Estimation of Games under No Regret[*]

Niccolò Lomys[†]   Lorenzo Magnolfi[‡]   Camilla Roncoroni[§]

November 14, 2021

Abstract

We develop a method to estimate a game's primitives in complex dynamic environments. Because of the environment's complexity, agents may not know or understand some key features of their interaction. Instead of equilibrium assumptions, we impose an asymptotic ε-regret (ε-AR) condition on the observed play. According to ε-AR, the time average of the counterfactual increase in past payoffs, had each agent changed each past play of a given action with its best replacement in hindsight, becomes small in the long run. We first prove that the time average of play satisfies ε-AR if and only if it converges to the set of Bayes correlated ε-equilibrium predictions of the stage game. Next, we use the static limiting model to construct a set estimator of the parameters of interest. The estimator's coverage properties directly arise from the theoretical convergence results. The method applies to panel data as well as to cross-sectional data interpreted as long-run outcomes of learning dynamics. We apply the method to pricing data in an online marketplace. We recover bounds on the distribution of sellers' marginal costs that are useful to inform policy experiments.

Keywords: Empirical Games; Bayes (Coarse) Correlated Equilibrium; Learning in Games; Regret Minimization; Partial Identification; Incomplete Models.

JEL Classification: C1; C5; C7; D4; D8; L1; L8.

[†]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

Examining data generated by strategic interaction through the lens of equilibrium models is a standard approach in empirical economics. This approach is used, with appropriate modifications, both for data in which the interaction is observed over time and for cross-sectional data on outcomes across different markets. Applications include the study of firms' pricing, entry, investment and technology adoption decisions, and bidding in auctions, both over time and in a cross-section.[1]

While appropriate in some settings, the equilibrium approach may be unsatisfactory in others. For instance, in many empirically relevant dynamic environments, agents may find it hard to know or understand some key features of the strategic interaction. Leading examples include real-time pricing in online user-to-user marketplaces and bidding in sponsored search auctions. We call these *complex dynamic environments*. These are characterized by three salient features. First, agents may not know their opponents' identities and incentives and may not have a prior on their private information and beliefs. Second, monitoring others' actions may be hard or impossible. Third, it may be difficult to know how the environment evolves over time. Due to complexity or lack of information, such environments are hard to predict and to model, and agents may be unable to form correct beliefs on their opponents' behavior. Therefore, although agents try to learn and adapt to their strategic environment, they may not behave according to a static or dynamic equilibrium notion.[2] Thus, a researcher that observes the full strategic interaction may want to relax equilibrium assumptions.

Additionally, what is modeled as cross-sectional data generated by independent play of static games may instead arise from an interaction that happens over time, with linkages across periods. We call these *long-run outcomes*. A salient example is entry in markets, in which firms decide to enter different markets at different times. As the researcher cannot typically determine the exact timing of different entry decisions, these data are commonly treated as cross-sectional and interpreted as long-run equilibrium outcomes (Ciliberto and Tamer, 2009).[3] In

---

[1]For seminal contributions, see Berry, Levinsohn, and Pakes (1995) for pricing, Bresnahan and Reiss (1991) for entry, Ackerberg and Gowrisankaran (2006) for technology adoption, and Guerre, Perrigne, and Vuong (2000) for auctions.

[2]This is not because agents are boundedly rational, but rather because the environment's characteristics make equilibrium play hard. Unsurprisingly, there is an increasing reliance on specialized pricing algorithms in these environments (see, e.g., Chen, Mislove, and Wilson, 2016).

[3]Specifically, 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. The objective of our econometric analysis is to infer long-run relationships between the exogenous variables in the data

such contexts, imposing equilibrium assumptions on a static game-theoretic structure ignores that players may have had different information at different times, and adapted their play by learning across markets.

In this paper, we develop a method to estimate a game's primitives in settings in which agents interact over time, encompassing both complex dynamic environments and long-run outcomes. Our method deemphasizes an equilibrium approach and instead pursues an adaptive approach based on the minimization of ex-post regret notions. In particular, we build on the idea that learning outcomes can be used as an alternative to solution concepts when analyzing the outcomes of game dynamics. This strategy is consistent with recent advances in robust analysis in economic theory and computer science (see references in Section 1.1).

In practice, we model agents interacting over discrete periods in an incomplete information environment (complete information is a special case of our setting). We only assume that agents' behavior satisfies the minimal long-run optimality condition of *asymptotic ε-regret* (hereafter, ε-AR). According to ε-AR, in the long run, the time average payoff of each type of each agent becomes ε-close to the payoff that type of that agent would have obtained by changing each past play of a given action with its best replacement in hindsight.[4] Intuitively, ε-AR requires the no ex post regret property of Nash equilibrium to hold, approximately, over sequences of play with respect to a specific benchmark policy in hindsight; however, ε-AR does not require agents' actions to be stable and independent. Importantly, ε-AR is a weaker requirement than equilibrium: whenever agents play any equilibrium of the underlying stage game, they also minimize their regrets.

Regret minimization (i.e., outperforming in hindsight a given benchmark strategy) is the leading criterion to evaluate performance in complex and unstructured dynamic environments, as reflected by a large literature in game theory and computer science (see references in Section 1.1). The ε-AR condition is satisfied by many procedures, once they are adapted to the play of games with incomplete information. This class includes simple adaptive heuristics, fictitious-play-like dynamics, calibrated learning, more sophisticated learning rules involving experimentation, specialized learning algorithms, and equilibrium play; several regret-minimizing procedures can also be cast as reinforcement learning rules.[5]

---

and the market structure that we observe at some point in time, without trying to explain how firms reached the observed equilibrium."

[4]We discuss alternative related notions of regret throughout the paper.

[5]See Section 3.1 for a brief overview of regret-minimizing procedures, their properties, and references to the original contributions.

The ε-AR condition can be satisfied in environments of arbitrary complexity and with almost no knowledge of the environment—each agent need not know more than his set of actions and (an estimate of) his realized payoff at the end of each period. However, we neither exclude that agents know or observe more nor assume that they coordinate on a specific regret minimization procedure. We only assume that each agent learns to interact in/adapt to the environment sufficiently well for ε-AR to hold. In this sense, ours is an incomplete model (Tamer, 2003).

Under the ε-AR assumption, we develop an empirical strategy to estimate the structural parameters of the stage game. We do so in two steps. First, we show that the time average of play satisfies ε-AR if and only if it converges to the set of Bayes correlated ε-equilibrium (hereafter, ε-BCE) predictions of the stage game (not necessarily to an element in this set). The ε-BCE notion extends Bayes Correlated Equilibrium (Bergemann and Morris, 2016) by requiring incentive constraints to only hold approximately. To establish the convergence result, we generalize to incomplete information environments earlier results on dynamic foundations for correlated equilibrium (Aumann, 1974, 1987) in games with complete information (e.g., Foster and Vohra, 1997; Hart and Mas-Colell, 2000, 2001; Fudenberg and Levine, 1995, 1999).

In the second step, we build on the theoretical convergence result to recover the parameters with a set estimator based on the restrictions implied by the static notion of ε-BCE. The method applies to panel data as well as to cross-sectional data interpreted as long-run outcomes of learning dynamics. In contrast to the existing literature that uses the BCE notion to develop methods to make inferences in static games (Magnolfi and Roncoroni, 2021; Syrgkanis, Tamer, and Ziani, 2021; Gualdani and Sinha, 2020; Bergemann, Brooks, and Morris, 2021), we do not pursue robustness to assumptions on information. Rather, we use ε-BCE restrictions for estimating dynamic interactions, or their long-run outcomes, in complex environments under weak assumptions on behavior, thus interpreting data generated by a dynamic interaction with a static limiting model.

Based on our convergence result, we show that the estimator almost surely contains the true parameter value after a sufficient number of periods. Moreover, by leveraging the features of the data-generating process, we show how to obtain tighter bounds on parameters by taking the intersection of set estimators computed at different points in time when panel data are available. We also construct theoretical bounds on how informative the estimated set of parameters can be. We provide Monte Carlo evidence for a repeated binary pricing game. Our method

produces tight estimated sets of parameters describing the distribution of sellers' marginal costs.

Our method is broadly applicable. For instance, it is well suited to study pricing behavior in e-commerce marketplaces. A sizable portion of online retail happens via decentralized platforms where sellers set prices independently. Examples of these environments are Amazon and Walmart marketplaces, eBay, and Etsy. In an empirical application, we use our method to recover the distribution of sellers' marginal cost on Swappa—an online marketplace for used cellphones and other portable electronic devices. Because of the complexity and the availability of pricing algorithms, relying on the ε-AR assumption is an appealing approach to recover the distribution of sellers' marginal costs in this environment. Our method produces estimates of seller-specific average cost which are plausible in this economic environment.

The primitives that we recover can inform policy exercises and counterfactuals. The distribution of marginal costs allows researchers to evaluate the degree of competitiveness of pricing in the platform by comparing prices to a Bertrand equilibrium and a perfect collusion benchmark. Moreover, the distribution of marginal costs enables simulations of the price path according to widely used algorithms in e-commerce applications. This counterfactual provides useful inputs to the design decision of the platform, which may want to make these algorithms available (or mandatory) to sellers. Finally, there is a growing literature on pricing algorithms and their effects on market outcomes (see, e.g., Assad, Calvano, Calzolari, Clark, Denicolò, Ershov, Johnson, Pastorello, Rhodes, Xu, and Wildenbeest, 2021, and references therein). Evaluating algorithmic pricing in this context, where primitives are obtained with an algorithm-independent method, would contribute to the ongoing debate.

1.1 Related Literature

We contribute to a recent literature that estimates empirical models of learning agents (for a survey, see Aguirregabiria and Jeon, 2020). Complementary to model-based and belief-based approaches (e.g., Doraszelski, Lewis, and Pakes, 2018; Aguirregabiria and Magesan, 2020), our method is based on regret-minimizing agents who need not specify a model of other agents' behavior, nor need to form beliefs about their information or behavior. This is in the spirit of incomplete models (Tamer, 2003; Haile and Tamer, 2003; Ciliberto and Tamer, 2009) and results in set-valued estimators of parameters.

A recent strand of the computer science literature explores the connection between regret minimization and empirical work. In an online auction environment, Nekipelov, Syrgkanis, and Tardos (2015) characterize and perform inference on the set of valuations consistent with a given level of regret, but without relying on an equilibrium concept. Nisan and Noti (2017a,b) evaluate a similar approach with experimental data and propose adjustments to the ε-regret estimation procedure. Noti and Syrgkanis (2021) find that econometric approaches based on regret minimization compare favorably to both machine learning methods and standard equilibrium-based econometrics when used to predict bids in online ad auctions. In contrast to these papers, our general econometric approach leverages convergence results to interpret the data through the lens of the static equilibrium notion of ε-BCE. Thus, our method easily applies to any underlying stage game and does not require relying on the specific features of a given empirical application (e.g., of the online auction environment or the online pricing game).

Regret minimization relative to some benchmark has been used as a criterion to develop robust approaches to decision making (e.g., Savage, 1951, 1972) and optimal treatment choice (e.g., Manski, 2004, 2007, 2021; Stoye, 2007, 2009; Schlag, 2007; Hirano and Porter, 2009). It is also a central idea in robust contracting and robust mechanism design (e.g., Hurwicz and Shapiro, 1978; Bergemann and Schlag, 2008, 2013; Renou and Schlag, 2011; Chassang, 2013; Caldentey, Liu, and Lobel, 2017; Beviá and Corchón, 2019; Guo and Shmaya, 2019, 2021; Chassang and Kapon, 2021; Braverman, Mao, Schneider, and Weinberg, 2018; Deng, Schneider, and Sivan, 2019; Camara, Hartline, and Johnsen, 2020). These papers use regret-based objectives to characterize optimal decision rules, treatment rules, or mechanisms in complex environments in which agents find it hard to form prior beliefs. In contrast, we use the minimization of a regret-based objective to impose a minimal optimality condition on dynamic behavior with the aim of inferring economic primitives in complex environments.

A large and growing literature at the intersection of economics and computer science studies learning and regret minimization. In particular, regret minimization is a leading approach in sequential decision problems (or online learning) and multi-armed bandit problems (see, e.g., Foster and Vohra, 1999; Cesa-Bianchi and Lugosi, 2006; Shalev-Shwartz, 2011; Bubeck and Cesa-Bianchi, 2012; Slivkins, 2019; Lattimore and Szepesvári, 2020; Hazan, 2021, and references therein) and a central idea in multiagent learning and learning in games (see, e.g., Vohra and Wellman, 2007; Nisam, Roughgarden, Tardos, and Vazirani, 2007; Roughgarden,

2016; Fudenberg and Levine, 1998; Young, 2004; Hart and Mas-Colell, 2013, and references therein). We contribute to this literature by providing novel results on the convergence properties of regret-minimizing dynamics in games with incomplete information.

1.2 Road Map

In Section 2, we present the model and the econometric problem we aim at providing an answer to. In Section 3, we formalize the asymptotic ε-regret property and study convergence under this property to the set of ε-BCE of the underlying stage game. In Section 4, we develop our econometric approach. In Section 5, we present our empirical application. In Section 6, we discuss additional results and extensions. In Section 7, we conclude. Proofs and omitted details are in the Appendices.

2 Model

In this section, we first describe the general model. Next, we present the econometric problem we aim at providing an answer to.

2.1 Basic Setup

There is a finite set $\mathcal{I}$ of I players, $\mathcal{I} := \{1, \ldots, I\}$, and i is a typical player. A *basic game G* consists of: (i) a finite set of payoff states, Θ; (ii) for each player i, a finite set of actions A_i, where we define $A := A_1 \times \cdots \times A_I$, and a payoff function $u_i : A \times \Theta \to \mathbb{R}$; and (iii) a full-support probability distribution $\psi \in \Delta_{++}(\Theta)$. Thus, $G := \left(\Theta, (A_i, u_i)_{i=1}^{I}, \psi\right)$. An *information structure S* consists of: (i) for each player i, a finite set of signals (or types) T_i, where we define $T := T_1 \times \cdots \times T_I$; and (ii) a signal distribution $\pi : \Theta \to \Delta(T)$. Thus, $S := \left((T_i)_{i=1}^{I}, \pi\right)$. We refer to the pair (G, S) as the *(incomplete information) stage game*.

We denote by θ, a_i, a, t_i, and t typical elements of sets Θ, A_i, A, T_i, and T (or of their subsets). We write $\pi(\cdot \,|\, \theta)$ for the probability distribution on T when the payoff state is θ and $\pi(t \,|\, \theta)$ for the probability of signal profile t when the payoff state is θ. We denote by a_{-i} a profile of actions for players other than i, i.e., $a_{-i} := (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_I)$, and by t_{-i} a profile of signals for players other than i.

Players interact over discrete time periods $n \in \mathbb{N} := \{1, 2, \ldots\}$ in some instance of game (G, S).

Assumption 1. *We maintain the following assumptions on what players observe about the game, on the timing of events, and on the sequence of payoff states.*

1. *Each player i knows his own set of actions A_i.*

2. *Within each period n, the following occurs in sequence:*

   (i) *State θ^n realizes;*

   (ii) *A profile of signals $(t_1^n, \ldots, t_I^n)$ 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;*

   (iv) *Payoffs realize and each player i observes (an estimate of) his own realized payoff $u_i(a^n, \theta^n)$.*

3. *Almost all paths of the process for payoff states $(\theta^n)_{n \in \mathbb{N}}$ have a limiting empirical distribution $\psi^L \in \Delta_{++}(\Theta^L)$ for some $\Theta^L \subseteq \Theta$.*

We denote by $(G, S)^\infty$ the game defined by (G, S) and Assumption 1 and refer to it as the *dynamic game*. The history of $(G, S)^\infty$ at the end of period N is the random sequence $(\theta^1, t^1, a^1, \ldots, \theta^N, t^N, a^N)$. Let $H^N := (A \times T \times \Theta)^N$ be the set of all possible histories at the end of period N, $H := \cup_{N \geq 1} H^N$ the set of all finite histories, and $H^\infty := (A \times T \times \Theta)^\infty$ the set of all possible infinite histories. We refer to $((\theta^n, t^n, a^n))_{n \in \mathbb{N}} \in H^\infty$ as a *sequence of states, signals, and actions from $(G, S)^\infty$*.

Let $T_i^L := \{t_i \in T_i : \pi(t_i, t_{-i} \mid \theta) > 0 \text{ for some } \theta \in \Theta^L\}$, $T^L := T_1^L \times \cdots \times T_I^L$, and π^L be the restriction of π to Θ^L. We refer to game (G^L, S^L), with $G^L := (\Theta^L, (A_i, u_i)_{i=1}^I, \psi^L)$ and $S^L := ((T_i^L)_{i=1}^I, \pi^L)$, as the *limiting stage game*.

When Θ is a singleton, (G, S) is a game with *complete information*, and so is (G^L, S^L). Moreover, many stochastic processes for payoff states satisfy part 1 of Assumption 1. Examples include: an i.i.d. process for payoff states, a perfectly persistent payoff state (in which case (G^L, S^L) is a game with complete information), payoff states that follow a Markov chain with limiting distribution, payoff states that follow a periodic Markov chain. In particular, the stochastic environment need not be stationary. Thus, the results that we develop in this paper apply to a wide variety of stochastic environments with incomplete information as well as to complete information environments.

We note however that the setting described in Assumption 1 does not include an important class of dynamic environments used in the applied literature. These are the environments with a payoff state that evolves endogenously as a function not only of the current state, but also of players' actions in the current period (e.g., Rust, 1987; Ericson and Pakes, 1995; Pakes, Ostrovsky, and Berry, 2007;

Aguirregabiria and Mira, 2007; Bajari, Benkard, and Levin, 2007; Pesendorfer and Schmidt-Dengler, 2008). We leave an extension of our method to these settings as future work.

We only assume that each player knows his own set of actions and receives as feedback (an estimate of) his own realized payoff at the end of each period. We do not make any further knowledge or monitoring assumption. Players may observe, know, or understand more, but need not do so. In particular:

- The (stage or dynamic) game need not be knowledge or common knowledge among players. More specifically: (i) players need not know how many and which opponents they are facing; (ii) players need not know what their opponents' payoff functions are; (iii) players need not know what the state space of the process for payoff states is or how the process evolves over time (i.e., they need not know Θ, ψ, Θ^L, or ψ^L)—in particular, players need not have a prior, let alone a common one; (iv) players need not know how payoff states are mapped into signals (i.e., they need not know π).

- Players need not observe the realized state θ^n at the end of each period.

- Players need not observe the profile of actions a^n that has been played at the end of each period (i.e., players need not monitor their opponents' actions).

Our minimal assumptions on what players know about the environment are not meant to capture a form of bounded rationality; indeed, we do not exclude that agents know more than what is specified by Assumption 1. Rather, these minimal assumptions capture the idea that the environment may be complex enough that players find it hard or impossible to know or understand some key features of their strategic interaction.

2.2   Econometric Problem

Given a dynamic game $(G, S)^\infty$, a researcher knows the corresponding limiting stage game up to some structural parameters. More formally, the limiting stage game belongs to a parametric class of games

$$\left\{ \left(G^L(\lambda^G), S^L(\lambda^S) \right) \right\}_{\lambda \in \Lambda}$$

indexed by structural parameters $\lambda := (\lambda^G, \lambda^S) \in \Lambda$, where Λ is a non-empty set. The true structural parameters in the data generating process are

$$\lambda_0 := (\lambda_0^G, \lambda_0^S)$$

and are unknown to the researcher.

We consider the two following data environments.

- Dataset D1: for some positive integer K, the researcher observes a realized sequence of actions $(a^n)_{n=1}^K \in A^K$ from the dynamic game $(G, S)^\infty$.

- Dataset D2: for some positive integer K, the researcher observes an empirical distribution of actions $q^K \in \Delta(A)$, defined pointwise by

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

for all $a \in A$, from the dynamic game $(G, S)^\infty$ (possibly without information about the timing of actions).[6]

Note that dataset D2 can be constructed from dataset D1, but not viceversa.[7]

Our goal is to develop a method to recover λ_0 under Assumption 1 when the researcher has access to dataset D1 or D2. We will do so by imposing an asymptotic ε-regret property on sequences $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ and by exploiting novel theoretical results on the convergence of regret-minimizing dynamics in games with incomplete information, which we develop in Section 3.

2.3 Illustration: A Two-Seller Pricing Game

The pricing game we use as illustrative example is a simplified version of that in our empirical application in Section 5.

Basic Setup. There are two sellers of a differentiated good, $i = 1, 2$. In each period n, each seller i privately observes the marginal cost t_i^n of selling a unit of his good and then sets a price $p_i^n \in \{\underline{p}, \overline{p}\}$, where $0 < \underline{p} < \overline{p}$. In each period n, if prices are $p^n := (p_1^n, p_2^n)$, seller i faces demand $g_i(p_i^n, p_{-i}^n)$ for his good. Seller i's profit in period n is $g_i(p_i^n, p_{-i}^n)(p_i^n - t_i^n)$. In period $n = 1$, marginal costs are drawn from a probability distribution $\tilde{\psi}$ with finite support $\widehat{T}$.

In the language of our general model, we have the following. The set of players is $\mathcal{I} = \{1, 2\}$. In the basic game G, we have $\Theta = \widehat{T}^2$, $A_1 = A_2 = \{\underline{p}, \overline{p}\}$, and $\psi = \tilde{\psi}$, and $u_i((p_i, p_{-i}), \theta) = g_i(p_i, p_{-i})(p_i - t_i)$. In the information structure S,

---

[6]That is, $q^N(a)$ is the empirical frequency of the action profile a in the first K periods.

[7]The assumption that only actions are observable to the researcher is the most common in the empirical literature. However, it is not the only one. For instance, Bergemann and Morris (2013) consider a setting where both actions and payoff states are observable to the researcher.

we have $T = \Theta$ and $\pi \colon \Theta \to \Delta(T)$ is such that $\pi(t \,|\, \theta) = 1$ if and only if $t = \theta$. Regarding the relationship between the stage game (G, S) and the limiting stage game (G^L, S^L), consider the following (non-exhaustive) possibilities:

- If Θ is a singleton (i.e., $\Theta = T_1 \times T_2 = \{(t_1, t_2)\}$), then $(G, S) = (G^L, S^L)$ and both are games with complete information.

- If $|\Theta| > 1$ and marginal costs are drawn in period $n = 1$ from $\tilde{\psi}$ and perfectly persistent thereafter—let $(\tilde{t}_1, \tilde{t}_2)$ denote the realization of the two sellers' marginal costs in period $n = 1$—then $(G, S) \neq (G^L, S^L)$, (G, S) is a game with incomplete information, and (G^L, S^L) is a game with complete information. In particular, $\Theta^L = T_1^L \times T_2^L = \{(\tilde{t}_1, \tilde{t}_2)\}$, ψ^L is trivial, and π^L is the restriction of π to Θ^L.

- If $|\Theta| > 1$ and marginal costs are i.i.d. over time, then $(G^L, S^L) = (G, S)$ and both are games with incomplete information.

- Suppose $|\Theta| > 1$ and the process for payoff states $(\theta^n)_{n \in \mathbb{N}}$ is an irreducible and aperiodic Markov chain with initial distribution $\tilde{\psi}$ and stationary distribution ψ^L. If $\tilde{\psi} = \psi^L$, then $(G^L, S^L) = (G, S)$; If $\tilde{\psi} \neq \psi^L$, then $(G^L, S^L) \neq (G, S)$.

Each seller i knows the prices he can set and, in each period, observes his own marginal cost of selling and receives as feedback (an estimate of) his own realized profit at the end of the period. We do not make any further knowledge or monitoring assumption. Sellers may or may not observe, know, or understand more about the pricing game being played. For instance, sellers may or may not know the demand function, their opponent's profit function, the distribution of marginal costs, and how payoff states evolve over time.

Econometric Problem. Suppose the researcher can estimate "offline" the demand function. For some positive integer K, the researcher observes: (i) a realized sequence of prices $(p^n)_{n=1}^{K}$ (dataset D1); or (ii) an empirical distribution of prices q^K, where $q^K(p) := \frac{1}{K} \sum_{n=1}^{K} \mathbb{1}_{\{p\}}(p^n)$ for all $p \in \{\underline{p}, \overline{p}\}^2$ (dataset D2). The researcher aims at recovering the distribution of the two sellers' marginal costs in the limiting stage game, that is, $\lambda_0^G = \psi^L$.

3 Asymptotic ε-Regret and Static ε-Equilibria

In this section, we first formalize the notions of regret and asymptotic ε-regret property. Next, we study convergence to the set of ε-BCE of the limiting stage

game under the asymptotic $\boldsymbol{\varepsilon}$-regret property. To simplify the exposition, in this section, given a dynamic game $(G, S)^\infty$, we denote the corresponding limiting stage game simply as (G^L, S^L), omitting the indexing with the true structural parameter λ_0. Since the results in this section are purely game-theoretical, no confusion will arise.

3.1 Regrets and the Asymptotic $\boldsymbol{\varepsilon}$-Regret Property

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

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

Let a_i be the last action played by player i with signal t_i up to time N. For each action $a_i' \in A_i$, let $V_i(a_i, a_i', t_i, N)$ be the average *counterfactual payoff* player i with signal t_i would have obtained had he played a_i' instead of a_i every time in the past that he actually played a_i. That is,

$$V_i(a_i, a_i', t_i, N) := \frac{1}{N} \sum_{n=1}^{N} v_i^n(a_i, a_i', t_i),$$

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

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

Definition 1 (Regret)**.** *For all $i \in \mathcal{I}$, $t_i \in T_i$, and $a_i, a_i' \in A_i$, the regret of player i with signal t_i for action a_i' with respect to action a_i before play at time $N + 1$ is denoted by $R_i(a_i, a_i', t_i, N)$ and defined by*

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

$R_i(a_i, a_i', t_i, N)$ 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 a_i was played in the past, the different action a_i'.

Notation. Let $\boldsymbol{\varepsilon} := (\varepsilon_i(a_i, a_i', t_i))_{i \in \mathcal{I}, a_i, a_i' \in A_i, t_i \in T_i^L}$ denote the vector that specifies, for all $i \in \mathcal{I}$, $a_i, a_i' \in A_i$, and $t_i \in T_i^L$, a non-negative real number $\varepsilon_i(a_i, a_i', t_i)$. We write $\boldsymbol{\varepsilon} = 0$ if $\varepsilon_i(a_i, a_i', t_i) = 0$ for all $i \in \mathcal{I}$, $a_i, a_i' \in A_i$, and $t_i \in T_i^L$. We write $\boldsymbol{\varepsilon}' \geq \boldsymbol{\varepsilon}$ if $\varepsilon_i'(a_i, a_i', t_i) \geq \varepsilon_i(a_i, a_i', t_i)$ for all $i \in \mathcal{I}$, $a_i, a_i' \in A_i$, and $t_i \in T_i^L$. The expressions $\boldsymbol{\varepsilon}' = \boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}' > \boldsymbol{\varepsilon}$ are defined analogously.

Definition 2 (Asymptotic ε-Regret Property). *Fix an ε. A sequence $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the asymptotic ε-regret (hereafter, ε-AR) property if*

$$\limsup_{N \to \infty} R_i(a_i, a_i', t_i, N) \leq \varepsilon_i(a_i, a_i', t_i) \tag{1}$$

for all $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i, a_i' \in A_i$. We refer to dynamics satisfying the ε-AR property as ε-regret dynamics.

The ε-AR property is a minimal optimality conditions for the play of $(G, S)^\infty$ capturing the idea that players learn in/adapt to their complex strategic environment. If the ε-AR property holds, the average regret experienced by each player for any of his signals for not having replaced all past plays of a given action with an arbitrary different action, separately for each of the player's actions, is "close" to vanish in the long-run; the meaning of close is quantified by the values specified by the vector ε.

Intuitively, the ε-AR property requires the no ex post regret property of Nash equilibrium to hold, approximately, over sequences $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ with respect to a specific benchmark policy in hindsight; in contrast to Nash equilibrium, however, the ε-AR property does not require players' actions to necessarily be stable and independent. Importantly, the ε-AR property is a weaker requirement than equilibrium play: whenever players reach or play any Bayes Nash or Bayes Correlated ε-equilibrium of (G^L, S^L), the resulting sequence $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the ε-AR property (see Section 3.2 for the formal statement).

Assuming that players minimize their regrets amounts to the following two observations. First, players are sufficiently sophisticated and rational to use any information they have about the environment to succeed in minimizing their regrets. Second, players may not reach a fully cooperative equilibrium of the dynamic game (if such equilibria exist and have positive regret).[8] Such failure to reach a fully cooperative equilibrium, however, need not depend on player's bounded rationality; rather, important economic environments are so complex and hard to predict that reaching a fully cooperative equilibrium of the dynamic game may be too hard or even impossible.[9]

---

[8]In many repeated games, if players are sufficiently patient, there are cooperative (or collusive) equilibria which Pareto-dominate all equilibria of the stage game and have positive regrets.

[9]In the computer science literature, recent results on the price of anarchy show that ε-regret dynamics have "near-optimal" welfare properties in many games (e.g., Blum, Hajiaghayi, Ligett, and Roth, 2008; Roughgarden, 2009; Hartline, Syrgkanis, and Tardos, 2015; Caragiannis, Kaklamanis, Kanellopoulos, Kyropoulou, Lucier, Paes Leme, and Tardos, 2015). Although they are are developed in different contexts than ours, such arguments suggest that ε-regret

The vector ε measures, for each of the player's signals and actions, how much each player's behavior departs from perfect regret minimization. Consistent with the previous observations, such departures can have two explanations: (i) players are not exact optimizers; (ii) players are sustaining some form of cooperation beyond that allowed by exact regret minimization despite the environment's complexity.

Regret minimization (i.e., outperforming in hindsight a given benchmark strategy) is the leading criterion in economics and computer science to evaluate performance in complex dynamic environments. The earliest regret-minimizing procedures can be traced back to Blackwell (1956a,b) and Hannan (1957). Since then, a large literature on the topic has developed at the intersection of economics and computer science. Once adapted to the play of games with incomplete information—with each player computing and minimizing his own regrets signal-by-signal—many well-known procedures or rules of behavior for the play of game $(G, S)^{\infty}$ generate a sequence $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ that has the ε-AR property. The class of regret-minimizing procedures—hereafter referred to as ε-AR algorithms—is very large, and its review goes beyond the scope of this paper (we refer to, e.g., Foster and Vohra, 1999; Cesa-Bianchi and Lugosi, 2006; Nisam et al., 2007; Roughgarden, 2016, for surveys of this literature).[10]

In general, ε-AR algorithms have desirable behavioral and computational properties. Importantly—and consistently with our Assumption 1—there are (simple) ε-AR algorithms that work in environments of arbitrary complexity and with almost no knowledge of the environment. More specifically, there are ε-AR algorithms that only require to know a player's set of actions and receive (an estimate of) the player's realized payoff as feedback at the end of each period, but not

dynamics have good welfare properties in our setting as well. The price of anarchy measures how the efficiency of a system degrades due to selfish behavior of its agents in the worst case.

[10]A (largely) non-exhaustive list of ε-AR algorithms includes: repeated Bayes Nash equilibrium play and repeated Bayes Correlated equilibrium play of (G^L, S^L) (see Section 3.2); simple adaptive heuristics, such as regret matching and its generalizations (Hart and Mas-Colell, 2000, 2001; Hart, 2005); fictitious-play-like dynamics (Fudenberg and Levine, 1995, 1999); calibrated learning (Foster and Vohra, 1997, 1998); specialized learning algorithms, such as greedy algorithms, weighted majority, multiplicative weights, and polynomial weights algorithms (e.g., Littlestone and Warmuth, 1994; Freund and Schapire, 1997, 1999; Cesa-Bianchi, Mansour, and Stoltz, 2007; Arora, Hazan, and Kale, 2012), and follow the perturbed leader and follow the regularized leader algorithms (e.g., Kalai and Vempala, 2005; Shalev-Shwartz, 2011). There are also ε-AR algorithms that allow for bounded memory (Lehrer and Solan, 2009) and that are dynamically consistent (Schlag and Zapechelnyuk, 2017). Finally, Several regret-minimizing procedures can also be cast into a form that is similar to reinforcement learning rules (e.g., Bush and Mosteller, 1955; Roth and Erev, 1995; Börgers and Sarin, 1997, 2000; Erev and Roth, 1998; Young, 2004).

(necessarily) that players observe $((a^n, \theta^n))_{n \in \mathbb{N}}$. Such algorithms are *model- and belief-free*, as they do not require players to: (i) know their opponents' payoff functions or signals; (ii) have (prior) beliefs about the information environment; (iii) to build a model of (or develop beliefs about) their opponents' play, and to reply optimally to such play.

Assuming that the ε-AR property holds, however, does not exclude that players know, understand, or observe more than the bare minimum needed to minimize their regrets—they just need not to. In other words, the ε-AR property neither requires nor excludes sophisticated behavior, strong knowledge assumptions, and rich informational feedback to players.

Although the class of ε-AR algorithms is very large, we pursue an *algorithm-independent approach*. That is, we neither assume that players adopt a specific regret minimization procedure nor assume that they coordinate on the same regret minimization procedure.

Probability Space. Typically, ε-AR algorithms involve randomization (i.e., require a player, say i, to play an element of $\Delta(A_i)$ in some or all periods). Thus, there are three sources of randomness in the model: (i) the process of payoff states $(\theta^n)_{n \in \mathbb{N}}$; (ii) the sequence of signal distributions $(\pi(\cdot \mid \theta^n))_{n \in \mathbb{N}}$; (iii) the randomization induced by ε-AR algorithms. These three sources naturally induce a probability measure, which we denote by $\mathbb{P}$, on the set of all finite histories H. By the Kolmogorov extension theorem, this probability measure uniquely extends to H^∞. Hereafter, whenever we use the expression "almost sureley" we mean almost surely with respect to probability measure $\mathbb{P}$.

3.2 Convergence under the Asymptotic ε-Regret Property

The relevant space of uncertainty in the limiting stage game (G^L, S^L) is $A \times T^L \times \Theta^L$. We write ν for a typical element of $\Delta(A \times T^L \times \Theta^L)$. The notion of Bayes correlated ε-equilibrium is defined through the restrictions we impose on ν.

Definition 3 (Bayes Correlated ε-Equilibrium)**.** *Fix an ε. The probability distribution $\nu \in \Delta(A \times T^L \times \Theta^L)$ is a Bayes Correlated ε-Equilibrium (hereafter, ε-BCE) of (G^L, S^L) if the two following properties hold.*

1. ν is consistent for (G^L, S^L); that is, for all $t \in T^L$ and $\theta \in \Theta^L$, we have

$$\sum_a \nu(a, t, \theta) = \pi^L(t \mid \theta) \psi^L(\theta).$$

2. ν *is* ε-*obedient for* (G^L, S^L); *that is, for all* $i \in \mathcal{I}$, $t_i \in T_i^L$, *and* $a_i \in A_i$, *we have*

$$\sum_{a_{-i}, t_{-i}, \theta} \big[u_i\big((a_i', a_{-i}), \theta\big) - u_i\big((a_i, a_{-i}), \theta\big)\big] \nu\big((a_i, a_{-i}), (t_i, t_{-i}), \theta\big) \leq \varepsilon_i(a_i, a_i', t_i)$$

for all $a_i' \in A_i$.

We denote by $E(\varepsilon)$ the set of ε-BCE of (G^L, S^L).

Consistency is a feasibility constraint which requires the marginal of ν on the exogenous variables T^L and Θ^L to be consistent with the elements of game (G^L, S^L). Obedience is an incentive constraint. A probability distribution ν is *obedient* if any 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, play their part of the realized action profile.

For any ε, the set $E(\varepsilon)$ is convex. When $\varepsilon = 0$, we have Bayes correlated equilibrium (hereafter, BCE). The BCE notion is due to Bergemann and Morris (2016) and is an incomplete information version of correlated equilibrium (Aumann, 1974, 1987). When Θ^L is a singleton, the notion of ε-BCE reduces to that of correlated ε-equilibrium for a complete information game.

Definition 4 (Empirical Distribution). *Let* $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ *be a sequence of states, signals, and actions from* $(G, S)^\infty$. *For all* $N \in \mathbb{N}$, *the empirical distribution* $Z^N \in \Delta(A \times T \times \Theta)$ *is defined pointwise by*

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

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. Note that q^K defined in (2.2) is the marginal on A of the empirical distribution Z^K for all $K \in \mathbb{N}$. That is, $q^K(a) = \sum_{t, \theta} Z^K(a, t, \theta)$ for all $a \in A$.

The next theorem establishes the following results: the sequence of empirical distributions converges almost surely to $E(\varepsilon)$ if and only if the sequence of states, signals, and actions from $(G, S)^\infty$ has the ε-AR property almost surely.

Theorem 1 (Convergence of ε-Regret Dynamics). *Fix an* ε. *The sequence of states, signals, and actions* $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ *from* $(G, S)^\infty$ *has the* ε-*AR property 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* $E(\varepsilon)$.

When the payoff state is perfectly persistent, if $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has ε-AR almost surely, then $(Z^N)_{N \in \mathbb{N}}$ converges almost surely to the set of correlated ε-equilibria of the complete information game (G^L, S^L).

Although only instrumental in this paper, Theorem 1 may be of independent interest as it provides dynamic foundations for the static equilibrium notion of ε-BCE (Bergemann and Morris, 2016). The BCE notion is central for the literature on information design (e.g., Aumann, Maschler, and Stearns, 1995; Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019; Kamenica, 2019) and for that on robust prediction in games with incomplete information (e.g., Bergemann and Morris, 2013, 2016). Theorem 1 generalizes to games with incomplete information earlier work on the dynamic foundations for correlated equilibrium in games with complete information (e.g., Foster and Vohra, 1997; Hart and Mas-Colell, 2000, 2001; Fudenberg and Levine, 1995, 1999).

Recent contributions in computer science (Hartline, Syrgkanis, and Tardos, 2015; Caragiannis, Kaklamanis, Kanellopoulos, Kyropoulou, Lucier, Paes Leme, and Tardos, 2015) also study convergence properties of regret-minimizing dynamics in incomplete information environments. There are motivational, technical, and conceptual differences between their setting and ours. First, whereas they focus on price of anarchy and efficiency results, our emphasis is on connecting ε-regret dynamics to the inference problem of an outside observer (see the next sections). Second, Hartline et al. (2015) study a private values setting in which private information is independent across players and time, and Caragiannis et al. (2015) restrict attention to generalized second price auctions and allow for valuations that are correlated across bidders; in contrast, we work with general games and allow for private information that is correlated across players and over time as well as for common values. Third, they adopt notions of players' payoffs and regrets that are different than ours. Fourth, they study convergence to the coarse analog of the set of BCE as (some definitions) in Forges (1993); in contrast, we study convergence to the set of BCE as in Bergemann and Morris (2016) (or to its coarse analog, see Section 6.1). Because of these and other contrasts, our analyses are complementary.

We now define restrictions on the set of actions implied by ε-BCE.

Definition 5 (ε-BCE Prediction). *Fix an ε. A probability distribution $q \in \Delta(A)$ is an ε-BCE prediction if there exists $\nu \in E(\varepsilon)$ such that*

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

17

for all $a \in A$. We denote by $Q(\varepsilon)$ the set of ε-BCE predictions of (G^L, S^L).

The following is an immediate corollary of Theorem 1.

Corollary 1 (Convergence of the Empirical Distribution of Actions). *Fix an ε. If the sequence of states, signals, and actions $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has ε-AR almost surely, then the sequence of empirical distributions of actions $(q^N)_{N \in \mathbb{N}}$ converges almost surely to $Q(\varepsilon)$.*

Remark 1. We clarify the convergence notion in Theorem 1 and Corollary 1. We do so for the convergence of $(q^N)_{N \in \mathbb{N}}$ to $Q(\varepsilon)$; similar observations apply to the convergence of $(Z^N)_{N \in \mathbb{N}}$ to $E(\varepsilon)$. First, the almost sure convergence of $(q^N)_{N \in \mathbb{N}}$ is to $Q(\varepsilon)$, not necessarily to a point in that set. Second, it is the empirical distribution of actions that becomes essentially an ε-BCE prediction, not necessarily the actual play. Third, define the distance between q^N and $Q(\varepsilon)$ as

$$d\big(q^N, Q(\varepsilon)\big) := \inf_{q \in Q(\varepsilon)} \big\| q^N - q \big\|,$$

where $\| \cdot \|$ is the Euclidean norm; then, the almost sure convergence of $(q^N)_{N \in \mathbb{N}}$ to $Q(\varepsilon)$ means that

$$\mathbb{P}\left(\lim_{N \to \infty} d\big(q^N, Q(\varepsilon)\big) = 0 \right) = 1.$$

That is, the following statement holds $\mathbb{P}$-almost surely: for any $\delta > 0$, there exists a positive integer N' such that, for each $N > N'$, there exists $\tilde{q}^N \in Q(\varepsilon)$ with $\|q^N - \tilde{q}^N\| < \delta$. In other words, the sequence of empirical distributions of actions $(q^N)_{N \in \mathbb{N}}$ eventually enters any neighborhood of $Q(\varepsilon)$ and stays there forever. Equivalently, the following statement holds $\mathbb{P}$-almost surely: for any $\varepsilon' > \varepsilon$, there exists a positive integer N' such that $q^N \in Q(\varepsilon')$ for all $N > N'$. In other words, $\mathbb{P}$-almost surely, there is a finite time N' after which the empirical distribution of actions is always an ε'-BCE prediction of (G^L, S^L).

3.3 Illustration: A Two-Seller Pricing Game

Consider again the two-seller pricing game introduced in Section 2.3. Hereafter, for ease of exposition we assume that marginal costs are i.i.d. across sellers and over time, so that $(G, S) = (G^L, S^L)$.

Regrets and the ε-Regret Property. Since each seller has two actions, the ε-AR property can be stated as follows: as the number of periods grows, the average

actual payoff of each seller becomes ε-close to that of having set the best fixed price in hindsight for each of the seller's marginal cost. Formally, for all i and t_i, the average factual profit that seller i with marginal cost t_i has obtained up to time N is

$$U_i(t_i, N) = \frac{1}{N} \sum_{n=1}^{N} g_i(p_i^n, p_{-i}^n)(p_i^n - t_i^n) \mathbb{1}_{\{t_i\}}(t_i^n).$$

Moreover, for each price p_i, the average counterfactual profit seller i with marginal cost t_i would have obtained had he set price p_i in all periods up to time N is

$$V_i(p_i, t_i, N) = \frac{1}{N} \sum_{n=1}^{N} g_i(p_i, p_{-i}^n)(p_i - t_i^n) \mathbb{1}_{\{t_i\}}(t_i^n).$$

The regret of seller i with marginal cost t_i for price p_i before setting a price at time $N + 1$ is

$$R_i(p_i, t_i, N) = \max \left\{ V_i(p_i, t_i, N) - U_i(t_i, N), 0 \right\}.$$

$R_i(k, t_i, N)$ is a measure of the time average regret experienced by seller i with marginal cost t_i at period N for not having set price p_i in all past periods up to N.

Let $\boldsymbol{\varepsilon}$ be a vector that specifies, for all i, p_i, and t_i, a non-negative real number $\varepsilon_i(p_i, t_i)$. A sequence of marginal costs and prices $((t^n, p^n))_{n \in \mathbb{N}}$ has the $\boldsymbol{\varepsilon}$-AR property if

$$\limsup_{N \to \infty} R_i(p_i, t_i, N) \leq \varepsilon_i(p_i, t_i)$$

for all i, t_i, and p_i.

For illustrative purposes, suppose sellers play the pricing game using *regret matching*. This is a particularly simple regret-minimizing procedure (due to Hart and Mas-Colell, 2000) for the repeated play of a complete information stage game. We adapt the procedure to repeated play of the incomplete information pricing game. Since each seller has two actions, regret matching prescribes setting each price p_i in period $N + 1$ with a probability that is proportional to the vector of regrets. Formally, let $\gamma_i^{N+1}(p_i, t_i, N)$ denote the probability of setting price p_i in period $N+1$ by seller i with marginal cost t_i; then, regret matching prescribes that

$$\gamma_i^{N+1}(p_i, t_i, N) = \frac{R_i(p_i, t_i, N)}{\sum_{p_i'} R_i(p_i', t_i, N)}$$

for all p_i. Play is arbitrary in the first period and when all regrets are zero. We use regret matching to generate the simulated paths of play that we present next.

Convergence under the ε-Regret Property. Let $p := (p_1, p_2)$ denote a typical price profile and $t := (t_1, t_2)$ a typical marginal cost profile. The empirical distribution of prices and marginal costs at time N, denoted by Z^N, is defined pointwise as

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

for all (p,t). In the context of our pricing game, since each seller has two actions, Theorem 1 can be stated as follows. A sequence of marginal costs and prices $((t^n, p^n))_{n \in \mathbb{N}}$ has the ε-AR property almost surely if and only if, as $N \to \infty$, the sequence of empirical distributions $(Z^N)_{N \in \mathbb{N}}$ satisfies the following properties:

1. Consistency for (G, S): for all t, we have

$$\lim_{N \to \infty} \sum_p Z^N(p,t) = \psi(t)$$

   almost surely.

2. ε-obedience for (G, S): for all i and t_i, we have

$$\limsup_{N \to \infty} \sum_{p,t_{-i}} \left[g_i(p_i', p_{-i})(p_i' - t_i) - g_i(p)(p_i - t_i) \right] Z^N(p, (t_i, t_{-i})) \leq \varepsilon_i(p_i', t_i)$$

   for all p_i' almost surely.

Hereafter, we consider the following parametric specification of the game. The demand for seller i's good is

$$g_i(p_i^n, p_{-i}^n) := M \frac{\exp(\alpha p_i^n)}{1 + \exp(\alpha p_i^n) + \exp(\alpha p_{-i}^n)}. \tag{2}$$

Here, $M > 0$ is the market size and $\exp(\alpha p_i^n)/(1 + \exp(\alpha p_i^n) + \exp(\alpha p_{-i}^n))$ is the probability that a consumer buys from seller i when prices are (p_i^n, p_{-i}^n). The parameter $\alpha < 0$ captures consumers' price sensitivity, so that i's probability of selling is decreasing in p_i^n and increasing p_{-i}^n. Marginal costs are distributed according to a discretized and truncated Normal distribution with mean $\mu > 0$, variance $\sigma > 0$, and finite support $\widehat{T}$. In addition, we consider the following numerical specification of the game. Sellers set prices in $\{\underline{p}, \overline{p}\} = \{3, 10\}$. Demand parameters are $M = 1$ and $\alpha = -1/3$. The distribution of marginal costs has mean $\mu = 3$, variance $\sigma = 1$, and is truncated at $\underline{t} = 0$ and $\overline{t} = 6$.

To illustrate the restrictions on actions implied by ε-BCE, we parametrize the vector $\boldsymbol{\varepsilon}$ as a function of a non-negative real number ε. In particular, for all i,

p_i, p_i' with $p_i \neq p_i'$, and t_i, we set $\varepsilon_i(p_i', t_i)$ equal to a fraction ε of an upper bound on seller i's maximum payoff difference from a deviation from p_i to p_i' when his signal is t_i. This procedure is described in detail in Appendix B and results in an intuitive scaling of $\boldsymbol{\varepsilon}$. With a slight abuse of notation, we write the resulting set of $\boldsymbol{\varepsilon}$-BCE predictions as $Q(\varepsilon)$ and plot it in Figure 1 for three values of ε. As expected, the set grows as ε grows, but remains quite small even for values of ε as high as 0.1.

Figure 1: Sets of $\boldsymbol{\varepsilon}$-BCE Predictions.

(a) $\varepsilon = 0.02$.

(b) $\varepsilon = 0.05$.

(c) $\varepsilon = 0.1$.
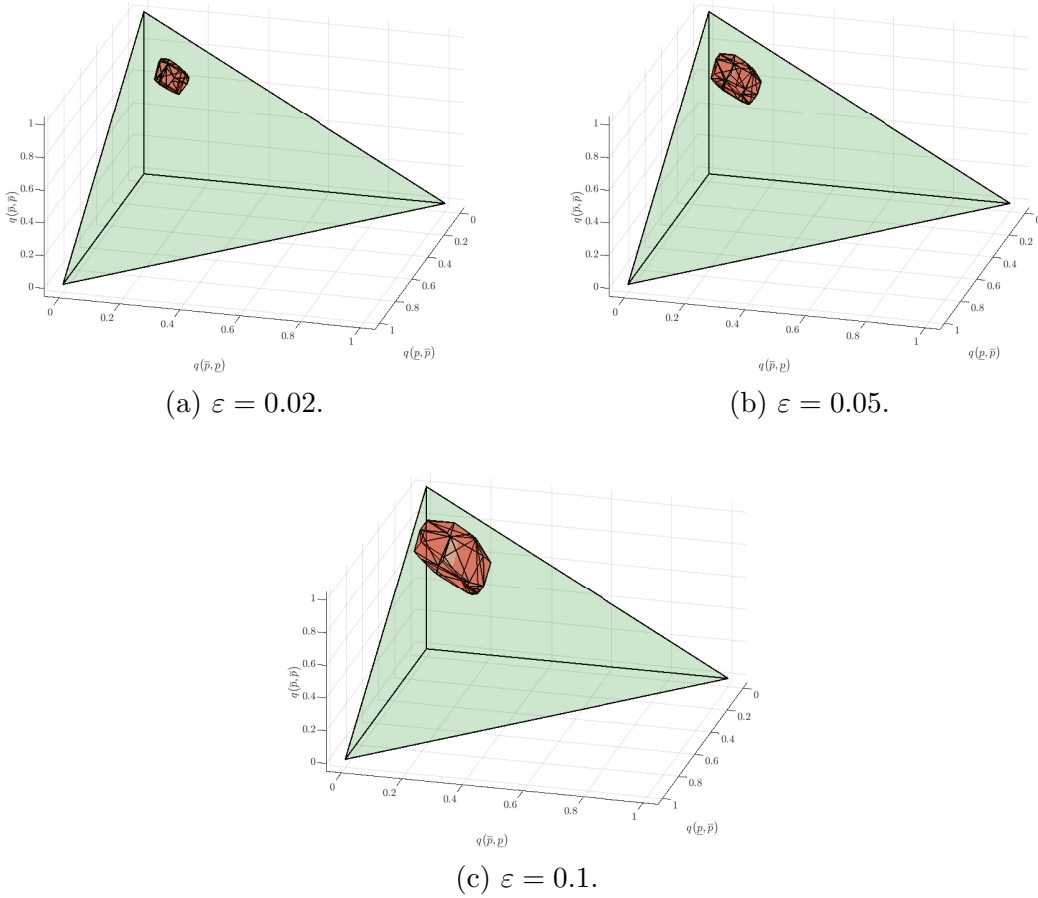
The light-red convex sets correspond to $Q(\varepsilon)$ for $\varepsilon = 0.02$, $\varepsilon = 0.05$, and $\varepsilon = 0.1$.

With the parametric and numerical specification above, we generate a sequence of marginal costs and prices $((t^n, p^n))_{n=1}^{K}$ by letting sellers set prices using regret matching. Figure 2 represents the empirical distribution of prices corresponding to $K = 400$ iterations of regret matching. Each blue dot is a snapshot of the empirical distribution of prices at a point along the path. The light-red convex set represent set $Q(\varepsilon)$ for $\varepsilon = 0.05$. The convergence of the empirical distribution

of prices to $Q(\varepsilon)$ provides an illustration of Corollary 1.

Figure 2: Convergence.
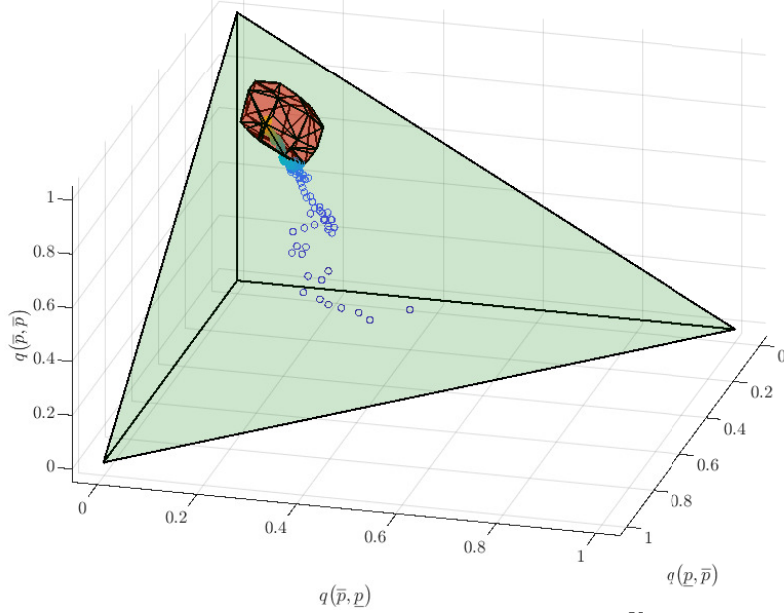
We represent a path of the empirical distribution of prices $\left(q^N\right)_{N=1}^K$ generated by $K = 400$ iterations of regret matching. Blue dots correspond to the empirical distribution of prices at a point along the path; the light-red convex set corresponds to $Q(\varepsilon)$ for $\varepsilon = 0.05$.

4   Econometrics

In this section, we build on the theoretical results we established in Section 3 to develop an empirical strategy that addresses the econometric problem presented in Section 2.

4.1   Identification and Estimation

4.1.1   Empirical Model

We start by laying out the empirical model. In particular, the next assumption summarizes the data-generating process and the observables.

Assumption 2. *The empirical model and the observables are summarized by 1–4 below.*

1. *A game $(G, S)^\infty$ is played over time under Assumption 1.*

2. *The corresponding limiting stage game belongs to a parametric class of games $\left\{\left(G^L(\lambda^G), S^L(\lambda^S)\right)\right\}_{\lambda \in \Lambda}$ indexed by structural parameters $\lambda := (\lambda^G, \lambda^S) \in \Lambda$,*

where Λ is a non-empty set. The true structural parameters in the data generating process are $\lambda_0 := (\lambda_0^G, \lambda_0^S)$.

3. *The researcher has access to either of the following datasets:*

   - *Dataset $\mathrm{D}1$: for some positive integer K, the researcher observes a realized sequence of actions $(a^n)_{n=1}^K \in A^K$ from the dynamic game $(G, S)^\infty$.*
   - *Dataset $\mathrm{D}2$: for some positive integer K, the researcher observes an empirical distribution of actions $q^K \in \Delta(A)$ from the dynamic game $(G, S)^\infty$ (possibly without information about the timing of actions).*

4. *The sequence of states, signals, and actions $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the $\boldsymbol{\varepsilon}$-AR property almost surely for some $\boldsymbol{\varepsilon}$.*

Recall that dataset D2 can be constructed from dataset D1, but not viceversa, as noted in Section 2.2.

In the rest of this section, we investigate how we can recover λ_0 under Assumption 2. Before, however, we note the following. As specified by Assumption 2, our empirical model is incomplete (in the sense of Tamer, 2003). This is so because although we assume that players learn to interact in/adapt to the environment sufficiently well for the ε-AR property, we leave unspecified other elements of the model. In particular, we do not exclude that players know, understand, or observe more than the bare minimum needed for the ε-AR property to hold. Moreover, we assume neither that players adopt a specific regret minimization procedure nor that they coordinate on the same one (i.e., we adopt an algorithm-independent approach).

4.1.2 Bayes Correlated ε-Equilibrium and Restrictions on Parameters

For any $\varepsilon' \geq \varepsilon$, we denote by $Q(\lambda; \varepsilon')$ the set of ε'-BCE predictions of game $(G^L(\lambda^G), S^L(\lambda^S))$ with structural parameters $\lambda := (\lambda^G, \lambda^S) \in \Lambda$. Under Assumption 2, in our empirical model, data are not sampled from a limiting "population" distribution of the observable actions, i.e., a fixed limiting $q_0 \in \Delta(A)$. Instead, Corollary 1 only ensures that the sequence of empirical distributions of actions $(q^N)_{N \in \mathbb{N}}$ converges almost surely to $Q(\lambda_0; \varepsilon)$, not necessarily to a point in that set. Hence, the data are not generated from the repetition of identical experiments. As a consequence, the standard notion of identified set is not meaningful in our context.

Despite these features, and although we do not maintain that the data are generated by equilibrium play, for any $\varepsilon' > \varepsilon$, the set of parameters compatible with a

given distribution of actions $q \in \Delta(A)$ as a ε'-BCE prediction of $(G^L(\lambda^G), S^L(\lambda^S))$, given by $\{\lambda \in \Lambda : q \in Q(\lambda; \varepsilon')\}$ is a key ingredient to recover valid bounds on the structural parameters in our setting, as we next show. We further discuss the special features of our environment in Section 4.2.1, after we develop our empirical strategy below.

4.1.3 Recovering Bounds on Parameters

We start by proposing a simple estimation strategy that uses the observed empirical distribution of actions q^K to recover the structural parameters λ. This can be done using the *"plug-in"* estimator.

Definition 6 (Plug-in Estimator). *Fix an ε' and let q^K be the observed empirical distribution of actions at time K. The plug-in estimator is denoted by $\hat{\Lambda}^K(\varepsilon')$ and defined by*

$$\hat{\Lambda}^K(\varepsilon') := \left\{\lambda \in \Lambda : q^K \in Q(\lambda; \varepsilon')\right\}. \tag{3}$$

The plug-in estimator $\hat{\Lambda}^K(\varepsilon')$ can be computed from the data under both dataset D1 and dataset D2.

If dataset D1 is available, we can leverage the dynamic nature of the data-generating process for the estimation of the primitives. Intuitively, we construct estimators $\hat{\Lambda}^K(\varepsilon')$ for successive values of K, where $K > N$ for some positive integer N, and then consider their intersection. At the limit, when iterated infinitely many times, this procedure delivers the *intersection of plug-in estimators*.

Definition 7 (Intersection of Plug-in Estimators). *Fix an ε', let N be a positive integer, and let $\left(q^K\right)_{K>N}$ be a sequence of observed empirical distributions of actions. The intersection of plug-in estimators is denoted by $\inf \hat{\Lambda}^N(\varepsilon')$ and defined by*

$$\inf \hat{\Lambda}^N(\varepsilon') := \bigcap_{K>N} \hat{\Lambda}^K(\varepsilon'). \tag{4}$$

In the next theorem, we characterize properties of these estimation strategies.

Theorem 2 (Properties of $\hat{\Lambda}^K(\varepsilon')$ and $\inf \hat{\Lambda}^N(\varepsilon')$). *Under Assumption 2, the following holds almost surely. For any $\varepsilon' > \varepsilon$, there exists a positive integer N' such that:*

1. *$\lambda_0 \in \hat{\Lambda}^K(\varepsilon')$ for all $K > N'$;*

2. *$\lambda_0 \in \inf \hat{\Lambda}^N(\varepsilon')$ for all $N > N'$.*

The first part of Theorem 2 establishes that, for any $\varepsilon' > \varepsilon$, the static equilibrium notion of ε'-BCE provides a valid restriction for the estimation of interactions that satisfy the optimality condition captured by the ε-AR property. More specifically, the theorem establishes that, under Assumption 2, the restrictions implied by the ε'-BCE notion lead to estimating a set which almost surely contains the true parameter in the data-generating process. 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-AR.

The second part of the theorem pushes this further. If the sequence of actions is observable (dataset D1), for any $\varepsilon' > \varepsilon$, the true parameter λ_0 is contained almost surely in the set $\inf \hat{\Lambda}^N(\varepsilon')$ for some sufficiently large N. This result is the foundation for a feasible procedure that constructs estimated sets of parameters as the intersection of plug-in estimators $\hat{\Lambda}^{K'}(\varepsilon')$ for several values of K. This procedure leverages the dynamic nature of the data-generating process to produce bounds that are tighter than we would obtain by only considering $\hat{\Lambda}^K(\varepsilon')$ for a fixed K as

$$\inf \hat{\Lambda}^N(\varepsilon') \subseteq \hat{\Lambda}^K(\varepsilon') \quad \text{for all } K > N.$$

Although the intersection bounds obtained by using the intersection of plug-in estimators are by construction smaller than those obtained with a single plug-in estimator at any point in time, the extra informativeness of this procedure is not always without drawbacks. To see this, let K be the highest time index in dataset D1. Moreover, suppose that, for some $N' < K$, $q^N \notin Q(\lambda_0; \varepsilon')$ for some $N \leq N'$, whereas $q^N \in Q(\lambda_0; \varepsilon')$ for $N >'$ (i.e., convergence to the set $Q(\lambda_0; \varepsilon')$ has occurred by time $N' + 1$, but not before). Then, if $N < N'$, it could very well be the case that $\lambda_0 \notin \inf \hat{\Lambda}^N(\varepsilon')$, whereas $\lambda_0 \in \hat{\Lambda}^K$. That is, if the intersection on the right-hand side of (3) is taken over some N such that $q^N \notin Q(\lambda_0; \varepsilon')$, it could well be the case that the intersection of plug-in estimators does not contain the true parameter λ_0, whereas the plug-in estimator constructed using the time average of play at time K (the latest period in the dataset) does. This issue is related to how large of an ε' we want to use, and to the rate of convergence and the sample size. We return to these ideas in Sections 4.2.2 and 4.2.3.

4.1.4 Theoretical Bounds

Although the width of the estimated sets $\hat{\Lambda}^K(\varepsilon')$ and $\inf \hat{\Lambda}^N(\varepsilon')$ depends on details of the model and of the unknown data-generating process, we may construct the-

oretical bounds on $\hat{\Lambda}^K(\varepsilon')$ and $\inf \hat{\Lambda}^N(\varepsilon')$ based on the properties in Assumption 2. To this aim, we first define two sets of parameters that are compatible with ε'-BCE predictions.

The *outer recoverable set* consists of all parameters compatible with at least one ε'-BCE prediction for the game characterized by λ_0.

Definition 8 (Outer Recoverable Set). *Fix an ε'. The outer set of recoverable parameters is denoted by $\Lambda^{OR}(\lambda_0; \varepsilon')$ and defined by*

$$\Lambda^{OR}(\lambda_0; \varepsilon') := \bigcup_{q \in Q(\lambda^0; \varepsilon')} \{\lambda \in \Lambda : q \in Q(\lambda; \varepsilon')\}.$$

The *inner recoverable set* consists of all parameters compatible with all ε'-BCE prediction for the game characterized by λ_0.

Definition 9 (Inner Recoverable Set). *Fix an ε'. The inner set of recoverable parameters is denoted by $\Lambda^{IR}(\lambda_0; \varepsilon')$ and defined by*

$$\Lambda^{IR}(\lambda_0; \varepsilon') := \bigcap_{q \in Q(\lambda^0; \varepsilon')} \{\lambda \in \Lambda : q \in Q(\lambda; \varepsilon')\}.$$

The inner and outer recoverable sets, although not estimable from the data as they depend on the true parameter λ_0, represent theoretical bounds on the informativeness of our estimation strategy. This is established by the next theorem.

Theorem 3 (Theoretical Bounds on $\hat{\Lambda}^K(\varepsilon')$ and $\inf \hat{\Lambda}^N(\varepsilon')$). *Under Assumption 2, the following holds almost surely. For any $\varepsilon' > \varepsilon$, there exists a positive integer N' such that:*

1. *$\hat{\Lambda}^K(\varepsilon') \subseteq \Lambda^{OR}(\lambda_0; \varepsilon')$ for all $K > N'$;*

2. *$\Lambda^{IR}(\lambda_0; \varepsilon') \subseteq \inf \hat{\Lambda}^N(\varepsilon')$ for all $N > N'$.*

Suppose that dataset D1 is available and that we develop an estimation strategy based on the intersection of sets $\hat{\Lambda}^N(\varepsilon')$ for $K > N$. In the best-case scenario, the data-generating process will be such that $(q^K)_{K>N}$ will trace out the full set of ε'-BCE predictions. In this case, the bounds on λ_0 implied by our estimation strategy may come close to the theoretical bounds of the inner recoverable set. Conversely, if $(q^K)_{K>N}$ does not move around much in the set of ε'-BCE predictions, or only dataset D2 is available, the bounds on λ_0 implied by our estimation strategy will still lie within the outer recoverable set of parameters; that is, the outer recoverable set is a worse-case scenario bound for our estimation strategy.

4.2 Discussion and Further Econometric Details

4.2.1 Discussion of the Empirical Strategy

The empirical environment described in Assumption 2 departs from both standard identification analysis and standard estimation setups. In identification analysis, the researcher observes the joint distribution of the observables—in our setting, that would correspond to a $q_0 \in \Delta(A)$. For estimation, the researcher typically observes a finite sample of N i.i.d. draws from q_0. The identified set of parameters is thus characterized as a function of q_0,[11] and a corresponding estimator instead is a function of the data sample. An important exception is Epstein, Kaido, and Seo (2016), who study inference in complete information static games. While they maintain that payoff types are i.i.d. draws, they note that when the equilibrium selection rule is left unspecified as in Ciliberto and Tamer (2009), the observables may not be i.i.d. draws as the equilibrium selection may vary across markets. Importantly, they develop inferential tools to estimate the identified set under these assumptions. Similar to Epstein et al. (2016), we do not observe i.i.d. draws from q_0, but rather a sequence of actions. Hence, the standard notion of an identified set is not relevant for our model. Due to this, we focus on the true parameter λ_0—similar to some results in Ciliberto and Tamer (2009).

Our estimation strategy reflects the special features of our environment. Typically, in the literature on partially identified models (see, e.g., Molinari, 2020), estimators are shown to have a weak consistency property with respect to the identified set, or confidence sets have certain coverage properties. Instead, we prove in Theorem 2 a strong coverage property of our estimator relative to λ_0. The strong (a.s.) nature of our results derives from the theoretical convergence property of the data.[12] Moreover, the special nature of sampling—where the dynamic data-generating process does not converge to a single distribution—can be leveraged to treat successive distributions of the observables q^N as different data points. This motivates our intersection results in Theorem 2.

---

[11]See, for instance, Beresteanu, Molchanov, and Molinari (2011) and Galichon and Henry (2011) on how to characterize the identified set in models of discrete games.

[12]Notice that we also maintain, similar to the literature on the estimation of games, that unobservable payoff shifters are i.i.d. draws. Although we could construct an inferential strategy based on this property, as in Epstein et al. (2016), we choose instead to build on the strong convergence properties deriving from theory.

4.2.2 Choosing the Value of ε'

The value of ε' captures how much players' behavior departs from perfect regret minimization. As discussed in Section 3.1, such departures can have two explanations: (i) players are not exact optimizers; (ii) players are sustaining some form of cooperation beyond that allowed by exact regret minimization (despite the environment's complexity). Therefore, the question arises of how a researcher should set ε' for estimation.

In our estimation strategy, we treat ε' as a tuning parameter. There is a clear trade-off when choosing ε'. On the one hand, the larger ε' is, the more robust we are with respect to the perfect regret minimization assumption. On the other hand, the larger ε' is, the less informative the estimated bounds on parameter are; this is so because for any $\varepsilon'' > \varepsilon'$,

$$\hat{\Lambda}^K(\varepsilon') \subseteq \hat{\Lambda}^K(\varepsilon'') \qquad \text{and} \qquad \inf \hat{\Lambda}^N(\varepsilon') \subseteq \inf \hat{\Lambda}^N(\varepsilon'')$$

for all K and N. The approach we follow is to profile how estimated bounds $\hat{\Lambda}^K(\varepsilon')$ and $\inf \hat{\Lambda}^N(\varepsilon')$ change as ε' varies. In practice, the choice of ε' can be guided by comparing the available sample size to a theoretically-informed guess of how many time periods periods players need to satisfy the ε'-AR property. We discuss this idea in the next section.

4.2.3 Rate of Convergence and Sample Size

Fix a value of ε'. It is natural to ask how many observations K a researcher needs to construct a plug-in estimator $\hat{\Lambda}^K(\varepsilon')$ that contains almost surely the true parameter λ_0. In our setting, this question is equivalent to the question of how long it takes for the sequence of empirical distributions of actions $(q^N)_{N \in \mathbb{N}}$ to converge almost surely to $Q(\lambda_0; \varepsilon')$. The answer to the latter question depends on: (i) how long it takes for the joint empirical distribution of signals and states to converge to the true distribution $\pi^L(\cdot \mid \cdot)\psi^L(\cdot)$; (ii) the specific ε-AR algorithm that players follow to play game $(G, S)^\infty$. Hence, we cannot provide sharp rate-of-convergence results without selecting a specific ε-AR algorithm (i.e., under an algorithm-independent approach).

A large literature in computer science and economics, however, has pointed out that many ε-AR algorithms have good convergence rates (i.e., they are fast, and more so the more accurate is players' knowledge of the environment or the informational feedback they receive). In addition, such literature characterizes the convergence rates for most ε-AR algorithms. In particular, for such algorithms,

it is possible to determine upper bounds on how far the empirical distribution of actions is from perfect regret minimization as a function of the number of periods N and other parameters of the environment. Even in an algorithm-independent approach, such theoretical results can be useful to guide the choice of ε' for estimation purposes as a function of the available sample size.

4.2.4 Joint Estimation of λ_0 and ε

An alternative approach is to construct a set estimator that recovers bounds for both λ_0 and ε_0, where ε_0 is the true value of ε in the data-generating process. Let D denote the dimension of vector ε and let q^K be the observed empirical distribution of actions at time K. The plug-in estimator of $(\lambda_0, \varepsilon_0)$, which we denote by $\hat{\Lambda}^K$, consists of all (λ, ε) compatible with the distribution of actions q^K as a ε-BCE prediction of $(G^L(\lambda^G), S^L(\lambda^S))$. That is,

$$\hat{\Lambda}^K := \left\{ (\lambda, \varepsilon) \in \Lambda \times \mathbb{R}_+^D : q^K \in Q(\lambda; \varepsilon) \right\}.$$

The projection on Λ of $\hat{\Lambda}^K$ at a fixed value of ε' returns the plug-in estimator $\hat{\Lambda}^K(\varepsilon')$.

4.3 Computation

The estimation strategy described in Section 4.1 relies on the set estimator $\hat{\Lambda}^N(\varepsilon')$. Although the definition of $\hat{\Lambda}^N(\varepsilon')$ in (3) is not immediately implementable, a few steps proposed in Magnolfi and Roncoroni (2021) allow to efficiently compute this set. In particular, denote by $b^\intercal$ the transpose of $b \in \mathbb{R}^{|A|}$ and by $B^{|A|} := \{ b \in \mathbb{R}^{|A|} : b^\intercal b \leq 1 \}$ the closed unit ball centered at $0_{|A|} \in \mathbb{R}^{|A|}$. The following result holds true.

Proposition 1. *Fix an ε' and let $g(\cdot; q^N, \varepsilon') \colon \Lambda \to \mathbb{R}$ be the function defined pointwise by*

$$g(\lambda; q^N, \varepsilon') := \max_{b \in B^{|A|}} \min_{q \in Q(\lambda; \varepsilon')} b^\intercal \left(q^N - q \right). \tag{5}$$

Then,

$$\hat{\Lambda}^N(\varepsilon') = \left\{ \lambda \in \Lambda : g\left(\lambda; q^N, \varepsilon' \right) = 0 \right\}. \tag{6}$$

According to Proposition 1, we can characterize the estimator $\hat{\Lambda}^N(\varepsilon')$ as the zero level set of the criterion function $g(\cdot; q^N, \varepsilon')$. Hence, computing the set estimator $\hat{\Lambda}^N(\varepsilon')$ amounts to evaluating the function $g(\cdot; q^N, \varepsilon')$ over an appropriately chosen finite subset (a grid) of Λ. The computation of $g(\cdot; q^N, \varepsilon')$ can be further

simplified by replacing the inner constrained minimization problem in (6) by its dual, which consists of a linear constrained maximization problem. This step makes it is possible to check whether a given value of λ belongs to the estimator $\hat{\Lambda}^N(\varepsilon')$ by solving a single linear constrained maximization problem. We refer to Appendix C for further computational details.

4.4   Illustration: A Two-Seller Pricing Game

We illustrate the estimation strategy in the context of the two-seller pricing game. We first generate data under the same parametric and numerical specification as in Section 3.3: sellers set prices in $\{3, 10\}$, demand parameters are $M = 1$ and $\alpha = -1/3$, and marginal costs are i.i.d. across players and over time according to a discretized Normal distribution with $\mu_0 = 3$, $\sigma_0 = 1$, and truncated at $\underline{t} = 0$ and $\bar{t} = 6$. Players set prices every period using regret matching, which we iterate for $K = 400$ periods to obtain the sequence of empirical distributions of prices $\left(q^N\right)_{N=1}^K$, $K = 400$.

We use the sequence of empirical distributions of prices $\left(q^N\right)_{N=1}^K$, $K = 400$, depicted in Figure 2 to estimate the structural parameter $\lambda^G = \sigma$, the variance parameter of the distribution of marginal costs.[13] Similar to Section 3.3, we parametrize ε' as a function of a non-negative real number ε' and denote—with slight abuse of notation—the criterion function as $g(\sigma; q^N, \varepsilon')$ and the corresponding estimator as $\hat{\Lambda}^N(\varepsilon')$. We compute this function for three separate values of N and $\varepsilon' = 0.05$; the resulting estimator $\hat{\Lambda}^N(0.05)$ is the set of values of σ for which $g(\sigma; q^N, 0.05) = 0$.

We report in Panel A of Table 1 the set estimator $\hat{\Lambda}^N(\varepsilon')$ for three values of ε' and three values of ε'. The simulations show that in this example $\hat{\Lambda}^N(\varepsilon')$ is already quite small for $N = 400$ and $\varepsilon' = 0.05$. Except for one case (discussed below), our set estimators include the true value $\sigma = 1$. Moreover, the intersection estimator has the potential of considerably shrinking the estimated set of parameters. When ε' and N are small, $\hat{\Lambda}^N(\varepsilon')$ and $\hat{\Lambda}^N(\varepsilon')$ do not contain the true value $\sigma = 1$. This illustrates the discussion in Sections 4.1 and 4.2 of the potential issues when considering the intersection of plug-in estimators for a small values of N and ε', the choice of ε', and its relation to the number of available observations.

We report in Panel B of Table 1 the inner and outer recoverable set. In this case, the inner recoverable set coincides with the true parameter. The outer

---

[13]Many other parametrizations of this game are possible. For instance, λ could describe sellers' payoff functions, seller-specific perfectly persistent marginal costs, or the information structure.

Figure 3: Estimation of σ.
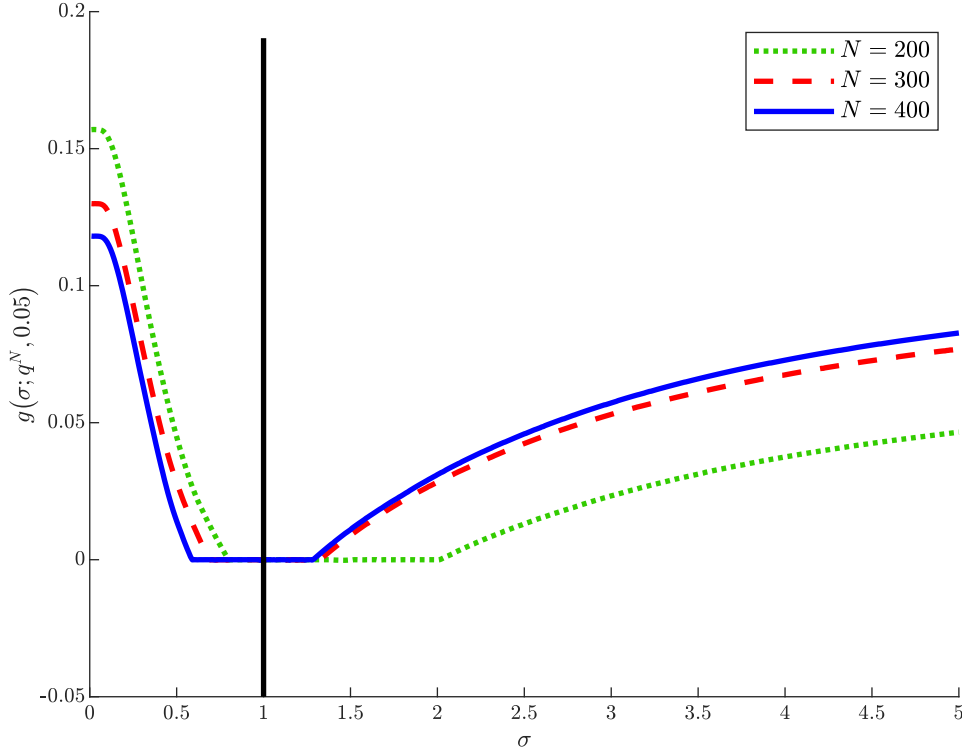
For $\varepsilon' = 0.05$, we represent the function $g(\sigma; q^N, \varepsilon')$ which characterizes the set estimator $\hat{\Lambda}^N(\varepsilon')$. Each of the three lines corresponds to a different value of N. The function is computed over a grid of 500 values. See Appendix C for computational details.

recoverable set gives worst-case bounds that are informative for $\varepsilon' = 0.02$ and $\varepsilon' = 0.05$. For $\varepsilon' = 0.1$, which correspond to a large departure from perfect regret minimization, the estimated sets remain informative, but the outer recoverable set indicates that worst-case bounds can be largely uninformative.

5   Application: Pricing in an Online Platform

We use our method to study sellers' pricing behavior on Swappa, an online marketplace for smartphones and other consumer electronics with around $100 million in sales in 2018. Swappa is an appealing empirical environment to apply our method to for at least two reasons. First, pricing on Swappa is a good example of a *complex dynamic environment*: sellers set prices over time and need to potentially keep track of a large and fast-evolving set of competitors and potential competitors. Thus, the payoffs of different pricing strategies are hard to predict. Second, this environment has broader relevance for many online decentralized

Table 1: Set Estimators and Recoverable Sets.

Panel A: Set Estimators	$\varepsilon' = 0.02$	$\varepsilon' = 0.05$	$\varepsilon' = 0.1$
$\hat{\Lambda}^{200}(\varepsilon')$	[1.03, 1.84]	[0.80, 2.01]	[0.51, 2.40]
$\hat{\Lambda}^{300}(\varepsilon')$	[0.89, 1.23]	[0.69, 1.32]	[0.44, 1.48]
$\hat{\Lambda}^{400}(\varepsilon')$	[0.76, 1.21]	[0.59, 1.28]	[0.37, 1.43]
inf $\hat{\Lambda}^{200}(\varepsilon')$	[1.03, 1.21]	[0.80, 1.28]	[0.51, 1.43]
Panel B: Recoverable Sets	$\varepsilon' = 0.02$	$\varepsilon' = 0.05$	$\varepsilon' = 0.1$
$\Lambda^{IR}(\sigma_0; \varepsilon')$	$\{1\}$	$\{1\}$	$\{1\}$
$\Lambda^{OR}(\sigma_0; \varepsilon')$	[0.50, 2.40]	[0.42, 2.71]	[0.30, 13.25]

Panel A reports set estimators $\hat{\Lambda}^N(\varepsilon')$ for three values of N, and inf $\hat{\Lambda}^N(\varepsilon')$ for $N = 200$; each set estimator is computed for three values of ε'. Panel B reports $\Lambda^{IR}(\sigma_0; \varepsilon')$ and $\Lambda^{OR}(\sigma_0; \varepsilon')$ computed over a grid of 500 values of σ for three values of ε'. See Appendix C for computational details.

platforms and marketplaces, where pricing is a decision of individual sellers. Examples include Amazon marketplace and eBay buy-it-now listings, which account for several hundreds of billions in annual sales. In contrast to other marketplaces, Swappa has desirable features: first, sellers do not compete with the platform (as in, e.g., Amazon and Walmart marketplaces); second, sellers do not have access to algorithmic pricing tools during our sample period.

In this environment, we use our method to recover the distribution of sellers' marginal costs. This primitive is an essential input to 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?

In the rest of the section, we first describe the empirical setting and the data; then we introduce the empirical model; finally, we discuss estimation results.

5.1 Empirical Setting and Data

Swappa is a user-to-user online marketplace where sellers list new or used electronic devices. Listings go through an approval process whereby the platform checks that the device is ready for activation. There is no listing fee, but buyers pay a flat fee to the platform upon purchase; the fee is included in the buyer's purchase price. Buyers shop on the website by first selecting the device they are interested in and then choosing among different listings for individual items. Figure 4 shows a screenshot from the website, which displays the set of available items at a point in time.

We collect data from Swappa by scraping the website for about three consecu-

Figure 4: A Screenshot from the Swappa Website.

The figure shows the web page that Swappa users see when selecting a device (in this case, Apple iPhone X).

tive months from November 2019 to February 2020. While many types of devices are on sale on Swappa, we only collect data for the ten most common models of iPhones, as these tend to be the thickest markets on the website. For each listing—corresponding to a phone on sale—we collect information on price and seller and product characteristics at hourly intervals until the item is sold. The hourly price observations are aggregated at the daily level by taking the average. Over the full sample period, we observe a total of 12,741 product listings owned by 2,436 sellers.

Sellers on the website greatly differ in the number of listings that they make over our sample period. While there is a long tail of sellers that only list one device in our sample, we focus on more experienced sellers. We do so because we want to model pricing behavior over time for sellers that make multiple pricing decisions. Hence, in the following analysis, we focus in particular on the largest sellers. The largest seller has 763 listings over the sample period; considering sellers that have at least 130 listings over the sample period yields a set of 15 sellers.[14] These sellers are typically small firms that acquire used cellphones, refurbish them, and resell them. A device's marginal cost for these firms mainly consists of the acquisition and refurbishment cost (including labor and parts).

5.2 Empirical Model

Sellers on the marketplace face a trade-off: lower prices increase the probability of selling a device but lead to lower margins. For each seller i we assume that, for each device j seller i has on sale, the seller makes a pricing decision every day d that the listing is active.[15] Hereafter, we refer to each device-day pair (j, d) as a different period n (i.e., $n = (j, d)$).

To construct the set of competitors $-i$ for every device-day (j, d) of seller i, we assume that the set of competing devices includes all the devices of the same model as that of device j that are available on day d on Swappa. Seller i's payoff in period n is defined as

$$u_i(p_i^n, p_{-i}^n, t_i) := g_i(p_i^n, p_{-i}^n)(p_i^n - t_i^n),$$

where p_i^n and p_{-i}^n denote the prices of seller i and its competitors in period n, the function $g_i(p_i^n, p_{-i}^n)$ maps profiles of prices (p_i^n, p_{-i}^n) in period n into probabilities of making a sale for seller i in period n, and t_i^n denotes seller i's marginal cost in period n.

---

[14]We further document heterogeneity across sellers in pricing behavior in Appendix D.2

[15]In Appendix D.3, we present evidence on the timing of pricing decisions and argue that daily timing fits well the environment we study.

Although we consider a few separate markets, in which sellers list relatively homogeneous items, several factors are likely to affect pricing behavior. For instance, the phones on sale may vary in storage capacity and condition (ranging from "mint" to "fair"). In principle, accounting for this observable heterogeneity in the model is straightforward. In practice, our data is limited, suggesting that we may want to learn from sellers' pricing behavior across different devices. Hence, to make the model estimable with our method, we adopt a few simplifying assumptions. Let $n = (j, d)$ be a device-day pair. We assume that marginal costs t_i^n can be decomposed into two additive components, or $t_i^n = \chi_{m(j)d} + \zeta_i^n$, where $\chi_{m(j)d}$ is an "average valuation" of a device j with characteristics $m(j)$ on day d, and ζ_i^n is a seller-specific cost shock in period n. We interpret $\chi_{m(j)d}$ as the market price of a device j with characteristics $m(j)$ on other platforms or online marketplaces on day d, and ζ_i^n as an idiosyncratic cost component, capturing the specific acquisition cost or valuation of the seller. This assumption is similar to the additive valuation assumption in Haile, Hong, and Shum (2003) and Wildenbeest (2011).

We further assume that $\chi_{m(j)d}$ can be estimated from price data, or $p_i^n = \chi_{m(j)d} + \rho_i^n$, where $\mathbb{E}[\rho_i^n \mid \chi_{m(j)d}] = 0$.[16] Hence, the pricing residual ρ_i^n captures the seller's pricing behavior for item j in period d, and we can rewrite payoffs as

$$\pi_i(\rho_i^n, \rho_{-i}^n, \zeta_i^n) := \tilde{g}_i(\rho_i^n, \rho_{-i}^n, \zeta_i^n)(\rho_i^n - \zeta_i^n),$$

where $\tilde{g}_i$ models probabilities of sale for device i. Moreover, we assume that, for all i, i', if $(\rho_i^n, \rho_{-i}^n) = (\rho_{i'}^n, \rho_{-i'}^n)$, then $\tilde{g}_i(\rho_i^n, \rho_{-i}^n) = \tilde{g}_{i'}(\rho_{i'}^n, \rho_{-i'}^n)$, as sale probabilities only depend on the tuples of pricing residuals (ρ_i^n, ρ_{-i}^n) and $(\rho_{i'}^n, \rho_{-i'}^n)$. Taken together, our assumptions on payoffs greatly simplify the pricing problem by suppressing the dependence on item characteristics. Although the assumptions are strong, they preserve the basic incentive structure and make the empirical environment tractable by yielding an empirical game that is close to the general model.

The pricing game that sellers play on Swappa is potentially large, including a fast-evolving set of many competing sellers. To reduce the dimensionality of the game, we assume that seller i's payoff—through the function $\tilde{g}_i$—only depends on the aggregate state $\overline{\rho}_{-i}^n$ in period n. The aggregate state $\overline{\rho}_{-i}^n$ captures prices of

---

[16]To operationalize this approach, we express $\chi_{m(j)d}$ as a linear index of observable covariates and estimate the linear model

$$p_i^n = x_{m(j)d}^\mathsf{T}\beta + \rho_i^n,$$

where β is a vector of parameters, and $x_{m(j)d}^\mathsf{T}$ is the transpose of vector $x_{m(j)d}$, which includes a "trusted seller" indicator, the seller's rating, warranty length, an indicator for additional warranty, and fixed effects for storage capacity level, condition, and date-device. Further details are in Appendix D.4.1.

sellers other than i in period n. This is in the spirit of the oblivious equilibrium approach of Weintraub, Benkard, and Van Roy (2008) and of aggregative- and large-games approaches (see, e.g., Jensen, 2018; Gradwohl and Kalai, 2021).

For each seller i, we estimate the distribution of marginal costs ζ_i^n by using the method described in Section 4. We denote by $\psi_i(\cdot; \lambda_{0i}) \in \Delta_{++}(T_i)$ the distribution of seller i's marginal costs (parametrized by λ_{0i}). Moreover, we denote by $\tilde{\psi}_i(\cdot) \in \Delta_{++}(\Theta_i)$ the distribution of prices of sellers other than i (that can be estimated from the data).

We assume that, for each seller i, the sequence of aggregate states, marginal costs, and prices $((\overline{\rho}_{-i}^n, \zeta_i^n, \rho_i^n))_{n \in \mathbb{N}}$ has the $\boldsymbol{\varepsilon}$-AR property almost surely. The empirical distribution of prices, marginal costs, and states at time N for seller i, denoted by Z_i^N, is defined pointwise as

$$Z_i^N(\rho_i, \zeta_i, \overline{\rho}_{-i}) := \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\{\rho_i\}}(\rho_i^n) \mathbb{1}_{\{\zeta_i\}}(\zeta_i^n) \mathbb{1}_{\{\overline{\rho}_{-i}\}}(\overline{\rho}_{-i}^n)$$

for all $(\rho_i, \zeta_i, \overline{\rho}_{-i}) \in (A_i \times T_i \times \Theta_i)$. By Theorem 1, since $((\overline{\rho}_{-i}^n, \zeta_i^n, \rho_i^n))_{n \in \mathbb{N}}$ has the ε-AR property almost surely as $N \to \infty$, the sequence of empirical distributions $(Z_i^N)_{N \in \mathbb{N}}$ satisfies the following properties:

1. Consistency:

$$\lim_{N \to \infty} \sum_{\rho_i} Z_i^N(\rho_i, \zeta_i, \overline{\rho}_{-i}) = \psi_i(\zeta_i; \lambda_{0i}) \psi_i(\overline{\rho}_{-i}),$$

almost surely.

2. $\boldsymbol{\varepsilon_i}$-obedience: for all ζ_i and ρ_i, we have

$$\limsup_{N \to \infty} \sum_{\rho_{-i}} \left[\left[\tilde{g}_i(\rho_i', \rho_{-i}, \zeta_i)(\rho_i' - \zeta_i) - \tilde{g}_i(\rho_i, \rho_{-i}, \zeta_i)(\rho_i - \zeta_i) \right] \right.$$
$$\left. \times Z_i^N(\rho_i, \zeta_i, \overline{\rho}_{-i}) Z_i^N(\rho_i, \zeta_i, \overline{\rho}_{-i}) \right] \leq \varepsilon_i(\rho_i, \rho_i', t_i)$$

for all ρ_i' almost surely.

Our estimation strategy is based on the restriction implied by consistency and $\boldsymbol{\varepsilon_i}$-obedience in the empirical pricing game we consider. To make the empirical model estimable, we perform some further simplifications. First, we adopt a coarse discretization of pricing residuals to reduce the computational burden of the method. In particular, we split the observations of ρ_i^n into positive and negative values and construct ρ_H and ρ_L as the median observations in each of the

two bins. We also discretize $\overline{\rho}_{-i}^n$ and assign a value $\overline{\rho}_H$ if the majority of seller i's competitors in period n are priced at ρ_H and a value $\overline{\rho}_L$ otherwise. With this discretization, the function $\tilde{g}_i$ takes four possible values, which we estimate directly from the data. In particular, for any pair $(\rho_i, \overline{\rho}_{-i})$, we compute $\tilde{g}_i(\rho_i, \overline{\rho}_{-i})$ as the average probability of selling a device for all observations with price $\rho_i = \rho$ when the aggregate state is $\overline{\rho}_{-i}$.

Second, similar to our illustrative example, we assume that each seller i's marginal costs ζ_i^n is i.i.d. according to a truncated Normal distribution with parameters μ_i and σ_i. This specification leaves the parameters to be seller dependent and allows for variability in marginal costs for a seller across periods (or, equivalently, across device-days). In the next section, we discuss the estimation of μ_i while fixing σ_i at different values; that is, $\lambda_{0i} := \mu_i(\sigma_i)$. The parameter σ_i captures the variability across periods of seller i's (residualized) marginal costs; thus, small (resp., large) values of σ_i corresponds to a relatively small (resp., large) variability across periods of seller i's marginal costs.

5.3   Estimation Results

For each seller i, we obtain an empirical distribution of prices q_i^N, and define a function $g_i\big(\mu_i; q_i^N, \varepsilon', \sigma_i\big)$, where the corresponding estimated sets $\hat{\Lambda}^N(\varepsilon', \sigma_i)$ contain the zeros of function g_i. The real number ε' parametrizes $\boldsymbol{\varepsilon'}$ as in Sections 3.3 and 4.4. We also express σ_i as the ratio of $p_H - p_L$, as to provide an intuitive scale.

We first fix $\varepsilon' = 0.05$ and profile the estimated sets for μ_i as a function of σ_i. Figure 5 represents the estimated set $\hat{\Lambda}^N(0.05, \sigma_i)$ for the four largest sellers in our sample (by the total number of listings) in different shades of blue, and for different values of σ_i on the vertical axis. The figure shows that similar values of average ζ_i rationalize the data for different sellers. The values of μ_i are in USD. Such values are negative because ζ_i is the deviation between sellers' marginal cost and average valuation. Thus, our results indicate that μ_i is in the order of \$20, suggesting that sellers are able to acquire and refurbish phones for around \$20 less than the average valuation. Higher values of σ_i enlarge the set $\hat{\Lambda}^N(0.05, \sigma_i)$, as a broader range of average parameters can rationalize the data for higher values of the variance parameter. However, for small variations in marginal cost within sellers, the estimated sets are small.

Figure 6 further explores heterogeneity across sellers, representing the estimated set $\hat{\Lambda}^N(0.05, 0.05)$ for the nine largest sellers in the sample. The figure indicates more heterogeneity than was apparent from just focusing on the four

Figure 5: Estimation Results I.
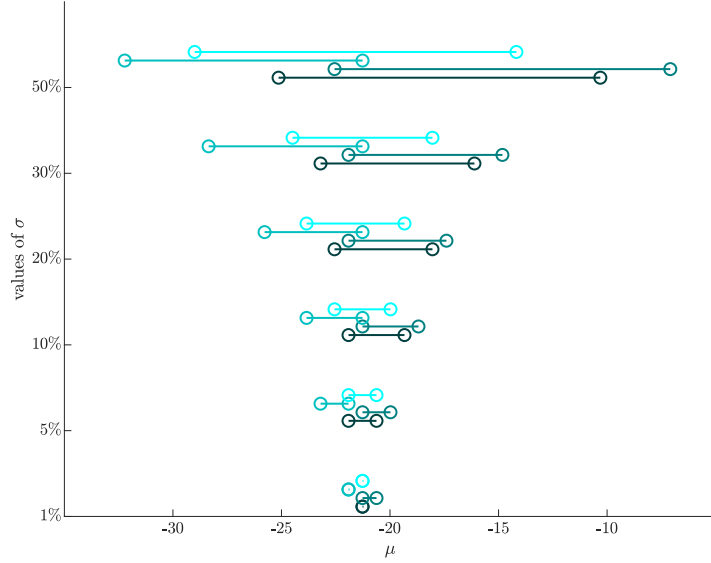
For $\varepsilon' = 0.05$, we represent the estimated set $\hat{\Lambda}^N(\varepsilon, \sigma_i)$ for the four largest sellers in the sample in different shades of blue, and for different values of σ_i on the vertical axis.

largest sellers. Figure 7 finally profiles the estimated set $\hat{\Lambda}^N(\varepsilon', 0.05)$ for different values of ε'. Unsurprisingly, larger values of ε' generate larger estimated sets. However, the estimated sets remain informative even for $\varepsilon' = 0.2$, which corresponds to a substantial departure from perfect regret minimization.

Figure 6: Estimation Results II.

For $\varepsilon' = 0.05$ and $\sigma_i = 0.05$, we represent the estimated set $\hat{\Lambda}^N(\varepsilon', \sigma_i)$ for the nine largest sellers in the sample.

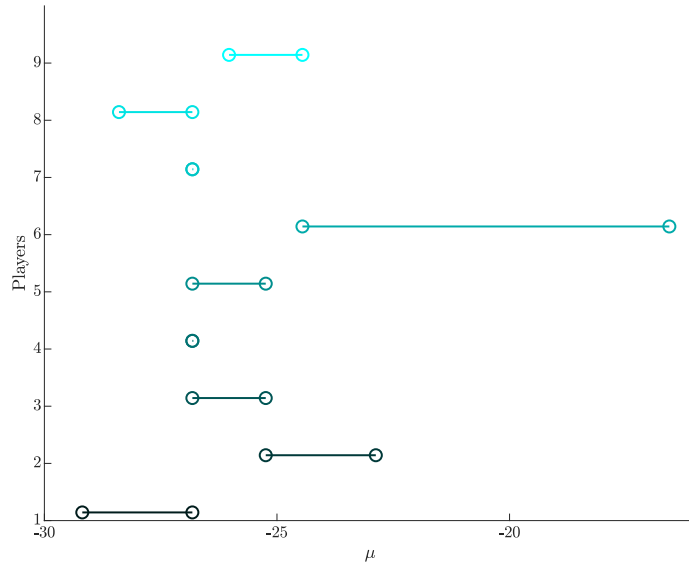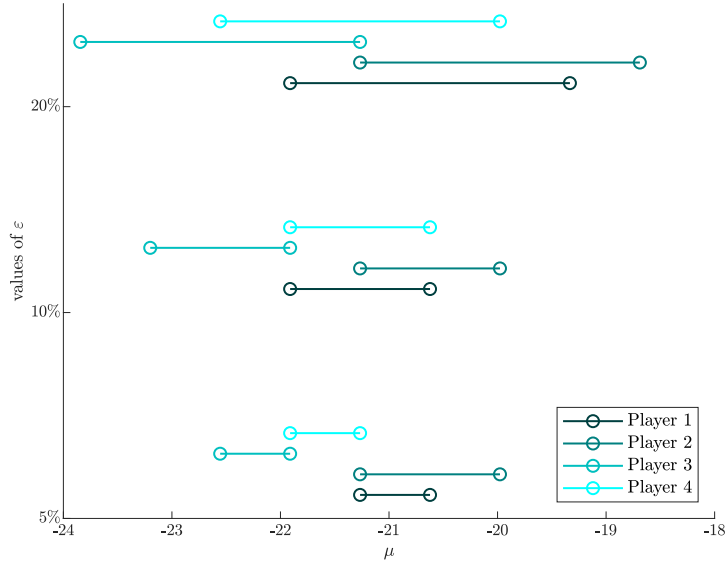The primitive that we recover enables two kinds of policy exercises. First, we

Figure 7: Estimation Results III.

For $\sigma_i = 0.05$, we represent the estimated set $\hat{\Lambda}^N(\varepsilon', \sigma_i)$ for the four largest sellers in the sample in different shades of blue, and for different values of ϵ on the vertical axis.

can use the distributions of sellers' marginal costs to evaluate how the current level of pricing compares with a Bertrand equilibrium and a perfect collusion benchmark. This measures the degree of current competitiveness of pricing in this online marketplace. Second, the distributions of sellers' marginal costs allow us to simulate pricing according to popular algorithms that are widely used in e-commerce applications. While some online marketplaces have made available these pricing tools to sellers, the effect that they have on prices is a topic of frontier research. This policy exercise could contribute meaningfully to this debate.

6 Additional Results and Discussion

6.1 Alternative Regret Notions

There are alternative notions of regret and asymptotic ε-regret property than those introduced by Definitions 1 and 2. A well-known alternative is that of *external regret*.[17]. In words, a sequence $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the asymptotic ε-external regret property if the time average of the counterfactual increase in past payoffs, had each player (i, t_i) played the best fixed action in hindsight, becomes ε-close to vanish in the long run. Thus, external regrets are a coarser measure of regret than internal regrets.

---

[17]As opposed to *internal regret*, which correspond to the notion in Definition 1. For additional regret notions see, e.g., Greenwald and Jafari (2003) and Lehrer (2003)

In Appendix E, we provide analogous results to those in Section 3 for the external regret notion. In particular, we show that a sequence of states, signals, and actions from $(G, S)^\infty$ has the asymptotic ε-external regret property almost surely if and only if the sequence of empirical distributions converges almost surely to the set of Bayes coarse correlated ε-equilibria of the limiting stage game. The notion of Bayes coarse correlated ε-equilibrium (hereafter, ε-BCCE) is new to this paper. It can be interpreted as an incomplete information version of coarse correlated equilibrium (Hannan, 1957; Moulin and Vial, 1978; Young, 2004) or as the coarse analog of the ε-BCE notion. Since the ε-BCCE is defined using coarse incentive constraints than those used to define the ε-BCE, the set of ε-BCCE of the limiting stage game is a superset of its set of ε-BCE.

With the appropriate changes, our estimation approach can be implemented under the assumption that the sequence of states, signals, and actions from $(G, S)^\infty$ has the asymptotic ε-external regret property. In that case, the ε'-BCCE notion would provide valid restrictions to estimate the structural parameters of interest.

We develop our econometric approach under the asymptotic ε-external regret property for two reasons. First, Blum and Mansour (2007) provide a "black box reduction" to convert any asymptotic ε-external regret algorithm into a ε-internal regret algorithm. Thus, whenever agents can satisfy the asymptotic ε-internal regret, they can also satisfy the asymptotic ε-external regret property. Second, for games with complete information, Greenwald and Jafari (2003) show that the tightest game-theoretic solution concept to which regret minimizing algorithms can (provably) converge is correlated equilibrium. For regret notions that are more refined than internal regrets, convergence properties cannot be established. Thus, our approach leverages the most informative restrictions implied by regret minimization while, at the same time, taking advantage of the possibility of using a limiting simple equilibrium restriction for estimation purposes. Because of the "black-box reduction," however, more informative restrictions do not come with additional assumptions on players knowledge or understanding of the environment.

6.2 Asymptotic ε-Regret and Orders on Information Structures

There are (at least) two natural orderings on information structures: an "incentive ordering" and a "statistical ordering". Roughly speaking, we have the following.[18]

---

[18]We refer to Bergemann and Morris (2016) for the formal definitions.

- 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[19] 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 address the following robustness question for our empirical exercise. Fix a limiting basic game $G^L(\lambda_0^G)$. Suppose the researcher assumes that the information structure is S^L, but the true information structure is $\tilde{S}^L$, for some $\tilde{S}^L$ that is individually sufficient for S^L, the exact $\tilde{S}^L$ being unknown to the researcher. Can the researcher still recover valid bounds on λ_0^G under the ε-AR assumption?

The answer to the previous question is positive. Let $E(\varepsilon)$ be the set of ε-BCE of $(G^L(\lambda_0^G), S^L)$ and let $\tilde{E}(\varepsilon)$ be the set of ε-BCE of $(G^L(\lambda_0^G), \tilde{S}^L)$. Suppose that the sequence of states, signals, and actions has the ε-AR property almost surely. Thus, by Theorem 1, the sequence of empirical distributions converges almost surely to $\tilde{E}(\varepsilon)$. As $\tilde{S}^L$ is individually sufficient for S^L, by Bergemann and Morris (2016)'s equivalence result, $\tilde{S}^L$ is also more incentive constrained than S^L, and so $\tilde{E}(\varepsilon) \subseteq E(\varepsilon)$. But then, the sequence of empirical distributions converges almost surely also to $E(\varepsilon)$. As a result, the bounds on λ_0^G under the ε-AR assumption remain valid, although they might not be as sharp as those one would obtain under the correct specification of the information structure.

The previous discussion also suggests that the results in this paper provide a novel interpretation of estimates obtained under BCE. While BCE has been adopted for estimation to weaken assumptions on information (Magnolfi and Roncoroni, 2021; Syrgkanis et al., 2021; Gualdani and Sinha, 2020), our results imply that BCE identified sets are still valid even if the data are not generated by BNE,

---

[19]A combination of the two information structures is a new information structure in which a pair of signals—one from each information structure—is observed and the marginal probability over signals from each of the original information structures corresponds to the original distribution over signals for that information structure.

insofar the ε-regret property holds. Thus, robustness with respect to assumptions on equilibrium play comes at no additional cost—in terms of informativeness of the estimation procedure—when one pursues robustness with respect to assumptions on information.

7   Conclusion

We develop a new approach to estimate a game's primitives when players interact in a complex dynamic environment. Because of the environment's characteristics, players may not know or understand key aspects of the interaction. This setup motivates a departure from standard equilibrium assumptions. In contrast, we pursue an adaptive approach that builds on the idea that learning outcomes can be used as an alternative to solution concepts when analyzing game dynamics. We impose an asymptotic ε-regret property on the observed play. Under the asymptotic ε-regret property, we first prove that the empirical distribution of actions converges to the set of Bayes correlated ε-equilibrium predictions of the underlying limiting stage game. Then, we use this static equilibrium notion to construct set-valued estimators of the parameters of interest. The econometric properties of these estimators directly arise from the theoretical convergence results under the asymptotic ε-regret property. We also construct theoretical bounds on how informative the estimated set of parameters can be. The method applies to panel data as well as to cross-sectional data interpreted as long-run outcomes of learning dynamics.

Monte Carlo evidence for a repeated binary pricing game shows that our estimation strategy delivers informative bounds on parameters for a simple two-seller pricing game. In an empirical application, we use the method to recover the distribution of sellers' marginal costs from data on pricing behavior on Swappa—an e-commerce platform for used cellphones and other portable electronic devices. The distribution of sellers' marginal costs is a key input to inform policy exercises and market-design counterfactuals. Our method produces estimates of seller-specific average costs that are plausible in this economic environment.

This paper leaves open several avenues for future research. We mention a few. First, the methods we develop have a counterpart for single-agent dynamics. A natural extension of our approach can be used to study the estimation of repeated discrete choice models in complex environments. Second, the foundational assumptions of our approach, the ε-AR property, may be formally tested using either experimental or non-experimental data. Such a test would help to better

define the scope of empirical environments where our method applies. Finally, in parallel work in progress, we show that coarse correlated ε-equilibrium restrictions on payoffs provide useful restrictions to robustly identify the structural primitives in infinitely repeated games with perfect monitoring or imperfect public monitoring under equilibrium assumptions. Our approach is robust in two respects. First, the method does not rely on equilibrium selection assumptions besides subgame perfection in repeated games with public monitoring and perfect public equilibrium in repeated games with imperfect public monitoring. Second, the method does not require the analyst to fully specify the informativeness of the game's monitoring structure or agents' patience.

A Proofs for Sections 3.2 and 4

A.1 Proof of Theorem 1

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

Consistency. Pick any $(t, \theta) \in T \times \Theta'$. Note the following:

$$
\begin{aligned}
\sum_a \nu(a, t, \theta) &= \sum_a \lim_{l\to\infty} Z^{N_l}(a, t, \theta) \\
&= \lim_{l\to\infty} \sum_a Z^{N_l}(a, t, \theta) \\
&= \lim_{l\to\infty} \left[\frac{\sum_a Z^{N_l}(a, t, \theta)}{\sum_{a,t} Z^{N_l}(a, t, \theta)} \sum_{a,t} Z^{N_l}(a, t, \theta) \right] \\
&= \lim_{l\to\infty} \frac{\sum_a Z^{N_l}(a, t, \theta)}{\sum_{a,t} Z^{N_l}(a, t, \theta)} \lim_{l\to\infty} \sum_{a,t} Z^{N_l}(a, t, \theta) \\
&= \lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} \lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)}{N_l}.
\end{aligned}
\tag{7}
$$

The ratio

$$
\frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)}
\tag{8}
$$

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 $\pi(\cdot \mid \theta^n)$, (8) is the empirical frequency of $\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)$ conditionally independent observations from $\pi(\cdot \mid \theta)$. Moreover, since almost all paths of process $(\theta^n)_{n\in\mathbb{N}}$ have a limiting empirical distribution ψ^L with full support on Θ^L, $\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n) \to \infty$ as $l \to \infty$ almost surely. Thus, by the strong law of large numbers,

$$
\lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{t\}}(t^n)\mathbb{1}_{\{\theta\}}(\theta^n)}{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)} = \pi^L(t \mid \theta) \quad \text{a.s.}
\tag{9}
$$

Moreover, again because almost all paths of process $(\theta^n)_{n\in\mathbb{N}}$ have a limiting empirical distribution ψ^L with full support on Θ^L

$$
\lim_{l\to\infty} \frac{\sum_{n=1}^{N_l} \mathbb{1}_{\{\theta\}}(\theta^n)}{N_l} = \psi^L(\theta) \quad \text{a.s.}
\tag{10}
$$

Together, (7), (9), and (10) give

$$
\sum_a \nu(a, t, \theta) = \pi^L(t \mid \theta)\psi^L(\theta) \quad \text{a.s.}
\tag{11}
$$

As $(t, \theta) \in T^L \times \Theta^L$ was arbitrarily chosen, we conclude from (11) that ν is almost surely consistent for (G^L, S^L).

ε-obedience. To begin, note the following:

$$V_i(a_i, a_i', N, t_i) - U_i(t_i, N)$$

$$= \frac{1}{N} \sum_{n=1}^{N} \left[u_i\big((a_i', a_{-i}^n), \theta^n)\big) - u_i\big((a_i^n, a_{-i}^n), \theta^n)\big) \right] \mathbb{1}_{\{a_i\}}(a_i^n) \mathbb{1}_{\{t_i\}}(t_i^n)$$

$$= \frac{1}{N} \sum_{\theta} \sum_{n=1}^{N} \left[u_i\big((a_i', a_{-i}^n), \theta^n)\big) - u_i\big((a_i^n, a_{-i}^n), \theta^n)\big) \right] \mathbb{1}_{\{a_i\}}(a_i^n) \mathbb{1}_{\{t_i\}}(t_i^n) \mathbb{1}_{\{\theta\}}(\theta^n) \tag{12}$$

$$= \sum_{a_{-i}, t_{-i}, \theta} \left[u_i\big((a_i', a_{-i}), \theta\big) - u_i\big((a_i, a_{-i}), \theta\big) \right] Z^N\big((a_i, a_{-i}), (t_i, t_{-i}), \theta\big).$$

Now pick any $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i, a_i' \in A_i$. As $\limsup_{N \to \infty} R_i(a_i, a_i', t_i, N) \leq \varepsilon_i(a_i, a_i', t_i)$ a.s., by definition of $R_i(a_i, a_i', t_i, N)$, we also have

$$\limsup_{N \to \infty} \left[V_i(a_i, a_i', t_i, N) - U_i(t_i, N) \right] \leq \varepsilon_i(a_i, a_i', t_i) \quad \text{a.s.} \tag{13}$$

Then, by (12) and (13),

$$\limsup_{N \to \infty} \sum_{a_{-i}, t_{-i}, \theta} \left[u_i\big((a_i', a_{-i}), \theta\big) - u_i\big((a_i, a_{-i}), \theta\big) \right] Z^N\big((a_i, a_{-i}), (t_i, t_{-i}), \theta\big) \leq \varepsilon_i(a_i, a_i', t_i) \tag{14}$$

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

$$\lim_{l \to \infty} \sum_{a_{-i}, t_{-i}, \theta} \left[u_i\big((a_i', a_{-i}), \theta\big) - u_i\big((a_i, a_{-i}), \theta\big) \right] Z^{N_l}\big((a_i, a_{-i}), (t_i, t_{-i}), \theta\big)$$

$$= \sum_{a_{-i}, t_{-i}, \theta} \lim_{l \to \infty} \left[u_i\big((a_i', a_{-i}), \theta\big) - u_i\big((a_i, a_{-i}), \theta\big) \right] Z^{N_l}\big((a_i, a_{-i}), (t_i, t_{-i}), \theta\big) \tag{15}$$

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

Together, (14) and (15) give

$$\sum_{a_{-i}, t_{-i}, \theta} \left[u_i\big((a_i', a_{-i}), \theta\big) - u_i\big((a_i, a_{-i}), \theta\big) \right] \nu\big((a_i, a_{-i}), (t_i, t_{-i}), \theta\big) \leq \varepsilon_i(a_i, a_i', t_i) \quad \text{a.s.} \tag{16}$$

As $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i, a_i' \in A_i$ were arbitrarily chosen, we conclude from (16) that ν is almost surely *ε-obedient* for (G^L, S^L).

45

[$\Longleftarrow$] Now suppose $(Z^N)_{N\in\mathbb{N}}$ converges almost surely to $E(\varepsilon)$ for some ε. Pick any $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i, a_i' \in A_i$. By ε-obedience,

$$\limsup_{N\to\infty} \sum_{a_{-i},t_{-i},\theta} \left[u_i\big((a_i',a_{-i}),\theta\big) - u_i\big((a_i,a_{-i}),\theta\big)\right] Z^N\big((a_i,a_{-i}),(t_i,t_{-i}),\theta\big) \leq \varepsilon_i(a_i,a_i',t_i) \quad (17)$$

holds almost surely. By (12) and (17),

$$\limsup_{N\to\infty} \left[V_i(a_i,a_i',t_i,N) - U_i(t_i,N)\right] \leq \varepsilon_i(a_i,a_i',t_i) \quad \text{a.s.,}$$

which implies

$$\limsup_{N\to\infty} R_i(a_i,a_i',t_i,N) \leq \varepsilon_i(a_i,a_i',t_i) \quad \text{a.s.}$$

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

A.2   Proof of Theorem 2

Proof of part 1.   For any $\varepsilon' > \varepsilon$, by definition of $\hat{\Lambda}^K(\varepsilon')$, we have:

$$\lambda_0 \in \hat{\Lambda}^K(\varepsilon') \iff q^K \in Q(\lambda_0;\varepsilon'). \tag{18}$$

Under Assumption 2, the sequence $(q^K)_{N\in\mathbb{N}}$ converges almost surely to $Q(\varepsilon;\lambda)$ as $K \to \infty$ (by Theorem 1 and Corollary 1). Then, by Remark 1, almost surely, for any $\varepsilon' > \varepsilon$, there exists N' such that $q^K \in Q(\lambda_0;\varepsilon')$ for all $K > N'$. Combining this fact with (18) gives the desired result. ∎

Proof of part 2.   It follows from part 1 and the definition of $\inf \hat{\Lambda}^N(\varepsilon')$ in (4). ∎

A.3   Proof of Theorem 3

Proof of part 1.   For any $\varepsilon' > \varepsilon$, by definition of $\hat{\Lambda}^K(\varepsilon')$ and of $\Lambda^{OR}(\lambda_0;\varepsilon')$, we have:

$$q^K \in Q(\lambda_0;\varepsilon') \iff \left\{\lambda \in \Lambda : q^K \in Q(\lambda;\varepsilon')\right\} \subseteq \Lambda^{OR}(\lambda_0;\varepsilon') \iff \hat{\Lambda}^K(\varepsilon') \subseteq \Lambda^{OR}(\lambda_0;\varepsilon'). \tag{19}$$

Under Assumption 2, almost surely, for any $\varepsilon' > \varepsilon$, there exists N' such that $q^K \in Q(\lambda_0;\varepsilon')$ for all $K > N'$ (by Theorem 1 and Corollary 1, see proof of Theorem 2). Combining this with (19) gives the desired result. ∎

Proof of part 2.   Under Assumption 2, almost surely, for any $\varepsilon' > \varepsilon$, there exists N' such that $q^K \in Q(\lambda_0;\varepsilon')$ for all $K > N'$ (by Theorem 1 and Corollary 1, see proof of Theorem 2). Thus, by definition of $\Lambda^{IR}(\lambda_0;\varepsilon')$ and $\inf \hat{\Lambda}^N(\varepsilon')$, almost surely, for any $\varepsilon' > \varepsilon$, there exists

N' such that

$$\Lambda^{IR}(\lambda_0; \boldsymbol{\varepsilon}') \subseteq \inf \hat{\Lambda}^N(\boldsymbol{\varepsilon}')$$

for all $N > N'$, which gives the desired result. ∎

A.4   Proof of Proposition 1

First, since $Q(\lambda; \boldsymbol{\varepsilon}')$ is a non-empty, closed, and convex set, it has the following (support-function) characterization:

$$q^N \in Q(\lambda; \boldsymbol{\varepsilon}') \iff \left\langle b^\intercal q^N - \sup_{q \in Q(\lambda;\boldsymbol{\varepsilon}')} b^\intercal q \leq 0 \quad \text{for all } b \in B^{|A|} \right\rangle. \tag{20}$$

Second, since $Q(\lambda; \boldsymbol{\varepsilon}')$ is also bounded and $b^\intercal q$ is continuous, (20) is equivalent to

$$q^N \in Q(\lambda; \boldsymbol{\varepsilon}') \iff \left\langle b^\intercal q^N - \max_{q \in Q(\lambda;\boldsymbol{\varepsilon}')} b^\intercal q \leq 0 \quad \text{for all } b \in B^{|A|} \right\rangle. \tag{21}$$

Third, since $b^\intercal q^N - \max_{q \in Q(\lambda;\boldsymbol{\varepsilon}')} b^\intercal q = \min_{q \in Q(\lambda;\boldsymbol{\varepsilon}')} b^\intercal (q^N - q)$ and $\min_{q \in Q(\lambda;\boldsymbol{\varepsilon}')} b^\intercal (q^N - q)$ evaluated at $b = 0_{|A|} \in B$ is equal to 0, (21) is equivalent to

$$q^N \in Q(\lambda; \boldsymbol{\varepsilon}') \iff \max_{b \in B^{|A|}} \min_{q \in Q(\lambda;\boldsymbol{\varepsilon}')} b^\intercal (q^N - q). \tag{22}$$

The desired result follows from (22), the definition of $\hat{\Lambda}^N(\boldsymbol{\varepsilon}')$ in (3), and the definition of $g(\cdot; q^N, \boldsymbol{\varepsilon}')$ in (5). ∎

B   Details for Illustrative Examples

Parametrization of $\boldsymbol{\varepsilon}$.   We describe how we parametrize $\boldsymbol{\varepsilon}$ as a function of a non-negative real number ε in the illustrative examples of Sections 3.3 and 4.4. Fix $i \in \mathcal{I}, a_i, a_i' \in A_i$, and $t_i \in T_i$, and consider $\boldsymbol{\varepsilon}$-obedience:

$$\sum_{a_{-i}, t_{-i}} \left[u_i\big((a_i', a_{-i}), t_i; \lambda\big) - u_i\big((a_i, a_{-i}), t_i; \lambda\big) \right] \nu\big((a_i, a_{-i}), (t_i, t_{-i})\big) \leq \varepsilon_i(a_i, a_i', t_i).$$

The left-hand side of the inequality provides a natural scale for $\varepsilon_i(a_i, a_i', t_i)$. To ensure computational tractability and not have $\varepsilon_i(a_i, a_i', t_i)$ depend on equilibrium objects (i.e., on ν), we consider the following upper bound:

$$\sum_{a_{-i}, t_{-i}} \left[u_i\big((a_i', a_{-i}), t_i; \lambda\big) - u_i\big((a_i, a_{-i}), t_i; \lambda\big) \right] \nu\big((a_i, a_{-i}), (t_i, t_{-i})\big)$$

$$\leq \sum_{a_{-i}, t_{-i}} \left| u_i\big((a_i', a_{-i}), t_i; \lambda\big) - u_i\big((a_i, a_{-i}), t_i; \lambda\big) \right| \nu\big((a_i, a_{-i}), (t_i, t_{-i})\big)$$

$$\leq \max_{a_{-i}} \left| u_i\big((a_i', a_{-i}), t_i; \lambda\big) - u_i\big((a_i, a_{-i}), t_i; \lambda\big) \right| \left(\sum_{a_{-i}, t_{-i}} \nu\big((a_i, a_{-i}), (t_i, t_{-i})\big) \right)$$

$$\leq \max_{a_{-i}} \left| u_i\big((a_i', a_{-i}), t_i; \lambda\big) - u_i\big((a_i, a_{-i}), t_i; \lambda\big) \right| \psi(t_i; \lambda),$$

which is computable for any guess of the parameter λ. Thus, to compute the function $g(\lambda; q^N, \varepsilon)$ we set

$$\varepsilon_i(a_i, a_i', t_i; \lambda) = \varepsilon \max_{a_{-i}} \left| u_i\big((a_i', a_{-i}), t_i; \lambda\big) - u_i\big((a_i, a_{-i}), t_i; \lambda\big) \right| \psi(t_i; \lambda),$$

where $\varepsilon \in [0, 1]$. Intuitively, this parametrization scales $\varepsilon_i(a_i, a_i', t_i; \lambda)$ as a fraction ε of an upper bound on player i's maximum payoff difference from a deviation from a_i to a_i' when his signal is t_i.

Although the parametrization introduces a dependence of ε on λ, this does not have consequences on the estimates: in simulations, we obtain qualitatively identical bounds if we set $\lambda = \lambda_0$. Alternatively, the procedure could be implemented for a fixed guess of λ.

C   Computational Appendix

In this appendix, we present the details for the computation of the function $g(\cdot; q^N, \varepsilon') \colon \Lambda \to \mathbb{R}$, defined pointwise by

$$g\big(\lambda; q^N, \varepsilon'\big) := \max_{b \in B^{|A|}} \min_{q \in Q(\lambda; \varepsilon')} b^\mathsf{T}\big(q^N - q\big), \tag{P0}$$

which characterizes the set estimator $\hat{\Lambda}^N(\varepsilon')$. We consider the case of our illustrative example and empirical application in which the limiting stage game has independent private values and is the same as the stage game: $\Theta = T$, $u_i(a, \theta; \lambda) = u_i(a, t_i; \lambda)$, and $\pi \colon \Theta \to \Delta(T)$ is such that $\pi(t \mid \theta) = 1$ if and only if $t = \theta$. The computational procedure, however, plainly applies to the general setting of Sections 2–4.

Rewriting P0.   To begin, rewrite problem (P0) as a constrained optimization problem by making explicit the constraint that $q \in Q(\lambda; \varepsilon')$. In particular, the optimization problem (P0) is equivalent to the following problem:

$$\max_{b \in \mathbb{R}^{|A|}} \min_{\substack{q \in \mathbb{R}_+^{|A|} \\ \nu \in \mathbb{R}_+^{|A| \cdot |T|}}} b^\mathsf{T}\big(q^N - q\big) \tag{P1}$$

subject to

$$b^\mathsf{T} b - 1 \leq 0,$$

$$q(a) - \sum_t \nu(a, t) = 0 \quad \forall a \in A,$$

$$\sum_a \nu(a, t) - \psi(t; \lambda) = 0 \quad \forall t \in T,$$

$$\sum_{a,t} \nu(a, t) - 1 = 0,$$

$$\sum_{a_{-i}, t_{-i}} \left[u_i(a_i', a_{-i}, t_i; \lambda) - u_i(a, t_i; \lambda) \right] \nu(a, t_i, t_{-i}) - \varepsilon \leq 0 \quad \forall i \in \mathcal{I}, a_i, a_i' \in A_i, t_i \in T_i.$$

The first constraint is equivalent to $b \in B^{|A|}$. The second, third, and fifth constraints, together, are equivalent to $q \in Q(\lambda; \varepsilon')$; in particular, the third and fifth constraints correspond to the ε'-BCE restrictions on ν (consistency and ε'-obedience), and the second constraint requires q to be the ε'-BCE prediction corresponding to ν. The fourth constraint is superfluous—it is implied by the third constraint and the fact that $\psi(\cdot; \lambda)$ is a probability distribution; we add the constraint explicitly to help computation.

Vectorization. Next, since problem (P1) is linear, we write it in matrix form. We represent the probability distribution $\psi(\cdot; \lambda)$ as a $|A| \times 1$ vector with elements $\psi(a^m; \lambda)$ for $1 \leq m \leq |A|$; analogously, we represent the probability distributions q and q^N as $|A| \times 1$ vectors. We represent a probability distribution $\nu \in \Delta(A \times T)$ as an $|A| \times |T|$ matrix with entries $\nu(a^m, t^n)$ for $1 \leq m \leq |A|$ and $1 \leq n \leq |T|$. We define $v := \text{vec}(\nu)$ to be the vectorization of the matrix representation of ν; the linear transformation $\text{vec}(\nu)$ stacks the columns of the matrix representation of ν on top on one other to obtain the $d_v \times 1$ vector v, where $d_v := |A| \cdot |T|$. Moreover, we define $z_1 := q^N - q$, $z_2 := v$, and the $d_z \times 1$ vector

$$z := \begin{bmatrix} z_1 \\ z_2 \end{bmatrix},$$

where $d_z := |A| \cdot (1 + |T|)$. Hereafter, we denote by 0_d $d \times 1$ vectors whose components are all zeros.

The equality constraints in (P1) are linear. Thus, we write them in matrix form as

$$M_{\text{eq}} z = y,$$

where: (i) M_{eq} is the appropriately defined $d_{\text{eq}} \times d_z$ matrix of coefficients, with $d_{\text{eq}} := |A| + |T| + 1$; (ii) y is the $d_{\text{eq}} \times 1$ vector defined by

$$y := \begin{bmatrix} q^N \\ \psi(\cdot; \lambda) \\ 1 \end{bmatrix}.$$

The inequality constraints in (P1) are also linear. Thus, we write them in matrix form as

$$M_{\text{ineq}} z \leq \varepsilon',$$

where: (i) M_{ineq} is the appropriately defined $d_{\text{ineq}} \times d_z$ matrix of coefficients, with $d_{\text{ineq}} :=$ $\sum_i (|A_i| \cdot (|A_i| - 1) \cdot |T_i|)$; (ii) ε' is the appropriately ordered $d_{\text{ineq}} \times 1$ vector.

Finally, since $Q(\lambda; \varepsilon')$ is a subset of the $(|A| - 1)$-dimensional simplex,

$$\max_{b \in B^{|A|}} \min_{q \in Q(\lambda; \varepsilon')} b^{\mathsf{T}}(q^N - q) = \max_{\tilde{b} \in B^{|A|-1}} \min_{q \in Q(\lambda; \varepsilon')} \begin{bmatrix} \tilde{b} \\ 0 \end{bmatrix}^{\mathsf{T}} (q^N - q),$$

where $B^{|A|-1}$ is the closed unit ball centered at $0_{|A|-1} \in \mathbb{R}^{|A|-1}$.

Therefore, we can now rewrite problem (P1) in the following equivalent form:

$$\max_{\tilde{b} \in \mathbb{R}^{|A|-1}} \min_{\substack{z_1 \in \mathbb{R}^{|A|} \\ z_2 \in \mathbb{R}_+^{d_v}}} \begin{bmatrix} \tilde{b} \\ 0_{d_v+1} \end{bmatrix}^{\mathsf{T}} z \tag{P2}$$

subject to

$$\tilde{b}^{\mathsf{T}} \tilde{b} \leq 1,$$
$$M_{\text{eq}} z = y,$$
$$M_{\text{ineq}} z \leq \varepsilon'.$$

Duality. Finally, we replace the inner linear constrained minimization problem in (P2) by its dual to obtain the following linear constrained maximization problem:

$$\max_{\substack{\tilde{b} \in \mathbb{R}^{|A|-1} \\ \ell_{\text{eq}} \in \mathbb{R}^{d_{\text{eq}}} \\ \ell_{\text{ineq}} \in \mathbb{R}_+^{d_{\text{ineq}}}}} - \begin{bmatrix} y \\ \varepsilon' \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \ell_{\text{eq}} \\ \ell_{\text{ineq}} \end{bmatrix} \tag{P3}$$

subject to

$$\tilde{b}^{\mathsf{T}} \tilde{b} \leq 1,$$
$$[M^{\mathsf{T}}]_{1:|A|} \begin{bmatrix} \ell_{\text{eq}} \\ \ell_{\text{ineq}} \end{bmatrix} = - \begin{bmatrix} \tilde{b} \\ 0 \end{bmatrix},$$
$$[M^{\mathsf{T}}]_{(|A|+1):d_z} \begin{bmatrix} \ell_{\text{eq}} \\ \ell_{\text{ineq}} \end{bmatrix} \geq 0_{d_v},$$

where

$$M := \left[\begin{array}{c} M_{\mathrm{eq}} \\ M_{\mathrm{ineq}} \end{array} \right],$$

the dual variables associated to the constraints of (P2) are the $d_{\mathrm{eq}} \times 1$ vector ℓ_{eq} and the $d_{\mathrm{ineq}} \times 1$ vector ℓ_{ineq}, and $[M^{\mathsf{T}}]_{r:s}$ denotes the matrix consisting of rows $r, r+1, \ldots, s$ of matrix M^{T}. Let $d_M := d_{\mathrm{eq}} + d_{\mathrm{ineq}}$ denote the number of rows of M. By strong duality, problem (P3) has the same value as problem (P2). Problem (P3) can be efficiently computed by using standard solvers.

D Empirical Application: Further Details

D.1 Sample of Listings

We describe in this appendix the sample of listings in our data. In total, we have 12,741 product listings by 2,436 sellers. We only collect data on Apple iPhone products. This is because there is a large difference between the popularity of iPhone and other cell phones on Swappa during the period that we study. Over the 80-day scraping period, different generations of iPhone vary in terms of their popularity. As shown in Figure 8, the number of listings for each product exceeded 100 for each day during our observation period, reaching a maximum of nearly 500.

Figure 8: Product Listings per Day.
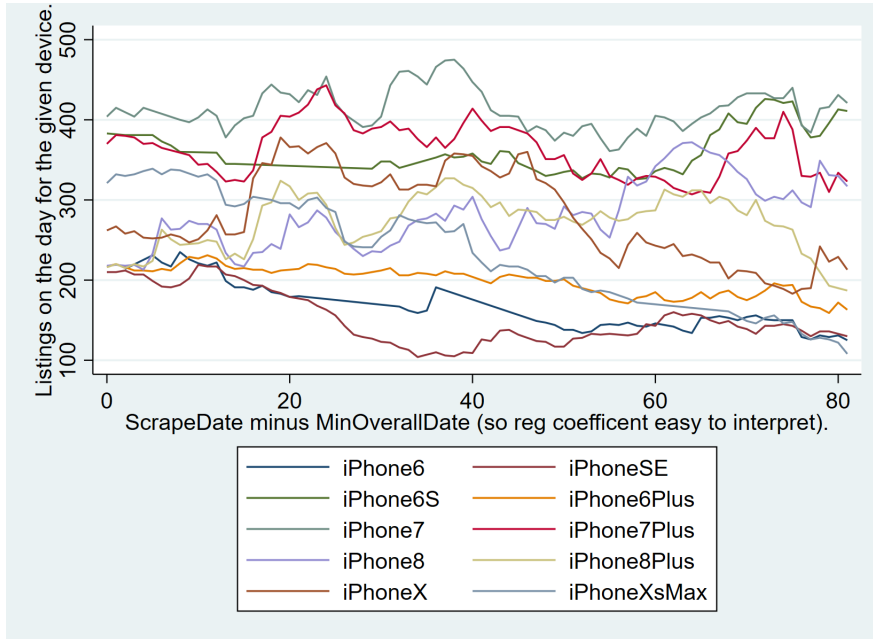
D.2 Heterogeneity across Sellers

We present in this appendix some evidence on seller's characteristics and heterogeneity across sellers. Table 2 shows summary statistics of characteristics of the 2,436 sellers in our data. Sellers have on average of 4.7 active listings, with a long right tail, since few sellers are categorized as business sellers or "trusted" by the platform.

Table 2: Sellers' Characteristics.

Characteristic	N	Mean	Sd	Explanation
i_Promoted	2,436	0.091	0.288	Indicator for listings being promoted
ListActive	2,436	4.755	54.10	Number of active listings for a seller
i_Trusted	2,436	0.025	0.155	Indicator for if a seller is trusted
i_BusinessSeller	2,436	0.006	0.078	Indicator for if a seller is a business

We further analyze the characteristics of the 59 sellers with more than 50 active listings in Table 3: the average number of active listings is 176.3 in this subsample.

Table 3: Characteristics for Sellers with >50 Active Listings.

Characteristic	N	Mean	Sd	Explanation
i_Promoted	59	0.254	0.439	Indicator for the listing being promoted
ListActive	59	176.3	303.4	Number of active listings for the sellers
i_Trusted	59	0.288	0.457	Indicator for if a seller is trusted
i_BusinessSeller	59	0.153	0.363	Indicator for if a seller is a business

We refine the set of sellers by showing the top 15 sellers in Table 4 by number of listings over the sample period. In this set, two sellers (after Tigerphones and GM DEALS) have a markedly different profile, since they seem to sell less than 5% of their listings, while other sellers typically manage to sell more than 90% of their listings over the period of our study. This is due to the fact that these sellers seem to use the same listings to sell multiple devices of the same model, thus interfering with our data collection exercise. Therefore, we exclude these two sellers from the analysis.

Table 4: Large Sellers—Total Listings and Sales.

Seller Name	Seller Total Sold	Seller Total Listings
Sellworld	597	763
Wisephone ER	385	558

Continued on next page

Table 4 – *Continued from previous page*

Seller name	Seller total sold	Seller total listings
TPSC	290	462
Certified Gadgets	288	426
Sam H.	178	425
GM DEALS	2	385
Tigerphones	0	363
Pangea Deals	42	302
We Buy Mobile Corp.	17	261
CellPros	173	237
Super Value Cellular	223	229
Noori	153	202
Smartphone Hub	179	198
Winston V.	156	193
cellularreturns	155	172

D.3 Pricing Patterns

We collect price data for each listing every hour until it is sold or closed. While we assume in Section 5 that pricing decisions by sellers are daily, pricing patterns display a considerable heterogeneity in the data. As in previous literature (Ellison, Snyder, and Zhang, 2018), we observe substantial inertia in prices, and only half of the listings change prices during our sample period. Most of the price change strategies in the product listings are "gradual decrease". We exclude outlier price changes of more than $100, and plot the histogram of price changes in Figure 9.

Figure 10 shows the frequency of price changes. As can be seen from the figure, most price changes occur within 15 days. Within 15 days, the frequency of price changes clearly shows a tendency to decrease with time. After 15 days, the willingness to change prices becomes less strong. Therefore, we can assume that sellers who are willing to adjust their prices tend to start adopting a price reduction strategy soon after they put their products on the shelves.

Table 5 shows the days it takes for each listing to sell. The table shows that the average number of days for a device to sell is less than 15 on average, with the longest being 14.52 days for the iPhone 6plus and the shortest being 6.613 days for the iPhone 6. This is highly consistent with the frequency of price changes: most sellers tend to sell their devices within 15 days by adjusting the price over time.

Figure 9: Difference between Last and First Price.

Figure 10: Price Changing Frequency.
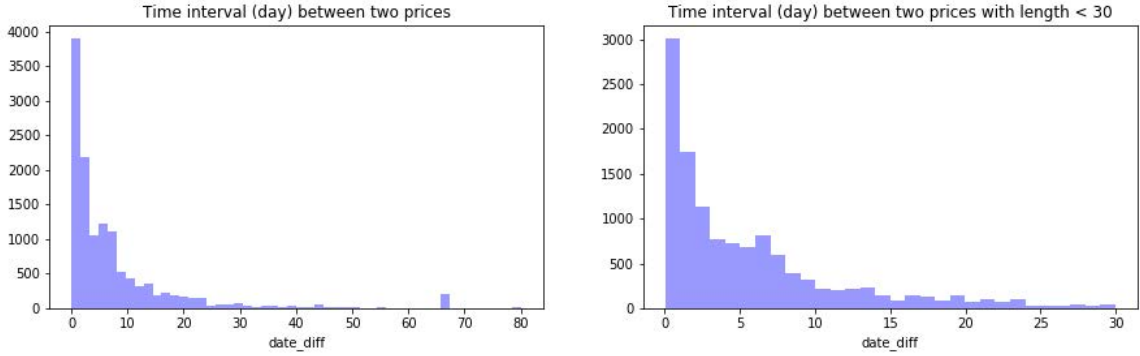
D.4 Homogenization and Sale Probabilities

In this appendix we describe the homogenization of device prices and the estimation of sale probabilities.

D.4.1 Homogenization

Similar to Haile et al. (2003) and Wildenbeest (2011), we seek to homogenize different pricing games by removing the effect of observable characteristics. To do so, we estimate the following regression using daily pricing observations:

$$p_i^n = x_{m(j)d}^\intercal \beta + \rho_i^n, \tag{23}$$

where p_i^n is the price that seller i chooses for device-day n, and $x_{m(j)d}$ are covariates. These include a "trusted seller" indicator, seller rating, warranty length, indicator for other non-

Table 5: Time from Original Listing to Sale.

Variables	N	mean	min	max	p25	p50	p75
iPhone 6	284	6.613	0	16	2	4	12
iPhone SE	631	13.57	0	27	7	14	19
iPhone 6S	416	8.358	0	22	3	7	13
iPhone 6 Plus	554	14.52	0	28	7	13	24
iPhone 7	2,396	9.988	0	29	5	9	14
iPhone 7 Plus	1,663	13.79	0	29	8	14	19
iPhone 8	2,813	12.49	0	26	6	13	20
iPhone 8 Plus	2,478	12.97	0	29	6	12	20
iPhone X	2,514	13.24	0	29	7	12	19
iPhone Xs Max	115	6.852	0	14	3	7	10

standard warranty, whether the listing is promoted, storage, condition and date $\times$ device fixed effects.

Table 6 shows the regression results over the full sample, and in three subsamples which describe different website traffic conditions (busy, medium, slow). The subsamples are constructed by obtaining device-model specific measures of traffic by counting the number of listings that day, and then denoting as busy, medium and slow, respectively, the observations that lie in the 0-33rd, 34th-66th, and 67th-100th percentiles of the distribution of listings.

The results in Table 6 show mostly intuitive correlations. Higher sellers' ratings result in higher prices, but the trusted seller indicator is ceteris-paribus associated with lower prices. Warranty length has no impact, but the presence of additional warranty seems to positively impact prices. Promoted listings have a small price premium. Interestingly, some of the indicators show a U-shape with the change of traffic conditions. For example, the coefficients of seller rating, other warranty and promoted listing variables are the smallest at medium and increase significantly at both busy and slow. Other variables show an inverted U-shape with traffic conditions. For instance, the coefficients of the trusted seller and the are largest at medium and smaller at busy and slow.

After we obtain the regression results, we compute the pricing residuals ρ_i^n and show their distributions for different devices in Figure 11. Some devices, such as iPhone XS Max, seem to have higher variance in their residual prices than others.

For each listing, sellers choose a "high" price $\bar{\rho}$ when the corresponding residual ρ_i^n is positive. The Figure 12 shows (device-by-device) how the fraction of high-priced products – according to our discretized definition — evolves over time. By and large, the fraction of high-priced devices hovers around 50%, with considerable variation day-to-day and across devices. From Figure 13 we also observe considerable dispersion in sellers' strategies.

Table 6: Regression Results in Different Device Traffic Situation.

Trusted Seller	-19.317***	-18.683***	-20.240***	-18.172***
	(-62.83)	(-37.05)	(-37.26)	(-32.08)
Seller Rating	27.562***	32.382***	18.709	30.103***
	(4.39)	(35.00)	(1.31)	(4.42)
Warranty Length	-0.020		-0.070	0.016
	(-0.58)		(-0.89)	(0.42)
Other Warranty	7.528***	10.668***	2.259	7.300***
	(7.27)	(5.79)	(1.13)	(4.69)
Promoted Listing	3.687***	5.141***	2.796***	3.360***
	(12.72)	(10.57)	(5.65)	(6.50)
Constant	185.481***	163.422***	246.506***	154.698***
	(5.91)	(36.44)	(3.45)	(4.55)
Observations	27,881	10,765	7,973	9,143
R-Squared	0.989	0.989	0.989	0.990
Website Traffic	All	Busy	Medium	Slow

t-statistics in parentheses, *** p<0.01, ** p<0.05, * p<0.1
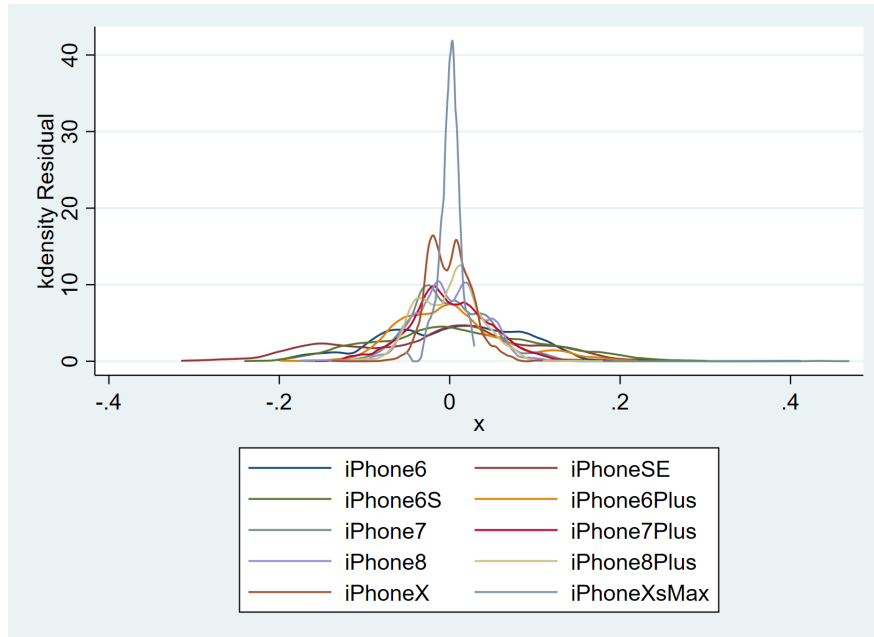
Figure 11: Residuals by Model.

56

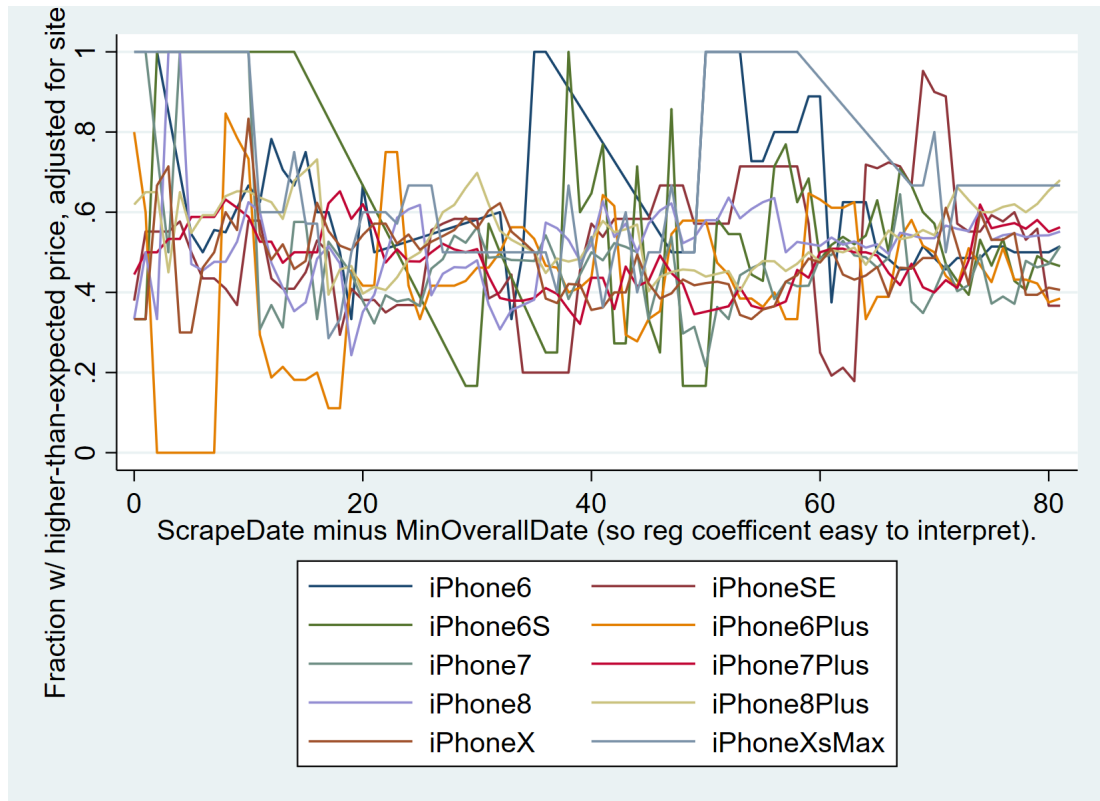Figure 12: Fraction of Higher-than-Expected Price.

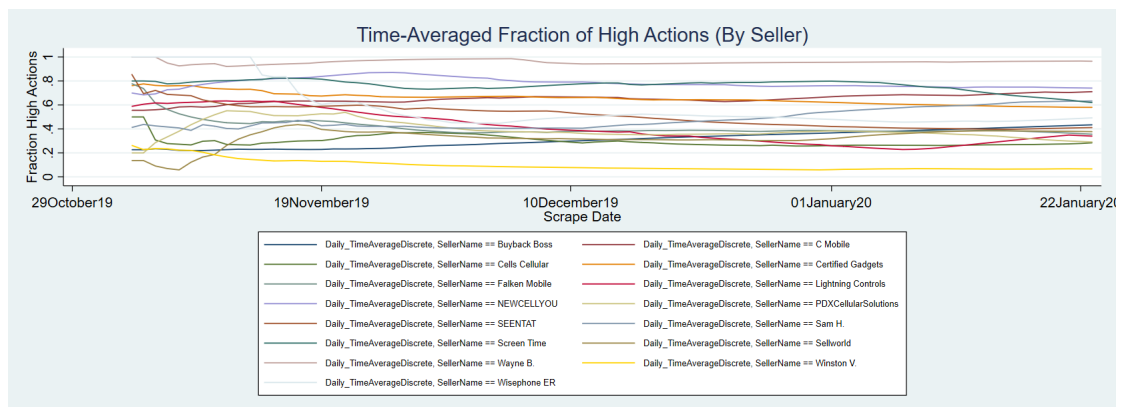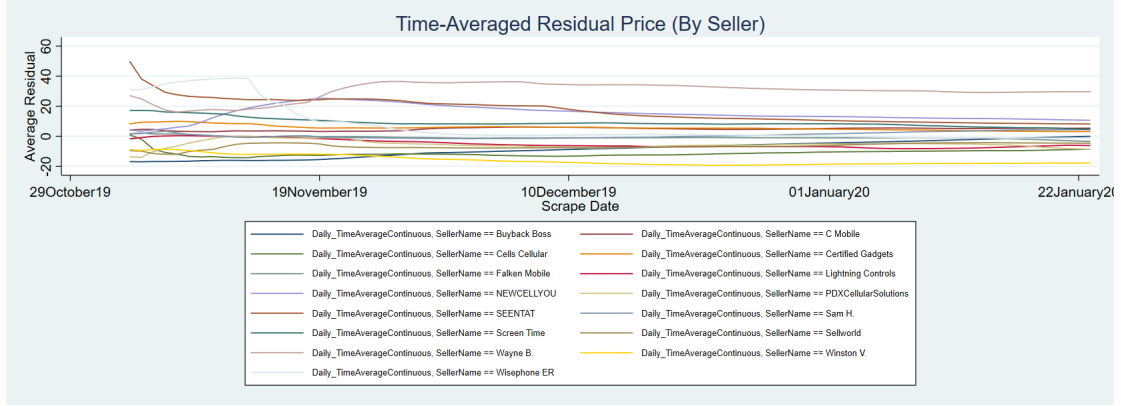Figure 13: Fraction of High Actions (by Seller).

Figure 14: Time-Averaged Residual Price.

D.4.2 Sale Probability

We describe in this appendix the estimation of the probability of sale $\tilde{g}$. We do so by simply computing

$$\tilde{g}_i(\rho_i^n, \rho_{-i}^n) = \#(Sold\ at\ (\rho_i^n, \rho_{-i}^n))/\#(Listings\ at\ (\rho_i^n, \rho_{-i}^n)).$$

In this equation, we estimate the probability of a listing selling at a given level of own and competitors' residualized prices (ρ_i^n, ρ_{-i}^n) as the share of listings sold over the total number of listings. Table 7 shows these estimates for all levels of own price and competitor residual prices.

Table 7: Sale Probabilities by Price Bin.

Own Price Bin	Competitor	Probability of Sale
High	High	.0349
High	Low	.0328
Low	High	.0767
Low	Low	.0636

E   External Regrets and Bayes Coarse Correlated Equilibrium

In this section, we first formalize the notions of external regrets, asymptotic ε^c-external regret property, and Bayes coarse correlated ε^c-equilibrium. Next, we study convergence to the set of Bayes coarse correlated ε^c-equilibria of the limiting stage game under the asymptotic ε^c-external regret property.

Remark 2. Hereafter, we refer to the regret notion introduced by Definition 1 as *internal regret.*

E.1 External Regrets and the ε^c-External Regret Property

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

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

Definition 10 (External Regret). *For all $i \in \mathcal{I}$, $t_i \in T_i$ and $a_i' \in A_i$, the external regret of player i with signal t_i for action a_i' before play at time $N+1$ is denoted by $R_i^{ext}(a_i', t_i, N)$ and defined by*

$$R_i^{\text{ext}}(a_i', t_i, N) := \max\big\{ V_i^{\text{ext}}(a_i', t_i, N) - U_i(t_i, N), 0 \big\}.$$

$R_i^{\text{ext}}(a_i', t_i, N)$ is a measure of the time average regret experienced by player i with signal t_i at period N for not having played action a_i' in all past periods up to N. When each player has at most two actions, external regrets coincide with internal regrets; otherwise, external regrets are a coarser measure of regret than internal regrets.

Notation. Let $\varepsilon^{\mathbf{c}} := (\varepsilon_i^c(a_i', t_i))_{i \in \mathcal{I}, a_i' \in A_i, t_i \in T_i^L}$ denote the vector that specifies, for all $i \in \mathcal{I}$, $a_i' \in A_i$, and $t_i \in T_i^L$, a non-negative real number $\varepsilon_i^c(a_i', t_i)$. We write $\varepsilon^{\mathbf{c}} = 0$ if $\varepsilon_i^c(a_i', t_i) = 0$ for all $i \in \mathcal{I}$, $a_i' \in A_i$, and $t_i \in T_i^L$. For any ε and $\varepsilon^{\mathbf{c}}$, with some abuse of notation, we write $\varepsilon^{\mathbf{c}} = \varepsilon$ if $\varepsilon_i^c(a_i', t_i) = \varepsilon_i(a_i, a_i', t_i)$ for all $i \in \mathcal{I}$, $a_i, a_i' \in A_i$, and $t_i \in T_i^L$.

Definition 11 (Asymptotic $\varepsilon^{\mathbf{c}}$-External Regret Property). *Fix an $\varepsilon^{\mathbf{c}}$. A sequence $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the asymptotic $\varepsilon^{\mathbf{c}}$-external regret (hereafter, $\varepsilon^{\mathbf{c}}$-AER) property if*

$$\limsup_{N \to \infty} R_i^{\text{ext}}(a_i', t_i, N) \leq \varepsilon_i^c(a_i', t_i) \tag{24}$$

for all $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i' \in A_i$.

To develop intuition, let us refer to player i with signal t_i as "player (i, t_i)." A sequence $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the asymptotic ε-internal regret property if the time average of the counterfactual increase in past payoffs, had each player (i, t_i) changed each past play of a given action with its best replacement in hindsight, becomes ε-close to vanish in the long run. In contrast, a sequence $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the asymptotic $\varepsilon^{\mathbf{c}}$-external regret property if the time average of the counterfactual increase in past payoffs, had each player (i, t_i) played the best fixed action in hindsight, becomes $\varepsilon^{\mathbf{c}}$-close to vanish in the long run. Clearly, when $\varepsilon = \varepsilon^{\mathbf{c}}$, if $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the asymptotic ε-internal regret property, then it also has the asymptotic $\varepsilon^{\mathbf{c}}$-external regret property.

E.2 Bayes Coarse Correlated ε^c-Equilibrium

Definition 12 (Bayes Coarse Correlated ε^c-Equilibrium)**.** *Fix an ε^c. The probability distribution $\nu \in \Delta(A \times T^L \times \Theta^L)$ is a Bayes Coarse Correlated ε^c-Equilibrium (hereafter, ε^c-BCCE) of (G^L, S^L) if the two following properties hold.*

1. ν is consistent for (G^L, S^L) (see Definition 3).

2. ν is coarsely ε^c-obedient for (G^L, S^L); that is, for all $i \in \mathcal{I}$ and $t_i \in T_i^L$, we have

$$\sum_{a, t_{-i}, \theta} \left[u_i\big((a_i', a_{-i}), \theta\big) - u_i(a, \theta) \right] \nu\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon_i^c(a_i', t_i)$$

for all $a_i' \in A_i$.

We denote by $E^c(\varepsilon^c)$ the set of ε^c-BCCE of (G^L, S^L).

Like the ε-BCE notion, the notion of Bayes coarse correlated ε^c-equilibrium is defined through the restrictions we impose on ν. What distinguishes the two equilibrium notions is the form of incentive constraint—obedience *versus* coarse obedience. 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.

For any ε^c, the set $E^c(\varepsilon^c)$ is convex. If $\varepsilon = \varepsilon^c$, we have $E(\varepsilon) \subseteq E^c(\varepsilon^c)$. When each player has at most two actions, if $\varepsilon = \varepsilon^c$, ε-obedience coincides with ε^c-coarse obedience, and so $E(\varepsilon) = E^c(\varepsilon^c)$.

When $\varepsilon^c = 0$, we have the notion of Bayes coarse correlated equilibrium (hereafter, BCCE). The BCCE notion is new to this paper. It can be interpreted as an incomplete information version of coarse correlated equilibrium (Hannan, 1957; Moulin and Vial, 1978; Young, 2004) or as the coarse analog of the BCE notion. When Θ^L is a singleton, the notion of ε^c-BCCE reduces to that of coarse correlated ε^c-equilibrium for a complete information game.

E.3 Convergence of ε^c-External Regret Dynamics

The next theorem establishes the following results: the sequence of empirical distributions converges almost surely to $E^c(\varepsilon^c)$ if and only if the sequence of states, signals, and actions from $(G, S)^\infty$ has the ε^c-AER property almost surely. Thus, the theorem provides dynamic foundations for the static equilibrium notion of ε^c-BCCE.

Theorem 4 (Convergence of ε^c-Regret Dynamics)**.** *Fix an ε^c. The sequence of states, signals, and actions $((\theta^n, t^n, a^n))_{n \in \mathbb{N}}$ from $(G, S)^\infty$ has the ε^c-AER property 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 $E^c(\varepsilon^c)$.

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

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

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

$$V_i^{\text{ext}}(a_i', t_i, N) - U_i(t_i, N) = \frac{1}{N} \sum_{n=1}^N \left[u_i((a_i', a_{-i}^n), \theta^n)) - u_i((a_i^n, a_{-i}^n), \theta^n)) \right] \mathbb{1}_{\{t_i\}}(t_i^n)$$

$$= \frac{1}{N} \sum_\theta \sum_{n=1}^N \left[u_i((a_i', a_{-i}^n), \theta^n)) - u_i(a^n, \theta^n) \right] \mathbb{1}_{\{t_i\}}(t_i^n) \mathbb{1}_{\{\theta\}}(\theta^n) \qquad (25)$$

$$= \sum_{a, t_{-i}, \theta} \left[u_i((a_i', a_{-i}), \theta) - u_i(a, \theta) \right] Z^N(a, (t_i, t_{-i}), \theta).$$

Now pick any $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i' \in A_i$. As $\limsup_{N \to \infty} R_i^{\text{ext}}(a_i', t_i, N) \le \varepsilon_i^c(a_i', t_i)$ a.s., by definition of $R_i^{\text{ext}}(a_i', t_i, N)$, we also have

$$\limsup_{N \to \infty} \left[V_i^{\text{ext}}(a_i', t_i, N) - U_i(t_i, N) \right] \le \varepsilon_i^c(a_i', t_i) \quad \text{a.s.} \qquad (26)$$

Then, by (25) and (26),

$$\limsup_{N \to \infty} \sum_{a, t_{-i}, \theta} \left[u_i((a_i', a_{-i}), \theta) - u_i(a, \theta) \right] Z^N(a, (t_i, t_{-i}), \theta) \le \varepsilon_i^c(a_i', t_i) \quad \text{a.s.} \qquad (27)$$

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

$$\lim_{l \to \infty} \sum_{a, t_{-i}, \theta} \left[u_i((a_i', a_{-i}), \theta) - u_i(a, \theta) \right] Z_l^N(a, (t_i, t_{-i}), \theta)$$

$$= \sum_{a, t_{-i}, \theta} \lim_{l \to \infty} \left[u_i((a_i', a_{-i}), \theta) - u_i(a, \theta) \right] Z_l^N(a, (t_i, t_{-i}), \theta) \qquad (28)$$

$$= \sum_{a, t_{-i}, \theta} \left[u_i((a_i', a_{-i}), \theta) - u_i(a, \theta) \right] \nu(a, (t_i, t_{-i}), \theta).$$

Together, (27) and (28) give

$$\sum_{a, t_{-i}, \theta} \left[u_i((a_i', a_{-i}), \theta) - u_i(a, \theta) \right] \nu(a, (t_i, t_{-i}), \theta) \le \varepsilon_i^c(a_i', t_i) \quad \text{a.s.} \qquad (29)$$

As $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i' \in A_i$ were arbitrarily chosen, we conclude from (29) that ν is almost surely coarsely ε^c-obedient for (G^L, S^L).

[⟸]   Now suppose $(Z^N)_{N \in \mathbb{N}}$ converges almost surely to $E^c(\boldsymbol{\varepsilon^c})$ for some $\boldsymbol{\varepsilon^c}$. Pick any $i \in \mathcal{I}$, $t_i \in T_i^L$, and $a_i' \in A_i$. By coarse $\boldsymbol{\varepsilon^c}$-obedience,

$$\limsup_{N \to \infty} \sum_{a, t_{-i}, \theta} \big[u_i\big((a_i', a_{-i}), \theta\big) - u_i(a, \theta) \big] Z^N\big(a, (t_i, t_{-i}), \theta\big) \leq \varepsilon_i^c(a_i', t_i) \quad \text{a.s.} \qquad (30)$$

By (25) and (30),

$$\limsup_{N \to \infty} V_i^{\text{ext}}(a_i', t_i, N) - U_i(t_i, N) \leq \varepsilon_i^c(a_i', t_i) \quad \text{a.s.,}$$

which implies

$$\limsup_{N \to \infty} R_i^{\text{ext}}(a_i', t_i, N) \leq \varepsilon_i^c(a_i', t_i) \quad \text{a.s.}$$

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

References

ACKERBERG, D. A. AND G. GOWRISANKARAN (2006): "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *The RAND Journal of Economics*, 37, 738–761.

AGUIRREGABIRIA, V. AND J. JEON (2020): "Firms' Beliefs and Learning: Models, Identification, and Empirical Evidence," *Review of Industrial Organization*, 56, 203–235.

AGUIRREGABIRIA, V. AND A. MAGESAN (2020): "Identification and Estimation of Dynamic Discrete Games When Players' Beliefs are Not in Equilibrium," *The Review of Economic Studies*, 87, 582–625.

AGUIRREGABIRIA, V. AND P. MIRA (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75, 1–53.

ARORA, S., E. HAZAN, AND S. KALE (2012): "The Multiplicative Weights Update Method: A Meta-Algorithm and Applications," *Theory of Computing*, 8, 121–164.

ASSAD, S., E. CALVANO, G. CALZOLARI, R. CLARK, V. DENICOLÒ, D. ERSHOV, J. JOHNSON, S. PASTORELLO, A. RHODES, L. XU, AND M. WILDENBEEST (2021): "Autonomous Algorithmic Collusion: Economic Research and Policy Implications," *Oxford Review of Economic Policy*, 37, 459–478.

AUMANN, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

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

AUMANN, R. J., M. MASCHLER, AND R. E. STEARNS (1995): *Repeated Games with Incomplete Information*, MIT Press.

BAJARI, P., C. L. BENKARD, AND J. LEVIN (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75, 1331–1370.

BERESTEANU, A., I. MOLCHANOV, AND F. MOLINARI (2011): "Sharp Identification Regions in Models with Convex Moment Predictions," *Econometrica*, 79, 1785–1821.

BERGEMANN, D., B. A. BROOKS, AND S. MORRIS (2021): "Counterfactuals with Latent Information," *The American Economic Review (forthcoming)*.

BERGEMANN, D. AND S. MORRIS (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

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

——— (2019): "Information Design: A Unified Perspective," *Journal of Economic Literature*, 57, 44–95.

BERGEMANN, D. AND K. H. SCHLAG (2008): "Pricing without Priors," *Journal of the European Economic Association*, 6, 560–569.

——— (2013): "Robust Monopoly Pricing," *Journal of the European Economic Association*, 146, 2527–2543.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica: Journal of the Econometric Society*, 841–890.

BEVIÁ, C. AND L. CORCHÓN (2019): "Contests with Dominant Strategies," *Economic Theory*, 1–19.

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

——— (1956b): "Controlled Random Walks," *Proceedings of the International Congress of Mathematicians*, 3, 336–338.

BLUM, A., M. HAJIAGHAYI, K. LIGETT, AND A. ROTH (2008): "Regret Minimization and the Price of Total Anarchy," in *Proceedings of the fortieth annual ACM symposium on Theory of computing*, 373–382.

BLUM, A. AND Y. MANSOUR (2007): "From External to Internal Regret." *Journal of Machine Learning Research*, 8.

BÖRGERS, T. AND R. 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.

BRAVERMAN, M., J. MAO, J. SCHNEIDER, AND M. WEINBERG (2018): "Selling to a No-Regret buyer," in *Proceedings of the 2018 ACM Conference on Economics and Computation*, 523–538.

BRESNAHAN, T. F. AND P. C. REISS (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 977–1009.

BUBECK, S. AND N. CESA-BIANCHI (2012): "Regret Analysis of Stochastic and Non-stochastic Multi-Armed Bandit Problems," *Foundations and Trends in Machine Learning*, 5, 1–122.

BUSH, R. R. AND F. MOSTELLER (1955): "Stochastic Models for Learning." .

CALDENTEY, R., Y. LIU, AND I. LOBEL (2017): "Intertemporal Pricing Under Minimax Regret," *Operations Research*, 65, 104–129.

CAMARA, M. K., J. D. HARTLINE, AND A. JOHNSEN (2020): "Mechanisms for a No-Regret Agent: Beyond the Common Prior," in *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 259–270.

CARAGIANNIS, I., C. KAKLAMANIS, P. KANELLOPOULOS, M. KYROPOULOU, B. LUCIER, R. PAES LEME, AND E. TARDOS (2015): "Bounding the Inefficieny of Outcomes in Generalized Second Price Auctions," *Journal of Economic Theory*, 157, 343–388.

CESA-BIANCHI, N. AND G. LUGOSI (2006): *Prediction, Learning, and Games*, Cambridge University Press.

CESA-BIANCHI, N., Y. MANSOUR, AND G. STOLTZ (2007): "Improved Second-Order Bounds for Prediction with Expert Advice," *Machine Learning*, 66, 321–352.

CHASSANG, S. (2013): "Calibrated Incentive Contracts," *Econometrica*, 81, 1935–1971.

CHASSANG, S. AND S. KAPON (2021): "Prior-Free Dynamic Allocation Under Limited Liability," *Theoretical Economics (forthcoming)*.

CHEN, L., A. MISLOVE, AND C. WILSON (2016): "An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace," in *Proceedings of the 25th International Conference on World Wide Web*, 1339–1349.

CILIBERTO, F. AND E. TAMER (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77, 1791–1828.

DENG, Y., J. SCHNEIDER, AND B. SIVAN (2019): "Strategizing Against No-Regret Learners," *arXiv preprint arXiv:1909.13861*.

DORASZELSKI, U., G. LEWIS, AND A. PAKES (2018): "Just Starting Out: Learning and Equilibrium in a New Market," *The American Economic Review*, 108, 565–615.

ELLISON, S. F., C. SNYDER, AND H. ZHANG (2018): "Costs of Managerial Attention and Activity as a Source of Sticky Prices: Structural Estimates from an Online Market," *National Bureau of Economic Research*.

EPSTEIN, L., H. KAIDO, AND K. SEO (2016): "Robust Confidence Regions for Incomplete Models," *Econometrica*, 84, 1799–1838.

EREV, I. AND A. E. ROTH (1998): "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategies," *The American Economic Review*, 88, 848–881.

ERICSON, R. AND A. PAKES (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62, 53–82.

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

FOSTER, D. P. AND R. V. VOHRA (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, Y. AND R. E. SCHAPIRE (1997): "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, 55, 119–139.

——— (1999): "Adaptive Game Playing Using Multiplicative Weights," *Games and Economic Behavior*, 29, 79–103.

FUDENBERG, D. AND D. 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.

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

GALICHON, A. AND M. HENRY (2011): "Set Identification in Models with Multiple Equilibria," *The Review of Economic Studies*, 78, 1264–1298.

GRADWOHL, R. AND E. KALAI (2021): "Large Games: Robustness and Stability," *Annual Review of Economics*, 13.

GREENWALD, A. AND A. JAFARI (2003): "A General Class of No-Regret Learning Algorithms and Game-Theoretic Equilibria," in *Learning Theory and Kernel Machines*, Springer, 2–12.

GUALDANI, C. AND S. SINHA (2020): "Identification and Inference in Discrete Choice Models with Imperfect Information," *Working Paper*.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525–574.

GUO, Y. AND E. SHMAYA (2019): "Robust Monopoly Regulation," *Working Paper*.

——— (2021): "Project Choice from a Verifiable Proposal," *Working Paper*.

HAILE, P., H. HONG, AND M. SHUM (2003): "Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions," *National Bureau of Economic Research*.

HAILE, P. A. AND E. TAMER (2003): "Inference with an incomplete model of English auctions," *Journal of Political Economy*, 111, 1–51.

HANNAN, J. (1957): "Approximation to Bayes Risk in Repeated Play," in *Contributions to the Theory of Games, Vol. 3*, ed. by M. Dresher, A. Tucker, and P. Wolfe, Princeton University Press.

HART, S. (2005): "Adaptive Heuristics," *Econometrica*, 73, 1401–1430.

HART, S. AND A. MAS-COLELL (2000): "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, 68, 1127–1150.

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

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

HARTLINE, J., V. SYRGKANIS, AND E. TARDOS (2015): "No-Regret Learning in Bayesian Games," *NIPS 2015: Advances in Neural Information Processing Systems*, 3061–3069.

HAZAN, E. (2021): *Introduction to Online Convex Optimization*, MIT Press (forthcoming).

HIRANO, K. AND J. R. PORTER (2009): "Asymptotics for Statistical Treatment Rules," *Econometrica*, 77, 1683–1701.

HURWICZ, L. AND L. SHAPIRO (1978): "Incentive Structures Maximizing Residual Gain under Incomplete Information," *The Bell Journal of Economics*, 9, 180–191.

JENSEN, M. K. (2018): "Aggregative Games," in *Handbook of Game Theory and Industrial Organization, Volume I*, Edward Elgar Publishing.

KALAI, A. AND S. VEMPALA (2005): "Efficient Algorithms for Online Decision Problems," *Journal of Computer and System Sciences*, 71, 291–307.

KAMENICA, E. (2019): "Bayesian Persuasion and In formation Design," *Annual Review of Economics*, 11, 249–272.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian Persuasion," *The American Economic Review*, 101, 2590–2615.

LATTIMORE, T. AND C. SZEPESVÁRI (2020): *Bandit Algorithms*, Cambridge University Press.

LEHRER, E. (2003): "A Wide Range No-Regret Theorem," *Games and Economic Behavior*, 42, 101–115.

LEHRER, E. AND E. SOLAN (2009): "Approachability with Bounded Memory," *Games and Economic Behavior*, 66, 995–1004.

LITTLESTONE, N. AND M. K. WARMUTH (1994): "The Weighted Majority Algorithm," *Information and Computation*, 108, 212–261.

MAGNOLFI, L. AND C. RONCORONI (2021): "Estimation of Discrete Games with Weak Assumptions on Information," *Working Paper*.

MANSKI, C. F. (2004): "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 72, 1221–1246.

——— (2007): "Minimax-Regret Treatment Choice with Missing Outcome Data," *Journal of Econometrics*, 139, 105–115.

——— (2021): *Social Choice with Partial Knowledge of Treatment Response*, Princeton University Press.

MOLINARI, F. (2020): "Microeconometrics with Partial Identification," *Handbook of Econometrics*, 7, 355–486.

MOULIN, H. AND J.-P. 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, D., V. SYRGKANIS, AND E. TARDOS (2015): "Econometrics for Learning Agents," in *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, 1–18.

NISAM, N., T. ROUGHGARDEN, E. TARDOS, AND V. V. VAZIRANI (2007): *Algorithmic Game Theory*, Cambridge University Press.

NISAN, N. AND G. NOTI (2017a): "An Experimental Evaluation of Regret-Based Econometrics," in *Proceedings of the 26th International Conference on World Wide Web*, 73–81.

——— (2017b): "A "Quantal Regret" Method for Structural Econometrics in Repeated Games," *arXiv preprint arXiv:1702.04254*.

NOTI, G. AND V. SYRGKANIS (2021): "Bid Prediction in Repeated Auctions with Learning," in *Proceedings of the Web Conference 2021*, 3953–3964.

PAKES, A., M. OSTROVSKY, AND S. BERRY (2007): "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit examples)," *the RAND Journal of Economics*, 38, 373–399.

PESENDORFER, M. AND P. SCHMIDT-DENGLER (2008): "Asymptotic Least Squares Estimators for Dynamic Games," *The Review of Economic Studies*, 75, 901–928.

RENOU, L. AND K. H. SCHLAG (2011): "Implementation in Minimax Regret Equilibrium," *Games and Economic Behavior*, 71, 527–533.

ROTH, A. E. AND I. 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, T. (2009): "Intrinsic Robustness of the Price of Anarchy," in *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, 513–522.

——— (2016): *Twenty Lectures on Algorithmic Game Theory*, Cambridge University Press.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 999–1033.

SAVAGE, L. J. (1951): "The Theory of Statistical Decision," *Journal of the American Statistical Association*, 46, 55–67.

——— (1972): *The Foundations of Statistics*, Courier Corporation.

SCHLAG, K. H. (2007): "Eleven: Designing Randomized Experiments under Minimax Regret," *Working Paper*.

SCHLAG, K. H. AND A. ZAPECHELNYUK (2017): "Dynamic Benchmark Targeting," *Journal of Economic Theory*, 169, 145–169.

SHALEV-SHWARTZ, S. (2011): "Online Learning and Online Convex Optimization," *Foundations and Trends in Machine Learning*, 4, 107–194.

SLIVKINS, A. (2019): "Introduction to Multi-Armed Bandits," *Foundations and Trends in Machine Learning*, 1, 1–286.

STOYE, J. (2007): "Minimax Regret Treatment Choice with Incomplete Data and Many Treatments," *Econometric Theory*, 23, 190–199.

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

SYRGKANIS, V., E. TAMER, AND J. ZIANI (2021): "Inference on Auctions with Weak Assumptions on Information," *Working Paper*.

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

VOHRA, R. AND M. WELLMAN (2007): "Special Issue on Foundations of Multi-Agent Learning," *Artificial Intelligence*, 171, 363–452.

WEINTRAUB, G. Y., C. L. BENKARD, AND B. VAN ROY (2008): "Markov Perfect Industry Dynamics with Many Firms," *Econometrica*, 76, 1375–1411.

WILDENBEEST, M. R. (2011): "An Empirical Model of Search with Vertically Differentiated Products," *The RAND Journal of Economics*, 42, 729–757.

YOUNG, H. P. (2004): *Strategic Learning and Its Limits*, Oxford University Press.