Estimating Demand from eBay Prices

Christopher P. Adams*

September 6, 2006

Abstract

This paper presents results for identification and estimation of the value distribution from eBay auction prices. The paper presents results for eBay type auctions with independent private values and unobserved participation. It is first shown that the distribution of values is identified from observing the distribution of prices and knowing the distribution of potential bidders. The main identification result presents conditions for which the distribution of values and the distribution of potential bidders are simultaneously identified. Not surprisingly, the intuition is similar to the standard results for identifying demand from observed equilibrium prices. The estimation method suggested by the identification results is used to estimate the value distribution for the "C5" Chevrolet Corvette sold on eBay. The results suggest that a simple OLS model on prices will over estimate the mean value of the item. The estimation results are then used to calculate the optimal reserve price for these cars. The estimated optimal reserves are compared to the actual reserves. Actual hidden reserves are set much higher than actual non-hidden reserves. The evidence suggests sellers set Buy-It-Nows and hidden reserves optimally to account for re-listing opportunities.

*Thanks to Patrick Bajari, George Deltas, Dan Hosken, Laura Hosken, Matt Lewis, Rob McMillan, Marc Rysman, Mike Sandfort, Dave Schmidt, Unjy Song, Steve Tenn, Bill Vogt, Abe Wilkelgren, Chuck Wood, Pai-Ling Yin, Robert Zeithammer, participants at the 2006 IIOC Meetings in Boston and the 2nd Statistical Challenges in E-Commerce Research Symposium and other FTC colleagues for helpful discussions. Thanks also to two anonymous referees. Thanks to eBay for providing the data. Thanks to Peter Newberry for excellent research assistance. Note that this paper does not necessarily represent the views of the Federal Trade Commission or any individual Commissioners. All remaining errors are my own. Email: cadams@ftc.gov

1 Introduction

EBay and eBay type auctions are an economic phenomena. EBay is fast becoming a major distribution channel in everything from Beanie Babies to Humvees (Lucking-Reiley (2000); Cohen (2002)).[1] The market place raises a lot of questions for retailers and manufacturers. For eBay retailers there is the question of where to set the reserve price. For manufacturers there are questions about the demand for various products and their features. EBay has also become a rich source of data for economists, other social scientists and even computer scientists (Bajari and Hortacsu (2004); Resnick et al. (2003)). Recently, eBay began making data available to the general public through APIs.[2] This paper shows how data from eBay can be used to answer questions of interest to retailers, manufacturers, economists and other social scientists. In particular the paper presents a new result which shows that the value distribution (demand) is identified from observed auction prices and the observation of some auction characteristic such as auction length. Using this result the paper estimates the demand for C5 Chevrolet Corvettes and these estimates are used to calculate optimal reserve prices. The paper finds that some sellers set their reserves at the appropriate level considering the dynamic programming problem posed by allowing sellers to relist unsold items. The result that (at least some) eBay sellers are optimally solving for dynamic programming problem in setting reserves and fixed prices contrasts with recent empirical work stating sellers do not set prices optimally (Levitt (2006)) and NFL coaches do not optimally solve dynamic programming problems (Romer (2006)).

EBay's basic auction format is something of a combination between an ascending bid auction (an English auction) and a second price sealed bid auction. All bidders may enter a "proxy" bid and the price remains equal to the second highest proxy bid.[3] Theoretically, bidders could choose to bid

[1] My colleague, Laura Hosken, even bought her wedding dress on eBay!

[2] Application program interfaces allow third part vendors to use eBay data and information.

[3] There are a number of circumstances in which the price differs from the second highest proxy bid. In particular, if there is only one bid above the reserve the price is set equal to

as if they are in standard English auction or they may enter their highest willingness to pay and let eBay's computers take care of the rest.[4] Here we will call the highest proxy bid entered by a particular bidder in a particular auction their "bid" and below we will show that this bid is equal the bidder's willingness to pay (unless it gets censored by auction price).

In 2003, eBay was selling over 1,000 cars per day and that number was growing at a rate of 20 per day. In 2003 there were hundreds of thousands of cars sold and millions of bidders using the market place. One of eBay's biggest sellers is the Chevrolet Corvette. The paper uses prices and car characteristics for the C5 Corvette. The C5 is the fifth version of the Corvette. This car was sold between 1997 and 2003. This car was chosen for the analysis because there are a relatively large number of auctions and the car is relatively homogenous. The data includes information on car characteristics such as the cars condition (new or used), mileage, model year, style (convertible) and car color.

The first section of the paper describes the basic model of behavior on eBay (due to Song (2005)).[5] Athey and Haile (2002) show that the value distribution is identified from the observed auction prices if the number of bidders in the auction is known and randomly determined.[6] Song (2005) points out while the number of actual bidders on eBay is known it is not randomly determined. In particular the current posted price leads to selection bias with the observed bidders having values greater than the current posted price. Song (2005) suggests that there is a set of "potential" bidders which are randomly determined but for which the number is unknown. Song (2005) presents an identification result for the case where the price and one other order statistic is observed but the number of potential bidders is unknown.[7]

---

the reserve.

[4]A bidder may enter the "minimum allowable bid" into the proxy bidding system, and in this way replicate bidding in an ascending bid English auction.

[5]See Bajari and Hortacsu (2003) for analysis of common value eBay auctions. Such auctions are probably a better description of collectible auctions.

[6]See Rezende (2002) for a straightforward estimation procedure for this case.

[7]Song (2005) shows that if we observe a second order statistic then we can use the Athey and Haile (2002) result to identify the value distribution that is conditional on the

This paper generalizes a result in Athey and Haile (2002) by showing the value distribution is identified from observed prices when the distribution of the number of potential bidders is known. While this may be of some technical interest, censoring in the data may make observation of this distribution difficult.[8] The paper's main result presents conditions for which the distribution of the number of potential bidders and the value distribution are simultaneously identified. The intuition comes from the standard result for estimating the demand side of the market from observing equilibrium prices. In that case it is necessary to have observable characteristics which vary with demand and not supply and observable characteristics which vary with supply and not demand. Similarly, here we need observable characteristics which vary with the number of bidders in the auction but not the value of the item. The trick, as always, is finding such instruments! This basic identification result is generalized to the case where there is observed and unobserved item heterogeneity.

The paper presents a parametric estimator based on the identification results. The paper assumes that the probability distribution over the number of potential bidders is a simple function of the auction's length and other observable characteristics of the auction. The estimator uses maximum likelihood techniques to estimate the parameters determining the relationship between the number of bidders and auction length as well as parameters of the value distribution. The value distribution is assumed to have a mean that is linear in observed item characteristics and a log-normal distribution.[9] This model is estimated on eBay price data for new and used C5 Chevrolet Corvettes. Comparing the results from the order statistics model with a standard OLS model the paper shows that the OLS model over estimates the mean value of the item by more than $20,000.

value being above the second order statistic, and from that the whole value distribution is identified. For example if we observe the second and third highest order statistics then we can identify the value distribution conditional on the value being above the third highest order statistic as we know the number of bidders with that conditional valuation, i.e. 2.

[8] It may be possible to use information such as page hits to estimate this distribution.

[9] See Giray et al. (2006) for alternative parametric specifications.

The results are used to calculate the optimal reserve which is compared to the actual reserves used. This is similar to the exercise in Haile and Tamer (2003) however this paper also analyzes the optimal reserve assuming the seller contemplates an infinite sequence of possible auctions in which to sell the car. In the case of the timber auctions analyzed in Haile and Tamer (2003) there is only one seller and one reserve price. Further the seller is a government agency. Here we have hundreds of private individuals and firms setting different reserves for different cars. The paper reports a large difference between the reserves set by those using hidden reserves and those using non-hidden reserves.[10] In particular hidden reserves are around $25,000 while non-hidden reserves are around $8,000. Economic theory also predicts two optimal reserves. If the seller is only considering one auction the estimated optimal reserve for these cars should be set at around $6,000.[11] If the seller considers the option to offer the car again in a week, the estimated optimal reserve for the C5 Corvettes is between $22,000 and $26,000. It may be coincidental but that optimal static reserve is similar to actual reserve used by non-hidden reserve sellers and the optimal dynamic reserve is similar to the actual reserve used by hidden reserve uses. Further, sellers that have both Buy-It-Nows and hidden reserves set their Buy-It-Nows almost exactly equal to the estimated optimal reserve determined by the dynamic programming problem. This dichotomy is consistent with eBay's own advice which states, "[No Reserve Price Auctions are] effective if you are looking to sell your vehicle in a hurry and getting the right price is less critical."[12]

The paper proceeds as follows. Section 2 presents the basic model. Section 3 presents the main non-parametric identification results. Section 4 presents a generalization to auctions with item heterogeneity. Section 5 presents a parametric estimator based on the identification results. Section 6 discusses the data used. Section 7 presents demand estimation results for

[10] EBay sellers must set a starting price and they also have the option of setting a hidden reserve. In auctions with a hidden reserve bidders are told whether or not the reserve has been met.

[11] Calculated using the estimates presented below.

[12] See "How to Sell a Vehicle" at http://www.pages.motors.ebay.com/howto/selling/listingB.html.

Chevrolet Corvettes. Section 8 derives and calculates the optimal reserves for C5 Corvette auctions based on the estimation results. The section compares the optimal reserves to the actual reserves used in these auctions. Section 9 concludes.

2 The Model and Notation

The model and notation closely follow Song (2005). It is a single eBay auction for a single item. There are N "potential" bidders and M observed bidders in the auction ($N \geq M$). As we are only going to consider auctions with two or more bidders (potential or observed) let $p_n = \Pr(N = n | N \geq 2)$.[13] The model is of a symmetric private information auction (Assumption 1), where each bidder knows the probability distribution over the number of bidders in the auction and the distribution from which the bidders draw their values (Assumption 2). Assumption 3 is made for simplicity.[14]

Assumption 1 *Each potential bidder's valuation V^i is an independent draw from $F(.)$, where $V^i \in [\underline{v}, \bar{v}]$.*

Assumption 2 *Each potential bidder knows p_n, $F(.)$ and their own value V^i.*

Assumption 3 *There is no minimum bid and there is no minimum increment.*

The auction lasts for the interval of time $[0, \tau]$ and each bidder is assumed to have a "last opportunity" to bid, although they don't have to bid at that "last opportunity" (Assumption 4). This last opportunity is drawn from a distribution (G) which may differ across bidders. Note that "last opportunity" refers to the last time the bidder can enter her proxy bid. This is a fairly general assumption as it does not require all bidders to have the

[13] Note that the highest two potential bidders are never censored in this model.

[14] This assumption is relaxed in the application. The paper presents an estimator that accounts for minimum bid requirements and other selection issues.

same last opportunity or even have it drawn from the same distribution. Evidence from a number of studies suggest there is a distribution of bid timing although the most bids come in the last minute of the auction (Bajari and Hortacsu (2003); Adams et al. (2006)).

Assumption 4 *Each potential bidder i is assumed to have a "last opportunity" to bid, $t^i \in [0, \tau]$, which is a random variable, such that the distribution of t^i is denoted $G^i(.)$.*

The following lemma due to Song (2005) states that bidders will (have) bid their value at their last opportunity if they are not censored.

Lemma 1 *(Song (2005)) Let C_t be the "cut off" price at time t, where $C_t = B_t^{(M-1:M)}$. Given Assumptions (1 - 4), in every Bayes Nash equilibrium, every bidder whose value for the item is greater than C_{t^i} at their last opportunity (t^i), will choose $B^i_{t^i} = V^i$ if they have not already done so.*

As eBay is a second price auction, the current price or "cut off" price is equal to the current second highest bid. It is straightforward to see that it must be optimal for each bidder to bid her value at her last opportunity. By assumption she has no chance to bid later than her last opportunity so it is a dominant strategy to bid her value.

In each auction, we assume that the amount of the lowest of the two highest bids, $B_\tau^{(M-1:M)}$ or the price, is observed. This price is denoted V_2. Note that from above, the price in an eBay auction equals the value of the *potential* bidder with the second highest value.

In regards to entry into the auction, that decision is endogenous in that only bidders with a positive expected value of entering will enter the auction. This doesn't really mean anything as the cost of entry for each bidder is either assumed to be 0 or infinity and is exogenously determined.[15] So

[15] Think of a bidder logging on to eBay at a particular date and time and either having an auction in which to bid (cost of entry is 0) or not having such an auction (cost of entry is infinity).

the probability distribution over the number of potential bidders (p_n) is determined exogenously. This assumption contrasts to the entry assumption in Bajari and Hortacsu (2003), who use endogenous entry and a zero-profit condition as part of their identification strategy. Assumption 5 states that the number of bidders in the auction (N) is independent of the values of bidders in that auction (V^i).

Assumption 5 *Let $N \perp V^i$ for all $V^i \in [\underline{v}, \bar{v}]$ and all $N \in \{1, 2, ...\}$.*

The next lemma generalizes the result that the value distribution is identified from the observed auction prices ($\{V_2\}$) when the number of bidders (N) is known (Athey and Haile (2002)). Here the lemma shows that the value is distribution is identified when the auction prices ($\{V_2\}$) are observed and the probability distribution over the number of bidders (p_n) is known.

Lemma 2 *Given Assumptions (1 - 5), if p_n is known for all $n \in \{2, 3, ...\}$ and $\{V_2\}$ is observed then $F(.)$ is identified.*

Proof. The proof has three steps. The first step defines an approximation of $F(.)$ denoted $F_K(.)$. In this step the set $[\underline{v}, \bar{v}]$ is discretized into a set of K elements $\{v_1, v_2, ..., v_K\} \subset [\underline{v}, \bar{v}]$. The second step shows $F_K(.)$ is identified. The third step shows that the approximation converges to $F(.)$ as $K \to \infty$.

Step 1. Let $[\underline{v}, \bar{v}]$ be segmented in to K disjoint sets of equal length such that the union is equal to the original set. Let $v_k = \underline{v} + \frac{k-1}{K}(\bar{v} - \underline{v})$ and

$$f_K(v_k) = \int_{v=v_k}^{v_{k+1}} f(v)dv \tag{1}$$

and

$$F_K(v_k) = \begin{cases} 0 & \text{if } k = 1 \\ \sum_{h=1}^{k-1} f_K(v_h) & \text{if } k > 1 \end{cases} \tag{2}$$

That is, we will approximate $[\underline{v}, \bar{v}]$ with a discrete set of K elements such that $v_k \in \{v_1, v_2, ..., v_K\}$ where $v_1 = \underline{v}$ and $v_K = \underline{v} + \frac{K-1}{K}(\bar{v} - \underline{v})$. Further, note the observed sequence of prices $\{V_2\}$ will also be approximated in this way, where $\{V_{2K}\}$ denotes the observed sequence from discrete set $\{v_1, v_2, ..., v_K\}$.

Step 2. This step is further broken down into three steps: the initial step, the induction step and the result.

Step 2.1. From Step (1), $F_K(v_1) = 0$ (see Equation (2)).

Step 2.2. This is the induction step. Let $F_K(v_k)$ be known. Consider the probability of observing a particular price $\{V_{2K}\}$ given Assumption (5) (Athey and Haile (2002)).

$$\Pr(V_{2K} = v_k | N \geq 2) = \sum_{n=2}^{\infty} p_n \frac{n! \Pr(v_k = V_{2K}) \Pr(v_k \geq V_{2K})(\Pr(v_k \leq V_{2K}))^{n-2}}{(n-2)!} \tag{3}$$

That is, given the discretization (Step (1)), the probability of observing a particular price equal to v_k in a second price auction is equal to the probability someone draws that value, someone else draws a value equal to or above it and all the other bidders have values equal to or less than the price. From Step (1), this can be rewritten as

$$\Pr(V_{2K} = v_k | N \geq 2) = \sum_{n=2}^{\infty} p_n \frac{n! f_K(v_k)(1 - F_K(v_k))(F_K(v_k) + f_K(v_k))^{n-2}}{(n-2)!} \tag{4}$$

Note that by assumption we have $F_K(v_k)$ is known. For notational convenience let $x_k = \Pr(V_{2K} = v_k | N \geq 2)$. That is x_k is the *observed* (large sample) probability of observing a price ($V_{2K} = v_k$), where we only consider those cases with at least two bidders. Given this notation we can rewrite the equation as

$$x_k = \sum_{n=2}^{\infty} p_n \frac{n! f_K(v_k)(1 - F_K(v_k))(F_K(v_k) + f_K(v_k))^{n-2}}{(n-2)!} \tag{5}$$

Let $h(f_K(v_k))$ equal the RHS of the equation. Given h is monotonic in $f_K(v_k)$ we have $f_K(v_k) = h^{-1}(x_k)$. As x_k, $F_K(v_k)$, p_n and n are known, $F_K(v_{k+1})$ is identified as $F_K(v_{k+1}) = F_K(v_k) + f_K(v_k)$.

Step 2.3. Given Steps (2.1) and (2.2), by induction $F_K(v_k)$ is identified for all $k \in \{1, 2, ..., K\}$.

Step 3. For a given $v \in [\underline{v}, \bar{v}]$ let $v_k \in \{v_1, v_2, ..., v_K\}$ be the closest to v in the set. From Equation (2) and Equation (1)

$$F_K(v_k) = \int_{\underline{v}}^{v_k} f(v) dv = F(v_k) \tag{6}$$

As $K \to \infty$, then $v_k \to v$ and thus $F(v_k) \to F(v)$ and so $F(.)$ is identified. Q.E.D.

Lemma 2 shows that if we know the probability distribution over the number of bidders then the underlying value distribution is identified. The result follows from breaking up the state space into countable segments and using induction to show that an approximation for $F(.)$ is identified.

This seems like a very useful result as it is much less restrictive than the result in Athey and Haile (2002). Unfortunately, it is not clear that it is possible to determine the distribution over the number of potential bidders given that bidders can have their existence censored. The next section shows that given certain data we can generalize this result to the case where the probability distribution is unknown. Note that there may also be other methods of estimating the distribution of potential bidders. For example, it may be possible to use site statistics such as "hits" on a particular auction.

3 Identification

The following assumption is critical to the main result of this section. Assume that auctions have some observable and numeric characteristic $t \in \mathcal{T}$, where $\mathcal{T}$ is a countable set with T elements. An example of $\mathcal{T}$ may be the set of auction lengths or the time of day the auction ends or an auction's length conditional on the time of day the auction ends.[16]

Assumption 6 *Let p_n and $F(.)$ have the following properties:*

1. $p_n(t) \neq p_n(t')$ *for all* $t \neq t' \in \mathcal{T}$ *and all* n.
2. $F(v|t) = F(v|t') = F(v)$ *for all* $t, t' \in \mathcal{T}$ *and* $v \in [\underline{v}, \bar{v}]$

[16] eBay auctions last for 3, 5, 7 or 10 days.

The assumption states that elements of $\mathcal{T}$ cause the distribution over the number of bidders to vary (for all n) but do not affect the distribution of bidder values. For example, as the number of days of the auction is increased this may lead to more bidders in the auction but it does not change the distribution of values for the item being sold.

The next assumption is a technical requirement on the probability distribution functions p_n. This assumption is sufficient for convergence of the Taylor Series used in the proof of the proposition.

Assumption 7 *Let $p_n(t)$ have the following properties:*

1. *$p_n(t) \in C^\infty$ for all $t \in \mathcal{T}$ and all n.*
2. *As $m \to \infty$, $p_n^{(m)}(t) \to 0$ for all $t \in \mathcal{T}$ and all n.*

The notation C^∞ refers to the fact that p_n is infinitely continuous (all derivatives are continuous) on $\mathcal{T}$. The second part of the assumption states that as the derivative index gets large the derivative gets small. There is no intuition for either condition, they are technical requirements for convergence of the Taylor Series approximation used in the proof. The following proposition shows that given these assumptions and the existence of the observed characteristics in $\mathcal{T}$ the distribution over the number of bidders is identified and thus the value distribution is identified from observed auction prices ($\{V_2\}$).

Proposition 1 *Given that Assumptions (1 - 7) hold and $\mathcal{T}$ and $\{V_2\}$ are observed, as $T \to \infty$, $F(.)$ is identified.*

Proof. The proof has five steps. The first step repeats the first step in Lemma 2 to create an approximation $F_K(.)$ of $F(.)$. The second step defines the main equation. The third step shows that when T is large enough, a polynomial approximation of $p_n(t)$ is identified for all n less than some $\bar{N}$ and all $t \in \mathcal{T}$. The fourth step shows that as T gets large $p_n(t)$ is identified for all n and all $t \in \mathcal{T}$. The fifth step invokes Lemma 2 which states that if

$p_n(t)$ is identified for all n and some t, then $F(.)$ is identified.

Step 1. Let $[\underline{v}, \bar{v}]$ be segmented in to K disjoint sets of equal length such that the union is equal to the original set. Let $v_k = \underline{v} + \frac{k-1}{K}(\bar{v} - \underline{v})$ and

$$f_K(v_k) = \int_{v=v_k}^{v_{k+1}} f(v)dv \tag{7}$$

and

$$F_K(v_k) = \begin{cases} 0 & \text{if } k = 1 \\ \sum_{h=1}^{k-1} f_K(v_h) & \text{if } k > 1 \end{cases} \tag{8}$$

That is, we will approximate $[\underline{v}, \bar{v}]$ with a discrete set of K elements such that $v_k \in \{v_1, v_2, ..., v_K\}$ where $v_1 = \underline{v}$ and $v_K = \underline{v} + \frac{K-1}{K}(\bar{v} - \underline{v})$. Further, note the observed sequence of prices $\{V_2\}$ will also be approximated in this way, where $\{V_{2K}\}$ denotes the observed sequence from discrete set $\{v_1, v_2, ..., v_K\}$.

Step 2. Given the discretization consider the probability of observing a price $V_{2K} = v_1$,

$$\Pr(V_{2K} = v_1 | N \geq 2) = \sum_{n=2}^{\infty} p_n \frac{n! \Pr(v_k = v_1) \Pr(v_k \geq v_1)(\Pr(v_k \leq v_1)^{n-2}}{(n-2)!} \tag{9}$$

Noting $\Pr(v_k \geq v_1) = 1$ and $\Pr(v_k \leq v_1) = \Pr(v_k = v_1)$, and letting $x_1 = \Pr(V_{2K} = v_1 | N \geq 2)$, from Equation (1) and (2) this equation can be rewritten as

$$x_1 = \sum_{n=2}^{\infty} p_n \frac{n! f_K(v_1)^{n-1}}{(n-2)!} \tag{10}$$

Step 3. Let us approximate the distribution over the number of bidders such that there exists an $\bar{N}$ that is the largest possible number of potential bidders. Let this approximation be denoted $p_{n\bar{N}}$. That is $p_{n\bar{N}} = 0$ for all $n > \bar{N}$. Given that $N(N-1)f_K(v)^{N-1} \to 0$ as $N \to \infty$ for large $\bar{N}$,[17] we can rewrite Equation (10)

$$x_1 = \sum_{n=2}^{\bar{N}} p_{n\bar{N}} \frac{n! f_K(v_1)^{n-1}}{(n-2)!} \tag{11}$$

[17] $N(N-1)f_K(v)^{N-1} = N^2 f_K(v)^{N-1} - N f_K(v)^{N-1}$ and $N^a x^N = \exp\{a \ln(N) + N \ln(x)\}$ and $\frac{da \ln(N)}{dN} = \frac{a}{N}$ and $\frac{dN \ln(x)}{dN} = \ln(x)$, for large N and $x < 1$, $a \ln(N) + N \ln(x) \to -\infty$ and $N^a x^N \to 0$.

By Assumption 6 we have

$$x_1(t) = \sum_{n=2}^{\bar{N}} p_{n\bar{N}}(t) \frac{n! f_K(v_1)^{n-1}}{(n-2)!} \tag{12}$$

For any $t \in \mathcal{T}$. Further let there exists a Jth order polynomial denoted $p_{n\bar{N}J}(t) = q_{n0} + tq_{n1} + t^2 q_{n2} + ... + t^J q_{nJ}$ that approximates $p_{n\bar{N}}(t)$ for all $t \in \mathcal{T}$ and $n \leq \bar{N}$. Replacing $p_{n\bar{N}}(t)$ with $p_{n\bar{N}J}(t)$, by Assumption 7 for large J and large $\bar{N}$ we have

$$x_1(t) = \sum_{n=2}^{\bar{N}} (q_{n0} + tq_{n1} + t^2 q_{n2} + ... + t^J q_{nJ}) \frac{n! f_K(v_1)^{n-1}}{(n-2)!} \tag{13}$$

By adding up we can write

$$\begin{aligned} x_1(t) &= 2(1 - \textstyle\sum_{n=3}^{\bar{N}} (q_{n0} + tq_{n1} + t^2 q_{n2} + ... + t^J q_{nJ})) f_K(v_1) \\ &+ \textstyle\sum_{n=3}^{\bar{N}} (q_{n0} + tq_{n1} + t^2 q_{n2} + ... + t^J q_{nJ}) \frac{n! f_K(v_1)^{n-1}}{(n-2)!} \end{aligned} \tag{14}$$

rearranging we have

$$x_1(t) = 2 f_K(v_1) + \sum_{n=3}^{\bar{N}} \sum_{j=0}^{J} q_{nj} t^j n(n-1) \left(f_K(v_1)^{n-1} - \frac{2}{n(n-1)} f_K(v_1) \right) \tag{15}$$

Let $T = 1 + (\bar{N} - 2) \times (J + 1)$ and we can write out a system of equations for different observations ($t \in \mathcal{T}$)

$$\begin{aligned} &x_1(t_1) = 2 f_K(v_1) + \textstyle\sum_{n=3}^{\bar{N}} \sum_{j=0}^{J} q_{nj} t_1^j n(n-1) \left(f_K(v_1)^{n-1} - \frac{2}{n(n-1)} f_K(v_1) \right) \\ &... \\ &x_1(t_T) = 2 f_K(v_1) + \textstyle\sum_{n=3}^{\bar{N}} \sum_{j=0}^{J} q_{nj} t_T^j n(n-1) \left(f_K(v_1)^{n-1} - \frac{2}{n(n-1)} f_K(v_1) \right) \end{aligned} \tag{16}$$

In matrix form this can be written as $X = AQ$ where

$$X = \begin{bmatrix} x_1(t_1) \\ x_1(t_2) \\ ... \\ x_1(t_T) \end{bmatrix} \tag{17}$$

and is a $T \times 1$ matrix,

$$A = \begin{bmatrix} 2 & 6 & 6t_1 & ... & 6t_1^J & ... & \bar{N}(\bar{N}-1) & \bar{N}(\bar{N}-1)t_1 & ... & \bar{N}(\bar{N}-1)t_1^J \\ ... \\ 2 & 6 & 6t_T & ... & 6t_T^J & ... & \bar{N}(\bar{N}-1) & \bar{N}(\bar{N}-1)t_T & ... & \bar{N}(\bar{N}-1)t_T^J \end{bmatrix} \quad (18)$$

and is a $T \times T$ matrix, and

$$Q = \begin{bmatrix} f_K(v_1) \\ q_{30}(f_K(v_1)^2 - \frac{1}{3}f_K(v_1)) \\ q_{31}(f_K(v_1)^2 - \frac{1}{3}f_K(v_1)) \\ ... \\ q_{3J}(f_K(v_1)^2 - \frac{1}{3}f_K(v_1)) \\ ... \\ q_{\bar{N}0}\left(f_K(v_1)^{\bar{N}-1} - \frac{2}{(\bar{N}(\bar{N}-1))}f_K(v_1)\right) \\ ... \\ q_{\bar{N}J}\left(f_K(v_1)^{\bar{N}-1} - \frac{2}{(\bar{N}(\bar{N}-1))}f_K(v_1)\right) \end{bmatrix} \quad (19)$$

which is a $T \times 1$ matrix. So we have that $Q = A^{-1}X$. From this we see that $Q_1 = f_K(v_1) = [A^{-1}X]_1$ and then value of $f_K(v_1)$ can be used to solve for the q_{nj}'s. Therefore for a given J and $\bar{N}$ and $T = 1 + (\bar{N} - 2) \times (J + 1)$, $p_{n\bar{N}J}(t)$ is identified for all $n \leq \bar{N}$ and all $t \in \mathcal{T}$.

Step 4. Letting $J \rightarrow \infty$ and by Assumption 7, $p_{n\bar{N}J}(t) \rightarrow p_n\bar{N}(t)$ for all $t \in \mathcal{T}$ and all n. Letting $\bar{N} \rightarrow \infty$, $p_{n\bar{N}}(t) \rightarrow p_n(t)$ is identified for all n and $t \in \mathcal{T}$.

Step 5. As $T \rightarrow \infty$, $p_n(t)$ is identified for all n and any $t \in \mathcal{T}$, by Lemma 2, $F(.)$ is identified. Q.E.D.

The proposition states that as the number of states (T) get arbitrarily large then the probability distribution over the number of bidders can be approximated arbitrarily closely and thus (from Lemma 2) the value distribution is identified. The proof is based on the intuition that varying auction length, the probability that there are n bidders varies but the value distribution does not. So any observed changes in price is coming from changes in the number of bidders. As the distribution is non-parametric the richer the variation the more detailed the description of the distribution. Unfortunately, auction length (at least on eBay) is not continuous and in fact only has 4 states. However, the number of bidders varies exogenously over the time of day that the auction ends then we can condition auction length on the time of day and generate a much larger number of states.[18] The researcher needs to be careful to choose a set of observable states for which the number of bidders varies exogenously and the value distribution does not vary.

4 Auction Heterogeneity

Traditionally the demand for differentiated products is estimated by assuming that product choices can be mapped into observed and unobserved product characteristics and using hedonic regression models (Berry et al. (1995)). The major argument for doing this is that there are often so many different products that it is not possible to identify the demand for each product without using information about the demand for similar products. As discussed below part of the identification strategy in this literature is to use this variation in product characteristics. The following results suggest that a similar model can be used with eBay type data.

Assumption 8 *Let V_j^i be distributed $F(., X_j)$ where X_j is a J dimensional vector of observed item characteristics.*

Assumption 8 states that the value distribution is some general function of the set of observed characteristics of the item X_j. An example would be a random coefficients model (Berry et al. (1995); Nevo (2000)). Note that in Berry et al. (1995) and more recently Bajari and Benkard (2005) there is a unobserved component of the product characteristics, this issue is discussed below.

Corollary 1 *If Assumptions (1 - 7) and Assumption 8 hold, then if $\{V_2\}$, $\mathcal{T}$ and X_j are observed for all items j, then $F(., X_j)$ is identified.*

[18] eBay provides the time that the auction ends to the second.

Proof. By Proposition 1, given X_j, $F(.|X_j)$ is identified. Therefore, for all X_j, $F(., X_j)$ is identified. Q.E.D.

Corollary 1 shows that it is straightforward to generalize Proposition 1 to a hedonic model. Note, that for simplicity it is assumed that the distribution of bidders is not a function of the item characteristics. It seems reasonable that this assumption can be relaxed although one needs to remain aware of exclusion restrictions necessary for identification. The rest of the section considers a model with unobserved item heterogeneity.

Assumption 9 *Let $V_j^i = v_{ij} + \xi_j$ where v_{ij} is distributed $F(., X_j)$ and $\xi_j \in [\underline{\xi}, \bar{\xi}]$, X_j is a J dimensional vector of observable characteristics of the item, and ξ_j is a characteristic of the item observed by the bidder and unobserved by the researcher.*

Note that ξ_j is constant across auctions for the same item. That is it represents some unobserved characteristic about the car or MP3 player rather than some unobserved characteristic that is auction specific.[19] An important simplification in Assumption 9 is that unobserved item heterogeneity enters the value function additively. This assumption makes identification straightforward, but one may be concerned that it is unnecessarily simple (Bajari and Benkard (2005)).

Proposition 2 *If Assumptions (1 - 7) and Assumption 9 hold, then if $\{V_2\}$, T and X_j are observed for all items j, then $F(., X_j)$ and ξ_j are identified.*

Proof. Step 1. Let $G(., X_j)$ be the distribution of $v_{ij} + \xi_j$. For a given item j by Corollary 1, $G_j(., X_j)$ is identified. Step 2. Let there be two items j and k such that $X_j = X_k$, then for some $a > b$ such that $G_j(a, X_j) = G_k(b, X_j)$, $a - b = \xi_j - \xi_k$. When this difference is equal to $\bar{\xi} - \underline{\xi}$, $\xi_j = \bar{\xi}$ and $\xi_k = \underline{\xi}$, and so $F(., X_j)$ and ξ_j are identified. Q.E.D.

[19] Adams (2004) suggests that it is necessary to observe individual bidders across auctions in order to account for unobserved auction heterogeneity.

Proposition 2 shows that Corollary 1 can be generalized to the case of unobserved item heterogeneity. Identification comes from comparing the distributions for items with similar observed characteristics.

5 Estimation Model

This section presents a parametric maximum likelihood estimator based on the results presented in the previous section. The main distributional assumption is that the error structure on the item's value is log-normal. The probability distribution over the number of potential bidders is assumed to be a function of auction length. The structural assumption is that the bids of the two highest bidders are equal to each bidder's value for the item. Further, for bidders that bid below the final price, their value for the item is less than the final price (Haile and Tamer (2003)).

Let bidder i's value function for item j be represented by the following function

$$\ln(v_{ij}) = \beta \ln(X_j) + \epsilon_{ij} \tag{20}$$

where v_{ij} is i's value for item j, X_j is a set of observable characteristics of item j, and ϵ_{ij} are unobservable characteristics of the bidder and the item. Let ϵ_{ij} be distributed $N(0, \sigma)$. The structural assumption gives $\ln(b_{ij}) = \ln(v_{ij})$. From above the distribution of prices is given by (assuming two or more potential bidders)

$$f_{V_2}(V_2|N \geq 2) = f(V_2)(1 - F(V_2)) \left(\sum_{n=2}^{\infty} p_n \frac{n! F^{n-2}(V_2)}{(n-2)!} \right) \tag{21}$$

where V_2 is the auction ending price and n is the potential number of bidders. Given the restrictions above the likelihood function of a set of observed prices is

$$\ln L = \sum_{k=1}^{K} \left(-\ln(\sigma) + \ln(\phi(z_k)) + \ln(1 - \Phi(z_k)) + \ln \left(\sum_{n=2}^{\infty} p_n \frac{n! \Phi^{n-2}(z_k)}{(n-2)!} \right) \right) \tag{22}$$

where $z_k = \frac{\ln(V_2) - \beta \ln(X_j)}{\sigma}$.

In this case our data is further censored by the existence of starting prices and hidden reserve prices.[20] We only observe auctions where the price met the highest of these two.[21] For simplicity the analysis further restricts the data to the case where the two highest bids are above the highest of the starting price or the secret reserve price. Given this restriction, the likelihood function becomes

$$\begin{aligned}\ln L &= \textstyle\sum_{k=1}^{K} - \ln(\sigma) - \ln(1 - \Phi(z_{kc})) + \ln(\phi(z_k)) + \ln(1 - \Phi(z_k)) \\ &+ \ln\left(\textstyle\sum_{n=2}^{\infty} p_n \frac{n!\Phi^{n-2}(z_K)}{(n-2)!}\right)\end{aligned} \tag{23}$$

where $z_{kc} = \frac{\ln(S)-\beta\ln(X_j)}{\sigma}$ and S is the highest of the starting price and the secret reserve price.

In the actual implementation things are simplified further by assuming that there are either 5 or 12 potential bidders, where the

$$\Pr(N = 12|X_j) = \frac{\exp\{\gamma_0 + \gamma_1 T_j\}}{1 + \exp\{\gamma_0 + \gamma_1 T_j\}} \tag{24}$$

where $\gamma T_j = \gamma_0 + \gamma_1 T_j + \gamma_T X_j$ and T_j is the length of auction j. This assumption is made in order to make the implementation and estimation simpler. The cost is that the model is less likely to provide accurate estimates.[22] Given this assumption the actual likelihood function that is estimated is

$$\begin{aligned}\ln L &= \textstyle\sum_{k=1}^{K} - \ln(\sigma) + \ln(1 - \Phi(z_{kc})) + \ln(\phi(z_k)) + \ln(1 - \Phi(z_k)) \\ &+ \ln\left(\frac{1}{1+\exp\{\gamma T_k\}} 20\Phi^3(z_k) + \frac{\exp\{\gamma T_k\}}{1+\exp\{\gamma T_k\}} 121\Phi^{11}(z_k)\right)\end{aligned} \tag{25}$$

As suggested above, identification relies on exclusion restrictions. In particular auction length is assumed to affect the number of bidders but not how bidders value the item. Note other observable characteristics may affect both the number of observed bidders and how those bidders value the item.

[20] Assume that no hidden reserve corresponds to a hidden reserve equal to 0.

[21] This was due to a miscommunication between myself and eBay.

[22] It would be straightforward but more cumbersome to allow alternative parameterizations including the Poisson distribution (Giray et al. (2006)).

6 Data

The paper uses prices on new and used C5 Corvettes sold on eBay in 2003 for which we have a valid Vehicle Identification Number (VIN).[23] VINs and car characteristics like make, model, model year, mileage and color are entered into the eBay Motors database by the seller. The data is not always accurately entered and VINs are not always provided. A valid VIN gives information about the car including make, model, year, engine type, and model style. By using only observations with a valid VIN we can corroborate information about the car that the seller includes on the eBay site, and eliminate those observations with conflicting information.

The analysis uses information from 705 C5 Corvette auctions. The C5 is the fifth version of the Corvette. This restriction allows the analysis to concentrate on a version of the Corvette that is relatively new and relatively homogenous. The original data covers three years of auctions but the majority occurred in 2003. Again in order to reduce the amount of unobserved heterogeniety the data set was restricted to only those auctions that occurred in 2003.

The data includes only "successful" auctions which are auctions in which there was at least one bidder above the starting price or the reserve price (if there was one). I further selected the data to include just auctions where there were at least two bidders above the starting price or the reserve price. Auctions in which the car sold with the Buy-It-Now option were dropped. The model estimated below accounts for selection of auctions above the starting price (or reserve price) but does not account for the use of the Buy-It-Now option.

Table 1 presents summary statistics of the C5 Corvettes in the sample. Used Corvettes account for 97% of the C5 Corvettes sold on eBaymotors.com. Here, we use the seller entered condition which is either new or used. The average C5 has 39,000 miles and just under four years old. Note that the mileage information is seller entered and we drop observations above 306,000 miles, all of which have nonsensical entries such as 999999999. In regards to

[23]See Adams et al. (2006) for more on Corvettes sold on eBay.

Variable	
New (Percentage)	3
Mileage (Mean in Miles)	38,584
Age (Mean in Years)	4
Automatic (Percentage)	60
Convertible (Percentage)	33
Color Red (Percentage)	20
Color Black (Percentage)	23
Number of Observations	705

Table 1: Summary Statistics of Car Characteristics for C5 Corvettes Sold in 2003

other characteristics: 60% are automatic, one third are convertibles, 20% are red Corvettes and 23% are black.

Table 2 provides summary statistics for the observable characteristics of the auctions. The average price for a C5 Corvette sold on eBay in 2003 is close to \$26,500. The average starting price is \$7,463. Around one third of the auctions include a hidden reserve and the average hidden reserve is \$25,000. Sellers of cars have an average feedback score of 245 and only 5% have negative scores. Note that this number is calculated as of a particular date in March 2004. The mode auction lasts for seven days and most auctions occur in the second half of the year.

7 Estimation Results

Table 3 presents results from OLS on the price data and results from the Order Statistic approach with an instrument for the number of bidders. The columns headed with β are the point estimates for the model parameters associated with the observed characteristics. The standard errors are presented in the column headed "SE".[24] The Order IV column presents results

[24] These are simple standard errors.

Variable	
Price (Mean Dollars)	26,455
Starting Price (Mean Dollars)	7,463
Hidden Reserve (Percentage)	34
Hidden Reserve (Mean Dollars (if positive))	24,984
Seller Feedback Negative as of March 2004 (Percentage)	5
Seller Feedback as of March 2004 (Mean Aggregate Score)	245
Auction Length - 3 Days (Percentage)	11
Auction Length - 5 Days (Percentage)	22
Auction Length - 7 Days (Percentage)	47
Auction Length - 10 Days (Percentage)	19
First Quarter 2003 (Percentage)	16
Second Quarter 2003 (Percentage)	28
Third Quarter 2003 (Percentage)	28
Fourth Quarter 2003 (Percentage)	25
Number of Observations	705

Table 2: Summary Statistics of Auction Characteristics for C5 Corvettes Sold in 2003

from the empirical model presented above (Equation (25)). The model is estimated with maximum likelihood and includes two equations and an estimate of the variance (σ). The first equation for mean valuation is linear in observables and identical to the OLS version. The second equation for the probability distribution over the number of bidders (5 or 12) is linear in obserable characteristics including the number of days of the auction and the month of the auction. Note that the number of days of the auction is only included in this equation and is not included in the equation determining the mean valuation.

The table highlights the main difference with estimates from the two approaches. The order statistics approach gives an estimate for the mean of the value distribution that is substantially lower than the OLS estimate suggests. In this case the difference from the point estimate ($e^{10.98} - e^{10.53}$) is a little over \$20,000. Note further that the 95% confidence intervals of the constant estimates from the two models do not overlap. That is, OLS may substantially overestimate the average value of the C5 Corvette.

The results suggest that the simple OLS model captures most of the change in in the mean value due to observable characteristics. However, the OLS model provides a biased estimate of the overall mean value of the item. Note the estimated model of the order statistics approach is quite simple. Still, these results provide some flavor for how to estimate demand when only auction prices and auction characteristics are observed. These results provide the mean and variance of the (log) normal distribution and thus one could determine the distribution of valuations for a particular C5 Corvette. For example a new manual red convertible has a mean of \$19,930 while two standard deviations gives \$8,434 and \$47,099.

8 Optimal vs Actual Reserves

On eBay sellers have the option to set the reserve price of their car. They have two choices. First, the seller must set a positive starting price. The first bid in the auction must be above this price but as eBay is a second price

	OLS		Order IV	
Variable	β	SE	β	SE
Mean Valuation				
Constant	10.98	0.10	10.53	0.09
Log Mileage (Log of Miles)	-0.03	0.01	-0.04	0.01
Log Age (Log of Years)	-0.36	0.03	-0.36	0.03
New (Dummy)	-0.59	0.08	-0.74	0.09
Automatic (Dummy)	-0.06	0.02	-0.06	0.02
Convertible (Dummy)	0.15	0.02	0.14	0.02
Color Black (Dummy)	-0.10	0.03	-0.08	0.03
Color Burgundy (Dummy)	-0.26	0.06	-0.24	0.06
Color Red (Dummy)	-0.05	0.03	-0.03	0.03
Color Silver (Dummy)	-0.05	0.03	-0.03	0.03
Color White (Dummy)	-0.03	0.04	-0.03	0.03
Negative Feedback (Dummy)	-0.07	0.05	-0.08	0.05
Log Feedback (Log of Aggregate Score)	-0.01	0.01	-0.01	0.01
σ	-	-	0.43	0.01
Month Dummies	Yes		Yes	
Probability of 12 Bidders				
Constant	-	-	2.75	0.95
Auction Length (Number of Days)	-	-	-0.12	0.11
Month Dummies	-	-	Yes	
Adjusted R-Squared	.39			
Log Likelihood			1331.19	
Number of Auctions	705		705	

Table 3: Estimates for C5 Chevrolet Corvette Sold in 2003

auction the ending price may be equal to the starting price. Second, the seller (for a relatively small fee) can set a hidden reserve price. This reserve is not listed but the bidders are informed about whether this reserve has been met. Bidders can bid below this reserve but the car will only sell if there is at least one bidder above the hidden reserve. Again the ending price may be equal to the hidden reserve. If the hidden reserve and the starting price are assumed to be equivalent (which they are in a private values setting) then auction theory posits an optimal reserve for the seller to choose.

This section compares the actual reserves to the optimal reserves predicted by the theory (given the results of the previous section). Sellers that use hidden reserves set them quite high, around $24,000, while sellers that use the starting price only, set the reserve at around $7,000. Economic theory suggests two possible reserves. If the seller only has one auction in which to sell the item the theory (and the coefficient estimates) suggest an optimal reserve of $5,400, a number that is much lower than actual hidden reserves and somewhat lower than the starting prices without hidden reserves. However, economic theory also posits a reserve where the seller accounts for the option to resell in a later auction. In this case the theory and estimates suggest an optimal reserve of $22,000 a number which is somewhat lower than the actual mean for sellers using hidden reserves but much higher than the reserve for sellers using only starting prices.

The seller's problem is to choose a reserve price that maximizes expected revenue. For simplicity assume there are no (or small) selling costs.[25] On eBay the price is equal to the reserve when there is one (only) bidder above the reserve. Consider only auctions with at least one bidder. In the case of a single auction expected revenue from the reserve is $S(1 - F(S))$ where S is the reserve price and $F(.)$ is the value distribution. That is, expected revenue equals the reserve (S) times the probability the highest bidder's value is above the reserve $(1 - F(S))$ (Haile and Tamer (2003)). Assuming the maximum is given by the first order conditions, the optimal reserve (S^*)

[25] Given the average price of a C5 Corvette this seems reasonable. Note that buyers generally pay shipping costs.

in a single auction is defined by the following equation.

$$S^* = \frac{1 - F(S^*)}{f(S^*)} \tag{26}$$

where $f(.)$ is the density function for value distribution.[26]

If the seller is considering relisting then it is reasonable to assume the seller faces a infinite sequence of possible auctions. The seller's problem is

$$\max_S V(S) = S(1 - F(S)) + F(S)\delta V(S) \tag{27}$$

where $V(S)$ is the continuation value (the option value of selling the item in the current auction) and δ is the seller's discount rate. This problem can be rewritten as

$$\max_S \frac{S(1 - F(S))}{1 - \delta F(S)} \tag{28}$$

The first order condition can be written to give

$$S^* = \frac{1 - (1 + \delta)F(S^*) + \delta F^2(S^*)}{f(S^*)(1 - \delta)} \tag{29}$$

If conditions are such that S^* solves the sellers maximization problem (Equation (28)) then this equation can be used to determine the optimal reserve.[27] It is straightforward to see that when $\delta = 0$ the S^* from the dynamic programming problem is equal to the solution to the single auction problem.

Table 4 presents a comparison between the actual reserves set by the sellers and the optimal reserves suggested by the theory for C5 Corvettes. The table also presents the Buy-It-Nows which are discussed further below.[28]

[26]Thanks to George Deltas for helping with this analysis. See Deltas and Jeitschko (2006) and Haile and Tamer (2003) for more discussion of optimal reserves with citations and information on the conditions which are required for the first order condition determines the optimum.

[27]See Deltas and Jeitschko (2006) and Haile and Tamer (2003) and a dynamic programming text for the necessary conditions.

[28]A Buy-It-Nows is eBay's fixed-price option that allows bidders to "win" the auction immediately if certain conditions are met. Note that not all sellers are qualified to offer Buy-It-Nows.

	Actual	Optimal
Static		
All (mean) ($)	13,405	5,739
All (median) ($)	10,000	5,597
Hidden Reserve (mean) ($)	24,984	5,825
Hidden Reserve (median) ($)	25,000	5,653
No Hidden Reserve (mean) ($)	7,467	5,884
No Hidden Reserve (median) ($)	1,000	5,597
Seller Feedback 100 or greater (mean) ($)	8,448	5,825
Seller Feedback 100 or greater (median) ($)	1,000	5,541
Dynamic		
All (mean) ($) (Interest 10%, 24 periods)	13,405	22,026
All (mean) ($) (Interest 12%, 24 periods)	13,405	21,163
All (mean) ($) (Interest 8%, 24 periods)	13,405	22,925
All (mean) ($) (Interest 10%, 12 periods)	13,405	18,958
Hidden Reserve (mean) ($) (Interest 10%, 24 periods)	24,984	22,248
Hidden Reserve (mean) ($) (Interest 10%, 52 periods)	24,984	25,848
Buy-It-Now		
BIN User (mean) ($) (Interest 10%, 52 periods)	28,642	26,370
BIN User and Hidden Reserve (mean) ($) (Interest 10%, 52 periods)	26,506	25,084

Table 4: Actual and Optimal Reserves for C5 Chevrolet Corvette Sold in 2003

The actual reserves are the mean (or median) of the highest of the starting price or the hidden reserve. In the static (single auction) case the optimal reserve is calculated using Equation (26) assuming a log-normal distribution using the coefficient estimates for the Order IV model presented in Table 3 evaluated at the mean (or median) value for each variable.[29] In the dynamic case the optimal reserve is calculated using Equation (29) assuming a log-normal distribution using coefficient estimates for the Order IV model with assumptions on the seller's discount rate and the time to the next auction. The various assumptions are stated in parenthesis. It is assumed that the discount rate is equal to $\delta = \frac{1}{1+i}$ where i is the interest rate. The discount rate accounts for both the seller's time preferences and the cost of relisting the car. As we don't observe the seller's time preferences the table considers interest rates equal to 10%, 8% and 12%. Note that the number reported in parenthesis is the annual equivalent to the periodic rate. The number of periods refers to the time between auctions where 24 means there is a little over two weeks between each auction.[30] The table presents results for all cars evaluating actual reserves at the mean and for the mean car, as well as at the median and for the median car. Actual and optimal reserves are also presented for three subsets of cars. Those with a hidden reserve, those with no hidden reserve and those where the seller has an aggregate feedback score that is greater than 100.[31] In order to account for possible differences in cars the optimal reserve is calculated for the mean (or median) of the values for each subset of cars.

The results suggest sellers choose reserves which are substantially higher than the optimal static reserve. This is particularly striking for sellers that use hidden reserves. These reserves are set very high and much higher than other reserves. The mean actual reserve for cars with hidden reserves is $25,000 (so is the median) which is much higher than the optimal static re-

[29]I set all the color dummies and month dummies to 0.

[30]The modal car auction lasts 7 days.

[31]This second subset may capture sellers who are more experienced on eBay as feedback score is highly correlated with the number of transactions that the seller has participated in.

serve of around \$6,000. The actual hidden reserves are only slightly higher than the optimal dynamic reserve of around \$22,000 for 24 periods and slightly lower the the optimal reserve of \$26,000 for 52 periods.[32] Sellers that do not use hidden reserves and sellers which have aggregate feedback scores that greater equal to 100 set reserves that are similar to the optimal static reserve. The mean actual reserves is around \$8,000 and the median is \$1,000. The mean and the median of the actual reserves straddle the optimal static reserve estimate which is between \$5,500 and \$5,900. The difference between hidden reserves and non-hidden reserves is not explained by observable differences in the cars because the optimal reserves are calculated for each subset of cars.

What about Buy-It-Nows? Auction theory suggests that the optimal reserve should be the same as the optimal price for a monopolist (Haile and Tamer (2003)). Assuming the seller is a monopoly for their particular C5 Corvette then the seller should set the Buy-It-Now at exactly the same as the reserve. Table 4 presents the actual and optimal Buy-It-Nows for all those that use Buy-It-Nows and for those that also use hidden reserves.[33] The optimal Buy-It-Now is calculated assuming relisting opportunities (Equation 29). For sellers that use both the Buy-It-Now and the hidden reserve the Buy-It-Now has an average that is around \$26,500 and less than \$500 more than the optimal dynamic Buy-It-Now. This result suggest that sellers that are using the Buy-It-Now and the hidden reserves are contemplating future auctions and are assuming that they will have the opportunity to relist the item.[34]

There are number explanations for why actual and estimated optimal reserves may differ. The sample used to calculate the actual reserves only includes auctions for which there was a sale.[35] The estimates used to calculate the optimal reserves may be incorrect due to mispecification. Another possi-

[32] There is also some variation depending on the interst rate although it is quite small.

[33] Note the sample does not include auctions in which the car was sold with the Buy-It-Now.

[34] Thanks to Laura Hosken for suggesting this test and pointing out this possibility.

[35] This suggests that mean (and median) estimates of the actual may be biased down.

bility is that sellers are choosing the wrong reserve, this may be particularly true for inexperienced sellers.

What is most puzzling is the difference between the hidden reserves and the non-hidden reserves. Is it a coincidence that those that choose to have a hidden reserve choose a number that is close to the optimal dynamic reserve while sellers that choose to have no hidden reserve choose a reserve close to the optimal static reserve? Are there two types of sellers? Are there sellers that want to sell immediately and are only contemplating a single auction and sellers that want the highest price and are accounting for future opportunities to sell the item. The evidence on the Buy-It-Nows suggest that this may be what is happening. Sellers that use both Buy-It-Nows and hidden reserves set their Buy-It-Nows at number that is almost identical to the dynamic optimal reserve, which is exactly as theory would predict. Further, circumstantial evidence comes from eBay's own advice which suggests sellers looking to sell their car in a hurry refrain from using hidden reserves.[36] An alternative explanation is that there are unobservable characteristics of the car that account for the higher hidden reserves.[37] Another is that different sellers have different "own use" values for the car and this explains the differences in the actual reserves used.[38]

9 Conclusion

This paper develops on ideas presented in Song (2005) and Athey and Haile (2002), and suggests an alternative method for identifying demand in single eBay auctions. Athey and Haile (2002) shows that in certain auctions demand can be identified from observing the price and the number of bidders. Unfortunately, in general eBay auctions it is not possible to observe

[36]See discussion in the introduction.

[37]In the estimation above it is assumed that reserves are set exogenously, or least not determined by unobserved characteristics of the car which will be correlated with the car's end price.

[38]Thanks to George Deltas for pointing out this possibility. Haile and Tamer (2003) calculate out the optimal static reserve with a positive own use value.

the number of bidders. Above I show that the value distribution is identified if the probability distribution over the number of bidders is known. However, given censoring in eBay type auctions it is not clear that it is possible identify that distribution. Song (2005) shows that for a certain set of eBay auctions it is possible to identify demand even when the number of bidders is unknown if the distribution of *two* order statistics are observed. However, in many cases it is not possible to observe two order statistics in eBay auctions. The paper presents the necessary conditions to simultaneously identify both the value distribution and the distribution over the number of potential bidders. The intuition is similar to solving the "simultaneity problem". These distributions are identified if there are observable characteristics of the auctions that vary with one (the distribution of potential bidders) but not the other (the value distribution). I show that the basic identification results can be generalized to the case where there is observed and unobserved item characteristics.

A parametric estimator based on the identification results is estimated on C5 Chevrolet Corvettes. Comparing the OLS results to the results from the order statistics approach, the latter provides for a much lower estimate of the mean value of the car. These results are used to calculate optimal reserve prices which are compared with actual reserve prices. The results suggest that those sellers using hidden reserves set them slightly higher than the optimal dynamic reserve while sellers that do not use hidden reserves set them slightly higher than the optimal static reserve. It is not clear whether this is coincidence and why there would be a such a large difference between the amounts of the two types of reserves. Indirect evidence suggests it is not a coincidence. Sellers that use both Buy-It-Now and hidden reserves set their Buy-It-Nows at the optimal amount given the dynamic considerations. This result contrasts with recent empirical work suggesting that sellers do not set optimal prices (Levitt (2006)) and NFL coaches do *not* optimally solve dynamic programming problems (Romer (2006)).

References

Christopher P. Adams. Identifing demand in ebay auctions. FTC Working Paper 272, June 2004.

Christopher P. Adams, Laura Hosken, and Peter W. Newberry. ‘vettes and lemons on ebay. SSRN http://ssrn.com/abstract=880780, February 2006.

Susan Athey and Philip Haile. Identification in standard auction models. *Econometrica*, 70(6):2170–2140, 2002.

Patrick Bajari and C. Lanier Benkard. Demand estimation with heterogenous consumer and unobserved product characteristics: A hedonic approach. *Journal of Political Economy*, 113(6):1239–1276, 2005.

Patrick Bajari and Ali Hortacsu. Winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *RAND Journal of Economics*, 34(2), 2003.

Patrick Bajari and Ali Hortacsu. Economic insights from internet auctions. *Journal of Economic Literature*, 42:457–486, 2004.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 60(4):889–917, 1995.

Adam Cohen. *The Perfect Store: Inside eBay.* Little, Brown and Company, first edition, 2002.

George Deltas and Thomas Jeitschko. Auction hosting site pricing and competition. University of Illinois, February 2006.

Tugba Giray, Kevin Haker, and Robin Sickles. Estimating consumer surplus in ebay computer monitor auctions. Rice University, February 2006.

Philip A. Haile and Elie T. Tamer. Inference with incomplete model of english auctions. *Journal of Political Economy*, 111(1), 2003.

Steve Levitt. An economist sells bagels: A case study in profit maximization. NBER Working Paper W12152, 2006.

David Lucking-Reiley. Auctions on the internet: What's being auctioned and how? *The Journal of Industrial Economics*, 68(3):227–252, September 2000.

Aviv Nevo. A practicioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics and Management Strategy*, 9(4): 513–548, 2000.

Paul Resnick, Richard Zeckhauser, John Swanson, and Kate Lockwood. The value of reputation on ebay: A controlled experiment. University of Michigan, May 2003.

Leonardo Rezende. Auction econometrics by least squares. University of Illinois, November 2002.

David Romer. Do firms maximize? evidence from professional football. *Journal of Poltical Economy*, 114(2):340–365, April 2006.

Unjy Song. Nonparametric estimation of an ebay auction model with an unknown number of bidders. Discussion Paper 05-14, University of British Columbia, 2005.