

Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models*

Ignacio Esponda
(WUSTL)

Demian Pouzo
(UC Berkeley)

August 18, 2015

Abstract

We develop an equilibrium framework that relaxes the standard assumption that people have a correctly-specified view of their environment. Each player is characterized by a (possibly misspecified) subjective model, which describes the set of feasible beliefs over payoff-relevant consequences as a function of actions. We introduce the notion of a Berk-Nash equilibrium: Each player follows a strategy that is optimal given her belief, and her belief is restricted to be the best fit among the set of beliefs she considers possible. The notion of best fit is formalized in terms of minimizing the Kullback-Leibler divergence, which is endogenous and depends on the equilibrium strategy profile. Standard solution concepts such as Nash equilibrium and self-confirming equilibrium constitute special cases where players have correctly-specified models. We provide a learning foundation for Berk-Nash equilibrium by extending and combining results from the statistics literature on misspecified learning and the economics literature on learning in games.

*We thank Vladimir Asriyan, Pierpaolo Battigalli, Larry Blume, Aaron Bodoh-Creed, Sylvain Chassang, Emilio Espino, Erik Eyster, Drew Fudenberg, Yuriy Gorodnichenko, Philippe Jehiel, Stephan Lauermaun, Natalia Lazzati, Kristóf Madarász, Matthew Rabin, Ariel Rubinstein, Joel Sobel, Jörg Stoye, several seminar participants, and four referees for very helpful comments. Esponda: Olin Business School, Washington University in St. Louis, 1 Brookings Drive, Campus Box 1133, St. Louis, MO 63130, iesponda@wustl.edu; Pouzo: Department of Economics, UC Berkeley, 530-1 Evans Hall #3880, Berkeley, CA 94720, dpouzo@econ.berkeley.edu.

1 Introduction

Economic models provide a framework to understand complex environments. Most economists recognize that the simplifying assumptions underlying our models are often wrong. But, despite recognizing that our models are likely to be misspecified, the standard approach in economics is to assume that the economic agents themselves have a correctly specified view of their environment. In this paper, we introduce an equilibrium framework that relaxes this standard assumption and allows the modeler to postulate that economic agents have a subjective and possibly incorrect view of their world.

An objective game represents the true environment faced by the agent (or players, in the case of several interacting agents). Payoff relevant states and privately observed signals are realized according to some objective probability distribution. Each player observes her own private signal and then players simultaneously choose actions. The action profile and the realized payoff-relevant state determine a consequence for each player, and consequences determine payoffs. This objective description of the environment is fairly standard in economics.

While it is also standard to implicitly assume that players know the objective game, we deviate from this practice by assuming that each player has a subjective model that represents her own view of the environment. Formally, a subjective model is a set of probability distributions over own consequences as a function of a player's own action and information. A key feature is that we allow the subjective model of one or more players to be misspecified, which roughly means that the set of subjective distributions does not include the true, objective distribution. For example, firms might incorrectly believe that sales depend only on their own price and not also on the price of other firms. Or a consumer might perceive a nonlinear price schedule to be linear and, therefore, respond to average, not marginal, prices. Or traders might not realize that the value of trade is partly determined by the terms of trade.

A *Berk-Nash equilibrium* is defined to be a strategy profile such that, for each player, there exists a belief with support in that player's subjective model that satisfies two conditions. First, the strategy must be optimal (in a static sense) given the belief. Second, the belief puts probability one on the set of subjective distributions over consequences that are "closest" to the true distribution, where the true distribution is determined by the objective game and the actual strategy profile. The notion of

“closest” is given by a weighted version of the Kullback-Leibler divergence, also known as relative entropy, that we define formally in the main text.

Berk-Nash equilibrium includes both standard and boundedly rational solution concepts in a common framework, such as Nash, self-confirming (e.g., Battigalli [1987], Fudenberg and Levine [1993a], Dekel et al. [2004]), fully cursed (Eyster and Rabin, 2005), and analogy-based expectation equilibrium (Jehiel [2005], Jehiel and Koessler, 2008). For example, suppose that the game is *correctly specified* (which means that the support of each player’s prior contains the true, objective distribution) and that the game is *strongly identified* (which means that there is always a unique distribution—whether or not correct—that matches the observed data). Then Berk-Nash equilibrium is equivalent to Nash equilibrium. If the strong identification assumption is dropped, then Berk-Nash is a self-confirming equilibrium. In addition to unifying previous approaches, the framework provides a systematic approach for both extending cases previously analyzed in the literature and considering new types of misspecifications.

We provide a foundation for Berk-Nash equilibrium (and the corresponding use of Kullback-Leibler divergence as a measure of “distance”) by studying a dynamic environment with a fixed number of players playing the objective game repeatedly. Players believe that the environment is stationary and start with a prior over a set of subjective distributions over consequences. In each period, they play the objective game and use the observed consequences to update their beliefs according to Bayes’ rule. The main objective is to characterize limiting behavior when players behave optimally but learn with a possibly misspecified subjective model.¹

The main result is that, if players’ behavior converges, then it converges to a Berk-Nash equilibrium of the game. A converse of the main result, showing that we can converge to any Berk-Nash equilibrium of the game for some initial (non-doctrinaire) prior, does not hold. But we do obtain a positive convergence result by relaxing the assumption that players exactly optimize. We show that, for any Berk-Nash equilibrium, there exists a policy rule that is myopic and *asymptotically optimal* (in the sense that optimization mistakes vanish with time) under which convergence to equilibrium occurs with probability one.

There is a longstanding interest among economists in studying the behavior of agents who hold misspecified views of the world. Early examples come from such di-

¹In the case of multiple agents, the environment need not be stationary, and so we are ignoring repeated game considerations where players take into account how their actions affect others’ future play. We extend the results to a population model with a continuum of agents in Section 5.

verse fields as industrial organization, mechanism design, psychology and economics, macroeconomics, and information economics (e.g., Arrow and Green [1973], Kirman [1975], Sobel [1984], Nyarko [1991], Sargent [1999], Kagel and Levin [1986]), although many times there is no explicit reference to a problem of misspecified learning. Most of the literature, however, focuses on particular settings, and there has been little progress in developing a unified framework. Our treatment unifies both “rational” and “boundedly rational” approaches, thus emphasizing that modeling the behavior of misspecified players does not constitute a large departure from the standard framework.

Arrow and Green [1973] provide a general treatment and make a distinction between an objective and subjective game. Their framework, though, is more restrictive than ours in terms of the types of misspecifications that players are allowed to have. Moreover, they do not establish existence of equilibrium and they do not provide a learning foundation for equilibrium.² Recently, Spiegel [2014] introduced a framework that uses Bayesian networks to analyze decision making under imperfect understanding of correlation structures.

Our paper is also related to the bandit (e.g., Rothschild [1974], McLennan [1984], Easley and Kiefer [1988]) and the self-confirming equilibrium literatures, which point out that agents might optimally end up with incorrect beliefs if feedback is incomplete and experimentation is costly.³ We also allow for beliefs to be incorrect due to insufficient feedback, but our main contribution is to allow for misspecified learning. When players have misspecified models, beliefs may be incorrect and endogenously depend on own actions even if players persistently experiment with all actions; thus, an equilibrium framework is needed to characterize steady-state behavior even in single-agent settings.⁴

²Misspecified models have also been studied in contexts that are outside the scope of our paper either because the decision problem is dynamic (we focus on the repetition of a static problem) or because interactions are mediated by a price system. Examples include the early literature on rational expectations with misspecified players (e.g., Blume and Easley [1982], Bray [1982], and Radner [1982]), the macroeconomics literature on bounded rationality (e.g., Sargent [1993], Evans and Honkapohja [2001]), a behavioral finance literature that studies under and over-reaction to information (e.g., Barberis et al., 1998), and a literature that formalizes psychological biases and studies related applications (e.g., Rabin [2002], Rabin and Vayanos [2010], Spiegel [2013]).

³In the macroeconomics literature, the term “self-confirming equilibrium” is sometimes used in a broader sense to include cases where agents have misspecified models (e.g., Sargent, 1999).

⁴The literature on self-confirming equilibrium considers two interesting extensions, which are also potentially applicable to our framework: refinements that restrict beliefs by allowing players to introspect about other players’ motivations (e.g., Rubinstein and Wolinsky, 1994), and non-Bayesian models of updating that capture ambiguity aversion (Battigalli et al., 2012).

From a technical point of view, our results extend and combine results from two literatures. First, the idea that equilibrium is a result of a learning process is taken from the literature on learning in games. This literature studies explicit learning models in order to justify Nash and self-confirming equilibrium (e.g., Fudenberg and Kreps [1988], Fudenberg and Kreps [1993], Fudenberg and Kreps [1995], Fudenberg and Levine [1993b], Kalai and Lehrer [1993]).⁵ In particular, we follow Fudenberg and Kreps [1993] in making the assumption that payoffs are perturbed, à la Harsanyi [1973], to guarantee that behavior is continuous in beliefs and, therefore, to justify how players might learn to play mixed strategy equilibria. We also rely on an idea by Fudenberg and Kreps [1993] to prove the converse of the main result. We extend this literature to account for the possibility that players learn with models of the world that are misspecified even in steady state.

Second, we rely on and contribute to the statistics literature that studies the consistency of Bayesian updating and characterizes limiting beliefs. In decision problems with correctly-specified models, the standard approach is to use martingale convergence theorems to prove that beliefs converge (e.g., Easley and Kiefer, 1988). This result guarantees convergence of beliefs from a subjective point of view, which is, unfortunately, not useful for our results because beliefs might still not converge in an objective sense when the agent has a misspecified model. Thus, we take a different route and follow the statistics literature on misspecified learning. This literature characterizes limiting beliefs in terms of the Kullback-Leibler divergence (e.g., Berk [1966], Bunke and Milhaud [1998]). We extend this statistics literature to the case where agents are not only passively learning about their environment but are also actively learning by taking actions.

We present the framework and several examples in Section 2, discuss the relationship to other solution concepts in Section 3, and provide a learning foundation in Section 4. We discuss assumptions and extension in Section 5 and conclude in Section 6.

⁵See Fudenberg and Levine (1998, 2009) for a survey of this literature.

2 The framework

2.1 The environment

A (simultaneous-move) **game** $\mathcal{G} = \langle \mathcal{O}, \mathcal{Q} \rangle$ is composed of a (simultaneous-move) objective game \mathcal{O} and a subjective model \mathcal{Q} . We now describe each of these two components in detail.

OBJECTIVE GAME. A (simultaneous-move) **objective game** is a tuple

$$\mathcal{O} = \langle I, \Omega, \mathbb{S}, p, \mathbb{X}, \mathbb{Y}, f, \pi \rangle,$$

where: I is the finite set of players (a special case is a single-agent decision problem); Ω is the set of payoff-relevant states; $\mathbb{S} = \times_{i \in I} \mathbb{S}^i$ is the set of profiles of signals, where \mathbb{S}^i is the set of signals of player i ; p is a probability distribution over $\Omega \times \mathbb{S}$, and, for simplicity, it is assumed to have marginals with full support; $\mathbb{X} = \times_{i \in I} \mathbb{X}^i$ is a set of profiles of actions, where \mathbb{X}^i is the set of actions of player i ; $\mathbb{Y} = \times_{i \in I} \mathbb{Y}^i$ is a set of profiles of (observable) consequences, where \mathbb{Y}^i is the set of consequences of player i ; $f = (f^i)_{i \in I}$ is a profile of feedback or consequence functions, where $f^i : \mathbb{X} \times \Omega \rightarrow \mathbb{Y}^i$ maps outcomes in $\Omega \times \mathbb{X}$ into consequences of player i ; and $\pi = (\pi^i)_{i \in I}$, where $\pi^i : \mathbb{X}^i \times \mathbb{Y}^i \rightarrow \mathbb{R}$ is the payoff function of player i .⁶ For simplicity, we prove the results for the case where all of the above sets are finite. In Online Appendix F, we provide technical assumptions under which the results extend to non-finite Ω and \mathbb{Y} .

The timing of the objective game is as follows: First, a state and a profile of signals are drawn according to the probability distribution p . Second, each player privately observes her own signal. Third, each player simultaneously chooses an action. Finally, each player observes her consequence and obtains a payoff. Note that the setup implicitly assumes that players observe at least their own payoffs and actions.

A strategy of player i is a mapping $\sigma^i : \mathbb{S}^i \rightarrow \Delta(\mathbb{X}^i)$. The probability that player i chooses action x^i after observing signal s^i is denoted by $\sigma^i(x^i | s^i)$. A strategy profile is a vector of strategies $\sigma = (\sigma^i)_{i \in I}$; let Σ denote the space of all strategy profiles.

Fix an objective game. For each strategy profile σ , there is an **objective distribution** over player i 's consequences, Q_σ^i , where, for each $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, $Q_\sigma^i(\cdot |$

⁶While it is redundant to have π^i also depend on x^i , it helps to reduce notation, particularly in applications.

$s^i, x^i) \in \Delta(\mathbb{Y}^i)$ is defined as follows:

$$Q_{\sigma}^i(y^i | s^i, x^i) = \sum_{\{(\omega, x^{-i}): f^i(x^i, x^{-i}, \omega) = y^i\}} \sum_{s^{-i}} \prod_{j \neq i} \sigma^j(x^j | s^j) p(\omega, s^{-i} | s^i). \quad (1)$$

The objective distribution represents the true distribution over consequences, conditional on a player's own action and signal, given the objective game and a strategy profile followed by the players.

SUBJECTIVE MODEL. The subjective model represents the set of distributions over consequences that players consider possible a priori. For a fixed objective game, a **subjective model** is a tuple

$$\mathcal{Q} = \langle \Theta, (Q_{\theta})_{\theta \in \Theta} \rangle,$$

where: $\Theta = \times_{i \in I} \Theta^i$, where Θ^i is the parameter space of player i ; and $Q_{\theta} = (Q_{\theta^i}^i)_{i \in I}$, where $Q_{\theta^i}^i : \mathbb{S}^i \times \mathbb{X}^i \rightarrow \Delta(\mathbb{Y}^i)$ represents the conditional distribution over player i 's consequences parameterized by $\theta^i \in \Theta^i$; we denote the conditional distribution by $Q_{\theta^i}^i(\cdot | s^i, x^i)$.⁷

While the objective game represents the true environment, the subjective model represents the players' perception of their environment. This separation between objective and subjective models, which is often implicit in standard treatments of games, is crucial in this paper.

A special case of a subjective model is one where each agent understands the objective game being played but is uncertain about the distribution over states, the consequence function, and (in the case of multiple players) the strategies of other players. In that case, agent i 's subjective game is parameterized by $p_{\theta^i}^i, f_{\theta^i}^i, \sigma_{\theta^i}^{-i}$ and a set Θ^i , and the subjective model $Q_{\theta^i}^i$ is derived by replacing these primitives in equation (1).

By directly defining $Q_{\theta^i}^i$ as a primitive, we stress two points. First, this is the object that is needed to characterize behavior in our setting. Moreover, by working with general subjective distributions over consequences, we allow for more general types of misspecifications, where players do not even have to understand the structural elements that determine their payoff relevant consequences.

⁷For simplicity, this definition of a subjective model assumes that players know the distribution over their own signals.

We maintain the following assumptions about the subjective model.⁸

Assumption 1. For all $i \in I$: (i) Θ^i is a compact subset of an Euclidean space, (ii) $Q_{\theta^i}^i(y^i | s^i, x^i)$ is continuous as a function of $\theta^i \in \Theta^i$ for all $(y^i, s^i, x^i) \in \mathbb{Y}^i \times \mathbb{S}^i \times \mathbb{X}^i$, (iii) For every $\theta^i \in \Theta^i$, there exists a sequence $(\theta_n^i)_n$ in Θ^i such that $\lim_{n \rightarrow \infty} \theta_n^i = \theta^i$ and, for all n , $Q_{\theta_n^i}^i(y^i | s^i, x^i) > 0$ for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, $y^i \in f^i(x^i, \mathbb{X}^{-i}, \omega)$, and $\omega \in \text{supp}(p(\cdot | s^i))$.

The last of these assumptions plays two roles. First, it rules out a stark form of misspecification by guaranteeing that there exists at least one parameter value that can rationalize every feasible observation. Second, it requires that if indeed a feasible event is deemed impossible by some parameter value, then that parameter value is not isolated in the sense that there are nearby parameter values that consider every feasible event to be possible. In Section 5, we show that equilibrium may fail to exist and steady-state behavior need not be characterized by equilibrium without this continuity assumption.

2.2 Examples

We illustrate our environment by using several examples that had previously not been integrated into a common framework. Also, previous work considered special cases with relatively simple misspecifications, in which the “right” equilibrium beliefs are fairly intuitive (e.g., simple averages of the data). Our framework, in contrast, provides a characterization of equilibrium beliefs essentially for all types of misspecified models, and, as some of the examples illustrate, this characterization is not necessarily straightforward in more complex settings.

Example 2.1. *Monopolist with unknown demand.* A monopolist faces demand function $y = f(x, \omega)$, where $x \in \mathbb{X}$ is the price chosen by the monopolist and ω is a random shock, distributed according to $p \in \Delta(\Omega)$. The monopolist observes sales y , but not the random shock. The monopolist’s payoff function is $\pi(x, y) = xy$ (i.e., there are no costs). The monopolist does not know the true demand function f . It believes that demand is

$$y = f_{\theta}(x, \omega) = a - bx + \omega, \tag{2}$$

⁸As usual, the superscript $-i$ denotes a profile where the i ’th component is excluded

where $\theta = (a, b) \in \Theta$ is a parameter vector and ω follows a standard normal distribution. Then $Q_\theta(\cdot | x)$ is a normal distribution with mean $a - bx$ and unit variance. \square

Example 2.2. *Non-linear pricing.* An agent chooses effort $x \in \mathbb{X}$ and then income $z = x + \omega$ is realized, where ω is a random shock with zero mean. The agent pays taxes $t = \tau(z)$, where $\tau(\cdot)$ is a nonlinear tax schedule. The agent observes $y = (z, t)$ and obtains payoff $\pi(x, z, t) = z - t - c(x)$, where $c(x)$ is the cost of effort x . The agent correctly understands how effort translates into income. She believes, however, that she faces a linear tax schedule: $t = \theta z + \varepsilon$, where $\theta \in \Theta$ is the constant tax rate and ε follows a standard normal distribution and measures uncertain aspects of the schedule (such as eligible deductions). This model captures an agent who fails to respond to the *marginal* tax rate. \square

Example 2.3. *Misspecified market structure.* Two firms compete by simultaneously choosing prices. A state $\omega = (s, \varepsilon)$ is a vector of costs $s = (s^1, s^2)$ and demand shocks $\varepsilon = (\varepsilon^1, \varepsilon^2)$. After a state is drawn, each firm i privately observes her own cost s^i and chooses a price x^i . The quantity sold by firm i is given by the demand system

$$\ln q^i = \alpha^* + \beta^* \ln x^j - \gamma^* \ln x^i + \varepsilon^i,$$

where $\gamma^* > 1$ is the demand elasticity (in absolute value). A consequence for firm i is $y^i = (s^i, q^i)$. The payoff of firm i is $\pi^i(x^i, y^i) = (x^i - s^i)q^i$. Each firm $i = 1, 2$ (incorrectly) believes that they are a monopolist in this market and that the demand function they face is

$$\ln q^i = \alpha^i - \gamma^i \ln x^i + \varepsilon^i, \tag{3}$$

where ε^i follows a standard normal distribution and $\theta^i = (\alpha^i, \gamma^i)$ parameterizes the subjective model. \square

Example 2.4. *Regression to the mean.* An instructor observes the initial performance s of a student and decides whether to praise or criticize him, $x \in \{C, P\}$. The student then performs again and the instructor observes his final performance, s' . The truth is that performances $y = (s, s')$ are independent, standard normal random variables. The instructor's payoff is $\pi(x, s, s') = s' - c(x, s)$, where $c(x, s) = \kappa |s| > 0$ if either $s > 0, x = C$ or $s < 0, x = P$, and, in all other cases, $c(x, s) = 0$.⁹

⁹Formally, a state is $\omega = (s, s')$ and the feedback function is $y = f(x, \omega) = \omega$.

The interpretation is that the instructor bears a (reputation) cost from lying that is increasing in the size of the lie, where lying is defined as either criticizing an above-average performance or praising a below-average performance. Because the instructor has no influence on performance, it is optimal to praise if $s > 0$ and to criticize if $s < 0$. The instructor, however, does not admit the possibility of regression to the mean and believes that

$$s' = s + \theta_x + \varepsilon,$$

where ε has a standard normal distribution, and $\theta = (\theta_C, \theta_P) \in \Theta$ parameterizes her perceived influence on the performance of the student.¹⁰ \square

Example 2.5. Classical monetary policy. There are two players, the government (G) and the public (P). The government chooses monetary policy x^G and the public chooses inflation forecasts x^P . Inflation, e , and unemployment, U , are determined as follows:¹¹

$$\begin{aligned} e &= x^G + \varepsilon_e \\ U &= u^* - \lambda(e - x^P) + \varepsilon_U, \end{aligned} \tag{4}$$

where $\lambda \in (0, 1)$ and where $\omega = (\varepsilon_e, \varepsilon_U)$ are random shocks with a full support distribution and $Var(\varepsilon_e) > 0$. Thus, inflation is determined by the government's action and a random term. And unemployment is determined by surprise inflation according to a Phillips curve, where $u^* > 0$ is the natural rate of unemployment. Realized inflation and unemployment, but not the error terms, are observed by both the public and the government. The government's payoff is $\pi(x^G, e, U) = -(U^2 + e^2)$. For simplicity, we focus on the government's problem and assume that the public has correct beliefs and chooses $x^P = x^G$. Under the *classical subjective model*, the government believes (correctly) that its policy x^G affects inflation, but it does not realize that unemployment is affected by *surprise* inflation:

$$\begin{aligned} e &= x^G + \varepsilon_e \\ U &= \theta_1 - \theta_2 e + \varepsilon_U. \end{aligned}$$

¹⁰A model that allows for regression to the mean is $s' = \alpha s + \theta_x + \varepsilon$; in this case, the agent would correctly learn that $\alpha = 0$ and $\theta_x = 0$ for all x .

¹¹Formally, a state is $\omega = (\varepsilon_e, \varepsilon_U)$ and the feedback function $y = (e, U) = f(x^G, x^P, \omega)$ is given by the system of equations (4).

The subjective model is parameterized by $\theta = (\theta_1, \theta_2)$. \square

Example 2.6. Trade with adverse selection. A buyer and a seller simultaneously submit a (bid) price $x \in \mathbb{X}$ and an ask price $a \in \mathbb{A}$, respectively. If $a \leq x$, then trade takes place at price x , and the buyer obtains payoff $v - x$, where v is the buyer’s value of the object. If $a > x$, then no trade takes place and each player receives 0. At the time she makes an offer, the buyer does not know her value or the ask price of the seller. The seller’s ask price and the buyer’s value are drawn from $p \in \Delta(\mathbb{A} \times \mathbb{V})$.¹²

We consider two different feedback functions to illustrate the importance of making explicit what players observe about the outcome of the game. Under *perfect feedback*, the buyer observes the realized ask price and her own value. Under *partial feedback*, the buyer observes the ask price, but she only observes her own value if she trades. Finally, suppose that A and V are actually correlated but that the buyer naively believes that they are independent. This is formally captured by letting $Q_\theta = \theta$ and $\Theta = \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$. \square

References for examples. *Example 2.1:* Nyarko [1991] studies a special case and shows that a steady state does not exist, although he does not allow for mixed strategies. *Example 2.2:* Sobel [1984] considers a similar misspecification to capture a consumer who responds to average, not marginal, pricing. *Example 2.3:* Arrow and Green [1973] and Kirman [1975] study other examples in which firms have incorrect perceptions about market structure. *Example 2.4:* The story of the instructor who does not understand regression to the mean is taken from Tversky and Kahneman [1973]; we are not aware of previous attempts to formalize the underlying misspecification. *Example 2.5:* The monetary policy example is based on Sargent (1999, Chapter 7). *Example 2.6:* This misspecification is first discussed in the lemons context by Kagel and Levin [1986]. It has been generalized by Eyster and Rabin [2005], Jehiel and Koessler [2008], and Esponda [2008].

2.3 Definition of equilibrium

DISTANCE TO TRUE MODEL. In equilibrium, we will require players’ beliefs to put probability one on the set of subjective distributions over consequences that are “clos-

¹²The typical story is that there is a population of sellers each of whom follows the weakly dominant strategy of asking for her valuation; thus, the ask price is a function of the seller’s valuation and, if buyer and seller valuations are correlated, then the ask price and buyer valuation are also correlated.

est” to the objective distribution. In order to describe the right notion of “closest”, we need some additional definitions. The following function, which we call the **weighted Kullback-Leibler divergence** (wKLD) function of player i , is a weighted version of the standard Kullback-Leibler divergence in statistics (Kullback and Leibler, 1951). It represents a non-symmetric measure between the objective distribution over i 's consequences given a strategy profile $\sigma \in \Sigma$ and the distribution as parameterized by $\theta^i \in \Theta^i$:¹³

$$K^i(\sigma, \theta^i) = \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_{\sigma^i}(\cdot | s^i, x^i)} \left[\ln \frac{Q_{\sigma^i}^i(Y^i | s^i, x^i)}{Q_{\theta^i}^i(Y^i | s^i, x^i)} \right] \sigma^i(x^i | s^i) p_{S^i}(s^i). \quad (5)$$

The set of closest parameter values of player i given σ is the set

$$\Theta^i(\sigma) \equiv \arg \min_{\theta^i \in \Theta^i} K^i(\sigma, \theta^i).$$

The interpretation is that $\Theta^i(\sigma) \subset \Theta^i$ is the set of parameter values that player i can believe to be possible after observing feedback consistent with strategy profile σ .

Remark 1. (a) The use of the Kullback-Leibler divergence to measure distance is not an arbitrary assumption. We show in Section 4 that this is the right notion of distance in a learning model with Bayesian players. (b) Because the wKLD function is weighted by a player's own strategy, it will place no restrictions on beliefs about outcomes that only arise following out-of-equilibrium actions (beyond the restrictions imposed by Θ).

The next result collects some useful properties of the wKLD. We cannot directly apply the Theorem of the Maximum to prove this result because the wKLD function K^i may take infinite values. Instead, the proof relies crucially on Assumption 1 to obtain upper hemicontinuity of $\Theta^i(\cdot)$.

Lemma 1. (i) For all $\sigma \in \Sigma$, $\theta^i \in \Theta^i$, and $i \in I$, $K^i(\sigma, \theta^i) \geq 0$, with equality holding if and only if $Q_{\theta^i}^i(\cdot | s^i, x^i) = Q_{\sigma^i}^i(\cdot | s^i, x^i)$ for all (s^i, x^i) such that $\sigma^i(x^i | s^i) > 0$. (ii) For every $i \in I$, $\Theta^i(\cdot)$ is non-empty, upper hemicontinuous, and compact valued.

Proof. See the Appendix. □

¹³The notation E_Q denotes expectation with respect to the probability measure Q . Also, we use the convention that $-\ln 0 = \infty$ and $0 \ln 0 = 0$.

OPTIMALITY. In equilibrium, we will require each player to choose a strategy that is optimal given her beliefs. A strategy σ^i for player i is **optimal** given $\mu^i \in \Delta(\Theta^i)$ if $\sigma^i(x^i | s^i) > 0$ implies that

$$x^i \in \arg \max_{\bar{x}^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}^i(\cdot | s^i, \bar{x}^i)} [\pi^i(\bar{x}^i, Y^i)] \quad (6)$$

where, for all i and (s^i, x^i) ,

$$\bar{Q}_{\mu^i}^i(\cdot | s^i, x^i) = \int_{\Theta^i} Q_{\theta^i}^i(\cdot | s^i, x^i) \mu^i(d\theta^i)$$

is the distribution over consequences of player i induced by μ^i .¹⁴

DEFINITION OF EQUILIBRIUM. We propose the following solution concept.

Definition 1. A strategy profile σ is a **Berk-Nash equilibrium** of game \mathcal{G} if, for all players $i \in I$, there exists $\mu^i \in \Delta(\Theta^i)$ such that

- (i) σ^i is optimal given μ^i , and
- (ii) $\mu^i \in \Delta(\Theta^i(\sigma))$, i.e., if $\hat{\theta}^i$ is in the support of μ^i , then

$$\hat{\theta}^i \in \arg \min_{\theta^i \in \Theta^i} K^i(\sigma, \theta^i).$$

Definition 1 places two types of restrictions on equilibrium behavior: (i) optimization given beliefs, and (ii) endogenous restrictions on beliefs. For comparison, notice that the definition of a Nash equilibrium is identical to Definition 1 except that condition (ii) is replaced with the condition that $\bar{Q}_{\mu^i}^i = Q_{\sigma}^i$; in other words, players must have correct beliefs in a Nash equilibrium.

EXISTENCE OF EQUILIBRIUM. The standard proof of existence of Nash equilibrium cannot be used to show existence of a Berk-Nash equilibrium because the corresponding version of a best response correspondence is not necessarily convex valued. To prove existence, we first perturb the payoffs of the game and establish that equilibrium exists in the perturbed game. We then consider a sequence of equilibria of

¹⁴Note that, even in the case where $Q_{\theta^i}^i$ is derived by replacing subjective primitives $p_{\theta^i}^i$, $f_{\theta^i}^i$, and $\sigma_{\theta^i}^{-i}$ in equation (1), so that player i correctly believes that every other player j mixes independently, we still allow player i to have correlated beliefs about her opponents' strategies, as in Fudenberg and Levine [1993a].

perturbed games where the perturbations go to zero and establish that the limit is a Berk-Nash equilibrium of the (unperturbed) game.¹⁵ The part of the proof that is not standard is establishing existence of equilibrium in the perturbed game. The perturbed best response correspondence is still not necessarily convex valued. Our approach is to characterize equilibrium as a fixed point of a *belief correspondence* and show that this correspondence satisfies the requirements of a generalized version of Kakutani's fixed point theorem.

Theorem 1. *Every game has at least one Berk-Nash equilibrium.*

Proof. See the Appendix. □

2.4 Examples: Finding a Berk-Nash equilibrium

We illustrate the definition by finding Berk-Nash equilibria for some of the examples in Section 2.2.

Example 2.1, continued from pg. 7. *Monopolist with unknown demand.*

Suppose that the monopolist can choose a price of 2 or 10, i.e., $\mathbb{X} = \{2, 10\}$ and that the true demand function is given by equation (2), with $\theta^0 = (a^0, b^0) = (40, 5)$. Figure 1 shows the true demand parameter value and the set of parameter values $\Theta = [12, 32] \times [1, 3]$ that the monopolist considers possible. In particular, $\theta^0 \notin \Theta$ and, therefore, we say that the monopolist has a *misspecified* model. The dashed line in Figure 1 depicts optimal behavior: price 10 is uniquely optimal to the left, price 2 is uniquely optimal to the right, and the monopolist is indifferent for parameter values on the dashed line.

Let $\sigma = (\sigma(2), \sigma(10))$ denote the monopolist's strategy. Because this is a single-agent problem, the objective distribution does not depend on σ ; hence, we denote it by $Q_{\theta^0}(\cdot | x)$ and note that it is a normal distribution with mean $\mu_{\theta^0}(x) = a^0 - b^0x$ and unit variance. Similarly, for $\theta = (a, b)$, $Q_{\theta}(\cdot | x)$ is a normal distribution with mean $\mu_{\theta}(x)$ and unit variance. It follows from equation (5) that

¹⁵The idea of perturbations and the strategy of the existence proof date back to Harsanyi [1973]; Selten [1975] and Kreps and Wilson [1982] also used these ideas to prove existence of perfect and sequential equilibrium, respectively.

$$\begin{aligned}
K(\sigma, \theta) &= \sum_{x \in \{2, 10\}} \sigma(x) \frac{1}{2} E_{Q_{\theta^0}(\cdot|x)} [(Y - \mu_\theta(x))^2 - (Y - \mu_{\theta^0}(x))^2] \\
&= \sum_{x \in \{2, 10\}} \sigma(x) \frac{1}{2} (\mu_{\theta^0}(x) - \mu_\theta(x))^2.
\end{aligned} \tag{7}$$

The derivatives are

$$\frac{\partial K(\sigma, \theta)}{\partial a} = \sum_{x \in \{2, 10\}} \sigma(x) (\mu_{\theta^0}(x) - \mu_\theta(x)) \tag{8}$$

$$\frac{\partial K(\sigma, \theta)}{\partial b} = \sum_{x \in \{2, 10\}} \sigma(x) (\mu_{\theta^0}(x) - \mu_\theta(x)) x. \tag{9}$$

First, consider pure strategies. If $\sigma = (0, 1)$ (i.e., the price is $x = 10$), the first order condition becomes $\mu_{\theta^0}(10) = \mu_\theta(10)$. Any parameter value $\theta \in \Theta$ such that the mean of Y under θ and $x = 10$ is equal to the true mean under $x = 10$ minimizes equation (7). In the left panel of Figure 1, these are the parameter values that lie on the line with slope 10 ($a^0 - b^0 10 = a - b 10$) and belong to Θ ; i.e., the line segment AB . Thus, if the monopolist were to choose $\sigma = (0, 1)$ in equilibrium, then its beliefs must have support in the segment AB . But this segment lies to the right of the dashed line, where it is *not* optimal to set a price of 10. Thus, $\sigma = (0, 1)$ is not an equilibrium. A similar argument establishes that $\sigma = (1, 0)$ is not an equilibrium: If it were, then the monopolist would believe that the parameter value is D . But D lies to the left of the dashed line, where it is not optimal to set a price of 2.

Finally, consider a totally mixed strategy σ . Because expressions (8) and (9) cannot be simultaneously equal to zero, the parameter value that minimizes $K(\sigma, \theta)$ lies on the boundary of Θ . In fact, a bit of algebra shows that, for each totally mixed σ , there is a unique minimizer $\theta_\sigma = (a_\sigma, b_\sigma)$ characterized as follows. If $\sigma(2) \leq 3/4$, the minimizer is in the segment BC : $b_\sigma = 3$ and $a_\sigma = 16\sigma(2) + 20$ is such that equation (8) is zero. The left panel of Figure 1 depicts an example where the unique minimizer $\theta_{\hat{\sigma}}$ under strategy $\hat{\sigma}$ is given by the tangency between the contour lines of $K(\hat{\sigma}, \cdot)$ and the feasible set Θ .¹⁶ If $\sigma(2) \in [3/4, 15/16]$, then $\theta_\sigma = C$ is the northeast vertex of Θ . Finally, if $\sigma(2) > 15/16$, the minimizer is in the segment CD : $a_\sigma = 32$ and

¹⁶After some algebra, it can be shown that $K(\sigma, \theta) = (\theta - \theta^0)' M_\sigma (\theta - \theta^0)$, where M_σ is a weighting matrix that depends on σ . In particular, the contour lines of $K(\sigma, \cdot)$ are ellipses.

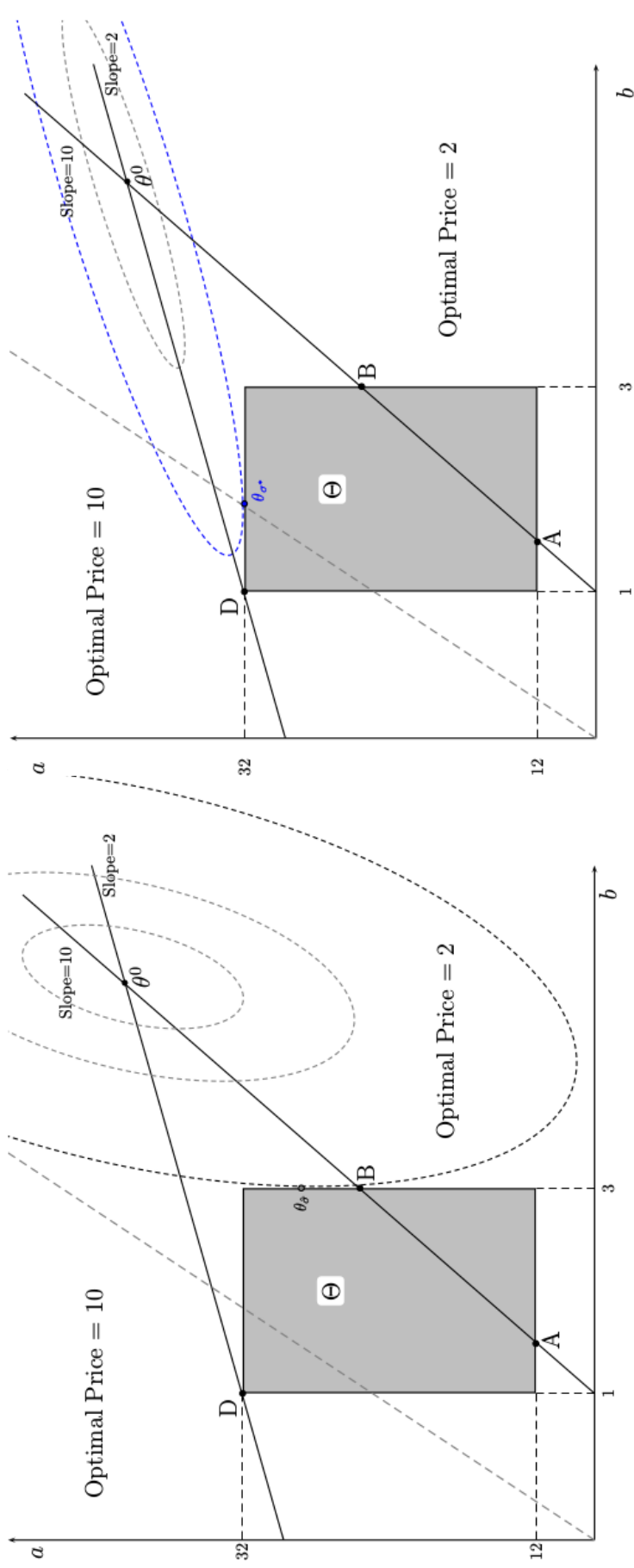


Figure 1: Monopolist with misspecified demand function. In the left panel, the parameter value that minimizes wKL divergence given strategy $\hat{\sigma}$ is θ_{σ^*} . The right panel illustrates the unique Berk-Nash equilibrium strategy: σ^* is optimal given θ_{σ^*} (because θ_{σ^*} lies on the indifference line), and θ_{σ^*} minimizes the wKL divergence given σ^* .

$b_\sigma = (420 - 416\sigma(2))/(100 - 96\sigma(2))$ is such that equation (9) is zero.

Because the monopolist is mixing, optimality requires that the equilibrium belief θ_σ lies on the dashed line. The unique Berk-Nash equilibrium is $\sigma^* = (23/24, 1/24)$, and its supporting belief, $\theta_{\sigma^*} = (32, 8/3)$, is given by the intersection of the dashed line and line segment CD , as depicted in the right panel of Figure 1. Note that, in this example, it is *not* the case that the equilibrium belief about the mean of Y is correct. Thus, an approach that had focused on fitting the mean, rather than minimizing K , would have led to the wrong conclusion.¹⁷ \square

Example 2.2, continued from pg. 8. *Non-linear pricing.* For any pure strategy x and parameter value θ ,

$$\begin{aligned} K(x, \theta) &= E \left[\ln \frac{Q(T | Z)p(Z | x)}{Q_\theta(T | Z)p(Z | x)} \mid X = x \right] \\ &= -\frac{1}{2} E [(\tau(Z) - \theta Z)^2 \mid X = x] + C, \end{aligned}$$

where E denotes the true conditional expectation and C is a constant that does not depend on θ . It is straightforward to check that $\theta(x) = E[(x + W)\tau(x + W)] / E[(x + W)^2]$ is the unique parameter value that minimizes $K(x, \theta)$. In particular, the intuitive conjecture that the agent believes that the constant marginal tax rate is given by the expected average tax rate, $E[\tau(x + W)/(x + W)]$, is incorrect in this problem.¹⁸

A strategy x^* is a Berk-Nash equilibrium if and only if $x = x^*$ maximizes $(1 - \theta(x^*))x - c(x)$. In contrast, the optimal strategy x^{opt} maximizes $x - E[\tau(x + W)] - c(x)$. For example, suppose that $\tau(z) = z^2$ is progressive, $c(\cdot)$ is increasing and strictly convex, and ω follows a normal distribution with zero mean and variance σ^2 . Then there is a unique Berk-Nash equilibrium and unique optimal strategy. Moreover, there is a threshold variance $\bar{\sigma}^2$ of the error term such that the Berk-Nash equilibrium strategy is lower than the optimal one for all higher variances and higher than optimal for all lower variances. Thus, despite a progressive tax schedule, naiveté can either decrease or increase the incentives to put effort. \square

Example 2.2, continued from pg. 8. *Misspecified market structure.* For

¹⁷The example also illustrates the importance of allowing for mixed strategies for existence of Berk-Nash equilibrium, even in single-agent settings. As an antecedent, Esponda and Pouzo [2011] argue that this is the reason why mixed strategy equilibrium cannot be purified in a voting application.

¹⁸This intuitive conjecture would be correct, however, if the agent had the following “random-coefficient” misspecified subjective model: $\tau = (\theta + \varepsilon)z$.

convenience, we focus on the case where prices and costs are chosen from a compact interval and players follow pure strategies, where $\sigma^i(s^i)$ is the price chosen by firm i given cost s^i . Optimality implies that each firm i follows strategy

$$\sigma^i(s^i) = [\gamma^i / (\gamma^i - 1)] s^i. \quad (10)$$

Since the error term is normally distributed, the minimizer of the wKLD function is given by the ordinary least squares estimand of equation (3). Thus, for all $\sigma = (\sigma^1, \sigma^2)$,

$$\begin{aligned} \gamma^i(\sigma) &= - \frac{Cov(\ln \sigma^i(S^i), \alpha^* + \beta^* \ln \sigma^j(S^j) - \gamma^* \ln \sigma^i(S^i) + \varepsilon^i)}{Var(\ln \sigma^i(S^i))} \\ &= \gamma^* - \beta^* \frac{Cov(\ln \sigma^i(S^i), \ln \sigma^j(S^j))}{Var(\ln \sigma^i(S^i))}. \end{aligned} \quad (11)$$

By replacing (10) into (11), the equilibrium belief is given by

$$\gamma^i = \gamma^* - \beta^* \frac{Cov(\ln S^i, \ln S^j)}{Var(\ln S^i)}, \quad (12)$$

and the unique equilibrium strategy is obtained by replacing (12) into (10).¹⁹ Moreover, (12) shows that firms estimate demand elasticity with a bias that depends on the sign of $\beta^* Cov(\ln S^i, \ln S^j)$. For example, suppose that $\beta^* > 0$, so that the products are substitutes, and that $Cov(\ln S^i, \ln S^j) > 0$. Then firms believe that demand is less elastic compared to the true elasticity. The intuition is that, when a firm chooses a higher price, it is because its costs are higher. But then the competitor's cost is also likely to be higher, so the other firm is also likely to choose a higher price. Because products are substitutes, the increase in the price of the other firm mitigates the fall in demand due to the increase in own price. This under-estimation of elasticity leads firms to set higher prices compared to the Nash equilibrium, which is unique and given by the dominant strategy $\sigma^i(s^i) = [\gamma^* / (\gamma^* - 1)] s^i$. Note that disregarding the competition leads to biased behavior in the Berk-Nash equilibrium, even though actions are strategically independent and the (correct) best response does not depend on the competitor's choice. \square

Example 2.4, continued from pg. 8. *Regression to the mean.* It is straightforward to check that optimal strategies are characterized by a cutoff: Let

¹⁹This is true as long as we make an assumption on the primitives that makes $\gamma^i > 1$.

$\sigma \in \mathbb{R}$ represent the strategy where the instructor praises an initial performance if it is above σ and criticizes it otherwise. The wKLD function is

$$K(\sigma, \theta) = \int_{-\infty}^{\sigma} E \left[\ln \frac{\varphi(S_2)}{\varphi(S_2 - (\theta_C + s_1))} \right] \varphi(s_1) ds_1 + \int_{\sigma}^{\infty} E \left[\ln \frac{\varphi(S_2)}{\varphi(S_2 - (\theta_P + s_1))} \right] \varphi(s_1) ds_1,$$

where φ is the density of the standard normal distribution and expectations are with respect to the true distribution. It is straightforward to show that, for each σ , the unique parameter vector that minimizes $K(\sigma, \cdot)$ is

$$\begin{aligned} \theta_C(\sigma) &= E[S_2 - S_1 \mid S_1 < \sigma] \\ &= 0 - E[S_1 \mid S_1 < \sigma] > 0 \end{aligned}$$

and, similarly, $\theta_P(\sigma) = 0 - E[S_1 \mid S_1 > \sigma] < 0$. The intuition is that the instructor is critical for performances below a threshold and, therefore, the mean performance conditional on a student being criticized is lower than the unconditional mean performance; thus, a student who is criticized delivers a better next performance in expectation. Similarly, a student who is praised delivers a worse next performance in expectation.

By following a strategy cutoff σ , the instructor believes, after observing initial performance $s_1 > 0$, that her expected payoff is $s_1 + \theta_C(\sigma) - \kappa s_1$ if she criticizes and $s_1 + \theta_P(\sigma)$ if she praises. By optimality, the cutoff makes her indifferent between praising and criticizing. Thus,

$$\sigma^* = \frac{1}{\kappa} (\theta_C(\sigma^*) - \theta_P(\sigma^*)) > 0$$

is the unique equilibrium cutoff. An instructor who ignores regression to the mean will have incorrect beliefs about the influence of her feedback on the student's performance; in particular, she is excessively critical in equilibrium because she incorrectly believes that criticizing a student improves performance and that praising a student worsens it. Moreover, as the reputation cost $\kappa \rightarrow 0$, meaning that instructors care only about performance and not about lying, $\sigma^* \rightarrow \infty$: instructors only criticize (as in Kahneman and Tversky's (1973) story). \square

3 Relationship to other solution concepts

We show that Berk-Nash equilibrium includes several solution concepts (both standard and boundedly rational) as special cases.

3.1 Properties of games

CORRECTLY-SPECIFIED GAMES. In Bayesian statistics, a model is correctly specified if the support of the prior includes the true data generating process. The extension to single-agent decision problems is straightforward. For example, the monopoly problem in Example 2.1 is not correctly specified because $\theta^0 \notin \Theta$. In games, however, we must account for the fact that the objective distribution over consequences (i.e., the true model) depends on the strategy profile. We say that a game is correctly specified if players do not a priori rule out any of the objective distributions that might possibly arise in equilibrium.²⁰

Definition 2. A game is **correctly specified given σ** if, for all $i \in I$, there exists $\theta^i \in \Theta^i$ such that $Q_{\theta^i}^i(\cdot \mid s^i, x^i) = Q_{\sigma}^i(\cdot \mid s^i, x^i)$ for all for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$; otherwise, the game is **misspecified given σ** . A game is **correctly specified** if it is correctly specified for all σ ; otherwise, it is **misspecified**.

IDENTIFICATION. From the point of view of the agent, what matters is identification of the distribution over consequences Q_{θ} , not the parameter θ . If the model is correctly specified, then the true Q_{θ^0} is trivially identified.²¹ Of course, this is not true if the model is misspecified, because the true distribution will never be learned. But we want a definition that captures the same spirit: If two distributions are judged to be equally a best fit (given the true distribution), then we want these two distributions to be identical; otherwise, we cannot identify which distribution is a best fit.

The fact that our agents also take actions introduces an additional nuance to the definition of identification. We can either ask for identification of the distribution over

²⁰A more precise terminology would say that a game is correctly specified *in steady state*. The reason is that, in the dynamic model, players believe that they face a stationary environment while in fact the environment may be non-stationary outside the steady state.

²¹Similarly, the true parameter θ^0 is identified if we impose the condition of identifiability of the parameter, i.e., $Q_{\theta} = Q_{\theta^0}$ implies that $\theta = \theta^0$. But, as mentioned earlier, it suffices to consider identification of distributions, not the parameter.

consequences for those actions that are taken by the agent (i.e., on the path of play) or for all actions (i.e., on and off the path). We refer to the former as weak identification (because it highlights the *potential* to identify the distribution over consequences, if *all* actions are taken) and to the latter as strong identification.

Definition 3. A game is **weakly identified given σ** if, for all $i \in I$: if $\theta_1^i, \theta_2^i \in \Theta^i(\sigma)$, then $Q_{\theta_1^i}^i(\cdot | s^i, x^i) = Q_{\theta_2^i}^i(\cdot | s^i, x^i)$ for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$ such that $\sigma^i(x^i | s^i) > 0$; if the condition is satisfied for *all* $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, then we say that the game is **strongly identified given σ** . A game is **[weakly or strongly] identified** if it is [weakly or strongly] identified for all σ .

Note that whether or not a game is correctly specified or identified depends on the primitives of the game, which includes the feedback or consequence functions. In particular, two games that are identical except for the feedback functions may differ in terms of being correctly specified and identified.

It is straightforward to show that a correctly specified game is weakly identified. We show in Online Appendix A that a correctly specified game is strongly identified if it also satisfies the property that consequences are perceived to be independent of own actions.

3.2 Relationship to Nash and self-confirming equilibrium

The next result shows that Berk-Nash equilibrium is equivalent to Nash equilibrium when the game is both correctly specified and strongly identified.

Proposition 1. (i) *Suppose that the game is correctly specified given σ and that σ is a Nash equilibrium of its objective game. Then σ is a Berk-Nash equilibrium of the (objective and subjective) game.*

(ii) *Suppose that σ is a Berk-Nash equilibrium of a game that is correctly specified and strongly identified given σ . Then σ is a Nash equilibrium of the corresponding objective game.*

Proof. (i) Let σ be a Nash equilibrium and fix any $i \in I$. Then σ^i is optimal given Q_σ^i . Because the game is correctly specified given σ , there exists $\theta_*^i \in \Theta^i$ such that $Q_{\theta_*^i}^i = Q_\sigma^i$ and, therefore, by Lemma 1, $\theta_*^i \in \Theta^i(\sigma)$. Thus, σ^i is also optimal given

$Q_{\theta_*^i}^i$ and $\theta_*^i \in \Theta^i(\sigma)$, so that σ is a Berk-Nash equilibrium. (ii) Let σ be a Berk-Nash equilibrium and fix any $i \in I$. Then σ^i is optimal given $\bar{Q}_{\mu^i}^i$, for some $\mu^i \in \Delta(\Theta^i(\sigma))$. Because the game is correctly specified given σ , there exists $\theta_*^i \in \Theta^i$ such that $Q_{\theta_*^i}^i = Q_\sigma^i$ and, therefore, by Lemma 1, $\theta_*^i \in \Theta^i(\sigma)$. Moreover, because the game is strongly identified given σ , any $\hat{\theta}^i \in \Theta^i(\sigma)$ satisfies $Q_{\hat{\theta}^i}^i = Q_{\theta_*^i}^i = Q_\sigma^i$. Then σ^i is also optimal given Q_σ^i . Thus, σ is a Nash equilibrium. \square

The next example illustrates that Proposition 1 can be used to conclude that the equilibrium of an incorrect model is Nash.

Example 2.5, continued from pg. 9. Classical monetary policy. We show that, despite appearances, the game is correctly specified. Fix a strategy x_*^P for the public. Note that $U = u^* - \lambda(x^G - x_*^P + \varepsilon_e) + \varepsilon_U$, whereas the government believes that $U = \theta_1 - \theta_2(x^G + \varepsilon_e) + \varepsilon_U$. Thus, by choosing $\theta^* = (\theta_1^*, \theta_2^*)$ such that $\theta_1^* = u^* + \lambda x_*^P$ and $\theta_2^* = \lambda$, it follows that the distribution over $Y = (U, e)$ parameterized by θ^* coincides with the objective distribution given x_*^P . Moreover, the model is strongly identified: since $Var(\varepsilon_e) > 0$, it is easy to check that θ^* is the unique minimizer of the wKL divergence. Thus, Proposition 1 implies that Berk-Nash equilibrium is equivalent to Nash equilibrium. In particular, the equilibrium policies are the same whether or not the government realizes that unemployment is driven by surprise, not actual, inflation.²² \square

The next result shows that a Berk-Nash equilibrium is also a self-confirming equilibrium (SCE) in games that are correctly specified, but not necessarily strongly identified. Recall that, in a SCE, beliefs must be correct “on the equilibrium path” but may be incorrect off equilibrium.²³

Proposition 2. *Suppose that the game is correctly specified given σ , and that σ is a Berk-Nash equilibrium. Then σ is also a self-confirming equilibrium (SCE).²⁴*

²²Sargent [1999] derived this result for a government that does OLS-based learning, which is a special case of our example when errors are normally distributed.

²³A strategy profile σ is a SCE if, for all players $i \in I$, σ^i is optimal given \hat{Q}_σ^i , where $\hat{Q}_\sigma^i(\cdot | s^i, x^i) = Q_\sigma^i(\cdot | s^i, x^i)$ for all (s^i, x^i) such that $\sigma^i(x^i | s^i) > 0$. This definition is slightly more general than the typical one, e.g., Dekel et al. [2004], because it does not restrict players to believe that consequences are driven by other players’ strategies.

²⁴A converse of Proposition 2 holds if the set \mathcal{Q} of possible views of the world imposes no a priori restrictions on off equilibrium beliefs.

Proof. Fix any $i \in I$ and let $\hat{\theta}^i$ be any element in the support of μ^i , where μ^i is player i 's belief supporting the choice of Berk-Nash equilibrium strategy σ^i . Because the game is correctly specified given σ , there exists $\theta_*^i \in \Theta^i$ such that $Q_{\theta_*^i}^i = Q_\sigma^i$ and, therefore, by Lemma 1, $K^i(\sigma, \theta_*^i) = 0$. Thus, it must also be that $K^i(\sigma, \hat{\theta}^i) = 0$. By Lemma 1, it follows that $Q_{\hat{\theta}^i}(\cdot | s^i, x^i) = Q_\sigma^i(\cdot | s^i, x^i)$ for all (s^i, x^i) such that $\sigma^i(x^i | s^i) > 0$. In particular, σ^i is optimal given $Q_{\hat{\theta}^i}$, and $Q_{\hat{\theta}^i}$ satisfies the desired self-confirming restriction. \square

The novelty in our paper is to consider games that are *not* correctly specified. For such games, as illustrated by the examples in the paper, a Berk-Nash equilibrium is not necessarily a Nash or self-confirming equilibrium. The reason is that, in misspecified games, beliefs can be incorrect even on the equilibrium path.

3.3 Relationship to fully cursed and ABEE

An **analogy-based game** satisfies the following four properties:

(i) The information structure is finite and partitional: The state space Ω has a finite number of elements; we denote the true distribution over Ω by p_Ω . In addition, for each i , there is a partition \mathcal{S}^i of Ω , and the element of \mathcal{S}^i that contains ω (i.e., the signal of player i in state ω) is denoted by $s^i(\omega)$.²⁵

(ii) For each i , $f^i(x, \omega) = (x^{-i}, \omega)$ for all (x, ω) , i.e., each player gets perfect feedback about (x^{-i}, ω) .

(iii) For each i , there exists a partition of Ω , denoted by \mathcal{A}^i , and the element of \mathcal{A}^i that contains ω is denoted by $\alpha^i(\omega)$.

(iv) $(Q_{\theta^i}^i)_{\theta^i \in \Theta^i}$ is the set of all joint probability distributions over $\mathbb{X}^{-i} \times \Omega$ that satisfy

$$Q_{\theta^i}^i(x^{-i}, \omega | s^i(\omega'), x^i) = Q_{\Omega, \theta^i}^i(\omega | s^i(\omega')) Q_{\mathbb{X}^{-i}, \theta^i}^i(x^{-i} | \alpha^i(\omega)).$$

In other words, every player i believes that x^{-i} and ω are independent conditional on the partition, i.e., $Q_{\mathbb{X}^{-i}, \theta^i}^i(x^{-i} | \omega) = Q_{\mathbb{X}^{-i}, \theta^i}^i(x^{-i} | \alpha^i(\omega))$ for all $\omega \in \Omega$.

Two special cases are noteworthy. If $\mathcal{A}^i = \{\Omega\}$ for all i , then each player believes that the actions of other players are independent of the state of the world. If $\mathcal{A}^i = \mathcal{S}^i$

²⁵This assumption is made only to facilitate comparison to the original definition of an analogy-based expectation equilibrium (Jehiel [2005], Jehiel and Koessler [2008]).

for all i , then each player believes that the actions of other players are independent of the state, conditional on their own private information.

Definition 4. (Jehiel and Koessler, 2008) A strategy profile σ is an analogy-based expectation equilibrium (ABEE) if for all $i \in I$, $\omega \in \Omega$, and x^i such that $\sigma^i(x^i | s^i(\omega)) > 0$,

$$x^i \in \arg \max_{\bar{x}^i \in \mathbb{X}^i} \sum_{\omega' \in \Omega} p_{\Omega}(\omega' | s^i(\omega)) \sum_{x^{-i} \in \mathbb{X}^{-i}} \bar{\sigma}^{-i}(x^{-i} | \omega') \pi^i(\bar{x}^i, x^{-i}, \omega'),$$

where $\bar{\sigma}^{-i}(x^{-i} | \omega') = \sum_{\omega'' \in \Omega} p_{\Omega}(\omega'' | \alpha^i(\omega')) \prod_{j \neq i} \sigma^j(x^j | s^j(\omega''))$.

Proposition 3. *In an analogy-based game, σ is a Berk-Nash equilibrium if and only if it is an ABEE.*

Proof. See the Appendix. □

Proposition 3 shows the equivalence of Berk-Nash and ABEE for games that satisfy the four properties above. As mentioned by Jehiel and Koessler [2008], ABEE is equivalent to Eyster and Rabin's (2005) fully cursed equilibrium in the special case where $\mathcal{A}^i = \mathcal{I}^i$ for all i .²⁶ In particular, Proposition 3 provides a misspecified-learning foundation for these solution concepts. Jehiel and Koessler [2008] discuss an alternative foundation for ABEE, where players receive feedback that is coarse and aggregated over past play, and, therefore, multiple beliefs are consistent with this feedback. Under this different feedback structure, ABEE can be viewed as a natural selection of the set of self-confirming equilibrium.

Example 2.6, continued from pg. 10. Trade with adverse selection. As a benchmark, the Nash equilibrium (NE) price maximizes

$$\Pi^{NE}(x) = \Pr(A \leq x) (E[V | A \leq x] - x).$$

In Online Appendix B, we show that x^* is a Berk-Nash equilibrium price if and only if $x = x^*$ maximizes an **equilibrium belief function** $\Pi(x, x^*)$ which represents

²⁶For experimental evidence of this type of naiveté, see the review by Kagel and Levin [2002] and the recent work by Charness and Levin [2009], Ivanov et al. [2010], and Esponda and Vespa [2013].

the belief about expected profit from choosing any price x under a steady-state x^* . The function Π depends on the feedback/misspecification assumptions. With perfect feedback and subjective model $\Theta = \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$, this is an example of an analogy-based game. The buyer learns the true marginal distributions of A and V and believes the joint distribution is given by the product of the marginal distributions. Therefore, the buyer's equilibrium belief function is

$$\Pi^{CE}(x) = \Pr(A \leq x) (E[V] - x), \quad (13)$$

and, in particular, does not depend on the equilibrium strategy x^* . Berk-Nash equilibrium coincides with fully cursed equilibrium in this case.

With partial feedback and subjective model $\Delta(\mathbb{A}) \times \Delta(\mathbb{V})$, the price offered by the buyer affects the sample of valuations that she observes. Also, the buyer does not realize that this selected sample would change if she were to change her price. Suppose that the buyer's behavior has stabilized to some price x^* . Then, the buyer's equilibrium belief function is

$$\Pi^{BE}(x, x^*) = \Pr(A \leq x) (E[V | A \leq x^*] - x). \quad (14)$$

Berk-Nash equilibrium coincides with naive behavioral equilibrium (Esponda, 2008) in this case.

Next, suppose that there is perfect feedback and consider a more general misspecified model. There is a partition of the set \mathbb{V} into k "analogy classes" $(\mathbb{V}_j)_{j=1, \dots, k}$, where $\cup_j \mathbb{V}_j = \mathbb{V}$ and $\mathbb{V}_i \cap \mathbb{V}_j = \emptyset$ for all $i \neq j$. The buyer believes that (A, V) are independent conditional on $V \in \mathbb{V}_i$, for each $i = 1, \dots, k$. The parameter space is $\Theta_A = \times_j \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$, where, for a parameter value $\theta = (\theta_1, \dots, \theta_k, \theta_{\mathbb{V}}) \in \Theta_A$, $\theta_{\mathbb{V}}$ parameterizes the marginal distribution over \mathbb{V} and, for each $j = 1, \dots, k$, $\theta_j \in \Delta(\mathbb{A})$ parameterizes the distribution over \mathbb{A} conditional on $V \in \mathbb{V}_j$. This is an example of an analogy-based game. Beliefs are as in a cursed equilibrium conditional on each analogy class, and so the equilibrium belief function is²⁷

$$\Pi^{ABEE}(x) = \sum_{j=1}^k \Pr(V \in \mathbb{V}_j) \{ \Pr(A \leq x | V \in \mathbb{V}_j) (E[V | V \in \mathbb{V}_j] - x) \}. \quad (15)$$

²⁷Note the well known fact that analogy-based expectation equilibrium with a single analogy class is equivalent to (fully) cursed equilibrium.

Berk-Nash equilibrium coincides with analogy-based expectation equilibrium in this case.

Finally, our setup allows us to extend ABEE to the case of partial feedback. Under natural refinements, we show in Online Appendix B that the corresponding equilibrium belief function is

$$\Pi^{BEA}(x, x^*) = \sum_{i=j}^k \Pr(V \in \mathbb{V}_j) \{ \Pr(A \leq x \mid V \in \mathbb{V}_j) (E[V \mid V \in \mathbb{V}_j, A \leq x^*] - x) \}. \quad \square$$

4 Equilibrium foundation

In this section, we provide a learning foundation for equilibrium. We follow Fudenberg and Kreps [1993] in considering games with (slightly) perturbed payoffs because, as they highlight in the context of providing a learning foundation for mixed-strategy Nash equilibrium, behavior need not be continuous in beliefs in the unperturbed game. Thus, even if beliefs were to converge, behavior would not necessarily settle down in the unperturbed game. Perturbations guarantee that if beliefs converge, then behavior also converges.

4.1 Perturbed game

A **perturbation structure** is a tuple $\mathcal{P} = \langle \Xi, P_\xi \rangle$, where: $\Xi = \times_{i \in I} \Xi^i$ and $\Xi^i \subseteq \mathbb{R}^{\#x^i}$ is a set of payoff perturbations for each action of player i ; $P_\xi = (P_{\xi^i})_{i \in I}$, where $P_{\xi^i} \in \Delta(\Xi^i)$ is a distribution over payoff perturbations of player i that is absolutely continuous with respect to the Lebesgue measure, satisfies $\int_{\Xi^i} \|\xi^i\| P_{\xi^i}(d\xi^i) < \infty$, and is independent from the perturbations of other players. A **perturbed game** $\mathcal{G}^{\mathcal{P}} = \langle \mathcal{G}, \mathcal{P} \rangle$ is composed of a game \mathcal{G} and a perturbation structure \mathcal{P} .

The timing of a perturbed game $\mathcal{G}^{\mathcal{P}}$ coincides with the timing of its corresponding (unperturbed) game \mathcal{G} , except for two modifications. First, before taking an action, each player not only observes a signal s^i but now she also privately observes a vector of own payoff perturbations $\xi^i \in \Xi^i$, where $\xi^i(x^i)$ denotes the perturbation corresponding to action x^i . Second, her payoff given action x^i and consequence y^i is $\pi^i(x^i, y^i) + \xi^i(x^i)$.

A strategy σ^i for player i is **optimal in the perturbed game** given $\mu^i \in \Delta(\Theta^i)$

if, for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, $\sigma^i(x^i | s^i) = P_\xi(\xi^i : x^i \in \Psi^i(\mu^i, s^i, \xi^i))$, where

$$\Psi^i(\mu^i, s^i, \xi^i) \equiv \arg \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}^i(\cdot | s^i, x^i)} [\pi^i(x^i, Y^i)] + \xi^i(x^i).$$

In other words, if σ^i is an optimal strategy, then $\sigma^i(x^i | s^i)$ is the probability that x^i is optimal when the state is s^i and the perturbation is ξ^i , taken over all possible realizations of ξ^i .

The definition of Berk-Nash equilibrium of a perturbed game \mathcal{G}^P is analogous to Definition 1, with the only difference that optimality must be required with respect to the *perturbed* game.

4.2 Learning foundation

We fix a perturbed game \mathcal{G}^P and consider a setting where players repeatedly play the corresponding objective game at each moment in time $t = 0, 1, 2, \dots$, and where the time- t state and signals, (ω_t, s_t) , are independently drawn from the same distribution p every period. In addition, each player i has a prior μ_0^i with full support over her (finite-dimensional) parameter space, Θ^i .²⁸ At the end of each period t , each player uses Bayes' rule and the information obtained in that period (her own signal, action, and consequence) to update her beliefs. Players believe that they face a stationary environment and myopically maximize the current period's expected payoff.

Let $B^i : \Delta(\Theta^i) \times \mathbb{S}^i \times \mathbb{X}^i \times \mathbb{Y}^i \rightarrow \Delta(\Theta^i)$ denote the Bayesian operator of player i : for all $A \subseteq \Theta$ Borel measurable and all (μ^i, s^i, x^i, y^i) ,

$$B^i(\mu^i, s^i, x^i, y^i)(A) = \frac{\int_A Q_{\theta^i}^i(y^i | s^i, x^i) \mu^i(d\theta)}{\int_\Theta Q_{\theta^i}^i(y^i | s^i, x^i) \mu^i(d\theta)}.$$

Note that Bayesian updating is well defined by Assumption 1.

Without loss of generality, we restrict behavior at time t to depend on the belief at time t and the new information received at time t , i.e., the signal and payoff perturbation.

Definition 5. A **policy** of player i is a sequence of functions $\phi^i = (\phi_t^i)_t$, where

²⁸We restrict attention to finite dimensional parameter spaces because, otherwise, Bayesian updating need not converge to the truth for most priors and parameter values even in correctly specified statistical settings (Freedman [1963], Diaconis and Freedman [1986]).

$\phi_t^i : \Delta(\Theta^i) \times \mathbb{S}^i \times \Xi^i \rightarrow \mathbb{X}^i$. A policy ϕ^i is **optimal** if $\phi_t^i \in \Psi^i$ for all t . A policy profile $\phi = (\phi^i)_{i \in I}$ is **optimal** if ϕ^i is optimal for all $i \in I$.

Let $\mathbb{H} \subseteq (\mathbb{S} \times \Xi \times \mathbb{X} \times \mathbb{Y})^\infty$ denote the set of observable histories, where any history $h = (s_0, \xi_0, x_0, y_0, \dots, s_t, \xi_t, x_t, y_t, \dots) \in \mathbb{H}$ must satisfy the feasibility restriction: for all $i \in I$, $y_t^i = f^i(x_t^i, x_t^{-i}, \omega)$ for some $\omega \in \text{supp}(p(\cdot | s_t^i))$ for all t . Let $\mathbf{P}^{\mu_0, \phi}$ denote the (objective) probability distribution over \mathbb{H} that is induced by the primitives of the game, the priors $\mu_0 = (\mu_0^i)_{i \in I}$ —which partly determine the initial actions—, and the policy profiles $\phi = (\phi^i)_{i \in I}$. Let $(\mu_t)_t$ denote the sequence of beliefs $\mu_t : \mathbb{H} \rightarrow \times_{i \in I} \Delta(\Theta^i)$ such that, for all $t \geq 1$ and all $i \in I$, μ_t^i is the posterior at time t defined recursively by $\mu_t^i(h) = B^i(\mu_{t-1}^i(h), s_{t-1}^i(h), x_{t-1}^i(h), y_{t-1}^i(h))$ for all $h \in \mathbb{H}$, where $s_{t-1}^i(h)$ is player i 's signal at $t-1$ given history h , and similarly for $x_{t-1}^i(h)$ and $y_{t-1}^i(h)$.

Definition 6. The **sequence of intended strategy profiles** given policy profile $\phi = (\phi^i)_{i \in I}$ is the sequence $(\sigma_t)_t$ of random variables $\sigma_t : \mathbb{H} \rightarrow \times_{i \in I} \Delta(\mathbb{X}^i)^{\mathbb{S}^i}$ such that, for all t and all $i \in I$,

$$\sigma_t^i(h)(x^i | s^i) = P_\xi(\xi^i : \phi_t^i(\mu_t^i(h), s^i, \xi^i) = x^i) \quad (16)$$

for all $(x^i, s^i) \in \mathbb{X}^i \times \mathbb{S}^i$.

An intended strategy profile σ_t describes how each player would behave at time t for each possible signal; it is random because it depends on the players' beliefs at time t , μ_t , which in turn depend on the past history.

One reasonable criteria to claim that the players' behavior stabilizes is that their intended behavior stabilizes with positive probability (cf. Fudenberg and Kreps, 1993).

Definition 7. A strategy profile σ is **stable [or strongly stable]** under policy profile ϕ if the sequence of intended strategies, $(\sigma_t)_t$, converges to σ with positive probability [or with probability one], i.e.,

$$\mathbf{P}^{\mu_0, \phi} \left(\lim_{t \rightarrow \infty} \|\sigma_t(h) - \sigma\| = 0 \right) > 0 \text{ [or } = 1 \text{]}$$

The next result establishes that, if behavior stabilizes to a strategy profile σ , then, for each player i , beliefs become increasingly concentrated on $\Theta^i(\sigma)$. The proof clarifies

the origin of the wKLD function in the definition of Berk-Nash equilibrium. Lemma 2 extends results from the statistics of misspecified learning (Berk [1966], Bunke and Milhaud [1998]) by considering a setting where there is active learning (i.e., agents learn from data that is endogenously generated by their own actions). Three new issues arise with active learning: (i) Previous results need to be extended to the case of non-i.i.d. and endogenous data; (ii) It is not obvious that steady-state beliefs can be characterized based on steady-state behavior, independently of the path of play (this is where Assumption 1 plays an important role; See Section 5 for an example); (iii) We allow KL divergence to be non-finite so that players can believe that other players follow *pure* strategies.²⁹

Lemma 2. *Suppose that, for a policy profile ϕ , the sequence of intended strategies, $(\sigma_t)_t$, converges to σ for all histories in a set $\mathcal{H} \subseteq \mathbb{H}$ such that $\mathbf{P}^{\mu_0, \phi}(\mathcal{H}) > 0$. Then, for all open sets $U^i \supseteq \Theta^i(\sigma)$,*

$$\lim_{t \rightarrow \infty} \mu_t^i(U^i) = 1$$

$\mathbf{P}^{\mu_0, \phi}$ -a.s. in \mathcal{H} .

Proof. It is sufficient to establish that $\lim_{t \rightarrow \infty} \int_{\Theta^i} d^i(\sigma, \theta^i) \mu_{t+1}^i(d\theta^i) = 0$ a.s. in \mathcal{H} , where $d^i(\sigma, \theta^i) = \inf_{\hat{\theta}^i \in \Theta^i(\sigma)} \|\theta^i - \hat{\theta}^i\|$. Fix $i \in I$ and $h \in \mathbb{H}$. Then

$$\begin{aligned} \int_{\Theta^i} d^i(\sigma, \theta^i) \mu_{t+1}^i(d\theta^i) &= \frac{\int_{\Theta^i} d^i(\sigma, \theta^i) \prod_{\tau=1}^t Q_{\theta^i}^i(y_\tau^i | s_\tau^i, x_\tau^i) \mu_0^i(d\theta^i)}{\int_{\Theta^i} \prod_{\tau=1}^t Q_{\theta^i}^i(y_\tau^i | s_\tau^i, x_\tau^i) \mu_0^i(d\theta^i)} \\ &= \frac{\int_{\Theta^i} d^i(\sigma, \theta^i) \prod_{\tau=1}^t \frac{Q_{\theta^i}^i(y_\tau^i | s_\tau^i, x_\tau^i)}{Q_{\sigma_\tau}^i(y_\tau^i | s_\tau^i, x_\tau^i)} \mu_0^i(d\theta^i)}{\int_{\Theta^i} \prod_{\tau=1}^t \frac{Q_{\theta^i}^i(y_\tau^i | s_\tau^i, x_\tau^i)}{Q_{\sigma_\tau}^i(y_\tau^i | s_\tau^i, x_\tau^i)} \mu_0^i(d\theta^i)} \\ &= \frac{\int_{\Theta^i} d^i(\sigma, \theta^i) e^{tK_t^i(h, \theta^i)} \mu_0^i(d\theta^i)}{\int_{\Theta^i} e^{tK_t^i(h, \theta^i)} \mu_0^i(d\theta^i)}, \end{aligned}$$

where the first line is well-defined by Assumption 1, the second line is well-defined because $\mathbf{P}^{\mu_0, \phi}(\mathcal{H}) > 0$ implies that all the terms we divide by are positive, and where we define $K_t^i(h, \theta^i) = -\frac{1}{t} \sum_{\tau=1}^t \ln \frac{Q_{\sigma_\tau}^i(y_\tau^i | s_\tau^i, x_\tau^i)}{Q_{\theta^i}^i(y_\tau^i | s_\tau^i, x_\tau^i)}$.³⁰ For any $\alpha > 0$, define $\Theta_\alpha^i(\sigma) \equiv$

²⁹For example, if player 1 believes that player 2 plays A with probability θ and B with $1 - \theta$, if we allow both $\theta = 0$ and $\theta = 1$ then the KL divergence will be infinity at $\theta = 1$ if player 2 plays B with positive probability.

³⁰If, for some θ^i , $Q_{\theta^i}^i(y_\tau^i | s_\tau^i, x_\tau^i) = 0$ for some $\tau \in \{1, \dots, t\}$, then we define $K_t^i(h, \theta^i) = -\infty$ and $\exp\{tK_t^i(h, \theta^i)\} = 0$.

$\{\theta^i \in \Theta^i : d^i(\sigma, \theta^i) < \alpha\}$. Then, for all $\varepsilon > 0$ and $\eta > 0$,

$$\int_{\Theta^i} d^i(\sigma, \theta^i) \mu_{t+1}^i(d\theta^i) \leq \varepsilon + C \frac{A_t^i(h, \sigma, \varepsilon)}{B_t^i(h, \sigma, \eta)}, \quad (17)$$

where $C \equiv \sup_{\theta_1^i, \theta_2^i \in \Theta^i} \|\theta_1^i - \theta_2^i\| < \infty$ (because Θ^i is bounded) and where

$$A_t^i(h, \sigma, \varepsilon) = \int_{\Theta^i \setminus \Theta_\varepsilon^i(\sigma)} e^{tK_t^i(h, \theta^i)} \mu_0^i(d\theta^i)$$

and

$$B_t^i(h, \sigma, \eta) = \int_{\Theta_\eta^i(\sigma)} e^{tK_t^i(h, \theta^i)} \mu_0^i(d\theta^i).$$

The proof concludes by showing that for every (sufficiently small) $\varepsilon > 0$, there exists $\eta_\varepsilon > 0$ such that $\lim_{t \rightarrow \infty} A_t^i(h, \sigma, \varepsilon) / B_t^i(h, \sigma, \eta_\varepsilon) = 0$. This result is achieved in several steps. First, for all $\varepsilon > 0$, define $K_\varepsilon^i(\sigma) = \inf \{K^i(\sigma, \theta^i) \mid \theta^i \in \Theta^i \setminus \Theta_\varepsilon^i(\sigma)\}$ and $\alpha_\varepsilon = (K_\varepsilon^i(\sigma) - K_0^i(\sigma)) / 3$. By continuity of $K^i(\sigma, \cdot)$, there exists $\bar{\varepsilon}$ and $\bar{\alpha}$ such that $0 < \alpha_\varepsilon \leq \bar{\alpha} < \infty$ for all $\varepsilon \leq \bar{\varepsilon}$. From now on, let $\varepsilon \leq \bar{\varepsilon}$. It follows that

$$K^i(\sigma, \theta^i) \geq K_\varepsilon^i(\sigma) > K_0(\sigma) + 2\alpha_\varepsilon \quad (18)$$

for all θ^i such that $d^i(\sigma, \theta^i) \geq \varepsilon$. Also, by continuity of $K^i(\sigma, \cdot)$, there exists $\eta_\varepsilon > 0$ such that

$$K^i(\sigma, \theta^i) < K_0^i(\sigma) + \alpha_\varepsilon / 2 \quad (19)$$

for all $\theta^i \in \Theta_{\eta_\varepsilon}^i$.

Second, let $\hat{\Theta}^i = \{\theta^i \in \Theta^i : Q_{\theta^i}^i(y_\tau^i \mid s_\tau^i, x_\tau^i) > 0 \text{ for all } \tau\}$ be the set of parameter values that can be rationalized by the observed history, and let $\hat{\Theta}_{\eta_\varepsilon}^i(\sigma) = \hat{\Theta}^i \cap \Theta_{\eta_\varepsilon}^i(\sigma)$. We now show that $\mu_0^i(\hat{\Theta}_{\eta_\varepsilon}^i(\sigma)) > 0$. By Lemma 1, $\Theta^i(\sigma)$ is nonempty. Pick any $\theta^i \in \Theta^i(\sigma)$. By Assumption 1, there exists $(\theta_n^i)_n$ in Θ^i such that $\lim_{n \rightarrow \infty} \theta_n^i = \theta^i$ and $Q_{\theta_n^i}^i(y^i \mid s^i, x^i) > 0$ for all $y^i \in f^i(\Omega, x^i, \mathbb{X}^{-i})$ and all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$. By continuity of $K^i(\sigma, \cdot)$, there exists θ_n^i such that $d^i(\sigma, \theta_n^i) < .5\eta_\varepsilon$. By continuity of Q and $K^i(\sigma, \cdot)$, there exists an open set U around θ_n^i such that $U \subseteq \hat{\Theta}_{\eta_\varepsilon}^i(\sigma)$. By full support, $\mu_0^i(\hat{\Theta}_{\eta_\varepsilon}^i(\sigma)) > 0$.

Next, in Claim 2.1 in the Appendix we use a law of large numbers argument for

non-iid random variables to establish that

$$\lim_{t \rightarrow \infty} K_t^i(h, \theta^i) = -K^i(\sigma, \theta^i) \quad (20)$$

for all $\theta^i \in \Theta^i$, a.s. in \mathcal{H} . Thus,

$$\begin{aligned} \liminf_{t \rightarrow \infty} B_t^i(h, \sigma, \eta_\varepsilon) e^{t(K_0^i(\sigma) + \frac{\alpha_\varepsilon}{2})} &\geq \liminf_{t \rightarrow \infty} \int_{\hat{\Theta}_{\eta_\varepsilon}^i(\sigma)} e^{t(K_0^i(\sigma) + \frac{\alpha_\varepsilon}{2} + K_t^i(h, \theta^i))} \mu_0^i(d\theta^i) \\ &\geq \int_{\hat{\Theta}_{\eta_\varepsilon}^i(\sigma)} e^{\lim_{t \rightarrow \infty} t(K_0^i(\sigma) + \frac{\alpha_\varepsilon}{2} - K^i(\sigma, \theta^i))} \mu_0^i(d\theta^i) \\ &= \infty, \end{aligned} \quad (21)$$

a.s. in \mathcal{H} , where the first line follows because we do not integrate over the complement of $\hat{\Theta}^i$ and exp is a positive function, the second line follows from Fatou's Lemma and (20), and the third line follows from (19) and the fact that $\mu_0^i(\Theta_{\eta_\varepsilon}^i(\sigma)) > 0$.

Next, we consider the term $A_t^i(h, \sigma, \varepsilon)$. Claims 2.2 and 2.3 in the Appendix imply that there exists T such that, for all $t \geq T$,

$$K_t^i(h, \theta^i) < -(K_0^i(\sigma) + (3/2)\alpha_\varepsilon)$$

for all $\theta^i \in \Theta^i \setminus \Theta_\varepsilon^i(\sigma)$, a.s. in \mathcal{H} . Thus,

$$\begin{aligned} \lim_{t \rightarrow \infty} A_t^i(h, \sigma, \varepsilon) e^{t(K_0^i(\sigma) + \alpha_\varepsilon)} &= \lim_{t \rightarrow \infty} \int_{\Theta^i \setminus \Theta_\varepsilon^i(\sigma)} e^{t(K_0^i(\sigma) + \alpha_\varepsilon + K_t^i(h, \theta^i))} \mu_0^i(d\theta^i) \\ &\leq \mu_0^i(\Theta^i \setminus \Theta_\varepsilon^i(\sigma)) \lim_{t \rightarrow \infty} e^{-t\alpha_\varepsilon/2} \\ &= 0 \end{aligned} \quad (22)$$

a.s. in \mathcal{H} . Equations (21) and (22) imply that $\lim_{t \rightarrow \infty} A_t^i(h, \sigma, \varepsilon)/B_t^i(h, \sigma, \eta_\varepsilon) = 0$ a.s.- $\mathbf{P}^{\mu_0, \phi}$. \square

Lemma 2 only implies that the *support* of posteriors converges, but posteriors need not converge. We can always find, however, a subsequence of posteriors that converges. By continuity of behavior in beliefs and the assumption that players are myopic, the stable strategy profile must be statically optimal. Thus, we obtain the following characterization of the set of stable strategy profiles when players follow optimal policies.

Theorem 2. *Suppose that a strategy profile σ is stable under an optimal policy profile for a perturbed game. Then σ is a Berk-Nash equilibrium of the perturbed game.*

Proof. Let ϕ denote the optimal policy function under which σ is stable. By Lemma 2, there exists $\mathcal{H} \subseteq \mathbb{H}$ with $\mathbf{P}^{\mu_0, \phi}(\mathcal{H}) > 0$ such that, for all $h \in \mathcal{H}$, $\lim_{t \rightarrow \infty} \sigma_t(h) = \sigma$ and $\lim_{t \rightarrow \infty} \mu_t^i(U^i) = 1$ for all $i \in I$ and all open sets $U^i \supseteq \Theta^i(\sigma)$; for the remainder of the proof, fix any $h \in \mathcal{H}$. For all $i \in I$, compactness of $\Delta(\Theta^i)$ implies the existence of a subsequence, which we denote as $(\mu_{t(j)}^i)_j$, such that $\mu_{t(j)}^i$ converges (weakly) to μ_∞^i (the limit could depend on h). We conclude by showing, for all $i \in I$:

(i) $\mu_\infty^i \in \Delta(\Theta^i(\sigma))$: Suppose not, so that there exists $\hat{\theta}^i \in \text{supp}(\mu_\infty^i)$ such that $\hat{\theta}^i \notin \Theta^i(\sigma)$. Then, since $\Theta^i(\sigma)$ is closed (by Lemma 1), there exists an open set $U^i \supset \Theta^i(\sigma)$ with closure \bar{U}^i such that $\hat{\theta}^i \notin \bar{U}^i$. Then $\mu_\infty^i(\bar{U}^i) < 1$, but this contradicts the fact that $\mu_\infty^i(\bar{U}^i) \geq \limsup_{j \rightarrow \infty} \mu_{t(j)}^i(\bar{U}^i) \geq \lim_{j \rightarrow \infty} \mu_{t(j)}^i(U^i) = 1$, where the first inequality holds because \bar{U}^i is closed and $\mu_{t(j)}^i$ converges (weakly) to μ_∞^i .

(ii) σ^i is optimal for the perturbed game given $\mu_\infty^i \in \Delta(\Theta^i)$:

$$\begin{aligned} \sigma^i(x^i | s^i) &= \lim_{j \rightarrow \infty} \sigma_{t(j)}^i(h)(x^i | s^i) \\ &= \lim_{j \rightarrow \infty} P_\xi(\xi^i : x^i \in \Psi^i(\mu_{t(j)}^i, s^i, \xi^i)) \\ &= P_\xi(\xi^i : x^i \in \Psi^i(\mu_\infty^i, s^i, \xi^i)), \end{aligned}$$

where the second line follows from optimality of ϕ^i and the (standard) fact that Ψ^i is single-valued with respect to μ^i , a.s.- P_ξ ,³¹ and the third line follows from a standard continuity argument provided in Claim A.4 in the Appendix. \square

4.3 A converse result

Theorem 2 provides our main justification for focusing on Berk-Nash equilibria: any strategy profile that is not an equilibrium cannot represent the limiting behavior of optimizing players. Theorem 2, however, does not imply that behavior will stabilize. In fact, it is well known that there are cases where optimal behavior will not converge

³¹ Ψ^i is single-valued a.s.- P_ξ because the set of ξ^i such that $\#\Psi^i(\mu^i, s^i, \xi^i) > 1$ is of dimension lower than $\#\mathbb{X}^i$ and, by absolute continuity of P_ξ , this set has measure zero.

to Nash equilibrium, which is a special case of Berk-Nash equilibrium.³² Thus, some assumption needs to be relaxed in order to prove convergence for general games.

An insight due to Fudenberg and Kreps [1993] is that a converse for the case of Nash equilibrium can be obtained by relaxing optimality and allowing players to make vanishing optimization mistakes.

Definition 8. A policy profile ϕ is **asymptotically optimal** if there exists a positive real-valued sequence $(\varepsilon_t)_t$ with $\lim_{t \rightarrow \infty} \varepsilon_t = 0$ such that, for all $i \in I$, all $(\mu^i, s^i, \xi^i) \in \Delta(\Theta^i) \times \mathbb{S}^i \times \Xi^i$, all t , and all $x^i \in \mathbb{X}^i$,

$$E_{Q_{\mu^i}(\cdot|s^i, x_t^i)} [\pi^i(x_t^i, Y^i)] + \xi^i(x_t^i) \geq E_{Q_{\mu^i}(\cdot|s^i, x^i)} [\pi^i(x^i, Y^i)] + \xi^i(x^i) - \varepsilon_t$$

where $x_t^i = \phi_t^i(\mu^i, s^i, \xi^i)$.

Fudenberg and Kreps' (1993) insight is to suppose that players are convinced early on that the equilibrium strategy is the right one to play, and continue to play this strategy unless they have strong enough evidence to think otherwise. And, as they continue to play the equilibrium strategy, the evidence increasingly convinces them that it is the right thing to do. This idea, however, need not work in the case of Berk-Nash equilibrium. The reason is that, even if players always follow the same strategy, beliefs need not converge if the model is misspecified. The following example shows this point in a context without actions.

Example (Berk, 1966). An unbiased coin is tossed every period and the agent believes that the probability of heads is either 1/4 or 3/4, but not 1/2. The agent, who takes no actions in this example, observes the outcome of each coin toss and updates her (non-doctrinaire) prior. In this case, both 1/4 and 3/4 are equidistant to the true distribution 1/2, and it is straightforward to show that the agent's beliefs do not settle down. \square

In this example, there is a failure of weak identification (see Definition 3). We can show, however, that if the game is weakly identified, then Lemma 2 and Fudenberg and Kreps' (1993) insight can be combined to obtain the following converse of Theorem 2.

³²Jordan [1993] shows that non-convergence is robust to the choice of initial conditions; Benaïm and Hirsch [1999] replicate this finding for the perturbed version of Jordan's game. In the game-theory literature, general global convergence results have only been obtained in special classes of games—e.g. zero-sum, potential, and supermodular games (Hofbauer and Sandholm, 2002).

Theorem 3. *Suppose that σ is an equilibrium of a perturbed game that is weakly identified given σ . Then there exists a profile of priors with full support and an asymptotically optimal policy profile ϕ such that σ is strongly stable under ϕ .*

Proof. See Online Appendix C.³³ □

5 Discussion

IMPORTANCE OF ASSUMPTION 1. The following example illustrates that equilibrium may fail to exist and steady-state behavior need not be characterized by equilibrium if Assumption 1 does not hold. A single agent chooses action $x \in \{A, B\}$ and obtains an outcome $y \in \{0, 1\}$. The agent's model is parameterized by $\theta = (\theta_A, \theta_B)$, where $Q_\theta(y = 1 \mid A) = \theta_A$ and $Q_\theta(y = 1 \mid B) = \theta_B$. The true model is $\theta^0 = (1/4, 3/4)$. The agent, however, is misspecified and considers only $\theta_1 = (0, 3/4)$ and $\theta_2 = (1/4, 1/4)$ to be possible, i.e., $\Theta = \{\theta_1, \theta_2\}$. In particular, Assumption 1(iii) fails for parameter value θ_1 .

Suppose that A is uniquely optimal for parameter value θ_1 and B is uniquely optimal for θ_2 (further details about payoffs are not needed). Then a Berk-Nash equilibrium does not exist. If A is played with positive probability, then θ_1 yields a wKL divergence of infinity (i.e., θ_1 cannot rationalize $y = 1$ given A) and θ_2 is the best fit. But then A is not optimal. If B is played with probability 1, then θ_1 is the best fit; but then B is not optimal.

In addition, Lemma 2 fails: Suppose that the path of play converges to pure strategy B . The best fit given B is θ_1 , but it is not necessarily true that the posterior converges weakly to a degenerate probability distribution on θ_1 . The reason is that it is possible that, along the path of play, the agent tried action A and observed $y = 1$, in which case the posterior would immediately assign probability 1 to θ_2 .

This example is a bit contrived, and, in general, Assumption 1(iii) seems fairly mild. For example, this assumption (hence the results in the paper) would be satisfied here if, for some $\bar{\varepsilon} > 0$, the parameter values $\theta = (\varepsilon, 3/4)$, for all $0 < \varepsilon \leq \bar{\varepsilon}$, were also included in Θ .

FORWARD-LOOKING AGENTS. In the dynamic model, we assumed for simplicity that all players are myopic and maximize current period's payoffs given current beliefs.

³³Note that the requirement that the priors have full support makes the statement non trivial.

In Online Appendix D, we extend Theorem 2 to the case of non-myopic players who solve a dynamic optimization problem with beliefs as a state variable. A key fact used in the proof of Theorem 2 is that myopically optimal behavior is continuous in beliefs. Non-myopic optimal behavior is also continuous in beliefs, but the issue is that it may not coincide with myopic behavior in the steady state if players still have incentives to experiment. We prove the extension by requiring that the game is weakly identified, which guarantees that players have no incentives to experiment in steady state.

STATIONARITY OF SUBJECTIVE MODELS. An important assumption in our framework is that players act as if they were facing a stationary environment. But, except in single-agent settings, the environment may not be stationary if other players are learning and experimenting with different strategies. There are a few reasons, however, why the assumption of stationary subjective models is sensible. The assumption is exactly correct in steady state and approximately correct near the steady state. Also, it is exactly correct in the construction we rely on to prove Theorem 3. Moreover, stationary models are a relatively simple and parsimonious way for people to model their environment. Each player in our framework has a stationary model of other players’ actions, but does not have a model to *explain* why other players take certain actions. If, in contrast, one player had a model of how other players make decisions, then the equilibrium concept would likely differ from Nash (or any related concept, such as Berk-Nash). For example, if one player correctly believes that all other players are like the players assumed in this paper, then she can influence their behavior and the right notion of steady state would be more akin to a Stackelberg equilibrium. Whether Nash (and Berk-Nash) or Stackelberg (and related concepts) are appropriate or not is likely to depend on the application. While our paper focuses on extending Nash to misspecified settings, it would be interesting to do the same for Stackelberg and related concepts.³⁴

LARGE POPULATION MODELS. Our framework assumes that there is a fixed number of players but, by focusing on stationary subjective models, rules out aspects of “repeated games” where players attempt to influence each others’ play. Alternatively, we can derive our results in a setting in which there is a population of a large number of agents in the role of each player $i \in I$. In this case, agents have negligible incentives to

³⁴If more than one player had a model of how other players make decisions, then we would have the well-known infinite regress problem.

influence each other’s play, and it is essentially optimal for agents to neglect repeated game considerations. We discuss some variants of population models that differ in the matching technology and feedback (the formal details are in Online Appendix E). The right variant of population model will depend on the specific application.³⁵

Single pair model: Each period a single pair of players is randomly selected from each of the i populations to play the game. At the end of the period, the signals, actions, and outcomes of their own population are revealed to everyone.³⁶ Steady-state behavior in this case corresponds exactly to the notion of Berk-Nash equilibrium described in the paper.

Random matching model: Each period, all players are randomly matched and observe only feedback from their own match. We now modify the definition of Berk-Nash equilibrium to account for this random-matching setting. The idea is similar to Fudenberg and Levine’s (1993) definition of a heterogeneous self-confirming equilibrium. Now each agent in population i can have different experiences and, hence, have different beliefs and play different strategies in steady state.

For all $i \in I$, define

$$BR^i(\sigma^{-i}) = \{ \sigma^i : \sigma^i \text{ is optimal given } \mu^i \in \Delta(\Theta^i(\sigma^i, \sigma^{-i})) \}.$$

Note that σ is a Berk-Nash equilibrium if and only if $\sigma^i \in BR^i(\sigma^{-i})$ for all $i \in I$.

Definition 9. A strategy profile σ is a **heterogeneous Berk-Nash equilibrium** of game \mathcal{G} if, for all $i \in I$, σ^i is in the convex hull of $BR^i(\sigma^{-i})$.

Intuitively, a heterogeneous equilibrium strategy σ^i is the result of convex combinations of strategies that belong to $BR^i(\sigma^{-i})$; the idea is that each of these strategies is followed by a segment of the population i .³⁷

³⁵In some cases, it may be unrealistic to assume that players are able to observe the private signals of previous generations, so some of these models might be better suited to cases with public, but not private, information.

³⁶Alternatively, we can think of different incarnations of players born every period who are able to observe the history of previous generations.

³⁷Unlike the case of heterogeneous self-confirming equilibrium, a definition where each action in the support of σ is supported by a (possibly different) belief would not be appropriate here. The reason is that $BR^i(\sigma^{-i})$ might contain only mixed, but not pure strategies (e.g., Example 1).

Random-matching model with population feedback: Each period all players are randomly matched; at the end of the period, each player in population i observes the signals, actions, and outcomes of their own population. Define

$$\bar{BR}^i(\sigma^i, \sigma^{-i}) = \{\hat{\sigma}^i : \hat{\sigma}^i \text{ is optimal given } \mu^i \in \Delta(\Theta^i(\sigma^i, \sigma^{-i}))\}.$$

Definition 10. A strategy profile σ is a **heterogeneous Berk-Nash equilibrium with population feedback** of game \mathcal{G} if, for all $i \in I$, σ^i is in the convex hull of $\bar{BR}^i(\sigma^i, \sigma^{-i})$.

The main difference when players receive population feedback is that their beliefs no longer depend on their own strategies but rather on the aggregate population strategies.

LACK OF PAYOFF INFORMATION. In the setup, players are assumed to observe their own payoffs. It is straightforward to relax this assumption by simply changing the definition of optimality. Currently, the fact that payoffs are observed implies that any belief $\mu^i \in \Delta(\Theta^i)$ of player i yields a unique expected utility function $x^i \mapsto E_{\bar{Q}_{\mu^i(\cdot|s^i, x^i)}} \pi^i(x^i, Y^i)$. If payoffs are not observed, then a belief μ^i may yield a set of expected utility functions, since an observed consequence may be associated with several possible, unobserved payoffs. In that case, optimality given μ^i should be defined as choosing a strategy that is best for any of the possible expected utility functions associated with μ^i .³⁸

EQUILIBRIUM STABILITY AND REFINEMENTS. Theorems 2 and 3 leave open the possibility of refinements. One natural refinement is to require exact, not asymptotic, optimality, and to ask whether certain equilibria can be reached with positive probability. Another possible refinement is to follow Harsanyi [1973] and rule out those equilibria that are not regular in the sense that they might not be approachable by equilibrium sequences of perturbed games. These ideas have been extensively studied in related contexts (e.g., Benaim and Hirsch [1999], Fudenberg and Levine [1993b],

³⁸Of course, the definition of identification would also need to be changed to require not only that there is a unique distribution that matches the observed data, but also that this unique distribution implies a unique expected utility function (the latter condition is satisfied in our setup by the assumption that realized payoffs are observed).

Doraszelski and Escobar [2010]), and we illustrate them in Online Appendix G in the context of Example 2.1.

CONTINGENT PLANS AND EXTENSIVE-FORM GAMES. In the timing of the objective game, player i first observes a signal s^i , then takes an action x^i , and finally observes a consequence (which can be interpreted as an ex post signal) y^i . We can consider an alternative timing where player i commits to a signal-contingent plan of action (i.e., a strategy) and observes both the realized signal s^i and the consequence y^i ex post. These two timing choices are equivalent in our setup. This equivalence may suggest that Berk-Nash equilibrium is also applicable to extensive-form games. This is correct provided that players compete by choosing contingent plan of actions and know the extensive form of the game. But the right approach is less clear if players have a misspecified view of the extensive form (for example, they may not even know the set of strategies available to them) or if players play the game sequentially (for example, we would need to define and update beliefs at each information set). The extension to extensive-form games is left for future work.

ORIGIN OF MISSPECIFIED MODELS. We view our decision to take players' misspecifications as given and characterize the resulting behavior as a first step towards being able to endogenize the subjective model. Some explanations for why agents may have misspecified models include the use of heuristics (Tversky and Kahneman, 1973), complexity (Aragones et al., 2005), the desire to avoid over-fitting the data (Al-Najjar [2009], Al-Najjar and Pai [2013]), and costly attention (Schwartzstein, 2009).

It would be interesting to allow for non-Bayesian agents who are aware of the possibility of misspecification and conduct tests to detect it. These tests, which would impose additional restrictions on beliefs, might provide an alternative way to endogenize misspecifications.³⁹

NON-BAYESIAN LEARNING. We showed that the assumption of Bayesian updating implies that the appropriate notion of “distance” in the definition of equilibrium is the (weighted) Kullback-Leibler divergence. It would be interesting to explore how other, non-Bayesian assumptions on the belief updating process yield other notions of distance.

³⁹Such additional restrictions on beliefs are imposed, for example, by Arrow and Green [1973] and Esponda [2008].

RELATIONSHIP TO BOUNDED RATIONALITY LITERATURE. As we showed, several instances of bounded rationality can be formalized via misspecified learning. A recent literature has emphasized other instances of bounded rationality in solution concepts that are potentially related to learning, such as sampling equilibrium (Osborne and Rubinstein [1998], Spiegel [2006]), the inability to recognize patterns (Piccione and Rubinstein [2003], Eyster and Piccione [2013]), valuation equilibrium (Jehiel and Samet, 2007), sparse Nash equilibrium (Gabaix, 2012), and cursed expectations equilibrium (Eyster et al., 2013). Of course, many instances of bounded rationality do not seem naturally fitted to misspecified learning, such as the literature that studies biases in information processing due to computational complexity (e.g., Rubinstein [1986], Salant [2011]), bounded memory (e.g., Wilson, 2003), or self-deception (e.g., Bénabou and Tirole [2002], Compte and Postlewaite [2004]).

6 Conclusion

We propose and provide a foundation for an equilibrium framework that allows agents to have misspecified views of their environment. Our framework unifies both standard solution concepts (Nash and self-confirming equilibrium) and a recent literature on bounded rationality and misspecified learning. It also provides a systematic approach to studying certain aspects of bounded rationality, that, we hope, stimulates further developments in this area.

References

- N. Al-Najjar. Decision makers as statisticians: Diversity, ambiguity and learning. *Econometrica*, 77(5):1371–1401, 2009.
- N. Al-Najjar and M. Pai. Coarse decision making and overfitting. *Journal of Economic Theory*, forthcoming, 2013.
- C.D. Aliprantis and K.C. Border. *Infinite dimensional analysis: a hitchhiker’s guide*. Springer Verlag, 2006.
- E. Aragonès, I. Gilboa, A. Postlewaite, and D. Schmeidler. Fact-free learning. *American Economic Review*, 95(5):1355–1368, 2005.
- K. Arrow and J. Green. Notes on expectations equilibria in bayesian settings. *Institute for Mathematical Studies in the Social Sciences Working Paper No. 33*, 1973.
- N. Barberis, A. Shleifer, and R. Vishny. A model of investor sentiment. *Journal of financial economics*, 49(3):307–343, 1998.
- P. Battigalli. *Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali*. Università Bocconi, Milano, 1987.
- P. Battigalli, S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci. Selfconfirming equilibrium and model uncertainty. Technical report, 2012.
- Roland Bénabou and Jean Tirole. Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915, 2002.
- M. Benaim and M.W. Hirsch. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behavior*, 29(1-2):36–72, 1999.
- Michel Benaim. Dynamics of stochastic approximation algorithms. In *Seminaire de Probabilités XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pages 1–68. Springer Berlin Heidelberg, 1999.
- R.H. Berk. Limiting behavior of posterior distributions when the model is incorrect. *The Annals of Mathematical Statistics*, 37(1):51–58, 1966.
- P. Billingsley. *Probability and Measure*. Wiley, 1995.
- L.E. Blume and D. Easley. Learning to be rational. *Journal of Economic Theory*, 26(2):340–351, 1982.
- M. Bray. Learning, estimation, and the stability of rational expectations. *Journal of economic theory*, 26(2):318–339, 1982.

- O. Bunke and X. Milhaud. Asymptotic behavior of bayes estimates under possibly incorrect models. *The Annals of Statistics*, 26(2):617–644, 1998.
- G. Charness and D. Levin. The origin of the winner’s curse: a laboratory study. *American Economic Journal: Microeconomics*, 1(1):207–236, 2009.
- Olivier Compte and Andrew Postlewaite. Confidence-enhanced performance. *American Economic Review*, pages 1536–1557, 2004.
- E. Dekel, D. Fudenberg, and D.K. Levine. Learning to play bayesian games. *Games and Economic Behavior*, 46(2):282–303, 2004.
- P. Diaconis and D. Freedman. On the consistency of bayes estimates. *The Annals of Statistics*, pages 1–26, 1986.
- Ulrich Doraszelski and Juan F Escobar. A theory of regular markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification. *Theoretical Economics*, 5(3):369–402, 2010.
- R. M. Dudley. *Real Analysis and Probability*. Cambridge University Press, 2002.
- R. Durrett. *Probability: Theory and Examples*. Cambridge University Press, 2010.
- D. Easley and N.M. Kiefer. Controlling a stochastic process with unknown parameters. *Econometrica*, pages 1045–1064, 1988.
- I. Esponda. Behavioral equilibrium in economies with adverse selection. *The American Economic Review*, 98(4):1269–1291, 2008.
- I. Esponda and D. Pouzo. Learning foundation for equilibrium in voting environments with private information. *working paper*, 2011.
- I. Esponda and E.I. Vespa. Hypothetical thinking and information extraction in the laboratory. *American Economic Journal: Microeconomics*, *forthcoming*, 2013.
- G. W. Evans and S. Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton University Press, 2001.
- E. Eyster and M. Rabin. Cursed equilibrium. *Econometrica*, 73(5):1623–1672, 2005.
- E. Eyster, M Rabin, and D. Vayanos. Financial markets where traders neglect the informational content of prices. *Working Paper*, 2013.
- Erik Eyster and Michele Piccione. An approach to asset-pricing under incomplete and diverse perceptions. *Econometrica*, 81(4):1483–1506, 2013.
- D.A. Freedman. On the asymptotic behavior of bayes’ estimates in the discrete case. *The Annals of Mathematical Statistics*, 34(4):1386–1403, 1963.

- D. Fudenberg and D. Kreps. Learning mixed equilibria. *Games and Economic Behavior*, 5:320–367, 1993.
- D. Fudenberg and D.M. Kreps. A theory of learning, experimentation, and equilibrium in games. Technical report, mimeo, 1988.
- D. Fudenberg and D.M. Kreps. Learning in extensive-form games i. self-confirming equilibria. *Games and Economic Behavior*, 8(1):20–55, 1995.
- D. Fudenberg and D.K. Levine. Self-confirming equilibrium. *Econometrica*, pages 523–545, 1993a.
- D. Fudenberg and D.K. Levine. Steady state learning and nash equilibrium. *Econometrica*, pages 547–573, 1993b.
- D. Fudenberg and D.K. Levine. *The theory of learning in games*, volume 2. The MIT press, 1998.
- D. Fudenberg and D.K. Levine. Learning and equilibrium. *Annual Review of Economics*, 1:385–420, 2009.
- X. Gabaix. Game theory with sparsity-based bounded rationality. *Working Paper*, 2012.
- J.C. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1):1–23, 1973.
- M. W. Hirsch, S. Smale, and R. L. Devaney. *Differential Equations, Dynamical Systems and An Introduction to Chaos*. Elsevier Academic Press, 2004.
- J. Hofbauer and W.H. Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.
- A. Ivanov, D. Levin, and M. Niederle. Can relaxation of beliefs rationalize the winner’s curse?: an experimental study. *Econometrica*, 78(4):1435–1452, 2010.
- P. Jehiel. Analogy-based expectation equilibrium. *Journal of Economic theory*, 123(2):81–104, 2005.
- P. Jehiel and F. Koessler. Revisiting games of incomplete information with analogy-based expectations. *Games and Economic Behavior*, 62(2):533–557, 2008.
- P. Jehiel and D. Samet. Valuation equilibrium. *Theoretical Economics*, 2(2):163–185, 2007.
- J. S. Jordan. Three problems in learning mixed-strategy nash equilibria. *Games and Economic Behavior*, 5(3):368–386, 1993.

- J.H. Kagel and D. Levin. The winner's curse and public information in common value auctions. *The American Economic Review*, pages 894–920, 1986.
- J.H. Kagel and D. Levin. *Common value auctions and the winner's curse*. Princeton University Press, 2002.
- E. Kalai and E. Lehrer. Rational learning leads to nash equilibrium. *Econometrica*, pages 1019–1045, 1993.
- A. P. Kirman. Learning by firms about demand conditions. In R. H. Day and T. Groves, editors, *Adaptive economic models*, pages 137–156. Academic Press, 1975.
- D. M. Kreps and R. Wilson. Sequential equilibria. *Econometrica*, pages 863–894, 1982.
- S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- H. J. Kushner and G. G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer Verlag, 2003.
- A. McLennan. Price dispersion and incomplete learning in the long run. *Journal of Economic Dynamics and Control*, 7(3):331–347, 1984.
- Y. Nyarko. Learning in mis-specified models and the possibility of cycles. *Journal of Economic Theory*, 55(2):416–427, 1991.
- Y. Nyarko. On the convexity of the value function in bayesian optimal control problems. *Economic Theory*, 4(2):303–309, 1994.
- M.J. Osborne and A. Rubinstein. Games with procedurally rational players. *American Economic Review*, 88:834–849, 1998.
- M. Piccione and A. Rubinstein. Modeling the economic interaction of agents with diverse abilities to recognize equilibrium patterns. *Journal of the European economic association*, 1(1):212–223, 2003.
- D. Pollard. *A User's Guide to Measure Theoretic Probability*. Cambridge University Press, 2001.
- M. Rabin. Inference by believers in the law of small numbers. *Quarterly Journal of Economics*, 117(3):775–816, 2002.
- M. Rabin and D. Vayanos. The gambler's and hot-hand fallacies: Theory and applications. *The Review of Economic Studies*, 77(2):730–778, 2010.
- R. Radner. *Equilibrium Under Uncertainty*, volume II of *Handbook of Mathematical Economics*, chapter 20. North-Holland Publishing Company, 1982.

- M. Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185–202, 1974.
- A. Rubinstein and A. Wolinsky. Rationalizable conjectural equilibrium: between nash and rationalizability. *Games and Economic Behavior*, 6(2):299–311, 1994.
- Ariel Rubinstein. Finite automata play the repeated prisoner’s dilemma. *Journal of economic theory*, 39(1):83–96, 1986.
- Y. Salant. Procedural analysis of choice rules with applications to bounded rationality. *The American Economic Review*, 101(2):724–748, 2011.
- T. J. Sargent. *Bounded rationality in macroeconomics*. Oxford University Press, 1993.
- T. J. Sargent. *The Conquest of American Inflation*. Princeton University Press, 1999.
- J. Schwartzstein. Selective attention and learning. *working paper*, 2009.
- R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International journal of game theory*, 4(1):25–55, 1975.
- J. Sobel. Non-linear prices and price-taking behavior. *Journal of Economic Behavior & Organization*, 5(3):387–396, 1984.
- R. Spiegler. The market for quacks. *Review of Economic Studies*, 73:1113–1131, 2006.
- R. Spiegler. Placebo reforms. *The American Economic Review*, 103(4):1490–1506, 2013.
- R. Spiegler. Bayesian networks and boundedly rational expectations. *Working Paper*, 2014.
- T. Tversky and D. Kahneman. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5:207–232, 1973.
- A. Wilson. Bounded memory and biases in information processing. *Working Paper*, 2003.

Appendix

Notation: Let $\mathbb{Z}^i = \{(s^i, x^i, y^i) \in \mathbb{S}^i \times \mathbb{X}^i \times \mathbb{Y}^i : y^i = f^i(x^i, x^{-i}, \omega), x^{-i} \in \mathbb{X}^{-i}, \omega \in \text{supp}(p(\cdot | s^i))\}$ be the set of feasible signals, actions, and consequences of player i . For each $z^i = (s^i, x^i, y^i) \in \mathbb{Z}^i$, define $\bar{P}_\sigma^i(z^i) = Q_\sigma^i(y^i | s^i, x^i) \sigma^i(x^i | s^i) p_{S^i}(s^i)$. We sometimes abuse notation and write $Q_\sigma^i(z^i) \equiv Q_\sigma^i(y^i | s^i, x^i)$, and similarly for $Q_{\theta^i}^i$. We begin by proving a few claims.

Claim A.1. *There exists $\theta_*^i \in \Theta^i$ and $K < \infty$ such that, for all $\sigma \in \Sigma$, $K^i(\sigma, \theta_*^i) \leq K$.*

Proof. By Assumption 1 and finiteness of \mathbb{Z}^i , there exists $\theta_*^i \in \Theta$ and $\alpha \in (0, 1)$ such that $Q_{\theta_*^i}^i(z^i) \geq \alpha$ for all $z^i \in \mathbb{Z}^i$. Thus, for all $\sigma \in \Sigma$, $K(\sigma, \theta_*) \leq -E_{\bar{P}_\sigma^i}[\ln Q_{\theta_*^i}^i(Z^i)] \leq -\ln \alpha$. \square

Claim A.2. *Fix any $\theta^i \in \Theta^i$ and $(\sigma_n)_n$ such that $Q_{\theta^i}^i(z^i) > 0$ for all $z^i \in \mathbb{Z}^i$ and $\lim_{n \rightarrow \infty} \sigma_n = \sigma$. Then $\lim_{n \rightarrow \infty} K^i(\sigma_n, \theta^i) = K^i(\sigma, \theta^i)$.*

Proof. Note that

$$\begin{aligned} K^i(\sigma_n, \theta^i) - K(\sigma, \theta^i) &= \sum_{z^i \in \mathbb{Z}^i} (\bar{P}_{\sigma_n}^i(z^i) \ln Q_{\sigma_n}^i(z^i) - \bar{P}_\sigma^i(z^i) \ln Q_\sigma^i(z^i)) \\ &\quad + \sum_{z^i \in \mathbb{Z}^i} (\bar{P}_\sigma^i(z^i) - \bar{P}_{\sigma_n}^i(z^i)) \ln Q_{\theta^i}^i(z^i). \end{aligned} \quad (23)$$

The first term in the RHS of (23) converges to zero because $\lim_{n \rightarrow \infty} \sigma_n = \sigma$, Q_σ is continuous, and $x \ln x$ is continuous for all $x \in [0, 1]$. The second term converges to zero because $\lim_{n \rightarrow \infty} \sigma_n = \sigma$, \bar{P}_σ^i is continuous, and $\ln Q_{\theta^i}^i(z^i) \in (-\infty, 0]$ for all $z^i \in \mathbb{Z}^i$. \square

Claim A.3. *K^i is (jointly) lower semicontinuous: Fix any $(\theta_n^i)_n$ and $(\sigma_n)_n$ such that $\lim_{n \rightarrow \infty} \theta_n^i = \theta^i$, $\lim_{n \rightarrow \infty} \sigma_n = \sigma$. Then $\liminf_{n \rightarrow \infty} K^i(\sigma_n, \theta_n^i) \geq K(\sigma, \theta^i)$.*

Proof. Note that

$$K^i(\sigma_n, \theta_n^i) - K(\sigma, \theta^i) = \sum_{z^i \in \mathbb{Z}^i} (\bar{P}_{\sigma_n}^i(z^i) \ln Q_{\sigma_n}^i(z^i) - \bar{P}_\sigma^i(z^i) \ln Q_\sigma^i(z^i)) \\ + \sum_{z^i \in \mathbb{Z}^i} (\bar{P}_\sigma^i(z^i) \ln Q_{\theta_n^i}^i(z^i) - \bar{P}_{\sigma_n}^i(z^i) \ln Q_{\theta_n^i}^i(z^i)). \quad (24)$$

The first term in the RHS of (24) converges to zero (same argument as equation 23). The proof concludes by showing that, for all $z^i \in \mathbb{Z}^i$,

$$\liminf_{n \rightarrow \infty} -\bar{P}_{\sigma_n}^i(z^i) \ln Q_{\theta_n^i}^i(z^i) \geq -\bar{P}_\sigma^i(z^i) \ln Q_{\theta^i}^i(z^i). \quad (25)$$

Suppose $\liminf_{n \rightarrow \infty} -\bar{P}_{\sigma_n}^i(z^i) \ln Q_{\theta_n^i}^i(z^i) \leq M < \infty$ (if not, (25) holds trivially). Then either (i) $\bar{P}_{\sigma_n}^i(z^i) \rightarrow \bar{P}_\sigma^i(z^i) > 0$, in which case (25) holds with equality, or (ii) $\bar{P}_{\sigma_n}^i(z^i) \rightarrow \bar{P}_\sigma^i(z^i) = 0$, in which case (25) holds because its RHS is zero (by convention that $0 \ln 0 = 0$) and its LHS is always nonnegative. \square

Claim A.4. *Suppose that ξ^i is a random vector in $\mathbb{R}^{\#\mathbb{X}^i}$ with an absolutely continuous probability distribution P_ξ . Then, for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$,*

$$\sigma^i(\mu^i)(x^i | s^i) = P_\xi \left(\xi^i : x^i \in \arg \max_{\bar{x}^i \in \mathbb{X}^i} E_{Q_{\mu^i}^i(\cdot | s^i, \bar{x}^i)} [\pi^i(\bar{x}^i, Y^i)] + \xi^i(\bar{x}^i) \right)$$

is continuous as a function of μ^i .

Proof. Fix $s^i \in \mathbb{S}^i$ and $x^i \in \mathbb{X}^i$ and define $g^i(\theta^i, \bar{x}^i) = E_{Q_{\theta^i}^i(\cdot | s^i, \bar{x}^i)} [\pi^i(x^i, Y^i) - \pi^i(\bar{x}^i, Y^i)]$, $G^i(\mu^i, \bar{x}^i) = \int_{\Theta^i} g^i(\theta^i, \bar{x}^i) \mu^i(d\theta^i)$ and $\Delta \xi^i(\bar{x}^i) = \xi^i(\bar{x}^i) - \xi^i(x^i)$ for all $\bar{x}^i \in \mathbb{X}^i$. Because $g^i(\cdot, \bar{x}^i)$ is continuous and bounded, then $G^i(\cdot, \bar{x}^i)$ is continuous under weak convergence. Then, for every $\varepsilon > 0$, there exists N_ε such that, for all $n \geq N_\varepsilon$,

$$|\sigma^i(\mu_n^i)(x^i | s^i) - \sigma^i(\mu^i)(x^i | s^i)| \leq P_\xi (\Delta \xi^i(\bar{x}^i) \leq G^i(\mu^i, \bar{x}^i) + \varepsilon \forall \bar{x}^i) \\ - P_\xi (\Delta \xi^i(\bar{x}^i) \leq G^i(\mu^i, \bar{x}^i) - \varepsilon \forall \bar{x}^i) \\ \leq \sum_{\bar{x}^i \in \mathbb{X}^i} P_\xi (G^i(\mu^i, \bar{x}^i) - \varepsilon \leq \Delta \xi^i(\bar{x}^i) \leq G^i(\mu^i, \bar{x}^i) + \varepsilon).$$

By absolute continuity of P_ξ , the above expression goes to zero as ε goes to zero.

□

Proof of Lemma 1. Part (i). Note that

$$\begin{aligned}
K^i(\sigma, \theta^i) &= - \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_\sigma^i(\cdot | s^i, x^i)} \left[\ln \frac{Q_{\theta^i}^i(Y^i | s^i, x^i)}{Q_\sigma^i(Y^i | s^i, x^i)} \right] \sigma_i(x^i | s^i) p_{S^i}(s^i) \\
&\geq - \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} \ln \left(E_{Q_\sigma^i(\cdot | s^i, x^i)} \left[\frac{Q_{\theta^i}^i(Y^i | s^i, x^i)}{Q_\sigma^i(Y^i | s^i, x^i)} \right] \right) \sigma_i(x^i | s^i) p_{S^i}(s^i) \quad (26) \\
&= 0,
\end{aligned}$$

where Jensen's inequality and the strict concavity of $\ln(\cdot)$ imply the inequality in (26) as well as the fact that (26) holds with equality if and only if $Q_{\theta^i}^i(\cdot | s^i, x^i) = Q_\sigma^i(\cdot | s^i, x^i)$ for all (s^i, x^i) such that $\sigma_i(x^i | s^i) > 0$ (recall that, by assumption, $p_{S^i}(s^i) > 0$).

Part (ii). $\Theta^i(\sigma)$ is nonempty: By Claim A.1, there exists $K < \infty$ such that the minimizers are in the constraint set $\{\theta^i \in \Theta^i : K^i(\sigma, \theta^i) \leq K\}$. Because $K^i(\sigma, \cdot)$ is continuous over a compact set, a minimum exists.

$\Theta^i(\sigma)$ is uhc: Fix any $(\sigma_n)_n$ and $(\theta_n^i)_n$ such that $\lim_{n \rightarrow \infty} \sigma_n = \sigma$, $\lim_{n \rightarrow \infty} \theta_n^i = \theta^i$, and $\theta_n^i \in \Theta^i(\sigma_n)$ for all n . We want to show that $\theta^i \in \Theta^i(\sigma)$ (so that $\Theta^i(\cdot)$ has a closed graph and hence, by compactness of Θ^i , it is uhc). Suppose, in order to obtain a contradiction, that $\theta^i \notin \Theta^i(\sigma)$. Then, by Claim A.1, there exists $\hat{\theta}^i \in \Theta^i$ and $\varepsilon > 0$ such that $K^i(\sigma, \hat{\theta}^i) \leq K^i(\sigma, \theta^i) - 3\varepsilon$ and $K^i(\sigma, \hat{\theta}^i) < \infty$. By Assumption 1, there exists $(\hat{\theta}_j^i)_j$ with $\lim_{j \rightarrow \infty} \hat{\theta}_j^i = \hat{\theta}^i$ and, for all j , $Q_{\hat{\theta}_j^i}^i(z^i) > 0$ for all $z^i \in \mathbb{Z}^i$. We will show that there is an element of the sequence, $\hat{\theta}_J^i$, that “does better” than θ_n^i given σ_n , which is a contradiction. Because $K^i(\sigma, \hat{\theta}^i) < \infty$, continuity of $K^i(\sigma, \cdot)$ implies that there exists J large enough such that $|K^i(\sigma, \hat{\theta}_J^i) - K^i(\sigma, \hat{\theta}^i)| \leq \varepsilon/2$. Moreover, Claim A.2 applied to $\theta^i = \hat{\theta}_J^i$ implies that there exists $N_{\varepsilon, J}$ such that, for all $n \geq N_{\varepsilon, J}$, $|K^i(\sigma_n, \hat{\theta}_J^i) - K^i(\sigma, \hat{\theta}_J^i)| \leq \varepsilon/2$. Thus, for all $n \geq N_{\varepsilon, J}$,

$$\begin{aligned}
|K^i(\sigma_n, \hat{\theta}_J^i) - K^i(\sigma, \hat{\theta}^i)| &\leq |K^i(\sigma_n, \hat{\theta}_J^i) - K^i(\sigma, \hat{\theta}_J^i)| + |K^i(\sigma, \hat{\theta}_J^i) - K^i(\sigma, \hat{\theta}^i)| \\
&\leq \varepsilon
\end{aligned}$$

and, therefore,

$$K^i(\sigma_n, \hat{\theta}_J^i) \leq K^i(\sigma, \hat{\theta}^i) + \varepsilon \leq K^i(\sigma, \theta^i) - 2\varepsilon. \quad (27)$$

Suppose $K^i(\sigma, \theta^i) < \infty$. By Claim A.3, there exists $n_\varepsilon \geq N_{\varepsilon, J}$ such that $K^i(\sigma_{n_\varepsilon}, \theta_{n_\varepsilon}^i) \geq$

$K^i(\sigma, \theta^i) - \varepsilon$. This result, together with expression (27), implies that $K^i(\sigma_{n_\varepsilon}, \hat{\theta}_J^i) \leq K(\sigma_{n_\varepsilon}, \theta_{n_\varepsilon}^i) - \varepsilon$. But this contradicts $\theta_{n_\varepsilon}^i \in \Theta^i(\sigma_{n_\varepsilon})$. Finally, if $K^i(\sigma, \theta^i) = \infty$, Claim A.3 implies that there exists $n_\varepsilon \geq N_{\varepsilon, J}$ such that $K^i(\sigma_{n_\varepsilon}, \theta_{n_\varepsilon}^i) \geq 2K$, where K is the bound defined in Claim A.1. But this also contradicts $\theta_{n_\varepsilon}^i \in \Theta^i(\sigma_{n_\varepsilon})$.

$\Theta^i(\sigma)$ is compact: As shown above, $\Theta^i(\cdot)$ has a closed graph, and so $\Theta^i(\sigma)$ is a closed set. Compactness of $\Theta^i(\sigma)$ follows from compactness of Θ^i . \square

Proof of Theorem 1. We prove the result in two parts. Part 1. We show existence of equilibrium in the perturbed game (defined in Section 4.1). Let $\Gamma : \times_{i \in I} \Delta(\Theta^i) \rightrightarrows \times_{i \in I} \Delta(\Theta^i)$ be a correspondence such that, for all $\mu = (\mu^i)_{i \in I} \in \times_{i \in I} \Delta(\Theta^i)$,

$$\Gamma(\mu) = \times_{i \in I} \Delta(\Theta^i(\sigma(\mu))),$$

where $\sigma(\mu) = (\sigma^i(\mu^i))_{i \in I} \in \Sigma$ and is defined as

$$\sigma^i(\mu^i)(x^i | s^i) = P_\xi \left(\xi^i : x^i \in \arg \max_{\bar{x}^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}^i(\cdot | s^i, \bar{x}^i)} [\pi^i(\bar{x}^i, Y^i)] + \xi^i(\bar{x}^i) \right) \quad (28)$$

for all $(x^i, s^i) \in \mathbb{X}^i \times \mathbb{S}^i$. Note that if there exists a $\mu_* \in \times_{i \in I} \Delta(\Theta^i)$ such that $\mu_* \in \Gamma(\mu_*)$, then $\sigma_* \equiv (\sigma^i(\mu_*^i))_{i \in I}$ is an equilibrium of the perturbed game. We now show that such μ_* exists by checking the conditions of the Kakutani-Fan-Glicksberg fixed point theorem: (i) $\times_{i \in I} \Delta(\Theta^i)$ is compact, convex and locally convex Hausdorff: The set $\Delta(\Theta^i)$ is convex, and since Θ^i is compact $\Delta(\Theta^i)$ is also compact under the weak topology (Aliprantis and Border [2006], Theorem 15.11). By Tychonoff's theorem, $\times_{i \in I} \Delta(\Theta^i)$ is compact too. Finally, the set is also locally convex under the weak topology.⁴⁰; (ii) Γ has convex, nonempty images: It is clear that $\Delta(\Theta^i(\sigma(\mu)))$ is convex valued for all μ . Also, by Lemma 1, $\Theta^i(\sigma(\mu))$ is non-empty for all μ ; (iii) Γ has a closed graph: Let $(\mu_n, \hat{\mu}_n)_n$ be such that $\hat{\mu}_n \in \Gamma(\mu_n)$ and $\mu_n \rightarrow \mu$ and $\hat{\mu}_n \rightarrow \hat{\mu}$ (under the weak topology). By Claim A.4, $\mu^i \mapsto \sigma^i(\mu^i)$ is continuous. Thus, $\sigma_n \equiv (\sigma^i(\mu_n^i))_{i \in I} \rightarrow \sigma \equiv (\sigma^i(\mu^i))_{i \in I}$. By Lemma 1, $\sigma \mapsto \Theta^i(\sigma)$ is uhc; thus, by Theorem 17.13 in Aliprantis and Border [2006], $\sigma \mapsto \times_{i \in I} \Delta(\Theta^i(\sigma))$ is also uhc. Therefore, $\hat{\mu} \in \times_{i \in I} \Delta(\Theta^i(\sigma)) = \Gamma(\mu)$.

Part 2. Fix a sequence of perturbed games indexed by the probability of perturbations $(P_{\xi, n})_n$. By Part 1, there is a corresponding sequence of fixed points $(\mu_n)_n$, such

⁴⁰This last claim follows since the weak topology is induced by a family of semi-norms of the form: $\rho(\mu, \mu') = |E_\mu[f] - E_{\mu'}[f]|$ for f continuous and bounded for any μ and μ' in $\Delta(\Theta^i)$.

that $\mu_n \in \times_{i \in I} \Delta(\Theta^i(\sigma_n))$ for all n , where $\sigma_n \equiv \sigma^i(\mu_n^i, P_{\xi,n})(x^i | s^i)$ (see equation (28)), where we now explicitly account for the dependance on $P_{\xi,n}$. By compactness, there exist subsequences of $(\mu_n)_n$ and $(\sigma_n)_n$ that converge to μ and σ , respectively. Since $\sigma \mapsto \times_{i \in I} \Delta(\Theta^i(\sigma))$ is uhc, then $\mu \in \times_{i \in I} \Delta(\Theta^i(\sigma))$. We now show that if we choose $(P_{\xi,n})_n$ such that, for all $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P_{\xi,n}(\|\xi_n^i\| \geq \varepsilon) = 0, \quad (29)$$

then σ is optimal given μ in the unperturbed game—this establishes existence of equilibrium in the unperturbed game. Suppose not, so that there exists i, s^i, x^i, \hat{x}^i , and $\varepsilon > 0$ such that $\sigma^i(x^i | s^i) > 0$ and $E_{\bar{Q}_{\mu^i}^i(\cdot | s^i, x^i)}[\pi^i(x^i, Y^i)] + 4\varepsilon \leq E_{\bar{Q}_{\mu^i}^i(\cdot | s^i, \hat{x}^i)}[\pi^i(\hat{x}^i, Y^i)]$. By continuity of $\mu^i \mapsto \bar{Q}_{\mu^i}^i$ and the fact that $\lim_{n \rightarrow \infty} \mu_n^i = \mu^i$, there exists n_1 such that, for all $n \geq n_1$, $E_{\bar{Q}_{\mu_n^i}^i(\cdot | s^i, x^i)}[\pi^i(x^i, Y^i)] + 2\varepsilon \leq E_{\bar{Q}_{\mu_n^i}^i(\cdot | s^i, \hat{x}^i)}[\pi^i(\hat{x}^i, Y^i)]$. It then follows from (28) and (29) that $\lim_{n \rightarrow \infty} \sigma^i(\mu_n^i, P_{\xi,n})(x^i | s^i) = 0$. But this contradicts $\lim_{n \rightarrow \infty} \sigma^i(\mu_n^i, P_{\xi,n})(x^i | s^i) = \sigma^i(x^i | s^i) > 0$. \square

Proof of Proposition 3. In the next paragraph, we prove the following result: For all σ and $\bar{\theta}_\sigma^i \in \Theta^i(\sigma)$, it follows that

$$Q_{\Omega, \bar{\theta}_\sigma^i}^i(\omega' | s^i) = p_\Omega(\omega' | s^i) \quad (30)$$

for all $s^i \in \mathcal{S}^i, \omega' \in \Omega$ and

$$Q_{\mathbb{X}^{-i}, \bar{\theta}_\sigma^i}^i(x^{-i} | \alpha^i) = \sum_{\omega'' \in \Omega} p_\Omega(\omega'' | \alpha^i) \prod_{j \neq i} \sigma^j(x^j | s^j(\omega'')) \quad (31)$$

for all $\alpha^i \in \mathcal{A}^i, x^{-i} \in \mathbb{X}^{-i}$. Equivalence between Berk-Nash and ABEE is then immediate by using (30) and (31) and by noticing that expected utility of player i with signal s^i and beliefs $\bar{\theta}_\sigma$ is given by $\sum_{\omega' \in \Omega} Q_{\Omega, \bar{\theta}_\sigma^i}^i(\omega' | s^i) \sum_{x^{-i} \in \mathbb{X}^{-i}} Q_{\mathbb{X}^{-i}, \bar{\theta}_\sigma^i}^i(x^{-i} | \alpha^i(\omega')) \pi^i(\bar{x}^i, x^{-i}, \omega')$.

Proof of (30) and (31): note that the (negative of) wKL divergence of player i

given σ can be written, up to a constant, as

$$\begin{aligned}
&= \sum_{s^i} \sum_{\tilde{w}, \tilde{x}^{-i}} \ln (Q_{\Omega, \theta^i}^i(\tilde{w} | s^i) Q_{\mathbb{X}^{-i}, \theta^i}^i(\tilde{x}^{-i} | \alpha^i(\tilde{w}))) \prod_{j \neq i} \sigma^j(\tilde{x}^j | s^j(\tilde{w})) p(\tilde{w} | s^i) p(s^i) \\
&= \sum_{s^i} \sum_{\tilde{w}} \ln (Q_{\Omega, \theta^i}^i(\tilde{w} | s^i)) p_{\Omega}(\tilde{w} | s^i) p(s^i) \\
&\quad + \sum_{\tilde{x}^{-i}} \sum_{\alpha^i \in \mathcal{A}^i} \ln (Q_{\mathbb{X}^{-i}, \theta^i}^i(\tilde{x}^{-i} | \alpha^i)) \sum_{\tilde{w} \in \alpha^i} \prod_{j \neq i} \sigma^j(\tilde{x}^j | s^j(\tilde{w})) p_{\Omega}(\tilde{w}).
\end{aligned}$$

It is straightforward to check that any parameter value that maximizes the above expression satisfies (30) and (31). \square

Proof of Statements in Lemma 2. Here, we prove the Claims used in the proof of Lemma 2 in the text. For each $z^i \in \mathbb{Z}^i$, define $freq_t^i(z^i) = \frac{1}{t} \sum_{\tau=1}^t \mathbf{1}_{z^i}(z_{\tau}^i)$. The function K_t^i can be written as

$$K_t^i(h, \theta^i) = \kappa_{1t}^i(h) + \kappa_{2t}^i(h) + \kappa_{3t}^i(h, \theta^i),$$

where

$$\kappa_{1t}^i(h) = -\frac{1}{t} \sum_{\tau=1}^t \sum_{z^i \in \mathbb{Z}^i} (\mathbf{1}_{z^i}(z_{\tau}^i) - \bar{P}_{\sigma_{\tau}}^i(z^i)) \ln Q_{\sigma_{\tau}}^i(z^i),$$

$$\kappa_{2t}^i(h) = -\frac{1}{t} \sum_{\tau=1}^t \sum_{z^i \in \mathbb{Z}^i} \bar{P}_{\sigma_{\tau}}^i(z^i) \ln Q_{\sigma_{\tau}}^i(z^i),$$

and

$$\kappa_{3t}^i(h, \theta^i) = \sum_{z^i \in \mathbb{Z}^i} freq_t^i(z^i) \ln Q_{\theta^i}^i(z^i).$$

The statements made below hold almost surely in \mathcal{H} , but we omit this qualification in the proofs.

Claim 2.1. $\lim_{t \rightarrow \infty} K_t^i(h, \theta^i) = -K^i(\sigma, \theta^i)$ for all $\theta^i \in \Theta^i$, a.s. in \mathcal{H} .

Proof: First, consider $\kappa_{1t}^i(h)$. Define, for all $z^i \in \mathbb{Z}^i$,

$$l_t^i(h, z^i) = (\mathbf{1}_{z^i}(z_{\tau}^i) - \bar{P}_{\sigma_{\tau}}^i(z^i)) \ln Q_{\sigma_{\tau}}^i(z^i) \quad (32)$$

and $L_t^i(h, z^i) = \sum_{\tau=1}^t \tau^{-1} l_{\tau}^i(h, z^i)$. Fix any $z^i \in \mathbb{Z}^i$. We now show that $L_t^i(h, z^i)$

converges a.s. to an integrable, and, therefore, a.s. finite $L_\infty^i(h, z^i)$. In order to show this, we use martingale convergence results. First, we show that $(L_t^i(h, z^i))_t$ is a martingale with respect to $\mathbf{P}^{\mu_0, \phi}$. Let h^t denote the partial history until time t . Observe that

$$\begin{aligned} E_{\mathbf{P}^{\mu_0, \phi}(\cdot|h^t)} [L_{t+1}^i(h, z^i)] &= \sum_{\tau=1}^t \tau^{-1} l_\tau^i(h, z^i) + \frac{1}{t+1} E_{\mathbf{P}^{\mu_0, \phi}(\cdot|h^t)} [l_{t+1}^i(h, z^i)] \\ &= \sum_{\tau=1}^t \tau^{-1} l_\tau^i(h, z^i) + \frac{1}{t+1} (\ln Q_{\sigma_{t+1}}^i(z^i)) E_{\mathbf{P}^{\mu_0, \phi}(\cdot|h^t)} [\mathbf{1}_{z^i}(z_{t+1}^i) - \bar{P}_{\sigma_{t+1}}^i(z^i)] \\ &= L_t^i(h, z^i). \end{aligned}$$

Second, we show that $(L_t^i(\cdot, z^i))_t$ satisfies

$$\sup_t E_{\mathbf{P}^{\mu_0, \phi}} [|L_t^i|] \leq M \quad (33)$$

for some finite constant M . Note that

$$\begin{aligned} E_{\mathbf{P}^{\mu_0, \phi}} [(L_t^i(h, z^i))^2] &= E_{\mathbf{P}^{\mu_0, \phi}} \left[\sum_{\tau=1}^t \tau^{-2} (l_\tau^i(h, z^i))^2 + 2 \sum_{\tau' > \tau} \frac{1}{\tau' \tau} l_\tau^i(h, z^i) l_{\tau'}^i(h, z^i) \right] \\ &= \left\{ \sum_{\tau=1}^t \tau^{-2} E_{\mathbf{P}^{\mu_0, \phi}} [(l_\tau^i(h, z^i))^2] + \sum_{\tau' > \tau} \frac{2}{\tau' \tau} E_{\mathbf{P}^{\mu_0, \phi}} [l_\tau^i(h, z^i) l_{\tau'}^i(h, z^i)] \right\} \\ &= \sum_{\tau=1}^t \tau^{-2} E_{\mathbf{P}^{\mu_0, \phi}} [(l_\tau^i(h, z^i))^2] \\ &\leq \sum_{\tau=1}^t \tau^{-2} E_{\mathbf{P}^{\mu_0, \phi}} [(\ln Q_{\sigma_\tau}^i(z^i))^2 Q_{\sigma_\tau}^i(z^i)] \\ &\leq 1 \end{aligned}$$

where the third line follows from the fact that, for $\tau' > \tau$, $E_{\mathbf{P}^{\mu_0, \phi}} [l_\tau^i(h, z^i) l_{\tau'}^i(h, z^i)] = E_{\mathbf{P}^{\mu_0, \phi}} [l_\tau^i(h, z^i) E_{\mathbf{P}^{\mu_0, \phi}(\cdot|h^\tau)} [l_{\tau'}^i(h, z^i)]] = 0$ because $(l_t^i)_t$ is a martingale difference sequence; the fourth line follows from the law of iterated expectations and the fact

that

$$\begin{aligned} E_{\mathbf{P}^{\mu_0, \phi}(\cdot|_{h^{\tau-1}})} \left[(l_\tau^i(h, z^i))^2 \right] &= (\ln Q_{\sigma_\tau}^i(z^i))^2 \left(\bar{P}_{\sigma_\tau}^i(z^i) - (\bar{P}_{\sigma_\tau}^i(z^i))^2 \right) \\ &\leq (\ln Q_{\sigma_\tau}^i(z^i))^2 \bar{P}_{\sigma_\tau}^i(z^i) \\ &\leq (\ln Q_{\sigma_\tau}^i(z^i))^2 Q_{\sigma_\tau}^i(z^i); \end{aligned}$$

and the last inequality follows because $(\ln x)^2 x \leq 1$ for all $x \in [0, 1]$, where we use the convention that $(\ln 0)^2 0 = 0$. Therefore, (33) holds and, by Theorem 5.2.8 in Durrett [2010], $L_t^i(h, z^i)$ converges a.s.- $\mathbf{P}^{\mu_0, \phi}$ to a finite $L_\infty^i(h, z^i)$. Thus, by Kronecker's lemma (Pollard [2001], page 105)⁴¹, it follows that

$$\lim_{t \rightarrow \infty} \sum_{z^i \in \mathbb{Z}^i} \left\{ t^{-1} \sum_{\tau=1}^t \ln Q_{\sigma_\tau}^i(z^i) (\mathbf{1}_{z^i}(z_\tau^i) - \bar{P}_{\sigma_\tau}^i(z^i)) \right\} = 0$$

a.s.- $\mathbf{P}^{\mu_0, \phi}$. Therefore,

$$\lim_{t \rightarrow \infty} \kappa_{1t}^i(h) = 0. \quad (34)$$

a.s.- $\mathbf{P}^{\mu_0, \phi}$. Next, consider $\kappa_{2t}^i(h)$. The assumption that $\lim_{t \rightarrow \infty} \sigma_t = \sigma$ and continuity of $Q_\sigma^i \ln Q_\sigma^i$ in σ imply that

$$\lim_{t \rightarrow \infty} \kappa_{2t}^i(h) = - \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_\sigma(\cdot|_{s^i, x^i})} [\ln Q_\sigma^i(Y^i | s^i, x^i)] \sigma^i(x^i | s^i) p_{\mathbb{S}^i}(s^i). \quad (35)$$

Finally, consider $\kappa_{3t}^i(h, \theta^i)$. For all $z^i \in \mathbb{Z}^i$,

$$|freq_t(z^i) - \bar{P}_\sigma^i(z^i)| \leq \left| \frac{1}{t} \sum_{\tau=1}^t (\mathbf{1}_{z^i}(z_\tau^i) - \bar{P}_{\sigma_\tau}^i(z^i)) \right| + \left| \frac{1}{t} \sum_{\tau=1}^t (\bar{P}_{\sigma_\tau}^i(z^i) - \bar{P}_\sigma^i(z^i)) \right|. \quad (36)$$

The first term in the RHS of (36) goes to zero (the proof is essentially identical to the proof above that κ_{1t}^i goes to zero, where equation (32) is replaced by $l_t^i(h, z^i) = (\mathbf{1}_{z^i}(z_\tau^i) - \bar{P}_{\sigma_\tau}^i(z^i))$). The second term goes to zero because $\lim_{t \rightarrow \infty} \sigma_t = \sigma$ and \bar{P}_σ^i is

⁴¹This lemma implies that for a sequence $(\ell_t)_t$ if $\sum_\tau \ell_\tau < \infty$, then $\sum_{\tau=1}^t \frac{b_\tau}{b_t} \ell_\tau \rightarrow 0$ where $(b_t)_t$ is a non-decreasing positive real valued that diverges to ∞ . We can apply the lemma with $\ell_t \equiv t^{-1} l_t$ and $b_t = t$.

continuous. Thus, for all $\zeta > 0$, there exists \hat{t}_ζ such that, for all $t \geq \hat{t}_\zeta$,

$$|\text{freq}_t^i(z^i) - \bar{P}_\sigma^i(z^i)| < \zeta \quad (37)$$

for all $z^i \in \mathbb{Z}^i$. Therefore,

$$\lim_{t \rightarrow \infty} \kappa_{3t}^i(h, \theta^i) = \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_\sigma(\cdot | s^i, x^i)} [\ln Q_{\theta^i}^i(Y^i | s^i, x^i)] \sigma^i(x^i | s^i) p_{\mathbb{S}^i}(s^i). \quad (38)$$

Equations (34), (35), and (38) imply (20). \square

Claims 2.2 and 2.3: We need the following definition. For any $\xi > 0$, let $\Theta_{\sigma, \xi}^i$ be the set such that $\theta^i \in \Theta_{\sigma, \xi}^i$ if and only if $Q_{\theta^i}^i(y^i | s^i, x^i) \geq \xi$ for all (s^i, x^i, y^i) such that $Q_\sigma^i(y^i | s^i, x^i) \sigma^i(x^i | s^i) p_{\mathbb{S}^i}(s^i) > 0$.

Claim 2.2. *There exists $\xi^* > 0$ and T_{ξ^*} such that, for all $t \geq T_{\xi^*}$,*

$$K_t^i(h, \theta^i) < -(K_0^i(\sigma) + (3/2)\alpha_\varepsilon)$$

for all $\theta^i \notin \Theta_{\sigma, \xi}^i$.

Proof: Equations (34) and (35) imply that, for all $\gamma > 0$, there exists \hat{t}_γ such that, for all $t \geq \hat{t}_\gamma$,

$$\left| \kappa_{1t}^i(h) + \kappa_{2t}^i(h) + \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_\sigma(\cdot | s^i, x^i)} [\ln Q_\sigma^i(Y^i | s^i, x^i)] \sigma^i(x^i | s^i) p_{\mathbb{S}^i}(s^i) \right| \leq \gamma. \quad (39)$$

For all $\theta^i \notin \Theta_{\sigma, \xi}^i$, let $z_{\theta^i}^i$ be such that $\bar{P}_\sigma^i(z_{\theta^i}^i) > 0$ and $Q_{\theta^i}^i(z_{\theta^i}^i) < \xi$. By (37), there exists $t_{p_L^i/2}$ such for all $t \geq t_{p_L^i/2}$,

$$\begin{aligned} \kappa_{3t}^i(h, \theta^i) &\leq \text{freq}_t^i(z_{\theta^i}^i) \ln Q_{\theta^i}^i(z_{\theta^i}^i) \\ &\leq (p_L^i/2) \ln \xi, \end{aligned} \quad (40)$$

for all $\theta^i \notin \Theta_{\sigma, \xi}^i$, where $p_L^i = \min_{\mathbb{Z}^i} \{\bar{P}_\sigma^i(z^i) : \bar{P}_\sigma^i(z^i) > 0\}$. Then (39) and (40) imply

that, for all $t \geq t_1 \equiv \max\{t_{p_L^i/2}, \hat{t}_1\}$,

$$\begin{aligned} K_t^i(h, \theta^i) &\leq - \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_{\sigma(\cdot|s^i, x^i)}} [\ln Q_{\sigma}^i(Y^i | s^i, x^i)] \sigma^i(x^i | s^i) p_{S^i}(s^i) + 1 + (p_L^i/2) \ln \xi \\ &\leq \#\mathbb{Z}^i + 1 + (q_L^i/2) \ln \xi \end{aligned} \quad (41)$$

for all $\theta^i \notin \Theta_{\sigma, \xi}^i$, where the second line follows from the facts that

$$\begin{aligned} &- \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_{\sigma(\cdot|s^i, x^i)}} [\ln Q_{\sigma}^i(Y^i | s^i, x^i)] \sigma^i(x^i | s^i) p_{S^i}(s^i) \leq \\ &- \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} \sum_{y^i \in Y^i} (\ln Q_{\sigma}^i(y^i | s^i, x^i)) Q_{\sigma}^i(y^i | s^i, x^i) \end{aligned}$$

and $x \ln(x) \in [-1, 0]$ for all $x \in [0, 1]$. In addition, the fact that $K_0^i(\sigma) < \infty$ and $\alpha_\varepsilon \leq \bar{\alpha} < \infty$ for all $\varepsilon \leq \bar{\varepsilon}$ implies that the RHS of (41) can be made lower than $-(K_0^i(\sigma) + (3/2)\alpha_\varepsilon)$ for some sufficiently small ξ^* . \square

Claim 2.3. For all $\xi > 0$, there exists \hat{T}_ξ such that, for all $t \geq \hat{T}_\xi$,

$$K_t^i(h, \theta^i) < -(K_0^i(\sigma) + (3/2)\alpha_\varepsilon)$$

for all $\theta^i \in \Theta_{\sigma, \xi}^i \setminus \Theta_\varepsilon^i(\sigma)$.

Proof: For any $\xi > 0$, let $\zeta_\xi = -\alpha_\varepsilon / (\#\mathbb{Z}^i 4 \ln \xi) > 0$. Then, by (37), there exists \hat{t}_{ζ_ξ} such that, for all $t \geq \hat{t}_{\zeta_\xi}$,

$$\begin{aligned} \kappa_{3t}^i(h, \theta^i) &\leq \sum_{\{z^i: \bar{P}_\sigma^i(z^i) > 0\}} \text{freq}_t^i(z^i) \ln Q_{\theta^i}^i(z^i) \\ &\leq \sum_{\{z^i: \bar{P}_\sigma^i(z^i) > 0\}} (\bar{P}_\sigma^i(z^i) - \zeta_\xi) \ln Q_{\theta^i}^i(z^i) \\ &\leq \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_{\sigma(\cdot|s^i, x^i)}} [\ln Q_{\theta^i}^i(Y^i | s^i, x^i)] \sigma^i(x^i | s^i) p_{S^i}(s^i) - \#\mathbb{Z}^i \zeta_\xi \ln \xi, \\ &= \sum_{(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i} E_{Q_{\sigma(\cdot|s^i, x^i)}} [\ln Q_{\theta^i}^i(Y^i | s^i, x^i)] \sigma^i(x^i | s^i) p_{S^i}(s^i) + \alpha_\varepsilon/4 \end{aligned} \quad (42)$$

for all $\theta^i \in \Theta_{\sigma, \xi}^i$ (since $Q_{\theta^i}^i(z^i) \geq \xi$ for all z^i such that $\bar{P}_\sigma^i(z^i) > 0$). Then (39) and (42)

imply that, for all $t \geq \hat{T}_\xi \equiv \max\{\hat{t}_{\zeta_\xi}, \hat{t}_{\alpha_\varepsilon/4}\}$,

$$K_t^i(h, \theta^i) < -K^i(\sigma, \theta^i) + \alpha_\varepsilon/2 \quad (43)$$

for all $\theta^i \in \Theta_{\sigma, \xi}^i$. Finally, (43) and (18) imply the desired result. \square

Online Appendix

A Relationship between correctly specified game and identification

The following property says that each player believes (either correctly or not) that the distribution over her consequences does not depend on her actions. Thus, players believe that they will get feedback about their payoff-relevant consequences irrespective of the action they take. This condition rules out those types of incorrect beliefs that are mainly due to lack of experimentation (but not due to other reasons, such as having a misspecified model).⁴²

Definition 11. A game has **subjective consequences that are own-action independent** if the subjective model satisfies: for all $i \in I$ and all $\theta^i \in \Theta^i$, $Q_{\theta^i}^i$ does not depend on x^i , i.e., $Q_{\theta^i}^i(y^i | s^i, x^i) = Q_{\theta^i}^i(y^i | s^i, \hat{x}^i)$ for all x^i, \hat{x}^i, s^i, y^i .

Remark 2. For the special case of subjective models derived from primitives $p_{\theta^i}^i$, $f_{\theta^i}^i$, and $\sigma_{\theta^i}^{-i}$, an equivalent definition of subjective consequences that are own-action independent is that, for each player i and parameter value θ^i , the feedback function $f_{\theta^i}^i$ does not depend on x^i .

We conclude by establishing useful connections between these properties.

Proposition 4. *A game that is correctly specified is weakly identified. A game that is correctly specified and has subjective consequences that are own-action independent is strongly identified.*

Proof. Fix any σ and $i \in I$. Because the game is correctly specified, there exists $\theta_*^i \in \Theta^i$ such that $Q_{\theta_*^i}^i = Q_{\sigma}^i$. In particular, $K^i(\theta_*^i, \sigma) = 0$. By Lemma 1, $K^i(\theta^i, \sigma) \geq 0$ for all $\theta^i \in \Theta^i$; therefore, $\theta_*^i \in \Theta^i(\sigma)$. Now consider any $\hat{\theta}^i \in \Theta^i(\sigma)$. Since $K^i(\theta_*^i, \sigma) = 0$, it must also be true that $K^i(\hat{\theta}^i, \sigma) = 0$. Lemma 1 then implies that $Q_{\hat{\theta}^i}^i(\cdot | s^i, x^i) =$

⁴²Arrow and Green [1973] defined a similar condition and restricted their setup to satisfy this condition.

$Q_{\theta_*^i}(\cdot | s^i, x^i)$ for all (s^i, x^i) such that $\sigma^i(x^i | s^i) > 0$. Thus, a game that is correctly specified is also weakly identified. Now suppose, in addition, that the game has own-action independent feedback. Then it follows that $Q_{\hat{\theta}^i}^i$ does not depend on x^i and, therefore, $Q_{\hat{\theta}^i}^i(\cdot | s^i, x^i) = Q_{\theta_*^i}(\cdot | s^i, x^i)$ for all (s^i, x^i) . Thus, the game is strongly identified. \square

As illustrated by Example 2.5, a correctly specified game might be strongly identified even if it does not satisfy own-action independent feedback.

B Example: Trading with adverse selection

In this section, we provide the formal details for the trading environment in Example 2.6. Let $p \in \Delta(\mathbb{A} \times \mathbb{V})$ be the true distribution; we use subscripts, such as p_A and $p_{V|A}$, to denote the corresponding marginal and conditional distributions. Let $\mathbb{Y} = \mathbb{A} \times \mathbb{V} \cup \{\square\}$ denote the space of observable consequences, where \square will be a convenient way to represent the fact that there is no trade. We denote the random variable taking values in $\mathbb{V} \cup \{\square\}$ by \hat{V} . Notice that the state space in this example is $\Omega = \mathbb{A} \times \mathbb{V}$.

Partial feedback is represented by the function $f^P : \mathbb{X} \times \mathbb{A} \times \mathbb{V} \rightarrow \mathbb{Y}$ such that

$$f^P(x, a, v) = \begin{cases} (a, v) & \text{if } a \leq x \\ (a, \square) & \text{if } a > x \end{cases},$$

and full feedback by the function $f^F(x, a, v) = (a, v)$. In all cases, payoffs are given by the function $\pi : \mathbb{X} \times \mathbb{Y} \rightarrow \mathbb{R}$, where

$$\pi(x, a, v) = \begin{cases} v - x & \text{if } a \leq x \\ 0 & \text{otherwise} \end{cases}.$$

The objective distribution for the case of partial feedback, Q^P , is, for all $x \in \mathbb{X}$,

$$Q^P(a, v | x) = p(a, v)1_{\{x \geq a\}}(x) \tag{44}$$

for all $(a, v) \in \mathbb{A} \times \mathbb{V}$ and

$$Q^P(a, \square | x) = p_A(a)1_{\{x < a\}}(x) \tag{45}$$

for all $a \in \mathbb{A}$.

The objective distribution for the case of full feedback, Q^F , is, for all $x \in \mathbb{X}$,

$$Q^F(a, v | x) = p(a, v)$$

for all $(a, v) \in \mathbb{A} \times \mathbb{V}$ and $Q^F(a, \square | x) = 0$ for all $a \in \mathbb{A}$.

We suppose that the buyer knows the environment except for the distribution $p \in \Delta(\mathbb{A} \times \mathbb{V})$. Then, for any distribution in the subjective model, Q_θ , expected profits from choosing $x \in \mathbb{X}$ are perceived to be

$$E_{Q_\theta(\cdot|x)}[\pi(x, A, \hat{V})] = \sum_{(a,v) \in \mathbb{A} \times \mathbb{V}} 1_{\{x \geq a\}}(x) (v - x) Q_\theta(a, v | x). \quad (46)$$

We suppose that the buyer has either one of two misspecifications over p captured by the parameter spaces $\Theta_I = \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$ (i.e., independent beliefs) or $\Theta_A = \times_j \Delta(\mathbb{A}) \times \Delta(\mathbb{V})$ (i.e., analogy-based beliefs) defined in the main text. Thus, combining feedback and parameter spaces, we have four cases to consider, and, for each case, we write down the corresponding subjective model and the wKLD function.

Cursed equilibrium. Feedback is f^F and the parameter space is Θ_I . The subjective model is, for all $x \in \mathbb{X}$,

$$Q_\theta^C(a, v | x) = \theta_A(a)\theta_V(v)$$

for all $(a, v) \in \mathbb{A} \times \mathbb{V}$ and $Q_\theta^C(a, \square | x) = 0$ for all $a \in \mathbb{A}$, where $\theta = (\theta_A, \theta_V) \in \Theta_I$.⁴³ Note that this is an analogy-based game. From (46), expected profits from choosing $x \in \mathbb{X}$ are perceived to be

$$Pr_{\theta_A}(A \leq x) (E_{\theta_V}[V] - x), \quad (47)$$

where Pr_{θ_A} denotes probability with respect to θ_A and E_{θ_V} denotes expectation with respect to θ_V .

⁴³In fact, the symbol \square is not necessary for this example, but we keep it so that all feedback functions are defined over the same space of consequences.

Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is⁴⁴

$$\begin{aligned} K^C(x, \theta) &= E_{Q^F(\cdot|x)} \left[\ln \frac{Q^F(A, \hat{V} | x)}{Q_\theta^C(A, \hat{V} | x)} \right] \\ &= \sum_{(a,v) \in \mathbb{A} \times \mathbb{V}} p(a, v) \ln \frac{p(a, v)}{\theta_A(a)\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_A(x), \theta_V(x))$, where $\theta_A(x) = p_A$ and $\theta_V(x) = p_V$ is the unique parameter value that minimizes $K^C(x, \cdot)$. Together with (47), we obtain equation (13) in the main text.

Behavioral equilibrium (naive version). Feedback is f^P and the parameter space is Θ_I . The subjective model is, for all $x \in \mathbb{X}$,

$$Q_\theta^{BE}(a, v | x) = \theta_A(a)\theta_V(v)1_{\{x \geq a\}}(x)$$

for all $(a, v) \in \mathbb{A} \times \mathbb{V}$ and

$$Q_\theta^{BE}(a, \square | x) = \theta_A(a)1_{\{x < a\}}(x)$$

for all $a \in \mathbb{A}$, where $\theta = (\theta_A, \theta_V) \in \Theta_I$. From (46), expected profits from choosing $x \in \mathbb{X}$ are perceived to be exactly as in equation (47).

Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is

$$\begin{aligned} K^{BE}(x, \theta) &= E_{Q^P(\cdot|x)} \left[\ln \frac{Q^P(A, \hat{V} | x)}{Q_\theta^{BE}(A, \hat{V} | x)} \right] \\ &= \sum_{\{a \in \mathbb{A}: a > x\}} p_A(a) \ln \frac{p_A(a)}{\theta_A(a)} + \sum_{\{(a,v) \in \mathbb{A} \times \mathbb{V}: a \leq x\}} p(a, v) \ln \frac{p(a, v)}{\theta_A(a)\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_A(x), \theta_V(x))$, where $\theta_A(x) = p_A$ and $\theta_V(x)(v) = p_{V|A}(v | A \leq x)$ for all $v \in \mathbb{V}$ is the unique parameter value that minimizes $K^{BE}(x, \cdot)$. Together with (47), we obtain equation (14) in the main text.

Analogy-based expectations equilibrium. Feedback is f^F and the parameter space

⁴⁴In all cases, the extension to mixed strategies is straightforward.

is Θ_A . The subjective model is, for all $x \in \mathbb{X}$,

$$Q_\theta^{ABEE}(a, v | x) = \theta_j(a)\theta_V(v)$$

for all $(a, v) \in \mathbb{A} \times \mathbb{V}_j$, all $j = 1, \dots, k$, and $Q_\theta^{ABEE}(a, \square | x) = 0$ for all $a \in \mathbb{A}$, where $\theta = (\theta_1, \theta_2, \dots, \theta_k, \theta_V) \in \Theta_A$. Note that this is an analogy-based game. From (46), expected profits from choosing $x \in \mathbb{X}$ are perceived to be

$$\sum_{j=1}^k Pr_{\theta_V}(V \in \mathbb{V}_j) \{ Pr_{\theta_j}(A \leq x | V \in \mathbb{V}_j) (E_{\theta_V}[V | V \in \mathbb{V}_j] - x) \}. \quad (48)$$

Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is

$$\begin{aligned} K^{ABEE}(x, \theta) &= E_{Q^F(\cdot|x)} \left[\ln \frac{Q^F(A, \hat{V} | x)}{Q_\theta^{ABEE}(A, \hat{V} | x)} \right] \\ &= \sum_{j=1}^k \sum_{(a,v) \in \mathbb{A} \times \mathbb{V}_j} p(a, v) \ln \frac{p(a, v)}{\theta_j(a)\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_1(x), \dots, \theta_k(x), \theta_V(x))$, where $\theta_j(x)(a) = p_{A|V_j}(a | V \in \mathbb{V}_j)$ for all $a \in \mathbb{A}$ and $\theta_V(x) = p_V$ is the unique parameter value that minimizes $K^{ABEE}(x, \cdot)$. Together with (48), we obtain equation (15) in the main text.

Behavioral equilibrium (naive version) with analogy classes. It is natural to also consider a case, unexplored in the literature, where feedback f^P is partial and the subjective model is parameterized by Θ_A . Suppose that the buyer's behavior has stabilized to some price x^* . Due to the possible correlation across analogy classes, the buyer might now believe that deviating to a different price $x \neq x^*$ affects her valuation. In particular, the buyer might have multiple beliefs at x^* . To obtain a natural equilibrium refinement, we assume that the buyer also observes the analogy class that contains her realized valuation, whether she trades or not, and that $\Pr(V \in \mathbb{V}_j, A \leq x) > 0$ for all $j = 1, \dots, k$ and $x \in \mathbb{X}$.⁴⁵ We denote this new feedback

⁴⁵Alternatively, and more naturally, we could require the equilibrium to be the limit of a sequence of mixed strategy equilibria with the property that all prices are chosen with positive probability.

assumption by a function $f^{P^*} : \mathbb{X} \times \mathbb{A} \times \mathbb{V} \rightarrow \mathbb{Y}^*$ where $\mathbb{Y}^* = \mathbb{A} \times \mathbb{V} \cup \{1, \dots, k\}$ and

$$f^{P^*}(x, a, v) = \begin{cases} (a, v) & \text{if } a \leq x \\ (a, j) & \text{if } a > x \text{ and } v \in \mathbb{V}_j \end{cases}.$$

The objective distribution given this feedback function is, for all $x \in \mathbb{X}$,

$$Q^{P^*}(a, v | x) = p(a, v)1_{\{x \geq a\}}(x) \quad (49)$$

for all $(a, v) \in \mathbb{A} \times \mathbb{V}$ and

$$Q^{P^*}(a, j | x) = p_{A|V_j}(a | V \in \mathbb{V}_j)p_V(\mathbb{V}_j)1_{\{x < a\}}(x) \quad (50)$$

for all $a \in \mathbb{A}$ and $j = 1, \dots, k$.

The subjective model is, for all $x \in \mathbb{X}$,

$$Q_\theta^{BEA}(a, v | x) = \theta_j(a)\theta_V(v)1_{\{x \geq a\}}(x)$$

for all $(a, v) \in \mathbb{A} \times \mathbb{V}_j$ and $j = 1, \dots, k$, and

$$Q_\theta^{BEA}(a, j | x) = \theta_j(a) \left(\sum_{v \in \mathbb{V}_j} \theta_V(v) \right) 1_{\{x < a\}}(x)$$

for all $a \in \mathbb{A}$ and $j = 1, \dots, k$, where $\theta = (\theta_1, \theta_2, \dots, \theta_k, \theta_V) \in \Theta_A$. In particular, from (46), expected profits from choosing $x \in \mathbb{X}$ are perceived to be exactly as in equation (48).

Also, for all (pure) strategies $x \in \mathbb{X}$, the wKLD function is

$$\begin{aligned} K^{BEA}(x, \theta) &= E_{Q^{P^*}(\cdot|x)} \left[\ln \frac{Q^{P^*}(A, \hat{V} | x)}{Q_\theta^{BEA}(A, \hat{V} | x)} \right] \\ &= \sum_{\{(a,j) \in \mathbb{A} \times \{1, \dots, k\} : a > x\}} p_{A|V_j}(a | V \in \mathbb{V}_j)p_V(\mathbb{V}_j) \ln \frac{p_{A|V_j}(a | V \in \mathbb{V}_j)p_V(\mathbb{V}_j)}{\theta_j(a) \left(\sum_{v \in \mathbb{V}_j} \theta_V(v) \right)} \\ &\quad + \sum_{j=1}^k \sum_{\{(a,v) \in \mathbb{A} \times \mathbb{V}_j : a \leq x\}} p(a, v) \ln \frac{p(a, v)}{\theta_j(a)\theta_V(v)}. \end{aligned}$$

For each $x \in \mathbb{X}$, $\theta(x) = (\theta_1(x), \dots, \theta_k(x), \theta_V(x))$, where $\theta_j(x)(a) = p_{A|V_j}(a | V \in \mathbb{V}_j)$ for all $a \in \mathbb{A}$ and $\theta_V(x)(v) = p_{V|A}(v | V \in \mathbb{V}_j, A \leq x)p_V(\mathbb{V}_j)$ for all $v \in \mathbb{V}_j$, $j = 1, \dots, k$ is the unique parameter value that minimizes $K^{BEA}(x, \cdot)$. Together with (48), we obtain

$$\Pi^{BEA}(x, x^*) = \sum_{i=j}^k \Pr(V \in \mathbb{V}_j) \{ \Pr(A \leq x | V \in \mathbb{V}_j) (E[V | V \in \mathbb{V}_j, A \leq x^*] - x) \}. \quad (51)$$

A price x^* is an equilibrium if and only if $x = x^*$ maximizes (51).

C Proof of converse result: Theorem 3

Let $(\bar{\mu}^i)_{i \in I}$ be a belief profile that supports σ as an equilibrium. Consider the following policy profile $\phi = (\phi_t^i)_{i,t}$: For all $i \in I$ and all t ,

$$(\mu^i, s^i, \xi^i) \mapsto \phi_t^i(\mu^i, s^i, \xi^i) \equiv \begin{cases} \varphi^i(\bar{\mu}^i, s^i, \xi^i) & \text{if } \max_{i \in I} \|\bar{Q}_{\mu^i}^i - \bar{Q}_{\bar{\mu}^i}^i\| \leq \frac{1}{2C}\varepsilon_t \\ \varphi^i(\mu^i, s^i, \xi^i) & \text{otherwise,} \end{cases}$$

where φ^i is an arbitrary selection from Ψ^i , $C \equiv \max_I \{ \#\mathbb{Y}^i \times \sup_{\mathbb{X}^i \times \mathbb{Y}^i} |\pi^i(x^i, y^i)| \} < \infty$, and the sequence $(\varepsilon_t)_t$ will be defined below. For all $i \in I$, fix any prior μ_0^i with full support on Θ^i such that $\mu_0^i(\cdot | \Theta^i(\sigma)) = \bar{\mu}^i$ (where for any $A \subset \Theta$ Borel, $\mu(\cdot | A)$ is the conditional probability given A).

We now show that if $\varepsilon_t \geq 0$ for all t and $\lim_{t \rightarrow \infty} \varepsilon_t = 0$, then ϕ is asymptotically optimal. Throughout this argument, we fix an arbitrary $i \in I$. Abusing notation, let $U^i(\mu^i, s^i, \xi^i, x^i) = E_{\bar{Q}_{\mu^i}^i(\cdot | s^i, x^i)} [\pi^i(x^i, Y^i)] + \xi^i(x^i)$. It suffices to show that

$$U^i(\mu^i, s^i, \xi^i, \phi_t^i(\mu^i, s^i, \xi^i)) \geq U^i(\mu^i, s^i, \xi^i, x^i) - \varepsilon_t \quad (52)$$

for all (i, t) , all (μ^i, s^i, ξ^i) , and all x^i . By construction of ϕ , equation (52) is satisfied if $\max_{i \in I} \|\bar{Q}_{\mu^i}^i - \bar{Q}_{\bar{\mu}^i}^i\| > \frac{1}{2C}\varepsilon_t$. If, instead, $\max_{i \in I} \|\bar{Q}_{\mu^i}^i - \bar{Q}_{\bar{\mu}^i}^i\| \leq \frac{1}{2C}\varepsilon_t$, then

$$U^i(\bar{\mu}^i, s^i, \xi^i, \phi_t^i(\mu^i, s^i, \xi^i)) = U^i(\bar{\mu}^i, s^i, \xi^i, \varphi^i(\bar{\mu}^i, s^i, \xi^i)) \geq U^i(\bar{\mu}^i, s^i, \xi^i, x^i), \quad (53)$$

for all $x^i \in \mathbb{X}^i$. Moreover, for all x^i ,

$$\begin{aligned}
|U^i(\bar{\mu}^i, s^i, \xi^i, x^i) - U^i(\mu^i, s^i, \xi^i, x^i)| &= \left| \sum_{y^i \in \mathbb{Y}^i} \pi(x^i, y^i) \{ \bar{Q}_{\bar{\mu}^i}^i(y^i | s^i, x^i) - \bar{Q}_{\mu^i}^i(y^i | s^i, x^i) \} \right| \\
&\leq \sup_{\mathbb{X}^i \times \mathbb{Y}^i} |\pi^i(x^i, y^i)| \sum_{y^i \in \mathbb{Y}^i} | \{ \bar{Q}_{\bar{\mu}^i}^i(y^i | s^i, x^i) - \bar{Q}_{\mu^i}^i(y^i | s^i, x^i) \} | \\
&\leq \sup_{\mathbb{X}^i \times \mathbb{Y}^i} |\pi^i(x^i, y^i)| \times \#\mathbb{Y}^i \times \max_{y^i, x^i, s^i} | \bar{Q}_{\bar{\mu}^i}^i(y^i | s^i, x^i) - \bar{Q}_{\mu^i}^i(y^i | s^i, x^i) |
\end{aligned}$$

so by our choice of C , $|U^i(\bar{\mu}^i, s^i, \xi^i, x^i) - U^i(\mu^i, s^i, \xi^i, x^i)| \leq 0.5\varepsilon_t$ for all x^i . Therefore, equation (53) implies equation (52); thus ϕ is asymptotically optimal if $\varepsilon_t \geq 0$ for all t and $\lim_{t \rightarrow \infty} \varepsilon_t = 0$.

We now construct a sequence $(\varepsilon_t)_t$ such that $\varepsilon_t \geq 0$ for all t and $\lim_{t \rightarrow \infty} \varepsilon_t = 0$. Let $\bar{\phi}^i = (\bar{\phi}_t^i)_t$ be such that $\bar{\phi}_t^i(\mu^i, \cdot, \cdot) = \varphi^i(\bar{\mu}^i, \cdot, \cdot)$ for all μ^i ; i.e., $\bar{\phi}^i$ is a stationary policy that maximizes utility under the assumption that the belief is always $\bar{\mu}^i$. Let $\zeta^i(\mu^i) \equiv 2C \|\bar{Q}_{\mu^i}^i - \bar{Q}_{\bar{\mu}^i}^i\|$ and suppose (the proof is at the end) that

$$\mathbf{P}^{\mu_0, \bar{\phi}}(\lim_{t \rightarrow \infty} \max_{i \in I} |\zeta^i(\mu_t^i(h))| = 0) = 1 \quad (54)$$

(recall that $\mathbf{P}^{\mu_0, \bar{\phi}}$ is the probability measure over \mathbb{H} induced by the policy profile $\bar{\phi}$; by definition of $\bar{\phi}$, $\mathbf{P}^{\mu_0, \bar{\phi}}$ does not depend on μ_0). Then by the 2nd Borel-Cantelli lemma (Billingsley [1995], pages 59-60), for any $\gamma > 0$, $\sum_t \mathbf{P}^{\mu_0, \bar{\phi}}(\max_{i \in I} |\zeta^i(\mu_t^i(h))| \geq \gamma) < \infty$. Hence, for any $a > 0$, there exists a sequence $(\tau(j))_j$ such that

$$\sum_{t \geq \tau(j)} \mathbf{P}^{\mu_0, \bar{\phi}} \left(\max_{i \in I} |\zeta^i(\mu_t^i(h))| \geq 1/j \right) < \frac{3}{a} 4^{-j} \quad (55)$$

and $\lim_{j \rightarrow \infty} \tau(j) = \infty$. For all $t \leq \tau(1)$, we set $\varepsilon_t = 3C$, and, for any $t > \tau(1)$, we set $\varepsilon_t \equiv 1/N(t)$, where $N(t) \equiv \sum_{j=1}^{\infty} 1\{\tau(j) \leq t\}$. Observe that, since $\lim_{j \rightarrow \infty} \tau(j) = \infty$, $N(t) \rightarrow \infty$ as $t \rightarrow \infty$ and thus $\varepsilon_t \rightarrow 0$.

Next, we show that

$$\mathbf{P}^{\mu_0, \phi} \left(\lim_{t \rightarrow \infty} \|\sigma_t(h^\infty) - \sigma\| = 0 \right) = 1,$$

where $(\sigma_t)_t$ is the sequence of intended strategies given ϕ , i.e.,

$$\sigma_t^i(h)(x^i | s^i) = P_\xi (\xi^i : \phi_t^i(\mu_t^i(h), s^i, \xi^i) = x^i).$$

Observe that, by definition,

$$\sigma^i(x^i | s^i) = P_\xi \left(\xi^i : x^i \in \arg \max_{x^i \in X^i} E_{\bar{Q}_{\bar{\mu}^i(\cdot | s^i, \hat{x}^i)}} [\pi^i(\hat{x}^i, Y^i)] + \xi^i(\hat{x}^i) \right).$$

Since $\varphi^i \in \Psi^i$, it follows that we can write $\sigma^i(x^i | s^i) = P_\xi (\xi^i : \varphi^i(\bar{\mu}^i, s^i, \xi^i) = x^i)$. Let $H \equiv \{h : \|\sigma_t(h) - \sigma\| = 0, \text{ for all } t\}$. Note that it is sufficient to show that $\mathbf{P}^{\mu_0, \phi}(H) = 1$. To show this, observe that

$$\begin{aligned} \mathbf{P}^{\mu_0, \phi}(H) &\geq \mathbf{P}^{\mu_0, \phi} \left(\bigcap_t \{ \max_i \zeta^i(\mu_t) \leq \varepsilon_t \} \right) \\ &= \prod_{t=\tau(1)+1}^{\infty} \mathbf{P}^{\mu_0, \phi} \left(\max_i \zeta^i(\mu_t) \leq \varepsilon_t \mid \bigcap_{l < t} \{ \max_i \zeta^i(\mu_l) \leq \varepsilon_l \} \right) \\ &= \prod_{t=\tau(1)+1}^{\infty} \mathbf{P}^{\mu_0, \bar{\phi}} \left(\max_i \zeta^i(\mu_t) \leq \varepsilon_t \mid \bigcap_{l < t} \{ \max_i \zeta^i(\mu_l) \leq \varepsilon_l \} \right) \\ &= \mathbf{P}^{\mu_0, \bar{\phi}} \left(\bigcap_{t > \tau(1)} \{ \max_i \zeta^i(\mu_t) \leq \varepsilon_t \} \right), \end{aligned}$$

where the second line omits the term $\mathbf{P}^{\mu_0, \phi}(\max_i \zeta^i(\mu_t) < \varepsilon_t \text{ for all } t \leq \tau(1))$ because it is equal to 1 (since $\varepsilon_t \geq 3C$ for all $t \leq \tau(1)$); the third line follows from the fact that $\phi_{t-1}^i = \bar{\phi}_{t-1}^i$ if $\zeta^i(\mu_{t-1}) \leq \varepsilon_{t-1}$, so the probability measure is equivalently given by $\mathbf{P}^{\mu_0, \bar{\phi}}$; and where the last line also uses the fact that $\mathbf{P}^{\mu_0, \bar{\phi}}(\max_i \zeta^i(\mu_t) < \varepsilon_t \text{ for all } t \leq \tau(1)) = 1$. In addition, for all $a > 0$,

$$\begin{aligned} \mathbf{P}^{\mu_0, \bar{\phi}} \left(\bigcap_{t > \tau(1)} \{ \max_i \zeta^i(\mu_t) \leq \varepsilon_t \} \right) &= \mathbf{P}^{\mu_0, \bar{\phi}} \left(\bigcap_{n \in \{1, 2, \dots\}} \bigcap_{\{t > \tau(1) : N(t) = n\}} \{ \max_i \zeta^i(\mu_t) \leq n^{-1} \} \right) \\ &\geq 1 - \sum_{n=1}^{\infty} \sum_{\{t : N(t) = n\}} \mathbf{P}^{\mu_0, \bar{\phi}} \left(\max_i \zeta^i(\mu_t) \geq n^{-1} \right) \\ &\geq 1 - \sum_{n=1}^{\infty} \frac{3}{a} 4^{-n} = 1 - \frac{1}{a}, \end{aligned}$$

where the last line follows from (55). Thus, we have shown that $\mathbf{P}^{\mu_0, \phi}(H) \geq 1 - 1/a$ for all $a > 0$; hence, $\mathbf{P}^{\mu_0, \phi}(H) = 1$.

We conclude the proof by showing that equation (54) indeed holds. Observe that σ is trivially stable under $\bar{\phi}$. By Lemma 2, for all $i \in I$ and all open sets $U^i \supseteq \Theta^i(\sigma)$,

$$\lim_{t \rightarrow \infty} \mu_t^i(U^i) = 1 \quad (56)$$

a.s. - $\mathbf{P}^{\mu_0, \bar{\phi}}$ (over \mathbb{H}). Let \mathcal{H} denote the set of histories such that $x_t^i(h) = x^i$ and $s_t^i(h) = s^i$ implies that $\sigma^i(x^i | s^i) > 0$. By definition of $\bar{\phi}$, $\mathbf{P}^{\mu_0, \bar{\phi}}(\mathcal{H}) = 1$. Thus, it suffices to show that $\lim_{t \rightarrow \infty} \max_{i \in I} |\zeta^i(\mu_t^i(h))| = 0$ a.s. - $\mathbf{P}^{\mu_0, \bar{\phi}}$ over \mathcal{H} . To do this, take any $A \subseteq \Theta$ that is closed. By equation (56), for all $i \in I$, and almost all $h \in \mathcal{H}$,

$$\limsup_{t \rightarrow \infty} \int 1_A(\theta) \mu_{t+1}^i(d\theta) = \limsup_{t \rightarrow \infty} \int 1_{A \cap \Theta^i(\sigma)}(\theta) \mu_{t+1}^i(d\theta).$$

Moreover,

$$\begin{aligned} \int 1_{A \cap \Theta^i(\sigma)}(\theta) \mu_{t+1}^i(d\theta) &\leq \int 1_{A \cap \Theta^i(\sigma)}(\theta) \left\{ \frac{\prod_{\tau=1}^t Q_\theta^i(y_\tau^i | s_\tau^i, x_\tau^i) \mu_0^i(d\theta)}{\int_{\Theta^i(\sigma)} \prod_{\tau=1}^t Q_\theta^i(y_\tau^i | s_\tau^i, x_\tau^i) \mu_0^i(d\theta)} \right\} \\ &= \mu_0^i(A | \Theta^i(\sigma)) \\ &= \bar{\mu}^i(A), \end{aligned}$$

where the first line follows from the fact that $\Theta^i(\sigma) \subseteq \Theta^i$; the second line follows from the fact that, since $h \in \mathcal{H}$, the fact that the game is weakly identified given σ implies that $\prod_{\tau=1}^t Q_\theta^i(y_\tau^i | s_\tau^i, x_\tau^i)$ is constant with respect to θ for all $\theta \in \Theta^i(\sigma)$, and the last line follows from our choice of μ_0^i . Therefore, we established that a.s. - $\mathbf{P}^{\mu_0, \bar{\phi}}$ over \mathcal{H} , $\limsup_{t \rightarrow \infty} \mu_{t+1}^i(h)(A) \leq \bar{\mu}^i(A)$ for A closed. By the portmanteau lemma, this implies that, a.s. - $\mathbf{P}^{\mu_0, \bar{\phi}}$ over \mathcal{H} ,

$$\lim_{t \rightarrow \infty} \int_{\Theta} f(\theta) \mu_{t+1}^i(h)(d\theta) = \int_{\Theta} f(\theta) \bar{\mu}^i(d\theta)$$

for any f real-valued, bounded and continuous. Since, by assumption, $\theta \mapsto Q_\theta^i(y^i | s^i, x^i)$ is bounded and continuous, the previous display applies to $Q_\theta^i(y^i | s^i, x^i)$, and since y, s, x take a finite number of values, this result implies that $\lim_{t \rightarrow \infty} \|\bar{Q}_{\mu_t^i}^i - \bar{Q}_{\bar{\mu}^i}^i\| = 0$ for all $i \in I$ a.s. - $\mathbf{P}^{\mu_0, \bar{\phi}}$ over \mathcal{H} . \square

D Non-myopic players

In the main text, we proved the results for the case where players are myopic. Here, we assume that players maximize discounted expected payoffs, where $\delta^i \in [0, 1)$ is the discount factor of player i . In particular, players can be forward looking and decide to experiment. Players believe, however, that they face a stationary environment and, therefore, have no incentives to influence the future behavior of other players. We assume for simplicity that players know the distribution of their own payoff perturbations.

Because players believe that they face a stationary environment, they solve a (subjective) dynamic optimization problem that can be cast recursively as follows. By the Principle of Optimality, $V^i(\mu^i, s^i)$ denotes the maximum expected discounted payoffs (i.e., the value function) of player i who starts a period by observing signal s^i and by holding belief μ^i if and only if

$$V^i(\mu^i, s^i) = \int_{\Xi^i} \left\{ \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} [\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} [V^i(\hat{\mu}^i, S^i)]] \right\} P_{\xi}(d\xi^i), \quad (57)$$

where $\hat{\mu}^i = B^i(\mu^i, s^i, x^i, Y^i)$ is the updated belief. For all (μ^i, s^i, ξ^i) , let

$$\Phi^i(\mu^i, s^i, \xi^i) = \arg \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} [\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} [V^i(\hat{\mu}^i, S^i)]] .$$

We use standard arguments to prove the following properties of the value function.⁴⁶

Lemma 3. *There exists a unique solution V^i to the Bellman equation (57); this solution is bounded in $\Delta(\Theta^i) \times \mathbb{S}^i$ and continuous as a function of μ^i . Moreover, Φ^i is single-valued and continuous with respect to μ^i , a.s.- P_{ξ} .*

Proof. We first show that

$$\xi^i \mapsto \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} [\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} [H^i(B^i(\mu^i, s^i, x^i, Y^i), S^i)]]$$

is measurable for any $H^i \in L^\infty(\Delta(\Theta^i) \times \mathbb{S}^i)$ and any (μ^i, s^i) . It suffices to check that set of the form $\left\{ \xi^i : \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} [\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} [H^i(B^i(\mu^i, s^i, x^i, Y^i), S^i)]] < a \right\}$

⁴⁶Doraszelski and Escobar [2010] study a similarly perturbed version of the Bellman equation.

is measurable for any $a \in \mathbb{R}$. It is easy to see that this set is of the form

$$\cap_{x^i \in \mathbb{X}^i} \left\{ \xi^i : E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} \left[\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} \left[H^i(B^i(\mu^i, s^i, x^i, Y^i), S^i) \right] \right] < a \right\}.$$

Each set in the intersection is trivially (Borel) measurable, therefore the intersection (of finitely many) of them is also measurable.

We now define the Bellman operator $H^i \in L^\infty(\Delta(\Theta^i) \times \mathbb{S}^i) \mapsto T^i[H^i]$ where

$$T^i[H^i](\mu^i, s^i) \equiv \int_{\Xi^i} \left\{ \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} \left[\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} \left[H^i(B^i(\mu^i, s^i, x^i, Y^i), S^i) \right] \right] \right\} P_\xi(d\xi^i).$$

By our first result, the operator is well-defined. Moreover, since $\int \|\xi^i\| P_\xi(d\xi^i) < \infty$ and π^i is uniformly bounded, it follows that T^i maps $L^\infty(\Delta(\Theta^i) \times \mathbb{S}^i)$ into itself. By Blackwell's sufficient conditions, there exists a unique $V^i \in L^\infty(\Delta(\Theta^i) \times \mathbb{S}^i)$ such that $V^i = T^i[V^i]$.

In order to establish continuity of V^i , by standard arguments it suffices to show that T^i maps $C(\Delta(\Theta^i) \times \mathbb{S}^i)$ into itself, where $C(\Delta(\Theta^i) \times \mathbb{S}^i) \equiv \{f \in L^\infty(\Delta(\Theta^i) \times \mathbb{S}^i) : \mu^i \mapsto f(\mu^i, s^i) \text{ is continuous, for all } s^i\}$. Suppose that $H^i \in C(\Delta(\Theta^i) \times \mathbb{S}^i)$. Since $\mu^i \mapsto B^i(\mu^i, s^i, x^i, y^i)$ is also continuous for all (s^i, x^i, y^i) , by the Dominated convergence theorem, it follows that $\mu^i \mapsto \int_{\mathbb{S}^i} H^i(B^i(\mu^i, s^i, x^i, y^i), \hat{s}^i) p_{S^i}(d\hat{s}^i)$ is continuous, for all (s^i, x^i, y^i) . This result and the fact that $\theta^i \mapsto E_{Q_{\theta^i}^i(y^i|s^i, x^i)} \left[\int_{\mathbb{S}^i} H^i(B^i(\tilde{\mu}^i, s^i, x^i, y^i), \hat{s}^i) p_{S^i}(d\hat{s}^i) \right]$ is bounded and continuous (for a fixed $\tilde{\mu}^i$), readily implies that

$$\mu^i \mapsto E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} \left[E_{p_{S^i}} \left[H^i(B^i(\mu^i, s^i, x^i, Y^i), S^i) \right] \right]$$

is also continuous. This result and the fact that $\mu^i \mapsto E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} \left[\pi^i(x^i, Y^i) \right]$ is continuous ($\theta^i \mapsto \sum_{y^i \in \mathbb{Y}^i} \pi^i(x^i, y^i) Q_{\theta^i}^i(y^i|s^i, x^i)$ is continuous and bounded), imply that T^i maps $C(\Delta(\Theta^i) \times \mathbb{S}^i)$ into itself.

The fact that Φ^i single-valued *a.s.* $- P_\xi$, i.e., for all (μ^i, s^i) , $P_\xi(\xi^i : \#\Phi^i(\mu^i, s^i, \xi^i) > 1) = 0$, follows because the set of ξ^i such that $\#\Phi^i(\mu^i, s^i, \xi^i) > 1$ is of dimension lower than $\#\mathbb{X}^i$ and, by absolute continuity of P_ξ , this set has measure zero.

To show continuity of $\mu^i \mapsto \Phi^i(\mu^i, s^i, \xi^i)$, observe that, by the previous calculations, $(\mu^i, x^i) \mapsto E_{\bar{Q}_{\mu^i}(\cdot|s^i, x^i)} \left[\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} \left[V^i(\hat{\mu}^i, S^i) \right] \right]$ is continuous (under the product topology) for all s^i and *a.s.* $- P_\xi$. Also, \mathbb{X}^i is compact. Thus by the theorem of the maximum, $\mu^i \mapsto \Phi^i(\mu^i, s^i, \xi^i)$ is continuous, *a.s.* $- P_\xi$.

□

Without loss of generality, we restrict behavior to depend on the state of the recursive problem. Optimality of a policy is defined as usual (with the requirement that $\phi_t^i \in \Phi^i$ for all t).

Lemma 2 implies that the *support* of posteriors converges, but posteriors need not converge. We can always find, however, a subsequence of posteriors that converges. By continuity of dynamic behavior in beliefs, the stable strategy profile is dynamically optimal (in the sense of solving the dynamic optimization problem) given this convergent posterior. For weakly identified games, the convergent posterior is a fixed point of the Bayesian operator. Thus, the players' limiting strategies will provide no new information. Since the value of experimentation is non-negative, it follows that the stable strategy profile must also be myopically optimal (in the sense of solving the optimization problem that ignores the future), which is the definition of optimality used in the definition of Berk-Nash equilibrium. Thus, we obtain the following characterization of the set of stable strategy profiles when players follow optimal policies.

Theorem 4. *Suppose that a strategy profile σ is stable under an optimal policy profile for a perturbed and weakly identified game. Then σ is a Berk-Nash equilibrium of the game.*

Proof. Let ϕ denote the optimal policy function under which σ is stable. By Lemma 2, there exists $\mathcal{H} \subseteq \mathbb{H}$ with $\mathbf{P}^{\mu_0, \phi}(\mathcal{H}) > 0$ such that, for all $h \in \mathcal{H}$, $\lim_{t \rightarrow \infty} \sigma_t(h) = \sigma$ and $\lim_{t \rightarrow \infty} \mu_t^i(U^i) = 1$ for all $i \in I$ and all open sets $U^i \supseteq \Theta^i(\sigma)$; for the remainder of the proof, fix any $h \in \mathcal{H}$. For all $i \in I$, compactness of $\Delta(\Theta^i)$ implies the existence of a subsequence, which we denote as $(\mu_{t(j)}^i)_j$, such that $\mu_{t(j)}^i$ converges (weakly) to μ_∞^i (the limit could depend on h). We now show that $\mu_\infty^i \in \Delta(\Theta^i(\sigma))$. Suppose not, so that there exists $\hat{\theta}^i \in \text{supp}(\mu_\infty^i)$ such that $\hat{\theta}^i \notin \Theta^i(\sigma)$. Then, since $\Theta^i(\sigma)$ is closed (by Lemma 1), there exists an open set $U^i \supset \Theta^i(\sigma)$ with closure \bar{U}^i such that $\hat{\theta}^i \notin \bar{U}^i$. Then $\mu_\infty^i(\bar{U}^i) < 1$, but this contradicts the fact that $\mu_\infty^i(\bar{U}^i) \geq \limsup_{j \rightarrow \infty} \mu_{t(j)}^i(\bar{U}^i) \geq \lim_{j \rightarrow \infty} \mu_{t(j)}^i(U^i) = 1$, where the first inequality holds because \bar{U}^i is closed and $\mu_{t(j)}^i$ converges (weakly) to μ_∞^i .

Given that $\lim_{j \rightarrow \infty} \sigma_{t(j)} = \sigma$ and $\mu_\infty^i \in \Delta(\Theta^i(\sigma))$ for all i , it remains to show that, for all i , σ^i is optimal for the perturbed game given $\mu_\infty^i \in \Delta(\Theta^i)$, *i.e.*, for all (s^i, x^i) ,

$$\sigma^i(x^i \mid s^i) = P_\xi(\xi^i : \psi^i(\mu_\infty^i, s^i, \xi^i) = \{x^i\}), \quad (58)$$

where $\psi^i(\mu_\infty^i, s^i, \xi^i) \equiv \arg \max_{x^i \in \mathbb{X}^i} E_{\bar{Q}_{\mu_\infty^i}^i(\cdot | s^i, x^i)} [\pi^i(x^i, Y^i)] + \xi^i(x^i)$.

To establish (58), fix $i \in I$ and $s^i \in \mathbb{S}^i$. Then

$$\begin{aligned} \lim_{j \rightarrow \infty} \sigma_{t(j)}^i(h)(x^i | s^i) &= \lim_{j \rightarrow \infty} P_\xi (\xi^i : \phi_{t(j)}^i(\mu_{t(j)}^i, s^i, \xi^i) = x^i) \\ &= P_\xi (\xi^i : \Phi^i(\mu_\infty^i, s^i, \xi^i) = \{x^i\}), \end{aligned}$$

where the second line follows by optimality of ϕ^i and Lemma 3. This implies that $\sigma^i(x^i | s^i) = P_\xi (\xi^i : \Phi^i(\mu_\infty^i, s^i, \xi^i) = \{x^i\})$. Thus, it remains to show that

$$P_\xi (\xi^i : \Phi^i(\mu_\infty^i, s^i, \xi^i) = \{x^i\}) = P_\xi (\xi^i : \psi^i(\mu_\infty^i, s^i, \xi^i) = \{x^i\}) \quad (59)$$

for all x^i such that $P_\xi (\xi^i : \Phi^i(\mu_\infty^i, s^i, \xi^i) = \{x^i\}) > 0$. From now on, fix any such x^i . Since $\sigma^i(x^i | s^i) > 0$, the assumption that the game is weakly identified implies that $Q_{\theta_1^i}^i(\cdot | x^i, s^i) = Q_{\theta_2^i}^i(\cdot | x^i, s^i)$ for all $\theta_1^i, \theta_2^i \in \Theta(\sigma)$. The fact that $\mu_\infty^i \in \Delta(\Theta^i(\sigma))$ then implies that

$$B^i(\mu_\infty^i, s^i, x^i, y^i) = \mu_\infty^i \quad (60)$$

for all $y^i \in \mathbb{Y}^i$. Thus, $\Phi^i(\mu_\infty^i, s^i, \xi^i) = \{x^i\}$ is equivalent to

$$\begin{aligned} &E_{\bar{Q}_{\mu_\infty^i}^i(\cdot | s^i, x^i)} \left[\pi^i(x^i, Y^i) + \xi^i(x^i) + \delta E_{p_{S^i}} [V^i(\mu_\infty^i, S^i)] \right] \\ &> E_{\bar{Q}_{\mu_\infty^i}^i(\cdot | s^i, \tilde{x}^i)} \left[\pi^i(\tilde{x}^i, Y^i) + \xi^i(\tilde{x}^i) + \delta E_{p_{S^i}} [V^i(B^i(\mu_\infty^i, s^i, \tilde{x}^i, Y^i), S^i)] \right] \\ &\geq E_{\bar{Q}_{\mu_\infty^i}^i(\cdot | s^i, \tilde{x}^i)} \left[\pi^i(\tilde{x}^i, Y^i) + \xi^i(\tilde{x}^i) \right] + \delta E_{p_{S^i}} \left[V^i(E_{\bar{Q}_{\mu_\infty^i}^i(\cdot | s^i, \tilde{x}^i)} [B^i(\mu_\infty^i, s^i, \tilde{x}^i, Y^i)], S^i) \right] \\ &= E_{\bar{Q}_{\mu_\infty^i}^i(\cdot | s^i, \tilde{x}^i)} \left[\pi^i(\tilde{x}^i, Y^i) + \xi^i(\tilde{x}^i) \right] + \delta E_{p_{S^i}} [V^i(\mu_\infty^i, S^i)] \end{aligned}$$

for all $\tilde{x}^i \in \mathbb{X}^i$, where the first line follows by equation (60) and definition of Φ^i , the second line follows by the convexity⁴⁷ of V^i as a function of μ^i and Jensen's inequality, and the last line by the fact that Bayesian beliefs have the martingale property. In turn, the above expression is equivalent to $\psi(\mu_\infty^i, s^i, \xi^i) = \{x^i\}$. \square

E Population models

Using arguments similar to the ones in the text, it is now straightforward to conclude that the definition of heterogenous Berk-Nash equilibrium captures the steady state

⁴⁷See, for example, Nyarko [1994], for a proof of convexity of the value function.

of a learning environment with a population of agents in the role of each player. To see the idea, let each population i be composed of a continuum of agents in the unit interval $K \equiv [0, 1]$. A strategy of agent ik (meaning agent $k \in K$ from population i) is denoted by σ^{ik} . The aggregate strategy of population (i.e., player) i is $\sigma^i = \int_K \sigma^{ik} dk$.

Random matching model. Suppose that each agent is optimizing and that, for all i , (σ_t^{ik}) converges to σ^{ik} a.s. in K , so that individual behavior stabilizes.⁴⁸ Then Lemma 2 says that the support of beliefs must eventually be $\Theta^i(\sigma^{ik}, \sigma^{-i})$ for agent ik . Next, for each ik , take a convergent subsequence of beliefs μ_t^{ik} and denote it μ_∞^{ik} . It follows that $\mu_\infty^{ik} \in \Delta(\Theta^i(\sigma^{ik}, \sigma^{-i}))$ and, by continuity of behavior in beliefs, σ^{ik} is optimal given μ_∞^{ik} . In particular, $\sigma^{ik} \in BR^i(\sigma^{-i})$ for all ik and, since $\sigma^i = \int_K \sigma^{ik} dk$, it follows that σ^i is in the convex hull of $BR^i(\sigma^{-i})$.

Random matching model with population feedback. Suppose that each agent is optimizing and that, for all i , $\sigma_t^i = \int_K \sigma_t^{ik} dk$ converges to σ^i . Then Lemma 2 says that the support of beliefs must eventually be $\Theta^i(\sigma^i, \sigma^{-i})$ for any agent in population i . Next, for each ik , take a convergent subsequence of beliefs μ_t^{ik} and denote it μ_∞^{ik} . It follows that $\mu_\infty^{ik} \in \Delta(\Theta^i(\sigma^i, \sigma^{-i}))$ and, by continuity of behavior in beliefs, σ^{ik} is optimal given μ_∞^{ik} . In particular, $\sigma^{ik} \in \bar{BR}^i(\sigma^{-i})$ for all i, k and, since $\sigma^i = \int_K \sigma^{ik} dk$, it follows that σ^i is in the convex hull of $\bar{BR}^i(\sigma^{-i})$.

F Extension to non-finite space of consequences

In this section we extend Lemma 2 to the case where the state space Ω and the space of consequences of each player i , \mathbb{Y}^i , are non-finite, Borel subsets of an Euclidean space. The signal and action spaces continue to be finite. We assume that P_Ω is absolutely continuous with respect to the Lebesgue measure.⁴⁹

Let $\mathbb{Z}^i = \{(s^i, x^i, y^i) \in \mathbb{S}^i \times \mathbb{X}^i \times \mathbb{Y}^i : y^i = f^i(\omega, x^i, x^{-i}), x^{-i} \in \mathbb{X}^{-i}, \omega \in \text{supp}(p(\cdot | s^i))\}$ be the set of feasible signals, actions, and consequences of player i . For each $A \subseteq \mathbb{Y}^i$

⁴⁸We need individual behavior to stabilize; it is not enough that it stabilizes in the aggregate. This is natural, for example, if we believe that agents whose behavior is unstable will eventually realize they have a misspecified model.

⁴⁹It is straightforward to extend this section to the case where P_Ω is absolutely continuous with respect to some σ -finite measure ν .

Borel,

$$Q_\sigma^i(A | s^i, x^i) = \int_{\{(\omega, x^{-i}): f^i(\omega, x^i, x^{-i}) \in A\}} \mu_{\sigma^{-i}}(dx^{-i} | \omega) P_\Omega(d\omega | s^i)$$

for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, and $\omega \mapsto \mu_{\sigma^{-i}}(\cdot | \omega) \in \Delta(\mathbb{X}^{-i})$ with $\mu_{\sigma^{-i}}(x^{-i} | \omega) \equiv \sum_{s^{-i}} \prod_{j \neq i} \sigma^j(x^j | s^j) p(s^{-i} | \omega)$. Define $\bar{P}_\sigma^i(A, s^i, x^i) = Q_\sigma^i(A | s^i, x^i) \sigma^i(x^i | s^i) p_{S^i}(s^i)$.

For any measurable function $h : \mathbb{Y}^i \rightarrow \mathbb{R}$ and for any $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, let

$$\int_A h(y) Q_\sigma^i(dy | s^i, x^i) = \int_{\{(\omega, x^{-i}): f^i(\omega, x^i, x^{-i}) \in A\}} h(f^i(\omega, x^i, x^{-i})) \mu_{\sigma^{-i}}(dx^{-i} | \omega) P_\Omega(d\omega | s^i)$$

for any $A \subseteq \mathbb{Y}^i$. For the case $A = \mathbb{Y}^i$, the expression in the RHS becomes

$$\int_\Omega \sum_{x^{-i} \in \mathbb{X}^{-i}} h(f^i(\omega, x^i, x^{-i})) \mu_{\sigma^{-i}}(x^{-i} | \omega) P_\Omega(d\omega | s^i).$$

Suppose that the density of $Q_\sigma^i(\cdot | s^i, x^i)$ exists, and denote it by $q_\sigma^i(\cdot | s^i, x^i)$. We sometimes abuse notation and write $q_\sigma^i(z^i) \equiv q_\sigma^i(y^i | s^i, x^i)$, and similarly for $q_{\theta^i}^i$.

The following conditions are needed to extend the results.

Condition 1. (i) $\theta \mapsto q_\theta^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i)$ is continuous a.s.-Lebesgue and for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$; (ii) for any $\theta^i \in \Theta^i$, there exists an open ball $B(\theta^i)$, such that $\int \sup_{\theta' \in B(\theta^i)} |\ln q_{\theta'}^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i)| P_\Omega(d\omega | s^i) < \infty$, for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$; (iii) For any $\theta^i \in \Theta^i$, $\int |\ln q_{\theta^i}^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i)|^2 P_\Omega(d\omega | s^i) < \infty$, for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$.

Condition 2. (i) $\sigma \mapsto q_\sigma^i(\cdot | s^i, x^i)$ is continuous a.s.-Lebesgue for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$; (ii) there exists a $\bar{l} : \mathbb{Y}^i \rightarrow \mathbb{R}_+ \cup \{\infty\}$ such that for any $\sigma \in \Sigma$, $q_\sigma^i(\cdot | s^i, x^i) \max\{1, |\ln q_\sigma^i(\cdot | s^i, x^i)|\} \leq \bar{l}(\cdot)$ a.s.-Lebesgue and for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$, and $\int_{\mathbb{Y}^i} \bar{l}(y) dy < \infty$.

Roughly, these conditions impose continuity and uniform-integrability-type restrictions of the log densities. Condition 1(ii) is analogous to assumption 2.1 in Bunke and Milhaud [1998] and it implies that $\ln q_\theta^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i)$ is P_Ω -integrable for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$, for all $\theta \in \Theta^i$.⁵⁰ Condition 1(iii) is needed to establish (pointwise)

⁵⁰Note that this assumption rules out the case in which the wKLD function is infinity, and it was not required in the main text for the case of a finite set of consequences. A finite wKLD function might rule out the case where a player believes her opponent can follow any possible pure strategy, but this is not really an issue for practical purposes because we can restrict a player to believe that her opponent chooses each strategy with probability at least $\varepsilon > 0$ and then take $\varepsilon \rightarrow 0$.

LLN-type results for $\frac{1}{t} \sum_{\tau=1}^t (\ln q_{\theta}^i(y_{\tau}^i | s_{\tau}^i, x_{\tau}^i))$ using martingale convergence methods. Condition 2(ii) is also needed to ensure uniform integrability of $q_{\sigma}^i(\cdot | s^i, x^i) \ln q_{\sigma}^i(\cdot | s^i, x^i)$ and $q_{\sigma}^i(\cdot | s^i, x^i)$.

The next condition ensures that the family of subjective densities is stochastic equi-continuous. This condition is needed to show an uniform LLN-type result for $t^{-1} \sum_{\tau=1}^t \log q_{\theta}(y_{\tau} | s_{\tau}, x_{\tau})$ and it is standard.

Condition 3. For any $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ such that

$$\sup_{\|\theta - \theta'\| \leq \delta} |q_{\theta}^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i) - q_{\theta'}^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i)| \leq \varepsilon$$

a.s.- P_{Ω} and for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$.

In this section we let $K_t^i(h, \theta^i) = -\frac{1}{t} \sum_{\tau=1}^t \left(\ln \frac{q_{\sigma_{\tau}}^i(y_{\tau}^i | s_{\tau}^i, x_{\tau}^i)}{q_{\theta}^i(y_{\tau}^i | s_{\tau}^i, x_{\tau}^i)} \right)$ and $K^i(\sigma, \theta^i) = \int \left(\ln \frac{q_{\sigma}^i(y | s, x)}{q_{\theta}^i(y | s, x)} \right) \bar{P}_{\sigma}^i(dz)$. The main results of this section are the following two Lemmas.

Lemma 4. For any $\theta^i \in \Theta^i$ and any $\epsilon > 0$, there exists a $t(\epsilon, \theta^i)$ such that

$$|K_t^i(h, \theta^i) + K^i(\sigma, \theta^i)| \leq \epsilon$$

for all $t \geq t(\epsilon, \theta^i)$ a.s.- $\mathbf{P}^{\mu_0, \phi}$.

Henceforth, let $\varepsilon > 0$ and α_{ε} be as in the proof of Lemma 2

Lemma 5. There exists a $t(\alpha_{\varepsilon})$ such that, for all $t \geq t(\alpha_{\varepsilon})$

$$K_t^i(h, \theta^i) \leq -K_0(\sigma) - \frac{3}{2}\alpha_{\varepsilon}$$

for all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \varepsilon\}$.

These Lemmas extend the results in Claims 2.1 and 2.2-3 to the non-finite consequence space. Given these Lemmas, the extension of Lemma 2 to the non-finite consequence space follows from analogous steps to those in the Proof of Lemma 2 in the text and therefore a formal proof is omitted.

In order to proof these lemmas we need the following intermediate results (their proof are relegated to the end of the section). Let $\varphi_a(z^i, \theta) \equiv \ln q_\theta^i(y|s^i, x^i)1\{q_\theta^i(y^i|s^i, x^i) \geq a\} + \ln a 1\{q_\theta^i(y^i|s^i, x^i) < a\}$ for any $a \geq 0$ (we omit the i to ease the notational burden). Also, fix a $(\sigma_t)_t$ and σ as in the statement of Lemma 2.

Lemma 6. *For any $a > 0$:*

(i) $\theta \mapsto \varphi_a(z, \theta)$ is continuous a.s.- \bar{P}_σ^i .

(ii) There exists a mapping $(y^i, s^i, x^i) \mapsto \bar{\varphi}_a(y^i, s^i, x^i)$, such that $|\varphi_a(z^i, \theta)| \leq \bar{\varphi}_a(z^i)$ and $\int_{\mathbb{Z}^i} \bar{\varphi}_a(z^i) \bar{P}_\sigma^i(dz^i) < \infty$.

(iii) For any $\varepsilon > 0$, there exists a $\delta'(\varepsilon, a) > 0$ such that for any $\theta \in \Theta^i$

$$\sup_{\theta' \in \Theta^i \mid \|\theta - \theta'\| \leq \delta'(\varepsilon, a)} |\varphi_a(Z, \theta) - \varphi_a(Z, \theta')| < \varepsilon$$

a.s.- $\bar{P}_{\sigma_t}^i$ for all t .

(iv) For any $\theta^i \in \Theta^i$, $\int_{\mathbb{Z}^i} |\varphi_a(z^i, \theta)|^2 \bar{P}_\sigma^i(dz^i) < \infty$.

Lemma 7. *For any $\varepsilon > 0$ and any $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \varepsilon\}$, there exists a $t(\varepsilon, a, \theta)$ such that, for all $t \geq t(\varepsilon, a, \theta)$*

$$\frac{1}{t} \sum_{\tau=1}^t \varphi_a(Z_\tau, \theta) - \int \varphi_a(z, \theta) \bar{P}_{\sigma_\tau}^i(dz) \leq \varepsilon.$$

Lemma 8. *For any $\varepsilon > 0$ and any $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \varepsilon\}$, there exists a $t(\varepsilon, a, \theta)$ such that, for all $t \geq t(\varepsilon, a, \theta)$*

$$\frac{1}{t} \sum_{\tau=1}^t \int \varphi_a(z, \theta) \bar{P}_{\sigma_\tau}^i(dz) - \int \varphi_a(z, \theta) \bar{P}_\sigma^i(dz) \leq \varepsilon$$

Lemma 9. *For any $\varepsilon > 0$, there exists a $a(\varepsilon)$ such that, for all $0 < a \leq a(\varepsilon)$*

$$\int \varphi_a(z, \theta) \bar{P}_\sigma^i(dz) - \int \ln q_\theta^i(y|s, x) \bar{P}_\sigma^i(dz) \leq \varepsilon$$

for all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \varepsilon\}$.

Lemma 10. For any $\epsilon > 0$, there exists a $t^o(\epsilon)$ such that

$$\left| \frac{1}{t} \sum_{\tau=1}^t \ln q_{\sigma_\tau}^i(Y_\tau^i | S_\tau^i, X_\tau^i) - \int \ln q_\sigma^i(y^i | s^i, x^i) \bar{P}_\sigma^i(dz) \right| \leq \epsilon$$

for all $t \geq t^o(\epsilon)$.

Proof. [Proof of Lemma 4] Given Lemma 10 it only remains to show that For any $\theta^i \in \Theta^i$ and any $\epsilon > 0$, there exists a $t(\epsilon, \theta^i)$ such that

$$\left| \frac{1}{t} \sum_{\tau=1}^t \ln q_\theta^i(Y_\tau^i | S_\tau^i, X_\tau^i) - \int (\ln q_\theta^i(y | s, x)) \bar{P}_\sigma^i(dz) \right| \leq \epsilon$$

for all $t \geq t(\epsilon, \theta^i)$ a.s.- $\mathbf{P}^{\mu_0, \phi}$.

In order to show this, we use martingale convergence results analogous to those in the proof of Claim 2.1; thus we only present a sketch. Henceforth we omit the index i . Let $\ell_\tau(\theta) \equiv \ln q_\theta(Z_\tau, \theta) - \int \ln q_\theta(z) Q_{\sigma_\tau}(dz)$. And let $L_t(h, \theta) = \sum_{\tau=1}^t \tau^{-1} \ell_\tau(\theta)$. We now show that $L_t(h, \theta)$ converges a.s. to an integrable $L_\infty(h, \theta)$.

First, it is easy to show that $(L_t(h, \theta))_t$ is a martingale with respect to $\mathbf{P}^{\mu_0, \phi}$, i.e. $E_{\mathbf{P}^{\mu_0, \phi}(\cdot|h^t)} [L_{t+1}(h, \theta)] = L_t(h, \theta)$. Second, we show that $\sup_t E_{\mathbf{P}^{\mu_0, \phi}} [\max\{L_t(h, \theta), 0\}] < \infty$, by directly showing that $\sup_t E_{\mathbf{P}^{\mu_0, \phi}} [|L_t(h, \theta)|] < \infty$. By the Jensen inequality it is enough to bound $E_{\mathbf{P}^{\mu_0, \phi}} [(L_t(h, \theta))^2]$ uniformly in t . Moreover, by the same steps as in the proof of Claim 2.1 one can show that $E_{\mathbf{P}^{\mu_0, \phi}} [(L_t(h, \theta))^2] \leq \sum_{\tau=1}^t \tau^{-2} E_{\mathbf{P}^{\mu_0, \phi}} [(\ln q_\theta(Z_\tau))^2]$. And, by iterated expectations

$$\begin{aligned} E_{\mathbf{P}^{\mu_0, \phi}} [(\ln q_\theta(Z_\tau))^2] &= E_{\mathbf{P}^{\mu_0, \phi}} \left[\sum_{x^{-i}} \int_{\Omega} (\ln q_\theta(f(\omega, X^i, x^{-i}), S^i, X^i, \theta))^2 \mu_{\sigma_t^{-i}}(x^{-i} | \omega) P_\Omega(d\omega | S^i) \right] \\ &\leq \sum_{x^{-i}} E_{\mathbf{P}^{\mu_0, \phi}} \left[\int_{\Omega} (\ln q_\theta(f(\omega, X^i, x^{-i}), S^i, X^i, \theta))^2 P_\Omega(d\omega | S^i) \right] \end{aligned}$$

which is finite by condition 1(iii).

Therefore, $\sup_t E_{\mathbf{P}^{\mu_0, \phi}} [\max\{L_t(h, \theta), 0\}] < \infty$ holds and, by Theorem 5.2.8 in Durrett [2010], $L_t(h, \theta)$ converges a.s.- $\mathbf{P}^{\mu_0, \phi}$ to $L_\infty(h, \theta)$. Let H be the set of h for which $L_t(h, \theta)$ converges to $L_\infty(h, \theta)$ and $L_\infty(h, \theta) < \infty$. Since $L_\infty(h, \theta)$ is $\mathbf{P}^{\mu_0, \phi}$ -integrable,

H has probability 1 under $\mathbf{P}^{\mu_0, \phi}$. For any $h \in H$, by Kronecker's lemma (Pollard [2001], page 105)⁵¹, $\frac{1}{t} \sum_{\tau=1}^t \ell_\tau(\theta) \rightarrow 0$ and thus the desired result follows. \square

Proof. [Proof of Lemma 5] We show that: For any $\epsilon > 0$, there exists a $t(\epsilon)$ such that, for all $t \geq t(\epsilon)$

$$K_t^i(h, \theta^i) \leq - \inf_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} K^i(\sigma, \theta^i) + 2\epsilon$$

for all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$. This is sufficient since by continuity of $K^i(\sigma, \cdot)$, $K^i(\sigma, \theta^i) \geq K_0(\sigma) + 2\alpha_\epsilon$ for any $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$. Thus by choosing $\epsilon = \frac{\alpha_\epsilon}{4}$ it follows that

$$K_t^i(h, \theta^i) \leq -K_0(\sigma) - 2\alpha_\epsilon + 2\epsilon = -K_0(\sigma) - \frac{3}{2}\alpha_\epsilon.$$

Let $a \equiv a(0.25\epsilon)$ as in Lemma 9. We divide the proof in several steps and throughout the proof we omit the dependence on i to ease the notational burden.

STEP 1. Note that $\{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$ is compact. Consider the following open cover of the set: $\{\theta \mid \|\theta - \theta'\| < 2^{-j}\delta\}$ for all $\theta' \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$ and $j = 1, 2, \dots$. Where $\delta = \delta'(0.25\epsilon, a)$ and δ' is as in Lemma 6(iii). By compactness, there exists a finite sub-cover: $B_1 \equiv \{\theta \mid \|\theta - \theta_1\| < 2^{-j_1}\delta\}, \dots, B_K \equiv \{\theta \mid \|\theta - \theta_K\| < 2^{-j_K}\delta\}$ of K elements. Note that K only depends on δ and thus only on ϵ (note that $a = a(0.25\epsilon)$) (actually, it also depends on ϵ , but this is considered fixed).

STEP 2. We now show that: There exists a $t(\epsilon)$ such that, for all $t \geq t(\epsilon)$

$$\frac{1}{t} \sum_{\tau=1}^t (-\ln q_\theta(Y_\tau | S_\tau, X_\tau)) \geq \inf_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \int (-\ln q_\theta(y|s, x)) \bar{P}_\sigma^i(dz) - \epsilon$$

for all $t \geq t^*(\epsilon)$ and all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$. Or equivalently,

$$\frac{1}{t} \sum_{\tau=1}^t (\ln q_\theta(Y_\tau | S_\tau, X_\tau)) \leq \sup_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \int (\ln q_\theta(y|s, x)) \bar{P}_\sigma^i(dz) + \epsilon$$

⁵¹This lemma implies that for a sequence $(\ell_t)_t$ if $\sum_\tau \ell_\tau < \infty$, then $\sum_{\tau=1}^t \frac{b_\tau}{b_t} \ell_\tau \rightarrow 0$ where $(b_t)_t$ is a non-decreasing positive real valued that diverges to ∞ . We can apply the lemma with $\ell_t \equiv t^{-1}l_t$ and $b_t = t$.

for all $t \geq t^*(\epsilon)$ and all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$.

First note that $\ln q_\theta(y|s, x) \leq \varphi_a(z, \theta)$ for all $a \geq 0$. And thus using the construction in step 1,

$$\begin{aligned}
\sup_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \frac{1}{t} \sum_{\tau=1}^t \ln q_\theta(Y_\tau | S_\tau, X_\tau) &\leq \max_{k=1, \dots, K} \frac{1}{t} \sum_{\tau=1}^t \sup_{\theta \in B_k} \varphi_a(Z_\tau, \theta) \\
&\leq \max_{k=1, \dots, K} \frac{1}{t} \sum_{\tau=1}^t \varphi_a(Z_\tau, \theta_k) \\
&\quad + \max_{k=1, \dots, K} \frac{1}{t} \sum_{\tau=1}^t \sup_{\theta \in B_k} |\varphi_a(Z_\tau, \theta) - \varphi_a(Z_\tau, \theta_k)| \\
&\leq \max_{k=1, \dots, K} \frac{1}{t} \sum_{\tau=1}^t \left\{ \varphi_a(Z_\tau, \theta_k) - \int \varphi_a(z, \theta_k) \bar{P}_{\sigma_\tau}^i(dz) \right\} \\
&\quad + \max_{k=1, \dots, K} \frac{1}{t} \sum_{\tau=1}^t \int \varphi_a(z, \theta_k) \bar{P}_{\sigma_\tau}^i(dz) + 0.25\epsilon
\end{aligned}$$

where the last line follows from Lemma 6(iii) and the fact that for $\theta \in B_k$, $\|\theta - \theta_k\| < \delta$. By Lemma 7 there exist a $t_1(k, \epsilon) (= t(0.25\epsilon, a, \theta_k))$ such that

$$\frac{1}{t} \sum_{\tau=1}^t \left\{ \varphi_a(Z_\tau, \theta_k) - \int \varphi_a(z, \theta_k) \bar{P}_{\sigma_\tau}(dz) \right\} \leq 0.25\epsilon$$

for all $t \geq t_1(k, \epsilon)$ and all $k = 1, \dots, K$. Then

$$\begin{aligned}
\sup_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \frac{1}{t} \sum_{\tau=1}^t \ln q_\theta(Y_\tau | S_\tau, X_\tau) &\leq \max_{k=1, \dots, K} \left\{ \frac{1}{t} \sum_{\tau=1}^t \int \varphi_a(z, \theta_k) \bar{P}_{\sigma_\tau}(dz) - \int \varphi_a(z, \theta_k) \bar{P}_\sigma(dz) \right\} \\
&\quad + \max_{k=1, \dots, K} \left\{ \int \varphi_a(z, \theta_k) \bar{P}_\sigma(dz) \right\} + 0.5\epsilon
\end{aligned}$$

for all $t \geq t_1 \equiv \max_{k=1, \dots, K} t_1(k, \epsilon)$. By Lemma 8 there exist a $t_2(\epsilon) \geq t_1$ such that

$$\frac{1}{t} \sum_{\tau=1}^t \int \varphi_a(z, \theta_k) \bar{P}_{\sigma_\tau}(dz) - \int \varphi_a(z, \theta_k) \bar{P}_\sigma(dz) \leq 0.25\epsilon$$

for all $t \geq t_2(\epsilon)$ and all $k = 1, \dots, K$. Thus

$$\begin{aligned} \sup_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \frac{1}{t} \sum_{\tau=1}^t \ln q_{\theta}(Y_{\tau} | S_{\tau}, X_{\tau}) &\leq \max_{k=1, \dots, K} \left\{ \int \ln q_{\theta_k}(y | s, x) \bar{P}_{\sigma}(dz) \right\} \\ &+ \max_{k=1, \dots, K} \left\{ \int \varphi_a(z, \theta_k) \bar{P}_{\sigma}(dz) - \int \ln q_{\theta_k}(y | s, x) \bar{P}_{\sigma}(dz) \right\} + 0.75\epsilon \end{aligned}$$

for all $t \geq t_2(\epsilon)$. By Lemma 9, and our choice of a ,

$$\int \varphi_a(z, \theta_k) \bar{P}_{\sigma}^i(dz) - \int \ln q_{\theta_k}(y | s, x) \bar{P}_{\sigma}(dz) \leq 0.25\epsilon.$$

And thus

$$\sup_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \frac{1}{t} \sum_{\tau=1}^t \ln q_{\theta}(Y_{\tau} | S_{\tau}, X_{\tau}) \leq \max_{k=1, \dots, K} \left\{ \int \ln q_{\theta_k}(y | s, x) \bar{P}_{\sigma}(dz) \right\} + \epsilon.$$

Since $\{\theta_1, \dots, \theta_K\} \subseteq \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$ it follows that

$$\sup_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \frac{1}{t} \sum_{\tau=1}^t \ln q_{\theta}(Y_{\tau} | S_{\tau}, X_{\tau}) \leq \sup_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \int \ln q_{\theta}(y | s, x) \bar{P}_{\sigma}(dz) + \epsilon.$$

for all $t \geq t_2(\epsilon)$.

This result readily implies that:

$$\frac{1}{t} \sum_{\tau=1}^t (-\ln q_{\theta}(Y_{\tau} | S_{\tau}, X_{\tau})) \geq \inf_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \int (-\ln q_{\theta}(y | s, x)) \bar{P}_{\sigma}(dz) - \epsilon$$

for all $t \geq t_2(\epsilon)$ and all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$.

STEP 3. By Lemma 10: There exists a $t^o(\epsilon)$ such that

$$\left| \frac{1}{t} \sum_{\tau=1}^t \ln q_{\sigma_{\tau}}(Y_{\tau} | S_{\tau}, X_{\tau}) - \int \ln q_{\sigma}(y | s, x) \bar{P}_{\sigma}(dz) \right| \leq \epsilon$$

for all $t \geq t^o(\epsilon)$. Therefore, by this and the result in step 2:

$$\frac{1}{t} \sum_{\tau=1}^t \left(\ln \frac{q_{\sigma_{\tau}}(Y_{\tau} | S_{\tau}, X_{\tau})}{q_{\theta}(Y_{\tau} | S_{\tau}, X_{\tau})} \right) \geq \inf_{\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}} \int \left(\ln \frac{q_{\sigma}(y | s, x)}{q_{\theta}(y | s, x)} \right) \bar{P}_{\sigma}(dz) - 2\epsilon$$

for all $t \geq t(\epsilon) \equiv \max\{t_2(\epsilon), t^o(\epsilon)\}$ and all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$. \square

F.1 Proofs of Supplementary Lemmas

In order to prove the intermediate lemmas, we need the following claims:

Claim 1. K^i is (jointly) continuous and finite: Fix any $(\theta_n^i)_n$ and $(\sigma_n)_n$ such that $\lim_{n \rightarrow \infty} \theta_n^i = \theta^i$, $\lim_{n \rightarrow \infty} \sigma_n = \sigma$. Then $\lim_{n \rightarrow \infty} K^i(\sigma_n, \theta_n^i) = K(\sigma, \theta^i)$.

Proof. We first show that $K^i(\sigma, \theta) = \int_{\mathbb{Z}^i} \bar{P}_\sigma^i(dz^i) \ln q_\sigma^i(z^i) - \int_{\mathbb{Z}^i} \bar{P}_\sigma^i(dz^i) \ln q_{\theta^i}^i(z^i) < \infty$. Observe that for any $\theta \in \Theta^i$ and σ ,

$$\int \bar{P}_\sigma^i(dz^i) \ln q_{\theta^i}^i(z^i) = \sum_{s^i, x^i} \sum_{x^{-i}} \int_{\Omega} (-\ln q_\theta^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i)) \mu_{\sigma^{-i}}(x^{-i} | \omega) P_\Omega(d\omega | s^i) \sigma^i(x^i | s^i) p_{S^i}(s^i)$$

By condition 2(ii) this is clearly finite. Finally, observe that

$$\int_{\mathbb{Z}^i} \bar{P}_\sigma^i(dz^i) \ln q_{\theta^i}^i(z^i) = \sum_{x^i, s^i} \left\{ \int_{\mathbb{Y}^i} q_\sigma^i(y^i | s^i, x^i) \ln q_\sigma^i(y^i | s^i, x^i) dy^i \right\} \sigma_n^i(x^i | s^i) p_{S^i}(s^i),$$

which by condition 1(ii)(iii) is finite.

We now show continuity of K^i ,

$$\begin{aligned} K^i(\sigma_n, \theta_n^i) - K(\sigma, \theta^i) &= \int_{\mathbb{Z}^i} (\bar{P}_{\sigma_n}^i(dz^i) \ln q_{\sigma_n}^i(z^i) - \bar{P}_\sigma^i(dz^i) \ln q_\sigma^i(z^i)) \\ &\quad + \int_{\mathbb{Z}^i} (\bar{P}_\sigma^i(dz^i) \ln q_{\theta_n^i}^i(z^i) - \bar{P}_\sigma^i(dz^i) \ln q_{\theta^i}^i(z^i)). \end{aligned} \quad (61)$$

The first term in the RHS converges to zero by by analogous steps to those in the proof of Claim 2.

The proof concludes by showing that the second term in the RHS of (61) is such that:

$$\lim_{n \rightarrow \infty} \int \bar{P}_{\sigma_n}^i(dz^i) \ln q_{\theta_n^i}^i(z^i) = \int \bar{P}_\sigma^i(dz^i) \ln q_{\theta^i}^i(z^i). \quad (62)$$

Observe that for any $\theta \in \Theta^i$ and σ ,

$$\int \bar{P}_\sigma^i(dz^i) \ln q_{\theta^i}^i(z^i) = \sum_{s^i, x^i} \sum_{x^{-i}} \int_{\Omega} (-\ln q_\theta^i(f^i(\omega, x^i, x^{-i}) | s^i, x^i)) \mu_{\sigma^{-i}}(x^{-i} | \omega) P_\Omega(d\omega | s^i) \sigma^i(x^i | s^i) p_{S^i}(s^i).$$

Since $\theta_n^i \rightarrow \theta^i$, then eventually $\theta_n^i \in B(\theta^i)$ (where $B(\theta^i)$ is that of condition 1(ii)). Thus, by condition 1(ii) and the Dominated Convergence Theorem (DCT),

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{\Omega} (-\ln q_{\theta}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)) \mu_{\sigma^{-i}}(x^{-i}|\omega) P_{\Omega}(d\omega|s^i) \\ &= \int_{\Omega} \lim_{n \rightarrow \infty} (-\ln q_{\theta_n^i}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)) \mu_{\sigma^{-i}}(x^{-i}|\omega) P_{\Omega}(d\omega|s^i). \end{aligned}$$

We note that $\theta \mapsto -\ln q_{\theta}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)$ is continuous a.s.- P_{Ω} by condition 1(i) (by this condition continuity is a.s.-Lebesgue, but since P_{Ω} is absolutely continuous with respect to Lebesgue, continuity also holds a.s.- P_{Ω}) and it is also finite a.s.- P_{Ω} . Thus, it follows that the RHS of the previous display equals

$$\int_{\Omega} (-\ln q_{\theta}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)) \mu_{\sigma^{-i}}(x^{-i}|\omega) P_{\Omega}(d\omega|s^i).$$

Since $\#\mathbb{X} < \infty$ and $\#\mathbb{S}^i < \infty$ the desired result follows. \square

Claim 2. Fix any $\theta \in \Theta^i$ and $\sigma \in \Sigma$ and $(\sigma_n)_n$ with $\lim_{n \rightarrow \infty} \sigma_n = \sigma$. Then $\lim_{n \rightarrow \infty} K^i(\sigma_n, \theta) = K^i(\sigma, \theta)$.

Proof. Note that

$$\begin{aligned} K^i(\sigma_n, \theta^i) - K(\sigma, \theta^i) &= \int_{\mathbb{Z}^i} (\bar{P}_{\sigma_n}^i(dz^i) \ln q_{\sigma_n}^i(z^i) - \bar{P}_{\sigma}^i(dz^i) \ln q_{\sigma}^i(z^i)) \\ &\quad + \int_{\mathbb{Z}^i} (\bar{P}_{\sigma}^i(dz^i) - \bar{P}_{\sigma_n}^i(dz^i)) \ln q_{\theta^i}^i(z^i). \end{aligned} \tag{63}$$

The first term in the RHS of (63) can be written as

$$\begin{aligned} & \sum_{x^i, s^i} \left\{ \int_{\mathbb{Y}^i} (q_{\sigma_n}^i(y^i|s^i, x^i) \ln q_{\sigma_n}^i(y^i|s^i, x^i) - q_{\sigma}^i(y^i|s^i, x^i) \ln q_{\sigma}^i(y^i|s^i, x^i)) dy^i \right\} \sigma_n^i(x^i | s^i) p_{S^i}(s^i) \\ & \quad + \sum_{x^i, s^i} \left\{ \int_{\mathbb{Y}^i} (q_{\sigma}^i(y^i|s^i, x^i) \ln q_{\sigma}^i(y^i|s^i, x^i)) dy^i \right\} \{ \sigma^i(x^i | s^i) - \sigma_n^i(x^i | s^i) \} p_{S^i}(s^i) \\ & \equiv T_{1,n} + T_{2,n} \end{aligned}$$

Note that $\sigma \mapsto q_{\sigma}^i(y^i|s^i, x^i)$ is continuous (condition 2(i)) and so is $x \ln x$, thus

$$\lim_{n \rightarrow \infty} q_{\sigma_n}^i(y^i|s^i, x^i) \ln q_{\sigma_n}^i(y^i|s^i, x^i) = q_{\sigma}^i(y^i|s^i, x^i) \ln q_{\sigma}^i(y^i|s^i, x^i)$$

a.s.-Lebesgue and for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$. By condition 2(ii) there exists a $y \mapsto \bar{l}(y)$ such that for any $\sigma \in \Sigma$, $|q_\sigma^i(\cdot|s^i, x^i) \ln q_\sigma^i(\cdot|s^i, x^i)| \leq \bar{l}(\cdot)$ a.s.-Lebesgue and for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$ and $\int_{\mathbb{Y}^i} \bar{l}(y) dy < \infty$. Then, by applying the Dominated Convergence Theorem (DCT) for each (s^i, x^i) , it follows that

$$\lim_{n \rightarrow \infty} \int_{\mathbb{Y}^i} (q_{\sigma_n}^i(y^i|s^i, x^i) \ln q_{\sigma_n}^i(y^i|s^i, x^i) - q_\sigma^i(y^i|s^i, x^i) \ln q_\sigma^i(y^i|s^i, x^i)) dy^i = 0.$$

Thus, $T_{1,n} \rightarrow 0$. It is easy to see that $|T_{2,n}| \leq \#\mathbb{X}^i \int \bar{L}^i(y) dy \|\sigma^i - \sigma_n^i\|$ and thus also vanishes as $n \rightarrow \infty$.

The second term in the RHS of (63) can be written as

$$\sum_{x^i, s^i} \left\{ \int_{\mathbb{Y}^i} \ln q_{\theta^i}^i(y^i|s^i, x^i) (q_{\sigma_n}^i(y^i|s^i, x^i) \sigma_n^i(x^i | s^i) - q_\sigma^i(y^i|s^i, x^i) \sigma^i(x^i | s^i)) dy^i \right\} p_{S^i}(s^i).$$

By condition 2(i) $\lim_{n \rightarrow \infty} q_{\sigma_n}^i(y^i|s^i, x^i) \sigma_n^i(y^i|s^i) = q_\sigma^i(y^i|s^i, x^i) \sigma^i(y^i|s^i)$ a.s.-Lebesgue and for all $(s^i, x^i) \in \mathbb{S}^i \times \mathbb{X}^i$. By condition 2(ii) $|q_\sigma^i(y|s^i, x^i) \sigma^i(x^i|s^i)| \leq \bar{l}(y)$ a.s.-Lebesgue and for all (s^i, x^i) . Also, since $\ln q_{\theta^i}^i(y^i|s^i, x^i) = \ln q_{\theta^i}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)$ is integrable (see remark below condition 1), it follows that $|\ln q_{\theta^i}^i(y^i|s^i, x^i) q_\sigma^i(y^i|s^i, x^i) \sigma^i(x^i|s^i)| \leq |\ln q_{\theta^i}^i(y^i|s^i, x^i)| \bar{l}(y^i)$ and is integrable. Thus by the DCT, the previous display converges to 0. \square

We now present the proofs of the lemmas

Proof. [Proof of Lemma 6] (i) φ_a can be viewed as a composition function $\theta \mapsto \varphi_a(z, \theta) = F_a \circ q_\theta(z)$ where $F_a(t) = \ln(t) 1\{t \geq a\} + a 1\{t < a\}$. Clearly F_a is continuous and constant (equal to a) in a neighborhood of 0; this and condition 1(i) imply that φ_a is continuous a.s.-Lebesgue and for all $(s, x) \in \mathbb{S}^i \times \mathbb{X}^i$; thus is continuous a.s.- \bar{P}_σ^i (because P_Ω is absolutely continuous w.r.t. Lebesgue).

(ii) Note that

$$\varphi_a(z^i, \theta) \leq |\ln q_\theta^i(y^i|s^i, x^i)| = |\ln q_\theta^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)| \leq \sup_{\theta \in \Theta} |\ln q_\theta^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)| \equiv \bar{q}^i(\omega, s^i, x).$$

We now show that $\bar{q}^i(\cdot, s^i, x)$ is P_Ω -integrable for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$. Take $(B(\theta))_{\theta \in \Theta}$ where $B(\theta)$ is as in condition 1(ii). By compactness of Θ , there exists a finite sub-cover

$B(\theta_1), \dots, B(\theta_K)$ for some $K < \infty$. Thus

$$\int_{\Omega} \bar{q}^i(\omega, s^i, x) P_{\Omega}(d\omega) \leq \sum_{k=1}^K \int_{\Omega} \sup_{\theta \in B(\theta_k)} |\ln q_{\theta}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)| P_{\Omega}(d\omega)$$

and by condition 1(ii) this is finite. Since $\#\mathbb{X} < \infty$, this holds uniformly in $x \in \mathbb{X}$.

Also,

$$\varphi_a(z^i, \theta) \geq \ln a$$

and thus, $\hat{\varphi}_a(\omega, s^i, x) \equiv \max\{\bar{q}^i(\omega, s^i, x), -\ln a\}$ is such that

$$|\varphi_a(z^i, \theta)| \leq \hat{\varphi}_a(\omega, s^i, x)$$

for $z^i = (f^i(\omega, x^i, x^{-i}), s^i, x^i)$. Note that by condition 1(ii) and 2(ii),

$$\sum_{x^{-i}} \int_{\Omega} \hat{\varphi}_a(\omega, s^i, x) \mu_{\sigma^{-i}}(x^{-i}|\omega) P_{\Omega}(d\omega|s^i) = \sum_{x^{-i}} \int_{\Omega} \max\{\bar{q}^i(\omega, s^i, x), -\ln a\} \mu_{\sigma^{-i}}(x^{-i}|\omega) P_{\Omega}(d\omega).$$

For $a > 1$, $\max\{\bar{q}^i(\omega, s^i, x), -\ln a\} \leq \max\{\bar{q}^i(\omega, s^i, x), 0\} = \bar{q}^i(\omega, s^i, x)$; and for $a \leq 1$, $\max\{\bar{q}^i(\omega, s^i, x), -\ln a\} \leq \bar{q}^i(\omega, s^i, x) - \ln a$. Note that $\int_{\Omega} (-\ln a) P_{\Omega}(d\omega) < \infty$ is finite for all $a > 0$. Thus by this fact and integrability of $\bar{q}^i(\cdot, s^i, x)$, it follows that $\sum_{x^{-i}} \int_{\Omega} \hat{\varphi}_a(\omega, s^i, x) \mu_{\sigma^{-i}}(x^{-i}|\omega) P_{\Omega}(d\omega|s^i)$ is finite.

Finally, for any $z^i = (y^i, s^i, x^i)$, let $\bar{\varphi}_a(y^i, s^i, x^i) = \hat{\varphi}_a(\omega, s^i, x)$ for some (ω, x^{-i}) such that $f^i(\omega, x^i, x^{-i}) = y^i$ (there is always at least one pair). Since $\hat{\varphi} \geq 0$, clearly

$$\int_{\mathbb{Z}^i} \bar{\varphi}_a(z^i) P_{\sigma}^i(dz^i) \leq \sum_{x^{-i}} \int_{\Omega} \hat{\varphi}_a(\omega, s^i, x) \mu_{\sigma^{-i}}(x^{-i}|\omega) P_{\Omega}(d\omega|s^i)$$

and thus the desired result follows.

(iii) We first note that condition 3 states that:

$$|q_{\theta}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i) - q_{\theta'}^i(f^i(\omega, x^i, x^{-i})|s^i, x^i)| \leq \varepsilon'$$

for all θ and θ' such that $\|\theta - \theta'\| \leq \delta(\varepsilon')$, a.s.- P_{Ω} and for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$. This readily implies that

$$|q_{\theta}^i(y^i|s^i, x^i) - q_{\theta'}^i(y^i|s^i, x^i)| \leq \varepsilon'$$

for all θ and θ' such that $\|\theta - \theta'\| \leq \delta(\varepsilon')$, a.s.- $\bar{P}_{\sigma_t}^i$ and for all $(s^i, x) \in \mathbb{S}^i \times \mathbb{X}$ and all t .

Consider any θ and θ' such that $\|\theta - \theta'\| \leq \delta(\varepsilon')$ where $\delta(\varepsilon')$ is as in condition 3 and $\varepsilon' = a\varepsilon$. If $q_\theta(y|s, x) < a$ and $q_{\theta'}(y|s, x) < a$, then clearly $|\varphi_a(z, \theta) - \varphi_a(z, \theta')| < \varepsilon$. If $q_\theta(y|s, x) \geq a$ and $q_{\theta'}(y|s, x) \geq a$ (and suppose WLOG, $q_\theta(y|s, x) > q_{\theta'}(y|s, x)$), then

$$|\varphi_a(z, \theta) - \varphi_a(z, \theta')| \leq \varphi_a(z, \theta) - \varphi_a(z, \theta') \leq \frac{1}{q_{\theta'}(y|s, x)} (q_\theta(y|s, x) - q_{\theta'}(y|s, x)) \leq \frac{1}{a} a\varepsilon = \varepsilon.$$

(because $\ln(t) - \ln(t') \leq \frac{1}{t'}(t - t')$). Finally, $q_{\theta'}(y|s, x) < a$ and $q_\theta(y|s, x) \geq a$ (the reverse case is analogous), then

$$|\varphi_a(z, \theta) - \varphi_a(z, \theta')| \leq \varphi_a(z, \theta) - \varphi_a(z, \theta') \leq \frac{1}{a} (q_{\theta'}(y|s, x) - q_\theta(y|s, x)) \leq \frac{1}{a} a\varepsilon = \varepsilon$$

because $\varphi_a(z, \theta) = \ln q_\theta(y|s, x) \leq \ln a + \frac{1}{a}(q_\theta(y|s, x) - a)$ and $a = \varphi_a(z, \theta')$. Thus, the desired result follows from condition 3 and $\delta'(\varepsilon, a) \equiv \delta(a\varepsilon)$.

(iv) From our previous calculations $|\varphi_a(z^i, \theta)| \leq \max\{|\ln q_\theta^i(y^i|s^i, x^i)|, -\ln a\}$. Hence, it suffices to show that $\max\{|\ln q_\theta^i(\cdot|s^i, x^i)|^2, (\ln a)^2\}$ is integrable. This follows from similar algebra to the one in part (i) and condition 1(iii). \square

Proof. [Proof of Lemma 9] Clearly, $\{\theta \in \Theta \mid d(\sigma, \theta) \geq \varepsilon\}$ is compact. Let $G_a(\theta) \equiv \int \varphi_a(z, \theta) Q_\sigma(dz)$. Observe that $G_0(\theta) = \int \ln q_\theta(y|s, x) Q_\sigma(dz)$.

Note that $\varphi_a(z, \theta) \leq \varphi_{a'}(z, \theta)$ for all $a \leq a'$ and all (z, θ) , and thus $G_a(\theta) \leq G_{a'}(\theta)$.

We now show that G_0 is lower semi-continuous over $\{\theta \in \Theta \mid d(\sigma, \theta) \geq \varepsilon\}$, i.e.,

$$\liminf_{\theta \rightarrow \theta_0} G_0(\theta) \geq \int \liminf_{\theta \rightarrow \theta_0} \ln q_\theta(y|s, x) Q_\sigma(dz) = \int \ln q_{\theta_0}(y|s, x) Q_\sigma(dz).$$

This (in fact, continuity) follows from Claim 1.

G_a is also continuous for $a > 0$. This follows from the fact that $\theta \mapsto \varphi_a(z^i, \theta)$ is continuous by Lemma 6(i)(ii) and the DCT.

For any $\theta \in \Theta$,

$$\begin{aligned} G_a(\theta) &= \int \varphi_a(z, \theta) Q_\sigma(dz) = \int \ln q_\theta(y|s, x) 1\{q_\theta(y|s, x) \geq a\} Q_\sigma(dz) + \ln a \int 1\{q_\theta(y|s, x) \leq a\} Q_\sigma(dz) \\ &\leq \int \ln q_\theta(y|s, x) 1\{q_\theta(y|s, x) \geq a\} Q_\sigma(dz) \end{aligned}$$

because, for $a < 1$, $\ln a < 0$ and $\int 1\{q_\theta(y|s, x) \leq a\}Q_\sigma(dz) \geq 0$. Thus

$$\begin{aligned} \lim_{a \rightarrow 0} G_a(\theta) &\leq \lim_{a \rightarrow 0} \int \ln q_\theta(y|s, x) 1\{q_\theta(y|s, x) \geq a\} Q_\sigma(dz) \\ &= \int \lim_{a \rightarrow 0} \ln q_\theta(y|s, x) 1\{q_\theta(y|s, x) \geq a\} Q_\sigma(dz) \\ &= G_0, \end{aligned}$$

where the second line follows from the DCT and we can use this theorem because $|\ln q_\theta(y|s, x) 1\{q_\theta(y|s, x) \geq a\}| \leq |\ln q_\theta(y|s, x)|$ and this is integrable (see the remark below condition 1).

Since $G_a(\theta) \geq G_0(\theta)$ this shows that (*pointwise in*) θ , $G_a(\theta)$ converges to $G_0(\theta)$ as $a \rightarrow 0$. By Dini's theorem (Dudley [2002], Theorem 2.4.10), this convergence is uniform. That is, for any $\epsilon > 0$, there exists a $a(\epsilon)$ such that, for all $0 < a \leq a(\epsilon)$

$$\left| \int \varphi_a(z, \theta) Q_\sigma(dz) - \int \ln q_\theta(y|s, x) Q_\sigma(dz) \right| \leq \epsilon$$

for all $\theta \in \{\theta \in \Theta \mid d(\sigma, \theta) \geq \epsilon\}$. □

Proof. [Proof of Lemma 8] Fix any $\epsilon > 0$ and $a > 0$. Let

$$\begin{aligned} &\int \varphi_a(z, \theta) \bar{P}_{\sigma_t}(dz) - \int \varphi_a(z, \theta) \bar{P}_\sigma(dz) \\ &= \sum_{x,s} \int_{\mathbb{Y}} \varphi_a(y, s, x, \theta) Q_\sigma(dy|s, x) \{\sigma_t(x|s) - \sigma(x|s)\} p_S(s) \\ &- \sum_{x,s} \int_{\mathbb{Y}} \varphi_a(y, s, x, \theta) \{Q_\sigma(dy|s, x) - Q_{\sigma_t}(dy|s, x)\} \sigma_t(x|s) p_S(s) \\ &= T_{1t}(\theta) + T_{2t}(\theta). \end{aligned}$$

Regarding T_1 note that

$$\begin{aligned}
|T_{1t}(\theta)| &\leq \|\sigma_t - \sigma\| \sum_{x,s} \int_{\mathbb{Y}} |\varphi_a(y, s, x, \theta)| Q_\sigma(dy|s, x) p_S(s) \\
&\leq \|\sigma_t - \sigma\| \sum_{x,s} \int_{\mathbb{Y}} \bar{\varphi}_a(y, s, x) Q_\sigma(dy|s, x) p_S(s) \\
&\leq \|\sigma_t - \sigma\| C(a)
\end{aligned}$$

where the second and third lines follow from Lemma 6; with $C(a)$ some constant which might depend on a . We know that $\|\sigma_t - \sigma\| \rightarrow 0$, thus $T_{1t} \rightarrow 0$. The fact that $\int \varphi_a(z, \theta) \bar{P}_{\sigma_t}(dz)$ converges to $\int \varphi_a(z, \theta) \bar{P}_\sigma(dz)$, implies that $\frac{1}{t} \sum_{\tau=1}^t \int \varphi_a(z, \theta) \bar{P}_{\sigma_\tau}(dz)$ does too.

It suffices to show that $T_{2t}(\theta) \leq \epsilon$ for t sufficiently large; *pointwise* in $\theta \in \Theta$. First observe that

$$|T_{2t}(\theta)| \leq \sum_{x^i, s^i} \left| \int_{\mathbb{Y}} \varphi_a(y, s^i, x^i, \theta) \{Q_\sigma(dy|s^i, x^i) - Q_{\sigma_t}(dy|s^i, x^i)\} \right|$$

since actions and signals belong to a finite discrete set, it suffices to show that the RHS vanishes pointwise in s^i, x^i . Observe that

$$\int_{\mathbb{Y}} \varphi_a(y^i, s^i, x^i, \theta) Q_\sigma(dy^i|s^i, x^i) = \sum_{x^{-i}} \int_{\Omega} \varphi_a(f(\omega, x^i, x^{-i}), s^i, x^i, \theta) \mu_{\sigma^{-i}}(x^{-i}|\omega) P_\Omega(d\omega|s^i).$$

Thus

$$\begin{aligned}
|T_{2t}(\theta)| &\leq \left| \sum_{x^{-i}} \int_{\Omega} \varphi_a(f(\omega, x^i, x^{-i}), s^i, x^i, \theta) \{ \mu_{\sigma^{-i}}(x^{-i}|\omega) - \mu_{\sigma_t^{-i}}(x^{-i}|\omega) \} P_\Omega(d\omega|s^i) \right| \\
&\leq \sum_{x^{-i}} \sup_{\omega \in \Omega} \left| \mu_{\sigma^{-i}}(x^{-i}|\omega) - \mu_{\sigma_t^{-i}}(x^{-i}|\omega) \right| \int_{\Omega} |\varphi_a(f(\omega, x^i, x^{-i}), s^i, x^i, \theta)| P_\Omega(d\omega|s^i) \\
&\leq \sum_{x^{-i}} \left| \sum_{s^{-i}} \left\{ \prod_{j \neq i} \sigma^j(x^j | s^j) - \prod_{j \neq i} \sigma_t^j(x^j | s^j) \right\} \right| \int_{\Omega} |\varphi_a(f(\omega, x^i, x^{-i}), s^i, x^i, \theta)| P_\Omega(d\omega|s^i).
\end{aligned}$$

Clearly $\left| \sum_{s^{-i}} \left\{ \prod_{j \neq i} \sigma^j(x^j | s^j) - \prod_{j \neq i} \sigma_t^j(x^j | s^j) \right\} \right| \rightarrow 0$ for all x^{-i} and thus it suffices to show that $\int_{\Omega} |\varphi_a(f(\omega, x^i, x^{-i}), s^i, x^i, \theta)| P_\Omega(d\omega|s^i)$ is finite. By the proof of Lemma 6(ii), this quantity is bounded above by $\max\{\bar{q}^i(\omega, s^i, x), -\ln a\}$ which is integrable.

□

Proof. [Proof of Lemma 7] Let $\ell_\tau(\theta) \equiv \varphi_a(Z_\tau, \theta) - \int \varphi_a(z, \theta) \bar{P}_{\sigma_\tau}(dz)$. And let $L_t(h, \theta) = \sum_{\tau=1}^t \tau^{-1} \ell_\tau(\theta)$. It suffices to show that $L_t(h, \theta)$ converges a.s. to an integrable $L_\infty(h, \theta)$. By following the same steps as those in the proof of Lemma 4, it suffices to show that $E_{\mathbf{P}^{\mu_0, \phi}} [(\varphi_a(Z_\tau, \theta))^2] < \infty$. This follows from Lemma 6(iv). □

G Global stability: Example 2.1 (monopoly with unknown demand).

Theorem 3 says that all Berk-Nash equilibria can be approached with probability 1 provided we allow for vanishing optimization mistakes. In this appendix, we illustrate how to use the techniques of stochastic approximation theory to establish stability of equilibria under the assumption that players make no optimization mistakes. We present the explicit learning dynamics for the monopolist with unknown demand, Example 2.1, and show that the unique equilibrium in this example is globally stable. The intuition behind global stability is that switching from the equilibrium strategy to a strategy that puts more weight on a price of 2 changes beliefs in a way that makes the monopoly want to put less weight on a price of 2, and similarly for a deviation to a price of 10.

We first construct a perturbed version of the game. Then we show that the learning problem is characterized by a nonlinear stochastic system of difference equations and employ stochastic approximation methods for studying the asymptotic behavior of such system. Finally, we take the payoff perturbations to zero.

In order to simplify the exposition and thus better illustrate the mechanism driving the dynamics, we modify the subjective model slightly. We assume the monopolist only learns about the parameter $b \in \mathbb{R}$; i.e., her beliefs about parameter a are degenerate at a point $a = 32 \neq a^0$ and thus are never updated. Therefore, beliefs μ are probability distributions over \mathbb{R} , i.e., $\mu \in \Delta(\mathbb{R})$.

PERTURBED GAME. Let ξ be a real-valued random variable distributed according to P_ξ ; we use F to denote the associated cdf and f the pdf. The perturbed payoffs are

given by $yx - \xi 1\{x = 10\}$. Thus, given beliefs $\mu \in \Delta(\mathbb{R})$, the probability of optimally playing $x = 10$ is

$$\sigma(\mu) = F(8a - 96E_\mu[B]).$$

Note that the only aspect of μ that matters for the decision of the monopolist is $E_\mu[B]$. Thus, letting $m = E_\mu[B]$ and slightly abusing notation, we use $\sigma(\mu) = \sigma(m)$ as the optimal strategy.

BAYESIAN UPDATING. We now derive the Bayesian updating procedure. We assume that the prior μ_0 is given by a Gaussian distribution with mean and variance m_0, τ_0^2 .⁵²

We now show that, given a realization (y, x) and a prior $N(m, \tau^2)$, the posterior is also Gaussian with a mean and variance specified below. To do this, note that the posterior is given by (up to omitted constants)

$$\frac{e^{-0.5(y-a+bx)^2} e^{-0.5\frac{(b-m)^2}{\tau^2}}}{\int e^{-0.5(y-a+bx)^2} e^{-0.5\frac{(b-m)^2}{\tau^2}} db}.$$

After tedious but straightforward algebra one can show that this expression equals (up to omitted constants)

$$\exp \left\{ -0.5 \frac{\left(b - \frac{-(y-a)x + m\tau^{-2}}{(x^2 + \tau^{-2})} \right)^2}{(x^2 + \tau^{-2})^{-1}} \right\}.$$

Thus, the posterior is also Gaussian with mean $\frac{-(y-a)x + m\tau^{-2}}{(x^2 + \tau^{-2})}$ and variance $(x^2 + \tau^{-2})^{-1}$. Hence, in order to keep track of the evolution of the sequence of posteriors $(\mu_t)_t$, it suffices to keep track of the mean and variance, which evolve according to the following

⁵²This choice of prior is standard in Gaussian settings like ours. As shown below this choice simplifies the exposition considerably.

law of motion:

$$\begin{aligned}
m_{t+1} &= \frac{-(Y_{t+1} - a)X_{t+1} + m_t\tau_t^{-2}}{(X_{t+1}^2 + \tau_t^{-2})} \\
&= m_t \left(\frac{\tau_t^{-2}}{(X_{t+1}^2 + \tau_t^{-2})} \right) + \frac{-(Y_{t+1} - a)}{X_{t+1}} \left(\frac{X_{t+1}^2}{(X_{t+1}^2 + \tau_t^{-2})} \right) \\
&= m_t + \left(\frac{-(Y_{t+1} - a)}{X_{t+1}} - m_t \right) \left(\frac{X_{t+1}^2}{X_{t+1}^2 + \tau_t^{-2}} \right),
\end{aligned}$$

and

$$\tau_{t+1}^2 = \frac{1}{(X_{t+1}^2 + \tau_t^{-2})}.$$

NONLINEAR STOCHASTIC DIFFERENCE EQUATIONS AND STOCHASTIC APPROXIMATION.

As shown below, the Bayesian updating is fully characterized by the law of motion of $(m_t)_t$ and $(\tau_t^2)_t$. We now cast this law of motion as a nonlinear system of stochastic difference equations to which we can apply results from stochastic approximation theory.

It is convenient to define $r_{t+1} \equiv \frac{1}{t+1} (\tau_t^{-2} + X_{t+1}^2)$ and keep track of this variable as opposed to τ_{t+1}^2 . Note that, $r_{t+1} = \frac{1}{t+1} \sum_{s=1}^{t+1} x_s^2 + \frac{1}{t+1} \tau_0^{-2} = \frac{x_{t+1}^2}{t+1} + \frac{t}{t+1} r_t$. Therefore,

$$m_{t+1} = m_t + \frac{1}{t+1} \frac{X_{t+1}^2}{r_{t+1}} \left(\frac{-(Y_{t+1} - a)}{X_{t+1}} - m_t \right)$$

and

$$r_{t+1} = r_t + \frac{1}{t+1} (X_{t+1}^2 - r_t).$$

Let $\beta_t = (m_t, r_t)'$, $Z_t = (X_t, Y_t)$,

$$G(\beta_t, z_{t+1}) = \begin{bmatrix} \frac{x_{t+1}^2}{r_{t+1}} \left(\frac{-(y_{t+1} - a)}{x_{t+1}} - m_t \right) \\ (x_{t+1}^2 - r_t) \end{bmatrix}$$

and

$$\begin{aligned}
\mathbb{G}(\beta) &= \begin{bmatrix} \mathbb{G}_1(\beta) \\ \mathbb{G}_2(\beta) \end{bmatrix} = E_{P_\sigma} [G(\beta, Z_{t+1})] \\
&= \begin{bmatrix} F(8a - 96m) \frac{100}{r} \left(\frac{-(a_0 - a - b_0 10)}{10} - m \right) + (1 - F(8a - 96m)) \frac{4}{r} \left(\frac{-(a_0 - a - b_0 2)}{2} - m \right) \\ (4 + F(8a - 96m)96 - r) \end{bmatrix}
\end{aligned}$$

where P_σ is the probability over Z induced by σ (and $y = a^0 - b^0x + \omega$).

Therefore, the dynamical system can be cast as

$$\beta_{t+1} = \beta_t + \frac{1}{t+1}\mathbb{G}(\beta_t) + \frac{1}{t+1}V_{t+1}$$

with

$$V_{t+1} = G(\beta_t, Z_{t+1}) - \mathbb{G}(\beta_t).$$

Stochastic approximation theory (e.g., Kushner and Yin [2003]) implies, roughly speaking, that in order to study the asymptotic behavior of $(\beta_t)_t$ it is enough to study the behavior of the orbits of the following ODE

$$\dot{\beta}(t) = \mathbb{G}(\beta(t)).$$

CHARACTERIZATION OF THE STEADY STATES. In order to find the steady states of $(\beta_t)_t$, it is enough to find β^* such that $\mathbb{G}(\beta^*) = 0$. Let $H_1(m) \equiv F(8a-96m)10(-(a_0-a) + (b_0-m)10) + (1-F(8a-96m))2(-(a_0-a) + (b_0-m)2)$. Observe that $\mathbb{G}_1(\beta) = r^{-1}H_1(m)$ and that H_1 is continuous and $\lim_{m \rightarrow -\infty} H_1(m) = \infty$ and $\lim_{m \rightarrow \infty} H_1(m) = -\infty$. Thus, there exists at least one solution $H_1(m) = 0$. Therefore, there exists at least one β^* such that $\mathbb{G}(\beta^*) = 0$.

Let $\bar{b} = b_0 - \frac{a_0-a}{10} = 5 - \frac{4}{5} = \frac{21}{5}$ and $\underline{b} = b_0 - \frac{a_0-a}{2} = 1$, $\bar{r} = 4 + F(8a-96\underline{b})96$ and $\underline{r} = 4 + F(8a-96\bar{b})96$, and $\mathbb{B} \equiv [\underline{b}, \bar{b}] \times [\underline{r}, \bar{r}]$. It follows that $H_1(m) < 0$ for all $m > \bar{b}$, and thus m^* must be such that $m^* \leq \bar{b}$. It is also easy to see that $m^* \geq \underline{b}$. Moreover,

$$\begin{aligned} \frac{dH_1(m)}{dm} &= -96f(8a-96m)10(-(a_0-a) + (b_0-m)10) \\ &\quad + 96f(8a-96m)2(-(a_0-a) + (b_0-m)2) - F(8a-96m)100 - (1-F(8a-96m))4 \\ &= 96f(8a-96m) \{(-2(a_0-a) + (b_0-m)4) - (-10(a_0-a) + (b_0-m)100)\} \\ &\quad - 4 - 96F(8a-96m) \\ &= 96f(8a-96m) \{8(a_0-a) - (b_0-m)96\} - 4 - 96F(8a-96m). \end{aligned}$$

Thus, for any $m \leq \bar{b}$, $\frac{dH_1(m)}{dm} < 0$, because $m \leq \bar{b}$ implies $8(a_0-a) \leq (b_0-m)80 < (b_0-m)96$.

Therefore, on the relevant domain $m \in [\underline{b}, \bar{b}]$, H_1 is decreasing, thus implying that there exists only one m^* such that $H_1(m^*) = 0$. Therefore, there exists only one β^*

such that $\mathbb{G}(\beta^*) = 0$.

We are now interested in characterizing the limit of β^* as the perturbation vanishes, i.e. as F converges to $1\{\xi \geq 0\}$. To do this we introduce some notation. We consider a sequence $(F_n)_n$ that converges to $1\{\xi \geq 0\}$ and use β_n^* to denote the steady state associated to F_n . Finally, we use H_1^n to denote the H_1 associated to F_n .

We proceed as follows. First note that since $\beta_n^* \in \mathbb{B}$ for all n , the limit exists (going to a subsequence if needed). We show that $m^* \equiv \lim_{n \rightarrow \infty} m_n^* = \frac{8a}{96} = \frac{8}{3}$. Suppose not, in particular, suppose that $\lim_{n \rightarrow \infty} m_n^* < \frac{8a}{96} = \frac{8}{3}$ (the argument for the reverse inequality is analogous and thus omitted). In this case $\lim_{n \rightarrow \infty} 8a - 96m_n^* > 0$, and thus $\lim_{n \rightarrow \infty} F_n(8a - 96m_n^*) = 1$. Therefore

$$\lim_{n \rightarrow \infty} H_1^n(\beta_n^*) = 10(-a_0 - a) + (b_0 - m^*)10 = 10\left(-8 + \frac{70}{3}\right) > 0.$$

But this implies that there exists a N such that $H_1^n(\beta_n^*) > 0$ for all $n \geq N$ which is a contradiction.

Moreover, define $\sigma_n^* = F_n(8a - 96m_n^*)$ and $\sigma^* = \lim_{n \rightarrow \infty} \sigma_n^*$. Since $H_1^n(m_n^*) = 0$ for all n and $m^* = \frac{8}{3}$, it follows that

$$\sigma^* = \frac{-2\left(-8 + \left(5 - \frac{8}{3}\right)2\right)}{10\left(-8 + \left(5 - \frac{8}{3}\right)10\right) - 2\left(-8 + \left(5 - \frac{8}{3}\right)2\right)} = \frac{1}{24}.$$

GLOBAL CONVERGENCE TO THE STEADY STATE. In our example, it is in fact possible to establish that behavior converges with probability 1 to the unique equilibrium. By the results in Benaim [1999] Section 6.3, it is sufficient to establish the *global* asymptotic stability of β_n^* for any n , i.e., the basin of attraction of β_n^* is all of \mathbb{B} .

In order to do this let

$$L(\beta) = (\beta - \beta_n^*)' P (\beta - \beta_n^*)$$

for all β ; where $P \in \mathbb{R}^{2 \times 2}$ is positive definite and *diagonal* and will be determined

later. Note that $L(\beta) = 0$ iff $\beta = \beta_n^*$. Also

$$\begin{aligned}\frac{dL(\beta(t))}{dt} &= \nabla L(\beta(t))' \mathbb{G}(\beta(t)) \\ &= 2(\beta(t) - \beta_n^*)' P(\mathbb{G}(\beta(t))) \\ &= 2\{(m(t) - m_n^*) P_{[11]} \mathbb{G}_1(\beta(t)) + (r(t) - r_n^*) P_{[22]} \mathbb{G}_2(\beta(t))\}.\end{aligned}$$

Since $\mathbb{G}(\beta_n^*) = 0$,

$$\begin{aligned}\frac{dL(\beta(t))}{dt} &= 2(\beta(t) - \beta_n^*)' P(\mathbb{G}(\beta(t)) - \mathbb{G}(\beta_n^*)) \\ &= 2(m(t) - m_n^*) P_{[11]} (\mathbb{G}_1(\beta(t)) - \mathbb{G}_1(\beta_n^*)) \\ &\quad + 2(r(t) - r_n^*) P_{[22]} (\mathbb{G}_2(\beta(t)) - \mathbb{G}_2(\beta_n^*)) \\ &= 2(m(t) - m_n^*)^2 P_{[11]} \int_0^1 \frac{\partial \mathbb{G}_1(m_n^* + s(m(t) - m_n^*), r_n^*)}{\partial m} ds \\ &\quad + 2(r(t) - r_n^*)^2 P_{[22]} \int_0^1 \frac{\partial \mathbb{G}_2(m_n^*, r_n^* + s(r(t) - r_n^*))}{\partial r} ds\end{aligned}$$

where the last equality holds by the mean value theorem. Note that $\frac{d\mathbb{G}_2(m_n^*, r_n^* + s(r(t) - r_n^*))}{dr} = -1$ and $\int_0^1 \frac{d\mathbb{G}_1(m_n^* + s(m(t) - m_n^*), r_n^*)}{dm} ds = \int_0^1 (r_n^*)^{-1} \frac{dH_1(m_n^* + s(m(t) - m_n^*))}{dm} ds$. Since $r(t) > 0$ and $r_n^* \geq 0$ the first term is positive and we already established that $\frac{dH_1(m)}{dm} < 0$ for all m in the relevant domain. Thus, by choosing $P_{[11]} > 0$ and $P_{[22]} > 0$ it follows that $\frac{dL(\beta(t))}{dt} < 0$.

Therefore, we show that L satisfies the following properties: is strictly positive for all $\beta \neq \beta_n^*$ and $L(\beta_n^*) = 0$, and $\frac{dL(\beta(t))}{dt} < 0$. Thus, the function satisfies all the conditions of a Lyapounov function and, therefore, β_n^* is globally asymptotically stable for all n (see Hirsch et al. [2004] p. 194).