Berk-Nash Rationalizability*

Ignacio Esponda
UC Santa Barbara

Demian Pouzo
UC Berkeley

October 29, 2025

Abstract

We study learning in complete-information games, allowing the players' models of their environment to be misspecified. We introduce Berk–Nash rationalizability: the largest self-justified set of actions—meaning each action in the set is optimal under some belief that is a best fit to outcomes generated by joint play within the set. We show that, in a model where players learn from past actions, every action played (or approached) infinitely often lies in this set. When players have a correct model of their environment, Berk–Nash rationalizability refines (correlated) rationalizability and coincides with it in two-player games. The concept delivers predictions on long-run behavior regardless of whether actions converge or not, thereby providing a practical alternative to proving convergence or solving complex stochastic learning dynamics. For example, if the rationalizable set is a singleton, actions converge almost surely.

*We thank Attila Ambrus, Andreas Blume, Laurent Mathevet, and Kevin Reffet for helpful comments.
Esponda: iesponda@ucsb.edu. Pouzo: dpouzo@berkeley.edu

Contents

1 Introduction

A classic justification for equilibrium is dynamic: players observe past play, update beliefs, and adapt; if behavior settles, it should reflect some sort of equilibrium. In complete-information games, the standard notion is Nash equilibrium. For instance, in models of fictitious play—where each period's action is a best response to the empirical distribution of others' past actions—if the action profile itself settles down (i.e., becomes constant from some time onward), then it must be a Nash equilibrium. An analogous logic applies under misspecified learning: when agents update within a (possibly wrong) model class using data from play, if the action profile settles, it must settle to a Berk–Nash equilibrium. The key message is not that learning guarantees convergence, but that, conditional on stabilization of actions, the stabilized outcome must be an equilibrium of the relevant notion (Nash under correct specification; Berk–Nash under misspecification).

However, this logic is conditional: it has bite only if play actually stabilizes—and in many learning models, stabilization is not guaranteed or hard to assess. In fictitious play, paths can cycle rather than settle (Shapley (1964)), and positive convergence is typically proved only for special classes (e.g., Monderer and Shapley (1996), Hofbauer and Sandholm (2002)). In misspecified learning, many application-specific papers establish convergence under environment-tailored or parametric assumptions.[1] More general analyses—using stochastic approximation, differential inclusions, and martingale methods—characterize asymptotic behavior in broader settings (e.g., Fudenberg, Lanzani and Strack (2021); Esponda, Pouzo and Yamamoto (2021); Frick, Iijima and Ishii (2023); Murooka and Yamamoto (2023)). These tools are powerful but technically demanding and not always straightforward to apply; they may require favorable initial conditions, yield conclusions only with positive probability, or leave some environments uncovered. As a result, they do not always deliver clear long-run predictions, especially when actions fail to converge or convergence is hard to verify.

We take a complementary approach. Rather than assume or try to prove convergence, we ask what can be said about the actions that are played (or approached) infinitely often. To that end, we introduce *Berk-Nash rationalizability* for simultaneous-moves games of complete information. This solution concept is the largest self-justified set of actions: actions that justify themselves as best responses when beliefs are fit to a single common forecast of joint play. In the special case where each player's model is correctly specified and identified

[1] Examples include Nyarko (1991), Fudenberg, Romanyuk and Strack (2017), Heidhues, Kőszegi and Strack (2018, 2021), Bohren and Hauser (2021), He (2022), Ba and Gindin (2023), and Murooka and Yamamoto (2025) among others.

(so Berk–Nash equilibrium coincides with Nash), Berk–Nash rationalizability is analogous to correlated rationalizability but adds the restriction of a common joint forecast of play. This additional belief discipline refines correlated rationalizability, which permits separate player-by-player conjectures that need not derive from any shared joint forecast.

We show that, in a broad learning environment, every action played—or approached—infinitely often is Berk–Nash rationalizable. Past actions are publicly observed, and this common record underpins both belief updating and conjectures about others' play. Each player is Bayesian about the payoff-relevant parameters of a (possibly misspecified) model of the game and updates posteriors period by period. To forecast others' actions, players use fictitious-play-style rules: forecasts track empirical frequencies and can be interpreted as Bayesian updating over opponents' behavior. In each period, players choose a myopic best response to their current parameter beliefs and these forecasts.

The result delivers convergence-free predictions: it characterizes what can recur on the path even when actions or beliefs do not settle. When the Berk–Nash rationalizable set is a singleton, it implies almost-sure convergence; when it is larger, it provides tight bounds on long-run behavior. The concept is computationally tractable (iterative best-response under data-consistent beliefs), travels naturally from single-agent applications to games, and—under correct specification and identification—refines (correlated) rationalizability, sharpening predictions for $N > 2$ players.[2] For applied work on misspecification, it offers a usefool tool: one can study limiting behavior without solving dynamics, heavy regularity assumptions, or verifying convergence theorems, while maintaining a clear link to equilibrium analysis.

Rationalizability was introduced by Bernheim (1984) and Pearce (1984), with Pearce characterizing it explicitly as the largest set that is self-justified under the best-response operator.[3] In these classical formulations, a player's belief about opponents' play is a product measure (independence across opponents). Brandenburger and Dekel (1987) extended this to correlated rationalizability, allowing arbitrary joint beliefs over opponents' actions—the version we adopt. The standard epistemic foundation is common knowledge of rationality (Brandenburger and Dekel (1987) , Tan and da Costa Werlang (1988)); by contrast, we use the concept for its learning-based predictive content—bounding long-run behavior implied by Bayesian learning rather than modeling epistemic assumptions.

Milgrom and Roberts (1991) establish a link between adaptive learning and (correlated)

[2]For $N = 2$ players, our solution concept coincides with correlated rationalizability in the special class of correctly specified and identified games.

[3]Pearce called this property the best response property.

rationalizability: in complete-information games, under broad adaptive rules whose forecasts place vanishing probability on profiles not played infinitely often, play converges to (correlated) rationalizable outcomes. When the game is correctly specified and identified, our environment overlaps with theirs. We obtain a refinement by strengthening the adaptive discipline: we require forecasts to be consistent with empirical frequencies (a common, data-consistent joint forecast). This restriction has bite—there are examples where all actions are classically rationalizable, yet only one remains under our consistency requirement.[4] Moreover, if we relax the forecasting discipline to allow players to learn from different subsamples of past data (in the spirit of Milgrom–Roberts' adaptive rules), our solution concept in the special case of correctly specified and identified games collapses to (correlated) rationalizability. These ideas extend to misspecified settings, motivating the notion of Berk–Nash *weak* rationalizability as the version without the common restriction on beliefs.

Fudenberg and Kreps (1993) study stochastic fictitious play, where players' payoffs are perturbed each period so that, to an outside observer, intended play looks mixed. They show that if intended mixed strategies converge, the limit must be a Nash equilibrium—providing the first learning justification for mixed-strategy Nash equilibrium. By contrast, in standard (unperturbed) fictitious play, even if the empirical distribution of actions converges, it need not converge to a Nash distribution because a common history can induce correlation across players' actions. Payoff perturbations break this history-induced correlation, which is exactly what underpins the Fudenberg–Kreps result. We extend our framework to allow such payoff perturbations. This yields a version of Berk–Nash rationalizability in which limiting beliefs are fitted to independent forecasts of opponents' strategies (no correlation from shared history). The concept, however, does not collapse to Bernheim–Pearce independent rationalizability, because the shared data still impose common (now independence-compatible) belief restrictions across players.

Esponda and Pouzo (2016, henceforth EP16) introduce a misspecified learning framework to capture both systematic biases and limits to accurate representation of the environment

[4]For the special class of supermodular games, Milgrom and Roberts (1990) show that the set of rationalizable outcomes is bounded by the smallest and largest equilibria. Because equilibria are always rationalizable, imposing a common-forecast restriction as we do does not change these bounds.

due to complexity or informational constraints.[5] They define the *Berk–Nash equilibrium* (possibly in mixed strategies) and show that, with Fudenberg–Kreps–style payoff perturbations, convergence of intended strategies implies convergence to a Berk–Nash equilibrium. Esponda, Pouzo and Yamamoto (2021) study single-agent settings without perturbations and prove that convergence of the empirical action distribution (without payoff perturbations) leads to a generalized notion of Berk–Nash equilibrium. The same implication applies in our framework when payoff perturbations are absent.

The Berk–Nash rationalizable set arises via sequential elimination of non-rationalizable actions. Prior work uses iterative arguments to bound long-run behavior: Heidhues, Kőszegi and Strack (2018) obtain bounds that collapse to a singleton, and He (2022) adapts this approach to a specific multi-dimensional setting. Frick, Iijima and Ishii (2023) develop a general elimination procedure for beliefs and—using a prediction-accuracy preorder with a supermartingale construction—prove monotone elimination and eventual convergence of beliefs to the resulting set; their framework also applies to social learning, which we do not study. Their convergence result requires the existence of a continuous selection from beliefs to optimal actions, a condition typically violated with finite action sets or when optimal actions are not unique for a given belief. By contrast, our approach is at the action level, uses a simple argument based only on off-the-shelf extensions of Berk's asymptotic-belief characterization (the EP16 extension), accommodates finite action sets (requiring only upper hemicontinuity of the action correspondence—no continuous selection needed), and extends naturally to multiple agents.[6]

We focus on games of complete information. Extending the analysis to asymmetric information raises both conceptual and technical issues. Conceptually, one must decide what the long-run object is: a limit profile of realized states, signals, and actions, or a limit strategy profile mapping signals to actions. The latter is often more informative—for example, it can encode monotonicity in private signals, which the former cannot. Technically, forecasts of

[5]Examples that broadly fit and relate to misspecification include behavioral gametheoretic models Jehiel (2005); Eyster and Rabin (2005); market adverse selection with misspecified beliefs Esponda (2008); learning with selective attention Schwartzstein (2014); social learning and herding under misspecified inference Bohren (2016); Frick, Iijima and Ishii (2020); incorrect causal models Spiegler (2016); overconfidence Heidhues, Kőszegi and Strack (2018); misspecification in a a recursive general equilibrium model Molavi et al. (2019); narrativebased misspecification Eliaz and Spiegler (2020); biased social learning due to assortativity neglect and attribution errors, with implications for inequality and discrimination Frick, Iijima and Ishii (2022); Chauvin (2023); selective memory Fudenberg, Lanzani and Strack (2024); welfare comparisons under biased beliefs Frick, Iijima and Ishii (2024); mislearning from prices He and Libgober (2025); and agents misinterpreting their own motives Heidhues, Kőszegi and Strack (2023); among others.

[6]Frick, Iijima and Ishii (2023)'s supermartingale method is also useful elsewhere—for instance, to characterize convergence in "slow-learning" regimes where Berk–Nash tools are uninformative, and in Fudenberg, Lanzani and Strack (2021) to show stability of uniformly strict Berk–Nash equilibria.

others' play must be conditional on private signals; with a continuum of signals this calls for signal-conditional learning rules and, in effect, consistent nonparametric estimation of opponents' conditional behavior. We pursue this extension in a sequel (Esponda and Pouzo (2025)).

Because much of the convergence literature centers on single-agent setting—and our ideas are most transparent there—we begin with that case in Section 2. Section 3 develops the general simultaneous-moves, complete-information game framework. Section 4 shows that limit actions are Berk–Nash rationalizable when agents learn the game's parameters and forecast opponents' strategies. Section 5 relates our solution concept to existing notions of rationalizability. Section 6 presents extensions: an equivalent distributional version; payoff perturbations yielding a mixed-strategy variant that eliminates history-induced correlation in play (but not in beliefs); player-specific histories that lead to a weaker form of Berk–Nash rationalizability without joint belief restrictions; and discussions of the relationship with correlated equilibrium and curb sets.

2 Single-agent problem

Primitives. A single agent chooses an action $a \in \mathbb{A}$. Consequences take values in $\mathbb{Y}$. The true consequence kernel is $Q : \mathbb{A} \to \Delta\mathbb{Y}$. The agent does not necessarily know Q but entertains a parametric model $\{Q_\theta : \theta \in \Theta\}$, with $Q_\theta : \mathbb{A} \to \Delta\mathbb{Y}$. The payoff function is $\pi : \mathbb{A} \times \mathbb{Y} \to \mathbb{R}$.

We state assumptions explicitly in Section 3 where this environment is a special case. In particular, all sets are subsets of Euclidean spaces, and we require compactness of $\mathbb{A}$ and Θ and certain regularity and continuity assumptions on Q, Q_θ, and π but allow both discrete and continuous actions and consequences spaces. Throughout this section, we use the following example from Heidhues, Kőszegi and Strack (2018) to illustrate definitions, results, and their application.

Example (Returns to effort). A single agent chooses effort $a \in \mathbb{A} = [0, \infty)$. Outcomes are real-valued ($\mathbb{Y} = \mathbb{R}$). The true model is

$$y \;=\; \big(\alpha^* + a\big)\theta^* \;+\; \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0,1),$$

i.e., the true kernel is $Q(\cdot \mid a) = \mathcal{N}\big((\alpha^* + a)\theta^*,\, 1\big)$, where $\alpha^* \geq 0$ is true ability and $\theta^* > 0$ is the true return to effort. The agent entertains a parametric model $\{Q_\theta : \theta \in [0, \overline{\theta}]\}$ with

perceived ability $\alpha > 0$ (held fixed, not learned):

$$y \;=\; \big(\alpha + a\big)\theta \;+\; \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0,1),$$

so $Q_\theta(\cdot \mid a) = \mathcal{N}\big((\alpha + a)\theta,\, 1\big)$. Overconfidence corresponds to $\alpha > \alpha^*$ and underconfidence to $\alpha < \alpha^*$.

Payoffs are $\pi(a, y) = y - c(a)$, where c is differentiable and strictly convex, and satisfies $c(0) = c'(0) = 0$, and $c'(a) \to \infty$ as $a \to \infty$. Although $\mathbb{A}$ is unbounded, marginal cost diverges and returns are bounded ex ante (by $\overline{\theta}$, which we pick to be large enough so it does not bind), so the optimal effort lies in a finite interval, effectively making the feasible choice set compact. ♦

Optimal action correspondence Define the expected payoff of action a under model parameter θ by $U(a, \theta) := \int_{\mathbb{Y}} \pi(a, y)\, Q_\theta(dy \mid a)$. For a belief $\mu \in \Delta\Theta$, the optimal action correspondence is

$$F(\mu) \;:=\; \arg\max_{a \in \mathbb{A}} \int_\Theta U(a, \theta)\, \mu(d\theta).$$

Kullback–Leibler divergence. The *KL divergence* is a function $K : \Theta \times \mathbb{A} \to \mathbb{R}$, defined for any parameter value $\theta \in \Theta$ and action $a \in \mathbb{A}$ as

$$K(\theta, a) := \int_{\mathbb{Y}} \ln\left(\frac{q(y \mid a)}{q_\theta(y \mid a)}\right) q(y \mid a)\, \nu(dy). \tag{1}$$

Furthermore, for any action distribution $\sigma \in \Delta\mathbb{A}$, we define

$$\Theta^m(\sigma) := \arg\min_{\theta \in \Theta} \int K(\theta, a)\sigma(da).$$

The KL divergence measures the 'distance' between the true model Q and the parametric model Q_θ. The set $\Theta^m(\sigma)$ consists of the parameter values $\theta \in \Theta$ whose associated model Q_θ provides the best fit to the true model, in the sense of minimizing the expected Kullback–Leibler divergence given data generated by Q when actions are drawn according to σ.

Example (continued). The KL divergence is

$$K(\theta, a) = \frac{1}{2}\left((\alpha + a)\theta - (\alpha^* + a)\theta^*\right)^2.$$

There is a unique minimizer of the expected KL divergence for any $\sigma \in \Delta\mathbb{A}$, which is a

convex combination of the single-action minimizers: $\theta^m(\sigma) = \int \theta^m(\delta_a)\,\eta_\sigma(da)$, with weights $\eta_\sigma(da) \propto (\alpha + a)^2\,\sigma(da)$ and normalized to integrate to one. The single-action minimizer is

$$\theta^m(\delta_a) = \theta^* + \theta^*\,\frac{\alpha^* - \alpha}{\alpha + a}. \tag{2}$$

Overconfidence ($\alpha > \alpha^*$) implies underestimation of the returns to effort ($\theta^m(\delta_a) < \theta^*$, and therefore $\theta^m(\sigma) < \theta^*$); underconfidence ($\alpha < \alpha^*$) implies the reverse. The bias $|\theta^m(\delta_a) - \theta^*|$ falls with a and vanishes as $a \to \infty$, since large effort dilutes the impact of ability. ♦

Solution concept. For each Borel set of actions $A \subseteq \mathbb{A}$, we define the *best response set* as $\Gamma(A) := F(\cup_{\sigma \in \Delta A} \Delta\Theta^m(\sigma))$. Equivalently,

$$\Gamma(A) = \{a \in \mathbb{A} : \exists \sigma \in \Delta A, \mu \in \Delta\Theta^m(\sigma) \text{ such that } a \in F(\mu)\}.$$

In other words, the set $\Gamma(A)$ consists of all actions that the agent might choose (according to the optimal action correspondence F) when she assigns probability one to the set of models that provide the best fit under some action distribution with support in A. This concept aligns with our earlier interpretation of KL divergence. Specifically, if feedback about consequences arises from actions drawn from the set A, then the models that provide the best fit are those that minimize KL divergence for some action distribution supported on A. Consequently, the agent will follow actions that are optimal for beliefs that assign probability one to these best-fit models.

Example (continued). The optimal action corresponding to a degenerate belief δ_θ is $(c')^{-1}(\theta)$, since the agent chooses a to solve $\theta = c'(a)$. For any Borel set $A \subseteq \mathbb{A}$, the best response operator is

$$\Gamma(A) = \left\{(c')^{-1}(\theta^m(\sigma)) : \sigma \in \Delta A\right\}. \tag{3}$$

♦

We are now ready to define our solution concept for this environment.

Definition 1. *An action a is Berk–Nash rationalizable if there exists a set $A \subseteq \mathbb{A}$ such that $a \in A$ and $A \subseteq \Gamma(A)$.*

A Berk–Nash rationalizable action lies in a set A that is *self-justified* under the operator Γ, i.e., $A \subseteq \Gamma(A)$. Every action in such a set is rationalizable: it is optimal for some belief supported on parameter values that best fit data generated by a distribution over A—and different actions may be justified by different distributions. The set of all rationalizable

actions is the union of all self-justified sets. Later, we consider a dynamic model with Bayesian updating from past actions and outcomes, and show that every limit action is rationalizable. [7]

The solution concept typically used in the literature is that of a Berk–Nash equilibrium (EP16). In terms of our best response operator, an action $a \in \mathbb{A}$ is a *Berk–Nash equilibrium* if $a \in \Gamma(\{a\})$. An immediate implication is that a Berk–Nash equilibrium action is rationalizable, but the converse is not necessarily true. In particular, the set of rationalizable actions may be larger than the set of equilibrium actions, as we now illustrate.

Characterization. The operator Γ is monotone and, under our assumptions, maps compact sets into compact sets. A well-known implication is that the union of sets that self-justified under Γ (in our case, the Berk–Nash rationalizable set, $\mathcal{B}$) is the largest fixed point of Γ and can be obtained by iteratively applying Γ starting from the largest set $\mathbb{A}$. Let $\mathcal{B}^0 = \mathbb{A}$ and $\mathcal{B}^{k+1} = \Gamma(\mathcal{B}^k)$; then $\mathcal{B}^{k+1} \subseteq \mathcal{B}^k$ for all k and

$$\mathcal{B}^* \;=\; \bigcup\{\mathcal{B} \subseteq \mathbb{A} :\; \Gamma(\mathcal{B}) \subseteq \mathcal{B}\} \;=\; \bigcap_{k\geq 0} \mathcal{B}^k \;=\; \lim_{k\to\infty} \Gamma^k(\mathbb{A}). \tag{4}$$

Example (continued). For compact A, expression (3) for Γ can be simplified further. Since $\theta^m(\sigma)$ is a convex combination of the single-action minimizers, it lies in the convex hull of the set $\{\theta^m(\delta_a) : a \in A\}$. Moreover, because $\theta^m(\delta_a)$ is continuous on the compact set A, it attains its minimum and maximum, and this convex hull is simply the interval

$$\left[\min_{a\in A} \theta^m(\delta_a),\; \max_{a\in A} \theta^m(\delta_a)\right].$$

Applying $(c')^{-1}$ to this interval, we conclude that

$$\Gamma(A) = \left[(c')^{-1}\left(\min_{a\in A} \theta^m(\delta_a)\right),\; (c')^{-1}\left(\max_{a\in A} \theta^m(\delta_a)\right)\right]. \tag{5}$$

This step uses the fact that the image of a closed interval under a continuous, strictly increasing function, $(c')^{-1}$, is again a closed interval.

This characterization of Γ motivates the definition of a simpler mapping that acts directly

[7] Our definition is similar, but not exactly equivalent to rationalizability in a game where player 1 chooses actions and player 2 chooses model parameters. That interpretation would correspond to a larger operator $F\left(\Delta \bigcup_{\sigma\in\Delta A} \Theta^m(\sigma)\right)$, which allows arbitrary mixtures over model parameters fit to different action distributions. In contrast, our definition restricts attention to beliefs supported on $\Theta^m(\sigma)$ for some fixed $\sigma \in \Delta A$.

on actions. Define $T : \mathbb{A} \to \mathbb{A}$ by

$$T(a) := (c')^{-1}(\theta^m(\delta_a)),$$

and note that the set of fixed points of T coincides with the set of Berk–Nash equilibrium actions. We will show that the mapping T also characterizes the best response operator Γ.

Overconfidence ($\alpha > \alpha^*$). In this case, the function $\theta^m(\delta_a)$ is increasing in a, so T is increasing. An increasing function may have multiple fixed points, so multiple Berk–Nash equilibria are possible. For any interval $A = [L, H]$, it follows from (5) that

$$\Gamma(A) = [T(L), T(H)].$$

To characterize the limit of iterated best responses, define a sequence of intervals $A^k = [a^k_{\min}, a^k_{\max}]$ by

$$a^{k+1}_{\min} = T(a^k_{\min}), \quad a^{k+1}_{\max} = T(a^k_{\max}),$$

starting from $a^0_{\min} = 0$ and $a^0_{\max} = \bar{a}$, where $\bar{a}$ is an upper bound on optimal actions. Then $(a^k_{\min})_k$ increases, $(a^k_{\max})_k$ decreases, and both sequences converge. The limits

$$a^\infty_{\min} = \lim_{k\to\infty} a^k_{\min}, \quad a^\infty_{\max} = \lim_{k\to\infty} a^k_{\max}$$

are fixed points of T (hence, equilibria), and, by the characterization in (4), the limiting interval $[a^\infty_{\min}, a^\infty_{\max}]$ is the set of all rationalizable actions. This is the case in Figure 1a, where $a_S = a^\infty_{\min}$ is the smallest Berk–Nash equilibrium and $a_L = a^\infty_{\max}$ is the largest equilibrium. If instead T had a unique fixed point, the limit would be a singleton and rationalizability would coincide with equilibrium.

Underconfidence ($\alpha < \alpha^*$). In this case, the function $\theta^m(\delta_a)$ is decreasing in a, so T is decreasing. A decreasing function has at most one fixed point, so there is a unique Berk–Nash equilibrium. For any interval $A = [a_{\min}, a_{\max}]$, we have

$$\Gamma(A) = [T(a_{\max}), T(a_{\min})].$$

To analyze the dynamics of Γ, define

$$a^{k+1}_{\min} = T(a^k_{\max}), \quad a^{k+1}_{\max} = T(a^k_{\min}),$$

again starting from $[0, \bar{a}]$. Then $(a^k_{\min})_k$ increases, $(a^k_{\max})_k$ decreases, and both sequences

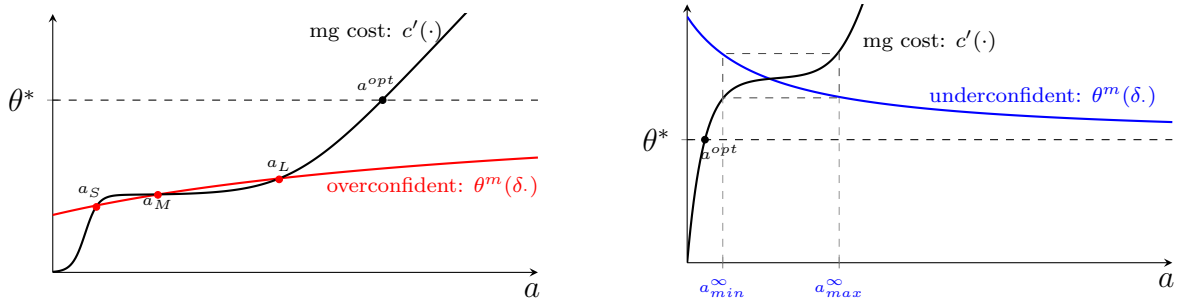
converge to

$$a^{\infty}_{\min} = \lim_{k\to\infty} a^{k}_{\min}, \quad a^{\infty}_{\max} = \lim_{k\to\infty} a^{k}_{\max},$$

By by the characterization in (4), the limiting interval $[a^{\infty}_{\min}, a^{\infty}_{\max}]$ is the set of rationalizable actions. The limits also satisfy

$$T(a^{\infty}_{\min}) = a^{\infty}_{\max}, \quad T(a^{\infty}_{\max}) = a^{\infty}_{\min},$$

so they form a 2-cycle of T and are fixed points of T^2. If T^2 has a unique fixed point, then it must also be a fixed point of T, and the limit is a singleton. In that case, rationalizability coincides with equilibrium. If T^2 has multiple fixed points, then $a^{\infty}_{\min}$ and $a^{\infty}_{\max}$ are the smallest and largest among them. This latter situation is illustrated in Figure 1b. ♦



(a) Overconfident agent

(b) Underconfident agent

Figure 1: Returns to effort example.
The optimal action a^{opt} is where the marginal cost curve intersects the true return θ^*. Berk–Nash equilibrium actions lie at intersections of the marginal cost curve and the KL-minimizing curve $\theta^m(\delta_\cdot)$. In the overconfident case, there are 3 Berk–Nash equilibria, a_S, a_M, and a_L, and the Berk–Nash rationalizable set is $[a_S, a_L]$; in the underconfident case, there is a unique Berk–Nash equilibrium, and the Berk–Nash rationalizable set is the 2-cycle interval $[a^{\infty}_{\min}, a^{\infty}_{\max}]$.

Limit actions are rationalizable. We consider an agent who learns by Bayesian updating and chooses myopically optimal actions over time. The agent starts from a full-support prior $\mu_0 \in \Delta\Theta$ and, at each discrete time $t = 1, 2, \ldots$:

- holds a belief $\mu_t \in \Delta\Theta$;
- chooses $a_t \in F(\mu_t)$ from the optimal action correspondence;
- observes consequence $y_t \sim Q(\cdot \mid a_t)$;
- updates μ_{t+1} by Bayes' rule.

An action $a \in \mathbb{A}$ is called a *limit action* of the sequence $a^\infty = (a_1, a_2, \ldots)$ if there exists a subsequence $(a_{t_k})_k$ such that $a_{t_k} \to a$ as $k \to \infty$. Equivalently, a is a limit action if, for every open neighborhood $U \subseteq \mathbb{A}$ of a, there are infinitely many times $t \in \mathbb{N}$ such that $a_t \in U$. When $\mathbb{A}$ is finite, an action is a limit action if and only if it is played infinitely often.

Theorem 1. *Almost surely, every limit action of a Bayesian agent is Berk–Nash rationalizable.*

Theorem 1 implies that asymptotic behavior is confined to Berk–Nash rationalizable actions: any non-rationalizable action cannot be a limit action and, in a finite action set, is played only finitely often (i.e., is eventually eliminated).

Example (continued). Proving convergence in these environments is far from trivial. In the *overconfident* case, when the Berk–Nash equilibrium (BNE) is unique, Heidhues, Kőszegi and Strack (2018) prove convergence to that equilibrium; in our setting this becomes an immediate corollary of Theorem 1, since the unique equilibrium is also the unique rationalizable action. With multiple equilibria, their result does not apply; nonetheless, our theorem delivers asymptotic bounds: in the overconfident case every limit action lies between the smallest and largest equilibrium. By contrast, Heidhues, Kőszegi and Strack (2021) cover *both* over- and underconfidence but analyze a *different* data-generating process in which the disturbance enters the production function (noise inside production) and impose Gaussian priors (more generally, a one-dimensional sufficient statistic). That specification does *not* encompass our baseline setting $y_t = Q(a_t, \theta) + \varepsilon_t$, where noise is outside production. The *underconfident* case in our example is particularly challenging: the Berk–Nash equilibrium is unique yet unstable in the sense of Esponda, Pouzo and Yamamoto (2021). Even so, Theorem 1 provides usable bounds: long-run actions must lie within the largest two-cycle, which contains the equilibrium. More generally, one could appeal to the stochastic-approximation approach in Esponda, Pouzo and Yamamoto (2021) to obtain bounds, extended to a continuum of actions by Murooka and Yamamoto (2023), but this route requires characterizing solutions of the relevant differential inclusions. ♦

Later in Section 3 we analyze a more general environment in which Theorem 1 appears as a special case. For intuition, we sketch the proof in this single-agent case, where the ideas are easiest to see. Take the set of limit actions, $\mathcal{A}$, and pick any $a \in \mathcal{A}$. From the infinite history, select a subsequence along which actions converge to the limit action a, and, by passing to a further subsequence if needed, the posteriors converge to μ and the empirical action distribution to σ. Because σ is a limit of empirical frequencies, it assigns probability only to actions that actually recur, so $\operatorname{supp}(\sigma) \subseteq \mathcal{A}$. We slightly extend the argument

of EP16 (itself an adaptation of Berk, 1966) to conclude that any limiting belief satisfies $\mu \in \Delta\Theta^m(\sigma)$ (best-fit parameters for σ). Since actions are optimal along the subsequence and the best-response correspondence F has a closed graph, limits of optimizers are optimal, so $a \in F(\mu)$. Hence, for any a in the set of limit actions $\mathcal{A}$, there exist $\sigma \in \Delta\mathcal{A}$ and $\mu \in \Delta\Theta^m(\sigma)$ with $a \in F(\mu)$, which shows $a \in \Gamma(\mathcal{A})$. Therefore, $\mathcal{A} \subseteq \Gamma(\mathcal{A})$ (i.e., the set of limit actions is self-justified under Γ) and, by definition, every limit action is Berk–Nash rationalizable.

3 Games

In this section we extend the framework to strategic environments with multiple decision makers. In particular, we study simultaneous-moves games of complete information.

3.1 Setup

There is a finite set of players I. For each $i \in I$, the individual action space is $\mathbb{A}^i$ and the joint action space is $\mathbb{A} := \times_{i \in I} \mathbb{A}^i$. Player i's consequences take values in $\mathbb{Y}^i$. The true environment is described by a mapping $Q^i : \mathbb{A} \to \Delta\mathbb{Y}^i$, which assigns to every action profile a probability distribution over player i's consequences. As before, player i does not know Q^i but entertains a parametric model $Q^i_{\theta^i} : \mathbb{A} \to \Delta\mathbb{Y}^i$ with parameter $\theta^i \in \Theta^i$. An action profile is an element $a = (a^i)_{i \in I} \in \mathbb{A}$, and $\sigma \in \Delta\mathbb{A}$ denotes an action-profile distribution. The payoff function of player i is $\pi^i : \mathbb{A}^i \times \mathbb{Y}^i \to \mathbb{R}$, where the distribution of $y^i \in \mathbb{Y}^i$ already incorporates the influence of all players' actions via Q^i.

We impose the following assumptions.[8]

Assumption (games). (i) *Spaces.* For each i, the sets $\mathbb{A}^i$ and Θ^i are nonempty compact subsets of Euclidean spaces; $\mathbb{Y}^i$ is a Borel subset of a Euclidean space. (ii) *Densities and a.e. continuity.* For each i there exists a Borel measure ν^i on $\mathbb{Y}^i$ such that, for every $a \in \mathbb{A}$, $Q^i(\cdot \mid a) \ll \nu^i$ with density $q^i(\cdot \mid a)$, and for every $\theta^i \in \Theta^i$, $Q^i_{\theta^i}(\cdot \mid a) \ll \nu^i$ with density $q^i_{\theta^i}(\cdot \mid a)$; moreover, for ν^i-a.e. y, the maps $a \mapsto q^i(y \mid a)$, $(\theta^i, a) \mapsto q^i_{\theta^i}(y \mid a)$, and $(a^i, y) \mapsto \pi^i(a^i, y)$ are continuous. (iii) *LR bound, density envelope, and integrability.* For each i there exist measurable $M^i : \mathbb{Y}^i \to [1, \infty)$ and $r^i : \mathbb{Y}^i \to [0, \infty)$ such that, for all $(\theta^i, a) \in \Theta^i \times \mathbb{A}$

[8] As usual, $L^p(\mathbb{Y}, \nu)$ denotes the space of all functions $f : \mathbb{Y} \to \mathbb{R}$ such that $\int |f(y)|^p \, \nu(dy) < \infty$. All Euclidean spaces, including $\mathbb{A}$ and Θ, are endowed with their Borel σ-algebra. The spaces $\Delta\mathbb{A}$ and $\Delta\Theta$ denote the sets of Borel probability measures on $\mathbb{A}$ and Θ, respectively, endowed with the topology of weak convergence.

and ν^i-a.e. y, one has $q^i_{\theta^i}(y \mid a)\, M^i(y)^{-1} \le q^i(y \mid a) \le q^i_{\theta^i}(y \mid a)\, M^i(y)$ and $q^i(y \mid a) \le r^i(y)$, and $\int_{\mathbb{Y}^i} M^i(y)\, r^i(y)\, \nu^i(dy) < \infty$. (iv) *Payoffs.* For ν^i-a.e. y, $(a^i, y) \mapsto \pi^i(a^i, y)$ is continuous; moreover, there exists $h^i_\pi : \mathbb{Y}^i \to [0, \infty)$ with $|\pi^i(a^i, y)| \le h^i_\pi(y)$ for all a^i and $\int_{\mathbb{Y}^i} h^i_\pi(y)\, M^i(y)\, r^i(y)\, \nu^i(dy) < \infty$.

These assumptions accommodate both discrete and continuous actions and consequences; ensure that best-fit parameters vary continuously with the action and that optimal actions vary continuously with beliefs; deliver a uniform law of large numbers; and guarantee that Bayesian updating never fails, since every realized consequence is admitted by the model.

3.2 Solution concept

Optimal action correspondence. Fix $i \in I$ and a (possibly correlated) distribution over others' actions $\beta^{-i} \in \Delta\mathbb{A}^{-i}$. For $\theta^i \in \Theta^i$, player i's expected utility is

$$U^i(a^i, \beta^{-i}, \theta^i) \;:=\; \int_{\mathbb{A}^{-i}} \int_{\mathbb{Y}^i} \pi^i(a^i, y^i)\, Q^i_{\theta^i}(dy^i \mid a^i, a^{-i})\, \beta^{-i}(da^{-i}). \tag{6}$$

Given a belief $\mu^i \in \Delta\Theta^i$, the optimal action correspondence is

$$F^i(\mu^i, \beta^{-i}) \;:=\; \arg\max_{a^i \in \mathbb{A}^i} \int_{\Theta^i} U^i(a^i, \beta^{-i}, \theta^i)\, \mu^i(d\theta^i).$$

Kullback–Leibler divergence. For $i \in I$, model $\theta^i \in \Theta^i$ and action profile $a \in \mathbb{A}$, define

$$K^i(\theta^i, a) \;:=\; \int_{\mathbb{Y}^i} \ln\!\Big(\frac{q^i(y^i \mid a)}{q^i_{\theta^i}(y^i \mid a)}\Big)\, q^i(y^i \mid a)\, \nu^i(dy^i), \quad \theta^i \in \Theta^i.$$

Let $\mathbb{Z}^i := \mathbb{S}^i \times \mathbb{A}^i$ and $\mathbb{Z} := \times_{j \in I} \mathbb{Z}^j$. For $\sigma \in \Delta\mathbb{A}$, define the set of best-fitting models for player i

$$\Theta^{m,i}(\sigma) \;:=\; \arg\min_{\theta^i \in \Theta^i} \int_{\mathbb{A}} K^i(\theta^i, a)\, \sigma(da).$$

Berk–Nash rationalizability. For every $A \subseteq \mathbb{A}$ (not necessarily a product set), we define

$$\Gamma(A) := \Big\{ a \in \mathbb{A} :\; \exists\, \sigma \in \Delta A \text{ s.t. } \forall i \in I,\; \exists\, \mu^i \in \Delta\Theta^{m,i}(\sigma) \text{ with } a^i \in F^i\big(\mu^i,\; \sigma^{-i}\big) \Big\}, \tag{7}$$

where $\sigma^{-i} \in \Delta A^{-i}$ denotes the marginal of σ over $\mathbb{A}^{-i}$.[9]

In other words, the set $\Gamma(A)$ consists of all action profiles that may be chosen when players assign probability one to the set of models that provide the best fit under some mixed action profile with support in A, and each player best responds to such beliefs.

The definition of Berk–Nash rationalizability is the same as in the single-agent case, except that Γ is now given by this generalized operator.

Definition 2. *An action profile a is* Berk–Nash rationalizable *if there exists $A \subseteq \mathbb{A}$ such that $a \in A \subseteq \Gamma(A)$.*

Moreover, an action profile a is a *Berk–Nash equilibrium* of the game if $a \in \Gamma(\{a\})$. In particular, equilibrium profiles are rationalizable, but the converse is not necessarily true.

Existence and characterization. The following result follows from standard arguments and the facts that Γ is nonempty-valued and maps closed sets into closed sets (see the Appendix).

Theorem 2 (Existence and characterization of rationalizable set)**.** *The set of Berk–Nash rationalizable signal-action profiles, $\mathcal{B} \subseteq \mathbb{A}$, is nonempty, compact, and is the largest fixed point of Γ. It can be obtained iteratively, as*

$$\mathcal{B} = \bigcap_{k=0}^{\infty} B^k \quad \textit{with} \quad B^0 = \mathbb{A} \quad \textit{and} \quad B^{k+1} = \Gamma(B^k).$$

Proof. See the Appendix. □

The theorem gives a simple recipe for computing the rationalizable set. Begin with all actions and repeatedly apply the operator Γ, removing anything that can't be justified given what remains. The actions that survive every round are exactly the BerkNash rationalizable actions, and this set is compact and nonempty.

[9] Formally, A^{-i} is the projection of A onto the opponents' coordinates: $A^{-i} := \{\, a^{-i} : \exists\, a^i \text{ with } (a^i, a^{-i}) \in A \,\}$. For any probability $\sigma \in \Delta A$, define σ^{-i} as the pushforward under that projection: for any $B \subseteq \times_{j \neq i} \mathbb{A}^j$, $\sigma^{-i}(B) = \sigma(\{\, a \in A : \; a^{-i} \in B \,\})$. For a product set $A = \times_{j \in I} A^j$, we have $A^{-i} = \times_{j \neq i} A^j$, and σ^{-i} is the usual marginal on the opponents' coordinates.

3.3 Example

There are three players, a manager (player 1) and two workers (players 2 and 3). Each player chooses effort $a^i \in [0, \infty)$ and receives outcome $y^i \in \mathbb{R}$. The technology is interdependent: the manager's effort affects only their own outcome and is productive if and only if the workers' efforts are sufficiently similar, while the workers' outcomes feature a multiplicative complementarity with the manager's effort.

Let $\alpha^* > 0$ be the true (baseline) team ability and $\theta^* > 0$ the true productivity index. Agents evaluate data through a parametric family that fixes a perceived team ability $\alpha > 0$ (fixed, not learned) and estimates $\theta \in [0, \bar{\theta}]$, with $\bar{\theta}$ large enough to be non-binding. We focus on overconfidence about team ability: $\alpha > \alpha^*$.

The true outcome equations are

$$y^1 = \big[\alpha^* + a^1 \, \mathbf{1}\{|a^2 - a^3| < k\}\big]\theta^* + \omega^1, \qquad y^i = \big[\alpha^* + a^i a^1\big]\theta^* + \omega^i, \quad i \in \{2, 3\},$$

where $\omega^i \sim \mathcal{N}(0, 1)$ and $\phi(x) = \mathbf{1}\{x < k\}$ for a threshold $k > 0$. Thus, the true model is $Q_1(\cdot \mid a) = \mathcal{N}\big([\alpha^* + a^1 \mathbf{1}\{|a^2 - a^3| < k\}]\theta^*, \, 1\big)$ for player 1, and $Q_i(\cdot \mid a) = \mathcal{N}\big([\alpha^* + a^i a^1]\theta^*, \, 1\big)$ for players $i \in \{2, 3\}$.

The perceived (misspecified) outcome equations are

$$y^1 = \big[\alpha + a^1 \, \mathbf{1}\{|a^2 - a^3| < k\}\big]\theta + \omega^1, \qquad y^i = \big[\alpha + a^i a^1\big]\theta + \omega^i, \quad i \in \{2, 3\}.$$

Thus, the perceived model (with fixed, not learned α and $\theta \in [0, \bar{\theta}]$) is $Q_{\theta,1}(\cdot \mid a) = \mathcal{N}\big([\alpha + a^1 \mathbf{1}\{|a^2 - a^3| < k\}]\theta, \, 1\big)$ for player 1, and $Q_{\theta,i}(\cdot \mid a) = \mathcal{N}\big([\alpha + a^i a^1]\theta, \, 1\big)$ for $i \in \{2, 3\}$.

Payoffs are $\pi_i(a, y^i) = y^i - c(a^i)$, where c is differentiable and strictly convex, and satisfies $c(0) = c'(0) = 0$, and $c'(a) \to \infty$ as $a \to \infty$. Because marginal returns are bounded, optimal efforts are also bounded.

For a degenerate action action profile δ_a with $a = (a^1, a^2, a^3)$, the KL-minimizing parameter value is

$$\begin{aligned} \theta^{m,1}(\delta_a) &= \theta^* \, \frac{\alpha^* + a^1 \mathbf{1}\{|a^2 - a^3| < k\}}{\alpha + a^1 \mathbf{1}\{|a^2 - a^3| < k\}}, \\ \theta^{m,i}(\delta_a) &= \theta^* \, \frac{\alpha^* + a^i a^1}{\alpha + a^i a^1}, \qquad i \in \{2, 3\}. \end{aligned}$$

For an action distribution σ on $[0, \infty)^3$, the KL-minimizer $\theta^{m,i}(\sigma)$ is a convex combination of the single-agent minimizers. As in the single-agent case, overconfidence ($\alpha > \alpha^*$) implies

underestimation: $\theta^{m,i}(\cdot) < \theta^*$ for all i.

Since the KL minimizer is unique, each player's belief degenerates at a single θ, so the best response is the unique action equating marginal cost to perceived marginal benefit. Hence, for any $A \subseteq [0,\infty)^3$,

$$\Gamma(A) = \Big\{ a \in [0,\infty)^3 : \ \exists\, \sigma \in \Delta A \text{ s.t. } \begin{array}{l} a^1 = (c')^{-1}\Big(\theta^{m,1}(\sigma)\ \mathbb{E}_\sigma\big[\mathbf{1}\{|a^2 - a^3| < k\}\big]\Big), \\ a^i = (c')^{-1}\Big(\theta^{m,i}(\sigma)\ \mathbb{E}_\sigma[a^1]\Big) \quad \text{for } i \in \{2,3\} \end{array} \Big\}.$$

For simplicity, we begin by providing an outer characterization of Γ without imposing common-σ consistency. Fix the workers' coordinates: $(a^2, a^3) \in [0,\infty)^2$. For any belief σ on $[0,\infty)^3$ define

$$\rho(\sigma) := \mathbb{E}_\sigma\big[\mathbf{1}\{|a^2 - a^3| < k\}\big] \in [0,1], \qquad a^1(\sigma) = (c')^{-1}\Big(\theta^{m,1}(\sigma)\,\rho(\sigma)\Big).$$

Starting from $a^1 \in [0,\infty)$, the manager's one-dimensional upper envelope evolves according to

$$M_1 = (c')^{-1}(\theta^*), \qquad M_{t+1} = (c')^{-1}\left(\theta^*\,\frac{\alpha^* + M_t}{\alpha + M_t}\right) \quad \text{for } t \geq 1,$$

which is monotone decreasing and converges to the unique fixed point M_∞ characterized by

$$c'\big(M_\infty\big) = \theta^*\,\frac{\alpha^* + M_\infty}{\alpha + M_\infty}.$$

Thus the manager's eventual image is contained in the interval $[0, M_\infty]$, and any point below M_∞ is obtained by a belief σ with $\rho(\sigma) < 1$, whereas M_∞ is reached by taking σ with $\rho(\sigma) = 1$ supported at $a^1 = M_\infty$.

Next, we fix player 1's coordinate to $[0, M_\infty]$ and iterate on players 2 and 3 separately. Let $m_1(\sigma) := \mathbb{E}_\sigma[a^1] \in [0, M_\infty]$. Player $i \in \{2,3\}$ best responds as

$$a^i(\sigma) = (c')^{-1}\big(\theta^{m,i}(\sigma)\, m_1(\sigma)\big).$$

With $a^1 \leq M_\infty$, a first pass gives the upper bound

$$N_1 = (c')^{-1}\big(\theta^* M_\infty\big).$$

Iterating with $a^1 \leq M_\infty$ and $a^i \leq N_t$ yields the one-dimensional map

$$N_{t+1} = (c')^{-1}\left(\theta^* M_\infty \frac{\alpha^* + M_\infty N_t}{\alpha + M_\infty N_t}\right),$$

which is monotone and converges to the unique fixed point N_∞ solving

$$c'(N_\infty) = \theta^* M_\infty \frac{\alpha^* + M_\infty N_\infty}{\alpha + M_\infty N_\infty}.$$

Hence the workers' eventual image is $[0, N_\infty]$; the lower endpoint is attained by beliefs with $m_1(\sigma) = 0$, and the upper endpoint by beliefs is supported only by $a^1 = M_\infty$ and $a^i = N_\infty$.

Suppose k is not too large relative to the (endogenous) difference between the actions of players 2 and 3, in particular suppose $k < N_\infty$. Then any belief σ supported on $[0, M_\infty] \times [0, N_\infty]^2$ must satisfy $\rho(\sigma) = 1$. In this case only M_∞ survives for player 1, and therefore only N_∞ survives for players 2 and 3. We end up with a unique Berk-Nash rationalizable profile $(M_\infty, N_\infty, N_\infty)$.

Of course, this conclusion does not use the full discipline of Γ. For example, a profile with $a^2 = 0$ and $a^3 = N_\infty$ cannot arise: to justify $a^2 = 0$, player 2 would have to believe that $a^1 = 0$ with probability one, and then player 3 must share the same belief, ruling out $a^3 = N_\infty$. In the appendix (Section 7.2), we show that the distance between players 2 and 3's actions is *less than or equal to*

$$k^* \;=\; N_\infty - c'^{-1}\Big(M_\infty\, \theta^*\, \frac{\alpha^*}{\alpha}\Big),$$

which implies that whenever $k > k^*$ the unique Berk–Nash rationalizable profile is $(M_\infty, N_\infty, N_\infty)$. In the special case where the players know the true ability ($\alpha = \alpha^*$), the bound simplifies to $k^* = 0$, so we have that $(M_\infty, N_\infty, N_\infty)$ is the unique Berk–Nash rationalizable profile for any $k > 0$.

4 Limit points are rationalizable

Each player i starts from a full-support prior $\mu_0^i \in \Delta\Theta^i$. At each period $t = 1, 2, \ldots$:

- Given posterior μ_t^i and a forecast (belief) $\tilde{\sigma}_t^{-i} \in \Delta\mathbb{A}^{-i}$ over opponents' actions, player

i chooses $a_t^i \in F^i(\mu_t^i,\ \tilde{\sigma}_t^{-i})$.[10]

- Consequences realize: for each i, $y_t^i \sim Q^i(\cdot \mid a_t, s_t^i)$.
- The action profile a_t is publicly observed; player i updates μ_{t+1}^i using Bayes' rule and the personal history $h_t^i := (a_{1:t},\ s_{1:t}^i,\ y_{1:t}^i)$.[11]

Forecasting opponents' actions. It remains to specify how player i forms a forecast $\tilde{\sigma}_t^{-i}$ about opponents' actions. A forecasting rule for player i is a sequence of Borel-measurable maps

$$a_{1:t-1} \mapsto \Psi_t^i(a_{1:t-1}) \in \Delta\mathbb{A}^{-i}$$

which produce at each period t a probability measure $\tilde{\sigma}_t^{-i} := \Psi_t^i(a_{1:t-1})$ over the opponents' actions, as a function of the publicly observed past actions $a_{1:t-1}$.

Given a realized sequence of opponents' actions $(a_\tau^{-i})_{\tau\geq 1}$, define the *empirical action distribution* up to time t by

$$\sigma_t^{-i}(B) \;=\; \frac{1}{t}\sum_{\tau=1}^{t} \mathbf{1}_B\big(a_\tau^{-i}\big), \qquad B \subseteq \mathbb{A}^{-i} \text{ Borel.}$$

Definition 3 (Pathwise subsequence consistency)**.** *A forecasting rule $(\Psi_t^i)_t$ is pathwise subsequence consistent if, for every realized action path $(a_t)_{t\geq 1}$ and every subsequence $(t_k)_k$ along which the empirical distribution $\sigma_{t_k} \Rightarrow \sigma \in \Delta\mathbb{A}$, the induced forecasts satisfy*

$$\tilde{\sigma}_{t_k}^{-i} \Rightarrow \sigma^{-i} \qquad \textit{for all } i \in I.$$

Pathwise subsequence consistency requires that, whenever observed frequencies stabilize along a subsequence, players' forecasts track those stabilized frequencies. We assume

[10] Formally, for each i and t, the action a_t^i is drawn from a Borel kernel $\phi_t^i(\cdot \mid \mu_t^i, \tilde{\sigma}_t^{-i})$ on $\mathbb{A}^i$ with $\phi_t^i\big(F^i(\mu_t^i, \tilde{\sigma}_t^{-i}) \mid \mu_t^i, \tilde{\sigma}_t^{-i}\big) = 1$. We impose no further restriction on ϕ_t^i beyond Borel measurability; it may vary with t. History affects actions only through $(\mu_t^i, \tilde{\sigma}_t^{-i})$. EP16 also consider forward-looking agents under a "weak identification" condition that drives experimentation incentives to zero in the long run; the same idea could be applied here.

[11] Formally, for any Borel set $S \subseteq \Theta$,

$$\mu_{t+1}^i(S) = \frac{\int_S q_\theta^i(y_t^i \mid a_t)\, \mu_t^i(d\theta)}{\int_{\Theta^i} q_\theta^i(y_t^i \mid a_t)\, \mu_t^i(d\theta)} \quad \text{for } Q^i(\cdot \mid a_t)\text{-a.e. } y_t^i.$$

Assumption (iii) implies $q_\theta^i(y_t^i \mid a_t) > 0$ $Q^i(\cdot \mid a_t)$-a.e. y_t^i for all $\theta \in \Theta^i$, so the denominator is positive and Bayes' rule is well defined.

throughout that each player's forecasting rule is pathwise subsequence consistent, and we refer to this environment as a game with Bayesian players and consistent forecasts.

Example (smoothed empirical forecasts). Fix $\alpha_0^i > 0$ and a full-support prior $\nu^i \in \Delta\mathbb{A}^{-i}$. Define

$$\tilde{\sigma}_t^{-i}(B) \;=\; \frac{\alpha_0^i}{\alpha_0^i + t - 1}\,\nu^i(B) \;+\; \frac{t-1}{\alpha_0^i + t - 1}\,\sigma_{t-1}^{-i}(B), \qquad B \subseteq \mathbb{A}^{-i}.$$

This retains full support for every finite t and puts vanishing weight on the prior as $t \to \infty$.

Proposition 1. *Smoothed empirical forecasts satisfy pathwise subsequence consistency.*

Proof. See the Appendix. □

This example matches the posterior predictive from Bayesian updating under a Dirichlet (or Dirichlet process) prior and specializes to empirical-frequency forecasting in fictitious play with a finite number of actions. Beyond Bayesian schemes, other forecasting rules, such as kernel- or shrinkage-based smoothers, also satisfy pathwise subsequence consistency, provided their deviation from the current empirical distribution vanishes over time. This "empirical-play" assumption also mirrors the standard learning-in-games approach used to justify Nash (and Berk–Nash) equilibrium: it avoids higher-order belief regress, such as player i forming conjectures about what models other players have, what those players think about the models of others, and so on. Instead, players simply treat the empirical distribution of past actions as their forecast.

Probability measures. Let $\mathbb{A} = \times_{i\in I}\mathbb{A}^i$ and $\mathbb{Y} = \times_{i\in I}\mathbb{Y}^i$. Let $\boldsymbol{P}_{(\mathbb{A}\times\mathbb{Y})^{\mathbb{N}}}$ denote the law on the space of infinite action-consequence sequences $(\mathbb{A}\times\mathbb{Y})^{\mathbb{N}}$ induced by the data-generating process and the players' belief, forecasting, and action rules. Write $\boldsymbol{P}_{\mathbb{A}^{\mathbb{N}}}$ for the marginal on $\mathbb{A}^{\mathbb{N}}$, and for every $a^\infty \in \mathbb{A}^{\mathbb{N}} := \mathbb{A}\times\mathbb{A}\times\ldots$, write $\boldsymbol{P}_{\mathbb{Y}^{\mathbb{N}}|\mathbb{A}^{\mathbb{N}}}(\cdot \mid a^\infty)$ for the conditional of $\mathbb{Y}^{\mathbb{N}}$ given a^∞.[12]

Asymptotic characterization of beliefs.

Lemma 1 (Asymptotic beliefs)**.** *Fix any infinite sequence of actions $a^\infty \in \mathbb{A}^{\mathbb{N}}$. Almost surely with respect to $\boldsymbol{P}_{\mathbb{Y}^{\mathbb{N}}}(\cdot \mid a^\infty)$, the following holds. Suppose there exists a subsequence $(t_k)_k$ such that the empirical measures σ_{t_k} converge weakly to σ. Then, for every player i*

[12]By disintegration, there exists a regular conditional probability (kernel) $\boldsymbol{P}_{\mathbb{Y}^{\mathbb{N}}|\mathbb{A}^{\mathbb{N}}}(\cdot \mid a^\infty)$ such that for all measurable $B \subseteq \mathbb{A}^{\mathbb{N}}$ and $C \subseteq \mathbb{Y}^{\mathbb{N}}$, $\boldsymbol{P}_{(\mathbb{A}\times\mathbb{Y})^{\mathbb{N}}}(B\times C) \;=\; \int_{a^\infty\in B} \boldsymbol{P}_{\mathbb{Y}^{\mathbb{N}}}(C \mid a^\infty)\,\boldsymbol{P}_{\mathbb{A}^{\mathbb{N}}}(da^\infty)$.

and every closed set $E \subseteq \Theta^i$ *with* $E \cap \Theta^{m,i}(\sigma) = \varnothing$*, there exist constants* $C > 0$*,* $\rho > 0$*, and an integer* K *such that, for all* $k \geq K$*,*

$$\mu^i_{t_k}(E) \;\leq\; C \exp\{-\rho\, t_k\}. \tag{8}$$

In particular, if the subsequence of posteriors $(\mu^i_{t_k})_k$ *converges to some* μ^i*, then* $\mu^i \in \Delta\Theta^{m,i}(\sigma)$*.*

Lemma 1 says that along any subsequence where the empirical action distribution converges, the posterior probability assigned to any set of models incompatible with the limiting action distribution converges to zero.[13] Consequently, any limit belief must be supported on the set of models that best explain the observed limiting behavior. The idea originates in Berk (1966)'s analysis of misspecified models under i.i.d. data and has since been extended to dynamic learning settings, including by EP16. Lemma 1 generalizes these results by allowing for a continuum of actions and by working with subsequences rather than requiring convergence of the empirical action distribution; the argument closely follows existing proofs and appears in the Supplemental Appendix.

Asymptotic characterization of signal-action profiles. An action profile $a \in \mathbb{A}$ is a *limit action profile* of $(a_t)_t \in \mathbb{A}^{\mathbb{N}}$ if there exists a subsequence $(a_{t_k})_k$ with $a_{t_k} \to a$. Equivalently, for every open neighborhood $U \subseteq \mathbb{A}$ of a, $a_t \in U$ for infinitely many t. When $\mathbb{A}$ is finite, a is a limit action profile if and only if it occurs infinitely often along the path.

Theorem 3. *Consider a game with Bayesian players and consistent forecasts. Almost surely with respect to* $\mathbf{P}_{\mathbb{A}^{\mathbb{N}}}$*, every limit signal-action profile is Berk–Nash rationalizable.*

Proof. By disintegration, there exists a full $\boldsymbol{P}_{\mathbb{A}^{\mathbb{N}}}$-measure set $A^\star \subseteq \mathbb{A}^{\mathbb{N}}$ such that, for each $a^\infty \in A^\star$, Lemma 1 holds for $\boldsymbol{P}_{\mathbb{Y}^{\mathbb{N}}|\mathbb{A}^{\mathbb{N}}}(\cdot \mid a^\infty)$-a.e.. Fix such a^∞ and choose y^∞ from its conditional probability-one set.

Let $\mathcal{Z}(a^\infty)$ denote the set of limit action profiles of a^∞. Pick any $a \in \mathcal{Z}(a^\infty)$ and a subsequence (t_k) with $a_{t_k} \to a$. By compactness, pass to a further subsequence (not relabeled) such that

$$\sigma_{t_k} \Rightarrow \sigma \in \Delta\mathbb{A} \quad \text{and} \quad \mu^i_{t_k} \Rightarrow \mu^i \in \Delta\Theta^i \quad \text{for each } i.$$

[13] The convergence is exponentially fast, but the speed of convergence is not necessary for the sequel.

A standard support argument yields $\sigma \in \Delta\mathcal{Z}(a^\infty)$. By pathwise subsequence consistency of forecasts, $\tilde{\sigma}^{-i}_{t_k} \Rightarrow \sigma^{-i}$ for each i. By Lemma 1 (applied to a^∞ and y^∞), $\mu^i \in \Delta\Theta^{m,i}(\sigma)$ for all i.

Since $a^i_{t_k} \in F^i(\mu^i_{t_k}, \tilde{\sigma}^{-i}_{t_k})$ for all k and F^i has a closed graph (upper hemicontinuous with closed values), we conclude

$$a^i \in F^i(\mu^i, \sigma^{-i}) \qquad \text{for each } i,$$

so $a \in \Gamma(\mathcal{Z}(a^\infty))$. As $a \in \mathcal{Z}(a^\infty)$ was arbitrary, every limit action profile is Berk–Nash rationalizable for all $a^\infty \in A^\star$, i.e., $\boldsymbol{P}_{\mathbb{A}^{\mathbb{N}}}$-a.s. □

Theorem 3 says that, with probability one, every limit (accumulation) point of play lies in the Berk-Nash rationalizable set. In contrapositive form: if an action profile is not rationalizable, it cannot arise as a limit point of play except on a probabilityzero set.

5 Relationship to rationalizability

To relate Berk–Nash equilibrium to classical notions of rationalizability (which implicitly assume players know the game and understand how action profiles map to consequences), we consider the special case where this is also true in our environment.

Definition 4 (Correct specification and identification)**.**

- Correct specification. *For each player i, there exists $\theta^i \in \Theta^i$ such that $Q^i_{\theta^i}(\cdot \mid a) = Q^i(\cdot \mid a)$ for all $a \in \mathbb{A}$.*
- Identification. *For each i and each $\sigma \in \Delta\mathbb{A}$, the set $\Theta^{m,i}(\sigma)$ is a singleton (the minimizer may depend on i and on σ).*

Proposition 2. *Consider a game that is correctly specified and identified. Then for every $A \subseteq \mathbb{A}$,*

$$\Gamma(A) = \{a \in \mathbb{A} : \exists \sigma \in \Delta A \text{ s.t. } \forall i,\ a^i \in BR^i(\sigma^{-i})\}, \tag{9}$$

where $BR^i(\sigma^{-i}) := \arg\max_{a^i \in \mathbb{A}^i} \int_{\mathbb{A}^{-i}} \int_{\mathbb{Y}^i} \pi^i(a^i, y^i) Q^i(dy^i \mid a^i, a^{-i}) \sigma^{-i}(da^{-i})$.

Proof. Correct specification says that for each i there exists θ^{i,θ^*} with $Q^i_{i,\theta^*} = Q^i$; in particular, $\theta^{i,\theta^*} \in \Theta^{m,i}(\sigma)$ for all σ. Identification says the minimizer is unique, so $\Theta^{m,i}(\sigma) = \{\theta^{i,\theta^*}\}$

for all σ, so the belief is degenerate at θ^{i,θ^*} and therefore $F^i(\delta_{\theta^{i,\theta^*}}, \sigma^{-i}) = BR^i(\sigma^{-i})$. The result then follows from the definition of Γ in (7). □

Bernheim–Pearce operator. By contrast, the *Bernheim–Pearce (correlated) operator* Γ_{BP} is defined for all product sets $A = \times_{i\in I} A^i$ as

$$\Gamma_{\mathrm{BP}}(A) \;:=\; \big\{\, a \in A :\; \forall i,\; \exists\, \sigma^{-i} \in \Delta A^{-i} \text{ s.t. } a^i \in BR^i(\sigma^{-i}) \,\big\}.$$

For both Γ and Γ_{BP}, the rationalizable set is the largest fixed point of the corresponding operator.[14] Our operator Γ differs in that it requires a *single* $\sigma \in \Delta\mathbb{A}$ whose marginals σ^{-i} simultaneously justify all players' best replies. This reflects learning from a shared history, which disciplines conjectures to be mutually consistent.

Theorem 4. *In correctly specified and identified games, the Berk–Nash rationalizable set is contained in the (correlated) rationalizable set. In two-player games, the two sets coincide.*

Proof. First, for any *product* set A we have $\Gamma(A) \subseteq \Gamma_{\mathrm{BP}}(A)$ (because a common $\sigma \in \Delta A$ yields marginals σ^{-i} that justify each a^i). Since Γ_{BP} maps product sets to product sets, $\Gamma^k_{\mathrm{BP}}(\mathbb{A})$ is a product set for every k. Proceeding by induction on k (so that $\Gamma^k(\mathbb{A}) \subseteq \Gamma^k_{\mathrm{BP}}(\mathbb{A})$), and using monotonicity of Γ together with the inclusion on product sets, for each $k \geq 0$:

$$\Gamma^{k+1}(\mathbb{A}) = \Gamma\big(\Gamma^k(\mathbb{A})\big) \;\subseteq\; \Gamma\big(\Gamma^k_{\mathrm{BP}}(\mathbb{A})\big) \;\subseteq\; \Gamma_{\mathrm{BP}}\big(\Gamma^k_{\mathrm{BP}}(\mathbb{A})\big) = \Gamma^{k+1}_{\mathrm{BP}}(\mathbb{A}),$$

where the first inclusion uses Γ's monotonicity and the inductive hypothesis, and the second uses that $\Gamma^k_{\mathrm{BP}}(\mathbb{A})$ is a product set and $\Gamma(\cdot) \subseteq \Gamma_{\mathrm{BP}}(\cdot)$ on product sets. Taking intersections over k gives

$$\bigcap_{k\geq 0} \Gamma^k(\mathbb{A}) \;\subseteq\; \bigcap_{k\geq 0} \Gamma^k_{\mathrm{BP}}(\mathbb{A}),$$

so the Berk–Nash rationalizable set is contained in the Bernheim–Pearce (correlated) rationalizable set.

Equality for two players. For $|I| = 2$ and any product set $A = A^1 \times A^2$, take $a \in \Gamma_{\mathrm{BP}}(A)$ with conjectures $\sigma^{-1} \in \Delta(A^2)$ and $\sigma^{-2} \in \Delta(A^1)$. The product measure $\bar\sigma := \sigma^{-2} \otimes \sigma^{-1} \in \Delta A$ has marginals $\bar\sigma^{-i} = \sigma^{-i}$, so $a^i \in BR^i(\bar\sigma^{-i})$ for $i = 1, 2$, hence $a \in \Gamma(A)$. Therefore $\Gamma_{\mathrm{BP}}(A) \subseteq \Gamma(A)$, and the two fixed-point sets coincide. □

[14] For Γ_{BP}, this is correlated rationalizability (Brandenburger and Dekel (1987)).

Our operator Γ enforces a single, mutually consistent forecast over joint play; Γ_{BP} allows player-by-player conjectures that need not come from a common joint distribution. With two players, any pair of marginal conjectures can be combined into a joint product distribution, so the restriction is without bite and the sets coincide. With three or more players, players' separate conjectures may be mutually inconsistent, so requiring a common joint forecast can only shrink the set, yielding containment.

Example (Berk–Nash vs. BP rationalizability). This example shows that Berk–Nash rationalizability can be a *strict* subset of BP rationalizability.[15]

Consider the 3-player game of Section 3.3. In the correctly specified case ($\alpha = \alpha^*$), for any $k > 0$ we argued there is a unique Berk–Nash rationalizable outcome. Because the model is correctly specified and identified, one can reach this conclusion far more simply by iterating Γ in (9). Indeed, players 2 and 3 face the same beliefs and therefore best-respond symmetrically, so $a_2 = a_3$. With this restriction, player 1's best response is uniquely pinned down at

$$M^* = (c')^{-1}(\theta^*),$$

and then players 2 and 3 best respond at

$$N^* = (c')^{-1}(\theta^* M^*).$$

Hence (M^*, N^*, N^*) is the unique Berk-Nash rationalizable profile (and therefore the unique Nash equilibrium).

By contrast, if $k < N^*$, the set of BP-rationalizable profiles is the whole box

$$[0, M^*] \times [0, N^*] \times [0, N^*].$$

The reason is that BP allows each player to justify a best response with potentially *different* beliefs. For instance, player 1 can choose $a^1 = 0$ rationalized by the belief that $(a^2, a^3) = (0, N^*)$; player 2 can choose $a^2 = 0$ rationalized by the belief that $a^1 = 0$; player 3 can choose $a^3 = N^*$ rationalized by the belief that $a^1 = M^*$; and player 1 can choose $a^1 = M^*$ rationalized by the belief that $(a^2, a^3) = (N^*, N^*)$. ♦

[15] In the appendix we present a game with a finite number of actions in which every action profile is BP-rationalizable but only a single one is Berk–Nash rationalizable.

6 Extensions

6.1 Rationalizable distributions

We defined rationalizability over actions. There is an equivalent definition in terms of distributions over actions, and this alternative definition captures the limit points of empirical-action distributions.

Distribution-level operator. For a set of action-profile distributions $\Sigma \subseteq \Delta\mathbb{A}$, define

$$\Phi(\Sigma) := \left\{\sigma \in \Delta\mathbb{A} : \forall a \in \operatorname{supp} \sigma,\ \exists\, \sigma_a \in \Sigma \text{ s.t. } \forall i,\ \exists\, \mu_a^i \in \Delta\Theta^{m,i}(\sigma_a) \text{ with } a^i \in F^i(\mu_a, \sigma_a^{-i})\right\}.$$

In other words, $\Phi(\Sigma)$ comprises those action-profile distributions whose every support action is justified by best-fitting beliefs formed relative to some reference distribution in Σ.

Analogously to the action-based notion, we define Berk–Nash rationalizability for distributions as follows:

Definition 5. *A distribution $\sigma \in \Delta\mathbb{A}$ is Berk–Nash rationalizable if there exists a set of action-profile distributions $\Sigma \subseteq \Delta\mathbb{A}$ such that $\sigma \in \Sigma \subseteq \Phi(\Sigma)$.*

Moreover, an action-profile distribution σ is a *generalized Berk–Nash equilibrium* of the game if $\sigma \in \Gamma(\{\sigma\})$. A generalized Berk–Nash equilibrium extends the baseline notion in two ways (see also Esponda, Pouzo and Yamamoto (2021) and Murooka and Yamamoto (2023)): (i) it allows cross-player correlation in play (distinct from Aumann's (1974) correlated equilibrium because only the marginals, not players' conditional strategies, matter), and (ii) it permits different action profiles to be supported by different beliefs. EP16 define a stricter version of Berk–Nash equilibrium that rules out both features, justified via Harsanyi-style independent (see Section 6.2). As usual, equilibrium profiles are rationalizable, but the converse is not necessarily true.

Versions of Theorems 2 and 3 hold for this case. In particular, the set of Berk–Nash rationalizable distributions is the largest fixed point of Φ and can be obtained by iterating Φ starting from $\Sigma_0 = \Delta\mathbb{A}$. Moreover, every limit empirical distribution of action profiles is Berk–Nash rationalizable. The proofs are entirely analogous to those for Γ.

Let $\mathcal{B} \subseteq \mathbb{A}$ and $\mathcal{M} \subseteq \Delta\mathbb{A}$ denote the set of Berk–Nash rationalizable actions and distributions, respectively. The following result shows the equivalence of the two approaches.

Proposition 3 (Equivalence of action and distributional approaches)**.**

$$\mathcal{M} \;=\; \Delta\mathcal{B}.$$

Proof. Step 1: $\Phi(\Delta\mathcal{B}) \subseteq \Delta\mathcal{B}$: Take $\sigma \in \Phi(\Delta\mathcal{B})$. By definition of Φ, for every $a \in \operatorname{supp}\sigma$ there exist $\sigma_a \in \Delta\mathcal{B}$ and, for each player i, a belief $\mu_a^i \in \Delta\Theta^{m,i}(\sigma_a)$ such that $a^i \in F^i(\mu_a^i, \sigma_a^{-i})$. By the definition of Γ, this implies $a \in \Gamma(\mathcal{B}) = \mathcal{B}$. Hence $\operatorname{supp}\sigma \subseteq \mathcal{B}$, i.e., $\sigma \in \Delta\mathcal{B}$.

Step 2: $\Delta\mathcal{B} \subseteq \Phi(\Delta\mathcal{B})$: Take $\sigma \in \Delta\mathcal{B}$ and let $a \in \operatorname{supp}\sigma$. Since $a \in \mathcal{B} = \Gamma(\mathcal{B})$, there exist $\sigma_a \in \Delta\mathcal{B}$ and, for each i, $\mu_a^i \in \Delta\Theta^{m,i}(\sigma_a)$ with $a^i \in F^i(\mu_a^i, \sigma_a^{-i})$. Thus the existential requirement in the definition of Φ is met for each support action a, implying $\sigma \in \Phi(\Delta\mathcal{B})$.

Combining both inclusions yields $\Phi(\Delta\mathcal{B}) = \Delta\mathcal{B}$. Since $\mathcal{M}$ is the largest Φ-fixed set, necessarily $\mathcal{M} = \Delta\mathcal{B}$. □

6.2 Mixed-strategy rationalizability

It is well known that fictitious play can converge to correlated distributions, and the same possibility arises in our setting. To rule out correlation across players, we follow Fudenberg and Kreps (1993) and introduce Harsanyi–style payoff perturbations independent across players and time. However, even with these perturbations, the common-σ restriction that defines Berk–Nash rationalizability remains. This underscores that common-belief restrictions stemming from a shared history and correlation in strategies induced by a shared history are distinct phenomena.

Payoff perturbations Given a (possibly correlated) distribution over opponents' actions $\beta^{-i} \in \Delta\mathbb{A}^{-i}$ and a belief $\mu^i \in \Delta\Theta^i$, the *baseline* expected utility of player for action a^i is $U^i(a^i, \beta^{-i}, \mu^i)$ defined in Section 3, equation (6). For the payoff shocks, let ε^i be an $\mathbb{A}^i \to \mathbb{R}$ random function with law P_{ε^i} supported on the set $C(A^i)$ of real-valued continuous functions, independent across players. Under suitable assumptions, the realized action satisfies

$$a^i \in \arg\max_{x \in \mathbb{A}^i} \left\{ U^i(x, \beta^{-i}, \mu^i) + \varepsilon^i(x) \right\} \quad \text{a.s.},$$

and the induced *intended mixed strategy* (choice kernel) on Borel sets $B \subseteq \mathbb{A}^i$ is well defined as follows:

$$\kappa^i(B \mid \mu^i, \beta^{-i}) := P_{\varepsilon^i}\left(\arg\max_{x \in \mathbb{A}^i} \{U^i(x, \beta^{-i}, \mu^i) + \varepsilon^i(x)\} \in B \right), \tag{10}$$

so $\kappa^i(\cdot \mid \mu^i, \beta^{-i}) \in \Delta\mathbb{A}^i$.

In addition, by joint continuity of U^i, the choice kernel is continuous: If $(\mu^i_n, \beta^{-i}_n) \to (\mu^i, \beta^{-i})$ in $\Delta\Theta^i \times \Delta\mathbb{A}^{-i}$, then

$$\kappa^i(\cdot \mid \mu^i_n, \beta^{-i}_n) \;\Rightarrow\; \kappa^i(\cdot \mid \mu^i, \beta^{-i}) \quad \text{in } \Delta\mathbb{A}^i.$$

For instance, in the special case of a finite number of actions, this follows from assuming P_{ε^i} has a continuous density on $\mathbb{R}^{A^i}$, and the choice kernel becomes, for each $a^i \in \mathbb{A}^i$

$$\kappa^i(\{a^i\} \mid \mu^i, \beta^{-i}) = P_{\varepsilon^i}\big(U^i(a^i, \beta^{-i}, \mu^i) + \varepsilon^i_{a^i} \geq U^i(b^i, \beta^{-i}, \mu^i) + \varepsilon^i_{b^i} \text{ for all } b^i \in \mathbb{A}^i\big)\,.$$

With Type-I extreme value shocks, this is the familiar multinomial logit.

Berk–Nash rationalizability for mixed strategies. We work on $\times_{i\in I}\Delta\mathbb{A}^i$. We refer to $\sigma^i \in \Delta\mathbb{A}^i$ as player i's *mixed strategy* and to $\sigma = (\sigma^i)_{i\in I}$ as the *mixed strategy profile*. We *abuse notation* by also writing σ (resp. σ^{-i}) for the associated product measure $\otimes_{i\in I}\sigma^i \in \Delta\mathbb{A}$ (resp. $\otimes_{j\neq i}\sigma^j \in \Delta\mathbb{A}^{-i}$) when the meaning is clear from context.

For any $\Sigma \subseteq \times_{i\in I}\Delta\mathbb{A}^i$, define

$$\Phi_M(\Sigma) := \Big\{\, \sigma \in \times_{i\in I}\Delta\mathbb{A}^i \;:\; \exists\, \hat\sigma \in \Sigma \text{ s.t. } \forall i \in I\; \exists\, \mu^i \in \Delta\Theta^{m,i}(\hat\sigma) \text{ with } \sigma^i = \kappa^i\big(\cdot \;\big|\; \mu^i,\, \hat\sigma^{-i}\big) \,\Big\}.$$

Thus, a mixed strategy belongs to $\Phi^*(\Sigma)$ exactly when there exists a *common* mixed strategy profile $\hat\sigma$ in Σ such that, for every player i, the i-th marginal σ^i is the random-utility best response $\kappa^i(\cdot \mid \mu^i, \hat\sigma^{-i})$ for some best-fitting belief $\mu^i \in \Delta\Theta^{m,i}(\hat\sigma)$.

Definition 6. *A mixed strategy profile $\sigma = (\sigma^i)_{i\in I}$ is* Berk–Nash rationalizable *if there exists $\Sigma \subseteq \times_{i\in I}\Delta\mathbb{A}^i$ such that $\sigma \in \Sigma \subseteq \Phi_M(\Sigma)$.*

Moreover, a mixed strategy profile $\sigma = (\sigma^i)_{i\in I}$ is a *Berk–Nash equilibrium* (see EP16) of the game if $\sigma \in \Phi_M(\{\sigma\})$. In particular, mixed-strategy equilibrium profiles are rationalizable, but the converse is not necessarily true.

The existence and characterization argument in Theorem 2 applies naturally to this setting: the set of Berk–Nash rationalizabile mixed-strategy profiles is nonempty and compact, and it can be obtained by repeated iteration of Φ_M starting from the largest set $\times_{i\in I}\Delta\mathbb{A}^i$.

The argument of Theorem 3 also applies naturally here to the intended strategy profile

$\kappa = (\kappa^i)_{i\in I}$ defined in (10). Define an intended mixed strategy profile $a \in \mathbb{A}$ as a *limit intended mixed strategy profile* of $(\kappa_t)_t$ if there exists a subsequence $(\kappa_{t_k})_k$ with $\kappa_{t_k} \to a$. Equivalently, for every open neighborhood $U \subseteq \times_{i\in I}\Delta\mathbb{A}^i$ of κ, $\kappa_t \in U$ for infinitely many t.

Theorem 5. *Consider a game with independent payoff perturbations. Almost surely, every limit intended mixed-strategy profile is a Berk–Nash rationalizable mixed-strategy profile.*

Proof. See the Appendix. □

Finally, we consider the special case of a correctly specified and identified game. By the usual argument, these assumptions imply that we can replace the belief μ^i in the definition of Φ_M with a degenerate belief at the truth, so that players best respond to correct beliefs.

Proposition 4. *Consider a game with independent payoff perturbations that is correctly specified and identified. Then for every $\Sigma \subseteq \times_{i\in I}\Delta\mathbb{A}^i$,*

$$\Phi_M(\Sigma) := \Big\{ \sigma \in \times_{i\in I}\Delta\mathbb{A}^i \ : \ \exists\, \hat\sigma \in \Sigma \text{ such that } \forall i, \sigma^i = br^i\big(\,\cdot \;\big|\; \hat\sigma^{-i}\big) \Big\}, \tag{11}$$

where br^i is player i's correct best response probabilistic function.[16]

We can contrast our operator to Bernheim and Pearce's original definition of (independent) rationalizability applied to mixed strategies. The BP-operator Φ_{BP} is, for any product set $\Sigma = \times_{i\in I}\Sigma^i \subseteq \times_{i\in I}\Delta\mathbb{A}^i$,

$$\Phi_{BP}(\times_{i\in I}\Sigma^i) := \Big\{ \sigma \in \times_{i\in I}\Delta\mathbb{A}^i \ : \ \forall i, \exists\, \hat\sigma^{-i} \in \times_{j\neq i}\Sigma^j \text{ s.t. } \sigma^i = br^i\big(\,\cdot \;\big|\; \hat\sigma^{-i}\big) \Big\}.$$

Independent payoff perturbations eliminate correlated strategies (and hence correlated conjectures about others' play), pushing outcomes to product form. Nevertheless, the common history still ties players' conjectures to a single reference profile: BP permits player-specific opponent references drawn independently from $\times_{j\neq i}\Sigma^j$, whereas Φ_M requires one $\hat\sigma \in \Sigma$ to underlie all conjectures simultaneously. This couples the coordinates and generally yields a smaller, non-product image (the restriction is vacuous only with two players).

Theorem 6. *In correctly specified and identified games with independent payoff perturbations, the Berk–Nash rationalizable set is contained in the (independent) rationalizable set. In two-player games, the two sets coincide.*

[16] Formally, for correct beliefs define $U^i(a^i, \sigma^{-i}) := \int_{\mathbb{A}^{-i}}\int_{\mathbb{Y}^i} \pi^i(a^i, y^i)\, Q^i(dy^i \mid a^i, a^{-i})\, \sigma^{-i}(da^{-i})$. Given payoff perturbations ε^i, the stochastic best response is the probability kernel $\mathrm{br}^i(B \mid \sigma^{-i}) := P_{\varepsilon^i}\big(\arg\max_{x\in\mathbb{A}^i}\{U^i(x, \sigma^{-i}) + \varepsilon^i(x)\} \in B\big)$ for all Borel $B \subseteq \mathbb{A}^i$.

6.3 Player-specific histories

We relax the common-history assumption by allowing each player to maintain a personal, possibly selective, *subsample* of periods. Formally, for each player i there is an increasing sequence of time indices $(\tau_n^i)_{n\geq 1} \subset \mathbb{N}$ with $\tau_n^i \to \infty$. Whenever a period $\tau = \tau_n^i$ lies in i's subsample, she observes the entire action profile a_τ from that period and her own realized consequence y_τ^i. Let $N_t^i := \max\{n : \tau_n^i \leq t\}$ denote the number of retained periods up to t, and assume $N_t^i \to \infty$.

Player i's *personal empirical distribution* over action profiles is

$$\hat{\sigma}_t^i \in \Delta\mathbb{A}, \qquad \hat{\sigma}_t^i(B) = \frac{1}{N_t^i} \sum_{n=1}^{N_t^i} \mathbf{1}_B(a_{\tau_n^i}) \quad \text{for Borel } B \subseteq \mathbb{A}.$$

Forecasts are assumed to be pathwise subsequence consistent *with respect to the player's own empirical distribution*: along any realized path and any subsequence (t_k) for which $\hat{\sigma}_{t_k}^i \Rightarrow \hat{\sigma}^i$, we have $\tilde{\sigma}_{t_k}^{-i} \Rightarrow (\hat{\sigma}^i)^{-i}$.

With player-specific histories, the rationalizability operator permits each player to anchor beliefs on (potentially) different empirical limits:

$$\Gamma_W(A) := \Big\{ a \in A : \ \forall i \in I, \ \exists\, \sigma^i \in \Delta A \text{ and } \mu^i \in \Delta\Theta^{m,i}(\sigma^i) \text{ such that } a^i \in F^i\big(\mu^i, \ (\sigma^i)^{-i}\big) \Big\}.$$

Definition 7 (Berk–Nash weak rationalizability)**.** *An action profile a is* weakly Berk–Nash rationalizable *if there exists $A \subseteq \mathbb{A}$ such that $a \in A \subseteq \Gamma_W(A)$.*

Theorem 7. *In a game with Bayesian players and consistent forecasts relative to their own empirical subsamples, almost surely every limit action profile is weakly Berk–Nash rationalizable.*

Proof. The proof is essentially identical to the proof of Theorem 3, so we only provide a sketch. Fix a realized path and a limit action profile a from the set of limit action profiles $\mathcal{Z}$. For each player i, pass to a subsequence (t_k) along which $\hat{\sigma}_{t_k}^i \Rightarrow \sigma^i$ and $\mu_{t_k}^i \Rightarrow \mu^i$. By pathwise subsequence consistency (relative to $\hat{\sigma}^i$), $\tilde{\sigma}_{t_k}^{-i} \Rightarrow (\sigma^i)^{-i}$. The usual Berk argument applied player by player yields $\mu^i \in \Delta\Theta^{m,i}(\sigma^i)$. Closed-graph properties of F^i then give $a^i \in F^i(\mu^i, (\sigma^i)^{-i})$ for each i, so $a \in \Gamma_W(\mathcal{Z})$, completing the adaptation. □

Correct specification and identification. Under correct specification and identification, beliefs collapse to that truth and $\Gamma_W(A)$ becomes

$$\Gamma_W(A) \;=\; \Big\{ a \in A : \; \forall i \in I, \; \exists \sigma^i \in \Delta A \text{ s.t. } a^i \in BR^i\big((\sigma^i)^{-i}\big) \Big\}.$$

This operator coincides with the Bernheim–Pearce (correlated) operator. In Γ_{BP}, one rationalizes a by, for each i, choosing any conjecture $\sigma^{-i} \in \Delta A^{-i}$ with $a^i \in BR^i(\sigma^{-i})$. Given such a marginal, one can extend it to some $\sigma^i \in \Delta A$ whose $(-i)$-marginal is σ^{-i}; hence $\Gamma_W(A) = \Gamma_{\text{BP}}(A)$.

Theorem 8. *In correctly specified and identified games, the Berk–Nash weakly rationalizable set coincides with the (correlated) rationalizable set.*

In the special case of correctly specified and identified games, players only need to forecast opponents' actions. Milgrom and Roberts (1991) define *adaptive forecasts*—beliefs that assign vanishing probability to action profiles that do not persist—and show that, under such forecasts, play converges to the set of (correlated) BP-rationalizable profiles. Our forecasting rule (placing vanishing probability on profiles not observed along the player's subsequence) is a special case of adaptive forecasts. Hence, in correctly specified and identified settings, their result implies convergence to the BP-rationalizable set. Equivalently, this conclusion follows here by combining Theorems 7 and 8.

6.4 Other solution concepts

Aumann's (1974) correlated equilibrium has two defining features: (i) it permits correlated play; and (ii) it imposes conditional best responses—whenever an action is used with positive probability, it must be optimal given the conditional distribution of opponents actions in those instances. Its learning foundations come from calibrated best responses (Foster and Vohra 1997), conditional universal consistency (Fudenberg and Levine 1999), and regret matching (Hart and Mas-Colell 2000), all of which implement this conditional best-response logic. By contrast, in our solution concept only the first feature is present: correlation may arise from common history, but players best respond to the marginal distribution of others' play rather than to conditional distributions. A natural extension is to allow conditional forecasts and best responses within our framework.

Following Basu and Weibull (1991), a set of action profiles A is *closed under rational behavior (curb)* if it is closed under Γ, i.e., $\Gamma(A) \subseteq A$, where Γ is the best-response operator

(and, more generally, may be taken to be the Berk–Nash operator used in our paper). A *minimal curb set* is a curb set that contains no proper curb subset; equivalently, these are the minimal fixed points of Γ. In contrast, (Berk-Nash) rationalizability calls a set A *self-justified* if $A \subseteq \Gamma(A)$. The union of all self-justified sets is the set of rationalizable action profiles, and it is characterized as the *largest* fixed point of Γ. As noted by Basu and Weibull (1991), minimal curb sets (minimal fixed points) and the rationalizable set (largest fixed point) are opposite ends of a spectrum. Our limiting results exclude actions outside the rationalizable set but do not rule out actions that are rationalizable yet not contained in any minimal curb set. An open question is to identify conditions—on the dynamics and on game primitives—under which our conclusions can be strengthened to imply convergence to minimal curb sets or to some other set that may be strictly contained in the rationalizable set.

7 Appendix

Lemma 2. *1. For each $i \in I$, the best-response correspondence $F^i(\mu^i, \sigma^{-i}) \subseteq \mathbb{A}^i$ is nonempty, compact-valued, and upper hemicontinuous in (μ^i, σ^{-i}).*

2. For each $i \in I$, the set of KL minimizers $\Theta^{m,i}(\sigma) \subseteq \Theta^i$ is nonempty, compact-valued, and upper hemicontinuous in σ.

Proof. (i) Fix i and define

$$\bar{U}^i(a^i, \theta^i, a^{-i}) := \int_{\mathbb{Y}^i} \pi^i(a^i, y^i)\, q^i_{\theta^i}(y^i \mid a^i, a^{-i})\, \nu^i(dy^i).$$

By Assumption 3.1(ii) the integrand is ν^i-a.e. continuous in (a^i, θ^i, a^{-i}). By Assumption 3.1(iv) and the LR bound in Assumption 3.1(iii), $|\pi^i(a^i, y^i)| \le h^i_\pi(y^i)$ and $q^i_{\theta^i} \le M^i\, q^i \le M^i r^i$, hence $|\pi^i q^i_{\theta^i}| \le h^i_\pi M^i r^i$ with $\int h^i_\pi M^i r^i\, d\nu^i < \infty$. Thus $\bar{U}^i$ is jointly continuous and uniformly bounded. For $(\mu^i, \sigma^{-i}) \in \Delta(\Theta^i) \times \Delta(\mathbb{A}^{-i})$ set

$$V^i(a^i, \mu^i, \sigma^{-i}) := \int_{\Theta^i} \int_{\mathbb{A}^{-i}} \bar{U}^i(a^i, \theta^i, a^{-i})\, \sigma^{-i}(da^{-i})\, \mu^i(d\theta^i).$$

Because $\bar{U}^i$ is bounded and continuous, $(a^i, \mu^i, \sigma^{-i}) \mapsto V^i(a^i, \mu^i, \sigma^{-i})$ is continuous under the weak topologies on $\Delta(\Theta^i)$ and $\Delta(\mathbb{A}^{-i})$. Since $\mathbb{A}^i$ is nonempty and compact (Assump-

tion 3.1(i)), Weierstrass yields nonemptiness and compactness of

$$F^i(\mu^i, \sigma^{-i}) := \arg\max_{a^i \in \mathbb{A}^i} V^i(a^i, \mu^i, \sigma^{-i}).$$

By Berges Maximum Theorem, $(\mu^i, \sigma^{-i}) \mapsto F^i(\mu^i, \sigma^{-i})$ is upper hemicontinuous.

(ii) By Assumption 3.1(iii), for ν^i-a.e. and all (θ^i, a), $|g^i(\theta^i, y, a)| = \left|\log(q^i/q^i_{\theta^i})\right| \leq \log M^i \leq M^i$ and $q^i \leq r^i$, with $\int M^i r^i \, d\nu^i < \infty$. Hence

$$|K^i(\theta^i, a)| = \Big| \int q^i \log \frac{q^i}{q^i_{\theta^i}} \, d\nu^i \Big| \leq \int M^i r^i \, d\nu^i < \infty,$$

uniformly in (θ^i, a). Let $(\theta^i_n, a_n) \to (\theta^i, a)$. By Assumption 3.1(ii), $q^i(\cdot \mid a_n) \to q^i(\cdot \mid a)$ and $q^i_{\theta^i_n}(\cdot \mid a_n) \to q^i_{\theta^i}(\cdot \mid a)$ a.e. If $q^i_{\theta^i} > 0$ then continuity of $(u, v) \mapsto u \log(u/v)$ on $\{u \geq 0, v > 0\}$ gives $q^i(\cdot \mid a_n) \log\big(q^i(\cdot \mid a_n)/q^i_{\theta^i_n}(\cdot \mid a_n)\big) \to q^i(\cdot \mid a) \log\big(q^i(\cdot \mid a)/q^i_{\theta^i}(\cdot \mid a)\big)$ a.e. If $q^i_{\theta^i} = 0$, then by the LR bound $q^i = 0$ as well and $|\log(q^i(\cdot \mid a_n)/q^i_{\theta^i_n}(\cdot \mid a_n))| \leq \log M^i \leq M^i$ while $q^i(\cdot \mid a_n) \to 0$, hence the product converges to 0. Moreover,

$$\big|q^i(\cdot \mid a_n) \log(q^i(\cdot \mid a_n)/q^i_{\theta^i_n}(\cdot \mid a_n))\big| \leq M^i r^i \in L^1(\nu^i),$$

so by the Dominated Convergence Theorem $K^i(\theta^i_n, a_n) \to K^i(\theta^i, a)$. Thus $(\theta^i, a) \mapsto K^i(\theta^i, a)$ is bounded and continuous. Consequently $(\theta^i, \sigma) \mapsto K^i(\theta^i, \sigma)$ is continuous on the compact set $\Theta^i \times \Delta\mathbb{A}$ (weak topology on $\Delta\mathbb{A}$). For each σ, Θ^i is compact and $\theta^i \mapsto K^i(\theta^i, \sigma)$ is continuous, so $\Theta^{m,i}(\sigma) := \arg\min_{\theta^i \in \Theta^i} K^i(\theta^i, \sigma)$ is nonempty and compact. Upper hemicontinuity of $\sigma \mapsto \Theta^{m,i}(\sigma)$ follows from Berge's Maximum Theorem. □

Lemma 3. *(i) Γ is nonempty valued; (ii) For any closed $A \subseteq \mathbb{A}$, the set $\Gamma(A)$ is closed.*

Proof. (i) Nonemptiness follows from the nonemptiness of Θ^m and F (Lemma 2 (1),(2)).
(ii) Let $(a_n)_n$ in $\Gamma(A)$ with $a_n \to a$. For each n, there exists $\sigma_n \in \Delta A$ such that for every $i \in I$ there is $\mu^i_n \in \Delta\Theta^{m,i}(\sigma_n)$ with $a^i_n \in F^i(\mu^i_n, \sigma^{-i}_n)$. Since A is closed in a compact set $\mathbb{A}$, it is also compact, so ΔA is also compact. Therefore (passing to a subsequence if necessary), $\sigma_n \Rightarrow \sigma \in \Delta A$. Because taking opponents' marginals is the pushforward by the (continuous) projection, weak convergence is preserved: for each i, $\sigma^{-i}_n \Rightarrow \sigma^{-i}$. By Lemma 2(ii), the correspondence $\Theta^{m,i}$ is upper hemicontinuous and compact-valued, so along a subsequence we may assume $\mu^i_n \to \mu^i \in \Delta\Theta^{m,i}(\sigma)$. By Lemma 2(i), the best-response correspondence F^i is upper hemicontinuous, so taking limits gives $a^i \in F^i(\mu^i, \sigma^{-i})$. Since this holds for all $i \in I$, we conclude $a \in \Gamma(A)$. Thus $\Gamma(A)$ is closed. □

7.1 Proof of Theorem 2

Let $\mathcal{P}(\mathbb{A})$ denote the power set of $\mathbb{A}$, and consider the partially ordered set $(\mathcal{P}(\mathbb{A}), \subseteq)$. Define the set of sets self-justified under Γ as $\mathcal{C} := \{A \subseteq \mathbb{A} : A \subseteq \Gamma(A)\}$. Then the set of Berk–Nash rationalizable actions is given by $\mathcal{B}' := \cup_{B \in \mathcal{C}} B$, which is the supremum of $\mathcal{C}$. Since Γ is monotone, Tarski's fixed point theorem implies that $\mathcal{B}'$ is the largest fixed point of Γ. The iterative characterization follows from Kleene's fixed point theorem, i.e., $\mathcal{B} = \mathcal{B}'$.

Monotonicity of Γ ensures that the sequence $\{B^k\}$ is nested. Moreover, by Lemma 3 in the appendix, Γ is nonempty-valued and maps closed sets to closed sets. So, starting from the feasible set $B^0 = \mathbb{A}$, each iterate B^k is nonempty and closed (hence, compact). The infinite intersection of nested, nonempty compact sets is nonempty, so $\mathcal{B} = \bigcap_{k \geq 0} B^k$ is nonempty and compact. □

7.2 Example in Section 3.3

We provide a tighter characterization of the iteration of Γ and prove the statement in the text about the threshold k^*.

Fix the box $[0, M] \times [0, N]^2$. For any belief σ supported in this box with $\mathbb{E}_\sigma[a^1] = \mu \in [0, M]$, and for $i \in \{2, 3\}$, set $X := a^1 a^i$. The KL-minimizer is

$$\theta^{m,i}(\sigma) = \theta^* \frac{\mathbb{E}_\sigma\big[(\alpha + X)(\alpha^* + X)\big]}{\mathbb{E}_\sigma\big[(\alpha + X)^2\big]} = \theta^* \frac{\alpha\alpha^* + (\alpha + \alpha^*)\,\mu_X + \nu_X}{\alpha^2 + 2\alpha\,\mu_X + \nu_X},$$

where $\mu_X := \mathbb{E}_\sigma[X]$ and $\nu_X := \mathbb{E}_\sigma[X^2]$. Under only the support and mean constraints,

$$0 \leq \mu_X \leq N\mu, \qquad 0 \leq \nu_X \leq N^2\,\mathbb{E}_\sigma\big[(a^1)^2\big] \leq N^2\,\mu M,$$

with the lower bounds attained by $a^i \equiv 0$, $a^1 \equiv \mu$, and the upper bounds by $a^i \equiv N$ and a twopoint distribution for $a^1 \in \{0, M\}$ with mean μ. Since $\theta^{m,i}(\sigma)$ is increasing in (μ_X, ν_X) for $\alpha > \alpha^*$,

$$\theta^{m,i}(\sigma) \;\in\; \left[\; \theta^* \frac{\alpha^*}{\alpha} \;,\; \theta^* \frac{\alpha\alpha^* + (\alpha + \alpha^*)\,N\mu + N^2\mu M}{\alpha^2 + 2\alpha\,N\mu + N^2\mu M} \;\right].$$

The worker best response is $a^i(\sigma) = (c')^{-1}\big(\mu\,\theta^{m,i}(\sigma)\big)$, hence

$$a^i(\sigma) \;\in\; \left[(c')^{-1}\big(\mu\,\theta^* \tfrac{\alpha^*}{\alpha}\big),\; (c')^{-1}\Big(\mu\,\theta^* \tfrac{\alpha\alpha^* + (\alpha+\alpha^*)\,N\mu + N^2\mu M}{\alpha^2 + 2\alpha\,N\mu + N^2\mu M}\Big)\right].$$

Define $\mathrm{diff}(\mu) := U(\mu) - L(\mu)$, where

$$L(\mu) = (c')^{-1}\Big(\mu\,\theta^*\tfrac{\alpha^*}{\alpha}\Big), \qquad U(\mu) = (c')^{-1}\Big(\mu\,\theta^*\,\frac{\alpha\alpha^* + (\alpha+\alpha^*)\,N\mu + N^2\mu M}{\alpha^2 + 2\alpha\,N\mu + N^2\mu M}\Big).$$

We show below that $\mathrm{diff}(\cdot)$ is nondecreasing. It follows that

$$k^* \;=\; \sup_{\mu\in[0,M]} \mathrm{diff}(\mu) \;=\; (c')^{-1}\Big(M\,\theta^*\tfrac{\alpha^*+NM}{\alpha+NM}\Big) \;-\; (c')^{-1}\Big(M\,\theta^*\tfrac{\alpha^*}{\alpha}\Big),$$

and in the text we use $M = M_\infty$ and $N = N_\infty$.

Proof that $\mathrm{diff}(\cdot)$ **is nondecreasing:** Let

$$L(\mu) = (c')^{-1}\big(\ell(\mu)\big), \quad \ell(\mu) = \mu\,\theta^*\tfrac{\alpha^*}{\alpha}, \qquad U(\mu) = (c')^{-1}\big(u(\mu)\big), \quad u(\mu) = \mu\,\theta^* r(\mu),$$

with

$$r(\mu) = \frac{\alpha\alpha^* + (\alpha+\alpha^*)N\mu + N^2 M\,\mu}{\alpha^2 + 2\alpha N\mu + N^2 M\,\mu}.$$

Define the inner gap

$$g(\mu) := u(\mu) - \ell(\mu) = \theta^*\mu\Big(r(\mu) - \tfrac{\alpha^*}{\alpha}\Big).$$

A direct calculation yields

$$r'(\mu) = \frac{N\,\alpha(\alpha-\alpha^*)(\alpha+NM)}{\big(\alpha^2 + 2\alpha N\mu + N^2 M\mu\big)^2} > 0 \quad (\alpha > \alpha^*),$$

hence $g'(\mu) \geq 0$, with equality only at $\mu = 0$. Thus $u(\mu) - \ell(\mu)$ is increasing on $(0, M]$. Since $(c')^{-1}$ is increasing, the desired result follows.

Proof of Proposition 1

Suppose $\sigma_{t_k} \Rightarrow \sigma$ on $\mathbb{A}$. The continuous mapping theorem applied to the projection onto $\mathbb{A}^{-i}$ gives $\sigma^{-i}_{t_k} \Rightarrow \sigma^{-i}$ and hence also $\sigma^{-i}_{t_k-1} \Rightarrow \sigma^{-i}$. Let $w_t := \frac{\alpha_0^i}{\alpha_0^i+t-1} \to 0$ and note that $\tilde\sigma^{-i}_t = w_t\nu^i + (1-w_t)\sigma^{-i}_{t-1}$. For any bounded continuous $f : \mathbb{A}^{-i} \to \mathbb{R}$,

$$\int f\,d\tilde\sigma^{-i}_{t_k} = w_{t_k}\int f\,d\nu^i \;+\; (1-w_{t_k})\int f\,d\sigma^{-i}_{t_k-1} \;\longrightarrow\; \int f\,d\sigma^{-i},$$

because $w_{t_k} \to 0$ and $\int f\,d\sigma^{-i}_{t_k-1} \to \int f\,d\sigma^{-i}$. Thus $\tilde\sigma^{-i}_{t_k} \Rightarrow \sigma^{-i}$.

References

Aumann, Robert J. 1974. "Subjectivity and correlation in randomized strategies." *Journal of mathematical Economics*, 1(1): 67–96.

Ba, Cuimin, and Alice Gindin. 2023. "A multi-agent model of misspecified learning with overconfidence." *Games and Economic Behavior*, 142: 315–338.

Basu, Kaushik, and Jörgen W Weibull. 1991. "Strategy subsets closed under rational behavior." *Economics Letters*, 36(2): 141–146.

Berk, Robert H. 1966. "Limiting behavior of posterior distributions when the model is incorrect." *The Annals of Mathematical Statistics*, 37(1): 51–58.

Bernheim, B Douglas. 1984. "Rationalizable strategic behavior." *Econometrica: Journal of the Econometric Society*, 1007–1028.

Bohren, J Aislinn. 2016. "Informational herding with model misspecification." *Journal of Economic Theory*, 163: 222–247.

Bohren, J Aislinn, and Daniel N Hauser. 2021. "Learning with heterogeneous misspecified models: Characterization and robustness." *Econometrica*, 89(6): 3025–3077.

Brandenburger, Adam, and Eddie Dekel. 1987. "Rationalizability and correlated equilibria." *Econometrica: Journal of the Econometric Society*, 1391–1402.

Chauvin, Kyle. 2023. "A misattribution theory of discrimination." working paper.

Eliaz, Kfir, and Ran Spiegler. 2020. "A model of competing narratives." *American Economic Review*, 110(12): 3786–3816.

Esponda, Ignacio. 2008. "Behavioral equilibrium in economies with adverse selection." *American Economic Review*, 98(4): 1269–1291.

Esponda, Ignacio, and Demian Pouzo. 2016. "Berk–Nash equilibrium: A framework for modeling agents with misspecified models." *Econometrica*, 84(3): 1093–1130.

Esponda, Ignacio, and Demian Pouzo. 2025. "Berk–Nash Rationalizability in Games with Asymmetric Information." Working paper.

Esponda, Ignacio, Demian Pouzo, and Yuichi Yamamoto. 2021. "Asymptotic behavior of Bayesian learners with misspecified models." *Journal of Economic Theory*, 195: 105260.

Eyster, Erik, and Matthew Rabin. 2005. "Cursed equilibrium." *Econometrica*, 73(5): 1623–1672.

Foster, Dean P, and Rakesh V Vohra. 1997. "Calibrated learning and correlated equilibrium." *Games and Economic Behavior*, 21(589): 40–55.

Frick, Mira, Ryota Iijima, and Yuhta Ishii. 2020. "Misinterpreting others and the fragility of social learning." *Econometrica*, 88(6): 2281–2328.

Frick, Mira, Ryota Iijima, and Yuhta Ishii. 2022. "Dispersed behavior and perceptions in assortative societies." *American Economic Review*, 112(9): 3063–3105.

Frick, Mira, Ryota Iijima, and Yuhta Ishii. 2023. "Belief convergence under misspecified learning: A martingale approach." *The Review of Economic Studies*, 90(2): 781–814.

Frick, Mira, Ryota Iijima, and Yuhta Ishii. 2024. "Welfare comparisons for biased learning." *American Economic Review*, 114(6): 1612–1649.

Fudenberg, Drew, and David K Levine. 1999. "Conditional universal consistency." *Games and Economic Behavior*, 29(1-2): 104–130.

Fudenberg, Drew, and David M Kreps. 1993. "Learning mixed equilibria." *Games and economic behavior*, 5(3): 320–367.

Fudenberg, Drew, Giacomo Lanzani, and Philipp Strack. 2021. "Limit points of endogenous misspecified learning." *Econometrica*, 89(3): 1065–1098.

Fudenberg, Drew, Giacomo Lanzani, and Philipp Strack. 2024. "Selective-Memory Equilibrium." *Journal of Political Economy*, 132(12): 3978–4020.

Fudenberg, Drew, Gleb Romanyuk, and Philipp Strack. 2017. "Active learning with a misspecified prior." *Theoretical Economics*, 12(3): 1155–1189.

Hart, Sergiu, and Andreu Mas-Colell. 2000. "A simple adaptive procedure leading to correlated equilibrium." *Econometrica*, 68(5): 1127–1150.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack. 2018. "Unrealistic expectations and misguided learning." *Econometrica*, 86(4): 1159–1214.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack. 2021. "Convergence in models of misspecified learning." *Theoretical Economics*, 16(1): 73–99.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack. 2023. "Misinterpreting yourself." *Available at SSRN 4325160*.

He, Kevin. 2022. "Mislearning from censored data: The gambler's fallacy and other correlational mistakes in optimal-stopping problems." *Theoretical Economics*, 17(3): 1269–1312.

He, Kevin, and Jonathan Libgober. 2025. "Higher-Order Beliefs and (Mis) learning from Prices." Working Paper.

Hofbauer, Josef, and William H Sandholm. 2002. "On the global convergence of stochastic fictitious play." *Econometrica*, 70(6): 2265–2294.

Jehiel, Philippe. 2005. "Analogy-based expectation equilibrium." *Journal of Economic theory*, 123(2): 81–104.

Milgrom, Paul, and John Roberts. 1990. "Rationalizability, learning, and equilibrium in games with strategic complementarities." *Econometrica: Journal of the Econometric Society*, 1255–1277.

Milgrom, Paul, and John Roberts. 1991. "Adaptive and sophisticated learning in normal form games." *Games and economic Behavior*, 3(1): 82–100.

Molavi, Pooya, et al. 2019. "Macroeconomics with learning and misspecification: A general theory and applications." *Unpublished manuscript*.

Monderer, Dov, and Lloyd S Shapley. 1996. "Potential games." *Games and economic behavior*, 14(1): 124–143.

Murooka, Takeshi, and Yuichi Yamamoto. 2023. "Convergence and Steady-State Analysis under Higher-Order Misspecification." Osaka School of International Public Policy, Osaka University.

Murooka, Takeshi, and Yuichi Yamamoto. 2025. "Bayesian Learning When Players Are Misspecified about Others." Institute of Social and Economic Research, The University of Osaka.

Nyarko, Yaw. 1991. "Learning in mis-specified models and the possibility of cycles." *Journal of Economic Theory*, 55(2): 416–427.

Pearce, David G. 1984. "Rationalizable strategic behavior and the problem of perfection." *Econometrica: Journal of the Econometric Society*, 1029–1050.

Schwartzstein, Joshua. 2014. "Selective attention and learning." *Journal of the European Economic Association*, 12(6): 1423–1452.

Shapley, Lloyd S. 1964. "Some Topics in Two-Person Games." In *Advances in Game Theory. Annals of Mathematics Studies*, , ed. Melvin Dresher, Lloyd S. Shapley and A. W. Tucker, 1–29. Princeton, NJ:Princeton University Press.

Spiegler, Ran. 2016. "Bayesian networks and boundedly rational expectations." *The Quarterly Journal of Economics*, 131(3): 1243–1290.

Tan, Tommy Chin-Chiu, and Sérgio Ribeiro da Costa Werlang. 1988. "The Bayesian foundations of solution concepts of games." *Journal of Economic Theory*, 45(2): 370–391.

Supplemental Appendix

S1 Proof of Lemma 1

In this Supplemental Appendix, we provide the proof of Lemma 1 for completeness. As remarked in the text, this proof is a straightforward extension of existing ones (e.g. Esponda, Pouzo and Yamamoto (2021)) to a continuum of actions and convergent subsequences instead of sequences.

The result holds for each player i. Henceforth, we drop the i superscript from the notation and focus on an individual player.

Define

$$L_t(a^\infty, y^\infty)(\theta) := t^{-1} \sum_{\tau=1}^{t} \ln \frac{q(y_\tau \mid a_\tau)}{q_\theta(y_\tau \mid a_\tau)}.$$

We will use the following fact:

Uniform SLLN. For any a^∞, there exists $\mathcal{Y}_{a^\infty}$ with $\mathbf{P}(\mathcal{Y}_{a^\infty} \mid a^\infty) = 1$ such that, for any $y^\infty \in \mathcal{Y}_{a^\infty}$ and for every $\varepsilon > 0$ there exists $T^1_\varepsilon(a^\infty, y^\infty)$ such that for all $t \geq T^1_\varepsilon(a^\infty, y^\infty)$,

$$\sup_{\theta \in \Theta} |L_t(a^\infty, y^\infty)(\theta) - \int K(\theta, a)\sigma_t(a^\infty)(da)| < \varepsilon. \tag{S1}$$

See Section S1.1 for a proof of this result.

From this point on, we fix any a^∞ and consider any $y^\infty \in \mathcal{Y}_{a^\infty}$. For simplicity, we drop (a^∞, y^∞) from the notation. Let $(\sigma_{t_k})_k$ be a subsequence converging to σ and let $E \subseteq \Theta$ be a closed set disjoint from $\Theta^m(\sigma)$. We rely on the following two results:

Approximate $(\sigma_{t_k})_k$ with its limit σ. Since K is continuous and Θ is compact, then for every $\varepsilon > 0$, there exists T^2_ε such that, for all k such that $t_k \geq T^2_\varepsilon$,

$$\sup_{\theta \in \Theta} |\int K(\theta, a)\sigma_{t_k}(da) - \int K(\theta, a)\sigma(da)| < \varepsilon.$$

E is well separated. Since K is continuous and E is closed in a compact set (hence, compact), there exists $\delta > 0$ such that for all $\theta \in E$,

$$\int K(\theta, a)\sigma(da) \geq K^*(\sigma) + \delta,$$

where $K^*(\sigma) := \min_{\theta\in\Theta} \int K(\theta, a)\sigma(da)$.

We now prove (8) in Lemma 1. For any $\xi > 0$, note that

$$\mu_{t_k}(E) = \frac{\int_E e^{-tL_{t_k}(\theta)}\mu_0(d\theta)}{\int_\Theta e^{-tL_{t_k}(\theta)}\mu_0(d\theta)} \leq \frac{\int_E e^{-t_k L_{t_k}(\theta)}\mu_0(d\theta)}{\int_{\Theta_\xi(\sigma)} e^{-t_k L_{t_k}(\theta)}\mu_0(d\theta)},$$

where $\Theta_\xi(\sigma) := \{\theta \in \Theta : \int K(\theta, a)\sigma(da) \leq K^*(\sigma) + \xi\}$.

Consider first the numerator. From UniformSLLN, the approximation of $(\sigma_{t_k})_k$ with its limit σ, and well separation of E, it follows that, for all k such that $t_k \geq \max\{T^1_\varepsilon, T^2_\varepsilon\}$, $L_{t_k}(\theta) \geq K^*(\sigma) + \delta - 2\varepsilon$ for all $\theta \in E$, and so

$$\int_E e^{-t_k L_{t_k}(\theta)}\mu_0(d\theta) \leq \mu_0(E)e^{-t_k(K^*(\sigma)+\delta-2\varepsilon)}. \tag{S2}$$

Next, consider the denominator. From UniformSLLN and the approximation of $(\sigma_{t_k})_k$ with its limit σ, it follows that, for all k such that $t_k \geq \max\{T^1_\varepsilon, T^2_\varepsilon\}$, $L_{t_k}(\theta) \leq K^*(\sigma)+\xi+2\varepsilon$ for all $\theta \in \Theta_\xi(\sigma)$, and so

$$\int_{\Theta_\xi(\sigma)} e^{-t_k L_{t_k}(\theta)}\mu_0(d\theta) \geq \mu_0(\Theta_\xi(\sigma))e^{-t_k(K^*(\sigma)+\xi+2\varepsilon)}. \tag{S3}$$

Combining (S2) and (S3) and setting $\varepsilon = \xi < \delta/10$, we obtain that for all k such that $t_k \geq \max\{T^1_\varepsilon, T^2_\varepsilon\}$,

$$\mu_{t_k}(E) \leq \frac{\mu_0(E)}{\mu_0(\Theta_\xi(\sigma))}e^{-t_k(\delta/2)}.$$

Since Θ is compact and $\theta \mapsto \int K(\theta, a)\sigma(da)$ is continuous, a minimizer $\theta^* \in \Theta$ exists and satisfies $\theta^* \in \Theta_\xi(\sigma)$. By continuity, there exists an open ball around θ^* contained in $\Theta_\xi(\sigma)$. Since μ_0 has full support, this ball has positive measure, so $\mu_0(\Theta_\xi(\sigma)) > 0$. Equation (8) in the statement of Lemma 1 then follows by setting $\rho = \delta/2 > 0$ and $C = \mu_0(E)/\mu_0(\Theta_\xi(\sigma))$. Since $\mu_0(E) \geq 0$ and $\mu_0(\Theta_\xi(\sigma)) > 0$, it follows that $C \geq 0$ and $C < \infty$.

Suppose, in addition, that $(\mu_{t_k})_k$ converges to μ. By equation (8), for any closed E such that $E \cap \Theta^m(\sigma) = \emptyset$, $\liminf_{k\to\infty} \mu_{t_k}(E) = 0$. By the Pormanteau lemma, $\mu(E) \leq \liminf_{k\to\infty} \mu_{t_k}(E) = 0$. Now consider any $\theta \in \Theta \setminus \Theta^m(\sigma)$. There exists an open neighborhood U_θ with closure $\bar{U}_\theta$ such that $\bar{U}_\theta \cap \Theta^m(\sigma) = \emptyset$. Since $\bar{U}_\theta$ is closed, $\mu(\bar{U}_\theta) = 0$, and so θ is not in the support of μ. Thus, $supp\, \mu \subseteq \Theta^m(\sigma)$.

S1.1 Uniform

Esponda, Pouzo and Yamamoto (2021) (henceforth, EPY 2021) proves the uniform SLLN (equation S1) for the case of a finite number of actions (see their Lemma 2). The proof for the case where $\mathbb{A}^i$ is a compact subset of Euclidean space is essentially identical, but we need to establish the following two claims uniformly for all actions.

Claim (uniform L^2 bound for g). Fix i, $\theta \in \Theta^i$, and $\varepsilon > 0$. Then

$$\sup_{a\in\mathbb{A}} \int_{\mathbb{Y}^i} \sup_{\theta'\in O(\theta,\varepsilon)} \left[g^i(\theta', y, a)\right]^2 Q^i(dy \mid a) \;<\; \infty.$$

Proof. For any a,

$$\int \sup_{\theta'\in O(\theta,\varepsilon)} \left[g^i(\theta', y, a)\right]^2 Q^i(dy \mid a) = \int \sup_{\theta'} \left|\log \frac{q^i(y \mid a)}{q^i_{\theta'}(y \mid a)}\right|^2 q^i(y \mid a)\, \nu^i(dy).$$

By the LR bound, $\sup_{\theta'}\left|\log(q^i/q^i_{\theta'})\right| \le \log M^i \le M^i$ a.e., hence

$$\le \int \left(\log M^i(y)\right)^2 q^i(y \mid a)\, \nu^i(dy) \;\le\; 2\int M^i(y)\, q^i(y \mid a)\, \nu^i(dy),$$

where we used $(\log t)^2 \le 2t$ for $t \ge 1$. By the density envelope, $q^i(\cdot \mid a) \le r^i$ a.e., so

$$\le 2\int M^i(y)\, r^i(y)\, \nu^i(dy),$$

which is finite by Assumption 3.1(iii) and independent of a. Taking the supremum over $a \in \mathbb{A}$ gives the claim. □

Claim (uniform oscillation bound, eq. (12) in EPY 2021). Fix i, $\theta \in \Theta^i$, and $\varepsilon > 0$. Then there exists $\delta(\theta, \varepsilon) > 0$ such that

$$\sup_{a\in\mathbb{A}} \mathbb{E}_{Q^i(\cdot|a)}\Big[\sup_{\theta'\in O(\theta,\delta(\theta,\varepsilon))} \left|g^i(\theta', Y, a) - g^i(\theta, Y, a)\right|\Big] \;<\; 0.25\,\varepsilon.$$

Proof. Let $H_r(y) := \sup_{\|\theta'-\theta\|\le r,\ a\in\mathbb{A}} \left|g^i(\theta', y, a) - g^i(\theta, y, a)\right|$. By a.e.-continuity on the compact set $\Theta^i \times \mathbb{A}$ (Assumption 3.1(ii)), $H_r(y) \downarrow 0$ for ν^i-a.e. y as $r \downarrow 0$. By the LR bound (Assumption 3.1(iii)), $H_r(y) \le 2|g^i(\theta, y, a)| \le 2\log M^i(y) \le 2M^i(y)$ a.e., and by the density

envelope $Q^i(dy \mid a) = q^i(y \mid a)\nu^i(dy) \le r^i(y)\nu^i(dy)$. Hence, for every a,

$$\mathbb{E}_{Q^i(\cdot|a)}\Big[\sup_{\theta' \in O(\theta,r)} |g^i(\theta', Y, a) - g^i(\theta, Y, a)|\Big] \le \int H_r(y)\, q^i(y \mid a)\, \nu^i(dy) \le \int H_r(y)\, r^i(y)\, \nu^i(dy).$$

Since $H_r \to 0$ a.e. and $H_r(\cdot)\, r^i(\cdot) \le 2M^i(\cdot)\, r^i(\cdot)$ with $\int 2M^i r^i\, d\nu^i < \infty$ (Assumption 3.1(iii)), the Dominated Convergence Theorem yields $\int H_r r^i\, d\nu^i \to 0$ as $r \downarrow 0$. Thus

$$\sup_{a \in \mathbb{A}}\ \mathbb{E}_{Q^i(\cdot|a)}\Big[\sup_{\theta' \in O(\theta,r)} |g^i(\theta', Y, a) - g^i(\theta, Y, a)|\Big] \le \int H_r r^i\, d\nu^i \ \underset{r \downarrow 0}{\longrightarrow}\ 0.$$

Choose $\delta(\theta, \varepsilon) > 0$ so that the right-hand side is $< 0.25\, \varepsilon$. □