Demand Estimation in Auction Platform Markets

Matthew Backus[*] Gregory Lewis[†]
University of Michigan Harvard University

[Preliminary and Incomplete - please do not cite]
July 3, 2009

Abstract

We provide a framework for estimating a demand system in auction platform markets with numerous persistent buyers, heterogeneous objects and unit demand. We construct a dynamic model of repeated second-price auctions with bidder entry and exit, in which consumers have independent and private valuation vectors over the full set of objects. We prove existence of equilibrium and characterize equilibrium strategies, and then provide sufficient conditions for the nonparametric identification of the joint distribution of private values. Nonparametric and semiparametric estimation procedures are proposed and tested by Monte Carlo simulation. Taking this model to a large dataset of auctions from eBay Motors, we find evidence that bidders do indeed substitute across auctions of similar products, and that their bidding strategies are responsive to the set of competing auctions at the time when they bid.

[*]Email: backus@umich.edu
[†]Email: glewis@fas.harvard.edu

1 Introduction

Auction platform markets are responsible for a substantial and growing share of trade in the US economy. E-commerce as a whole has been growing rapidly, with estimated growth of 13.9% in 2006, relative to 6.8% growth rates for offline trade in the same sectors.[1] Of this trade, 93% is accounted for by business-to-business transactions. It is harder to quantify the share held specifically by auction platforms, but in retail eBay alone has revenues of $36 billion from its auctions business in 2007[2]; Google realized $21 billion in revenue from its online advertising platform in 2008[3]; and 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 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 advertising market of a merger between Microsoft and Yahoo.

At first glance, auctions data is an extremely rich of information about demand. 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. For example, a buyer who loses in an auction for a particular Corvette may choose to bid in an auction for a Mustang. This is informative for demand, much in the same way that “second-choice” data is useful in Berry, Levinsohn, and Pakes (2004). Moreover, the choice sets available to buyers vary with high frequency as auctions come and go, which is also useful 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 try to 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.

[1]Source: US Census Bureau, E-Stats Report 2006. Includes only the manufacturing, merchant wholesale, retail and selected services sectors; excludes eMarketplaces, such as eBay.

[2]Source: eBay Annual Report for 2007

[3]Source: Google Annual Report for 2008

In this paper, we provide theory, identification and estimation results for estimating demand in auction platform markets with a large number of relatively short-lived buyers. We perform Monte Carlo simulations to show that our estimation approach works well in moderately sized samples. As an application, we look at a large panel of auctions on eBay Motors. We find reduced-form evidence that bidders substitute across models, and respond dynamically to the set of upcoming auctions.

A key assumption in the analysis will be that the market is *large*, in the sense that there are many bidders and similar objects up for auction. In this case, the true state of the market at any time will be complex and unknown to bidders. Rather than forming beliefs about the unknown true state, they may instead condition their behavior on a simpler publicly observable state vector and on their own private valuation. Similar assumptions underpin the equilibrium notion of Fershtman and Pakes (2009). An attractive feature of this approach is that determining what the relevant states are becomes an empirical question.

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 observations of bids placed when facing different sequences of upcoming auctions.

The theory makes a number of simple predictions, which we test in the next part of the paper using a dataset of over 50 000 auctions of Corvettes, Mustangs and Cameros from eBay Motors, the nation's largest used car market. We show widespread evidence that bidders bid sequentially in multiple auctions, often on objects that are not similar but not identical. For example, 7% of bidders on first generation Cameros also bid on first generation Mustangs. We also demonstrate that bidders are responsive to dynamic concerns, shading their bids in the current auction down when there are many upcoming auctions of identical or close substitute cars.

The paper is related to the literature on dynamic auctions and eBay. Jofre-Benet and Pesendorfer (2003) was the first paper to attack estimation in a dynamic auction game, though in a world with a small number of infinitely long-lived bidders. Subsequent to this, a number of papers have looked at dynamics on eBay specifically. Budish (2008) examines the optimality of eBay's market design with respect to the sequencing of sales and information revelation. 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're 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 completelely simultaneous. A second related literature is the large literature on demand estimation in durable goods markets (see for example Berry, Levinsohn, and Pakes (1995)). Finally, there is the academic literature on eBay itself — for an overview see Bajari and Hortacsu (2004).

The next section describes eBay and motivates the theoretical framework given in Section 3. Section 4 proves non-parametric identification. Section 5 describes our two different estimation approaches, while section 6 gives Monte Carlo simulations for those estimators. Section 7 is the application, and section 8 concludes.

2 Market Description

eBay is widely considered to be the world leader in online auctions. Various elements of its platform design, such as the use of proxy bidding agents, feedback scores and "buy-it-now" offers are widely copied.[4] In some sense, it is the dominant auction platform design. With that in mind, in this section we document the key features of the design, to motivate our theoretical model and estimation in later sections.

At any time, eBay hosts a large number of object listings from a variety of sellers. Buyers can browse these, either by navigating through categories delineated by the site, or by directly searching for key phrases. For example, in Figure 1 we show the results of picking the category "Chevrolet" and sub-category "Corvette" at a randomly chosen time of day. This brings up 1065 Corvette listings, here ordered by time until auction end.[5] These Corvettes are very heterogeneous, ranging from model year 1966 to 2009 on this page alone.

[4] See for example eBay competitors like eBid and uBid and Yahoo! Japan auctions.

[5] Last year eBay changed the default ordering to "best match", which is a proprietary search algorithm.

One might be tempted to think of these 1065 auctions as simultaneous, since they are all in progress at the time the search was conducted. However, these auctions all close at different times. For bidders who bid always late in the auction, say 5 minutes before close, this implies that the auctions are in fact almost *sequential*. Taking this to its logical limit, we can think of the market as divided into the "current auction" (the one closing right now now), the "upcoming auctions" (those closing later), and the "past auction" (those that closed before now). The set of upcoming auctions is continually evolving, as some auctions end and new listings are posted. At any point in time, buyers can see the current status of upcoming auctions that match any specific search term — for example, they could look at the car description, all past bids and current price in all auctions for 2005 Corvettes currently in progress on eBay. For this reason, buyers may condition their bids on the current auction on what else is out there (including the decision not to bid). Buyers can also look backwards, to see the closing prices and bid history of the auctions that closed in the last 20 days.

After having examined all the information available, a buyer may choose to place a bid in a particular auction. They are instructed to enter the maximum they are willing to pay for the item, and then eBay's proxy bidding system will bid up from the current standing price in standardized increments on their behalf until either their bid is the highest yet entered in the system, or an additional increment would take them over their maximum. For example, if bidder A enters a bid of $8000 on a vehicle where the standing price is $6000 and the highest bid placed in the system by a rival is $7000, then the system will update the standing price to $7100 ($7000 + $100 increment), and will record this bidder as the currently high bidder. On the other hand, if the highest bid placed were $9000, the price would jump to $8100 and the system would notify the bidder that he was outbid by a rival. Another key feature then is that high bidders are committed to the current auction, in the sense that if they bid in another auction and win both, they may end up with two objects — a bad idea in markets with unit demand and transactions costs.

Commitment can be costly, since bidders who wait can update on the auction state, and learn what other options are out there. Fear of premature commitment is perhaps one reason why bidders tend to bid late in auctions. On eBay Motors, 75% of bids are placed in the last day of the auction (Lewis 2007). Similar results have been found elsewhere on eBay and in other auction markets, and a wide range of alternative explanations for late bidding have been offered (see e.g. Roth and Ockenfels (2002)). At the close of the auction, the currently high bidder is the winner — provided the secret reserve has been met — and she pays the

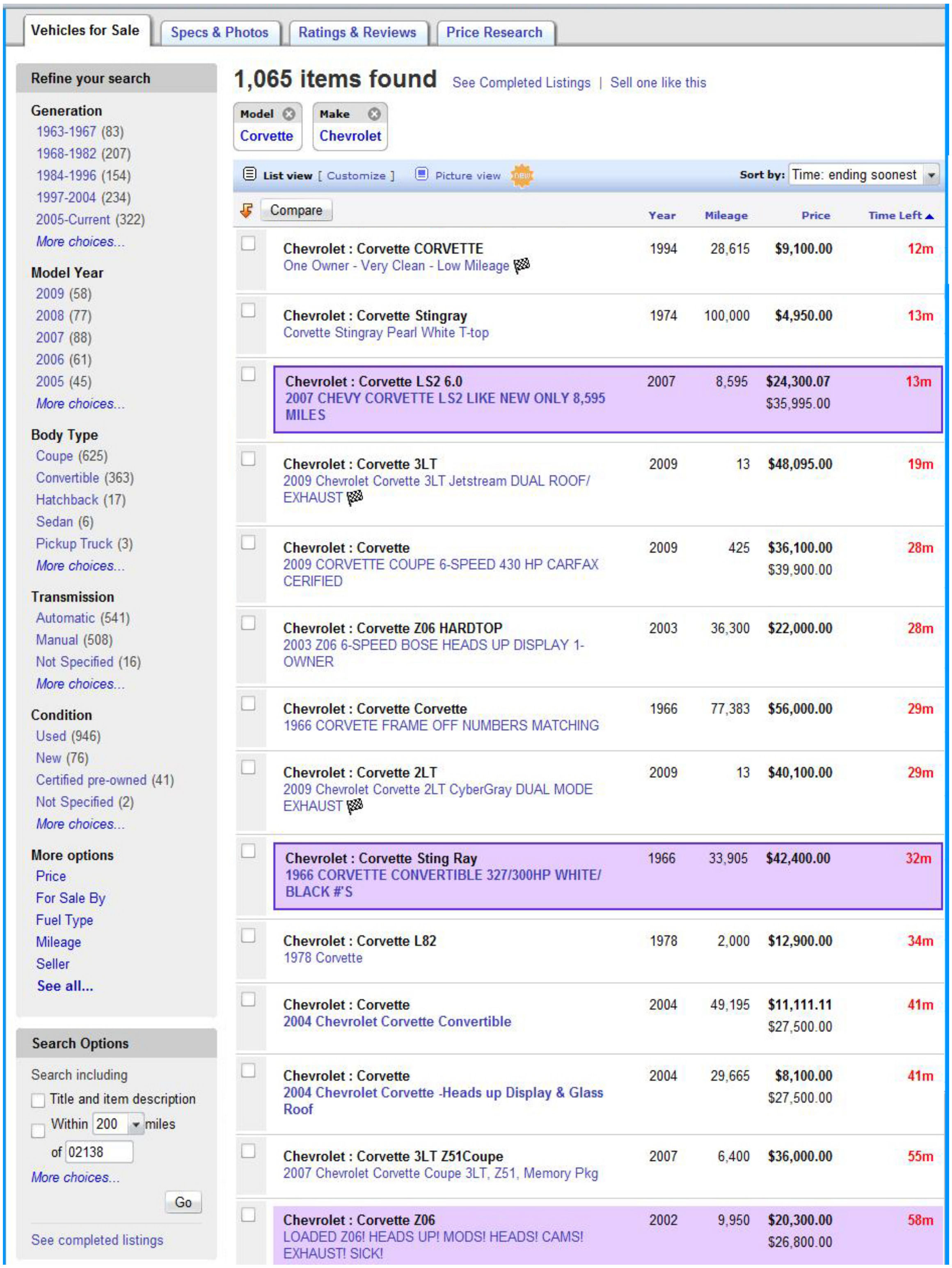

Figure 1: **Corvettes for sale on eBay** The figure shows the first page of the 1065 Corvette auctions open for bidding, at a randomly selected time of day.

current high bid. At that point, if the winner has unit demand, they will exit the market. The vast majority of bidders (over 99%) do so, suggesting that the unit demand assumption is reasonable. Losing bidders may continue to other auctions, or exit the market.

Overall then, the salient features of the market are that there are a large number of sequential auctions of heterogeneous objects, and a large number of potential bidders, all of whom can examine both the history of past auctions and look at the upcoming sequence of auctions before bidding; and a proxy bidding system that commits the current high bidder to the auction they have chosen.

3 Model and Dynamic Equilibrium

In this section, we formalize the operation of a generic auction platform market with unit demand as a dynamic game in discrete time. Every period, a second-price sealed bid auction may be held for a product, such as an iPod or a Zune. There are many different objects, and bidders have private valuations for each of the objects, so that bidder A, for example, may value an iPod much more highly than a Zune while bidder B may be indifferent. In addition, bidders are able to condition on a publicly observed state variable. For instance, they may be able to see which objects will be auctioned in subsequent periods and adjust their bids in anticipation of future supply. Following the auction the winning bidder and some of the losers exogenously exit the market, new bidders enter, and the public state changes.

There are many modeling choices embedded in this basic setup. Two of these are central and novel. First, we recognize that bidders in auction platform markets possess public and significant information about the state of the market when making their bids. They may be able to search for future upcoming auctions, as on eBay, or be aware of fluctuations in supply that influence the distribution of bidders today or the option value to staying in the market. Second, we allow for bidders to have private valuations that are correlated across objects. Taken together, these imply that the bid placed on object X today may depend on whether object Y, a close substitute, is being auctioned tomorrow. In the aggregate, prices and demand may depend heavily on the current state of the market.

This notion of the "state of the market" is an important one. It potentially summarizes the past, present and future: what happened in past auctions, what is being auctioned today, and what will be auctioned tomorrow. In some markets, bidders may not find it worthwhile to invest time in looking at future listings, or what happened yesterday. With that in mind, we will generally remain agnostic as to the nature of the state variable in the modeling below, except to assume that it is *public* (i.e. observed by all bidders) and *finite*.

Less importantly, we treat the proxy-bidding auction mechanism as a sealed-bid second price auction. This is a reasonable approximation given the late bidding, as noted in Bajari and Hortacsu (2004), and has become somewhat standard in this literature. The decision to work in discrete time is made purely to simplify the description of the game, which is as follows:

Bidders: Bidders have unit demand for a good in the set $\mathcal{A}$, where $|\mathcal{A}| = J$. They have a privately known vector of valuations x for the goods in $\mathcal{A}$, their type. They are risk neutral, and receive a payoff of $x_j - p$ for buying good j at price p. The payoff to losing in any period is normalized to zero. Bidders are impatient, with a common discount rate δ.

Market: Time is discrete with infinite horizon, $t = 1, 2 \cdots$. In any period t, the state of the market is described by a (vector-valued) variable $s_t \in S$, where S is compact. This state may capture information related to past supply or demand conditions, or encode information about future sequences of upcoming auctions. At a minimum it must identify the object currently under auction. Since the set of upcoming auctions is finite, and the state is assumed to depend on a finite history, the state variable is a finite vector.

Stage Game: In each period t, the following stage game is played. First, a sealed-bid second price auction is held for the current object a_t, in which all bidders participate. Then entry and exit of bidders takes place, suppliers post new objects, and the state variable is accordingly updated. These are described in more detail below.

Entry and Exit: At the end of every period, the winner is assumed to exit with certainty.[6] Losers exogenously exit the market with probability $\rho(s_t) \in (0, 1)$, receiving a payoff normalized to zero on exit. Simultaneously, $E_t \leq \overline{E}$ new bidders enter, where the distribution of E_t may depend on the state s_t. To ensure that the total number of participants in the market N_t doesn't explode, we assume $\exists\ \overline{N}$ such that no-one enters whenever $N_t \geq \overline{N}$.[7]. To ensure it doesn't permanently collapse, we assume that if $N < \overline{N}$, $\mathbb{P}(N_t \geq 2) > 0\ \forall s_t \in S$. Each entrant draws their valuation vector x identically and independently from a distribution F over a compact set $X \subseteq [0, \bar{\omega}]^J$.

[6] 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.

[7] This is not an assumption on the primitives; to get a more elegant assumption we would need to be more explicit in defining the state variable, which we want to avoid

Supply: Supply is essentially the rate at which different products appear on the auction market. With that in mind, we assume that at the end of period t, suppliers may list a new object with closing time $t+m$, where the auction duration m is fixed. The object is chosen according to a multinomial distribution over the discrete set $\mathcal{A} \cup \varnothing$ where $\varnothing$ is the event that no object is listed. The distribution may depend on the current state s_t, which allows for endogenous supply. This specification rules out the possibility of multiple items being listed in a single period. The idea here is to enforce auctions occurring in sequence, as we see online.

The assumptions made up to this point ensure that $\{s_t\}_{t=1}^{\infty}$ can be written as a first-order Markov Process.

Strategies and Beliefs: Bidders need to know who they're up against. Formally, in every period they should hold beliefs over the distribution of opposing types $\mathbf{x}_{-i}$. We restrict their beliefs to be measurable with respect to the current state s_t i.e. they believe that the state is sufficient for $\mathbf{x}_{-i}$. We also restrict bidders to symmetric stationary bidding strategies $\beta : X \times S \rightarrow \mathbb{R}^+$ that are measurable with respect to their private information and the state.

Equilibrium: We will look for a Markov Perfect Equilibrium (MPE) on the restricted strategy and belief spaces. This implies that the strategies must maximize payoffs given beliefs, and the beliefs must be consistent with equilibrium play. Usually, the MPE concept is used as an equilibrium refinement. Here, we have explicitly restricted strategies and beliefs: a player who was able to condition strategies or beliefs on the full history of the game may be able to improve his payoff.

We make these restrictions both because they simplify the problem and because they seem realistic. They simplify matters by making the game Markovian, allowing recursive representations of the bidder optimization problem. They also seem reasonable. While bidders could in principle learn from the whole history of the game, in practice the level of inference required seems daunting even to us. Instead, we substitute the idea that bidders might coordinate on some common features of the past, such as recent prices or participation, and form beliefs by asking what the steady-state distribution of opposing types is conditional on these statistics.

Now we characterize the equilibrium bidding strategies. First, we define the (ex-post) value function for a type x in state s:

$$v(x,s) = \max_b G_1(b|s)\left(x - E[B^1|B^1 < b]\right) + \delta(1 - G_1(b|s))(1-\rho(s)) \int v(x,s')t(s'|s,\mathbf{b})ds' \tag{1}$$

where $G_1(\cdot|s)$ is the distribution of the highest opposing bid today, and $t(s'|s,\mathbf{b})$ is the transition density of the state variable. In the special case where the state variables evolve exogenously, independent of the bids, we have $t(s'|s,\mathbf{b}) = t(s'|s)$. Now, fixing the opponent bids and the associated state transitions, we can take an FOC in the bidder problem to get the optimal strategies:

$$\begin{aligned} \beta(x,s) &= x_t - \widetilde{v}(x,s,b,b) + \frac{\partial}{\partial b} E\left[\widetilde{v}(x,s,b,B^1)|B^1 > b\right] \\ &\underbrace{=}_{\text{exogenous states}} x_t - \widetilde{v}(x,s) \end{aligned} \tag{2}$$

where x_t is the private valuation of the object auctioned at t, $\widetilde{v}(x,s,b,B) = \delta(1-\rho(s))E[v(x,s')|s,b,B^1]$ is the (ex-ante) continuation value of a type x who loses in period t in state s by bidding b against a higher winning bid of B^1. The partial derivative term is negative. In the special case of exogenous state transitions, we just write $\widetilde{v}(x,s)$.

To get some intuition for this result, 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 of course the weakly dominant strategy is to bid the value of the prize, which is just the value of the object less the continuation value. This is what we see in the case with exogenous state transitions.

But when state transitions are endogenous, the continuation value depends on how everyone in the auction bids. This has two implications. First, the bidder should try to avoid a winner's curse effect by recognizing that the "prize" may have a different value when he wins than when he loses. For example, if a high winning bid today leads to a state with soft competition tomorrow, then a bidder who bids aggressively today might regret marginally winning, because he has forgone the opportunity to take on a softer market tomorrow. Second, the bidder may try to manipulate tomorrow's state through his bid. Together these generate the first expression in (2).

To simplify the analysis in the rest of the paper, we introduce an assumption on the states:

Assumption 1 (**States**). *The states s are exogenous, finite and positive recurrent.*

Models which satisfy this assumption are those in which bidders bid based only on the identity or characteristics of the current object being auctioned, or where they look ahead a finite number of auctions. It rules out models in which bidders are backward-looking, trying to infer how competitive today's market is from yesterday's outcome. The exogenous states assumption is really used to simplify the theory; it implies that the stationary distribution of states is independent of the equilibrium strategies. We can do away with this in the estimation, which may be useful if learning is an important feature of the market. The finite states assumption is more important, since this underpins the nonparametric estimation approach outlined below.

An implication of the assumption is that a stationary distribution for the Markov process $\{s_t\}_{t=1}^{\infty}$ exists. However, in order to have well-defined beliefs, we need to also show that there is a stationary distribution of opposing types given any state and bidding strategies:

Lemma 1 (**Stationary Types**). *Given any bidding strategy β, a stationary distribution of types $\widetilde{F}|s$ exists for any $s \in S$.*

The proof relies on the fact that the distribution of types — both the number of participants and their valuations — also evolve as a Markov process. Then, taking a little more care since the type space is infinite, we can show that standard results imply the existence of a stationary distribution of types. Now, given well defined beliefs, we can show existence of a pure strategy MPE:

Theorem 1 (**Existence**). *A pure strategy MPE on the restricted strategy and belief spaces exists. Moreover, if the state is constant (i.e. $|S| = 1$), the equilibrium is unique.*

The proof is non-trivial. Proving existence requires overcoming the standard problem in auction settings that payoffs are not continuous in actions. Our approach is to use the characterization of the bidding strategies in (2) and to show the existence of a fixed point of the operator $\Gamma(\beta) = x_t - \widetilde{v}_\beta(x, s)$, where we are now explicit in noting the dependence of the value function on the strategies β. This is tricky because proving continuity of the operator requires proving continuity of the stationary type distributions in the strategies.

In the simple case where the state is constant, one can show this operator is a contraction mapping, and then simply apply the Banach fixed point theorem to obtain the result.

The key take home of this section is that 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. In the next section we develop nonparametric identification results for auction markets.

4 Nonparametric Identification

We give a nonparametric identification result in the spirit of Athey and Haile (2002). 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.

Theorem 2 (Identification). *If δ is known, and the econometrician observes the state, all bids and bidder identities generated by play under a fixed equilibrium, the valuation distribution F is non-parametrically identified. Moreover, the private valuation of any bidder observed bidding in every state $s \in S$ is identified.*

The idea of the proof is as follows. Call an individual bid vector "complete" if it includes a bid for every state $s \in S$. Not all bid vectors will be complete, due to exit, and the set of complete observations is a selected sample. Now suppose we could identify the valuations of those bidders with complete observations; and could also determine the probability that their observation was complete. Then by re-weighting the valuation density for complete bidders by the inverse of that probability, we would get the underlying valuation density.[8]

The key is showing that for complete observations there is an inversion from bids to valuations. Now it turns out that we can express their continuation values as the solution to a linear system of equations based on (1). The linear system has a unique solution, and then

[8] It is necessary that the probability of a complete observation is non-zero for all types; this is true on the interior of X.

from the vector of value functions we can deduce the valuations by re-arranging (2).[9] This is easiest to see in a simple example.

Example of Identification: There are two goods, so $\mathcal{A} = \{a_1, a_2\}$. The exit probability ρ is constant across states. Supply is binomial and independent of state, with q the probability of good 1. Bidders condition on the product identity in the current and next auction, implying four states: $1 = \{1,1\}$, $2 = \{1,2\}$, $3 = \{2,1\}$ and $4 = \{2,2\}$. Let the bids for a given bidder be $b_1 \cdots b_4$. Then the value function is:

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

where Q is the transition matrix between states. Substituting out x_i using the bidding function and rearranging gives:

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

Let $V = [v(x,1), v(x,2, v(x,3), v(x,4)]^T$, 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:

$$(I - \delta(1-\rho)Q)V = U$$

where ρ, Q and U can be estimated directly from the data. Then standard results imply $(I - \delta(1-\rho)Q)$ is invertible and therefore the existence of a unique solution for V. Then finally, to identify the valuation for good i, we look at the bid in states when good i was auctioned, and add the ex-ante continuation value to the bid:

$$x_1 = b_1 + \delta(1-\rho)\,(qV(1) + (1-q)V(2)) = b_2 + \delta(1-\rho)\,(qV(3) + (1-q)V(4))$$
$$x_2 = b_3 + \delta(1-\rho)\,(qV(1) + (1-q)V(2)) = b_4 + \delta(1-\rho)\,(qV(3) + (1-q)V(4))$$

[9] A similar argument was made in a different context by Pesendorfer and Schmidt-Dengler (2003).

Notice that the valuations are in fact *overidentified* in this case, where $|J| < |S|$. This provides a potential test of the theory.

5 Estimation Strategy

Suppose that the econometrician has a good dataset, observing all bids placed in each auction, the state variable and the identity of all bidders. 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 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 a semiparametric estimation approach based on simulated method of moments, as is used elsewhere for demand estimation in industrial organization and marketing. In that approach, we impose 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. In both cases, we proceed in two steps. The first step is identical for both approaches; the second varies depending on the method.

5.1 First Step: Transitions, exits and payments

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 either a kernel estimation approach or local linear regression.

5.2 Second Step: Nonparametric Approach

Drop incomplete observations, so that each observation now consists of an S-dimensional bid vector. For each complete observation i, use the first-stage estimates to construct a vector $\widehat{u}_i = [\widehat{u}_{i1}, \widehat{u}_{i2} \cdots \widehat{u}_{iS}]$, where $\widehat{u}_{is} = b_{is} - \widehat{\mathbb{E}}[B_1|B_1 < b, s]$. Then, as we show in the proof of Theorem 2, the value function for bidder i is $\widehat{v}_i = [\widehat{v}(x_i, 1), \widehat{v(x_i, 2)} \cdots \widehat{v}(x_i, S)]$ 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 $x_i = b_i + \delta(1-\rho)Tv_i$.

Now, the set of bidders with complete observations is a selected sample, and we need to correct for this. In the appendix, we show how to derive an expression for $P(A, x)$, the probability that a type x is observed bidding in the set of states in A, for $A \subseteq S$, in terms of previously estimated objects. Then we can estimate the type density $f(x)$ by kernel density estimation, taking the estimated $\widehat{x}_i$ as data and re-weighting by the inverse of $\widehat{P}(S, \widehat{x}_i)$.

5.3 Second Step: Semiparametric Approach:

The semiparametric approach proceeds in the opposite direction: instead of inverting bids to valuations, we take draws from a parameterized type distribution and simulate bids. The key to the simulation process is the value function representation in (1). Knowing $G_1(b|s)$, $\mathbb{E}[B_1|B_1 < b, s]$, $\rho(s)$ and Q, one could solve for the value function by iteration on the Bellman equation. This is computationally cheap since the iteration process obeys a contraction mapping. Since we don't know these objects, we replace them with our first-stage estimates. Then we can estimate the value functions $v(x, s)$ and via (2) obtain estimates of the bidding functions, $\beta(x, s)$.

We assume that the the type distribution has a known parametric form, so that the type density is $f(x; \theta)$. Then given any parameter θ we can sample a set of types x_i, and solve for their bids. We also solve for $P(A, x_i)$, and therefore get the probability that these bids appear in our sample for any subset of states A. Assuming we have some moment conditions $E[\gamma(y_i; \theta^0)] = 0$ formed from the bidder observations y_i and the predicted moments from the structural model, these moments can be simulated. We then search over the parameter space to find the parameter that minimizes a quadratic form in the simulated moment conditions.

If the chosen parametric distribution is multivariate normal, then there are natural candidates for the moment conditions. We can match the observed and simulated bids in each state, the square of those bids, and for bidders observed in multiple states, the product of bids across states. These correspond closely with the mean and variance-covariance of the type distribution.

This approach may be computationally demanding because for every new parameter update we must solve a dynamic optimization problem for a large sample of simulated types. We can speed this up using a Gibbs Sampling approach. 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 characteristics, 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 computed the simulated moments as weighted sums.

5.4 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 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 (e.g. McFadden (1974)).

To do this, we assume that valuations depend on the characteristics of the goods z_t, as well as on an idiosyncratic component:

$$x_{it} = z_t\beta_i + \epsilon_{it} \tag{4}$$

where we index individuals by i and auctions by t as before. The ϵ_{it} terms capture idiosyncratic preferences for the particular object being auctioned, while the β_i terms are the individual's type. This is pretty much a standard random coefficients demand specification, assuming no unobserved product heterogeneity.[10] Bidders differ in their tastes for the characteristics, and we would like to recover the distribution of the random coefficients. The characteristic space X has dimension $|X|$, where we assume $|X| \leq |J| \leq |S|$.

An implication of the characteristic-based approach is that the payoff-relevant state variables may change. For example, if consumers of digital cameras are assumed to have preferences only over the camera resolution, then if two different cameras may have the same mega-pixels, states in which those different cameras are up for auction may be pooled.

Since presumably this is a case where there are few complete observations, we take the SMM approach. The first stage is exactly as before: estimate transitions, exit and entry and per period payoffs just as before. In the second stage, the parametric specification must be of a (joint) distribution for the β_i (although, as often done elsewhere, they could be assumed independent), as well as a distribution for the idiosyncratic errors ϵ_{it}. Given this specification, one can simulate types and errors, which imply valuations x_{it} from (4), and through the first-stage estimates, bids. Basically, it's extremely similar to before, as well as being very much in the spirit of Berry, Levinsohn, and Pakes (1995).

[10] There are ways to deal with unobserved heterogeneity in this framework, but this would take us far afield.

6 Monte Carlo

We perform a simple Monte Carlo exercise to test both our nonparametric and SMM estimation approaches in small samples. 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 states. In each period, each of these goods is equally likely to be listed. The number of entrants N_t is always 3 each period, but exit is random with losing bidders exiting with probability $\rho = 0.25$, and winning bidders exiting with certainty. We start with 9 bidders. The discount rate δ is set to be 0.99.

Bidders have bivariate normal valuations for the goods, where X has mean $\mu = [100, 150]$ and covariance-matrix:

$$\Sigma = \begin{bmatrix} 100, 100 \\ 100, 400 \end{bmatrix}$$

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 converges in this case too. Then for each Monte Carlo iteration we simulate a dataset of 500 auctions, after discarding an initial 10 000 auctions as "burn-in". This amounts to 250 auctions per product, which seems like a moderate amount of data, especially given the volume of transactions online.[11]

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 1 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 second stage SMM approach, we (correctly) specify a multivariate normal distribution for the types, and try to match the mean and variance of bids on product 1, the mean and variance of bids on product 2, the covariance of bids for bidders observed bidding in both states, the fraction of bidders observed bidding only in state 1 and state 2 respectively, and the average duration of a bidder in the market. Since the number of moments (8) exceeds the number of parameters (5), we employ the identity matrix as the GMM weighting matrix.

[11] For comparison, we have 55000 auction observations in our car dataset below, with 15 models of classic cars, which is over 3000 auctions per product.

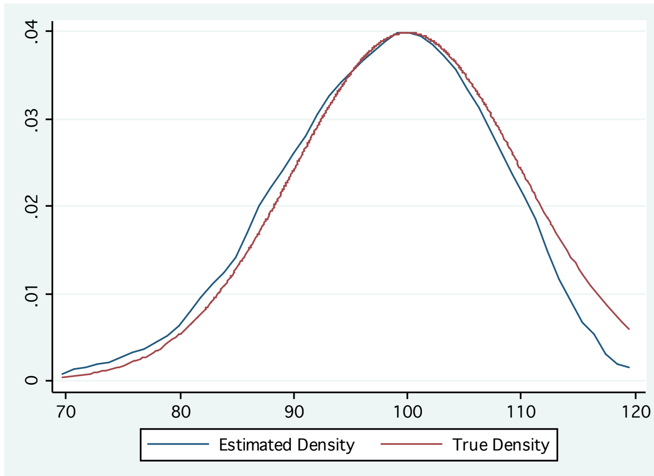

Figure 2: **Monte Carlo Simulation** The figure shows the true and estimated marginal density of valuations for product 1 from a randomly chosen Monte Carlo simulation of 500 auctions.

Our results are summarized in Table 1. As expected, we do extremely well in estimating the state transitions q and the probability of exit ρ, as these are just simple averages. More significantly, we do well on the non-parametric density estimates as well. The integrated squared error (ISE) of our marginal density estimates are on average 0.023 and 0.007 for products 1 and 2 respectively, which is small. Another way to see that we do a good job on the density estimate is to look at Figure 2, which shows the estimated and actual densities for a random iteration. The fit appears to be very good, although we have slightly underestimated the valuations (and in other samples will slightly overestimate them).

Turning to the parametric estimation results, we see that the SMM procedure estimates the mean valuations quite precisely. The covariance matrix is estimated with more error, although all estimates are still well within a standard deviation of the truth. Overall then, the nonparametric estimator seems preferable where it can be applied, but the SMM approach works well too.

7 Application: eBay Motors

7.1 Data

The main data source is a collection of auction webpages from completed used car auctions on eBay Motors. eBay Motors is the automobile arm of eBay, and the largest automotive site on the Internet, attracting 11 million unique visitors to its site each month[12] Every month, approximately 36000 vehicles are sold, a rate just slightly slower than a car a minute. A large share of those vehicles are classic sports-cars: Corvettes, Mustangs and Cameros.

We build a large dataset consisting of nearly 55000 auctions of these three car models. This data was obtained by downloading the auction webpages for these car models over an 8 month period, and then implementing a pattern matching algorithm to pull variables of interest from the webpage html code. We dropped observations with nonstandard or missing data; new or certified pre-owned cars, and cars under salvage title. This is done to rule out cars that are quite unlikely substitutes for the others in the market, such as those best sold for parts or those that were brand new.

Table 2 summarizes the variables in the dataset. In terms of car characteristics, the cars are old (20 years on average, including 1953 corvettes!); but have relatively low mileage for that age. Many have manual transmission. Few cars are sold under manufacturer warranty. The sellers are relatively experienced, with high feedback scores, and relatively high feedback percentage. The minimum bid is set relatively lower than the highest bid, and so most auctions receive bids, although only 26% of cars actually sell in any one auction, due to the use of secret reserves.

There are many auctions of these cars closing each day (on average 277 a day), so a buyer with relatively weak preferences across the model-years can substitute across auctions easily. Obviously, those with very precise preferences have fewer choices. In the data, we see that the average bidder bids in 2.7 auctions. Most bidders do not buy a car in the end, but 10.5% buy a single car and 0.7% buy more than one. Given that the fraction of buyers who buy more than 1 car is less than 1%, the unit demand assumption made earlier seems reasonable.

7.2 Substitution Patterns

One useful feature of auction market data is that some bidders are observed repeatedly. This allows us to get a basic sense of the extent to which substitution across products occurs simply by looking at the set of auctions a given bidder bids in. For example, if we observe that bidders who bid in an auction for a 1st generation mustang never bid in auctions for any other kind of car, we might conclude that there are no close substitutes for that car;

[12]Source: Nielsen//NetRatings.

whereas if many such bidders also bid on 2nd generation mustangs, we might think they were substitutes. This is useful for identifying demand, much in the same way as 2nd choice data is exploited in Berry, Levinsohn, and Pakes (2004).

With this in mind, we calculate a transition matrix across model-generations, as shown in Table 4. Each entry in the matrix is the probability of that a bidder who bids on the model-generation in the rows, subsequently bids on the model-generation in the columns, re-scaled so the probabilities sum to 1.[13] We find that most buyers that bid in multiple auctions tend to bid on the same car as the first time, but that the probabilities of that ranges between 50% (for a 2nd generation Camaro) to 75% (for a 1st generation Mustang). For most vehicles, there is a relatively close substitute that captures around 15% of the remaining bids. Usually, this substitute is a car of the same model and similar generation (either one older or newer). This provides some evidence that competition across vehicles (product competition) is important, as well as competition between similar products at different times (inter-temporal competition).

7.3 Auction Competition and Price Dynamics

Our next step is to try to quantify how big these effects are, in a reduced form way. One simple prediction is that a bidder considering a bid on a particular model of car, say model A, should reduce his bid if there are many other auctions for A currently being run on eBay that will close after the end of this auction. This comes straight from the theory: given the presence of many close substitutes, his continuation value is higher, and so he should reduce his bid accordingly. Similarly, if he regards model B as a good substitute for model A, when there are many auctions for B being run he should also bid lower.

To test this, we first create three new variables: own_{jt}, indicating the number of upcoming auctions of exactly the same model and generation of car j in the week following the closing of the auction; and $cross_{jt}$, indicating the number of upcoming auctions of the "closest substitute" car to j in that following week. To define which car was the closest substitute, we looked at the substitution matrix of Figure 4, and chose the vehicle which received the most bids from buyers whose initial bidding choice was the car they bid on. Lastly, we define $avgbidders_{jt}$, the average number of bidders on cars of type j in auctions in the *preceding*

[13] The re-scaling is necessary since some bidders bid on multiple cars, rather than just two.

week. This is one summary measure of the state of the market that bidders may use.[14]

We then ran the following hedonic regression:

$$\log b_{ijt} = x_{ijt}\beta + \alpha_1 \log own_{jt} + \alpha_2 \log cross_{jt} + \alpha_2 \log avgbidders_{jt} + \varepsilon_{ijt} \tag{5}$$

where x_{jt} is a vector of car characteristics for the product j sold in auction t, such as mileage and warranty status; and the terms own_{jt} and cross_{ijt} are defined as above. We use two slightly different specifications. In the first, we run it on only the winning bid, as it is hard to interpret the bids of lower ranked bidders as representing true valuations.[15] In the second we run this on all bids, recognizing the limitations of this approach, and compare the results from each.

The results are shown in Table 4. In columns (1) and (4), we include only own_{jt} and find significant and negative effects for both dependent variables: bidders shade their bids down by 0.15% for a 1% increase in the number of available upcoming auctions. Once we control for $cross_{jt}$ in (2) and (5), these effects weaken, but we find that for both identical and close substitutes, there is evidence of bid shading. Finally, in columns (3) and (6), we introduce $avgbidders_{jt}$ and find no significant evidence that this affects bidding strategies. Notice in particular that the coefficient shrinks considerably in the specification with all bids; the bigger coefficient in column (3) may just be driven by the fact that with more bidders, the highest order statistic is higher, rather than representing an actual shift in the way people bid for any valuation vectors. Overall then, it seems as though bidders are forward looking, and shade their bids down to account for their option value as the theory predicts. To back out their actual values then, it will be important to measure this latent option value. This will require a structural model, as detailed below.

7.4 Structural Estimation

[TO BE COMPLETED]

[14] We also tried average lagged prices, but these did not show up as even vaguely significant in any specification — they may be too noisy.

[15] For the argument, see Song (2004) or Lewis (2007).

8 Conclusion

[TO BE COMPLETED]

References

Adams, C. P. (2009): "Identification of the Joint Value Distribution from Auction Data," Working paper, FTC.

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., 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.

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

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574.

Haile, P. A., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *The Journal of Political Economy*, 111(1), 1–51.

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. Vickers.

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

Isaac, R. (1963): "A General Version of Doeblin's Condition," *Ann. Math. Statist.*, 34(2), 668–671.

Jofre-Benet, M., and M. Pesendorfer (2003): "Estimation of a Dynamic Auction Game," *Econometrica*, 71(5), 1443–1489.

Lewis, G. (2007): "Asymmetric Information, Online Disclosure and Adverse Selection: The Case of eBay Motors," Working paper, Harvard.

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

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.

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

Roth, A. E., and A. Ockenfels (2002): "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet," *American Economic Review*, 92(4), 1093–1103.

Rust, J. (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, ed. by R. Engle, and D.L.McFadden, vol. IV. Elsevier Science.

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

Song, U. (2004): "Nonparametric Estimation of an eBay Auction Model with an Unknown Number of Bidders," Working Paper, University of British Columbia.

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. V. Roy (2008): "Markov Perfect Industry Dynamics with Many Firms," *Econometrica*, 76(6), 1375–1411.

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

Appendix

Proof of Lemma 1: Let $y \in Y$ be the vector of types at any point, where $Y \equiv S \times \{*, [0, \bar{\omega}]\}^{J \cdot \bar{N}}$ and $*$ is a mass point and a placeholder for having fewer than $\bar{N}$ bidders in the market. Define for each set $A \in Y$, $l[A] = \{t : lim\ inf_n P^n(t, A) > 0\}$, where $P^n(t, A)$ is the probability of reaching A from t in n steps. Then from Isaac (1963) it suffices to prove existence of a stationary distribution that there is a probability measure m on the state space and a $\delta \in (0, 1)$ such that if $m(A) \geq \delta$, then $m(l[A]) > 0$.

This is trivially satisfied for our application because all points in the bidder type space are in the same recurrent class. To be precise, consider two arbitrary balls u, v in $[0, \bar{\omega}]^{m_U \cdot J}$ and $[0, \bar{\omega}]^{m_v \cdot J}$ respectively, for $m_u, m_v \leq \bar{N}$. To reach v from u in finite steps simply replace, each period, a bidder in u but not in v via the entry and exit process: with positive probability $\rho_s > 0$ each bidder exits, and bidders in v enter with positive probability as long as every type in v is in the support of F. Moreover, as exit is binomial, we can have more or less to adjust the number of bidders. As each of these events happens with positive probability and the number of bidders is finite, this demonstrates that we can reach any v from u in finitely many steps with positive probability. We have demonstrated that there are no transient states in type space and therefore $A \subseteq l[A]$ for A within the the ergodic class of public states and number of bidders. If $m(A) \geq \delta$, then $m(l[A]) > 0$. □

Proof of Theorem 1: If we can show a fixed point of the mapping $\Gamma(\beta) = x_t - \widetilde{v}_\beta(x, s)$, we are done; for if such strategies $\beta(x, s)$ exist, then from (2), they will constitute an equilibrium. Equivalently, we show a fixed point of the value functions under the mapping:

$$\Gamma(v(x, s)) = G_1(\beta(x, s))(x_s - \mathbb{E}[B_1 | B_1 \leq \beta(x, s), s]) + (1 - G_1(\beta(x, s)))\delta \sum_{t \in S} Q_{st} v(x, t)$$

where Q is the transition matrix. We look for a fixed point on the space of continuous functions on $[0, \bar{\omega}] \times S$ under the sup norm which is a Banach space. The Schauder fixed point theorem implies existence of a fixed point if the map Γ is continuous in $v(x, s)$. Notice that the stationary distribution of bids depends on the value function via (2), and so we establish continuity in two steps. First, we demonstrate weak convergence of the stationary distributions as $V_n(x, s)$ approaches some limit point $V_0(x, s)$ in the sup norm. Second, we use this to establish continuity of the distribution function of the first order statistic, which also yields continuity of the expectation.

Let $y \in Y \equiv S \times \{*, [0, \bar{\omega}]\}^{J \cdot \bar{N}}$ be an element of the full state space, consisting of the public state itself, the number of bidders, and their types. Let $T : Y \times Y \rightarrow [0, 1]$ be the transition distribution. Then the first claim is a straightforward application of Stokey, Lucas, and Prescott (1989) Theorem 12.13 if we can demonstrate that the transitions T on Y are converge weakly in distribution as $V_n \rightarrow V_0$. Consider the transitions pointwise; for a collection of bidders T sorts them according to $\beta(x, s)$ and the highest bidder exits, leaving us with the second through the Nth order statistic. This much is deterministic. The distribution of types the following period consists of these mass points each with probability $1 - \rho_s$ (accounting for exit) and additional entrants (with valuations given by F) the number of which is a random variable whose distribution may depend on the state. The transition operator T at a distribution μ over the full state space, then, is the μ-weighted integral over each of these distributions. Note that $V(x, s)$ appears in T only through $\beta(x, s)$, so it is sufficient to show the weak convergence of $\beta(x, s)$. Note moreover that T depends only on $\beta(x, s)$ to *rank* bidders.

We will use the fact that the stationary distribution over the full state space is continuous in bidder types. To see this, suppose otherwise and conjecture a mass point of active bidders at some point in type space. For this to be true, the *point* must be a member of the recurrent class, however this is impossible because the entry process assigns zero probability to point masses. Once that point in the type space is left (which eventually happens with probability one due to random exit), it is returned to with probability zero, and therefore it is transient, implying a contradiction. Thus because the stationary distribution over bidder types is continuous and the bidding function given by 2 is monotone, the distribution of the gap between the first and second order statistics of bids - call it $H(\cdot)$ - is continuous on $\mathbb{R}^+$.

This implies that for any $\epsilon > 0$, we can choose a $\gamma > 0$ such that $H(\gamma) < \epsilon$. Then let $\delta = \gamma/2$. So long as $|\beta_n - \beta_0| \leq \delta$ in the sup norm, the two bid functions predict the same winner - and therefore imply identical transitions T_n and T_0 - with probability $(1 - \epsilon)$. Taking ϵ to zero, we obtain weak convergence of T_n to T_0 as β_n goes to β_0.

Second, we obtain weak convergence of the density of the first order statistic by taking advantage of symmetry to write it down explicitly. For symmetric non-independent random variables,

$$g_Y(y_1,\ldots,y_n) = \begin{cases} n!f(y_1,\ldots,y_n) & \text{if } y_1 \geq y_2 \geq \ldots \geq y_n \\ 0 & \text{else} \end{cases}$$

As the transition function is independent of order, our stationary density is symmetric. Therefore weak convergence of the density of first order statistics follows from weak convergence in the parent density. However, weak convergence of the density implies continuity of the distribution, and therefore the Schauder fixed point theorem guarantees existence of an equilibrium. Finally, if $|S| = 1$, it is easy to show that the map $\gamma(\beta) = x - \widetilde{v}(x)$ is a contraction mapping; and then applying the Banach fixed point theorem instead of the Schauder theorem, we obtain the result. □

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 $u_s = G_1(b_s|s)(b_s - \mathbb{E}[B_1|B_1 < b_s, s])$. Q and ρ are identified from the data, δ is assumed known and u is identified for any bidder observed bidding in every state. From Rust (1994), $(I - \delta Q)^{-1}$ exists for $\delta < 1$. Given these objects, the ex-ante continuation value and therefore the valuation x is identified for all complete observations. Since every bidder on the interior of $[0, \bar{\omega}]^J$ loses with positive probability, all types in X are observed bidding in every state. Thus the density of valuations for complete observations $g(x)$ has full support on X. Moreover,

$$f(x) = \frac{g(x)}{P(S,x) \int \frac{g(x)}{P(s,x} dx}$$

where $P(S, x)$ is identified from the data and defined in (7) below. □

8.1 Additional Estimation Details

Let A be a subset of S. We want to 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 $B \subseteq S$. We can express this recursively:

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

where Q is the Markov transition matrix (recall, states are assumed finite). 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

Statistic	Truth	Mean Estimate	Std. Dev Estimates
		First Stage	
Product 1 listed tomorrow (q)	0.5	0.5018	0.0243
Probability of Exit (ρ)	0.25	0.2521	0.0070
		Second Stage: Nonparametric Approach	
Integrated Squared Error for product 1	—	0.0232	0.0046
Integrated Squared Error for product 2	—	0.0076	0.0014
		Second Stage: Parametric Approach	
Mean valuation for product 1	100	100.0426	0.2857
Mean valuation for product 2	150	150.0225	0.5186
Std deviation of valuation for product 1	100	100.6698	5.7778
Std deviation of valuation for product 2	400	407.3346	19.0271
Covariance in valuations	100	97.0895	10.9307

This table gives the outcomes of 100 monte carlo simulations from the model described in Section 6. The first two statistics are estimated in the first stage. The integrated squared errors for products 1 and 2 are defined as the sum of the square differences between the observed and estimated marginal valuation densities for products 1 and 2, and are outputs of the non-parametric procedure. The estimated means and variances-covariances come from the simulated method of moments procedure.

Table 2: Summary Statistics

	Mean	Std. Dev	Min	Max
Car Characteristics				
Miles	80718.451	97800.03	1	500000
Age (in years, relative to 2007)	20.369	14.31	0	54
Manual Transmission	0.401	0.49	0	1
Warranty	0.149	0.356	0	1
# of Options	4.701	4.731	0	25
Seller Characteristics				
Seller Feedback Score	181.056	800.174	-3	55207
Seller Feedback Percentage	98.476	5.739	0	100
Auction Characteristics and Outcomes				
Minimum Bid	6588.45	13119.825	0.01	1495000
Auctions with $>$ 1 bid	0.856	0.351	0	1
Fraction Sold	0.263	0.44	0	1
Highest Bid	14010.528	14698.928	0.99	860100
Auctions per day	277.801	64.858	165	460
Buyer Characteristics and Behavior				
# Auctions bid in	2.707	7.495	1	553
1 car bought	0.105	0.306	0	1
2 or more cars bought	0.007	0.084	0	1
N		54890		

This table provides summary statistics for the covariates used in the analysis.

Table 3: Substitution Patterns

Model		Corvette						Camero				Mustang			
Generation		1	2	3	4	5	6	1	2	3	4	1	2	3	4	5
Corvette	1	52.32	13.83	9.87	3.58	2.38	1.63	6.03	1.31	0.32	0.61	5.66	0.00	0.79	0.57	1.09
	2	6.67	54.72	12.36	3.37	3.10	1.80	7.91	1.15	0.71	0.66	5.57	0.09	0.39	0.52	1.01
	3	1.14	2.82	61.78	10.85	3.55	1.34	3.78	2.75	1.47	1.52	5.45	0.17	1.20	1.49	0.67
	4	0.32	0.93	13.93	56.23	8.68	1.40	1.65	1.47	2.13	3.59	2.68	0.18	1.80	3.86	1.16
	5	0.35	1.11	4.86	7.59	63.25	9.58	1.60	0.71	0.53	2.30	1.38	0.01	1.01	3.35	2.37
	6	0.37	1.14	2.99	3.24	14.93	65.10	2.01	0.29	0.53	1.36	1.34	0.00	0.60	2.00	4.09
Camaro	1	0.88	2.03	4.17	1.33	1.55	0.63	67.56	6.35	1.91	2.68	6.92	0.12	1.41	1.41	1.04
	2	0.18	0.49	5.14	2.60	1.01	0.15	13.54	50.03	7.77	4.56	7.29	0.52	3.97	2.18	0.55
	3	0.13	0.12	3.62	4.06	0.74	0.39	3.65	9.31	51.73	11.86	3.80	0.26	6.26	3.10	0.97
	4	0.18	0.34	2.88	4.45	2.62	1.27	3.58	3.57	8.26	54.85	2.16	0.06	3.91	10.16	1.71
Mustang	1	0.39	0.68	2.94	1.14	0.64	0.35	4.00	1.95	0.99	0.86	75.64	0.62	3.74	3.99	2.07
	2	0.00	0.00	2.02	0.89	0.06	0.40	1.53	4.33	2.23	1.29	17.72	45.28	15.51	6.57	2.15
	3	0.05	0.17	1.15	2.02	0.69	0.25	1.22	1.73	2.74	2.78	7.38	0.62	64.26	12.87	2.05
	4	0.05	0.11	1.34	2.17	2.03	0.60	1.09	0.94	1.50	4.70	5.31	0.27	9.26	60.66	9.97
	5	0.12	0.27	0.96	1.16	3.19	2.16	1.60	0.64	0.61	1.91	5.63	0.18	2.64	18.35	60.59

Substitution Patterns: Shows a transition matrix for buyers: the initial car bid on are the rows, the next car bid in are the columns, and each point in the matrix gives the probability of seeing that event in the data. The diagonal elements are large, indicating that most (but not all) times, a bidder who begins by bidding on a particular model-generation continues to do so in subsequent auctions.

Table 4: Bidding and Upcoming Auctions

	Log Bid					
	Highest bids only			All Bids		
	(1)	(2)	(3)	(4)	(5)	(6)
Log miles	-0.116**	-0.116**	-0.116**	-0.117**	-0.117**	-0.117**
	(0.0047)	(0.0047)	(0.0047)	(0.0047)	(0.0047)	(0.0047)
Manual transmission	0.232**	0.232**	0.231**	0.223**	0.223**	0.222**
	(0.0089)	(0.0089)	(0.0089)	(0.0102)	(0.0102)	(0.0103)
Log upcoming same auctions	-0.150**	-0.072**	-0.053	-0.154**	-0.070*	-0.068
	(0.0248)	(0.0278)	(0.0298)	(0.0308)	(0.0357)	(0.0380)
Log upcoming substitute auctions		-0.144**	-0.131**		-0.152**	-0.155**
		(0.0219)	(0.0226)		(0.0315)	(0.0333)
Log avg bidders last week			0.062			0.007
			(0.0322)			(0.0414)
Monday	-0.022	-0.021	-0.021	-0.004	-0.003	-0.002
	(0.0122)	(0.0122)	(0.0122)	(0.0149)	(0.0149)	(0.0149)
Tuesday	-0.019	-0.019	-0.018	-0.005	-0.004	-0.004
	(0.0123)	(0.0123)	(0.0125)	(0.0152)	(0.0152)	(0.0155)
Wednesday	-0.014	-0.013	-0.012	-0.011	-0.010	-0.008
	(0.0127)	(0.0127)	(0.0128)	(0.0153)	(0.0153)	(0.0155)
Thursday	-0.019	-0.018	-0.017	-0.023	-0.022	-0.021
	(0.0127)	(0.0127)	(0.0128)	(0.0157)	(0.0157)	(0.0160)
Friday	-0.033*	-0.032*	-0.031*	-0.026	-0.025	-0.024
	(0.0132)	(0.0132)	(0.0132)	(0.0164)	(0.0164)	(0.0164)
Saturday	-0.021	-0.021	-0.020	0.002	0.002	0.003
	(0.0127)	(0.0127)	(0.0127)	(0.0157)	(0.0157)	(0.0157)
Model-Year FE	yes	yes	yes	yes	yes	yes
R^2	0.6484	0.6487	0.6475	0.3751	0.3753	0.3736
N	36366	36366	35757	233642	233642	229665

Standard errors (clustered by seller) are given in parentheses. * denotes significance at 5%, ** at 1%. In all specifications, fixed effects for the car model-year are included. In (1)-(2), the dependent variable is log highest bid; in (3)-(4) it is the log bid, and all bids are included. The covariates of interest are the log upcoming same goods, which measures the log number of auctions of the same good upcoming in the next 7 days; and log upcoming substitute goods, which is the same measure for the closest substitute to the good being auctioned, as defined by the transition matrix above.