An Estimable Demand System for a Large Auction Platform Market[*]

Matthew Backus

University of Michigan

Gregory Lewis

Harvard University and NBER

October 13, 2010

Abstract

Economists have developed a range of empirically tractable demand systems for fixed price markets. But auction mechanisms also play an important part in allocating goods, and yet existing empirical auction techniques treat each auction in isolation, obscuring market interactions. Here we provide a framework for estimating a demand system in a large auction platform market with a dynamic population of buyers, heterogeneous objects and unit demand. We construct a model of repeated second-price auctions in which bidders have multidimensional private valuations, developing an equilibrium concept under which strategies reflect option values. We prove existence of this equilibrium and characterize the ergodic distribution of types. Having developed a demand system, we show that it is non-parametrically identified from panel data. Relatively simple nonparametric and semiparametric estimation procedures are proposed and tested by Monte Carlo simulation. Our analysis highlights the importance of both dynamic bidding strategies and panel data sample selection issues when analyzing these markets.

1 Introduction

Most goods and services are sold at fixed prices. Yet auctions are used as the allocation mechanism in a wide variety of contexts, including procurement and the granting of oil drilling and spectrum rights. Online auction platforms, in particular, have grown substantially in recent years. In retail eBay alone has revenues of \$36 billion from its auctions business in 2007[1], while Google realized \$21 billion in revenue from its online advertising platform in 2008[2]. More specialized auction sites such as DoveBid and IronPlanet have sold billions of dollars of used aviation and construction equipment respectively.

Given their importance in the modern economy, one would like to be able to estimate demand in these platform markets. This would allow us to answer questions of broad economic interest, such as how much welfare has been generated by these platforms; as well as narrower strategic questions, such as how a firm with a fixed inventory should set reserves and time sales to dynamically maximize its revenue. Demand estimation is often also a necessary first step for the evaluation of anti-trust issues, such as the potential impact on the search-keyword advertising market of a merger between Microsoft and Yahoo.

At first glance, auctions data is an extremely rich of information about demand. In auctions we actually observe a continuous bid which directly reflects willingness to pay, relative to the discrete choice we typically see in fixed price markets. Moreover, for any buyer we generally observe all the auctions that they bid in, which provides valuable information about which items they view as close substitutes. This is informative for demand, much in the same way that "second-choice" data is useful in Berry, Levinsohn, and Pakes (2004). As the choice sets available to buyers vary we may observe differing participation, which is helpful for identifying substitution patterns.

Yet the strengths of auction market data also pose some difficulties. As Hendricks and Porter (2007) note in their survey article, participants in auction markets are playing a complex dynamic game, where they must continuously adapt to the changing set of available auctions, and learn about rival's valuations. Most of the existing tools of structural auction econometrics are focused on independent auctions of homogenous objects, which limits their direct applicability to auction demand estimation, where bidders repeatedly interact across auctions, and substitution across products is important.

---

[1]Source: eBay Annual Report for 2007
[2]Source: Google Annual Report for 2008

In this paper, we develop an estimable demand system for an auction platform market with a large number of relatively short-lived, yet persistent, buyers. The first part of the paper outlines an intuitive and empirically tractable equilibrium concept — competitive Markov equilibrium — and characterizes the long-run distribution of bidders and their strategies. Next, we show that demand is non-parametrically identified from panel data. In the last part, we develop a nonparametric and a semiparametric estimation approach for backing out demand from bid data. We also show how to estimate a characteristic-based model of demand. These approaches are tested by Monte Carlo simulation and found to work well in moderately sized samples.

To begin, the theory section introduces a stylized model of an auction market, in which each period a good is sold by second-price sealed bid auction. There are a finite number of different goods that can be sold, and supply of these goods is exogenous. Bidders have multidimensional private valuations over the different goods. These valuations may be correlated. They have unit demand, and upon entry participate in every auction until they either win or randomly exit. The stage game is played each period over an infinite horizon.

The environment is complicated, because for any bidder the set of rival types is unknown, and Bayes-Nash equilibrium would imply that everyone simultaneously solves a filtration problem, using the observed history — possibly private and arbitrarily long — to infer the distribution of types. We simplify by developing an equilibrium concept in which bidders condition only on a coarser publicly observable state vector in forming beliefs about rival types, and that they take the state evolution as exogenous. We call this notion competitive Markov equilibrium, and argue that it is appropriate for large anonymous markets.

Then, if bidders have unit demand, we have a simple way to deal with dynamic concerns. Participation has an option value: the expected surplus from future auctions conditional on today's state. This option value is struck when a bidder wins an auction, so she shades her bids accordingly. The challenge of estimating private values is then estimating the long-run option value. We show that this is non-parametrically identified from observing both individual bidder time series and the full bid distribution across states.

The final part of the paper is concerned with estimation. We offer two approaches. One follows the nonparametric identification logic directly, showing that by looking at the time series of bidders who are observed bidding in every state we can back out their individual valuations. As is common with panel data, there are selection concerns. These bidders are

a selected sample, and to get the true distribution of valuations, it is necessary to re-weight the estimated density. We show how to do this.

A disadvantage of the nonparametric method is that it is data intensive, since if the state space is large, the set of bidders who bid in every state may be very small. By making a parametric assumption on the type distribution, we can make use of the remaining data. We first show how to estimate the bid function for any type directly from the data, and then argue that this allows us to simulate moments given any parameter vector. We can thus apply simulated GMM to consistently estimate the true parameter vector.

The paper is related to various strands of literature. Jofre-Benet and Pesendorfer (2003) was the first paper to attack estimation in a dynamic auction game, though in a world where private information was transient. Subsequent to this, a number of papers have looked at dynamics on the eBay platform specifically. Budish (2008) examines the optimality of eBay's market design with respect to the sequencing of sales and information revelation. In Said (2009), the author investigates efficiency and revenue maximization in a similar setup through the lens of dynamic mechanism design. Zeithammer (2006) developed a model with forward-looking bidders, and showed both theoretically and empirically that bidders shade down current bids in response to the presence of upcoming auctions of similar objects. Ingster (2009) develops a dynamic model of auctions of identical objects, and provides equilibrium characterization and identification results. Sailer (2006) estimates participation costs for bidders facing an infinite sequence of identical auctions. Relative to this literature, our main contribution is the focus on sequential auctions of heterogeneous objects, where bidders have multidimensional persistent private valuations. In short, we are focused on developing a demand system. A different approach has been taken in Adams (2009), who looks at the problem of nonparametric identification when auctions are completely simultaneous.

A second related literature is on alternative equilibrium notions for dynamic games. In coarsening the set of information bidders condition on, we are following the path of Krusell and Smith (1998). Weintraub, Benkard, and Van Roy (2008) develops the related notion of oblivious equilibrium. Fershtman and Pakes (2009) also emphasizes the importance of a finite state space. Finally, we build on the literature for estimating demand systems in durable goods markets (e.g. Berry (1994), Berry, Levinsohn, and Pakes (1995), Gowrisankaran and Rysman (2009)).

The next section introduces the theoretical framework, while section 3 proves non-parametric

identification. Section 4 describes our two different estimation approaches, while section 5 gives Monte Carlo simulations for those estimators. Section 6 concludes.

2   Model

Our aim in this section is to create an abstract model of a large auction market, and analyze it. The space of such models is vast, and we narrow in a number of ways. We consider a market in which similar products — such as iPods and Zunes — are sold by second-price sealed bid auctions. These auctions are held in discrete time, with one good auctioned per period over an infinite horizon. Since our focus is on demand, we assume for simplicity that supply is random and exogenous. Bidders are persistent with unit demand, and enter the market with private (possibly correlated) valuations for each of the objects. Winning bidders immediately exit, while losing bidders exit randomly. We aim at characterizing the long-run behavior of this dynamic system.

We have chosen this set of assumptions to match some features of the environment on eBay, whose platform design dominates online auctions. In any eBay category, there are many different products sold by auction to a large number of anonymous buyers.[3] Although these auctions typically last for many days, and thus overlap — so that at any given point in time there are many auctions occurring simultaneously — they finish at different ending times, in sequence. As Bajari and Hortacsu (2004) and Hendricks and Porter (2007) have noted, this timing, combined with the way the proxy bidding system works, imply that eBay is well approximated as sequence of second-price sealed bid auctions. Yet our intent is not to model eBay *per se* — and indeed we ignore some important features of the eBay environment — but rather to develop a reasonably motivated and rich abstract model and see what we can learn from the exercise.

2.1   Environment

We formalize the above description of the environment in what follows:

Bidders and Payoffs: Bidders have unit demand for a good in the set $\mathcal{J}$, where $|\mathcal{J}| = J$. Their demand is summarized by a privately known vector of valuations $x = (x_1, x_2 \cdots x_J)$,

---

[3]eBay hides the identity of the bidders by replacing parts of the username with asterixes.

their type. They are risk neutral, and receive a payoff of $x_j - p$ for buying a single good j at price p, and zero otherwise. Bidders are impatient, with a common discount rate δ.

Market: Time is discrete with infinite horizon, $t = 1, 2 \cdots$. In each period t, the following stage game is played. First, a sealed-bid second price auction is held for the current object j_t, in which all bidders present in the market may participate. Then entry and exit of bidders takes place, and suppliers post new objects. These are described in more detail below.

Auctions: At time t, object $j_t \in \mathcal{J}$ is auctioned. Bidders have the choice between not participating (action ϕ), and submitting a bid. The action space is thus $\mathcal{A} = \phi \cup \mathbb{R}^+$, where $\mathcal{A}$ is totally ordered under the usual ordering $<$ with $\phi < 0$. The highest bidder wins the auction and pays the second highest bid, or zero if his is the only bid submitted. Ties are broken randomly. If no-one participates, the item is not sold.

Entry and Exit: At the end of every period, the winner is assumed to exit with certainty.[4] Losers exogenously exit the market with probability $\rho \in (0, 1)$, receiving a payoff normalized to zero on exit. Simultaneously, E_t new bidders enter, where E_t is random with a distribution that depends on the total number of buyers in the previous period N_{t-1}. We assume that $E_t | N_{t-1}$ has strictly positive support on the finite set of integers $\{0, 1, 2 \cdots \overline{N} - N_{t-1}\}$. This ensures that the size of the market does not explode.[5] Each entrant draws their valuation vector x identically and independently from a distribution F with associated strictly positive density f, and support a compact set $\mathcal{X} = [0, \bar{x}]^J$.

Supply: Supply is essentially the rate at which different products appear on the auction market. At the end of period t, suppliers list a new object to be auctioned in period $t + 1 + k^f$, where k^f is the lead-time bidders have in observing future supply. The object to be auctioned is randomly chosen according to a multinomial distribution over the set of products $\mathcal{J}$.[6]

Information Sets and Bidding Strategies: New entrants are assumed to be able to view the history of the game for the last k^h periods.[7] Incumbent bidders may have observed more: at time t, a bidder i who entered the market time t_i can observe a "window" of past actions and current and upcoming auctions, from $t_i - k^h$ to $t + k^f$. The cases $k^h = 0$

---

[4]Since they have unit demand, they are indifferent about exiting in any period following a win. But even an $\varepsilon > 0$ participation cost would make exit optimal.

[5]With endogenous entry, one would expect a condition like this to hold: entry falls as the number of participants in the market increases, and so surplus falls. With exogenous entry, we must impose it.

[6]It is easy to extend the model to allow for fluctuating total supply by letting the the object be multinomial over $\mathcal{J} \cup \varnothing$, where $\varnothing$ is the event that nothing is listed.

[7]This mimics eBay, where the history of auctions held in the past 14+ days is public, though anonymized.

and $k^f = 0$ correspond to no observable public history and no knowledge of future supply, respectively. So a generic information set consists of a valuation x, the history h_{it} and the list of current and upcoming objects $\mathbf{j}_t = (j_t \cdots j_{t+k^f})$. A (pure) bid strategy is a mapping $\beta_i : \mathcal{X} \times \mathcal{H}^{it} \times \mathcal{J}^{k^f+1} \to \mathcal{A}$ from any information set to the action set.

2.2 Analysis

We analyze this environment in three parts. First, we motivate and define a new equilibrium concept called a competitive Markov equilibrium (CME) that we think is appropriate for long-run analysis of large markets. Under this equilibrium, we show that strategies take a simple and intuitive form: bidders bid their valuation for the good under auction, less their continuation value. Second, we characterize the long-run properties of the dynamic system for arbitrary strategies, showing that a stationary distribution over types exists. Finally, we combine these two pieces — equilibrium characterization and long-run dynamics — to show existence of a CME. We also argue that as the market becomes large, while holding the ratio of buyers to sellers fixed, the CME tends towards the anonymous equilibrium of the continuum game.

To motivate the concept, consider the decision problem of a bidder in this environment. He knows the recent history of the market, current supply and his own valuation. What he doesn't know is who else is in the market (his rivals, their valuations and their history), nor how they will bid. In addition, he must form expectations about future demand conditions, and should in principle worry about "leakage"; his bid today might reveal valuable information to future rivals. Addressing these issues with standard equilibrium concepts is problematic. Though Milgrom and Weber (2000) were able to provide an elegant equilibrium characterization of sequential auctions under certain information structures, their approach is essentially static and does not extend to the infinite horizon case.[8] In this environment, forming rational expectations about the play of opponents requires some notion of the long-run stationary distribution of types, but without placing some structure on the admissible strategies, the state space may grow without bound.

Technical objections aside, expecting this behavior from bidders in large markets seems unrealistic. Bidders on eBay don't worry about leakage, because they don't expect their

---

[8]They consider the problem of auctioning k identical objects to n bidders, where $n > k$, and the bidders all enter in the first period.

individual bids to be tracked by rivals. Rivals don't track them because the market is large, turnover is rapid, and there is little to be gained from the information. Basically, bidders expect that with this many auctions, the probability of meeting the same opponents in the future is low.

To capture this intuition, we assume that bidders believe that now — and in the future — they are to compete with a random draw from the long-run population of types. These beliefs may be conditioned on some simple and publicly observable "state" variables, such as recent prices, which may be informative as to whether demand is currently high or low. Bidders have rational expectations, and their beliefs will be correct in equilibrium. Importantly, they will take the state as given and its evolution as exogenous, though in fact since the market is finite, their actions may have some impact on the state transitions.

This kind of assumption was introduced in the macroeconomics literature by Krusell and Smith (1998) as a behavioral assumption, arguing that agents will make inferences based on simple functionals of all the information available in the environment, at least when doing so loses little information. The special case where agents ignore all current information and the state is constant or exogenously given has a complete information counterpart in the oblivious equilibrium concept of Weintraub, Benkard, and Van Roy (2008). We also provide a "large-market" justification of the assumption in the discussion below.

Formally, for any public history of actions $h_t = (\mathbf{a}_{t-1} \cdots \mathbf{a}_{t-k^h})$, let h_t^{anon} be the anonymized history (i.e. where identifiers on bid identity are removed). The associated space of anonymized histories is stationary, and denoted as $\mathcal{H}^{anon} \equiv \mathcal{A}^{k^h \times \overline{N}}$. Bidders also know the supply $\mathbf{j}_t \in \mathcal{J}^{k^f+1}$, implying the space of all anonymized public information is $\mathcal{H}^{anon} \times \mathcal{J}^{k^f+1}$. Define a *coarsening function T* as a (Borel) measurable function that finitely partitions the space into "states" $s \in \mathcal{S}$, where $|\mathcal{S}| = S$. We require that T partitions different current objects into different states, so that minimally the agent conditions on the object under auction — in math, $j_t \neq j_t' \Rightarrow T(h_t^{anon}, \mathbf{j_t}) \neq T(h_t^{anon'}, \mathbf{j_t}')$.

This coarsening T defines the state space, and therefore what bidders pay attention to. Bidders must also have a model of state transitions. Define Q as the transition matrix between states, with typical element $Q_{ij} = \mathbb{P}(s' = j | s = i)$, for s' the state tomorrow. Our notion of equilibrium requires that bidders correctly understand the distribution of competing types and the state transitions, and optimize against this:

Definition 1 (Competitive Markov Equilibrium). *A (symmetric) competitive Markov*

7

equilibrium (CME) with respect to a coarsening function T consists of:

    (i) *Correct beliefs about the ergodic distribution of opposing types conditional on any $s \in \mathcal{S}$; and about the state transition matrix Q*

    (ii) *Symmetric Markovian strategies $\beta(x, s)$ that maximize expected payoffs given beliefs.*

Let us unpack this a bit. The CME is extremely similar to a Bayes-Nash equilibrium, requiring that strategies are optimal given beliefs; and that beliefs are consistent with equilibrium play. The key differences are that here bidders condition only on the state s in forming beliefs, even when this is coarser than the public information available to them; and that they do not account for how their actions may influence the state transitions Q. This is the is the "competitive" part of the name, as it corresponds to the case in perfect competition where firms do not recognize that their joint production decisions determine the price. Here, bidders behave as though they were small, and do not endogenize the impact of their own actions on the future states. Another important difference is that this is a long-run concept: bidders believe they face draws from the ergodic distribution of types, which is a sensible belief only if the market does indeed converge to a long-run distribution and has been in operation for a while. This formulation avoids the issue of a prior on the initial type draw.

It turns out that under these assumptions, the equilibrium bidding strategy $\beta(x, s)$ has a intuitive and simple form. Temporarily putting aside questions of stationarity and existence, fix a CME. Bidders have well-defined beliefs about the distribution of types in any state, and given the equilibrium bid strategies, can also work out the distribution of highest opposing bids (i.e. the bid they need to beat to win). They also have rational expectations about state transitions. This allows us to define an (ex-post) value function for a type x in state s:

$$v(x, s) = \max_{b \in \mathcal{A}} G_1(b|s)\left(x_t - E[B^1|B^1 < b, s]\right) + (1 - G_1(b|s))\delta(1 - \rho)\sum_{s'=1}^{S} v(x, s')Q_{ss'} \quad (1)$$

where $G_1(\cdot|s)$ is the distribution of the highest opposing bid today given the state; x_t is the bidder's valuation of the object currently under auction, and Q is the equilibrium transition matrix. The first term in the value function is the probability of winning — the probability that the highest opposing bid is lower — times the surplus conditional on winning, equal to current valuation less expected payment. The second term is the probability of losing times the continuation value in that event. Non-participation (ϕ) implies certain loss, so

$G_1(\phi|s) = 0 \; \forall s$.

Now let $\widetilde{v}(x,s) = \delta(1-\rho)\sum_{s'=1}^{S} v(x,s')Q_{ss'}$ denote the discounted ex-ante value function (i.e. before exit and state transitions are determined). Maximizing the value function above, we get the optimal strategies:

Lemma 1 (Equilibrium Strategies). *In a symmetric CME, bidders bid their valuation less their ex-ante continuation value if it is positive; otherwise they don't participate:*

$$\beta(x,s) = \begin{cases} x_t - \widetilde{v}(x,s) & , \; x_t - \widetilde{v}(x,s) \geq 0 \\ \phi & , \; otherwise \end{cases} \tag{2}$$

Think about this process as a single auction, where the winner gets the object, and the losers are awarded a prize with value equal to the continuation value. Re-normalizing the prizes, it's like a standard second-price auction where the winner gets the object less continuation value, and losers get nothing. Then the weakly dominant strategy is to bid the value of the prize, which is just the value of the object less the continuation value.[9] In the case where the "prize" has negative value, there is no reason to participate.

The intuitive appeal of this characterization is that it reduces the bidder problem to forming some expectation of their continuation value, which should be informed by the state of the market (recent history and future supply). Under a CME, we require these expectations to be correct in the sense of matching the long-run behavior of the system.

Our next step is to analyze these dynamics. Since bidders enter and exit every period over an infinite horizon, if we kept track of specific identities the state-space would grow without bound. So for the long-run analysis, we ignore identity, and keep track of an anonymous $\overline{N}$-vector $\mathbf{x}_t$ of types currently in the market.[10]

We let the "true state" of the market $\omega_t \in \Omega \equiv \mathcal{X}^{\overline{N}} \times \mathcal{S}$ be defined by the anonymized vector of types $\mathbf{x}_t$ and the state s_t. For any symmetric strategy β, the true state evolves as a first

---

[9]As Budish (2008), Said (2009) and Zeithammer (2009) all note, this is not true in the full Bayes-Nash equilibrium, due to a winner's curse effect: on winning, a bidder learns that the remaining types had lower valuations, and therefore regrets winning now rather than later at a lower price. This effect is swept away in a CME, since bidders take the state evolution as exogenous.

[10]We use 0 as a placeholder when there are fewer than $\overline{N}$ bidders in the market. Type transitions occur first by removing exiting bidders and replacing them with the placeholder, and then adding entrants sequentially, starting from the first open placeholder. For example, suppose we have $\overline{N} = 3$, and there are two bidders with (unidimensional valuations) 1 and 2 respectively. Then we have $\mathbf{x}_t = (1,2,0)$; and if at the end of the period bidder 1 exits and two new bidders with values 3 and 4 respectively enter, we will have $\mathbf{x}_{t+1} = (3,2,4)$.

order Markov process, with the type transitions governed by the entry and exit rules. The state transitions are determined by the exogenous supply and the actions taken by the types in accordance with the strategies. Denote by $\mathcal{F}$ the Borel σ-field over Ω.

Lemma 2 (**Ergodic Distribution of True States**). *For any strategy β, there is a unique invariant measure μ_β on the measurable space $(\Omega, \mathcal{F})$, strongly converged to at uniform geometric rate from any initial measure μ_0. The conditional ergodic distribution of $\mathbf{x}_{-i}$ given s exists and is well-defined for any $s \in \mathcal{S}$.*

This says that the market "settles down" to a steady-state, regardless of the initial conditions, with a unique stationary distribution of types in each state. The intuition for this is that entry and supply are exogenous, and only the exit of winning bidders is endogenously determined. This has a limited influence on the long-run evolution of the market.[11] An implication of this is that in the long-run, given any strategies, agents have well-defined beliefs about the population of bidders they face. We can now look for a fixed point: strategies that are optimal given long-run beliefs; and beliefs that are consistent with the ergodic distributions induced by these strategies. We will call a strategy $\beta(x, s)$ monotone if in every state bids increase in the valuation of the good under auction, and decrease in the valuations of other objects. We say it is strictly monotone if the monotonicity is strict except for non-participating types.[12]

Theorem 1 (**Existence**). *For any coarsening function T, there exists a CME in continuous strictly monotone pure strategies. If there is only one product, the CME is unique.*

The proof is non-trivial.[13] One would like to exploit the characterization of the bidding strategies in (2), and show existence of a fixed point of the best response operator $\Gamma(\beta) = x_t - \widetilde{v}_\beta(x, s)$. But as is often the case in auction theory, Γ does not map continuous functions into continuous functions since if some types "play an atom" under β, then the continuation value $\widetilde{v}_\beta(x, s)$ may be discontinuous at some points.

So instead we take a different and well-trodden approach, discretizing the action space, and then applying the methodology of Reny (2008) based on contractible mappings to show a fixed point of the finite action game. Since whether bids increase or decrease in the valuations

---

[11]The formal proof uses a renewal argument, based on the idea that the mean hitting time to the "null set" where everyone has exited the market is finite.

[12]That is, $\beta(x, s)$ is strictly increasing in x_t except possibly where $\beta(x, s) = \phi$.

[13]Duffie, Geanakoplos, Mas-Colell, and McLennan (1994) establish existence for a broad class of stochastic games, but strategies may be mixed and a public coordination device is required.

depends on the state, we cannot directly apply his results, and must modify some parts of the proof.[14] A limiting argument as in Athey (2001) yields an equilibrium for the continuous action game. In the case where there is only one product, we can show that the Γ operator is a contraction mapping, implying a unique CME.

Finally, to conclude this section, we wish to argue informally that for large markets, this concept is a sensible approximation to a more standard equilibrium concept — in this case, anonymous equilibrium. One way to do this is to show that as the market becomes arbitrarily large, the set of CME and anonymous equilibria coincide. So consider the following continuum game, in the spirit of Satterthwaite and Shneyerov (2007). In each period, measure p_j of each of the goods $j \in \mathcal{J}$ is sold by second-price auction, where p_j is the multinomial supply probability. Current market participants are randomly assigned to an auction for one of the goods. A measure $E[E_t]$ of bidders enters each period, winners exit with certainty, and a measure ρ of the losers exit.

The steady-state buyer to item ratio is the same as in the finite market.[15] However, in the continuum game, the steady-state distribution of types in every period will be exactly the stationary distribution. Bidders always face a random draw of rival types, and so the anonymized history is uninformative. Similarly, since their next auction in every subsequent period is randomly drawn from the supply distribution, information about future auctions is also unhelpful. The only payoff relevant state variable is the current object under auction. This implies that in any CME or anonymous equilibrium, bidders will bid their valuation for the current object less their continuation value, where the continuation value is constant across anonymized histories or states. The set of CME and anonymous equilibria are equal.

So we learn that the key simplification that we make — even in the limit — is that bidders do not keep track of identity; for if they did, they could perhaps profitably condition on recent history if they happened to be randomly matched with other incumbent bidders. Our intuition for why anonymity is a reasonable assumption comes from large online markets, where bidders rarely expect to meet the same opponents again, and even when they do, are typically not aware of it. If you believe, as we do, that the anonymity assumption is reasonable in large markets, then the CME concept is attractive because it preserves the

---

[14]Reny (2008) allows for arbitrary partial orders on both the type and action spaces, but requires that under those orders, increasing types must take increasing actions. In our case, increasing types (e.g. a higher valuation for object 1) take some higher actions (bid higher in states where 1 is auctioned) and lower actions (bid lower in states where 2 is auctioned).

[15]In both cases, the steady-state ratio is given by $\frac{E[E_t]-(1-\rho)}{\rho}$.

limiting properties of other concepts while allowing for strategies that respond to endogenous fluctuations in state. On the other hand, if bidders pay attention to the specific actions of other bidders — as would be the case in a small or concentrated market — it is less sensible.

Lemmas 1 and 2 provide the main take homes from this section. First, in equilibrium bidders shade their bids down from their values, where the extent of shading depends on their continuation value in the current state. Notice immediately that this is starkly different from the "usual" model of second-price sealed bid auctions, where bids may be interpreted as valuations. Indeed, valuations are strictly higher than bids, implying that nonparametric estimates of the value distribution obtained by treating auctions as independent will be systematically biased upward. Second, the stationary distribution of types exists, but is different from the valuation distribution F, due to selection: bidders with low valuations will persist in the market for longer. So again, treating the auction data as a cross-section would be misleading. In the next section we develop nonparametric identification results for large auction markets that addresses both of these issues.

3 Nonparametric Identification

Equilibrium play implies a precise data generating process, with bidders entering, making bids and exiting according to the model. Suppose that one were to observe all the data produced in the course of equilibrium play, essentially consisting of the object auctioned in each period, the bids placed and the associated bidder identities. Could one then identify the underlying distribution of valuations, and thus recover demand?

We give a nonparametric identification result in the spirit of Athey and Haile (2002).[16] We think this is useful because it makes explicit the assumptions that are needed to identify the primitives of the dynamic game, as well as providing some guidance as to a sensible estimation strategy. For intuition, we will first work through a simple two good example.

---

[16] Similar identification arguments were made in Jofre-Benet and Pesendorfer (2003) and Pesendorfer and Schmidt-Dengler (2003). Our identification problem differs substantively since there are persistent latent variables — the valuations — whereas in their setup it is only the observable state variables that may persist.

3.1 Example

Suppose there are two goods, so $\mathcal{J} = \{j_1, j_2\}$. The exit probability ρ is constant across states. Supply is binomial and independent of state, with q the probability of good 1. Bidders coarsen the public information so that they condition only on the product identity in the current and next auction, implying four states: $1 = \{1, 1\}$, $2 = \{1, 2\}$, $3 = \{2, 1\}$ and $4 = \{2, 2\}$. Recall that an equilibrium bid strategy for a type x is a bid in each state, and let the equilibrium bids for a fixed type be $b_1 \cdots b_4$. Then the interim continuation value in state i, v_i, is given by:

$$v_i = G_1(b_i|i)\,(x_i - \mathbb{E}[B_1|B_1 < b_i, i]) + \delta(1-\rho)(1 - G_1(b_i|i)) \sum_{j=1}^{4} Q_{ij}\, v_j$$

where Q is the transition matrix between states. From the bidding function, we substitute out x_i as $b_i + \widetilde{v}_i$, where $\widetilde{v}_i$ is the ex-ante continuation value. Rearranging terms yields:

$$v_i - \delta(1-\rho) \sum_{j=1}^{4} Q_{ij}\, v_j = G_1(b_i|i)\,(b_i - \mathbb{E}[B_1|B_1 < b_i, i]) \tag{3}$$

Let $v = [v_1, v_2, v_3, v_4]^T$, and let u be given by:

$$u \equiv \begin{bmatrix} G_1(b_1|1)\,(b_1 - \mathbb{E}[B_1|B_1 < b_1, 1]) \\ G_1(b_2|2)\,(b_2 - \mathbb{E}[B_1|B_1 < b_2, 2]) \\ G_1(b_3|3)\,(b_3 - \mathbb{E}[B_1|B_1 < b_3, 3]) \\ G_1(b_4|4)\,(b_4 - \mathbb{E}[B_1|B_1 < b_4, 4]) \end{bmatrix}$$

i.e. the expected difference between bid and payment in any one period. Then (3) can be represented as the linear system:

$$(I - \delta(1-\rho)Q)\, v = u$$

where I is an $S \times S$ identity matrix. Next standard results imply that $(I - \delta(1-\rho)Q)$ is invertible, and therefore the existence of a unique solution for v (Stokey, Lucas, and Prescott 1989). Thus we have:

$$v = (I - \delta(1-\rho)Q)^{-1} u$$

13

Up to now, we have just been manipulating mathematical expressions. We now turn to the question of identification. Suppose that the econometrician knows the coarsening function T (i.e. the mapping from the public information, which is observed, to the state variables). Then each of the auctions can be correctly classified into the four states. Now fix a particular bidder in the dataset, who is observed participating in every state. For such a bidder, we can construct an S-length bid vector $b = b_1 \cdots b_4$ corresponding to their bid in each of the four states.[17] We shall call such bid vectors "complete". Plugging this vector into the above expression for u, we can recover the expected difference between payment and bid.

The transition matrix Q and exit probability ρ are identified directly from the data. So if the econometrician also knows the discount factor ρ, then the interim continuation values v of any type who bids b is identified. Now, we can move from v to the actual valuations x by again using the bid function expression:

$$x_1 = b_1 + \delta(1 - \rho)\left(qv_1 + (1 - q)v_2\right) = b_2 + \delta(1 - \rho)\left(qv_3 + (1 - q)v_4\right)$$
$$x_2 = b_3 + \delta(1 - \rho)\left(qv_1 + (1 - q)v_2\right) = b_4 + \delta(1 - \rho)\left(qv_3 + (1 - q)v_4\right)$$

where the expressions reflect the fact that good 1 is auctioned in states 1 and 2, while good 2 is auctioned in states 3 and 4. Thus for any bid vector b, we have identified the underlying valuation $x = (x_1, x_2)$, which is unique. In fact you can see that in this case, where $|J| < |S|$, the valuations are over-identified. This provides a potential test of the theory.

Given a random sample of complete bid-vectors, one could thus identify the valuation distribution F. But in the actual data the set of complete bid vectors is a selected sample, as some types will never participate in some states, and some types will persist longer, thus potentially being over-sampled. Assume for simplicity that all types participate in all states. So the key is to address the selection issue. As we show in the appendix, once the type x is identified for bid vector b, the probability that the bid vector is complete can also be identified.[18] Thus by inverting individual complete bid vectors to valuation vectors, and then re-weighting the density of these valuations by the inverse of the probability that they would be complete, we can identify the type density. This gives us demand.

---

[17]Of course, in the data they may have made many more than four bids; but as long as they have bid in every state, this construction is feasible.

[18]The idea is to recursively define the probability of being seen only in state 1, then in states 1 and 2....

3.2 Formal Result

The formal result simply summarizes what we have learnt from the example. The only difference is that we need to be a little careful about what we can identify for bidders who don't participate. Clearly, if a type never bids on product one, say, regardless of what the state is, then we cannot identify their valuation for product one. Thus for each subset B of $\mathcal{J}$, partition the type space X into $2^J - 1$ sets of types who bid only on products in B, regardless of the state. Let X_B be the $|B|$-dimensional random variable defined by restricting valuations X to products in B, with distribution function $F_B(x_B) \equiv P(X_B \leq x_B | X \in B)$. Then we have the following result:

Theorem 2 (**Identification**). *If the discount rate δ and the coarsening function T are known, the distribution F_B is non-parametrically identified for all $B \subseteq \mathcal{J}$. Moreover, the private valuation of any bidder observed bidding in every state $s \in S$ is identified.*

The conditions for identification are slightly stronger than is usual in these dynamic settings. Typically, it is necessary to know the discount rate.[19] But here the econometrician must also know the coarsening function. This is because the CME concept does not nail down bidder beliefs; it can accommodate a variety of models about which variables bidders pay attention to. The downside with this is that the econometrician must actually work out what those variables are in order to identify demand. Given these assumptions, we get pointwise identification for complete observations, via the same argument as in the example. We also get identification of the distribution of types who make a positive bid on every object.

In fact, if bidders are sufficiently forward looking, all types will make positive bids on every object in *some* state. To see this, notice that what causes a bidder to never bid on product j is the lost future surplus from winning other objects in future auctions. But if product j is to be auctioned every period for a long time — an event that will eventually happen — then the discounted future surplus from other objects will be very small, implying that they will bid on product j. Intuitively, variation in the set of upcoming auctions can be used to identify the valuations of bidders who may generally not participate.

---

[19]But as Ackerberg, Hirano, and Shahriar (2009) argue, if there are buy price auctions on the platform, it may be possible to identify discount rates from the data.

4 Estimation Strategy

Suppose that the conditions of the Theorem 2 are met, and the econometrician knows or can determine the discount rate and coarsening function. He also has a panel dataset, consisting of all the bids placed in each auction and bidder identify. Assume also for simplicity that all types bid on every good.[20] How should estimation proceed? We propose two different approaches. The first approach is nonparametric, following the logic of the identification section by inverting from observed bids to valuations. We look directly at the individual-level micro-data, treating the record of all bids placed by a given bidder as an observation. The individual-level data may differ in its dimensions: for some bidders, we may only see a single bid, while for others we may see many bids. The structural model implies that at most we should see S distinct bids by any one bidder, a different bid for every state. In the language of the section above, these S length bid vectors are "complete observations".

For complete observations, we can invert from the bid vector to a valuation vector via the first order condition provided we have estimates of the transition matrix and the distribution of opposing bids. This is very much like the approach of Guerre, Perrigne, and Vuong (2000). One important difference is that the set of complete observations is a selected sample of the bidders — bidders with high valuations are more likely to win and exit quickly, and therefore less likely to be observed bidding in every state. For this reason, it is necessary to re-weight the density of the estimated valuations in order to get an estimate of the type density.

The nonparametric approach is very clean and makes no parametric assumptions, but requires a fair number of complete observations. This may be impractical in markets with many states and high turnover in participants. Many bidders on eBay, for example, participate in only one or two auctions before either winning or giving up. We therefore also outline two additional approaches, where we impose successively more stringent parametric assumptions. First, we outline a semiparametric estimation approach based on simulated generalized method of moments, as is used elsewhere for demand estimation in industrial organization and marketing. There we assume a parametric structure on the distribution of types, and then choose parameters to match moments implied by the structural model with those observed in the data.

Even in that case, if there are a large number of products the model quickly becomes un-

---

[20]Where this fails, the estimation results presented here can be adapted with more work to account for non-participation; but it is important that non-participation is observable.

wieldy. So, following the literature (e.g. McFadden (1974)) we consider projecting product valuations onto characteristics. Instead of types being valuations for products, types are now random coefficients indicating the marginal value of product characteristics. In the standard specification we consider, this implies a linear structure for valuations in characteristics. Under reasonable distributional assumptions on the random coefficients, this linearity can be exploited and a very simple estimation procedure can be used.

Regardless of the approach — nonparametric, semiparametric or characteristic-based — there is a common first step in which a number of primitives are estimated.

4.1   Step 1: Estimate transitions, exit and payments

In the first step, we non-parametrically estimate the probability of winning with a bid of b in state s, $G_1(b|s)$; the expected payment conditional on winning, $\mathbb{E}[B_1|B_1 < b, s]$; the Markov transition matrix Q; the invariant measure over states π; and the probability of exit conditional on losing, $\rho = [\rho_1, \rho_2 \cdots \rho_S]$. This first step can be summarized as estimating elements of the per period payoffs and the transition probabilities, and is similar to that of both Bajari, Benkard, and Levin (2007) and Pakes, Ostrovsky, and Berry (2007) in their papers on dynamic games estimation.

All of these are conditional moments, and provided the conditioning variable is discrete — as the state variable is — we can consistently estimate the conditional moment from the relevant empirical analogue. So for example, to estimate an element of the transition matrix Q_{ij}, we have:

$$\widehat{Q_{ij}} = \frac{\sum_{t=1}^{T} 1(s_{t-1} = i)1(s_t = j)}{\sum_{t=1}^{T} 1(s_{t-1} = i)}$$

where $t = 1 \cdots T$ indexes auctions and $1(\cdot)$ is an indicator function. The only "difficult" object to estimate is $\mathbb{E}[B_1|B_1 < b, s]$ because for fixed s the conditioning variable b is continuous. This can be done state-by-state using any nonparametric approach, such as kernel density or sieve estimation.

4.2   Nonparametric Approach Step 2: Recover valuations

The key to the non-parametric approach is to treat the data as a sequence of (short) time series, one for each bidder. We restrict attention to complete observations, a subset of our

dataset consisting of S-dimensional bid vectors $b_i = (b_{i1} \ldots b_{iS})$.

For each observation i, we can use the first-stage estimates to construct a vector $\widehat{u}_i = (\widehat{u}_{i1}, \widehat{u}_{i2} \ldots \widehat{u}_{iS})$, where $\widehat{u}_{is} = b_{is} - \widehat{\mathbb{E}}[B_1|B_1 < b_{is}, s]$. Then the interim continuation value for bidder i, $\widehat{v}_i = (\widehat{v}_{i1} \ldots \widehat{v}_{iS})$, is the solution to the linear system $\widehat{v}_i = (I - \delta(1 - \rho)\widehat{Q})^{-1}\widehat{u}_i$. Moreover, we have from (2) that given a J-length sub-vector $\widetilde{b}_i$ of b_i consisting of bids on different objects, an associated sub-vector $\widetilde{\rho}$ of ρ, and an associated $J \times S$ submatrix $\widetilde{Q}$ of Q consisting of the transitions associated with the bids in $\widetilde{b}_i$, we get $x_i = \widetilde{b}_i + \delta(1 - \widetilde{\rho})\widetilde{Q}v_i$. Substituting in our estimates on the right hand side of this expression, we get an estimate $\widehat{x}_i$ of the valuation of each bidder.

The set of bidders with complete observations is a selected sample, and we need to correct for this. In the appendix, we derive an expression for $P(A, x)$, the probability that a type x is observed bidding only in the set of states in A, for $A \subseteq S$. This expression is identified from previously estimated objects. Now, the probability that a given type generates a complete observation is $P(S, x)$. We can thus correct for the selection bias by assigning a weight equal to $1/\widehat{P}(S, \widehat{x}_i)$ to each $\widehat{x}_i$, and then use weighted kernel density estimation to back out the type density $f(x)$.

We omit a formal analysis of the asymptotic properties of this estimator, both because it takes us into the realm of non-parametric estimation with dependent data and because we suspect that the semiparametric approach outlined below is more likely to be used in practice. Yet intuitively Lemma 2 guarantees that the data generating process quickly converges to an ergodic distribution, and so the asymptotics should be well-behaved. This is supported by the estimator's performance in our Monte Carlo experiments, below.

4.3 Semiparametric Approach Step 2a: Estimate bid function

The semiparametric approach goes in the opposite direction. Instead of inverting bids to valuations we take draws from a parametrized type distribution, simulate bids, and match the moments of the simulated bid distribution with those observed in the data. This places weaker demands on the data, as we need not observe a large sample of complete bid vectors.

The first part of the second step is working out how to simulate bids for a given type. Our idea is to solve for the optimal biding function in this environment. Bidders in this environment face a Markov Decision Problem (MDP): they need to choose a bid in each

18

state to maximize their payoff. As is the case in other environments, policy iteration will suffice to find the optimal bid vector for a type x.

To be more concrete, start with any initial strategy β_0, such as bidding the object's valuation ($\beta_0(x, s) = x_t$). Then from (1), one can solve for the ex-ante continuation value given that strategy by plugging in this bid into the probability of winning, expected payment on winning etc. From this estimated continuation value $v_0(x, s)$, we can define a new strategy $\beta_1(x, s) = x_t - v_0(x, s)$ and compute a new continuation value. Iterating in this way, we quickly converge to the optimal policy, since the iteration process obeys a contraction mapping.[21]

4.4 Semiparametric Approach Step 2b: Match Moments

Once we have bidding strategies, we are close to being able to calculate moments of the bid distribution in each state. Of course, we need to know not only how the types will bid, but also the stationary distribution of types. To close the model then, we form a parametric model for the type distribution $F_\theta \equiv F(X|\theta)$, where θ is a finite dimensional parameter. Let $\widetilde{F}$ be the stationary distribution of types. F_θ and $\widetilde{F}_\theta$ will typically be quite different. Our aim is to identify the true parameter vector θ_0, as this identifies demand.

A generic estimation approach is to simulate data from the structural model under θ, and compare the simulated and sample moments. Any minimum distance estimator that chooses θ to minimize the distance between thoughtfully chosen functionals of the observed and actual bid distributions will converge to the true θ_0, provided it is identified. To take a specific example, consider comparing a moment like $E[b|S = 1]$ — the mean bid in state 1 — across the sample and the parametric model. Dropping repeated bids in the same state by the same bidder, we can form the sample moment as a simple average. On the model side, we need to calculate $E_\theta[\beta(X, 1)]$, where the expectation is over the stationary distribution of types who bid in state 1, which we can estimate for fixed θ. The easiest way to do this is to simulate draws from F_θ, and compute the relevant bids using the estimated bid function.

Yet as noted earlier, the stationary distribution of bids in state 1 is not the type distribution F. So we need to re-weight the draws to consistently estimate $E_\theta[\beta(X, 1)]$. As we show in the appendix, one can use the first stage estimates to solve for the probability that in the stationary distribution any type x will be observed bidding in any subset of states $A \subseteq \mathcal{S}$.

---

[21]Here, for example, we see the benefit of assuming that bidders assume state transitions are exogenous. For if not, one would need a model of the counterfactual state transitions had another strategy been followed.

So the probability we see a bid in state 1 is just the complement of the probability of seeing only bids in A^c, and we can use to correctly re-weight and account for the selection issue.

Although any minimum distance estimator will do — and indeed in the Monte Carlo we use a particularly simple approach based on comparing the mean and variance of bids in a state, and the covariance across pairs of states — for the asymptotics it is easiest to appeal to standard results from the literature on GMM. Treat the data as a time series of auction observations $t = 1 \cdots T$. Then we can construct moment conditions based on the fact at the truth, sample and simulated moments should coincide. The asymptotic theory of Hansen (1982) applies, showing that provided the true parameter is identified and the environment is strictly stationary, the GMM estimation approach will recover the truth asymptotically. Note that Lemma 2 proves the stationarity of the environment, so that is satisfied. But we need also appeal to the results of Pakes and Pollard (1989) to argue that the simulation error has no impact on the asymptotic distribution of the estimator as the number of simulations grows large; and to either Andrews (1994) or Ai and Chen (2003) to argue that the non-parametric first stage does not preclude $\sqrt{N}$ consistency.

From a computational point of view, there are ways to speed up the estimation. One important bottleneck is that for every new parameter update and sample of bidders, we must solve a dynamic optimization problem for a large sample of simulated types. We can speed this up using an importance sampling approach.[22] Instead of drawing the types anew on each iteration and solving out for their bidding strategies, instead choose a set of R types initially to uniformly span some plausible region of X and compute their optimal bids. The choice of initial region is up to the researcher: one suggestion might be to regress prices on states, and then take the region of types spanned by the coefficient estimate plus four standard deviations on either side. This need only be done once. Then, to compute the simulated moments, we weight the types according to their relative likelihood under θ and compute the simulated moments as weighted sums.

4.5   Characteristic Space Approach

At the end of the day, we are trying to estimate the distribution of valuations over different products. As in the more general demand literature, this can be overly demanding of the data

---

[22]This idea was also used in Ackerberg (2003). See Ackerberg (2009) for a thorough discussion of this technique.

if the product space is large. Even after imposing a multivariate normal parametric structure, for example, we need to estimate a variance covariance matrix with $J(J+1)/2$ parameters. Given this, we may want to project valuations down onto product characteristics.

To do this, we assume that valuations depend on the characteristics of the goods $z_t \in \mathbb{R}^k$, as well as on tastes for the characteristics and an unobserved utility shock:

$$x_{ijt} = z_{jt}\alpha_i + \gamma_i \tag{4}$$

where we index individuals by i, products by j and auctions by t as before. The pair $(\alpha_i, \gamma_i) \in \mathbb{R}^{k+1}$ are the individual's type, assumed to be iid across individuals as before. This is similar to the random coefficients demand specification in the discrete choice literature. As usual, α_i captures individual-specific tastes for the various product characteristics. There are two important differences. First, γ_i replaces the totally idiosyncratic shock ε_{ijt}. The interpretation of γ_i is that it accounts for persistent differences across agents in their willingness-to-pay for any inside good. This could arise, for example, from income differences across agents.

Second, we have omitted any source of unobserved product heterogeneity — for example, an additive term ξ_j in the utility function. In principle a measurement error approach such as Li and Vuong (1998) could be used to generalize our estimation approach to this case, provided ξ_t was iid over time and not known by the agents in advance of period t. This introduces considerable complication, however, and we will not discuss it here.

Absent this problem, nothing from our previous semiparametric estimation approach need change. Given any type (α, γ), we can compute a valuation for each product in the dataset.[23] Then the model becomes identical to our earlier product-based model, and our estimation strategy is the same. Our first step is to recover the bid distributions, transition rules, and exit probability, so that we can compute the bid strategy for each type as before. Then given a parametric assumption on the joint distribution of (α, γ), we find the parameters that minimize the difference between a set of sample and simulated moments.

Yet under stronger assumptions we can use a much simpler estimation procedure. Suppose that agents condition only on the characteristics of the object under auction, so that z is the state variable. Assume also that supply is exogenous and uncorrelated over time, by which we

---

[23]We continue to assume there are a finite number of products, and therefore characteristic bundles, so the state space remains finite.

mean that the characteristics z_t of the goods to be auctioned are drawn independently each period from a fixed distribution with discrete support. Then the states evolve independently across periods, and the bidding function is:

$$\beta(\alpha, \gamma, z) = x - \widetilde{v}(x) = \alpha z + \gamma - \widetilde{v}(\alpha, \gamma)$$

where the first equality follows because the ex-ante continuation value is independent of the state, and the second follows by substituting in for the valuation.

Now, let the mean type be $(\mu_\alpha, \mu_\gamma) = (E_F[\alpha], E_F[\gamma])$. Then we can write bids as:

$$b_{it} = \alpha_i z_t + \gamma_i - \widetilde{v}(\alpha, \gamma) = \mu_0 + \mu_\alpha z_t + \varepsilon_{it} \tag{5}$$

for $\mu_0 = E[\gamma - \widetilde{v}(\alpha, \gamma)]$ and $\varepsilon_{it} = z_t(\alpha_i - \mu_\alpha) + \gamma - \widetilde{v}(\alpha, \gamma) - E[\gamma - \widetilde{v}(\alpha, \gamma)]$.

Since entry is exogenous, the types of new bidders will be uncorrelated with z_t. This implies the orthogonality condition $E[z_t \varepsilon_{it}] = 0$. So from (5) we can estimate the mean tastes for the characteristics μ_α by OLS regression of bids on characteristics, using a sample consisting only of the bids of entrants in the period they enter.[24] This hedonic regression may be all that is needed in some applications.

One might wonder why such a simple approach is possible here and not elsewhere. The key is that a characteristic-based model places strong and linear restrictions on the relationships between the valuations. In the product-based model, to make inferences about the joint valuation density for any pair of products (A,B), it was necessary to look at joint bids. By contrast, in characteristic space one can just look at the marginal distributions of bids on products A and B separately, and infer the mean taste for the characteristics. This implies that looking at the initial bids of new bidders — which is a sample of types directly from F — will suffice, and this allows us to completely abstract from the dynamic selection issues emphasized elsewhere in the paper.

---

[24]In practice, since we have conditional heteroscedasticity, using a feasible generalized least squares (FGLS) approach will be more efficient than OLS. This may be important in small samples.

5 Monte Carlo

We perform a simple Monte Carlo exercise to test both our nonparametric and parametric estimation approaches in small samples. In the simple case there are two goods, and bidders condition their bids only on the identity of the good and their private information, so there are two corresponding public states. In each period, each of these goods is equally likely to be listed. The number of entrants k is always 3 each period, but exit is random with losing bidders exiting with probability $\rho = 0.25$ and winning bidders exiting with certainty. The discount rate δ is set to 0.99. We consider alternative parameterizations for k and ρ. Bidders' private valuations are distributed bivariate normal with mean μ and covariance matrix Σ, which we also allow to vary in different Monte Carlo experiments.

Data is generated by first solving for the bidding function via policy iteration – though we have only been able to prove this converges in general for $|S| = 1$ (see Theorem 1), it has converged in all of our experiments to date. Then for each Monte Carlo iteration we simulate a dataset of 500 auctions, after discounting an initial 10,000 auctions as a "burn in". This amounts to, in expectation, 250 auctions per product, which seems like a moderate amount of data, especially given the volume of transactions online.

We run both estimation routines assuming that the econometrician knows the entry process, the discount rate δ and that the transitions between states are random rather than Markov. In the common first stage, we estimate the probability of exit ρ, and the probability of good one being listed, q. Then in the nonparametric second stage approach, we estimate the marginals of the type density using a Gaussian kernel density estimation approach with automatic bandwidth choice by cross-validation. In the parametric estimation, we (correctly) specify a multivariate normal distribution for the types, and take a minimum distance (MD) approach by matching the mean and variance of bids for each state in which a bidder may be observed, as well as the covariance between bids in each state for bidders observed in both.

Our results for the simple case are presented in Table 1. Experiment A is a baseline case with $k = 3$, $\rho = 1/4$, and symmetric μ and Σ. Experiment B slowed the rate of entry and random exit, allowing for fewer, but longer-lived bidders. Experiments C and D add positive and negative correlation between valuations, respectively, and finally experiment E adds positive correlation as well as significant asymmetry in the means and variance terms. Estimates from the structural models proposed in this paper appear in panel 1 for each experiment. Panel 2 presents, for comparison, estimates from a naive approach that treats each observed
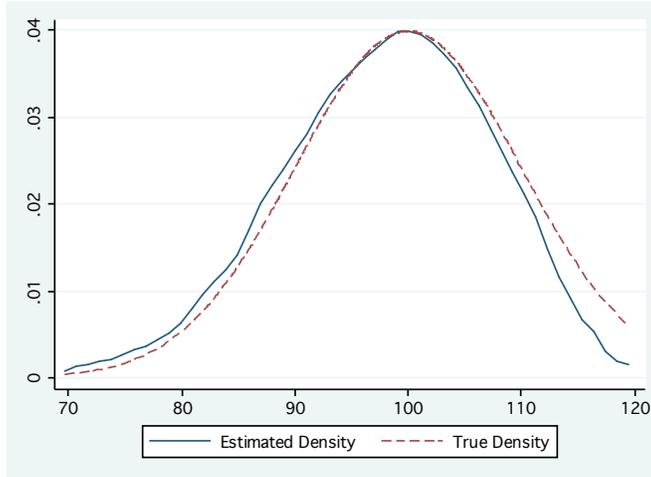
23

Figure 1: **Monte Carlo Simulation** The figure shows the true and estimated marginal density of valuations for product 1 for a randomly chosen Monte Carlo simulation of 500 auctions with specification E.

bid as a draw from the distribution of private valuations.

The supply parameter p and exit probability ρ are precisely estimated. Estimates of μ and Σ come from the parametric approach, and are found to be accurate and encouragingly precise for all parameterizations of the model. The naive approach gets both the means and the variances wrong in a statistically significant way. This bias derives from three sources that we have dealt with in our model: repeat bidding, selection by type, and bid shading according to the option value of losing.

In the far-right hand columns appear estimates of the mean integrated squared error (MISE), which is a measure of goodness-of-fit for our nonparametric approach.[25] We report MISE for the marginal distributions of types in both dimensions. In comparing these numbers, note that MISE is not a normalized measure, and is therefore not comparable across specifications. We can, however, compare the MISE generated by the structural approach with the MISE generated by the naive approach, and we see substantial improvement, with the structural model often doing a full order of magnitude better. For a graphical intuition of the fit achieved, Figure 1 depicts the true and the estimated marginal distribution of types for good 1 from experiment E of the simple model.

We also run a parallel set of Monte Carlo experiments for a forward-looking model, which is based on the identification example presented earlier in the paper. Now bidders pay

---

[25]For a candidate distribution $f_n(x)$ and a true distribution $f(x)$ this is calculated according to $MISE = \int (f_n(x) - f(x))^2 dF(x)$

attention to the good being auctioned in the next period, which means we now have four public states. Results are presented in Table 2. While the more complicated structure of the model has a cost in precision, both estimators continue to dramatically outperform the naive approach under all specifications. Note that the cost in precision is much more severe for the nonparametric approach– this is because, with four states instead of two, the sample of bidders for which we observe bids in every state is much smaller.

Finally, we run the same set of experiments for a backwards-looking variation on the model. In this version, the public state variable consists of the product being auctioned today as well as a dummy for whether there were more than 9 bidders in the previous auction. This is meant to capture inference regarding the level of demand. Results are presented in Table 3. It turns out that the level of demand only affects the bidding strategies of a small fraction of types significantly. Estimation performs roughly on par with that in the forward-looking model. As in that case, however, while estimates of σ_{12} are within two standard deviations of the truth, when the truth is nonzero the results consistently exhibit attenuation.

6   Conclusion

We have developed an estimable demand system for a large auction market. By defining and focusing on competitive Markov equilibria, we got an intuitive and tractable characterization of bidder strategies. We also showed existence of equilibrium, arguing that regardless of the initial conditions of the market, there is strong convergence to a stationary type distribution. A key result was that demand is identified from panel data, in which the same bidder is observed repeatedly bidding in different states.

Turning to estimation, we outlined two different approaches for recovering the distribution of types. These were tested by Monte Carlo simulation. From these exercises, we learned that accounting for dynamics and allowing for multiple products is important. Naïve estimates based on treating the data as a cross-section are systematically biased downwards, and are unable to account for correlation in the valuations for different products.

Understanding these basic empirical problems will be helpful for future research. We hope also that this is a first step towards obtaining satisfying models for other markets: those where bidders have market power, or non-unit demand. In future work, we would like to take this model to data and see how it performs.

References

ACKERBERG, D. A. (2003): "Advertising, learning, and consumer choice in experience good markets: an empirical examination," *International Economic Review*, 44(3), 1007—1040.

——— (2009): "A new use of importance sampling to reduce computational burden in simulation estimation," *Quantitative Marketing and Economics*, 7(4), 343—376.

ACKERBERG, D. A., K. HIRANO, AND Q. SHAHRIAR (2009): "Identification of Time and Risk Preferences in Buy Price Auctions," working paper, UCLA.

ADAMS, C. P. (2009): "Identifying Demand from Online Auction Data," Working Paper, Federal Trade Commission.

AI, C., AND X. CHEN (2003): "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71(6), 1795–1843.

ANDREWS, D. W. K. (1994): "Asymptotics for Semiparametric Econometric Models Via Stochastic Equicontinuity," *Econometrica*, 62(1), 43–72.

ATHEY, S. (2001): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69(4), 861—889.

ATHEY, S., AND P. A. HAILE (2002): "Identification of Standard Auction Models," *Econometrica*, 70(6), 2107–2140.

BAJARI, P., L. BENKARD, AND J. LEVIN (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75(5), 1331–1370.

BAJARI, P., AND A. HORTACSU (2004): "Economic Insights from Internet Auctions," *Journal of Economic Literature*, 42(2), 457–486.

BERRY, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics*, 25, 242–262.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *The Journal of Political Economy*, pp. 68—105.

BUDISH, E. (2008): "Sequencing and Information Revelation in Auctions for Imperfect Substitutes: Understanding eBay's Market Design," Working paper, Harvard University.

DUFFIE, D., J. GEANAKOPLOS, A. MAS-COLELL, AND A. MCLENNAN (1994): "Stationary Markov Equilibria," *Econometrica*, 62(4), 745—781.

EILENBERG, S., AND D. MONTGOMERY (1946): "Fixed point Theorems for Multi-Valued Transformations," *American Journal of Mathematics*, 68, 214—222.

FERSHTMAN, C., AND A. PAKES (2009): "Finite State Dynamic Games with Asymmetric Information: A Framework for Applied Work," Working Paper, Harvard University.

GOWRISANKARAN, G., AND M. RYSMAN (2009): "Dynamics of Consumer Demand for New Durable Goods," Working paper, University of Arizona.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029—1054.

HENDRICKS, K., AND R. PORTER (2007): "A Survey of Empirical Work in Auctions," in *Handbook of Industrial Organization*, ed. by M. Armstrong, and R. Porter, vol. III. North-Holland.

INGSTER, A. (2009): "A Structural Study of a Dynamic Auction Market," Working Paper, NYU Stern.

JOFRE-BENET, M., AND M. PESENDORFER (2003): "Estimation of a Dynamic Auction Game," *Econometrica*, 71(5), 1443–1489.

JOVANOVIC, B., AND R. W. ROSENTHAL (1988): "Anonymous Sequential Games," *Journal of Mathematical Economics*, 17(1), 77—87.

KRUSELL, P., AND A. A. SMITH (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867—896.

LI, T., AND Q. VUONG (1998): "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, 65, 135—165.

McFadden, D. (1974): "Conditional logit analysis of qualitative choice behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, pp. 105—142. Academic Press.

Meyn, S., and R. L. Tweedie (2009): *Markov Chains and Stochastic Stability.* Cambridge University Press.

Milgrom, P., and R. Weber (2000): "A Theory of Auctions and Competitive Bidding II," in *The Economic Theory of Auctions*, ed. by P. Klemperer. Edward Elgar Publishing Ltd.

Pakes, A., M. Ostrovsky, and S. Berry (2007): "Simple Estimators for the Parameters of Discrete Dynamic Games, with Entry/Exit Examples," *Rand Journal of Economics*, 38(2), 373—399.

Pakes, A., and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57(5), 1027–1057.

Pesendorfer, M., and P. Schmidt-Dengler (2003): "Identification and Estimation of Dynamic Games," Working Paper, LSE.

Reny, P. J. (2008): "On the Existence of Monotone Equilibria in Bayesian Games," Working Paper, University of Chicago.

Said, M. (2009): "Auctions with Dynamic Populations: Efficiency and Revenue Maximization," Working paper, Yale University.

Sailer, K. (2006): "Searching the eBay Marketplace," Working Paper, University of Munich.

Satterthwaite, M., and A. Shneyerov (2007): "Dynamic Matching, Two-Sided Incomplete Information and Participation Costs: Existence and Convergence to Perfect Competition," *Econometrica*, 75(1), 155–200.

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics.* Harvard University Press.

Weintraub, G., C. L. Benkard, and B. Van Roy (2008): "Markov Perfect Industry Dynamics with Many Firms," *Econometrica*, 76(6), 1375–1411.

WOLINSKY, A. (1988): "Dynamic Markets with Competitive Bidding," *The Review of Economic Studies*, 55(1), 71—84.

ZEITHAMMER, R. (2006): "Forward-looking bidding in online auctions," *Journal of Marketing Research*, 43(3), 462–476.

——— (2009): "Sequential auctions with information about future goods," Working paper, UCLA Anderson.

Appendix

Proof of Lemma 1: Conditional on a positive bid, the expected payoff function $\pi(b,s)$ has the form:

$$\pi(b,s) = \int_0^b (x_t - B^1)g(B^1|s)dB^1 + \left(\int_b^\infty g(B^1|s)dB^1\right)(1-\rho)\delta \sum_{s'=1}^S v(x,s')Q_{ss'}$$

since we show in Corollary 1 below that the highest bid has a density on $\mathbb{R}^+$. Writing $\widetilde{v}(x,s)$ for $(1-\rho)\delta \sum_{s'=1}^S v(x,s')Q_{ss'}$, and taking an FOC in b, we get $(x_t - b)g(B^1|s) - g(B^1|s)\widetilde{v}(x,s) = 0$, which has unique solution $b_t = x_t - \widetilde{v}(x,s)$. If this is negative, then the corner solution $b_t = 0$ delivers the highest payoff conditional on a positive bid; and it is easily shown that non-participation ϕ does better still because the probability of winning with a bid of zero is positive (e.g. there may be no other bidders in the market). $\qquad\square$

Proof of Lemma 2: By Theorem 11.12 in Stokey, Lucas, and Prescott (1989), uniform geometric convergence in total variation norm will be achieved if their "condition M" holds: there exists $\varepsilon > 0$ and $N \geq 1$ such that for all $A \in \mathcal{F}$, either $P^N(\omega, A) \geq \varepsilon$ $\omega \in \Omega$ or $P^N(\omega, A^c) \geq \varepsilon$ $\forall \omega \in \Omega$, for P^N the N-step transition probability. Let $s_0 = T(\mathbf{0}, \mathbf{1})$ and $\omega_0 = (\mathbf{0}, s_0)$ (i.e. ω_0 is a superset of the true state where there are no current participants, nor have been for the observable history; and upcoming supply is entirely of object 1). Let $p_0 = \mathbb{P}(E_t = 0|N_{t-1} = 0) > 0$, by assumption; let $q_0 = \mathbb{P}(j_t = 1)$; and let $N_0 = \max\{k^f, k^h\} + 1$. Then $P^{N_0}(\omega, \omega_0) \geq \rho^{\bar{N}}(p_0 q_0)^{N_0} > 0$ $\forall \omega \in \Omega$. Choosing $N = N_0$ and $\varepsilon = \rho^{\bar{N}}(p_0 q_0)^{N_0 - 1}$, condition M holds since either $\omega_0 \in A$ or $\omega_0 \in A^c$. Next, consider the marginal ergodic distribution on $\mathcal{X}^{\bar{N}-1} \times S$ produced by integrating out the first non-zero element of the type vector. Since S is finite, there is a well-defined conditional distribution of $\mathbf{x}_{-i}$ given s. $\qquad\square$

Proof of Theorem 1: Let the action space $\mathcal{A}_k$ be a finite grid $\{\phi, 0, \frac{\bar{x}}{k}, \frac{2\bar{x}}{k} \cdots \bar{x}\}^S$, with each dimension totally ordered as written. Endow A_k with the Euclidean metric denoted d, where ϕ is treated as having value $\frac{-\bar{x}}{k}$. Here, we define strategies as mappings from type to actions, $\beta : \mathcal{X} \to \mathcal{A}_k$, and write $\beta(x) = (\beta^1(x), \beta^2(x) \cdots \beta^S(x))$. For each s, let $\lambda(s) \in \mathcal{J}$ be the object under auction in state s; the mapping is well defined because of restriction (i) on the coarsening T. As in the text, we call β monotone if $\beta^s(x)$ is increasing in x_j if $j = \lambda(s)$ and decreasing otherwise. Following Reny (2008), let M denote the space of monotone functions

from $\mathcal{X}$ to $\mathcal{A}_k$, and define a metric δ on M by:

$$\delta(\beta, \beta') = \int_{\mathcal{X}} d(\beta(x), \beta'(x)) d\mu(x)$$

As shown there, (M, δ) is a compact absolute retract. Define the payoff function $u(a, t, \beta_{-i})$, equal to the value of the game for type t when playing action vector a, everyone else follows strategy β_{-i}, and the invariant measure over true states is determined as though *everyone* plays β_{-i}. As Reny (2008) shows, the interim payoff function $U(\beta, \beta_{-i}) = \int_{\mathcal{X}} u(\beta(t), t, \beta_{-i}) d\mu(x)$ is continuous in both arguments since the type distribution is atomless.

Let $\Gamma(\beta_{-i})$ be the best response correspondence; by the theorem of the maximum it is non-empty valued and upper-hemicontinuous. A fixed point of the correspondence, β^*, will be a CME, since then strategies will be optimal given that the invariant measure over types is generated by play of β^*. The existence of such a fixed point is guaranteed by the fixed point theorem due to Eilenberg and Montgomery (1946) if we can show that Γ is contractible-valued and $\Gamma(\beta_{-i}) \in M$ whenever $\beta_{-i} \in M$.

Fix $\beta_{-i} \in M$ and $x \in \mathcal{X}$, and let $\widetilde{v}(x, s)$ still denote the ex-ante continuation value given optimal actions. Then we can look at the optimal action state by state, and we have $b_s^* = \arg\max_{b \in \mathcal{A}_k} G_1(b|s)(x_{\lambda(s)} - E[B^1|B^1 \leq b, s]) + (1 - G_1(b|s))\widetilde{v}(x, s)$. Define $b_s^l = \max_b\{b : x_{\lambda(s)} - E[B^1|B^1 \leq b, s] - \widetilde{v}(x, s) > 0\}$ and $b_s^u = \min_b\{b : x_t - E[B^1|B^1 \leq b, s] - \widetilde{v}(x, s) < 0\}$. Comparing their payoffs directly, we see that b_s^l yields higher payoff than b_s^u; this proves that these two bids could not both be best responses. Next, at least one of these must be a best response — suppose $b < b_s^l$ is optimal, then $x_t - E[B^1|B^1 \leq b, s] - \widetilde{v}(x, s) > 0$; but since the payoff function is a weighted sum of $x_t - E[B^1|B^1 \leq b, s]$ and $\widetilde{v}(x, s)$, and $G_1(b|s)$ is (weakly) increasing in b for $\beta_{-i} \in M$, b_s^l must be weakly better still, and also a best response. The same argument holds on the other side. In sum, either b_s^l or b_s^u (but not both) is part of the optimal strategy, and the best response is unique unless $G_1(b|s)$ is constant over actions in some state (i.e. β_{-i} assigns no mass to these actions).

Two results follow. First, since b_s^l and b_s^u are increasing in $x_{\lambda(s)}$ and decreasing in $x_{-\lambda(s)}$ — since v increases in $x_{-\lambda(s)}$ — the best response to any $\beta_{-i} \in M$ is also in M. Second, given best responses b and b', the join $b \vee b'$ is also a best response, since the probability of winning must be constant over those actions, and so taking the coordinate-wise max doesn't change payoffs at all. Γ is thus join-closed and non-empty for $\beta_{-i} \in M$.

Finally, then, we need to construct a contraction on the best reply correspondence. We

slightly adapt Reny (2008) because actions are not increasing in type in all states. Fix β_{-i} and $(\beta, \beta_0) \in \Gamma(\beta_{-i})$. Let $F_i(x)$ be the marginal distribution of valuations for product i, and define functions $\Phi^s(x) = \frac{F_{\lambda(s)}(x) - \bar{F}_{-\lambda(s)}(x) + 1}{2}$, where $\bar{F}_{-\lambda(s)}(x)$ is the average valuation quantile of objects not auctioned at s. Then define

$$h^s(\tau, \beta)(x) = \begin{cases} \beta^s(x), & \text{if } \Phi^s(x) \leq |1 - 2\tau| \text{ and } \tau < 0.5 \\ \beta_0^s(x), & \text{if } \Phi^s(x) \leq |1 - 2\tau| \text{ and } \tau \geq 0.5 \\ \beta_0^s(x) \vee \beta^s(x), & \text{if } \Phi^s(x) > |1 - 2\tau| \end{cases}$$

Define $h(\tau, \beta)(x) = (h^1(\tau, \beta)(x), h^2(\tau, \beta)(x) \cdots h^S(\tau, \beta)(x))$. It can be verified that $h(\tau, \beta)$ is monotone, continuous and a best reply almost everywhere. This proves Γ contractible valued, and hence it has a fixed point on M, showing existence for the finite action game.

To complete the proof, it suffices to show that there is a limiting set of strategies in which no positive measure of types makes a bid $b \in \mathbf{R}^+$ in any state s (a positive measure of types bidding ϕ presents no problems). But we showed earlier that for any grid fineness k, either b_s^l or b_s^u is played by every type x against the equilibrium strategy β_k^*; taking a limit as $k \to \infty$, we get that strategy $\beta_\infty(x) = b$ for b solving $x_{\lambda(s)} - \tilde{v}(x, s) = 0$ is played in some limit equilibrium for all x. Fix this equilibrium, some state s and some bid $b \in \mathbf{R}^+$. Since $\beta_\infty(x)$ is strictly increasing in x_t, the set of types bidding b must be of dimension $J - 1$ and thus Lebesgue-negligible. Finally, since the type distribution is absolutely continuous with respect to Lebesgue measure, this implies the measure of types bidding b is zero.

For the uniqueness claim, define an operator $\Gamma(\beta)(x, s) = x - \tilde{v}_\beta(x, s)$ on the space of continuous strictly increasing pure strategies under the sup norm. Showing Γ a contraction mapping suffices to prove uniqueness:

$$\begin{aligned} \|\Gamma(\beta) - \Gamma(\tilde{\beta})\| &= \max_{s \in \mathcal{S}} \sup_{x \in \mathcal{X}} |\tilde{v}_\beta(x, s) - \tilde{v}_{\tilde{\beta}}(x, s)| \\ &\leq \max_{s \in \mathcal{S}} \sup_{x \in \mathcal{X}} |G_1^\beta(\beta(x, s))(x - \mathbb{E}[B_1 | B_1 \leq \beta(x, s), s]) \\ &\quad - G_1^{\tilde{\beta}}(\tilde{\beta}(x, s))(x - \mathbb{E}[B_1 | B_1 \leq \tilde{\beta}(x, s), s])| \\ &\leq \max_{s \in \mathcal{S}} \sup_{x \in \mathcal{X}} G_1(\beta(x, s)) \left| \mathbb{E}[B_1 | B_1 \leq \beta(x, s), s] - \mathbb{E}[B_1 | B_1 \leq \tilde{\beta}(x, s), s] \right| \\ &\leq \|\beta - \tilde{\beta}\| \end{aligned}$$

where in the third line we use the fact that $G_1^\beta(\beta(x,s)) = G_1^{\widetilde{\beta}}\left(\widetilde{\beta}(x,s)\right)$ for all x and s. This property holds because under both strategies the same winner is chosen in each auction (types are totally ordered for $J = 1$), and thus the transition functions and invariant measures over types are the same in both cases; implying the distribution functions are equal if evaluated at the bids of a fixed type x in any state s. □

Corollary 1. *In equilibrium the distribution $G(b|b > 0, s)$ is absolutely continuous for all s.*

Proof: Since the equilibrium bid function is strictly monotone on $b \geq 0$, and the max operator selecting the highest bid is continuous, it will suffice to show that the invariant measure on the type space μ_x is absolutely continuous with respect to Lebesgue measure. So fix a Lebesgue-negligible set A on $\mathcal{X}^{\bar{N}}$, and let $C = A \times \mathcal{S}$. Following Meyn and Tweedie (2009), let $L(x, C) = \lim_{N \to \infty} P^N(x, C)$. It suffices to show $L(x, C) = 0$. But to reach C, one must draw a sequence of types from A, and since the distribution of entrants F is absolutely continuous with respect to Lebesgue measure, if A is negligible, the probability of those draws is zero. Thus $L(x, c) = 0$, completing the proof. □

Proof of Theorem 2: Following precisely the logic of the identification example, we can write the vector of continuation values of a bidder who bids b as the solution to a linear system $V = (I - \delta(1 - \rho)Q)^{-1}u$, where Q is the transition matrix, $\rho = \rho_1, \rho_2 \cdots \rho_S$ is the vector of exit probabilities and $u = [u_1, u_2 \cdots u_S]$ is the S-length vector with terms of the form $u_s = G_1(b_s|s)(b_s - \mathbb{E}[B_1|B_1 < b_s, s])$. Notice that since $b_s = \phi$ is possible, we may have $u_s = 0$ for some s. Q and ρ are identified from the data, δ is assumed known and u is identified for any bidder observed bidding in every state. Also $(I - \delta Q)^{-1}$ exists for $\delta < 1$. Given these objects, the ex-ante continuation value is identified for all complete observations. Now fix any set $B \subset \mathcal{J}$. Then to get the valuation vector restricted to B, x_B, we simply add the continuation value to the bid for each of $|B|$ states where distinct products in B are auctioned, where the bids are positive on B by assumption. Moreover, any type in the interior of the support of F_B has a positive probability of generating a complete observation. So by re-weighting the density of valuations for complete observations, we recover F_B. The correct re-weighting factor is $P(S, x)$, identified from the data and defined in (7) below. □

Selection Correction Probabilities: Let A be a subset of S. We want to get the probability that any type x ends up submitting bids in the states in A. Define $p(x, s) = G_1(\beta(x, s)|s) + (1 - G_1(\beta(x, s)|s)) \rho(s)$, which is just the probability that a type x will exit the sample in state s, whether by winning or losing. Also define $P(B, x, s)$ to be the probability of a bidder

x who enters the sample in state s being observed bidding only in states $B \subseteq \mathcal{S}$. We can express this recursively:

$$P(B, x, s) = 1(s \in B) \left[p(x, s) + (1 - p(x, s)) \sum_{s' \in B} Q_{ss'} P(B, x, s') \right] \tag{6}$$

Then the probability of observing a bidder x in group A can be defined implicitly as:

$$P(A, x) = \sum_{s \in A} \pi(s) P(A, x, s) - \sum_{B \subset A} P(B, x) \tag{7}$$

where π is the invariant measure over states. The idea is simply that the probability of seeing bids for every state s in A is equal to the probability that the bidder stays within A less the probability that he stays in a strict subset of A.

Table 1: Monte Carlo Results: Simple Model

		p	ρ	μ_1	μ_2	σ_1^2	σ_2^2	σ_{12}	MISE1	MISE2
A	true	0.5	0.25	100	100	100	100	0		
1	est	0.49833	0.25234	99.9402	100.009	100.0492	98.6958	-0.64142	4.6255e-005	4.8577e-005
	std	0.023662	0.0075173	0.26534	0.26359	4.632	4.9464	4.4967	1.1449e-005	1.163e-005
2	est			97.1177	97.1304	72.1543	71.3238		0.0010196	0.001067
	std			.34195	0.31483	4.101	4.1475		0.00044665	0.00041746
B	true	0.5	0.125	100	100	100	100	0		
1	est	0.50225	0.12626	100.0734	99.9833	99.5718	100.3942	-0.64446	9.6628e-005	9.5621e-005
	std	0.024032	0.0053759	0.33802	0.31305	8.3594	7.7141	5.952	1.5998e-005	1.4511e-005
2	est			94.8977	94.9262	59.9129	59.975		0.0029342	0.0029148
	std			0.42577	0.42052	4.9428	4.1342		0.0004266	0.00034493
C	true	0.5	0.25	100	100	100	100	50		
1	est	0.49936	0.25133	100.0668	99.9853	100.1644	101.752	50.7979	7.9872e-005	8.1255e-005
	std	0.020988	0.0063798	0.23589	0.25533	3.9595	4.2876	4.3402	1.5956e-005	1.4507e-005
2	est			96.5085	96.4077	65.0887	66.731		0.001157	0.0012266
	std			0.32974	0.31951	3.3594	3.3033		0.00044747	0.0003483
D	true	0.5	0.25	100	100	100	100	-50		
1	est	0.49912	0.2512	100.0154	99.9708	99.6918	100.9662	-49.2951	3.4979e-005	3.5023e-005
	std	0.024874	0.0072957	0.36192	0.32671	5.4416	5.4446	5.0245	1.3121e-005	9.4455e-006
2	est			97.8661	97.8516	74.91	75.1188		0.00096236	0.00097535
	std			0.42579	0.40438	3.8811	3.9719		0.00041383	0.00039483
E	true	0.5	0.25	100	150	100	400	100		
1	est	0.50191	0.25225	99.9653	149.9225	98.8567	398.8675	99.1276	0.00010318	1.6378e-005
	std	0.021896	0.007705	0.23589	0.49806	4.7741	16.1838	9.5806	2.012e-005	3.1715e-006
2	est			95.7595	143.7828	61.4454	278.0839		0.0011407	0.00026994
	std			0.33151	0.72523	3.4637	13.7853		0.00052147	0.00011414

For each experiment A-E, panel 1 depicts structural estimates and panel 2 the naive estimates.

Experiment A: baseline case, with $k = 3$, $\rho = 1/4$, and symmetric μ and Σ.

Experiment B: longer-lived bidders, with $k = 2$ and $\rho = 1/8$.

Experiment C: positive correlation in valuations, with $\sigma_{12} = 50$.

Experiment D: negative correlation in valuations, with $\sigma_{12} = -50$.

Experiment E: Asymmetric μ and Σ, as well as positive correlation in valuations.

Table 2: Monte Carlo Results: Forward-Looking Model

		p	ρ	μ_1	μ_2	σ_1^2	σ_2^2	σ_{12}	MISE1	MISE2
A	true	0.5	0.25	100	100	100	100	0		
1	est	0.49987	0.24901	99.9341	99.808	94.8547	94.8593	-1.4892	0.00022831	0.00022285
	std	0.021162	0.0065131	0.2871	0.29919	6.4142	6.3335	4.9616	5.5919e-005	4.6794e-005
2	est			97.1755	97.025	71.8718	72.8582		0.00068577	0.00063888
	std			0.33698	0.33314	3.4844	3.7971		0.00034674	0.00033673
B	true	0.5	0.125	100	100	100	100	0		
1	est	0.49768	0.12487	99.9521	100.0425	97.5813	97.3092	-2.959	0.00042409	0.00041063
	std	0.023491	0.0057885	0.39795	0.40503	7.982	9.1703	6.1737	8.064e-005	7.7461e-005
2	est			94.9433	95.0142	59.7917	59.2406		0.0019602	0.0019706
	std			0.40746	0.40215	4.5596	4.3055		0.00032758	0.00026379
C	true	0.5	0.25	100	100	100	100	50		
1	est	0.4941	0.25058	99.819	99.8049	91.8484	92.6894	42.9742	0.0003049	0.0003044
	std	0.026691	0.0069024	0.3072	0.28907	5.9086	5.3654	5.2429	7.199e-005	5.835e-005
2	est			96.4719	96.4774	65.7957	65.7867		0.00083543	0.00081302
	std			0.30472	0.34834	3.3865	3.3406		0.00034527	0.00035448
D	true	0.5	0.25	100	100	100	100	-50		
1	est	0.50369	0.25079	99.949	100.0644	96.3282	97.4392	-44.4073	0.00023373	0.00023328
	std	0.023882	0.0069434	0.32751	0.321	8.1597	8.5975	5.6176	5.9192e-005	6.275e-005
2	est			97.7889	97.9007	75.2727	75.6323		0.00057612	0.00052678
	std			0.36284	0.37382	4.0511	4.6602		0.00029306	0.00032672
E	true	0.5	0.25	100	150	100	400	100		
1	est	0.50503	0.25052	99.7609	149.6732	92.5428	369.852	87.3218	0.00035134	6.9119e-005
	std	0.022175	0.0064735	0.32631	0.55967	6.1947	25.9437	11.3395	5.9314e-005	1.4589e-005
2	est			95.7264	143.7643	62.4126	277.382		0.00082781	0.00019845
	std			0.35323	0.61251	3.5549	14.2732		0.00041283	8.2194e-005

For each experiment A-E, panel 1 depicts structural estimates and panel 2 the naive estimates.

Experiment A: baseline case, with $k = 3$, $\rho = 1/4$, and symmetric μ and Σ.

Experiment B: longer-lived bidders, with $k = 2$ and $\rho = 1/8$.

Experiment C: positive correlation in valuations, with $\sigma_{12} = 50$.

Experiment D: negative correlation in valuations, with $\sigma_{12} = -50$.

Experiment E: Asymmetric μ and Σ, as well as positive correlation in valuations.

Table 3: Monte Carlo Results: Backward-Looking Model

		p	ρ	μ_1	μ_2	σ_1^2	σ_2^2	σ_{12}	MISE1	MISE2
A	true	0.5	0.25	100	100	100	100	0		
1	est	0.50225	0.24903	99.5455	99.476	90.5713	90.0723	0.81214	0.00030105	0.00032009
	std	0.022929	0.0066757	0.283	0.29668	7.5447	7.9162	4.8788	9.6654e-005	0.00010594
2	est			97.0564	97.0128	66.3069	66.7998		0.001306	0.001079
	std			0.34049	0.34873	3.3928	3.9655		0.00039424	0.00047592
B	true	0.5	0.125	100	100	100	100	0		
1	est	0.50543	0.126	99.4281	99.261	94.1747	93.1985	-0.99954	0.00057706	0.00060249
	std	0.018153	0.0053351	0.37785	0.35817	9.5136	10.7578	6.787	0.00015555	0.00015446
2	est			94.8876	94.8948	56.1622	55.311		0.0030883	0.0031265
	std			0.37063	0.38293	4.6828	4.2582		0.00045435	0.00048513
C	true	0.5	0.25	100	100	100	100	50		
1	est	0.5014	0.24894	99.4041	99.3418	89.3272	89.2724	41.2166	0.00035513	0.00037807
	std	0.021388	0.0060638	0.27722	0.26244	7.4923	7.6598	5.9907	9.662e-005	9.3648e-005
2	est			96.3922	96.3618	61.9956	61.7951		0.0013294	0.0013081
	std			0.34935	0.26843	3.4279	3.4378		0.00033777	0.00045434
D	true	0.5	0.25	100	100	100	100	-50		
1	est	0.50131	0.24898	99.7147	99.5963	92.185	90.9354	-40.2983	0.00031033	0.00033724
	std	0.024505	0.0061843	0.32289	0.30584	8.3661	7.6269	5.7865	0.00010573	9.9659e-005
2	est			97.8232	97.751	68.4931	68.5974		0.0011092	0.00095543
	std			0.35707	0.36555	3.4815	4.2274		0.00030701	0.00044662
E	true	0.5	0.25	100	150	100	400	100		
1	est	0.50115	0.25014	99.2897	148.9155	88.1829	363.6608	83.7978	0.00040269	8.7667e-005
	std	0.022705	0.0065597	0.28836	0.53591	8.3556	27.3134	11.0149	9.2293e-005	2.3743e-005
2	est			95.9237	143.5118	59.2387	258.8469		0.0012831	0.00028309
	std			0.32372	0.65581	3.7201	13.116		0.00040399	0.0001278

For each experiment A-E, panel 1 depicts structural estimates and panel 2 the naive estimates.

Experiment A: baseline case, with $k = 3$, $\rho = 1/4$, and symmetric μ and Σ.

Experiment B: longer-lived bidders, with $k = 2$ and $\rho = 1/8$.

Experiment C: positive correlation in valuations, with $\sigma_{12} = 50$.

Experiment D: negative correlation in valuations, with $\sigma_{12} = -50$.

Experiment E: Asymmetric μ and Σ, as well as positive correlation in valuations.