Unobserved Heterogeneity and Reserve Prices in Auctions

James W. Roberts*

June 28, 2013

Abstract

This article shows how reserve prices can be used to control for unobserved object heterogeneity to identify and estimate the distribution of bidder values in auctions. Reserve prices are assumed to be monotonic in the realization of unobserved heterogeneity, but not necessarily set optimally. The model is estimated using transaction prices from a used car auction platform to show that the platform enables sellers to capture a large fraction of the potential value from selling their vehicle. Individual sellers benefit mostly from access to a large set of buyers, but the magnitude depends on accounting for unobserved heterogeneity.

JEL CODES: D44, L20, L62

Keywords: Auctions, Unobserved Heterogeneity, Mechanism Design

*Department of Economics, Duke University and NBER. Contact: j.roberts@duke.edu.

This article has benefited from helpful conversations with and insight from Robert Porter, Igal Hendel, Ali Hortaçsu, Aviv Nevo, John Asker, Federico Bugni, Matthew Gentzkow, Phil Haile, Ben Handel, Joel Horowitz, Tom Hubbard, Rosa Matzkin, Ryan McDevitt, Kanishka Misra, Jason O’Connor, Mallesh Pai, Aaron Sojourner, Andrew Sweeting, Elie Tamer, Xun Tang, Glen Weyl and numerous seminar participants. Finally, I thank Joonsuk Lee for being so gracious in providing the data used in the article. Financial support from Duke University and the Center for the Study of Industrial Organization at Northwestern University is greatly appreciated. Any errors are my own.
1 Introduction

A seller can benefit in several ways from utilizing an auction platform to sell an asset. The seller may be able to reach a greater number of buyers, who now must compete against one another for the asset, and if the seller can set a reserve price, he can retain some bargaining power. Quantifying the relative importance of these channels in any particular setting requires estimating buyers’ latent demand for the good in order to, for example, measure the return to reaching an extra bidder or determine a seller’s optimal reserve price. Estimating this type of model primitive is often a central goal of empirical auction research, and in this way the field is similar to many other areas of applied economics. However, as it has been recently noted (Einav and Nevo (2007) and Pakes (2008)), the field has not kept pace with other empirical branches of economics in its treatment of unobserved auction-level heterogeneity. Until recently, much of the empirical auction literature assumed that the econometrician and auction participants have the same information about the items being auctioned. This assumption is likely violated in many settings.

When the auction literature does control for unobserved heterogeneity, it only uses information from bidder behavior. This article presents an alternative approach based on incorporating observed seller behavior. If bidders observe and respond to an object’s characteristic that the econometrician cannot observe, it is likely that the seller, who as the object’s owner is familiar with its intricacies, also observes this characteristic. I show how to use the behavior of the seller, such as the realizations of one of his choice variables, to recover information about what the econometrician does not observe. One of the seller’s important choice variables, which is often observed in data, is the reserve price (the price below which he will not sell the item). Different reserve prices for assets with the same observable attributes suggest that the items differ in value in ways that are not directly attributable to the observed characteristics. In this way, reserve prices can be an indication of the unobserved quality of an asset.

To formalize this intuition, I present a model of reserve price setting and bidding behavior
for an item in which agents observe some characteristic of the object that the econometrician does not. After showing that the model can be identified using information typically found in standard data sets, I estimate the model with data on weekly used car auctions in South Korea. In these auctions, used car owners bring their vehicles to an auction house that sells the cars through English button auctions. Although studying used car auctions is of independent interest due to their importance as an inventory sourcing mechanism for dealers (e.g. in 2007 U.S. used car dealers filled over 30% of their inventory via auctions), more relevant for this article is the fact that used cars likely differ in ways observable to bidders at the auction, but unobservable to the econometrician, thus making them a natural context in which to explore the impacts of controlling for unobserved heterogeneity.

There are several key identifying assumptions in the model. These include the assumptions that sellers and bidders both observe and respond to what the economist does not observe and that sellers do not have private information about the cars’ qualities that buyers value, but are not privy to. Several institutional features of the setting suggest that these assumptions are reasonable. Buyers are used car dealers who resell the vehicles to individuals after auction. Therefore it is unlikely that unobserved heterogeneity would affect a seller’s, but not a buyer’s, value for a car. Additionally, given that buyers are themselves dealers, they are trained in evaluating automobiles. Moreover, they are given ample time to inspect each car prior to auction. Therefore it is unlikely that sellers have private information valued by buyers. Nevertheless, before structurally estimating the model, I present empirical evidence that the data clearly support these, and other, maintained assumptions.

After estimating the distribution of bidder values, I explore the benefits to sellers from using the auction house as a platform to sell their vehicles. There are various channels through which sellers benefit from using an auction relative to, say, approaching buyers independently. First, at the auction house, buyers compete simultaneously for a vehicle. This benefits the seller if otherwise he would have to approach potential buyers one-on-one and engage in some costly bargaining process with them. Second, buyers are used car
dealers that must register with the auction house to attend its weekly auctions. In this way the seller is able to engage a greater number of buyers than would be possible on his own without bearing substantial costs. Access to a greater number of bidders is valuable to sellers given the classic result of Bulow and Klemperer (1996): no amount of bargaining power is worth as much to a seller as one extra bidder. Finally, reserve prices afford sellers some amount of bargaining power at the auction house.

Using the structural estimates, I quantify the impact of each of these channels on seller revenue. Relative to the first two channels, the third, reserve prices, has a relatively small impact on revenues. This is important since a great deal of literature focuses on the ability of sellers to set reserve prices.\(^1\) This literature is largely motivated by the fact that, for a fixed set of bidders, setting the correct reserve price is essential for optimal mechanism design. While this article lends quantitative support to the fact that setting an optimal reserve price among a fixed set of bidders improves seller revenues, it also provides evidence that the gain may be small relative to simply adding another bidder. It is natural to view auction houses, such as the one studied in this article, as conduits through which a seller can interact with a larger set of buyers than he could on his own.

I also explore the impact of controlling for unobserved heterogeneity when quantifying the relative importance of these channels. First, an individual seller’s optimal reserve price will tend to be overestimated when unobserved heterogeneity is ignored. This is sensible since ignoring unobserved heterogeneity causes too much variation in bids to be attributed to variation in values. Overestimating the dispersion in values tends to upward bias the reserve price. Second, given that reserve prices are estimated to be more valuable to a seller when unobserved heterogeneity is ignored, the value of adding another bidder relative to setting an optimal reserve with the original set of bidders, is estimated to be comparatively lower than when unobserved heterogeneity is controlled for. This implies that too much of

the gains to sellers from using the auction house are credited to setting the right reserve price, and too little is assigned to the larger pool of buyers. In this way the article is related to a larger discussion (e.g. Klemperer (2002)) about the relative importance to sellers of auction design (i.e. setting the correct reserve price) and more general economic forces (i.e. the amount of competition for the item), since it shows that unobserved heterogeneity can be important when quantifying these determinants of seller returns.

**Related Literature**

The often unstated assumption in the empirical auction literature is that the economists’ information set about the object being sold coincides perfectly with that of the bidders, aside from their private signals, and therefore that any variation in bids across observably similar objects derives solely from variation in bidder values. Recently there have been several attempts to deal with this issue and I now discuss some of them as they relate to this article.

Krasnokutskaya (2011) cleverly applies deconvolution techniques, the roots of which can be found in Li and Vuong (1998), and Li, Perrigne, and Vuong (2000) in the auction literature, to address unobserved heterogeneity. Like this article, Krasnokutskaya’s semi-parametric method, which has recently been applied by other researchers, such as in Asker (2010), allows the unobserved heterogeneity to vary within auctions with the same number of bidders. To allow for asymmetric bidders, data on multiple bids from the same auction is required. If symmetry is imposed and one has access to reserve prices, this method can be used when only the winning bid is observable. This, for example, is done in Decarolis (2009) when comparing different mechanisms used in Italian procurement auctions. A more flexible version of Krasnokutskaya’s model is presented in Hu, McAdams, and Shum (2013), but the authors require data on three bids in an auction. Krasnokutskaya’s approach is quite distinct from what the current article proposes. First, her method assumes that a bidder’s value (or cost since she is looking at procurement auctions, specifically, highway maintenance and paving contracts) is the product of a bidder’s private cost and a common unobserved com-
ponent/cost. She also introduces a less restrictive form in which the mean and variance of the cost distribution are affected by the common component in orthogonal ways by letting there be a two-dimensional common component: one affecting average cost and one affecting the individual cost deviations, though this requires the stronger assumption that both components of the common shock are distributed independently from one another as well as from the private signals. This article does not require specifying the functional relationship between private and common signals. Second, Krasnokutskaya requires bidders’ signals to be distributed independently from the common component, where as I allow signals to depend on the unobserved common component.

Several authors have adopted a fully parametric approach to dealing with unobserved heterogeneity in auctions. Athey, Levin, and Seira (2011), in their study of timber auctions, assume a parametric form for the distribution of bids conditional on the unobserved heterogeneity. They then assume a distribution for this unobserved heterogeneity and form the unconditional likelihood based on their parametric assumptions. They also assume, which I do not, that the unobserved heterogeneity is orthogonal to the observable characteristics of the item being auctioned. The authors readily admit that some of these assumptions are strong, in particular that the realization of unobserved heterogeneity is orthogonal to the observable item characteristics, but they defend such assumptions “in light of most empirical work on auctions, which makes the even stronger assumption that there is no unobserved heterogeneity at all.” Moreover, the authors provide substantial evidence that their parameterization is appropriate by comparing predicted and actual bid residuals.2

Roberts and Sweeting (forthcoming) provide another example of recent attempts to parametrically control for unobserved heterogeneity. In that article, the authors estimate the distribution of bidder values using simulated maximum likelihood with importance sampling. This enables them to allow for unobserved heterogeneity in values, as well as in entry costs.

---

2In another defense of the authors’ empirical work, their goal is to calibrate the distribution of bidder values in what are likely to be competitive (sealed-bid) auctions to then make out-of-sample predictions to assess the competitiveness of other potentially collusive (open-outcry) auctions.
and the degree of selection in the entry process. This is important since their focus is on a model of auction participation in which bidders with higher values may be more likely to submit bids. However, to implement this approach, functional form assumptions about the distribution of structural parameters must be made.

Evidenced by the aforementioned literature, the current attempts to deal with unobserved auction heterogeneity use only bidder behavior. This article incorporates seller behavior to extend the existing methodology, and in this way it is more tangentially related to the analysis of eBay Motors in Lewis (2011). In his model the seller has private information, which is signaled to the bidders, and the economist observes this signal. While my model is more restrictive, since I do not allow the seller to have this private information, I provide support for this presumption in the empirical application. Another important difference is that the information bidders have about the automobiles varies in Lewis (2011), thereby justifying a common values model, whereas in the application below there is no reason to think that bidders' information sets vary. I provide further justification for the private values approach below.

The method’s details are explained below, but as motivation I now reference a figure from a seminal article in the empirical auction literature: figure 3 from Laffont, Ossard, and Vuong (1995), which can be found on p. 978 in their Appendix A. The article looks at Dutch auctions of eggplants in France. The picture lists the auctions in chronological order and plots for each one the winning bid and reserve price, which clearly move in tandem. While in a Dutch auction, unlike the English auction studied below, changes in the winning bid might be induced solely by exogenous variation in reserve price. If this were the case, one would also expect variation in the fraction of items sold, which according to the authors does not happen. Therefore, while the auctions may appear similar to the econometrician, they likely differ in ways that are recognized and responded to by seller and bidder alike. This is the central intuition of the current article.

The similarity between this article and existing work in structural industrial organization
should be noted. Using reserve prices to control for unobserved heterogeneity in the bidding functions closely mirrors the conceptual framework of Olley and Pakes (1996) and was originally suggested to be of use in auctions by Athey and Haile (2006). Other related attempts to use this approach in the structural econometric auction literature are Haile, Hong, and Shum (2006) and Guerre, Perrigne, and Vuong (2009), which rely on monotonicity of the unobserved heterogeneity in the number of bidders. One advantage of using reserve prices instead of the number of bidders is that, in many data sets, the latter is unobservable, or at the least, imperfectly observable.³

The article proceeds as follows. Section 2 introduces the modeling framework used for estimation and addresses several questions of identification. Section 3 introduces the Korean automobile auction data. Section 4 discusses estimation, illustrates its results and brings these results to bear on the importance of controlling for unobserved heterogeneity when evaluating the relative impact on seller revenues of setting an optimal reserve price and increasing competition. Section 5 concludes. Propositions and their proofs are presented in the appendix and all tables and figures appear at the end.

2 The Model

In this section I explain my strategy for using variation in reserve prices to account for unobserved heterogeneity in auctions. To introduce notation, in any auction \(t\), there are \(N_t\) bidders with unit demands indexed by \(i\). In the empirical section I consider used car auctions. For each car there are some observed characteristics \(Z_t \in \mathbb{R}^K\) and an unobserved (to the econometrician) characteristic \(U_t \in \mathbb{R}\).⁴ Both the seller and the bidders observe \(Z_t\) and \(U_t\). The econometrician only observes \(Z_t\). At each auction bidders learn their independent and

³In addition, the focuses of these two articles are quite different from the current article. Haile, Hong, and Shum (2006) focus on developing a test for private versus common values. Guerre, Perrigne, and Vuong (2009) consider the identification of the first price auction model with risk averse bidders.

⁴The assumption that unobserved heterogeneity is single valued is a restriction that will be maintained throughout the empirical portion of the article. However, in the conclusion I suggest how, in future work, one might relax this assumption and still identify the model, should the necessary data be available.
private values $V \sim F_V(\cdot)$, which can depend upon $Z_t$ and $U_t$, but not on $N_t$. I am assuming, therefore, that bidders are symmetric in their value distribution. Let the auction format be any one in which bidders submit bid $B$ according to some function $B = \beta(Z, U, V)$.

Now, suppose that sellers set reserve prices $R$ according to $R = m(Z, U)$, where it is assumed that $m(\cdot)$ is invariant across auctions and that $N$ includes those bidders with values potentially less than $R$. If different auctions in the data are held by distinct sellers, this could be restrictive. However, the assumption could be relaxed should the data contain the necessary information about the sellers or contemporaneous auctions to allow it. In the empirical example below, I argue that the same function is used to set each vehicle’s reserve price.

Due to data limitations, in the empirical portion of the article I will take a semi-parametric approach to estimation in which I will assume that $m(\cdot)$ is a function known up to parameters. Using variation in reserve prices across auctions, we can recover realizations of unobserved heterogeneity. This can then be treated as known when using the bidding data to estimate the distribution of bidder values conditional on observables and this realization of unobserved heterogeneity.

However, as an extension to this approach, in this section I establish that in an environment with sufficient data, a more general approach might be taken. I begin by making the following assumption:

**Assumption S1** For all values $z$ of $Z$, $m(Z, U)$ is strictly increasing in $U$.

Sellers are not required to optimally set reserve prices. Instead, I assume that a higher realization of unobserved heterogeneity leads the seller to set a higher reserve price. This is an intuitive restriction on the function $m(\cdot)$ when the unobserved heterogeneity is given its normal interpretation as some index of unobserved quality. However, one way to gauge the stringency of the assumption is to ask whether it holds when sellers optimally select reserve prices. Proposition 1, below, establishes that Assumption S1 holds for a large class of bidder value distributions when sellers are presumed to set optimal reserve prices.
Before presenting the proposition, recall that from the literature on optimal auctions (e.g. Myerson (1981), Riley and Samuelson (1981)) the optimal reserve price, assuming that $F_V(\cdot)$ is strictly increasing and continuously differentiable, is given by

$$r^* = v_0 + \frac{1 - F_V(r^*)}{f_V(r^*)}$$

where $v_0$ is the seller’s value of keeping the car if it failed to sell at auction. For any realization of unobserved heterogeneity, $U$, if we define $\psi(t, U) = t - \frac{1 - F_V(t|U)}{f_V(t|U)}$, then the optimal reserve price $r^*$ is the solution to $\psi(r^*, U) = v_0$.

**Proposition 1.** If $V$ and $U$ are affiliated and $\psi(t, U)$ is increasing in $t$ for all $U$, then the optimal reserve price is increasing in $U$.

Intuitively, affiliation implies that higher realizations of $U$ make higher realizations of $V$ more likely. See Milgrom and Weber (1982) for the precise definition. Note that the regularity condition on the function $\psi(t, U)$ is the familiar one introduced in Myerson (1981). It is a high-level, although commonly imposed (see Bagnoli and Bergstrom (2005) and Ewerhart (2011) for references to its application in classic articles on labor economics, search theory, voting, mechanism design, regulation, etc.\(^5\)) assumption on model primitives that is satisfied for many parametric distributions often used in empirical work (e.g. uniform, normal, logistic, and with some restrictions on parameters, many other distributions).

Despite the result of Proposition 1, I will continue to allow sellers to set non-optimal reserve prices. This is because, provided that assumption S1 holds, the next proposition shows that with normalization assumptions on the functions $\beta(\cdot)$ and $m(\cdot)$, and conditional independence assumptions regarding the distributions of $U$ and $V$, the distributions of both unobserved heterogeneity and bidder signals are identified without assuming independence between $V$ and $U$. The result builds on the work of Matzkin (2003). For notation, decompose the observable characteristics into two sets of variables: $Z = (Z_0, Z_1)$.

\(^5\)Some of these examples make the stronger assumption of a monotone hazard rate.
**Proposition 2.** Let Assumption S1 hold and assume (i) \( \exists z_1 \) such that \( \forall U \) and \( \forall z_0, m(z_0, z_1, u) = u \) and (ii) \( U \) is independent of \( Z_1 \) conditional on \( Z_0 \). Then \( F_{V|Z,U}(\cdot) \) is identified from observations of the transaction price and the reserve price in first and second price auctions.

Therefore, using only data on reserve prices and winning bids, the distribution of bidders’ private values, conditional on observable auction covariates and the realization of unobserved heterogeneity, is identified.\(^6\) The proof of identification also suggests a two step estimation procedure whereby the econometrician first estimates the distribution of unobserved heterogeneity and the reserve price function \( m(\cdot) \). Then, the realization of the random variable \( U \) can be inferred since with \( m(\cdot) \) known, for any \( R \) and \( Z \), there is a unique solution to \( R = m(Z, U) \). Then, the distribution of \( V \) conditional on some auction covariates and the realization of \( U \), can be estimated using the distribution of winning bids for those auction covariates and the realization of \( U \).

In many applied settings, and indeed in the one studied below, implementing the fully non-parametric estimation procedure may be impractical. However, to illustrate the procedure, I provide a small simulation experiment in Appendix A.

### 3 Data Description

In this section I introduce the used car auction data. I will also show how the data support the contention that both bidders and sellers observe what the econometrician does not and justify the modeling assumptions made above.

Cars have been sold by auction for at least 60 years. The National Automobile Auction Association (according to the NAAA 2007 survey) estimates that over 9.5 million cars, valued at nearly $89 billion, were sold by auctions in the U.S. in 2007, a 2.2% increase over 2006.

\(^6\)One should note that the normalization required in the proposition is one that is implicitly assumed in many functional forms. For example, if we let \( Z = Z_1 \), both \( R = ZU \) and \( R = \gamma Z + U \) satisfy the restriction when \( z = 0 \).
and a 5% increase over 2005. The National Association of Automobile Dealers estimates that dealers acquired 32% of their used car inventory via auctions, up from less than 6% in 1982 (see Taylor (2007) and Kontos (2006) for more details). While the importance of used car auctions as an inventory sourcing mechanism makes their study of independent interest, automobile auctions are ideal to study unobserved heterogeneity: buyers and sellers likely are able to observe aspects of the cars, which affect their values, and thus their bids and reserve prices, that cannot be observed in even a rather extensive data set provided to the analyst. Examples of such unobserved qualities are rust, dents and tire quality.

Summary of Data

The used car auctions took place on a weekly basis at a particular auction house in Suwon, South Korea, about an hour south of Seoul, between October 10th, 2001 and November 27th, 2002. At this wholesale used car auction, both dealers and individuals may sell cars, but only dealers registered with the auction house may bid. During the 57 weeks covered, there were over 380 members registered with the auction house and this group was stable over the time period. Unlike the U.S., where it is common for dealers to sell both new and used cars, the dealers in this data set specialize in used cars only and, according to sources, these dealers are perceived to be sellers of high quality used cars. Also, car dealers in South Korea tend to be much smaller than their counterparts in the U.S. The auction house was set up by an affiliate of Daewoo, and its main competitor was a similar, but much smaller, auction house (set up by Hyundai) an hour away. The auction house charges the seller $50 for any car it auctions regardless of whether it sells and, in the event of sale, collects 2.2% of the transaction price from both the buyer and seller. The individual seller of the car can choose the reserve price, though in practice the auction house almost always provides guidance on this matter. Registered members receive information three days ahead of time about cars that are going to be auctioned. The information contains a description of the cars including make, model, mileage, age, etc. If a dealer wishes to place a bid, he must send an agent to
the auction house to be present at the time of the auction.

On auction day, bidders show up three or more hours before the auction to inspect the automobiles being auctioned that day. Once bidding begins, all bidders sit at tables in front of the car being auctioned. A screen at the front of the room displays the current bid for the car. The auction format used is a variant on the classic English button auction originally described in Cassady (1967) and analyzed in Milgrom and Weber (1982). The price of the item being auctioned rises continuously, and in order to remain in the auction a bidder must keep his finger on a button beneath the table in front of him. Reentry of bidders is allowed, and this simplifies the bidders’ strategies since they do not need to condition on which bidders are currently “in” or “out” of the auction. The price for the car being auctioned rises in $30 increments every three seconds, provided that at least two bidders have their buttons depressed. Bidders cannot observe when their competitors have dropped out of the auction. A bidder wins if he is the last bidder with his finger on the button and he pays the current bid at this moment. The winning bidder is responsible for transporting the automobile offsite.

Table 1 summarizes the data. While during the period of study there were 28,334 passenger cars auctioned, I focus only on the 4,943 automatic transmission, five passenger sedans that were not part of rental or government fleets prior to auction. I only observe the transaction price and the identity of the winner and the second highest bidder. Variable definitions appear below the table. There are three dominant makes sold at the auction house: Hyundai (in 42% of auctions), Daewoo (39% of auctions) and Kia (18% of auctions). Cars tend to sell for approximately $3,000. On average, more than five bidders place bids on a car. A person is counted as a bidder if they ever have their button depressed once the price has begun to rise.

Table 2 explores the factors affecting bids and the relationship between winning bids and reserve prices. The estimated impacts of the cars’ attributes are sensible and robust to different specifications. The last column of table 2 includes the log of the reserve price as
a covariate in an attempt to assess the co-movement of it and the winning bid, controlling
for all of the other likely determinants of price. There are several important observations to
make. First, the winning bid and reserve price are highly correlated (see specification 7 in
the table). Second, the overall fit of the model jumps from an $R^2$ of 0.90 to 0.98 when reserve
prices are added in specification 6. Moreover, when only the reserve price is controlled for, as
in specification 7, the $R^2$ is 0.97. This suggests that the reserve price does well in controlling
for unobserved variation in an item’s “quality”, subject to the caveat that the regressors are
now more collinear. Finally, the correlation of the residuals from regressions of the winning
bid and reserve price on the automobile covariates in specification 5 is 0.890, compared to a
raw correlation between the winning bid and reserve price of 0.983.

The empirical work in the article relies on auctions for which I have the winning bid
conditional on the reserve price being met. From these, co-movement between the winning
bid and the reserve price will be used to control for the unobserved heterogeneity. However,
it is important to know whether this is the case for auctions that fail to meet the reserve
price. For these auctions, it should also be the case that higher reserve prices tend to be
associated with higher maximum, though ultimately failing, bids since there is still likely
to be some unobserved characteristic affecting bidders’ bids and sellers’ reserve prices. To
this end, I was able to obtain data on the highest bid placed for an object that did not
sell (because the reserve price was not met) for a smaller time period: October 10th, 2001
to December 27th, 2001. During this time, 107 of the 1,117 cars failed to sell. I am able
to incorporate these auctions by using the second highest bid despite the fact it was less
than the reserve price. Table 3 gives regression results of both (logs of) highest bid and
average bid on the final three specifications in table 2 to show that bids and reserve prices
in auctions for cars that fail to sell respond in a similar manner to auction covariates as
they did in auctions ending in a sale. Also, the correlation in the residuals from regressions
of the highest bid and reserve prices on car characteristics in specification 1 in table 3 for
cars that failed to sell is 0.942, compared to a raw correlation between the highest bid and
reserve price of 0.985. The correlation in the residuals from regressions of the average bid and reserve prices on car characteristics in specification 4 in table 3 for cars that failed to sell is 0.943, compared to a raw correlation between the average bid and reserve price of 0.984. The results continue to support the similar possession of, and response to, information by bidders and sellers which is not available to the econometrician.

For external validity, we also might wonder if there is something different about the cars that do not sell. To test this, I estimated a logit model of whether the car failed to sell as a function of the cars’ characteristics and also the time of year. The only characteristic that seems to matter is that older cars are more likely to fail to sell. The results appear in table 4 below.

**Does the Model Fit the Data?**

Before moving on to the estimation, it remains to be shown that the data fit the model and its assumptions described so far. In order to validate the model, I will address whether I can support the assumptions that (i) conditional on \( Z \) and \( U \), bidders have independent private values, (ii) the correlation in bids and reserve price residuals is not a mechanical result, (iii) unobserved heterogeneity affects both the reserve price and the bids, (iv) bidders are symmetric and (v) sellers use the same reserve price setting function \( m(\cdot) \). While there may be other concerns about the model, the verification of these assumptions should greatly strengthen the confidence in the estimation results below.

There are two justifications for assuming independent private values (IPV) among the bidders, conditional on a car’s characteristics. First, bidders are dealers, and each has its own inventory, its own transportation cost and may cater to idiosyncratic local demands, thereby ruling out a pure common value component of the cars. In fact, bidders are often sent by a downstream purchaser to bid for a certain automobile on the downstream purchaser’s behalf, which lends further support to the IPV assumption. The second justification is that the information bidders have about a certain car is likely to be very similar. This auction
house was viewed as selling “high quality” used cars so that issues involving “Lemons” and other bidders potentially having information pertinent to another bidder’s value are not as relevant. As is well known, the assumption of IPV does not rule out correlation in bidders’ information or their values being commonly affected by characteristics of the item, provided that bidders do not have private information about the characteristics. So, the assumption that bidders all observe what the econometrician does not, and that this component affects their valuations, does not violate the assumption of IPV since bidders do not have private (from the other bidders) information about what the econometrician does not observe, in the same way they do not have private information about observed auction heterogeneity. This is in contrast to a market like eBay motors where buyers differ on many dimensions, which are correlated with their information sets about the auctioned item. For example, Lewis (2011) notes that a common value model is more appropriate for eBay motors because bidders who live in close proximity to the automobile being auctioned may be better able to inspect the car themselves than distant bidders.

In Section 3 I reported the correlation in residuals from regressions of the winning bid and reserve price on automobile characteristics. This attempts to formalize the intuition underlying the motivating picture from Laffont, Ossard, and Vuong (1995) that while auctions may appear similar to the econometrician, they likely differ in ways that are recognized and responded to by seller and bidder, via reserve prices and bids, alike. However, this replication may have been a mechanical result if the winning bidder tended to be the only bidder bidding above the reserve price. With a bid increment this would have guaranteed that the winning bid and the reserve price moved in lockstep. However, this is not the case because while I do not see any bids other than the winning bid, I do know the number of different bidders bidding on each car and how many were above the reserve price. On average, just under five bidders bid on each car and approximately four bid above the reserve price. Therefore, we can be confident that the co-movement in winning bids and reserve prices isn’t mechanically derived.
The identification and estimation of the model rely on the fact that there is nothing that enters the reserve price equation that does not enter the bidding function. Importantly, this rules out a seller’s private signal or the unobserved heterogeneity affecting only the seller’s resale opportunities (or in general his outside option) should the car not sell at auction. In regard to the private signal entering the reserve pricing function, as mentioned above, the auction house usually sets the reserve price. This means that it is not the seller who is setting it, and thus it is less likely that there is a private signal entering this function, which varies from car to car. I return to this in greater detail below.

One might also worry that, contrary to my assumptions, in the actual data generating process unobserved heterogeneity enters the reserve price function but not the bid function. However, buyer resale likely mitigates this concern. That is, since the buyers are used car dealers who will resell whatever vehicle they purchase at auction, if some characteristic of the car leads a seller to value it more, and thus set a higher reserve price, it is reasonable to believe that buyers too will bid more, since they eventually will sell the car to individuals who are also likely to value this characteristic.

To complement these arguments, we can test the hypothesis that there is something in which reserve prices, but not bids, are increasing, by looking at whether the probability of a failed sale due to a lack of bids above the reserve price increases in the reserve price. From the logit model of failure described above, I find that the reserve price is not a statistically, or economically, significant determinant of failure to sell. As the reserve price increases from from the 25th percentile to the 75th percentile, the probability of failure increases approximately 2 percentage points. This lends strong support to the assumption that nothing enters the reserve pricing function that does not also enter the bidding function.

The previous discussion dealt with the possibility that sellers may observe and respond to some information that bidders have (or may not have in the case of a private signal), but do not value. Another possibility is that sellers possess information that bidders value, but are unaware of. This is similar to the model studied in Lewis (2011). To address the
possibility that sellers in these auctions have information that bidders would like to know, one needs to consider how sellers behave in this informational environment. That is, how should the possession of private information that is valuable to bidders affect seller behavior? Cai, Riley, and Ye (2007) show that when bidders value a seller’s private information, it is incentive compatible for sellers to use their reserve price as a signal of this information: there is a separating equilibrium in which sellers with higher realizations of the private information set higher reserve prices. Moreover, there is a testable implication of this model in the data. Cai, Riley, and Ye (2007) also show that when reserve prices can be used as signals, the optimal reserve price is increasing in the number of bidders. When the log of reserve prices is regressed on automobile’s characteristics in specification 3 in table 2 and the number of bidders, the effect of the number of bidders on reserve prices is small, negative and statistically insignificant. This, along with the previous evidence that there is not a private signal which affects reserve prices but not bids, provides support for modeling reserve prices as functions of $Z$ and $U$ alone.

Another important assumption about bidders is that they are symmetric. Without this assumption, and without data on bidder identities, we cannot recover the distribution of bidder values using only the winning bid. The assumption of bidder symmetry is reasonable in this context as the used car dealers in these data tend to be uniformly relatively small and do not specialize in certain types of vehicles. However, it is possible that the sources of bidder inventories may differ in a way that makes some bidders stronger competitors than others at the auction. For example, if the auction house is the only source of cars for a dealer, that dealer may wish to pay higher prices for the same vehicle than another with multiple inventory sources would not. While there is some evidence that a few buyers tend to buy more vehicles at the auction house than others, 75% of bidders purchase 17 or fewer automobiles across 59 distinct auction days.

Finally, I have assumed throughout that all sellers use the same reserve pricing function $m(\cdot)$. This is potentially problematic since individual sellers are the ones bringing the auto-
mobiles to be sold. However, the majority of the time the auction house provides guidance on how to set the reserve price of the car (however, the data do not identify in which specific instances this occurs). To the extent that the auction house is an “expert” on the auction market, it then serves as a mechanism for translating car characteristics into reserve prices. Moreover, as the analysis will bear out, reserve prices appear to move in consistent ways across vehicles despite the fact that the identity of the sellers is changing. In general, if seller characteristics were observed, the reserve pricing function could vary with these attributes and thus differ across sellers.

4 Estimation

In this section I report estimates of the distribution of bidder values. As described in Section 2, I begin by specifying a flexible functional form for \( R = m(Z_1, U) \). \( Z_1 \) includes a car’s age and mileage as I will condition the analysis on the make and model of each vehicle. I assume that \( m(Z_1, U) = g(X; \alpha) + U \), where \( g(\cdot) \) is a known quadratic function, up to parameters \( \alpha \), of covariates \( X \) that are a car’s age, mileage, age squared, mileage squared, the interaction of age and mileage and a constant and I assume that \( E[UX] = 0 \) and \( E[XX'] = 0 \).

In the second stage, I consider all auctions with \( N \) potential bidders and characteristics \( Z_2 \). In practice, \( Z_2 = Z_1 \) when unobserved heterogeneity is ignored, and \( Z_2 = [Z_1 \hat{U}] \) when it is not. To avoid repeating notation, when I condition on \( Z \) below it is understood to stand for either \( Z_1 \) or \( Z_2 \) according to whether or not unobserved heterogeneity is controlled for. Below I outline how I specify the number of potential bidders and smooth across cars with differing characteristics. The distribution of winning bids for these auctions is related to the distribution of values for these cars as follows. Let \( v_{(i:N)} \) be the \( i \)th highest of \( N \) draws from \( F_V \) and denote its distribution by \( F_V{(i:N)} \). By ignoring the auctions’ tiny bid increments, we know that if \( F_{W|Z,N} \) is the distribution of winning bids across auctions with

\[^7\text{Including other covariates, like a car’s color, or using a cubic instead of a quadratic did not materially change the results.}\]
$N$ potential bidders, then $F_{W|Z,N} = F_{V(z,N)|Z}$. Furthermore (see Athey and Haile (2002) and more generally Arnold, Balakrishnan, and Nagaraja (1992)):

$$F_{V(i,N)|Z}(v) = \phi(F_{V|Z}; N) = \frac{N!}{(N-i)!(i-1)!} \int_{0}^{F_{V|Z}(v)} s^{N-i}(1-s)^{i-1} ds \quad (2)$$

where $\phi(\cdot)$ is a known and strictly monotonic function.

The estimation technique is a two step procedure. First $F_{V(z,N)}(v)$ must be obtained from the distribution of winning bids, which I estimate using the non-parametric kernel estimator for this distribution from Nadaraya (1964):

$$\hat{F}_{W|Z=z}(v) = \frac{\sum_{t=1}^{T} \tilde{k}_1 \left( \frac{v-W_t}{h} \right) k_2 \left( \frac{z-Z_t}{h} \right)}{\sum_{t=1}^{T} k_2 \left( \frac{z-Z_t}{h} \right)} \quad (3)$$

where $h$ is the bandwidth, $k_1 : \mathbb{R} \to \mathbb{R}$ and $k_2 : \mathbb{R}^L \to \mathbb{R}$ are kernel functions, and $\tilde{k}_1 = \int_{-\infty}^{0} k_1(s)ds$. I smooth across covariates using a Gaussian kernel proportional to the standard bandwidth suggested by Silverman (1986).

Once the distribution of winning bids is estimated, the estimate of $F_{V|Z=z}(v)$ is obtained via the known inverse mapping $\phi(\cdot)^{-1}$ from equation (2) above. Provided that we have a consistent estimator of the distribution of winning bids, since $\phi(\cdot)$ is a continuous function, the continuous mapping theorem implies that the estimator of $F_{V|Z=z}(v)$ will be consistent. Sufficient conditions (see, for example, Matzkin (2003)) for $\hat{F}_{W|Z=z}(v)$ to be a consistent estimator when $Z$ contains an estimated covariate (i.e. $Z = Z_2$) are that the kernel functions in equation (3) are Lipschitz continuous, $h^2 T \sqrt{T}/lnT \to \infty$ and $E|Z| < \infty$. These conditions are met in estimation.

As the above discussion makes clear, the estimate of the distribution of values depends on $N$. While some data sets include a natural measure of $N$, the set of potential bidders in these car auctions is not obvious.\footnote{Song (2004) suggests a method for avoiding this assumption, but this requires having data on more than one bid and being able to interpret all bids used in the estimation procedure as order statistics. Since I only have data on the winning bid, I am forced to take a stand regarding the number of potential bidders in an} One possibility is to assume that the number of observed bidders
in any auction is the set of potential bidders for that auction. This is problematic if some potential bidders do not bid. Therefore, I assume that the number of potential bidders is the maximum of the number of bidders bidding in any one auction. To support this assumption, I only pool across auction days if the number of bidders showing up at the auction house is similar on these days. There is not much variation within any month, but there is some variation across months. Therefore, I only pool across months with approximately the same number of total unique bidders. All the results below consider months 10/2001, 11/2001, 12/2001, 1/2002, 5/2002, 7/2002, 8/2002, 9/2002 and 10/2002, across which the average number of unique bidder is 62.6. I further restrict the set of cars to Hyundai Sonatas. To avoid outliers, I drop all auctions with more than eight bidders, which leaves over 92% of the auctions. I will thus assume $N = 8$. Table 5 presents summary statistics for this restricted sample.

Compared to the full sample of cars summarized in table 1, cars in this restricted sample tend to have moderately higher winning bids and mileage. Approximately the same number of bidders submit bids for these cars. Importantly, as was the case in the full sample, in this restricted sample there is a high degree of correlation in reserve prices and winning bids. If we control for the (log of) reserve price in a regression of the (log of) winning bid on age, mileage, color dummies and month dummies, the adjusted $R^2$ jumps from 0.773 to 0.971. In addition, the correlation in residuals from regressions of the winning bid and reserve price on these same covariates is 0.9343, compared to 0.9835 in the raw data for this sample.

To illustrate the estimation results, figure 1 presents estimates of the distribution of values for a seven year old car with 100,000 KMs both ignoring unobserved heterogeneity and assuming its realization of $U$ is at the 10th, 50th and 90th percentiles for this type of car. The solid distribution ignores unobserved heterogeneity. Clearly this increases the estimated variance of the distribution of values. This is expected since more of the variation in bids is attributed to variation in values. Focusing on the three distributions controlling for unob-

---

21
served heterogeneity, there is a clear ordering of the distributions. The value distributions for cars with higher realizations of $U$ maintain a first order stochastic dominance over those with lower realizations. For example, the estimated mean [with bootstrapped 95% confidence intervals based on 100 repetitions] of these distributions are, respectively, $1,919.61 \pm 1,937.11$, $2,426.20 \pm 2,448.65$, and $2,971.17 \pm 2,999.08$. This is intuitive since whatever the econometrician is not observing, the more favorable is this characteristic, the higher ought to be both the bids and the reserve prices.$^{9,10}$

**Optimal Reserve Price and Auction House Performance**

Auction platforms, such as the auction house considered in this article, are often attractive to individual sellers when it may be difficult to locate and negotiate with potential buyers. The ability to access a centralized, formal auction environment attracting numerous potential

---

$^9$To check the robustness of the estimates to my assumption regarding the number of bidders, I have also estimated value distributions assuming that $N = 7$ and $N = 9$. The results indicate that the estimates change, as they must, but not by very much. Moreover, the changes are intuitive. For example, if we assume $N = 7$ and estimate the distribution of values for seven year old automatic, five passenger, Hyundai Sonatas with 100,000 KMs, which are the basis for figure 1, the distributions are similar, but first order stochastically dominated by those assuming $N = 8$. This is intuitive since the winning bids are interpreted to be the second highest value from a smaller sample of bidders, and so the distribution they are inferred to be drawn from must contain higher values, on average. The opposite is true when it is assumed that $N = 9$. When it is assumed that $N = 7$, the means of the same value distributions are $2,168.82$, $2,542.19$, and $3,146.35$, respectively. When it is assumed that $N = 9$, the means of the same value distributions are $1,871.81$, $2,326.22$, and $2,788.02$, respectively.

$^{10}$An alternative way to approach estimation is to specify functional forms for the distribution of values and the number of potential bidders and to estimate these distributions’ parameters via maximum likelihood. While the advantage to this approach is that it does not require specifying the number of bidders, it does necessitate specifying functional forms for the distribution of potential bidders as well as the distribution of bidder values, which the current method avoids doing. Additionally, the identification of the parameters of each specified distribution is based purely off of functional form assumptions since the model will attempt to explain the data (the distribution of winning bids) by a mixture of second order statistics, with weights determined by the probability masses for different numbers of potential bidders. Nevertheless, I have attempted to implement such an approach. Based on the distribution of the log of winning bids looking approximately normal, I assumed that $F_V = \text{lognormal}(Z \beta, \sigma)$ and $Pr[N = n] = \text{Poisson}(\lambda)$. I allow the number of potential bidders to range from 2 to 20 and estimate the model using maximum likelihood, integrating out over the number of potential bidders. I find that the estimated mean of the distribution of bidder values for a seven year old Hyundai Sonata with 100,000 KMs is $1,758.45$, $2,184.41$ and $2,864.57$, when the realization of unobserved heterogeneity is at its 10th, 50th and 90th percentile, respectively. The estimated average number of potential bidders is 8.84. While this approach has its drawbacks, the fact that standard functional form choices for the distribution of values and the number of potential bidders produced estimates of the distribution of values and a mean number of bidders that are similar to those in the benchmark case above lends support to the estimation approach used throughout the article.
buyers is one reason platforms such as eBay, or well-known auction houses like Christie’s or Sotheby’s, are so popular among individual sellers of art, wine and other miscellaneous goods. When price discovery is important, these auctions allow a seller to avoid having to set a (potentially uninformed) fixed price or negotiate with buyers. In this section, I evaluate how well sellers do in capturing the value from selling their car at the auction house. I then explore the channels through which the auction house improves an individual’s ability to sell his automobile. There are at least three obvious ways by which a seller can benefit. The first channel is the simultaneous competition among buyers that an auction house generates. Related to this is a reduction in a seller’s cost of reaching buyers. The second channel is the attraction of additional potential buyers. When these buyers must compete in an auction, as they do at this platform, an additional bidder is worth more than any amount of negotiating power the seller may have with the original set of bidders. Third, the auction house’s reserve price enables the seller to retain some bargaining power with the set of bidders.

To begin, I estimate the fraction of the maximum potential value from selling their car that sellers capture by using the auction house, assuming that they could hold the optimal auction with the actual set of potential buyers from the auction house. It is well-known that a seller’s optimal mechanism with a fixed set of buyers is an open outcry auction with an optimal reserve price. However, in estimating the optimal reserve price, controlling for unobserved heterogeneity is crucial. To see the impact of ignoring unobserved heterogeneity on a common policy variable of interest, I estimate how the optimal reserve price would change based on whether the estimated distribution of values, which affects the optimal reserve price, controls for unobserved heterogeneity. According to equation (1), the optimal reserve price also depends on \( v_0 \). As is often the case, the data provide no direct measures for \( v_0 \). Thus, I follow the literature and estimate the optimal reserve prices for a variety of values.

11 Reserve prices have figured centrally in the empirical auction literature as they were some of the original policy variables of interest in such work as Paarsch (1997). Other examples of estimating optimal reserve prices include Mead, Schniepp, and Watson (1981), Mead, Schniepp, and Watson (1984), Paarsch (1997), Haile and Tamer (2003), Li and Perrigne (2003), and Aradillas-Lopez, Gandhi, and Quint (2013).

12 A similar technique is used in Haile and Tamer (2003). As noted there, unlike the requirement in Myerson (1981) that the inverse hazard rate is strictly monotone in \( \rho \), not \( U \), this procedure only requires...
of reasonable values for $v_0$ and let these values vary with the value of the car by specifying $v_0 = \gamma R$, for an assumed $\gamma$.

Table 6 presents comparisons of estimates of the optimal reserve prices controlling and not controlling for unobserved heterogeneity for different values of $\gamma$. Let $R^*_C$ and $R^*_NC$ be the optimal reserve prices when the distribution of values is estimated controlling and not controlling for unobserved heterogeneity, respectively. For each value of $\gamma$, table 6 shows the average of $R^*_C - R^*_NC$ across all cars, and $R^*_C - R^*_NC$ for cars with unobserved heterogeneity within 0.5 standard deviations of the median realization of unobserved heterogeneity across all cars. The results in table 6 are intuitive. Ignoring unobserved heterogeneity tends to overstate the dispersion in values and this will tend to bias the optimal reserve price upward.\footnote{Formally, a sufficient, but not necessary, condition for this result is that the estimated distribution of values ignoring unobserved heterogeneity hazard rate dominates that controlling for unobserved heterogeneity, provided the usual regularity condition (see Proposition 1) holds. To see this, let $\lambda_C(\cdot)$ and $\lambda_{NC}(\cdot)$ be the hazard rates of the distribution of values controlling and not controlling for unobserved heterogeneity, respectively. The assumptions of hazard rate dominance gives $\lambda_C(r) > \lambda_{NC}(r)$ for all $r$. Suppose that $r^*_C = v_0 + \frac{1}{\lambda_C(r^*_C)}$. Then hazard rate dominance delivers that $r^*_C - \frac{1}{\lambda_{NC}(r^*_C)} < v_0$. Regularity assures that for a value $r^*_NC$ such that $r^*_NC = v_0 + \frac{1}{\lambda_{NC}(r^*_NC)}$, it must be that $r^*_NC > r^*_C$.} Focusing on the second column, which is less subject to relatively high or low realizations of unobserved heterogeneity than the first column, the difference in optimal reserve prices tends to grow in $v_0$. In general, a reserve price is more valuable to a seller the greater is $v_0$ since the cost of failing to sell the item is lower. This effect, combined with the upward pressure that increased dispersion in values places on reserve prices, rationalizes the results that the difference in optimal reserve prices tends to grow in $v_0$.

The remaining columns in table 6 estimate the fraction of value the seller is able to capture by selling his car at the auction house. Revenues are estimated via simulation (using 10,000 simulations) with different reserve prices assuming that $N = 8$. In general, the reserve price set by the auction house will not be the individual seller’s optimal reserve price. This may occur for a variety of reasons including the fact that the auction house sells multiple automobiles per day and faces at least some competition from other auction
houses (in other words, the auction house, which sets the reserve price, is not a monopolist selling a single good). In addition, the seller must pay the auction house an up-front fee and a commission. Despite these deviations from the individual seller’s optimal mechanism, he does quite well when selling his automobile through the auction house. Focusing on the columns giving statistics of the fraction of value captured by a seller when the actual reserve prices are used, on average sellers obtain around 94% of the value they would if they could hold the optimal auction with this set of potential bidders. There is some variation in how well the sellers do, but 90% of them achieve more than 83% of the optimal value. These results are not very sensitive to assumptions about $\gamma$. Highlighting the importance of controlling for unobserved heterogeneity when evaluating this auction platform’s value to sellers, the next set of columns show that if the auction house used the “optimal” reserve prices based on ignoring unobserved heterogeneity, on average the sellers would capture even less of the maximum value than they do using the actual reserve prices.

I now explore the channels through which an auction platform, such as the one studied here, generates value for individual sellers and how ignoring unobserved heterogeneity impacts assessment of the importance of these channels. First, the auction house creates a centralized location and transparent price discovery mechanism through which sellers can easily interact with buyers. Absent the platform, the seller must reach and interact with individual buyers. This process is obviously costly for the seller and may be additionally taxing if they wish to avoid negotiations. Since it is an open question as to how a seller would reach a set of potential buyers if they could not sell at the auction house, and what would be his cost of doing so, I avoid imposing a specific sales process (such as, for example, a series of take-it-or-leave-it offers).

Instead, I focus on the fact that the auction house almost surely enlarges the set of buyers that a seller could reach on his own without bearing substantial costs. Specifically, I estimate $\pi_{R=0}^{N+1} - \pi_{R=0}^N$, where $\pi_{R=0}^N$ is the revenue from holding an auction without a reserve price with $N$ bidders. As mentioned above, $\pi_{R=0}^{N+1}$ can also be interpreted as an upper bound on a seller’s
revenue from negotiating with $N$ bidders on his own given that Bulow and Klemperer (1996) show that no amount of bargaining power with $N$ buyers is worth as much as adding one more buyer and holding an auction with no reserve price.

Finally, in practice sellers can set reserve prices at the auction house. This transfers some amount of the bargaining power, which the last channel assumed away, back to sellers. So the third, and final, channel through which sellers benefit comes from the seller setting a reserve price:

$$\pi_{R=R^*}^{N+1} - \pi_{R=0}^{N+1} = 0.$$

Table 7 shows a breakdown of the importance of these channels, both when unobserved heterogeneity is controlled for and when it is not, for various number of bidders, assuming here that $\gamma = 0$. The first row in each panel displays the average increase in seller revenue, based on 10,000 simulations, when an auction with $N+1$ bidders and an optimal reserve price is used as compared to when the seller holds an auction with $N$ buyers and no reserve price. The next rows in each panel of the table display the fraction of this difference attributable to adding another bidder and setting an optimal reserve price, respectively. Clearly the advantage of the auction with an extra bidder and an optimal reserve price decreases in the number of buyers, as should be expected. However, the fraction for which each channel is responsible is not constant. As $N$ increases, the first channel, the simultaneous competition created by the auction house, becomes relatively more important, while the third, the ability to set an optimal reserve price, becomes relatively inconsequential. Across $N$, the ability of the auction house to attract additional bidders is valued by sellers, and by a much greater amount than were they able to set an optimal reserve price. For example, when there are five bidders, the value of adding an additional bidder is worth approximately 17.5 times on average as much to the seller as being able to set an optimal reserve price with the original set. In other words, the relative effect of adding another bidder, as compared to setting an optimal reserve price, shows that the classic theoretical result of Bulow and Klemperer (1996) is economically significant in a relevant empirical setting.

The second panel of table 7 shows the importance of each channel when unobserved heterogeneity is not controlled for.
heterogeneity is ignored. The main difference centers on the importance of reserve prices. Consistent with the findings in table 6, when unobserved heterogeneity is ignored, reserve prices have a greater impact on seller revenues. Moreover, the relative importance of adding another bidder as compared to setting a reserve price is also understated. Now, when there are five bidders, the ability to add another bidder, even if no reserve price is set, is worth only 4.6 times as much as the ability to set an optimal reserve with the original set of bidders.

5 Conclusion

The existing empirical auction literature has improved our understanding of how a variety of auctions operate in practice and has answered questions of practical importance such as revenue maximizing design, the detection of collusion and the estimation of firm markups. Much of this literature, however, ignores unobserved heterogeneity. While several recent attempts have been made to address this empirical issue, the current article proposes a new methodology that relies on using information from the behavior of individuals other than bidders (who are the only auction participants currently used in efforts to control for unobserved heterogeneity), namely sellers, to help control for what the researcher does not observe in the data.

Oftentimes such seller behavior is available. The solution relies on a strict monotonicity assumption in the seller behavior, which is intuitively plausible and, while not required in the model, generally is the outcome of optimal seller behavior. Evidence from used car auctions clearly shows the importance of controlling for unobserved heterogeneity. Estimates of the optimal reserve price, which is a key tool of mechanism design and a frequently analyzed object in the empirical auction literature, will depend on whether unobserved heterogeneity is controlled for.

Given that the data come from individuals using an auction platform to sell their vehicles, a natural question is to what extent sellers are able to capture the value from selling their
automobile in this environment. On the one hand, the auction house will not generally set
the same reserve price that the individual would set if he could hold an optimal auction. On
the other hand, sellers clearly benefit from the auction house’s formal bidding environment,
which allows them to avoid the pains of negotiating, and gives them access to a large set
of buyers, who may be costly to interact with absent the platform. One way to gauge the
performance of the platform from a seller’s perspective is to estimate a seller’s value of the
platform’s auction and compare this to the seller’s value of holding the optimal auction with
the same set of bidders. By this benchmark, the auction platform allows sellers to capture
a large fraction, on average over 90%, of this maximum potential value. However, this likely
understates the benefit of the auction platform to individual sellers since it is costly for
sellers to locate and interact with the same set of buyers on their own. To support the claim
that access to more bidders particularly benefits sellers, I show that increasing the number
of bidders substantially improves seller revenues even if they cannot set any reserve price. I
also show that the benefit of an additional bidder relative to setting an optimal reserve price
depends on whether unobserved heterogeneity is controlled for, and if it is not, I find that
the relative gain will be understated.

One path for future work is to incorporate multi-dimensional unobserved heterogeneity.
I speculate that the identification result above can be generalized to this setting. Suppose,
for example, that the signal $U \in \mathbb{R}^{L^2}, L^1 > 1$. One could require that there exist as many
policy variables (denoted by what is now a vector $R$) as there are dimensions of $U, L^1$, and
that for each component of $R \in \mathbb{R}^{L^2}, L^2 \geq L^1$, say $R_t$, there exist functions $R_t = m_t(Z, U_t)$
which are monotonic in $U_t$. Then, if we allow the bids to depend on the entire vector $U$,
we can simply perform the “first part” of the above proof $L^1$ times to each of the functions
$m(\cdot)$ in order to uncover all of the realizations of unobserved heterogeneity, but this remains
to be shown. One setting that is potentially appropriate for such a model is eBay, where
sellers can choose not only reserve prices, but also auction length, whether or not to display
pictures of the item, whether or not to have the object insured, shipping charges, etc.
References


Auctions Under Unobserved Auction Heterogeneity,” Review of Economic Studies, 78(1),
293–327.

Laffont, J.-J., H. Ossard, and Q. Vuong (1995): “Econometrics of First Price Auc-
tions,” Econometrica, 63(4), 953–980.


Li, T., I. Perrigne, and Q. Vuong (2000): “Conditionally Independent Private Infor-

Li, T., and Q. Vuong (1998): “Nonparametric Estimation of the Measurement Error
Model Using Multiple Indicators,” Journal of Multivariate Analysis, 65(2), 139–165.


Mead, W. J., M. Schniepp, and R. Watson (1981): “The Effectiveness of Compe-
tition and Appraisals in the Auction Markets for National Forest Timber in the Pacific

——— (1984): “Competitive Bidding for U.S. Forest Service Timber in the Pacific North-

Econometrica, 50(5), 1089–1122.

6(1), 58–73.


### Tables and Figures

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>25\textsuperscript{th}-tile</th>
<th>Median</th>
<th>75\textsuperscript{th}-tile</th>
</tr>
</thead>
<tbody>
<tr>
<td>WINNING BID (in $)</td>
<td>3661.83</td>
<td>2889.37</td>
<td>1450.00</td>
<td>3260.00</td>
<td>5150.00</td>
</tr>
<tr>
<td>RESERVE PRICE (in $)</td>
<td>3361.26</td>
<td>2842.75</td>
<td>1200.00</td>
<td>3000.00</td>
<td>4800.00</td>
</tr>
<tr>
<td>WINNING BID/RESERVE PRICE</td>
<td>1.20</td>
<td>0.44</td>
<td>1.03</td>
<td>1.09</td>
<td>1.20</td>
</tr>
<tr>
<td>WINNING BID - RESERVE PRICE</td>
<td>300.57</td>
<td>294.89</td>
<td>90.00</td>
<td>240.00</td>
<td>450.00</td>
</tr>
<tr>
<td>NUMBER OF UNIQUE BIDDERS</td>
<td>5.37</td>
<td>2.40</td>
<td>3.00</td>
<td>5.00</td>
<td>7.00</td>
</tr>
<tr>
<td>AGE (in months)</td>
<td>73.60</td>
<td>28.79</td>
<td>56.00</td>
<td>76.00</td>
<td>95.00</td>
</tr>
<tr>
<td>MILEAGE (in 100,000 km’s)</td>
<td>0.96</td>
<td>0.48</td>
<td>0.64</td>
<td>0.95</td>
<td>1.25</td>
</tr>
</tbody>
</table>

**Table 1:** Auction summary statistics. There are 4,943 observations of each variable and the time period is 10/10/2001-11/27/2002. During the period of study there were 28,334 passenger cars auctioned. The types of car auctioned were sedans, hatchbacks, jeeps, small minivans, minibuses, small trucks and motorbikes. Unlike the U.S., SUVs were not prevalent in South Korea at this time. In an effort to maintain a homogenous data set, the sample summarized in this table consists of 5 passenger sedans with automatic transmissions that had been previously owned for “regular” individual purposes, as opposed to being part of government or rental company fleets. All variables are measured at the time of auction. Here AGE is the number of months since the car was originally manufactured and MILEAGE is the reading on the odometer in 100,000s of kilometers.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>-0.026***</td>
<td>-0.026***</td>
<td>-0.023***</td>
<td>-0.023***</td>
<td>-0.023***</td>
<td>-0.005***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0002)</td>
<td></td>
</tr>
<tr>
<td>MILEAGE</td>
<td>0.114***</td>
<td>0.126***</td>
<td>-0.032***</td>
<td>-0.037**</td>
<td>-0.030*</td>
<td>-0.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.02)</td>
<td>(0.012)</td>
<td>(0.018)</td>
<td>(0.017)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>NUMBER OF BIDDERS</td>
<td>0.034***</td>
<td>0.024***</td>
<td>0.022***</td>
<td>0.022***</td>
<td>0.027***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.0008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOG RESERVE PRICE</td>
<td></td>
<td></td>
<td></td>
<td>0.711***</td>
<td>0.871***</td>
<td>(0.005)</td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAKE DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MODEL DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MAKE/MODEL DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MAKE/MODEL DUMMIES×MILEAGE DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>COLOR DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MONTH DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>T</td>
<td>4,943</td>
<td>4,943</td>
<td>4,943</td>
<td>4,943</td>
<td>4,943</td>
<td>4,943</td>
<td>4,943</td>
</tr>
<tr>
<td>R²</td>
<td>0.623</td>
<td>0.633</td>
<td>0.889</td>
<td>0.896</td>
<td>0.903</td>
<td>0.979</td>
<td>0.966</td>
</tr>
</tbody>
</table>

Table 2: Winning bid determinants. OLS Regressions where the dependent variable is the log of the winning bid in each auction. The time period is from 10/10/2001-11/27/2002. MAKE/MODEL DUMMIES×MILEAGE DUMMIES means that indicators for the make and model of the vehicle are interacted with indicators for whether the car’s mileage is between the 25th and 75th percentile of mileage (0.64 and 1.25, respectively), or greater than the 75th percentile of mileage. Interacting the make and model dummies with a finer parsing of mileage does not materially affect any of the results. *, **, *** imply significance at the 10%, 5% and 1% level, respectively.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>-0.026***</td>
<td>-0.005***</td>
<td>-0.026***</td>
<td>-0.005***</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
<td>(0.0007)</td>
<td>(0.0005)</td>
<td>(0.0007)</td>
<td>(0.0005)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>MILEAGE</td>
<td>-0.024</td>
<td>0.013</td>
<td>-0.024</td>
<td>0.013</td>
<td>(0.027)</td>
<td>(0.027)</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.027)</td>
<td>(0.027)</td>
<td>(0.027)</td>
<td>(0.027)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>NUMBER OF BIDDERS</td>
<td>0.028***</td>
<td>0.031***</td>
<td>0.028***</td>
<td>0.031***</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>LOG RESERVE PRICE</td>
<td></td>
<td>0.708***</td>
<td>0.887***</td>
<td>0.708***</td>
<td>0.891***</td>
<td>(0.012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.012)</td>
<td>(0.005)</td>
<td>(0.012)</td>
<td>(0.005)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>MAKE DUMMIES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MODEL DUMMIES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MAKE/MODEL DUMMIES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MAKE/MODEL DUMMIES×MILEAGE DUMMIES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>COLOR DUMMIES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>MONTH DUMMIES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

|                  | 1,117        | 1,117        | 1,117        | 1,117        | 1,117        | 1,117        |
|                  | 0.910        | 0.980        | 0.959        | 0.910        | 0.980        | 0.954        |
| Dependent Variable| Max Bid      | Max Bid      | Max Bid      | Mean Bid     | Mean Bid     | Mean Bid     |

Table 3: Highest bid and mean bid determinants. OLS Regressions of the highest bid and mean bid submitted in auctions between 10/10/2001-12/27/2001. MAKE/MODEL DUMMIES×MILEAGE DUMMIES means that indicators for the make and model of the vehicle are interacted with indicators for whether the car's mileage is between the 25th and 75th percentile of mileage (0.64 and 1.25, respectively), or greater than the 75th percentile of mileage. Interacting the make and model dummies with a finer parsing of mileage does not materially affect any of the results. *, **, *** imply significance at the 10%, 5% and 1% level, respectively.
### Table 4: Failure to sell determinants. Logit model of failure to sell for automobiles auctioned between 10/10/2001-12/27/2001. 107 out of 1,117 cars failed to sell. Coefficient estimates reported. *, **, *** imply significance at the 10%, 5% and 1% level, respectively.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>0.008*</td>
<td>0.005</td>
<td>-0.002</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>MILEAGE</td>
<td>-0.911***</td>
<td>-0.859***</td>
<td>-0.854***</td>
<td>-0.841***</td>
</tr>
<tr>
<td></td>
<td>(0.311)</td>
<td>(0.322)</td>
<td>(0.326)</td>
<td>(0.328)</td>
</tr>
<tr>
<td>LOG RESERVE PRICE</td>
<td></td>
<td></td>
<td></td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.263)</td>
</tr>
<tr>
<td>MAKE DUMMIES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>MODEL DUMMIES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>COLOR DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>MONTH DUMMIES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>$T$</td>
<td>1,117</td>
<td>1,089</td>
<td>1,089</td>
<td>1,089</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.0133</td>
<td>0.0348</td>
<td>0.0671</td>
<td>0.0682</td>
</tr>
</tbody>
</table>

Table 5: Auction summary statistics for five passenger Hyundai Sonatas with automatic transmissions that sold during the months cited in the text. There are 454 observations of each variable. All variables are measured at the time of auction. Here AGE is the number of months since the car was originally manufactured, MILEAGE is the reading on the odometer in 100,000s of kilometers.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>25th-tile</th>
<th>Median</th>
<th>75th-tile</th>
</tr>
</thead>
<tbody>
<tr>
<td>WINNING BID (in $)</td>
<td>3574.38</td>
<td>2244.73</td>
<td>2410.00</td>
<td>3230.00</td>
<td>4670.00</td>
</tr>
<tr>
<td>RESERVE PRICE (in $)</td>
<td>3320.22</td>
<td>2227.41</td>
<td>2000.00</td>
<td>3000.00</td>
<td>4400.00</td>
</tr>
<tr>
<td>WINNING BID/RESERVE PRICE</td>
<td>1.16</td>
<td>0.34</td>
<td>1.02</td>
<td>1.07</td>
<td>1.17</td>
</tr>
<tr>
<td>WINNING BID - RESERVE PRICE</td>
<td>254.16</td>
<td>239.72</td>
<td>60.00</td>
<td>180.00</td>
<td>390.00</td>
</tr>
<tr>
<td>NUMBER OF UNIQUE BIDS</td>
<td>4.83</td>
<td>1.89</td>
<td>3.00</td>
<td>5.00</td>
<td>6.00</td>
</tr>
<tr>
<td>AGE (in months)</td>
<td>84.10</td>
<td>23.71</td>
<td>69.00</td>
<td>86.50</td>
<td>101.00</td>
</tr>
<tr>
<td>MILEAGE (in 100,000 km’s)</td>
<td>1.13</td>
<td>0.46</td>
<td>0.85</td>
<td>1.10</td>
<td>1.41</td>
</tr>
</tbody>
</table>
Table 6: Comparisons of the optimal reserve prices when unobserved heterogeneity is controlled for and when it is not. $\gamma$ is the assumed multiple to infer $v_0$: $\gamma R = v_0$, where $R$ is the observed reserve price. The first column shows the average of the difference in the optimal reserve price ignoring unobserved heterogeneity and the optimal reserve price controlling for unobserved heterogeneity. The second column shows this average difference for cars with unobserved heterogeneity within 0.5 standard deviations of the median realization of unobserved heterogeneity across all cars. The remaining set of columns show statistics of the fraction of value captured by a seller when the actual or “optimal” reserve price ignoring unobserved heterogeneity is used. Standard errors, which for percentage estimates are in percentage points, are based on 100 bootstrapped repetitions.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>Overall $R_C^* - R_{NC}^*$ Mean</th>
<th>Median $R_C^* - R_{NC}^*$ Mean</th>
<th>% of Value Captured with Actual Reserves</th>
<th>% of Value Captured with Sub-Optimal Reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma R = v_0$</td>
<td>$\gamma R = v_0$</td>
<td>Percentiles</td>
<td>Percentiles</td>
</tr>
<tr>
<td></td>
<td>$v_0$</td>
<td>$v_0$</td>
<td>Mean</td>
<td>$10^{th}$</td>
</tr>
<tr>
<td>0.00</td>
<td>$221.57$</td>
<td>$77.21$</td>
<td>$88.42%$</td>
<td>$83.69%$</td>
</tr>
<tr>
<td></td>
<td>(25.36)</td>
<td>(24.89)</td>
<td>(0.19)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>0.25</td>
<td>$258.52$</td>
<td>$90.20$</td>
<td>$90.66$</td>
<td>$86.69$</td>
</tr>
<tr>
<td></td>
<td>(23.16)</td>
<td>(25.52)</td>
<td>(0.18)</td>
<td>(0.38)</td>
</tr>
<tr>
<td>0.50</td>
<td>$239.72$</td>
<td>$98.45$</td>
<td>$92.23$</td>
<td>$86.98$</td>
</tr>
<tr>
<td></td>
<td>(21.35)</td>
<td>(21.01)</td>
<td>(0.20)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>0.75</td>
<td>$236.03$</td>
<td>$123.57$</td>
<td>$91.81$</td>
<td>$85.84$</td>
</tr>
<tr>
<td></td>
<td>(21.79)</td>
<td>(22.12)</td>
<td>(0.20)</td>
<td>(0.76)</td>
</tr>
</tbody>
</table>
Table 7: Channels by which sellers benefit from the auction. The table shows a breakdown of a seller’s value from auctioning a car at the auction house. The revenues from holding an auction without a reserve price with the same set of buyers is $\pi_{R=0}^N$, and with an optimal reserve price is $\pi_{R=R^*}^N$. Standard errors, which for percentage estimates are in percentage points, are based on 100 bootstrapped repetitions.
Figure 1: Estimated value distributions, not controlling and controlling for unobserved heterogeneity. Seven year old automatic, five passenger, Hyundai Sonatas with 100,000 KMs. The solid distribution is the estimated distribution of values ignoring unobserved heterogeneity. The other four distributions control for unobserved heterogeneity at the given percentiles.
A Monte Carlo Study

In this section I provide a small simulation study of the non-parametric two step estimation procedure. I generate reserve price and bidding data for \( T \) auctions and then repeat \( S \) times. For the reserve price, I set \( R = m(Z_0, Z_1, U) = Z_1(\gamma_0 Z_0 + \gamma_1 Z_1) + U \), where (i) \( Z_0 = z^H_0 \) with probability \( p \) and \( Z_0 = z^L_0 \) with probability \( 1 - p \), (ii) \( Z_1 \sim N(\mu^H_{Z_1}, \sigma_{Z_1}) \) when \( Z_0 = z^H_0 \) and \( Z_1 \sim N(\mu^L_{Z_1}, \sigma_{Z_1}) \) when \( Z_0 = z^L_0 \) and (iii) \( U \sim N(\mu^H_U, \sigma_U) \) when \( Z_0 = z^H_0 \) and \( U \sim N(\mu^L_U, \sigma_U) \) when \( Z_0 = z^L_0 \). Given this specification for \( m(\cdot) \), reserve prices are strictly increasing in \( U \) and \( u = m(z_0, 0, u) \) for all \( z_0 \) and \( u \). In addition, for this data generating process, \( U \) is independent of \( Z_1 \) conditional on \( Z_0 \). For each auction \( t = 1, ..., T \), I simulate an English button auction by setting \( N_t = 4 \), drawing values such that \( V_t \sim N(\alpha_0 + \alpha_1 U_t, \sigma_V) \) and set the winning bid \( W_t \) equal to the second highest value. Given this specification, bids are stochastically increasing in the realization of unobserved heterogeneity.

For each auction \( t \), I assume that I only observe \( \{R_t, Z_{0t}, Z_{1t}, W_t\} \). I experiment with \( T = 250 \) and \( T = 500 \). For both cases, \( S = 100 \), \( p = 0.5 \), \( z^L_0 = -1 \), \( z^H_0 = 1 \), \( \mu^H_{Z_1} = -2.5 \), \( \mu^L_{Z_1} = 2.5 \), \( \sigma_{Z_1} = 1 \), \( \mu^L_U = -4.5 \), \( \mu^H_U = 4.5 \), \( \sigma_U = 2.5 \), \( \gamma_0 = 1 \), \( \gamma_1 = 2 \), \( \alpha_0 = 100 \), \( \alpha_1 = 1 \) and \( \sigma_V = 5 \). The results for estimates of \( m(\cdot) \) and \( F_U \) appear in table 8.

To investigate the performance of the estimator of the distribution of values, I consider estimating the case where \( U = 0 \). I do this based on the first step estimates as well as assuming that I observe the realizations of \( U \) for each auction, which generally is infeasible. These cases are referred to as the “No Information” (NI) and “Full Information” (FI) cases, respectively. In general, the NI case compares favorably to the generally infeasible full information case. Additionally, increasing the sample size from 250 to 500 improves the quality of the estimates.
\[ T = 250 \quad |\text{Bias}| \quad \text{Variance} \quad \text{MSE} \quad |\text{Bias}| \quad \text{Variance} \quad \text{MSE} \]

<table>
<thead>
<tr>
<th></th>
<th>( T = 250 )</th>
<th>( T = 500 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( m )</td>
<td>0.4669</td>
</tr>
<tr>
<td></td>
<td>( F_U )</td>
<td>0.0168</td>
</tr>
<tr>
<td></td>
<td>( F_{V</td>
<td>U=0,^+/^-} )</td>
</tr>
<tr>
<td></td>
<td>( F_{V</td>
<td>U=0,^/-/+} )</td>
</tr>
</tbody>
</table>

Table 8: Simulation results. \( S = 100 \), \( p = 0.5 \), \( z_0^L = -1 \), \( z_0^H = 1 \), \( \mu_{Z_1}^H = -2.5 \), \( \mu_{Z_1}^L = 2.5 \), \( \sigma_{Z_1} = 1 \), \( \mu_U^L = -4.5 \), \( \mu_U^H = 4.5 \), \( \sigma_U = 2.5 \), \( \gamma_0 = 1 \), \( \gamma_1 = 2 \), \( \alpha_0 = 100 \), \( \alpha_1 = 1 \) and \( \sigma_V = 5 \). The second to last row presents estimates of the distribution of values assuming \( U = 0 \) based on the first stage estimates. The last row assumes that the actual realizations of \( U_t \) are observed when estimating the distribution of values conditional on \( U = 0 \).

### B Proofs

#### Proposition 1

**Proof.** Affiliation between \( V \) and \( U \) with joint density \( f \) provides that for all \( v_1 < v_2 \) and \( u_1 < u_2 \):

\[
f(v_1, u_2)f(v_2, u_1) < f(v_1, u_1)f(v_2, u_2)
\]

Rearranging and making use of the corresponding conditional distributions:

\[
\frac{f(v_2|u_1)f(u_1)}{f(v_1|u_1)f(u_1)} < \frac{f(v_2|u_2)f(u_2)}{f(v_1|u_2)f(u_2)}
\]

and therefore by integrating with respect to \( v_2 \) yields:

\[
\frac{1 - F(v_1|u_1)}{f(v_1|u_1)} < \frac{1 - F(v_1|u_2)}{f(v_1|u_2)}
\]

Therefore, if \( r_k^* \) is the solution when \( U = u_k \), \( k = 1, 2 \), we have \( \frac{1 - F(r_1^*|u_1)}{f(r_1^*|u_1)} < \frac{1 - F(r_2^*|u_2)}{f(r_2^*|u_2)} \). Since \( \psi(t, U) \) is increasing in \( t \) for all \( U \), it must be that \( r_1^* < r_2^* \).

Q.E.D.
Proposition 2

Proof. Beginning with reserve prices, we can identify the conditional distribution of unobserved heterogeneity as follows: $F_{R|(Z_0,Z_1)=(z_0,z_1)}(u) = Pr[R < u|(Z_0,Z_1) = (z_0,z_1)] = Pr[m(Z_0,Z_1,U) < u|(Z_0,Z_1) = (z_0,z_1)] = Pr[U < u|(Z_0,Z_1) = (z_0,z_1)] = Pr[U < u|Z_0 = z_0] = F_{U|Z_0=z_0}(u)$, where the third equality comes from the normalization of $m$ in assumption (i), and the fourth comes from the conditional independence assumption (ii). Also, we note that: $F_{U|Z_0=z_0}(u) = Pr[U < u|Z_0 = z_0] = Pr[U < u|(Z_0,Z_1) = (z_0,z_1)] = Pr[m(Z_0,Z_1,U) < m(z_0,z_1,u)|(Z_0,Z_1) = (z_0,z_1)] = Pr[R < m(z_0,z_1,u)|(Z_0,Z_1) = (z_0,z_1)] = F_{R|(Z_0,Z_1)=(z_0,z_1)}(m(z_0,z_1,u))$, where the second equality again uses independence and the third uses the monotonicity of $m(\cdot,U)$. Therefore, the function $m$ is identified since:

$$m(z_0,z_1,u) = F_{R|(Z_0,Z_1)=(z_0,z_1)}^{-1}(F_{R|(Z_0,Z_1)=(z_0,z_1)}(F_{R|(Z_0,Z_1)=(z_0,z_1)}^{-1}(u)))$$

Now, with data on reserve prices, $R$, and item characteristics $Z$, due to the strict monotonicity of $m(\cdot,U)$, for each $(R,Z)$ pair there will be a unique realization of $U$ satisfying $R = m(Z,U)$.

Let $W_0 = [Z,U]$. Using only transaction prices, we can identify $F(V|W_0)$ in a first price and second price auction using the arguments of Theorems 6 and 1, respectively, in Athey and Haile (2002).

Q.E.D.