Dynamic Foundations for Empirical Static Games*

Niccolò Lomys† Lorenzo Magnolfi‡ Camilla Roncoroni§

January 31, 2020

Abstract

We develop an approach to identification and estimation of a game's primitives when players interact repeatedly in a one-shot incomplete information game and the only restriction on behavior is an asymptotic no regret (ANR) condition. This property requires that the time average of the counterfactual increase in past payoffs, had different actions been played, becomes approximately zero in the long run. Well-known algorithms for the repeated play of a one-shot incomplete information game satisfy the ANR property. Under the ANR assumption, we (partially) identify the structural parameters of the one-shot game. We establish our result in two steps. First, we prove that the empirical distribution of play that satisfies ANR converges to the set of Bayes (coarse) correlated equilibrium predictions of the underlying one-shot game. To do so, we generalize to incomplete information environments prior results on dynamic foundations for equilibrium play in static games of complete information. Second, we show how to use the limiting model to obtain consistent estimates of the parameters of interest. We apply our method to data on pricing behavior in an online platform.

Keywords: Empirical Games; Incomplete Information; Bayes (Coarse) Correlated Equilibrium; Learning in Games; no regret; Partial Identification; Incomplete Models; Robust Predictions.

JEL Classification: C57; C70; L10.

---

*We thank Dirk Bergemann for excellent feedback that greatly improved the paper. We had valuable conversations with Tilman Börgers, Giacomo Calzolari, Mira Frick, Ryota Iijima, Rida Laraki, Alessandro Lizzeri, Juan Ortner, Larry Samuelson, and Eran Shmaya. We received helpful comments from seminar audiences at the University of Mannheim, EEA–ESEM 2019, the 2019 North American Winter Meeting of the Econometric Society, and the 2018 Lisbon Meetings in Game Theory and Applications. The first author acknowledges funding from the European Research Council Grant Nº 714147, and from the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

†Toulouse School of Economics, University of Toulouse Capitole; niccolo.lomys@tse-fr.eu.

‡Department of Economics, University of Wisconsin-Madison; magnolfi@wisc.edu.

§Department of Economics, University of Warwick; c.roncoroni@warwick.ac.uk.

1 Introduction

When data are generated by strategic interaction, an external analyst needs to specify a solution concept to interpret players' outcomes and leverage the observables to perform identification of the game's primitives. It is common practice to assume that equilibrium play is observed. This assumption is justified whenever players are able to form correct expectations[1] on the strategic environment and to (optimally) behave accordingly. However, for many real-world strategic environments, it is not obvious that behavior satisfies these requirements.

In this paper, we explicitly allow for the possibility that players may need time to adapt to the strategic environment. Instead of assuming that players choose equilibrium strategies at each point in time, we only impose a behavioral assumption describing a minimal optimality condition for the long-run outcome of players' interaction. More specifically, we model players as playing repeatedly a one-shot incomplete information game and assume that the long-run outcome of their interaction satisfies a property of "asymptotic no regret" (hereafter, ANR). The ANR property requires that the time average of the counterfactual increase in past payoffs, had different actions been played, becomes arbitrarily small in the long run. Intuitively, no matter how players are adapting, we assume that they eventually eliminate the regret for not having played differently in the past. After having imposed that players' behavior satisfies the ANR property, we derive the implications for the identification of the game's primitives.

We show that, under the ANR assumption, it is possible to partially identify the structural parameters of the one-shot incomplete information game. We do so in two steps. First, we show that the empirical distribution (i.e. the time average) of actions, signals, and states that satisfies ANR converges to the set of Bayes correlated equilibria (Bergemann and Morris (2016)), or to its coarse analog—depending on the specific regret notion we consider—of the underlying one-shot game (although, in general, converge is not to a particular point in this set). To establish this property, we generalize to games of incomplete information prior results on dynamic foundations for equilibrium play in static games of complete information (Hart and Mas-Colell (2013) and references therein). Second, we show how to use the limiting model to obtain consistent estimates of the parameters of interest. Our empirical approach is based on a behavioral assumption that has only implications for the predictions of a limiting model, as opposed to the full data generating process. Our approach gives rise to non-standard econometric issues, as it is not possible to fully characterize a single limit distribution of the observables, but only the set it belongs to. Yet, we show that we can use the limiting model to obtain a consistent estimator for the parameters of interest. Since behavior is not specified, our model is incomplete in the sense of Tamer (2003).

The ANR property is satisfied by a large class of well-known algorithms for the repeated play of the underlying one-shot game, once they are appropriately extended to

[1]Given their information sets, which are assumed to be correctly defined by the analyst.

games of incomplete information. This class includes simple adaptive heuristics, fictitious-play-like dynamics, more sophisticated learning rules involving active experimentation, calibrated learning, and several equilibrium dynamics. Since we do not fully specify what the behavior of players is or what they do to play according to this minimal long-run requirement, we depart from the current literature on empirical dynamic games that typically imposes the Markov perfect (or related) solution concept (for a review, see Ackerberg, Benkard, Berry and Pakes (2007)).

The ANR property is weaker than the static no-ex post-regret property of pure-strategy Nash equilibrium that is sometimes invoked to motivate the choice of modeling cross-sectional data as equilibrium outcomes of a static game. Indeed, this descriptive (in the sense of Pakes (2016)) interpretation of static models is often paired with the assumptions of complete information and pure-strategy Nash equilibrium. The rationale for these assumptions is that the no-ex post-regret property of pure-strategy Nash equilibria reflects the stable nature of long-run outcomes.[2] Although appropriate for some environments, the static notion of no-ex post regret is a strong requirement: our work is thus complementary to standard equilibrium models of strategic interaction and provides an alternative whenever Nash equilibrium does not represent an appropriate restriction on behavior. In fact, Nash equilibrium of the static game is neither a natural long-run outcome of many simple game dynamics (for a review, see Hart and Mas-Colell (2013)), nor easy to compute in large games.

We apply our method to data on pricing behavior by sellers in an online platform. In this environment, a large population of sellers interacts strategically over time. These sellers may (but do not need to) display a limited degree of sophistication and rely on ANR algorithms when setting prices. The large number of sellers in this environment generates a severe curse of dimensionality, which makes it impossible to apply standard approaches in the estimation of games literature. We instead reformulate the sellers' problem as an aggregative game (Nocke and Schutz (2018)), whereby a seller's profits only depend on rivals' prices through an aggregate statistic. With this approach, we significantly reduce the dimensionality, and can thus proceed to recover the distribution of sellers' marginal costs. This primitive is fundamental to investigate market-design counterfactuals.

Related Literature. Our work is related to a large and growing literature at the intersection of economics and computer science studying regret learning and regret minimization. The regret learning framework was originally developed for single-agent decision problems. However, a series of papers have pointed out significant implications of regret-based learning procedures in strategic settings. Seminal game theory papers in which players try to minimize regret include Hannan (1957), Foster and Vohra (1997, 1998, 1999), and Hart and Mas-Colell (2000). In particular, our work is related to Hart and Mas-Colell (2000, 2013), whose convergence results we generalize to incomplete informa-

[2] For instance, Ciliberto and Tamer (2009) argue as follows: "The idea behind cross-section studies is that in each market, firms are in a long-run equilibrium."

tion environments. Hartline, Syrgkanis and Tardos (2015) and Caragiannis, Kaklamanis, Kanellopoulos, Kyropoulou, Lucier, Paes Leme and Tardos (2015) offer related theoretical results from the viewpoint of computer science. In contrast to all these authors, the emphasis of our work is on connecting learning dynamics to the inference problem of an external observer. Moreover, because of motivation, modeling, and technical differences, our setting is different from that studied by the computer science literature. Therefore, our analyses of regret-based learning procedures in settings with incomplete information are complementary, but not equivalent. We refer to Section 3 for further discussion.

Recent contributions in computer science also offers connection to empirical work (Nekipelov, Syrgkanis and Tardos (2015) and Nisan and Noti (2017a,b)). Nekipelov et al. (2015) are the first to suggest the regret-based approach to econometrics. In an online auctions environment, they characterize (and perform inference on) the set of valuations consistent with a given level of regret, but do not rely on an equilibrium concept. Nisan and Noti (2017a,b) evaluate a similar approach in experimental data and propose adjustments to the no regret estimation procedure. We also perform inference on the distribution of payoff types when the path of play has an asymptotic no regret property. In contrast to these papers, however, we do so by leveraging convergence results to interpret the data through the lens of the static equilibrium notion of Bayes correlated equilibrium or its coarse analog.

More broadly, regret learning and regret minimization is now the leading approach in online learning (see Shalev-Shwartz (2011)) and in multi-armed bandit problems (see Bubeck and Cesa-Bianchi (2012) and Lattimore and Szepesvári (2019)). These concepts are also a cornerstone in the algorithmic game theory literature (see Nisam, Roughgarden, Tardos and Vazirani (2007), Shoam and Leyton-Brown (2008), and Roughgarden (2016)). An excellent discussion of regret minimization from the online machine learning perspective can be found in Cesa-Bianchi and Lugosi (2006). Regret minimization has also been used in designing treatment rules (e.g., Manski (2004) and Stoye (2009)) and in forecast aggregation (e.g., Arieli, Babichenko and Smorodinsky (2018) and Babichenko and Garber (2018)).

We are not the first to leverage on results in the literature on learning in games to perform empirical analysis. Lee and Pakes (2009) develop a learning-based procedure to compute counterfactuals in dynamic games. Several recent advances in the estimation of dynamic games investigate tractable and less restrictive empirical models (e.g., Doraszelski, Lewis and Pakes (2018)). Our paper proposes a valid descriptive approach that is complementary to these structural methods.

We also develop a novel approach to deal with the curse of dimensionality inherent in the estimation of large games. In our application, we reformulate the game in an aggregative form as in Nocke and Schutz (2018). Hence, players' payoffs only depend on their competitors' actions via an aggregate statistic. We thus achieve a dramatic reduction in the dimension of the (Bayes correlated) equilibrium distribution, without redefining the equilibrium concept as in Weintraub, Benkard and Van Roy (2008).

Magnolfi and Roncoroni (2019), Syrgkanis, Tamer and Ziani (2018), and Gualdani and Sinha (2019) also consider estimation under the assumption of Bayes correlated equilibrium behavior for models of discrete games, auctions, and single-agent decision problems respectively. Although this paper proposes the use of a similar estimation technique, the motivation is very different. In fact, Magnolfi and Roncoroni (2019), Syrgkanis et al. (2018) and Gualdani and Sinha (2019) exploit the link between equilibrium behavior and information to establish that Bayes correlated equilibrium allows to estimate games under weak assumptions on information. In this paper, instead, we motivate estimation under Bayes correlated equilibrium when the data are generated by repeated interaction with weak behavioral assumptions.

Road Map. In Section 2, we present the theoretical framework and formalize the notions of regret and asymptotic ε-regret for the repeated play of a one-shot incomplete information game. In Section 3, we study the convergence properties of ε-regret dynamics to the set of Bayes (coarse) correlated ε-equilibria of the one-shot game. In Section 4, we specialize the theoretical model to study what features of the underlying economic environment we can empirically recover, and how, under weak assumptions on behavior when the one-shot game is played repeatedly over time. In Section 5, we present our empirical application and outline our aggregative approach to sidestep the curse of dimensionality at the estimation stage. In Section 6, we further discuss our main theoretical and econometric results and present extensions and robustness checks. In Section 7, we conclude. In Appendix A, we discuss regret-minimizing algorithms for the repeated play of the one-shot incomplete information game. Omitted proofs are in Appendix B.

2 Model

2.1 Basic Setup

One-Shot Game. There is a finite set $\mathcal{I}$ of I players, $\mathcal{I} \coloneqq \{1, \dots, I\}$, and a finite set of payoff states, Θ. We write i for a typical player and θ for a typical state. A *basic game* G consists of: (i) for each player i, a finite set of actions A_i, where we define $A \coloneqq A_1 \times \cdots \times A_I$, and a payoff function $u_i \colon A \times \Theta \to \mathbb{R}$; and (ii) a full-support common prior $\psi \in \Delta_{++}(\Theta)$. Thus, $G \coloneqq \big((A_i, u_i)_{i=1}^I, \psi\big)$. An *information structure* S consists of: (i) for each player i, a finite set of signals (or types) T_i, where we define $T \coloneqq T_1 \times \cdots \times T_I$; and (ii) a full-support signal distribution $\pi \colon \Theta \to \Delta_{++}(T)$. Thus, $S \coloneqq \big((T_i)_{i=1}^I, \pi\big)$. Together, the pair (G, S) defines an *incomplete information game*. If Θ is a singleton, the game is one of *complete information*. We denote by a_{-i} a profile of actions for players other than i, i.e. $a_{-i} \coloneqq (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_I)$. Analogously, t_{-i} denotes a profile of signals for players other than i. We refer to (G, S) as the *one-shot game*.

Repeated Game. Hereafter, we assume that the one-shot game is played repeatedly over time. Time is discrete and periods are indexed by $n \in \mathbb{N} := \{1, 2, \dots\}$. We maintain the following assumptions on the sequence of payoff states, the timing of events, and on what players observe about the game being played.

Assumption 1. *The sequence of payoff states $(\theta^n)_{n\in\mathbb{N}}$ follows a time-homogeneous, irreducible, aperiodic, and positive recurrent Markov chain with stationary distribution $\psi \in \Delta_{++}(\Theta)$. Within each period $n \in \mathbb{N}$, the timing of events is as follows:*

(i) State θ^n realizes (unobserved by players);

(ii) A profile of signals $(t_1^n, \dots, t_I^n)$ is drawn from $\pi(\cdot \mid \theta^n)$;

(iii) After observing his signal t_i^n, each player i selects an action a_i^n and payoffs realize;

(iv) Each player i observes his own realized payoff $u_i(a^n, \theta^n)$ (but not necessarily the realized state θ^n and the profile of actions a^n that has been played).

We denote the infinite repetition of game (G, S) under Assumption 1 by $(G, S)^\infty$. We refer to $(G, S)^\infty$ as the *repeated game* and to $((a^n, t^n, \theta^n))_{n\in\mathbb{N}}$ as a *sequence of actions, signals, and states from* $(G, S)^\infty$.

Remark 1. We only assume that each player receives as feedback his own realized payoff at the end of each period. In particular, it is worth noting the following:

- The (one-shot or repeated) game need not be common knowledge among players;
- Players need not know what the state space Θ of process $(\theta^n)_{n\in\mathbb{N}}$ is and how the process evolves over time (i.e. players need not know ψ);
- Players need not know the signal distribution π.

To ease exposition, we assumed that the state space of process $(\theta^n)_{n\in\mathbb{N}}$ coincides with the set of payoff states of (G, S) and that the stationary distribution of $(\theta^n)_{n\in\mathbb{N}}$ coincides with the common prior of (G, S). The results in this paper also hold, modulo the obvious changes, when the state space of $(\theta^n)_{n\in\mathbb{N}}$ is some is some Θ' with $\emptyset \neq \Theta' \subset \Theta$, and its stationary distribution of is some $\psi' \in \Delta_{++}(\Theta')$.

Special instances of Assumption 1 are the case in which $(\theta^n)_{n\in\mathbb{N}}$ is an i.i.d. process and the case in which the payoff state is perfectly persistent (i.e. $\theta^n = \theta$ for some payoff state θ and all periods n).

2.2 Regrets and ε-Regret Dynamics

In this section, we formalize the notions of internal regret, external regret, and ε-regret dynamics for game $(G, S)^\infty$. The notions of external and internal regret for the repeated play of a one-shot complete information game are due to Hart and Mas-Colell (2000) (but see also Foster and Vohra (1997) and Fudenberg and Levine (1998, 1999a)).[3] We extend their notions to the repeated play of a one-shot incomplete information game in the natural way—with each player computing his own regrets signal-by-signal.

For all $i \in \mathcal{I}$ and $t_i \in T_i$, let $U_i(t_i, N)$ be the average payoff that player i with signal t_i has obtained up to time N; that is,

$$U_i(t_i, N) := \frac{1}{N} \sum_{n=1}^{N} u_i\big((a_i^n, a_{-i}^n), \theta^n)\big) \mathbb{1}_{\{t_i\}}(t_i^n).$$

Moreover, for each action $k \in A_i$, let $V_i^{ext}(k, t_i, N)$ be the average payoff player i with signal t_i would have obtained had he played k in all periods up to time N; that is,

$$V_i^{ext}(k, t_i, N) := \frac{1}{N} \sum_{n=1}^{N} u_i\big((k, a_{-i}^n), \theta^n)\big) \mathbb{1}_{\{t_i\}}(t_i^n).$$

Definition 1 (External Regret)**.** *For all $i \in \mathcal{I}$, $t_i \in T_i$ and $k \in A_i$, the external regret $R_i^{ext}(k, t_i, N)$ for action k before play at time $N + 1$ is defined by*

$$R_i^{ext}(k, t_i, N) := \max\left\{V_i^{ext}(k, t_i, N) - U_i(t_i, N), 0\right\}.$$

Now, let j be the last action played by player i with signal t_i up to time N. For each action $k \in A_i$, let $V_i^{int}(j, k, t_i, N)$ be the average payoff player i with signal t_i would have obtained had he played k instead of j every time in the past that he actually played j; that is,

$$V_i^{int}(j, k, t_i, N) := \frac{1}{N} \sum_{n=1}^{N} v_i^n(j, k, t_i),$$

where, for all $n \in \mathbb{N}$,

$$v_i^n(j, k, t_i) := \begin{cases} u_i\big((k, a_{-i}^n), \theta^n)\big) \mathbb{1}_{\{t_i\}}(t_i^n) & \text{if } a_i^n = j \\ u_i\big((a_i^n, a_{-i}^n), \theta^n)\big) \mathbb{1}_{\{t_i\}}(t_i^n) & \text{if } a_i^n \neq j \end{cases}.$$

Definition 2 (Internal Regret)**.** *For all $i \in \mathcal{I}$, $t_i \in T_i$ and $j, k \in A_i$, the internal regret $R_i^{int}(j, k, t_i, N)$ for action k with respect to action j before play at time $N+1$ is defined by*

$$R_i^{int}(j, k, t_i, N) := \max\left\{V_i^{int}(j, k, t_i, N) - U_i(t_i, N), 0\right\}.$$

$R_i^{ext}(k, t_i, N)$ is a measure of the time-average regret experienced by player i with signal t_i at period N for not having played action k in all past periods up to N. $R_i^{int}(j, k, t_i, N)$ is a measure of the time-average regret experienced by player i with signal t_i at period N for not having played, every time that j was played in the past, the different action k.

External regrets are based on the increase in the average payoff, if any, were player i with signal t_i to replace all past plays with the best fixed action in hindsight. In contrast, internal regrets are based on the increase in the average payoff, if any, were player i with

[3] Hart and Mas-Colell (2000) refer to external and internal regrets as unconditional and conditional regrets.

signal t_i to replace all past plays of a particular action with an arbitrary different action, separately for each of his actions. Thus, external regrets are a rougher measure of regret than internal regrets.

Definition 3 (ε-Regret Dynamics). *Let $\varepsilon \geq 0$. A sequence $((a^n, t^n, \theta^n))_{n\in\mathbb{N}}$ from $(G,S)^\infty$ has:*

(i) Asymptotic ε-external regret (hereafter, ε-AER) if

$$\limsup_{N\to\infty} R_i^{ext}(k, t_i, N) \leq \varepsilon$$

for all $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$.

(ii) Asymptotic ε-internal regret (hereafter, ε-AIR) if

$$\limsup_{N\to\infty} R_i^{int}(j, k, t_i, N) \leq \varepsilon$$

for all $i \in \mathcal{I}$, $t_i \in T_i$, and $j, k \in A_i$.

Asymptotic ε-regret notions (either external or internal) can be interpreted as a minimal long-run optimality conditions for the play of $(G,S)^\infty$. Under asymptotic ε-regret, the average regret experienced by each player for any of his signals for not having played differently in the past (i.e. according to the benchmark policy that depends on the notion of regret under consideration) is ε-close to vanish in the long-run.

The class of algorithms that players can follow in game $(G,S)^\infty$ and that generate a sequence of actions, signals, and states satisfying ε-AIR or ε-AER (hereafter, asymptotic ε-regret algorithms) is very large. Such class includes simple adaptive heuristics, fictitious-play-like dynamics, more sophisticated learning rules, such as calibrated learning and Bayesian learning, and also repeated equilibrium play of the one-shot game. We provide several examples of asymptotic ε-regret algorithms in Appendix A. For our empirical implementation, however, we do not take any stand on how players play game $(G,S)^\infty$. We only assume that players play $(G,S)^\infty$ sufficiently well for the sequence of actions, signals, and states to satisfy ε-AER or ε-AIR.

Before moving to the convergence properties of ε-regret dynamics, however, it is worth noting the following about asymptotic ε-regret algorithms. First, several asymptotic ε-regret algorithms are *simple*. In particular, though the notions of ε-AER and ε-AIR are defined on sequences $((a^n, t^n, \theta^n))_{n\in\mathbb{N}}$ from $(G,S)^\infty$, there exist asymptotic ε-regret algorithms that do not require players to observe $((a^n, \theta^n))_{n\in\mathbb{N}}$. That is, there exists asymptotic ε-regret algorithms that only require players' own utility as an input. Such algorithms are *model-free* and *uncoupled*. They are model-free because they do not require players to build a model of (or develop beliefs about) their opponents' play or to reply optimally to such play; they are uncoupled because they do not require players to know the payoff functions of the other players or their signals. Moreover, such algorithms do not require players to understand how the sequence of states $(\theta^n)_{n\in\mathbb{N}}$ evolves, its state space, and how signals are generated. Second, *sophisticated* asymptotic ε-regret algorithms also exist. Third, there are asymptotic ε-regret algorithms which are *Hannan-consistent*. Hannan-consistent algorithms unilaterally assure vanishing regrets for a player independently of what other the other players do. Thus, there is no need to impose that players coordinate on the same asymptotic ε-regret algorithm. To sum up, assuming that players' play of $(G,S)^\infty$ satisfies ε-AER or ε-AIR neither requires nor excludes sophisticated players or complex behavior.

3 Bayes Correlated ε-Equilibria and ε-Regret Dynamics

3.1 Equilibrium Notions for (G,S)

The relevant space of uncertainty in the incomplete information game (G,S) is $A \times T \times \Theta$. We write ν for a typical element of $\Delta(A \times T \times \Theta)$. The notions of Bayes correlated ε-equilibrium and Bayes coarse correlated ε-equilibrium are defined through the restrictions we impose on ν. Such restrictions are the following.

Definition 4 (Restrictions on ν)**.** *The probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is:*

(i) Consistent for (G,S) if, for all $t \in T$ and $\theta \in \Theta$, we have

$$\sum_a \nu(a,t,\theta) = \pi(t \,|\, \theta)\psi(\theta).$$

(ii) ε-obedient for (G,S) if, for all $i \in \mathcal{I}$, $t_i \in T_i$, and $a_i \in A_i$, we have

$$\sum_{a_{-i},t_{-i},\theta} \big[u_i\big((a_i',a_{-i}),\theta\big) - u_i\big((a_i,a_{-i}),\theta\big)\big]\nu\big((a_i,a_{-i}),(t_i,t_{-i}),\theta\big) \le \varepsilon \qquad (1)$$

for all $a_i' \in A_i$.

(iii) Coarsely ε-obedient for (G,S) if, for all $i \in \mathcal{I}$ and $t_i \in T_i$, we have

$$\sum_{a,t_{-i},\theta} \big[u_i\big((a_i',a_{-i}),\theta\big) - u_i(a,\theta)\big]\nu\big(a,(t_i,t_{-i}),\theta\big) \le \varepsilon \qquad (2)$$

for all $a_i' \in A_i$.

Consistency is a feasibility constraint which says that the marginal of distribution ν on the exogenous variables T and Θ is consistent with the description of game (G,S). Obedience and coarse obedience are incentive constraints. A probability distribution ν is *obedient* if any player i who knows ν and is told his action-signal pair (a_i,t_i) from a realization of ν weakly prefers to play a_i, given that the other players, who know their realized action-signal pair, are going to play their part of the realized action profile. A probability distribution ν is *coarsely obedient* if any player i who knows ν, is told his signal t_i, but not his action a_i, from a realization of ν, and is given a choice between (a) committing to whatever joint action profile (a_i, a_{-i}) has realized from ν, and (b) committing to a fixed action a_i', weakly prefers (a) to (b), given that the other players, who know their realized signal, but not their realized action, are committed to playing their part of whatever joint action has realized.

The form of obedience we impose on ν distinguishes the notion of *Bayes correlated ε-equilibrium* from that of *Bayes coarse correlated ε-equilibrium*, which we can now define.

Definition 5 (Bayes Correlated ε-Equilibrium and Bayes Coarse Correlated ε-Equilibrium)**.** *Let $\varepsilon \geq 0$. The probability distribution $\nu \in \Delta(A \times T \times \Theta)$ is:*

(i) A Bayes Correlated ε-Equilibrium (hereafter, ε-BCE) of (G, S) if it consistent and ε-obedient for (G, S).

(ii) A Bayes Coarse Correlated ε-Equilibrium (hereafter, ε-BCCE) of (G, S) if it consistent and coarsely ε-obedient for (G, S).

We denote by $BCE_{(G,S)}(\varepsilon)$ the set of ε-BCE of (G, S) and by $BCCE_{(G,S)}(\varepsilon)$ the set of ε-BCCE of (G, S).

For $\varepsilon = 0$, we have the notions of Bayes correlated equilibrium (hereafter, BCE) due to Bergemann and Morris (2013, 2016) and that of Bayes coarse correlated equilibrium (hereafter, BCCE). The BCE notion can be seen as an incomplete information version of correlated equilibrium (Aumann (1974, 1987)). The BCCE notion can be seen as an incomplete information version of coarse correlated equilibrium (Hannan (1957), Moulin and Vial (1978), and Young (2004)). We extend Bergemann and Morris (2013, 2016)'s notion of BCE to its coarse analog in the natural way. When Θ is a singleton, the definition of ε-BCE (resp., ε-BCCE) reduces to the definition of correlated ε-equilibrium (resp., coarse correlated ε-equilibrium) for a complete information game. For all $\varepsilon \geq 0$, it is straightforward to show that $BCE_{(G,S)}(\varepsilon)$ and $BCCE_{(G,S)}(\varepsilon)$ are convex sets and that $BCE_{(G,S)}(\varepsilon) \subseteq BCCE_{(G,S)}(\varepsilon)$.

The distribution $\nu \in \Delta(A \times T \times \Theta)$ is a Bayes Nash equilibrium (hereafter, BNE) action-signal-state distribution of (G, S) if there exists a BNE $\beta \coloneqq (\beta_1, \ldots, \beta_I)$ of (G, S), where $\beta_i \colon T_i \to \Delta(A_i)$ for all $i \in \mathcal{I}$, such that

$$\nu(a, t, \theta) = \psi(\theta)\pi(t \,|\, \theta) \prod_{i=1}^{I} \beta_i(a_i \,|\, t_i)$$

for all $a \in A$, $t \in T$, and $\theta \in \Theta$. It is straightforward to show that every BNE action-signal-state distribution of (G, S) is a BCE (and so a BCCE) of (G, S). Moreover, for all $\varepsilon \geq 0$, $BCE_{(G,S)}(\varepsilon)$ contains the convex hull of all BNE action-signal-state distributions of (G, S), and so does $BCCE_{(G,S)}(\varepsilon)$.

3.2 Convergence of ε-Regret Dynamics

Definition 6 (Empirical Distribution)**.** *Let $((a^n, t^n, \theta^n))_{n \in \mathbb{N}}$ be a sequence of actions, signals, and states from $(G, S)^\infty$. For every $N \in \mathbb{N}$, the empirical distribution $Z^N \in$*

$\Delta(A \times T \times \Theta)$ *is defined pointwise as*

$$Z^N(a,t,\theta) \coloneqq \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{\{a\}}(a^n)\mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n) \tag{3}$$

for all $(a,t,\theta) \in A \times T \times \Theta$.

That is, $Z^N(a,t,\theta)$ is the empirical frequency of the action-signal-state profile (a,t,θ) in the first N periods.

The next theorem establishes the following facts:

(i) A necessary and sufficient condition for the sequence of empirical distributions to converge almost surely to $BCE_{(G,S)}(\varepsilon)$ is that the sequence of actions, signals, and states has ε-AIR.

(ii) A necessary and sufficient condition for the sequence of empirical distributions to converge almost surely to $BCCE_{(G,S)}(\varepsilon)$ is that the sequence of actions, signals, and states has ε-AER.

Theorem 1 (Convergence of ε-Regret Dynamics)**.** *Let* $\varepsilon \geq 0$. *The sequence of actions, signals, and states* $((a^n,t^n,\theta^n))_{n\in\mathbb{N}}$ *from* $(G,S)^\infty$*:*

(i) Has ε*-AIR almost surely if and only if, as* $N \to \infty$*, the sequence of empirical distributions* $(Z^N)_{N\in\mathbb{N}}$ *converges almost surely to* $BCE_{(G,S)}(\varepsilon)$.

(ii) Has ε*-AER almost surely if and only if, as* $N \to \infty$*, the sequence of empirical distributions* $(Z^N)_{N\in\mathbb{N}}$ *converges almost surely to* $BCCE_{(G,S)}(\varepsilon)$.

Proof. We prove part (i). For the proof of part (ii), we refer to Appendix B.1.

[$\Longrightarrow$] Fix $\varepsilon \geq 0$ and suppose $((a^n,t^n,\theta^n))_{n\in\mathbb{N}}$ from $(G,S)^\infty$ has ε-AIR almost surely. Consider any subsequence $\left(Z^{N_l}\right)_{l\in\mathbb{N}}$ of $\left(Z^N\right)_{N\in\mathbb{N}}$ that converges almost surely to some $\nu \in \Delta(A\times T\times\Theta)$. We need to show that ν is almost surely an ε-BCE of (G,S), i.e. that ν is almost surely consistent and ε-obedient for (G,S).

Consistency. Pick any $(t,\theta) \in T \times \Theta$. Note the following:

$$\begin{aligned}
\sum_a \nu(a,t,\theta) &= \sum_a \lim_{l\to\infty} Z^{N_l}(a,t,\theta) \\
&= \lim_{l\to\infty} \sum_a Z^{N_l}(a,t,\theta) \\
&= \lim_{l\to\infty} \left[\frac{\sum_a Z^{N_l}(a,t,\theta)}{\sum_{a,t} Z^{N_l}(a,t,\theta)} \sum_{a,t} Z^{N_l}(a,t,\theta)\right] \qquad (4)\\
&= \lim_{l\to\infty} \frac{\sum_a Z^{N_l}(a,t,\theta)}{\sum_{a,t} Z^{N_l}(a,t,\theta)} \lim_{l\to\infty} \sum_{a,t} Z^{N_l}(a,t,\theta) \\
&= \lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} \lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)}{N_l}.
\end{aligned}$$

The ratio

$$\frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} \tag{5}$$

is the empirical frequency of the signal profile t when filtered at time steps where the state is θ. As the t^n's are drawn from $\pi(\cdot \,|\, \theta^n)$, (5) is the empirical frequency of $\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)$ conditionally independent observations from $\pi(\cdot \,|\, \theta)$. Moreover, as the Markov chain is recurrent, $\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n) \to \infty$ as $l \to \infty$ almost surely. Thus, by the strong law of large numbers,

$$\lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} = \pi(t \,|\, \theta) \qquad \text{a.s.} \tag{6}$$

Moreover, since the Markov chain is irreducible, positive recurrent, and has stationary distribution ψ,

$$\lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)}{N_l} = \psi(\theta) \qquad \text{a.s.} \tag{7}$$

Together, (4), (6), and (7) give

$$\sum_a \nu(a,t,\theta) = \pi(t \,|\, \theta)\psi(\theta) \qquad \text{a.s.} \tag{8}$$

As $(t,\theta) \in T \times \Theta$ was arbitrarily chosen, we conclude from (8) that ν is almost surely consistent for (G,S).

ε-obedience. To begin, note the following:

$$\begin{aligned}
&V_i^{int}(j,k,N,t_i) - U_i(t_i,N) \\
&= \frac{1}{N}\sum_{n=1}^{N} \big[u_i\big((k,a_{-i}^n),\theta^n\big) - u_i\big((a_i^n,a_{-i}^n),\theta^n\big)\big]\mathbb{1}_{\{j\}}(a_i^n)\mathbb{1}_{\{t_i\}}(t_i^n) \\
&= \frac{1}{N}\sum_{\theta\in\Theta}\sum_{n=1}^{N} \big[u_i\big((k,a_{-i}^n),\theta^n\big) - u_i\big((a_i^n,a_{-i}^n),\theta^n\big)\big]\mathbb{1}_{\{j\}}(a_i^n)\mathbb{1}_{\{t_i\}}(t_i^n)\mathbb{1}_{\{\theta\}}(\theta^n) \\
&= \sum_{a_{-i},t_{-i},\theta} \big[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\big] Z^N\big((j,a_{-i}),(t_i,t_{-i}),\theta\big).
\end{aligned} \tag{9}$$

Now pick any $i \in \mathcal{I}$, $t_i \in T_i$, and $j,k \in A_i$. As $\limsup_{N\to\infty} R_i^{int}(j,k,t_i,N) \le \varepsilon$ a.s., by definition of $R_i^{int}(j,k,t_i,N)$, we also have $\limsup_{N\to\infty} \big[V_i^{int}(j,k,t_i,N) - U_i(t_i,N)\big] \le \varepsilon$ a.s. But then, by (9),

$$\limsup_{N\to\infty} \sum_{a_{-i},t_{-i},\theta} \big[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\big] Z^N\big((j,a_{-i}),(t_i,t_{-i}),\theta\big) \le \varepsilon \qquad \text{a.s.} \tag{10}$$

Moreover, on the subsequence $\big(Z^{N_l}\big)_{l\in\mathbb{N}}$ we get

$$\begin{aligned}
&\lim_{l\to\infty} \sum_{a_{-i},t_{-i},\theta} \big[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\big] Z^{N_l}\big((j,a_{-i}),(t_i,t_{-i}),\theta\big) \\
&= \sum_{a_{-i},t_{-i},\theta} \lim_{l\to\infty} \big[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\big] Z^{N_l}\big((j,a_{-i}),(t_i,t_{-i}),\theta\big)
\end{aligned} \tag{11}$$

$$= \sum_{a_{-i},t_{-i},\theta} \left[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\right]\nu\big((j,a_{-i}),(t_i,t_{-i}),\theta\big).$$

Together, (10) and (11) give

$$\sum_{a_{-i},t_{-i},\theta} \left[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\right]\nu\big((j,a_{-i}),(t_i,t_{-i}),\theta\big) \leq \varepsilon \qquad \text{a.s.} \tag{12}$$

As $i \in \mathcal{I}$, $t_i \in T_i$, and $j,k \in A_i$ were arbitrarily chosen, we conclude from (12) that ν is almost surely ε-obedient for (G,S).

[$\Longleftarrow$] Now suppose $\left(Z^N\right)_{N\in\mathbb{N}}$ converges almost surely to $BCE_{(G,S)}(\varepsilon)$ for some $\varepsilon \geq 0$. Pick any $i \in \mathcal{I}$, $t_i \in T_i$, and $j,k \in A_i$. By ε-obedience,

$$\limsup_{N\to\infty} \sum_{a_{-i},t_{-i},\theta} \left[u_i\big((k,a_{-i}),\theta\big) - u_i\big((j,a_{-i}),\theta\big)\right] Z^N\big((j,a_{-i}),(t_i,t_{-i}),\theta\big) \leq \varepsilon \qquad \text{a.s.} \tag{13}$$

By (9) and (13),

$$\limsup_{N\to\infty} \left[V_i^{int}(j,k,t_i,N) - U_i(t_i,N)\right] \leq \varepsilon \qquad \text{a.s.,}$$

which implies

$$\limsup_{N\to\infty} R_i^{int}(j,k,t_i,N) \leq \varepsilon \qquad \text{a.s.}$$

by definition of $R_i^{int}(j,k,t_i,N)$. As $i \in \mathcal{I}$, $t_i \in T_i$, and $j,k \in A_i$ were arbitrarily chosen, the desired result follows. ■

The empirical exercise in Section 4 is based on the sufficiency parts of Theorem 1. If players' play in $(G,S)^\infty$ satisfies the minimal long-run optimality condition captured by ε-AIR (resp., ε-AER), Theorem 1 implies that the empirical distribution converges almost surely to $BCE_{(G,S)}(\varepsilon)$ (resp., $BCCE_{(G,S)}(\varepsilon)$). That is, from some time on, the empirical distribution is almost surely close to an ε-BCE of (G,S) (resp., ε-BCCE of (G,S)). The convergence here is to $BCE_{(G,S)}(\varepsilon)$ (resp., $BCCE_{(G,S)}(\varepsilon)$), not necessarily to a point in that set. Moreover, observe that it is the empirical distribution that becomes essentially an ε-BCE (resp., ε-BCCE), not necessarily the actual play.

Remark 2. Almost sure convergence of $\left(Z^N\right)_{N\in\mathbb{N}}$ to $BCE_{(G,S)}(\varepsilon)$ (resp., $BCCE_{(G,S)}(\varepsilon)$) means that the sequence $\left(Z^N\right)_{N\in\mathbb{N}}$ eventually enters any neighborhood of $BCE_{(G,S)}(\varepsilon)$ (resp., $BCCE_{(G,S)}(\varepsilon)$) and stays there forever. An equivalent way of stating this is the following: given any $\varepsilon' > \varepsilon$, there is a finite time $N(\varepsilon')$ after which the empirical distribution is always a ε'-BCE (resp., ε'-BCCE) of (G,S) almost surely; that is, $Z^N \in BCE_{(G,S)}(\varepsilon')$ (resp., $Z^N \in BCCE_{(G,S)}(\varepsilon')$) for all $N > N(\varepsilon')$ almost surely.

Remark 3. When the payoff state is perfectly persistent (i.e. $\theta^n = \theta$ for some payoff state θ and all periods n), if $((a^n,t^n,\theta^n))_{n\in\mathbb{N}}$ has ε-AIR (resp., ε-AER) almost surely, then the sequence of empirical distributions $(Z^N)_{N\in\mathbb{N}}$ converges almost surely to the set of correlated ε-equilibria (resp., coarse correlated ε-equilibria) of the complete information game with payoff state $\Theta = \{\theta\}$.

Hartline et al. (2015) also study convergence properties of no-regret learning dynamics in settings with incomplete information.[4] There are (at least) three main differences between our and their work. First, they study a setting where private information is independent across players (independent private values) and time. In contrast, in our setting, private information in arbitrarily correlated across players and and evolves over time according to a Markov chain. Second, they study convergence to the coarse analog of the agent normal form Bayes correlated equilibrium of Forges (1993). In contrast, we study convergence notion of Bayes correlated equilibrium in Bergemann and Morris (2016) and to its coarse analog. Third, they adopt a different definition of players' payoffs (and, consequently, regrets). Because of these and other technical differences, our results are not equivalent. The additional generality of our setting is necessary for the econometric and empirical analyses we perform in the next sections.

4 From the Model to the Data

Within the model developed in Section 2, we now turn to address the following question: when players interact repeatedly in an incomplete information game, what features of the underlying one-shot game can we recover under weak assumptions on behavior?

We leverage the convergence results established in Section 2 to recover features of the underlying one-shot game under the assumption that the data are the outcomes of an ongoing interaction over time. The literature on empirical games typically maintains that the observable actions results from equilibrium play in a cross-section of simultaneous games (e.g., Berry (1992), Tamer (2003), and Ciliberto and Tamer (2009)) or from fully rational dynamic equilibrium play in a panel of dynamic games (e.g., Ericson and Pakes (1995), Benkard (2004), Jofre-Bonet and Pesendorfer (2003), and Ryan (2012)). In contrast, we do not make any strong assumption on how players play the one-shot game in each period. We only assume that players play the one-shot game sufficiently well for the sequence of actions, signals, and states to satisfy, in the long-run, the minimal optimality condition captured by ε-AIR. The results of this section extend in the obvious way when we assume that the sequence of actions, signals, and states has ε-AER instead of ε-AIR.

Empirical Model. We start by laying out the main assumptions on the empirical model and on the observables.

Assumption 2. *The empirical model and the observables are summarized by (i)–(v) below.*

(i) The structural parameters $\lambda := \left(\lambda^G, \lambda^S\right)$ *that define the one-shot incomplete information game* $\left(G\left(\lambda^G\right), S\left(\lambda^S\right)\right)$ *belong to a compact set* Λ.

(ii) $\lambda_0 := \left(\lambda_0^G, \lambda_0^S\right)$ *are the true structural parameters in the data generating process.*

[4] Relatedly, Caragiannis et al. (2015) study convergence properties of no-regret learning dynamics in generalized second price auctions.

(iii) The game $\left(G(\lambda_0^G), S(\lambda_0^S)\right)$ *is played repeatedly over time under Assumption 1.*

(iv) The sequence of actions, signals, and states $((a^n, t^n, \theta^n))_{n\in\mathbb{N}}$ *from* $\left(G(\lambda_0^G), S(\lambda_0^S)\right)^\infty$ *has* ε*-AIR.*

(v) The econometrician only observes the realized sequence of actions $(a^n)_{n\in\mathbb{N}}$*.*

Although each player i observes the sequence $((a^n, t_i^n, \theta^n))_{n\in\mathbb{N}}$, in applied contexts outside analysts have typically less information than players. Therefore, we assume that the econometrician only observes the realized sequence of actions $(a^n)_{n\in\mathbb{N}}$. [5]

In the remaining part of this section, we investigate what we can recover of λ_0, and how, under Assumption 2.

Bayes Correlated Equilibrium and Restrictions on Parameters. To recover the structural parameter λ_0 under Assumption 2, we rely on the convergence results in Section 2, which motivate the adoption of our equilibrium restrictions. In particular, we show that estimation under ε-BCE allows to recover valid bounds on the structural parameter. We have already defined the notion of ε-BCE; denote by $BCE_{(G,S)}(\lambda;\varepsilon)$ the set of ε-BCE of the incomplete information game $(G(\lambda^G), S(\lambda^S))$ with structural parameters $\lambda := (\lambda^G, \lambda^S) \in \Lambda$. We now expand on the restrictions that ε-BCE implies for the structural parameters.

Definition 7 (ε-BCE Prediction)**.** *Let* $\varepsilon \geq 0$*. A probability distribution* $q \in \Delta(A)$ *is an* ε*-BCE prediction if there exists* $\nu \in BCE_{(G,S)}(\lambda;\varepsilon)$ *such that*

$$q(a) = \sum_{t,\theta} \nu(a,t,\theta)$$

for all $a \in A$*. The set of* ε*-BCE predictions for the game with structural parameters* λ *is denoted by* $Q_{(G,S)}(\lambda;\varepsilon)$*.*

Definition 8 (Identified Set)**.** *Let* $q \in \Delta(A)$ *be a distribution of actions. The set of parameters identified by* q *under the* ε*-BCE assumption, denoted by* $\Lambda^I_{(G,S)}(q;\varepsilon)$*, is*

$$\Lambda^I_{(G,S)}(q;\varepsilon) := \left\{\lambda \in \Lambda : q \in Q_{(G,S)}(\lambda;\varepsilon)\right\}.$$

In our model, we do not observe a limiting "population" distribution of the observables, i.e. a fixed limiting $q \in \Delta(A)$. To see this, let the *empirical distribution of actions* $q^N \in \Delta(A)$ be defined pointwise as

$$q^N(a) := \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{\{a\}}(a^n)$$

for all $a \in A$. Under Assumption 2, Theorem 1 only ensures that the sequence of empirical distributions $\left(Z^N\right)_{N\in\mathbb{N}}$ converges almost surely to $BCE_{(G,S)}(\lambda;\varepsilon)$ as $N \to \infty$.

[5] Although this assumption is the most common in the empirical literature, it is not the only one considered in the literature. For instance, Bergemann and Morris (2013) consider identification under BCE behavior in a model where both actions and payoff states, i.e. $((a^n, \theta^n))_{n\in\mathbb{N}}$, are observable to the econometrician.

Therefore, as q^N is the marginal on A of the empirical distribution Z^N for all $N \in \mathbb{N}$, i.e. $q^N(a) = \sum_{t,\theta} Z^N(a,t,\theta)$ for all $a \in A$, Theorem 1 only ensures that, as the sample size gets large, the empirical distribution of actions converges almost surely to the set $Q_{(G,S)}(\lambda_0;\varepsilon)$, not necessarily to a point in that set. To overcome this complication, instead of focusing on the identified set $\Lambda^I_{(G,S)}(q;\varepsilon)$, we consider the set of parameters that can be recovered when *any* $q \in Q_{(G,S)}(\lambda_0;\varepsilon)$ may describe the data, which leads to the next definition.

Definition 9 (Recoverable Set)**.** *The set of recoverable parameters under Assumption 2, denoted by* $\Lambda^R_{(G,S)}(\varepsilon)$*, is*

$$\Lambda^R_{(G,S)}(\varepsilon) \coloneqq \bigcup_{q \in Q_{(G,S)}(\lambda_0;\varepsilon)} \Lambda^I_{(G,S)}(q;\varepsilon).$$

The bounds imposed by $\Lambda^R_{(G,S)}(\varepsilon)$ are valid, in the sense that $\lambda_0 \in \Lambda^R_{(G,S)}(\varepsilon)$.[6]

Recovering Bounds on Parameters. Consider the "plug-in" estimator

$$\hat{\Lambda}^N_{(G,S)}(\varepsilon) \coloneqq \left\{\lambda \in \Lambda :\ q^N \in Q_{(G,S)}(\lambda;\varepsilon)\right\} = \Lambda^I_{(G,S)}\left(q^N;\varepsilon\right),$$

where q^N is the observed empirical distribution of N actions.

Theorem 2 (Properties of $\hat{\Lambda}^N_{(G,S)}$)**.** *Under Assumption 2, for any $\varepsilon' > \varepsilon$, the following statements hold almost surely as $N \to \infty$:*

(i) $\lambda_0 \in \hat{\Lambda}^N_{(G,S)}(\varepsilon')$.

(ii) $\hat{\Lambda}^N_{(G,S)}(\varepsilon) \subseteq \Lambda^R_{(G,S)}(\varepsilon')$.

Theorem 2 says that a static equilibrium notion, ε-BCE, provides an adequate behavioral restriction for the estimation of dynamic interactions that satisfy the long-run optimality condition described by the ε-AIR assumption. Part (i) of the theorem establishes that the restriction of ε-AIR leads to estimating a set of parameters which contains the true structure of the data generating process. Part (ii) describes bounds on this estimated set, which is contained within the (theoretical) recoverable set. The width of the bounds, in practice, will depend on the specific model and on the informativeness of the data. Despite data not being generated by the repetition of identical experiments, we bound structural parameters without statistical assumptions on the sampling process on top of the economic assumption of ε-no-internal regret in the limit.

Proof of Theorem 2. To establish part (i), fix any $\varepsilon' > \varepsilon$ and note that, by definition of $\hat{\Lambda}^N_{(G,S)}(\varepsilon')$,

$$\lambda_0 \in \hat{\Lambda}^N_{(G,S)}(\varepsilon') \Longleftrightarrow q^N \in Q_{(G,S)}(\lambda_0;\varepsilon'). \tag{14}$$

Under Assumption 2, by Theorem 1 we have that $\left(Z^N\right)_{N\in\mathbb{N}}$ converges almost surely to $BCE_{(G,S)}(\varepsilon;\lambda)$ as $N \to \infty$. Then, by Remark 2, there exists $N(\varepsilon')$ such that $Z^N \in$

[6]That $\lambda_0 \in \Lambda^R_{(G,S)}(\varepsilon)$ holds by construction of $\Lambda^R_{(G,S)}(\varepsilon)$.

$BCE_{(G,S)}(\varepsilon';\lambda)$ for all $N > N(\varepsilon')$ almost surely. As q^N is the marginal on A of Z^N for all $N \in \mathbb{N}$, it follows that $q^N \in Q_{(G,S)}(\lambda_0;\varepsilon')$ for all $N > N(\varepsilon')$ almost surely. Combining this fact with (14) gives the desired result.

To establish part (ii), fix any $\varepsilon' > \varepsilon$ and note that, by definition of $\Lambda_N(\varepsilon')$ and of $\hat{\Lambda}_R(\varepsilon')$,

$$q^N \in Q_{(G,S)}(\lambda_0;\varepsilon') \Longleftrightarrow \Lambda^I_{(G,S)}\big(q^N;\varepsilon'\big) \subseteq \Lambda^R_{(G,S)}(\varepsilon') \Longleftrightarrow \hat{\Lambda}^N_{(G,S)}(\varepsilon') \subseteq \Lambda^R_{(G,S)}(\varepsilon'). \tag{15}$$

Moreover, as $\varepsilon' > \varepsilon$,

$$\hat{\Lambda}^N_{(G,S)}(\varepsilon) \subseteq \hat{\Lambda}^N_{(G,S)}(\varepsilon'). \tag{16}$$

Under Assumption 2, by Theorem 1 we have that $q^N \in Q_{(G,S)}(\lambda_0;\varepsilon')$ for all $N > N(\varepsilon')$ almost surely (see proof of part (i)). Combining this with (15) and (16) gives the result. ■

5 Empirical Application

We use our method to perform inference on the distribution of sellers' marginal costs on Swappa, an online marketplace for used smartphones and other consumer electronics with around $100 million in sales in 2018. Sellers (typically individuals or small firms) can list individual devices on Swappa for free. Listings go through an approval process whereby the platform checks that the device is ready for activation. Whereas listing is free for sellers, buyers pay a fee to the platform. We focus on our analysis on the larger sellers on the platform. These are—for the vast majority—small firms that acquire used cellphones, refurbish them, and resell them. Marginal costs for these firms are mostly the cost of acquisition of used devices, and the costs (including labor and parts) of refurbishing them.

We consider Swappa an appealing empirical environment for our model for at least two reasons. First, it is a good example of a decentralized platform, where pricing is a decision of individual sellers. In this, it resembles many other platform and marketplaces, including Amazon marketplace and eBay buy-it-now listings. Second, although the sellers we consider are the larger agents on the platform, they are still small and their level of sophistication in pricing decisions is unclear.

Estimating sellers' marginal costs under a minimal restriction on behavior will allow us to consider a variety of market design questions. For instance, suppose that the platform were to offer pricing algorithms to sellers. What would be the resulting change in sales and surplus?

5.1 Data and Empirical Environment

Data collection is still in progress.

5.2 Empirical Model

Players in $\mathcal{I}$ are sellers of a differentiated good. In each period $n \in \mathbb{N}$, a set $\mathcal{I}^n \subseteq \mathcal{I}$ of sellers is called to play the following pricing game. Each seller $i \in \mathcal{I}^n$ sets price $p_i^n \in P$, where P is a finite set, and faces demand

$$g(p_i^n, p_{-i}^n) := m \frac{\exp(\alpha p_i^n)}{1 + \sum_{j \in \mathcal{I}^n} \exp(\alpha p_j^n)},$$

where m is the market size and α describes consumers' sensitivity to price. Parameters m and α are known to sellers. Seller i has unit marginal cost $t_i^n \in T$, where T is a finite set; this is seller i's private information. Seller i's payoff is

$$\tilde{u}_i(p_i^n, p_{-i}^n, t_i^n) := g(p_i^n, p_{-i}^n)(p_i^n - t_i^n).$$

In each period $n \in \mathbb{N}$, each seller $i \in \mathcal{I}$ is active with positive probability. Conditional on being active, each seller's unit marginal cost evolves according to an irreducible and positive recurrent Markov chain with stationary distribution $\psi(\cdot; \lambda) \in \Delta_{++}(T)$. These Markov chains are independent across sellers, so that, in each period, unit marginal costs are i.i.d. across sellers. Moreover, sellers' unit marginal costs are independent of the set of sellers that are active in a given period.

We work with an aggregative formulation of the pricing game, which is equivalent to the original game. This aggregative approach, which is in the spirit of Nocke and Schutz (2018), is helpful to sidestep curse of dimensionality in large games. For all $n \in \mathbb{N}$, $\mathcal{I}^n \subseteq \mathcal{I}$, $i \in \mathcal{I}^n$, and $p_i^n \in P$, define $f_i^n := \exp(\alpha p_i^n)$ and $\overline{f}^n := \sum_{j \in \mathcal{I}^n} f_j^n$. Note that $f_i^n \in F$ and $\overline{f}^n \in \overline{F}$, where F and $\overline{F}$ are finite sets defined in the obvious way. Hereafter, we suppress the subscript i since sellers are ex ante identical. A representative seller's payoff becomes

$$u\big(f^n, \overline{f}^n, t^n\big) = \left[m \frac{f^n}{1 + \overline{f}^n} \left(\frac{\log f^n}{\alpha} - t^n \right) \right].$$

Before setting f^n, each seller observes t^n and $o^n := |\mathcal{I}^n| \in O := \{1, \ldots, |\mathcal{I}|\}$, but not necessarily $\mathcal{I}^n$ (i.e. each seller observes the number of sellers, but not necessarily their identity). Hereafter, we assume that $(o^n)_{n \in \mathbb{N}}$ evolves according to an irreducible and positive recurrent Markov chain with stationary distribution $\phi \in \Delta_{++}(O)$. At the end of each period n, each seller in $\mathcal{I}^n$ observes his own realized payoff $u\big(f^n, \overline{f}^n, t^n\big)$ (but not necessarily $\overline{f}^n$).

Definition 10 (Aggregative Bayes Correlated Equilibrium)**.** *Let $\varepsilon \geq 0$. The probability distribution $\nu \in \Delta(F \times \overline{F} \times T \times O)$ is an Aggregative Bayes Correlated ε-equilibrium (hereafter, ε-ABCE) of the aggregative pricing game if it satisfies the following conditions:*

(i) ν is consistent; that is, for all $t \in T$ and $o \in O$, we have

$$\sum_{f, \overline{f}} \nu\big(f, \overline{f}, t, o\big) = \psi(t; \lambda)\phi(o).$$

(ii) ν *is* ε*-obedient; that is, for all* $t \in T$, $f \in F$, *and* $o \in O$ *we have*

$$\sum_{\overline{f}} \left[u\big((f', \overline{f} + (f' - f), t)\big) - u\big((f, \overline{f}), t\big) \right] \nu\big(f, \overline{f}, t, o\big) \leq \varepsilon.$$

for all $f' \in \mathcal{F}$.

We denote by $ABCE(\varepsilon; \lambda)$ *the set of* ε*-ABCE of the aggregative pricing game.*

For all $t \in T$, denote by $U(t, o, N)$ the average payoff that a seller with signal t playing the aggregative pricing game with o sellers has obtained up to time N; that is,

$$U(t, o, N) := \frac{1}{N} \sum_{n=1}^{N} u\big(f^n, \overline{f}^n, t^n)\big) \mathbb{1}_{\{t\}}(t^n) \mathbb{1}_{\{o\}}(o^n).$$

Next, let f be the last action played by the seller with signal t playing the aggregative pricing game with o sellers up to time N. For each action $f' \in F$, let $V^{int}(f, f', t, o, N)$ be the average payoff the seller with signal t playing the aggregative pricing game with o sellers would have obtained had he played f' instead of f every time in the past that he actually played f; that is,

$$V^{int}(f, f', t, o, N) := \frac{1}{N} \sum_{n=1}^{N} v^n(f, f', t, o),$$

where, for all $n \in \mathbb{N}$,

$$v^n(f, f', t, o) := \begin{cases} u\big((f', \overline{f}^n + (f' - f^n), t^n)\big) \mathbb{1}_{\{t\}}(t^n) \mathbb{1}_{\{o\}}(o^n) & \text{if } f^n = f \\ u\big((f^n, \overline{f}^n), t^n\big) \mathbb{1}_{\{t\}}(t^n) \mathbb{1}_{\{o\}}(o^n) & \text{if } f^n \neq f \end{cases}$$

For all $t \in T_i$ and $f, f' \in F$, the internal regret $R^{int}(f, f', t, o, N)$ for action f' with respect to action f before play at time $N + 1$ is defined by

$$R^{int}(f, f', t, o, N) := \max\left\{ V^{int}(f, f', t, o, N) - U(t, o, N), 0 \right\}.$$

Let $\varepsilon \geq 0$. A sequence $((f^n, \overline{f}^n, t^n, o^n))_{n \in \mathbb{N}}$ from the aggregative pricing game has ε-AIR if

$$\limsup_{N \to \infty} R^{int}(f, f', t, o, N) \leq \varepsilon$$

for all $t \in T$, $o \in O$ and $f, f' \in F$.

For every $N \in \mathbb{N}$, the empirical distribution $Z^N \in \Delta(F \times \overline{F} \times T \times O)$ is defined pointwise as

$$Z^N(f, \overline{f}, t, o) := \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\{f\}}(f^n) \mathbb{1}_{\{\overline{f}\}}(\overline{f}^n) \mathbb{1}_{\{t\}}(t^n) \mathbb{1}_{\{o\}}(o^n)$$

for all $(f, \overline{f}, t, o) \in F \times \overline{F} \times T \times O$.

Theorem 3. *Let* $\varepsilon \geq 0$. *If the sequence* $((f^n, \overline{f}^n, t^n, o^n))_{n \in \mathbb{N}}$ *from the aggregative pricing game has* ε*-AIR almost surely, then the sequence of empirical distributions* $(Z^N)_{N \in \mathbb{N}}$

converges almost surely to $ABCE(\varepsilon; \lambda)$ as $N \to \infty$.

The proof of Theorem 3 is similar to that of Theorem 1–(i) and is in Appendix B.2.

6 Additional Results and Discussion

6.1 Informational Robustness

There are at least two natural orderings on information structures: an "incentive ordering" and a "statistical ordering". Roughly speaking, we have the following.[7]

- Incentive ordering: an information structure is more incentive constrained than another if it gives rise to a smaller set of BCE.
- Statistical ordering: an information structure is individually sufficient for another if there exists a combined information structure where each player's signal from the former information structure is a sufficient statistic for the state and other players' signals in the latter information structure; individual sufficiency captures intuitively when one information structure contains more information than another.

Bergemann and Morris (2016) show that one information structure is more incentive constrained than another if and only if the former is individually sufficient for the latter. That is, the statistical ordering is equivalent to the incentive ordering.

Building on the latter equivalence, we can provide a robustness result for our empirical exercise. Fix a basic game G. Suppose the econometrician knows that players observe at least information structure S, but may observe more—in the sense that they may observe information structure S', for some S' that is individually sufficient for S, the exact S' being unknown to the econometrician. Can the econometrician recover valid bounds on the payoff structure of (G, S') under the ε-AIR assumption for the repeated play of the misspecified model (G, S)? The answer to this question is positive. Suppose that the sequence of actions, signals and states is generated by play in (G, S')—the true game—and has ε-AIR. Thus, by Theorem 1, the sequence of empirical distributions converges almost surely to $BCE_{(G,S')}(\varepsilon)$. As S' is individually sufficient for S, by Bergemann and Morris (2016)'s equivalence result, S' is also more incentive constrained than S, and so $BCE_{(G,S')}(\varepsilon) \subseteq BCE_{(G,S)}(\varepsilon)$. But then, the sequence of empirical distributions converges almost surely also to $BCE_{(G,S)}(\varepsilon)$. As a result, the bounds on the payoff structure under the ε-AIR assumption for the misspecified model remain valid for the true model, although they might not be as sharp as those one would obtain under the correct specification of the information structure. An analogous result can be established for ε-AER dynamics and the set of Bayes coarse correlated equilibria.

[7] We refer to Bergemann and Morris (2016) for the formal definitions and the discussion of other orderings on information structures.

6.2 How Long to Equilibrium?

In empirical applications, the question often arises of how many observations one needs to consistently estimate the parameters of interest. In our setting, this concern needs to be paired with an assessment of how long it takes for the empirical distribution to converge to $BCE_{(G,S)}(\varepsilon)$ (or $BCCE_{(G,S)}(\varepsilon)$).

The answer to the latter question depends on the particular ε-AIR (or ε-AER) algorithm that players follow in the repetition of the one-shot game (G, S). For instance, if players play a BNE or a BCE of (G, S) in each period, then the empirical distribution is in $BCE_{(G,S)}(\varepsilon)$ since period 1. When players follow regret-based algorithms, Hart and Mansour (2010) show that the rate of convergence to the set of correlated ε-equilibria of the underline complete information game is polynomial in the number of players; given our signal-by-signal extension of regret-based algorithms to incomplete information environments, one can show that the rate of convergence is longer—as each player now needs to accumulate experiences for each of their signal—but remains of the same order when players follow regret-based algorithms for the repeated play of (G, S).

In short, we cannot provide sharp rate-of-convergence results under weak assumptions on behavior (i.e. without selecting a specific ε-AIR (or ε-AER) algorithm). However, it is worth noting that converge to the set ε-BCE (or ε-BCCE) of (G, S) is, in general, faster that convergence to Nash or Bayes Nash equilibria (or related solution concepts).

7 Concluding Remarks

We propose an estimation strategy that is valid when data on strategic interaction are interpreted as the long-run result of a history of game plays. We model players as interacting repeatedly in a one-shot incomplete information game. We remain agnostic on how players play the game and only impose a minimal optimality condition for the long-run outcome of players' interaction.

In particular, we assume that play satisfies a property of "asymptotic no regret" (ANR). This condition requires that the time average of the counterfactual increase in past payoffs, had different actions been played, becomes approximately zero in the long run. A large class of dynamics, well studied both in economics and computer science, satisfies the ANR property. For example, ANR is satisfied by regret matching algorithms, calibrated learning, and variants of fictitious play. Moreover, the condition is trivially satisfied if the observed outcomes are the result of equilibrium play. Under the ANR assumption, we show partial identification of the structural parameters of the one-shot game. Identification relies on the result that the empirical distribution of play that satisfies ANR converges to the set of Bayes (coarse) correlated ε-equilibrium prediction of the underlying one-shot game. Consequently, we can use the limiting model to obtain consistent estimates of the parameters of interest.

This paper shows that no regret conditions, already a fundamental ingredient in optimization, machine learning, and algorithmic game theory, can allow economists to perform identification and estimation in empirical games under minimal restrictions on several aspects of the underlying interaction (notably, behavior and information). In future work, we plan to explore further the identifying power of regret minimization and its connections with equilibrium play in both single-agent and strategic environments.

A Asymptotic ε-Regret Algorithms: Examples

In this appendix, we provide some examples of algorithms that players can follow in game $(G, S)^\infty$ and that generate a sequence of actions, signals, and states satisfying ε-AIR or ε-AER (hereafter, ε-AIR and ε-AER algorithms). By Theorem 1, under ε-AIR (resp., ε-AER) algorithms the sequence of empirical distributions $(Z^N)_{N\in\mathbb{N}}$ converges almost surely to $BCE_{(G,S)}(\varepsilon)$ (resp., $BCCE_{(G,S)}(\varepsilon)$). Although we only sketch here a few examples, it is worth noting that the class of ε-regret algorithms is much larger.

A.1 Regret Matching and Generalizations

To begin with, we can generalize internal and external regret matching algorithms—due to Hart and Mas-Colell (2000) for the infinite repetition of a one-shot complete information game—to the incomplete information game $(G, S)^\infty$.

Let t_i be player i's signal in period $N+1$ and let j be the action played by player i the last time in the past he observed signal t_i. *Internal regret matching* stipulates that each action $k \neq j$ is played in period $N+1$ with a probability that is proportional to its internal regret $R_i^{int}(j, k, t_i, N)$ and, with the remaining probability, the same action j is played in period $N+1$. Formally, denote by $p_i^{N+1,int}(j, k, t_i, N)$ the probability of playing action k in period $N+1$ by player i with signal t_i, given that i has played action j the last time in the past he observed signal t_i. Then, internal regret matching prescribes that

$$p_i^{N+1,int}(j, k, t_i, N) = \begin{cases} cR_i^{int}(j, k, t_i, N) & \text{if } k \neq j \\ 1 - \sum_{k\neq j} cR_i^{int}(j, k, t_i, N) & \text{if } k = j \end{cases} \tag{17}$$

for some sufficiently small constant $c > 0$.[8] The play in the first period can be arbitrary.

In the same spirit, *external regret matching* stipulates that each action k is played in period $N+1$ with a probability that is directly proportional to the vector of external regrets. Formally, denote by $p_i^{N+1,ext}(k, t_i, N)$ the probability of playing action k in period $N+1$ by player i with signal t_i. Then, external regret matching prescribes that

$$p_i^{N+1,ext}(k, t_i, N) = \frac{R_i^{ext}(k, t_i, N)}{\sum_j R_i^{ext}(j, t_i, N)}$$

for all $k \in A_i$. Again, the play in the first period can be arbitrary, and so is play if all external regrets are zero.

Following the logic of Hart and Mas-Colell (2000), we can show that, if all players play according to internal (resp., external) regret matching in each period, then the sequence of actions, signals, and states has no-AIR (resp., no-AER) almost surely. The proof that all regrets vanish in the limit uses arguments suggested by Blackwell (1956)'s approachability.

Instead of the switching probability being proportional to $R_i^{int}(j, k, t_i, N)$, i.e. equal to

[8] The constant c must guarantee that (17) yields a probability distribution over A_i and, moreover, that the probability of j is strictly positive.

$cR_i^{int}(j, k, t_i, N)$, we may want to allow this switching probability to be given by a general function $f(R_i^{int}(j, k, t_i, N))$ of $R_i^{int}(j, k, t_i, N)$. If f is sign-preserving (i.e. $f(x) > 0$ for $x > 0$ and $f(0) = 0$) and Lipschitz continuous, we call the resulting algorithm *generalized internal regret matching* algorithm. Similarly, *generalized external regret matching* algorithms exist (see Hart and Mas-Colell (2001a), Cahn (2004), and Hart (2005)). Building on these authors' results, we can show that, if each type of each player plays according to a generalized internal (resp., external) regret matching algorithm in each period, then the sequence of actions, signals, and states has no-AIR (resp., no-AER) almost surely. In particular, different players may use different generalized regret matching algorithms.

Regret matching algorithms have some important properties, which we collect in the next remarks.

Remark 4. Regret matching algorithms are not very sophisticated—in fact, they are often referred to as *adaptive heuristics*. This is so at least on three grounds.

- Regret matching algorithms only require regrets as an input.[9] Thus, to play according to a regret matching algorithm, a player does not need to know the payoff functions of the other players or their signals (such property is called *uncoupledness*). In other words, regret matching algorithms achieve ε-AIR or ε-AER and only require each player to know what other players do, not what their objectives or private information are.
- Regret matching algorithms do not require players to build a model of (or develop beliefs about) their opponents' play or to reply optimally to such play. That is, regret matching algorithms are *model-free*.
- To play according to regret matching algorithms, players need not know ψ and π. That is, players need not understand how the sequence of states $(\theta^n)_{n\in\mathbb{N}}$ evolves (and so, in particular, players need not have a prior, let alone a common one) and how signals are generated.

Remark 5. A motivation for our work is to provide valid bounds on structural parameters under weak assumptions on behavior (see Section 4). Bounds on parameters are less sharp under the ε-AER assumption than under the ε-AIR assumption, as $BCE_{(G,S)}(\varepsilon) \subseteq BCCE_{(G,S)}(\varepsilon)$. However, the ε-AER assumption is more robust, at least in the following sense. Call an algorithm for the play of $(G, S)^\infty$ *Hannan-consistent* if it guarantees, for any algorithms other players may follow, that all the regrets of this player vanish in the limit with probability one.[10] That is, an algorithm is Hannan-consistent if it unilaterally assures vanishing regrets independently of what other players do. Hannan-consistent algorithms exist for both internal and external regrets (see Blackwell (1954), Hannan (1957), Fudenberg and Levine (1995, 1998), Foster and Vohra (1993, 1998, 1999), Freund and Schapire (1999), Hart and Mas-Colell (2000), and Young (2004)—their extension to

[9]This is also true for smooth variants of fictitious play, introduced in Section A.2.

[10]Hannan consistency is often referred to as *universal consistency*, after Fudenberg and Levine (1995).

incomplete information settings is straightforward. Importantly, however, external regret matching and its generalizations are Hannan-consistent, whereas internal regret matching and its generalizations are not. That is, external regret matching assures vanishing external regrets for any player who uses it irrespective of the behavior of the other players. In contrast, if a single player uses internal regret matching, there is no assurance that his internal regrets will vanish over time unless we assume that the other players use the same algorithm. As the previous remark points out, regret matching algorithms only require regrets as an input. It follows that to achieve ε-AER players need to know neither the payoff functions of the other players and their signals nor the algorithms that other players adopt. In contrast, to achieve ε-AIR: (i) if players do not know the payoff functions or the signals of other players, they need to know what regret-minimizing algorithm the other players adopt and to coordinate on this algorithm; (ii) if, instead, they do not know what other players do, they need to design more sophisticated algorithms to assure that their own internal regrets vanish—these algorithms requiring common knowledge of the underlying one-shot game.

Remark 6. Under internal regret matching, as internal regrets become small, so does the probability of switching (see (17)) to a different action. Hence, internal regret matching leads to longer and longer stretches of time in which the action is constant, and the play exhibits much *inertia* and infrequent switches (conditional on the player's signal). In particular, if all the regrets are zero, then a regret matching player will continue to play the same action of the previous period.

The full class of regret-based algorithms—for which, if played by all players, the sequence of actions, signals, and states has asymptotic ε-regret almost surely—is even larger. We refer to Hart (2005) for an extensive survey of regret-based algorithms for the repeated play of complete information games. Their extension to incomplete information games mimics that for internal and external regret matching and their generalizations outlined above. For instance, even procedures that fall in the class of some reinforcement learning algorithms (see, for example, Roth and Erev (1995), Börgers and Sarin (1997, 2000), and Erev and Roth (1998)), broadly defined as including those procedures whereby individuals react to past payoffs without full knowledge of the game, may lead to convergence results such as the ones we obtain in this paper. A relevant case is presented in Hart and Mas-Colell (2001b), who develop a modified regret matching algorithm, dubbed proxy-regret matching, for environments where each player initially knows only his own set of actions and is informed, after each period of play, only of his own realized payoff. However, not all regret-based algorithms are as simple as internal and external regret matching.

A.2 Calibrated Learning, Variants of Fictitious Play, and Other Dynamics

Previous work has identified several dynamics for the repeated play of a one-shot complete information game that converge to the set of correlated ε-equilibria of the underlying game. A notable example of such dynamics is *calibrated learning* by Foster and Vohra (1997)). Here, each player computes calibrated forecasts on the behavior of the other players, and then plays a best reply to these forecasts. Forecasts are calibrated if, roughly speaking, probabilistic forecasts and long-run frequencies are close: for example, an event must occur approximately $\pi\%$ of the times for which the forecast was a $\pi\%$ chance of the event. There are various ways to generate calibrated forecasts (see, among others, Foster and Vohra (1997, 1998, 1999), Foster (1997), Fudenberg and Levine (1999b), and Kakade and Foster (2008). Other notable examples are *conditional smooth fictitious play eigenvector strategies* (Fudenberg and Levine (1998, 1999a)) and *smooth conditional fictitious play* (Cahn (2004)), where each player i plays at each period a smoothed-out best reply to the distribution of the play of the opponents in those periods where i played the same action j as in the previous period.

It is possible to extend such dynamics to the repeated play of the one-shot incomplete information game (G, S) signal-by-signal, in the spirit of what we outline above for regret matching and its generalizations. Then we can then show that such dynamics converge to $BCE_{(G,S)}(\varepsilon)$ or $BCCE_{(G,S)}(\varepsilon)$. It follows from Theorem 1 that the sequence of actions, signals, and states has ε-AIR or ε-AER.

A.3 Repeated Equilibrium Play of (G, S)

Suppose that, in each period, players play a BNE of (G, S)—not necessarily the same. If so, the sequence of empirical distributions is in the convex hull of all BNE action-signal-state distributions of (G, S), which is contained in $BCE_{(G,S)}(\varepsilon)$, and so in $BCCE_{(G,S)}(\varepsilon)$. It follows from Theorem 1 that the sequence of actions, signals, and states has no-AIR and no-AER. Similarly, if players play a ε-BCE of (G, S) (resp., a ε-BCCE of (G, S)) in each period, then the sequence of actions, signals, and states has ε-AIR and ε-AER (resp., ε-AER).

B Proofs

B.1 Proof of Theorem 1–(ii)

[$\Longrightarrow$] Fix $\varepsilon \geq 0$ and suppose $((a^n, t^n, \theta^n))_{n\in\mathbb{N}}$ from $(G, S)^\infty$ has ε-AER almost surely. Consider any subsequence $\left(Z^{N_l}\right)_{l\in\mathbb{N}}$ of $\left(Z^N\right)_{N\in\mathbb{N}}$ that converges almost surely to some $\nu \in \Delta(A \times T \times \Theta)$. We need to show that ν is almost surely an ε-BCCE of (G, S), i.e. that ν is almost surely consistent and coarsely ε-obedient for (G, S).

Consistency. The proof of consistency is the same as for part (i) of Theorem 1.

Coarse ε-obedience. To begin, note the following:

$$\begin{aligned}
V_i^{ext}(k, t_i, N) - U_i(t_i, N) &= \frac{1}{N}\sum_{n=1}^{N}\left[u_i\big((k, a_{-i}^n), \theta^n)\big) - u_i\big((a_i^n, a_{-i}^n), \theta^n)\big)\right]\mathbb{1}_{\{t_i\}}(t_i^n) \\
&= \frac{1}{N}\sum_{\theta\in\Theta}\sum_{n=1}^{N}\left[u_i\big((k, a_{-i}^n), \theta^n)\big) - u_i(a^n, \theta^n)\right]\mathbb{1}_{\{t_i\}}(t_i^n)\mathbb{1}_{\{\theta\}}(\theta^n) \\
&= \sum_{a,t_{-i},\theta}\left[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\right]Z^N\big(a, (t_i, t_{-i}), \theta\big).
\end{aligned} \tag{18}$$

Now pick any $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$. As $\limsup_{N\to\infty} R_i^{ext}(k, t_i, N) \leq \varepsilon$ a.s., by definition of $R_i^{ext}(k, t_i, N)$, we also have $\limsup_{N\to\infty}\left[V_i^{ext}(k, t_i, N) - U_i(t_i, N)\right] \leq \varepsilon$ a.s. But then, by (18),

$$\limsup_{N\to\infty}\sum_{a,t_{-i},\theta}\left[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\right]Z^N\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon \qquad \text{a.s.} \tag{19}$$

Moreover, on the subsequence $\big(Z^{N_l}\big)_{l\in\mathbb{N}}$ we get

$$\begin{aligned}
&\lim_{l\to\infty}\sum_{a,t_{-i},\theta}\left[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\right]Z_l^N\big(a, (t_i, t_{-i}), \theta\big) \\
&= \sum_{a,t_{-i},\theta}\lim_{l\to\infty}\left[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\right]Z_l^N\big(a, (t_i, t_{-i}), \theta\big) \\
&= \sum_{a,t_{-i},\theta}\left[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\right]\nu\big(a, (t_i, t_{-i}), \theta\big).
\end{aligned} \tag{20}$$

Together, (19) and (20) give

$$\sum_{a,t_{-i},\theta}\left[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\right]\nu\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon \qquad \text{a.s.} \tag{21}$$

As $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$ were arbitrarily chosen, we conclude from (21) that ν is almost surely coarsely ε-obedient for (G, S).

[$\Longleftarrow$] Now suppose $\big(Z^N\big)_{N\in\mathbb{N}}$ converges almost surely to $BCCE_{(G,S)}(\varepsilon)$ for some $\varepsilon \geq 0$. Pick any $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$. By coarse ε-obedience,

$$\limsup_{N\to\infty}\sum_{a,t_{-i},\theta}\left[u_i\big((k, a_{-i}), \theta\big) - u_i(a, \theta)\right]Z^N\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon \qquad \text{a.s.} \tag{22}$$

By (18) and (22),

$$\limsup_{N\to\infty} V_i^{ext}(k, t_i, N) - U_i(t_i, N) \leq \varepsilon \qquad \text{a.s.,}$$

which implies

$$\limsup_{N\to\infty} R_i^{ext}(k, t_i, N) \leq \varepsilon \qquad \text{a.s.}$$

by definition of $R_i^{ext}(k, t_i, N)$. As $i \in \mathcal{I}$, $t_i \in T_i$, and $k \in A_i$ were arbitrarily chosen, the desired result follows. ■

B.2 Proof of Theorem 3

Fix $\varepsilon \geq 0$ and suppose $((f^n, \overline{f}^n, t^n, o^n))_{n\in\mathbb{N}}$ from the aggregative pricing game has ε-AIR almost surely. Consider any subsequence $\left(Z^{N_l}\right)_{l\in\mathbb{N}}$ of $\left(Z^N\right)_{N\in\mathbb{N}}$ that converges almost surely to some $\nu \in \Delta(F \times \overline{F} \times T \times O)$. We need to show that ν is almost surely an ε-ABCE of the aggregative pricing game, i.e. that ν is almost surely consistent and ε-obedient for the aggregative pricing game.

Consistency. Pick any $(t, o) \in T \times O$. Note the following:

$$\begin{aligned}\sum_{f,\overline{f}} \nu\big(f, \overline{f}, t, o\big) &= \sum_{f,\overline{f}} \lim_{l\to\infty} Z^{N_l}\big(f, \overline{f}, t, o\big) \\ &= \lim_{l\to\infty} \sum_{f,\overline{f}} Z^{N_l}\big(f, \overline{f}, t, o\big) \\ &= \lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{o\}}(o^n)}{N_l}.\end{aligned} \tag{23}$$

Under the assumptions of the theorem: (i) $(t^n)_{n\in\mathbb{N}}$ and $(o^n)_{n\in\mathbb{N}}$ are two irreducible and positive recurrent Markov chains with stationary distributions $\psi(\cdot;\lambda)$ and ϕ; (ii) t^n is independent of o^n for all n. Thus,

$$\lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{o\}}(o^n)}{N_l} = \psi(t;\lambda)\phi(o) \qquad \text{a.s.} \tag{24}$$

Together, (23) and (24) give

$$\sum_{(f,\overline{f})} \nu\big(f, \overline{f}, t, o\big) = \psi(t;\lambda)\phi(o) \qquad \text{a.s.} \tag{25}$$

As $(t, o) \in T \times O$ was arbitrarily chosen, we conclude from (25) that ν is almost surely consistent for the aggregative pricing game.

ε-obedience. To begin, note the following:

$$\begin{aligned}&V^{int}(f, f', t, o, N) - U(t, o, N) \\ &= \frac{1}{N}\sum_{n=1}^{N} \Big[u\big(f', \overline{f}^n + (f' - f^n), t^n, o^n\big) - u\big(f^n, \overline{f}^n, t^n, o^n\big)\Big] \mathbb{1}_{\{f\}}(f^n)\mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{o\}}(o^n) \\ &= \sum_{\overline{f}} \Big[u\big(f', \overline{f} + (f' - f), t, o\big) - u\big(f, \overline{f}, t, o\big)\Big] Z^N\big(f, \overline{f}, t, o\big).\end{aligned} \tag{26}$$

Now pick any $t \in T$, $o \in O$, and $f, f' \in F$. As $\limsup_{N\to\infty} R^{int}(f, f', t, o, N) \leq \varepsilon$ a.s., by definition of $R^{int}(f, f', t, o, N)$, we also have $\limsup_{N\to\infty} \big[V^{int}(f, f', t, o, N) - U(t, o, N)\big] \leq \varepsilon$ a.s. But then, by (26),

$$\limsup_{N\to\infty} \sum_{\overline{f}} \Big[u\big(f', \overline{f} + (f' - f), t, o\big) - u\big(f, \overline{f}, t, o\big)\Big] Z^N\big(f, \overline{f}, t, o\big) \leq \varepsilon \qquad \text{a.s.} \tag{27}$$

Moreover, on the subsequence $\left(Z^{N_l}\right)_{l\in\mathbb{N}}$ we get

$$\begin{aligned}
&\lim_{l\to\infty}\sum_{\overline{f}}\Big[u\big(f',\overline{f}+(f'-f),t,o\big)-u\big(f,\overline{f},t,o\big)\Big]Z^{N_l}\big(f,\overline{f},t,o\big)\\
&=\sum_{\overline{f}}\lim_{l\to\infty}\Big[u\big(f',\overline{f}+(f'-f),t,o\big)-u\big(f,\overline{f},t,o\big)\Big]Z^{N_l}\big(f,\overline{f},t,o\big) \qquad (28)\\
&=\sum_{\overline{f}}\lim_{l\to\infty}\Big[u\big(f',\overline{f}+(f'-f),t,o\big)-u\big(f,\overline{f},t,o\big)\Big]\nu\big(f,\overline{f},t,o\big).
\end{aligned}$$

Together, (27) and (28) give

$$\sum_{\overline{f}}\Big[u\big(f',\overline{f}+(f'-f),t,o\big)-u\big(f,\overline{f},t,o\big)\Big]\nu\big(f,\overline{f},t,o\big)\leq\varepsilon \qquad \text{a.s.} \qquad (29)$$

As $t\in T$, $o\in O$, and $f,f'\in F$ were arbitrarily chosen, we conclude from (29) that ν is almost surely ε-obedient for the aggregative pricing game. ■

References

Ackerberg, Daniel, C. Lanier Benkard, Steven Berry, and Ariel Pakes (2007), "Econometric Tools for Analyzing Market Outcomes." James J. Heckman and Edward E. Leamer eds., *Handbook of Econometrics*. Elsevier.

Arieli, Itai, Yakov Babichenko, and Rann Smorodinsky (2018), "Robust Forecast Aggregation." *Proceedings of the National Academy of Sciences*, 115, 67–96.

Aumann, Robert J. (1974), "Subjectivity and Correlation in Randomized Strategies." *Journal of Mathematical Economics*, 1, E12135–E12143.

——— (1987), "Correlated Equilibrium as an Expansion of Bayesian Rationality." *Econometrica*, 55, 1–18.

Babichenko, Yakov and Dan Garber (2018), "Learning of Optimal Forecast Aggregation in Partial Evidence Environments." *Working Paper*.

Benkard, Lanier (2004), "A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft." *The Review of Economic Studies*, 71, 581–611.

Bergemann, Dirk and Stephen Morris (2013), "Robust Predictions in Games with Incomplete Information." *Econometrica*, 11, 487—-522.

——— (2016), "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games." *Theoretical Economics*, 11, 487—-522.

Berry, Steven (1992), "Estimation of a Model of Entry in the Airline Industry." *Econometrica*, 60, 889–917.

Blackwell, David (1954), "Controlled Random Walks." *Proceedings of the International Congress of Mathematicians*, 3, 336–338.

——— (1956), "An Analog of the Minmax Theorem for Vector Payoffs." *Pacific Journal of Mathematics*, 6, 1–8.

Börgers, Tilman and Rajiv Sarin (1997), "Learning Through Reinforcement and Replicator Dynamics." *Journal of Economic Theory*, 77, 1—-14.

——— (2000), "Naive Reinforcement Learning With Endogenous Aspirations." *Internation Economic Review*, 41, 921—-950.

Bubeck, Sebastien and Nicolò Cesa-Bianchi (2012), "Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems." *Foundations and Trends in Machine Learning*, 5, 1–122.

Cahn, Amotz (2004), "General Procedures Leading to Correlated Equilibria." *International Journal of Game Theory*, 33, p. 21–40.

Caragiannis, Ioannis, Christos Kaklamanis, Panagiotis Kanellopoulos, Maria Kyropoulou, Brendan Lucier, Renato Paes Leme, and Éva Tardos (2015), "Bounding the Inefficieny of Outcomes in Generalized Second Price Auctions." *Journal of Economic Theory*, 157, 343—-388.

Cesa-Bianchi, Nicolò and Gàbor Lugosi (2006), "Prediction, Learning, and Games." *Cambridge University Press*.

Ciliberto, Federico and Elie Tamer (2009), "Market Structure and Multiple Equilibria in Airline Markets." *Econometrica*, 77, 1791–1828.

Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes (2018), "Just Starting Out: Learning and Equilibrium in a New Market." *The American Economic Review*, 108, 565–615.

Erev, Ido and Alvin E. Roth (1998), "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategies." *American Economic Review*, 88, 848—-881.

Ericson, Richard and Ariel Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *The Review of Economic Studies*, 62, 53–82.

Forges, Françoise (1993), "Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information." *Theory and Decision*, 77, 277–310.

Foster, Dean P. (1997), "Proof of Calibration via Blackwell's Approachability Theorem." *Games and Economic Behavior*, 29, 74–78.

Foster, Dean P. and Rakesh V. Vohra (1993), "A Randomization Rule for Selecting Forecasts." *Operations Research*, 41, 704–709.

——— (1997), "Calibrated Learning and Correlated Equilibrium." *Games and Economic Behavior*, 21, 40–55.

——— (1998), "Asymptotic Calibration." *Biometrika*, 85, 379–390.

——— (1999), "Regret in the On-Line Decision Problem." *Games and Economic Behavior*, 29, 7–35.

Freund, Yoav and Robert E. Schapire (1999), "Adaptive Game Playing Using Multiplicative Weights." *Games and Economic Behavior*, 29, 79–103.

Fudenberg, Drew and David K. Levine (1995), "Consistency and Cautious Fictitious Play." *Journal of Economic Dynamics and Control*, 19, 1065–1090.

——— (1998), "The Theory of Learning in Games." *MIT Press.*

——— (1999a), "Conditional Universal Consistency." *Games and Economic Behavior*, 29, 104–130.

——— (1999b), "An Easier Way to Calibrate." *Games and Economic Behavior*, 29, 131–137.

Gualdani, Cristina and Shruti Sinha (2019), "Identification and inference in discrete choice models with imperfect information." *Working Paper.*

Hannan, James (1957), "Approximation to Bayes Risk in Repeated Play." M. Dresher, A.W. Tucker, and P. Wolfe eds., *Contributions to the Theory of Games, Vol. 3.* Princeton University Press.

Hart, Sergiu (2005), "Adaptive Heuristics." *Econometrica*, 73, 1401–1430.

Hart, Sergiu and Yishay Mansour (2010), "How Long to Equilibrium? The Communication Complexity of Uncoupled Equilibrium Procedures." *Games and Economic Behavior*, 69, 107–126.

Hart, Sergiu and Andreu Mas-Colell (2000), "A Simple Adaptive Procedure Leading to Correlated Equilibrium." *Econometrica*, 68, 1127–1150.

——— (2001a), "A General Class of Adaptive Strategies." *Journal of Economic Theory*, 98, 26–54.

——— (2001b), "A Reinforcement Procedure Leading to Correlated Equilibrium." Gerard Debreu, Wilhelm Neuefeind, and Walter Trockel eds., *Economic Essays: A Festschrift for Werner Hildenbrand*. Springer.

——— (2013), "Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics." *World Scientific Series in Economic Theory*.

Hartline, Jason, Vasilis Syrgkanis, and Eva Tardos (2015), "No-Regret Learning in Bayesian Games." *NIPS 2015*.

Jofre-Bonet, Mireia and Martin Pesendorfer (2003), "Estimation of a Dynamic Auction Game." *Econometrica*, 71, 1443–1489.

Kakade, Sham M. and Dean P. Foster (2008), "Deterministic Calibration and Nash Equilibrium." *Journal of Computer and System Science*, 74, 115–130.

Lattimore, Tor and Csaba Szepesvári (2019), "Bandit Algorithms." *Preprint availtable at https://tor-lattimore.com/downloads/book/book.pdf*.

Lee, Robin S. and Ariel Pakes (2009), "Multiple Equilibria and Selection by Learning in an Applied Setting." *Economics Letters*, 104, 13–16.

Magnolfi, Lorenzo and Camilla Roncoroni (2019), "Estimation of Discrete Games with Weak Assumptions on Information." *Working Paper*.

Manski, Charles F. (2004), "Statistical Treatment Rules for Heterogeneous Populations." *Econometrica*, 72, 1221–1246.

Moulin, Hervé and Jean-Philippe Vial (1978), "Strategically Zero-Sum games: The Class of Games whose Completely Mixed Equilibria Cannot Be Improved Upon." *Internation Journal of Game Theory*, 7, 201–221.

Nekipelov, Denis, Vasilis Syrgkanis, and Eva Tardos (2015), "Econometrics for Learning Agents." *EC 2015*.

Nisam, Noam, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani (2007), "Algorithmic Game Theory." *Cambridge University Press*.

Nisan, Noam and Gali Noti (2017a), "An experimental evaluation of regret-based econometrics." *Proceedings of the 26th International Conference on World Wide Web*, 73–81, International World Wide Web Conferences Steering Committee.

——— (2017b), "A Quantal Regret Method for Structural Econometrics in Repeated Games." *Proceedings of the 2017 ACM Conference on Economics and Computation*, 123–123, ACM.

Nocke, Volker and Nicolas Schutz (2018), "Multiproduct-Firm Oligopoly: An Aggregative Games Approach." *Econometrica*, 86, 523–557.

Pakes, Ariel (2016), "Methodological Issues in Analyzing Market Dynamics." Frank Thuijsman and Florian Wagener eds., *Advances in Dynamic and Evolutionary Games: Theory, Applications, and Numerical Methods*. Springer International Publishing.

Roth, Alvin E. and Ido Erev (1995), "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term." *Games and Economic Behavior*, 8, 164—-212.

Roughgarden, Tim (2016), "Twenty Lectures on Algorithmic Game Theory." *Cambridge University Press*.

Ryan, Stephen P. (2012), "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica*, 80, 1019–1061.

Shalev-Shwartz, Shai (2011), "Online Learning and Online Convex Optimization." *Foundations and Trends in Machine Learning*, 4, 107–194.

Shoam, Yoav and Kevin Leyton-Brown (2008), "Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations." *Cambridge University Press*.

Stoye, Jörg (2009), "Minimax Regret Treatment Choice with Finite Samples." *Journal of Econometrics*, 151, 70–81.

Syrgkanis, Vasilis, Elie Tamer, and Juba Ziani (2018), "Inference on Auctions with Weak Assumptions on Information." *Working Paper*.

Tamer, Elie (2003), "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria." *The Review of Economic Studies*, 70, 147–165.

Weintraub, Gabriel Y, C Lanier Benkard, and Benjamin Van Roy (2008), "Markov Perfect Industry Dynamics with Many Firms." *Econometrica*, 76, 1375–1411.

Young, H. Peyton (2004), *Strategic Learning and Its Limits*. Oxford University Press.