $\subseteq$ Limit points $=$ Unif. B-NE.

Positive Attractiveness

- Equilibria are positively attractive if they have positive probability from any starting beliefs.

- Uniformly strict Berk-Nash equilibria are positively attractive under various types of misspecification:
  - Causation Neglect, where the agent mistakenly believes that their action does not affect the outcome distribution,
  - Subjective Bandits, where the agent thinks that the outcomes observed when playing one action are uninformative about the distribution induced by the others,
  - In supermodular environments, extremal equilibria are positively attractive.

Actions, Utilities, and Objective Outcome Distributions

- Every period $t \in \mathbb{N}$, the agent chooses an action a from the finite set A.

- Finite set of outcomes Y.

- Action a has two consequences:
  - Induces objective probability distribution over outcomes $p_a^* \in \Delta(Y)$;
  - Directly influences the agent's payoff through $u : A \times Y \to \mathbb{R}$.

Subjective Beliefs of the Agent

- Let $P := \times_{a \in A} \Delta(Y)$ be the space of all action-dependent outcome distributions.
- Elements $p \in P$, components p_a.
- The agent is Bayesian.
- They have a prior $\mu_0 \in \Delta(P)$.
- $\Theta := \mathrm{supp} \mu_0$ is the set of conceivable outcome distributions.
- The agent may be misspecified, i.e. $p^* \notin \Theta$.

Behavior of the Agent

- A history of length t is an $(a_\tau, y_\tau)_{\tau=0}^{t} = (a^t, y^t) \in A^t \times Y^t$.

- A (pure) policy $\pi : \bigcup_{t=0}^{\infty} A^t \times Y^t \to A$ specifies an action for every history.

- The agent wants to maximize expected discounted utility with discount factor $\beta \in [0, 1)$.

- $A^m(\mu) = \arg\max_{a \in A} \int_P \mathbb{E}_{p_a}[u(a, y)]\, d\mu(p)$ is the set of myopic best replies to belief μ.

- Two outcome distributions $p, p' \in \Theta$ are observationally equivalent under action a, written $p \sim_a p'$, if $p_a(y) = p'_a(y)$.
- Let $\mathcal{E}_a(p) \subseteq \Theta$ denote the outcome distributions in Θ that are observationally equivalent to p under a.
- Agents are not arbitrarily patient, so no reason to expect them to have much data about the consequences of every action.
- For each action a, let

$$\Theta(a) = \underset{p \in \Theta}{\operatorname{argmin}} \left(\sum_{y \in Y} p_a^*(y) \log p_a^*(y) - \sum_{y \in Y} p_a^*(y) \log p_a(y) \right)$$
$$= \underset{p \in \Theta}{\operatorname{argmin}} \left(- \sum_{y \in Y} p_a^*(y) \log p_a(y) \right)$$

denote the KL-minimizers (or *likelihood maximizers*)

Refinements of Berk-Nash Equilibrium

Definition (Uniform and Uniformly Strict Berk-Nash Equilibria)

Action a is a

(i) uniform Berk-Nash equilibrium if for every KL minimizing outcome distribution $p \in \Theta(a)$, there is a belief over the observationally equivalent distributions $\nu \in \Delta\left(\mathcal{E}_a(p)\right)$ such that $a \in A^m(\nu)$.

(ii) uniformly strict Berk-Nash equilibrium if $\{a\} = A^m(\nu)$ for every belief in $\nu \in \Delta(\Theta\left(a\right))$.

When the agent is correctly specified (i.e. $p^* \in \Theta$),

$$\text{Uniform B-NE} = \text{B-NE} = \text{Self-Confirming},$$

as p^*_a is the unique KL minimizer for a.

Technical Assumptions

- **Simplyfing assumption:** For all $p \in \Theta$, p and p^* are mutually absolutely continuous. This guarantees that no conceivable distribution is ruled out after a finite number of observations.
- Also assume that the prior μ_0 has subexponential decay: there is $\Phi : \mathbb{R}_+ \to \mathbb{R}$ such that for every $p \in \Theta$ and $\varepsilon > 0$ we have

$$\mu_0(B_\varepsilon(p)) \geq \Phi(\varepsilon)$$

  with

$$\lim \Phi(K/n) \exp(n) = \infty \qquad \forall K > 0.$$

- Priors with a density that is bounded away from 0 on their support, priors with finite support, and Dirichlet priors all have subexponential decay. Fudenberg, He, and Imhof [2017] show that Bayesian updating can behave oddly on priors w/o subexponential decay.

Only Uniform-Berk Nash Equilibria are Limit Actions

Theorem (Limit Actions are Uniform Berk-Nash Equilibria)

If actions converge to a $\in A$ with positive probability, a is a uniform Berk-Nash equilibrium.

- Other results on convergence to B-NE require myopia and either i.i.d. payoff shocks or a finite-support prior Esponda and Pouzo [2016], Frick, Iijima, and Ishii [2021], Bohren and Hauser [2021].
- Sharper conclusion: a limit action must be a best reply to all of the KL minimizers it induces.
- *Key lemma*: Beliefs of misspecified agents converge to the KL minimizers at a uniform rate.

Proof Sketch

- The agent's belief concentrates around the distributions that minimize the KL divergence from the empirical frequency at an exponential rate e^{Kt} that is uniform over the sample realizations.
- While playing a, the empirical frequency converges to p_a^*
- The difference between the empirical frequency and p_a^* is a random walk, and it oscillates in the direction of the different minimizers.
- By the Central Limit Theorem these oscillations die out at rate $\frac{1}{\sqrt{t}}$, which is slower than the exponential concentration of beliefs.

Proof Sketch

- So we can use an extension of the second Borel-Cantelli lemma for events that are not "too correlated" to show that infinitely often the beliefs concentrate around every minimizer.

- If a is not uniform B-NE, this induces the agent to switch to another action.

Non-convergence

- Nyarko [1991] shows by example that misspecified learning may not converge.
- There always exists a B-NE, but Fudenberg, Lanzani, and Strack [2021] shows there need not be a uniform BN-E.
- One case where they do exist is if the agent is correctly specified.
- If no equilibrium is uniform, actions cannot converge; this may be easier to check than directly verifying non-convergence.

Stability Notion

Definition (Stability)

A Berk-Nash equilibrium a is uniformly stable if for every $\kappa \in (0,1)$, there is an $\epsilon > 0$ such that for all initial beliefs $\nu \in \Delta(\Theta)$ such that $\nu(\Theta(a)) > 1 - \epsilon$, the action prescribed by *any* optimal policy converges to $a \in A$ with probability greater than $1 - \kappa$.

Theorem (Characterization theorem)

Action $a \in A$ is uniformly stable if and only if it is a uniformly strict Berk-Nash equilibrium.

- Theorem doesn't extend to strict BN-E that are not uniformly strict.
- In general there is a gap between uniformly strict BN-E and stability, but in "sufficiently rich" problems, this gap is absent (Theorem 3).

Proof Sketch for Uniformly Strict Implies Uniformly Stable

- Since a is a uniformly strict B-N equilibrium, a is the unique myopic best reply to every action-contingent outcome distribution p in a ball around the KL minimizers $\Theta(a)$.

- The agent needn't be myopic, and non-equilibrium actions can convey information, but when beliefs are concentrated around the minimizers, the subjective value of an alternative action can't much higher than its value against the most favorable minimizer, and since a is a uniformly strict BN-E the dynamic optimum policy is to play a.

- Then use the fact that a transformation of the odds-ratio between the non-KL minimizers and KL minimizers is a positive supermartingale (as in Frick, Iijima, and Ishii [2021]) to generalize the “active supermartingale” result of Fudenberg and Levine (1993) to misspecification.

- Use the Dubins upcrossing inequality to show that if this odd ratio starts sufficiently low, with an arbitrarily large probability it never crosses the threshold needed to switch action.

Positive Attractiveness

Definition (Positively Attracting)

Action $a \in A$ is positively attracting if for every optimal policy π

$$\mathbb{P}_\pi \left[\lim_{t \to \infty} a_t = a \right] > 0 \, .$$

Causation Neglect

- When the agent has causation neglect they believe that the distribution over outcomes is the same for all actions:

$$p_a = p_b \quad \forall a, b \in A, p \in \Theta.$$

Theorem

Suppose that the agent has causation neglect. If a is a uniformly strict Berk-Nash equilibrium, then it is positively attracting.

- *Example:* The agent is randomly matched with an opponent and believes they are playing a simultaneous move game, and they are uncertain about the distribution over strategies p in the opponents' population.
- In reality the opponents observe a noisy signal about the action taken by the agent before moving, so $p^*_a \neq p^*_b$ if $a = b$.

Sketch of the Proof of Positive Attractiveness

- The uniform consistency of beliefs guarantees that on every path of outcome realizations, beliefs concentrate around the empirical frequency.

- We use this concentration to show that if the empirical frequency is close to p_a^*, the beliefs concentrate around $\Theta(a)$.

- Causation neglect guarantees that the empirical frequency is a sufficient statistic.

- We combine this with our stability result to guarantee that once the beliefs get sufficiently close to the KL minimizers, the agent never switches to another action.

Summary

- All uniformly strict Berk Nash equilibria are uniformly stable, and only uniform Berk Nash equilibria can be limit points.

- Sufficient conditions for uniformly strict B-NE to be positively attracting under several forms of misspecification:
  - Causation Neglect;
  - Subjective Bandit Problems;
  - Extreme actions in Supermodular Environments.

- Missing: conditions that ensure convergence with probability 1.

References I

[1] J Aislinn Bohren and Daniel N Hauser. "Learning with heterogeneous misspecified models: Characterization and robustness." *Econometrica* (2021).

[2] Ignacio Esponda and Demian Pouzo. "Equilibrium in misspecified Markov decision processes." *Theoretical Economics* (2021).

[3] Ignacio Esponda and Demian Pouzo. "Berk–Nash equilibrium: A framework for modeling agents with misspecified models." *Econometrica* (2016).

[4] Mira Frick, Ryota Iijima, and Yuhta Ishii. "Welfare comparisons for biased learning." (2021).

[5] Drew Fudenberg and Giacomo Lanzani. "Which misperceptions persist?" *Theoretical Economics* (2023).

References II

[6] Drew Fudenberg, Giacomo Lanzani, and Philipp Strack. "Limit Points of Endogenous Misspecified Learning." *Econometrica* (2021).

[7] Drew Fudenberg and Jean Tirole. *Game theory*. MIT press, 1991.

[8] Tristan Gagnon-Bartsch, Matthew Rabin, and Joshua Schwartzstein. *Channeled attention and stable errors*. Harvard Business School, 2023.

[9] Irving L Glicksberg. "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points." *Proceedings of the American Mathematical Society* (1952).

[10] Yaw Nyarko. "Learning in mis-specified models and the possibility of cycles." *Journal of Economic Theory* (1991).