Bounds in Auctions with Unobserved Heterogeneity

Timothy B. Armstrong*
Stanford University

April 26, 2012

Abstract

Many empirical studies of auctions rely on the assumption that the researcher observes all variables that make auctions differ ex ante. When there is unobserved heterogeneity, the direction of the bias this causes is known only in a few restrictive examples. In this paper, I show that ignoring unobserved heterogeneity in a first price sealed bid auction with symmetric independent private values gives bounds on several quantities of economic interest under surprisingly general conditions. The results apply to certain quantities related to expectations of valuations, including bidder profits (which can be used to recover bid preparation costs in entry models) and the efficiency loss of assigning the object randomly. I then turn to estimation of these bounds, and show that, when only the winning bid is available, the rate of convergence can be slower than the square root of the number of auctions observed and depends on the number of bidders. These results apply more generally to estimation of functionals of a distribution from repeated observations of an order statistic and may be of independent interest. I apply these methods to bound the efficiency loss from replacing a set of procurement auctions for highway construction in Michigan with random assignment.

*email: timothya@stanford.edu. Thanks to Han Hong and Liran Einav for support and helpful discussions, and to Jon Levin, Jakub Kastl, the editor and three anonymous referees, and participants at the Stanford IO Seminar and Third Year Seminar for comments. I also thank Liran Einav and Ignacio Esponda for sharing the data from the Michigan Department of Transportation. All remaining errors are my own. This paper was written with generous support from a fellowship from the endowment in memory of B.F. Haley and E.S. Shaw through the Stanford Institute for Economic Policy Research.

1 Introduction

Since the seminal work of Guerre, Perrigne, and Vuong (2000), numerous empirical studies of auctions have used models of optimizing behavior to recover bidders' values. If the researcher observes outcomes of many auctions that differ only in observable variables, many standard auction models provide enough structure to completely recover a bidder's decision problem. The optimal bid as a function of the bidder's type can then be inverted to recover the distribution of bidder values from observed bids.

A crucial assumption of this approach is that the researcher is able to observer multiple auctions in which participants differ ex ante across auctions only in observable ways. This amounts to assuming that the only difference between the knowledge of the econometrician and auction participants about their environment comes from sampling error. In most cases, a more realistic assumption is that bidders have knowledge about aspects of the object being auctioned or the preferences of other bidders that vary across auctions and that the researcher cannot observe. This problem of unobserved heterogeneity makes estimates that do not take it into account inconsistent.

While methods have been proposed for consistently estimating bidder valuations under unobserved heterogeneity, these methods require data on multiple bids (typically, data on all bids is required, since a strict subset of bids can only be used if they are from the same bidders, rather than, say, the top three bids) or instruments that can be used to condition on it, as well as additional assumptions on the form of the unobserved heterogeneity. With only data on the winning bid or a single bidder, it is not clear what can be learned.

In this paper, I show that, in a symmetric first price sealed bid auction with independent private values, several economic primitives of interest in applications can be bounded using only data on a single bidder or data on the winning bid. The results apply to certain quantities related to expectations of valuations and include expected bidder profits, entry costs, and the long run loss in surplus of replacing the auction with random or arbitrary allocation. Many of these bounds are simply the naive estimates that do not take unobserved heterogeneity into account. When only data on a single bid is available and a model with no unobserved heterogeneity cannot be ruled out, these bounds are sharp. While the bounds confirm existing intuition about the direction of bias caused by certain types of unobserved heterogeneity, I show that they hold with surprising generality. The unobserved heterogeneity can take essentially any form, as long as the auction falls under the symmetric independent private values framework after conditioning on it.

I then turn to the estimation of these bounds. The estimation of the efficiency loss from random allocation requires the estimation of the mean of the bid distribution. While this is a trivial problem with data from a single bidder, I show that this problem becomes much more difficult with repeated observations of the winning bid. In contrast to the case of data on a single bidder or data on all bids, with only data on the winning bid, the mean of the bid distribution (or the value distribution) cannot always be estimated at a rate proportional to the square root of the number of auctions observed. I derive the rate of convergence, and conditions under which the limiting variable is normal. These depend on the number of bidders and the shape of the distribution of the winning bid. While I develop these results for first price auctions, they apply more generally to estimation of the mean or other statistics of a distribution from repeated observations of the greatest order statistic, and similar ideas will apply to other order statistics, such as repeated observations of the second largest value in an ascending auction. I give more detailed intuition in Section 3 and formal proofs in the appendix, but the basic intuition is that, unlike estimating a fixed quantile from repeated observations from an order statistic, estimating the mean requires knowledge of the tails, which are observed infrequently when only an order statistic is observed. The sample analogue requires placing increasing weight on observations at the tail, which can lead to nonstandard rates of convergence and asymptotic distributions. In this sense, this issue is related to the slow rates of convergence derived by Khan and Tamer (2007), Khan and Nekipelov (2008), Andrews and Schafgans (1998), and others for parameters that are "identified at infinity," and has also been pointed out recently in the context of ascending auctions by Menzel and Morganti (2011), who take a different approach (see the discussion below). It is important to note that, while this problem with slower rates of convergence when only observing the winning bid arises in estimating statistics that must be interpreted as bounds under unobserved heterogeneity, it is distinct from the problem of unobserved heterogeneity. Even in the case of no unobserved heterogeneity, where the bounds being estimated are equal to the actual parameters of interest, rates of convergence can be slower when only observing the winning bid.

I apply these results to bound the surplus loss from replacing a subset of Michigan highway procurement auctions with a mechanism that leads to random or arbitrary assignment. As documented by Decarolis (2010), various forms of "average bid" auctions, which lead to random allocation in equilibrium, are common in highway procurement auctions, making this a particularly interesting application for this counterfactual policy change. I also provide estimates that use the methods of Krasnokutskaya (2009), which require additional data and assumptions, to account for unobserved heterogeneity. While the bounds that use only the winning bid are somewhat conservative compared to the estimates proposed by Krasnokutskaya (2009), they are still informative. Thus, while other methods are more accurate when the data is available and the researcher is willing to make additional assumptions, the bounds in this paper are useful when such data is not available.

While some authors have proposed methods for dealing with unobserved heterogeneity in first price auctions using data on multiple bidders or additional data, this paper is the first, to my knowledge, that provides bounds with only observations of the winning bid or bids from a single bidder. Krasnokutskaya (2009) and Asker (2008) use an additive or multiplicative separability assumption on bidder values as a function of the unobserved heterogeneity, while Athey, Levin, and Seira (2011) use a similar approach with a fully parametric model. All of these papers require data on multiple bids and, at least in the nonparametric case of Krasnokutskaya (2009), it is not clear how to extend these methods so that a strict subset of order statistics can be used, such as the two highest bids. Haile, Hong, and Shum (2003) provide a method that uses the assumption that the unobserved heterogeneity can be conditioned out using an additional variable, such as the number of potential bidders.

For ascending auctions, Aradillas-Lopez, Gandhi, and Quint (2010) propose bounds on revenues under counterfactual reserve prices under unobserved heterogeneity using variation in the number of bidders. In that setting, the second largest value can be observed despite the unobserved heterogeneity, and the challenge is in bounding the joint distribution of the two largest values while allowing for correlation induced by the unobserved heterogeneity. In contrast, a key difficulty in estimating first price auctions with unobserved heterogeneity is that none of the values can be recovered from the distribution of a single bid. While some of the results in the present paper use bounds on the correlation of the values, bounds on functions of the equilibrium markups play an important role.

The irregular nature of some estimation problems with repeated observations of a single bid has also been pointed out recently in the context of second price auctions by Menzel and Morganti (2011). While those authors focus on second price auctions and derive upper bounds on the attainable rate of convergence for estimating statistics of the marginal value distribution, the present paper treats first price auctions and derives rates of convergence and asymptotic distributions for the sample analogues of these statistics, thereby giving a lower bound on the attainable rate of convergence (by demonstrating estimators that attain it). In this sense, these papers are complementary.

The results in this paper largely confirm the intuition that unobserved heterogeneity tends to make first price auctions look "less competitive" by increasing the variance of the observed bid distribution. Suppose that, conditional on the unobserved heterogeneity, the bid distribution is concentrated near a single value, but this value changes throughout the sample, spreading out the bid distribution observed to the researcher. On average, bidders know that other bidders will bid near their values, making it difficult to get away with large markups. However, the researcher observes many bids that are much lower than the given valuation and, not taking into account that these bids are from a different type of auction, calculates that the bidder could get away with a large markup. This leads the researcher to overestimate markups and bidder profits, and to overestimate valuations (or underestimate costs in a procurement auction). While the conclusions of this informal argument have been confirmed in some examples arising from particular applications (for example Krasnokutskaya, 2009, finds this in an application to a set of highway procurement auctions similar to the ones used in this paper), it has not been clear whether it applies in any generality to the estimation of any economic primitives of interest in practice. In this paper, I develop bounds on several economic primitives of interest in applications and show that they hold under surprisingly general conditions. The results show that the intuition described above holds in a very general setup when applied to certain quantities involving expectations of valuations, which lead to bounds on, among other things, bidder profits and the loss total surplus from assigning the object randomly.

The plan for the paper is as follows. Section 2 defines the problem being studied and presents the bounds. Section 3 presents estimators for these bounds and derives asymptotic results that can be used for inference. Section 4 applies these bounds to a set of procurement auctions for highway construction projects, Section 5 discusses ways in which the results could be extended along with inherent limitations of the methods used in this paper and Section 6 concludes. The proof of the main result of Section 3 is given in an appendix.

2 Setup and Bounds

In this section, I define the empirical model used throughout the paper, and derive population bounds on quantities of economic interest. Most of these bounds use naive estimates that estimate the value distribution under the assumption of no unobserved heterogeneity. While bounds on the entire value distribution require more assumptions, the naive estimates can be used to bound certain functions of the mean of the value distribution and its order statistics. I begin by deriving bounds on the mean of the value distribution and the average value of the highest bidder. Using similar ideas, I then bound the expected profit from entering an auction and the efficiency loss from replacing the auction with a lottery. First, I present the general model used throughout the paper and define some notation.

I treat a general version of the symmetric independent private values (IPV) model with unobserved heterogeneity. Krasnokutskaya (2009), and Athey and Haile (2002) consider versions of this model. Conditional on a random variable U observed to bidders but not the researcher, n bidders draw iid values $V_1, \ldots, V_n$ from a distribution $F_V(v|U)$. Here, n is treated as nonrandom, but the results could also be stated with n random and the assumptions holding conditional on n. Each bidder observes her value and the number of bidders and submits a sealed bid. The object is awarded to the highest bidder for a price equal to the bid submitted by that bidder. As shown by Milgrom and Weber (1982), this model has a symmetric Nash equilibrium in which bids are increasing in values. This leads to bids $B_1, \ldots, B_n$ being iid conditional on U.

Krasnokutskaya (2009) considers a version of this model with the additional additive separability assumption that $V_i = A_i + U$ for some independent random variables A_i that are also independent of U (a multiplicative separability assumption gives the same result). She shows that the distributions of U and the A_i's can be recovered if the bids of at least two bidders are observed and proposes a consistent method of estimating these distributions from observed bids. However, the model is not identified if only one bid is observed in each auction, and it is not known whether the top two bids, or even any larger set of order statistics short of the full sample, are enough to recover any of these distributions. Even if enough bids are observed to use this method, it may be desirable to have estimates that do not impose additive or multiplicative separability assumptions on the value distribution.

It should be noted that, despite the general form of unobserved heterogeneity considered in the present paper, unobserved heterogeneity is not the only possible source of correlation between bids. The IPV assumption conditional on unobserved heterogeneity means that participants that values are independent conditional on bidders' information sets. More generally, bidders may have correlation in values even after conditioning on everything they observe. In such cases, the distribution of valuations is typically modeled as satisfying the affiliation property (see Milgrom and Weber, 1982), a certain strong notion of positive dependence, although more general forms of dependence are also possible (see de Castro, 2010). One possible motivation for a model in which values are correlated is one in which values are drawn independently conditional on some variable U, but, unlike the model considered here, bidders observe only their own valuations, and not U (see Li, Perrigne, and Vuong, 2000).

With data on multiple bids in each auction, tests developed by Jun, Pinkse, and Wan (2010) or de Castro and Paarsch (2010) could be used to determine whether values in a model with no unobserved heterogeneity satisfy the affiliation property. However, a completely general model with no unobserved heterogeneity and correlated values that may not satisfy affiliation cannot be ruled out empirically with existing results. In any case, the bounds in this paper are best suited to situations with data sets where only a single bid is recorded in each auction, which precludes empirical examinations of the correlation structure of bids. See Section 5.4 for some discussion of the limitations of the IPV assumption.

Let $G_{B_1}(b|U)$ be the distribution of a given bid conditional on U, and let $G_{B_1}(b) = E\left(G_{B_1}(b|U)\right)$ denote the marginal distribution of a single bid. Let $B_{(1)} \leq \ldots \leq B_{(n)}$ denote order statistics of the bid distribution and let $G_{B_{(k)}}(b|U)$ and $G_{B_{(k)}}(b)$ denote the cdf of the kth order statistic conditional on U and marginal respectively. Let $B_{(1)} \leq \ldots \leq B_{(n)}$ denote order statistics of the bid distribution and let $G_{B_{(k)}}(b|U)$ and $G_{B_{(k)}}(b)$ denote the cdf of the kth order statistic conditional on U and marginal respectively. Since the bounds depend on naive estimates of the bid distribution under the assumption of no unobserved heterogeneity, it is useful to define notation for the estimated marginal bid distribution under this assumption using data on the highest order statistic. If U were observed, the identity $G_{B_1}(b|U)^n = G_{B_{(n)}}(b|U)$ could be used to back out the marginal bid distribution from the distribution of the highest bid. Using this identity and the (possibly false) assumption of no unobserved heterogeneity, the researcher can estimate $\tilde{G}(t) \equiv G_{B_{(n)}}(t)^{\frac{1}{n}}$. For all of these distributions, I use the corresponding lowercase letters to denote the density.

The bounds derived below will depend on the support of the bid distribution. Let $\bar{b}(U) = \sup[supp[B_i|U]]$ and $\bar{b} = \sup[supp[B_i]]$ be the suprema of the support of the conditional and marginal bid distributions respectively, and let $\underline{b}(U)$ and $\underline{b}$ be defined similarly for the lower bound of the support. Inequalities stated below that involve $\bar{b}$ are interpreted as holding with $\bar{b} = \infty$ when the upper support point is infinite, and will hold trivially in these cases with the convention that $\infty \leq \infty$. The same convention holds for inequality statements with inequalities involving expecations of random variables with infinite expectations.

If U were observed, values could be estimated from observed bids conditional on each value of U using the celebrated formula $V_i = B_i + \frac{G_{B_1}(B_i|U)}{(n-1)g_{B_1}(B_i|U)}$ of Guerre, Perrigne, and Vuong (2000). Most of the results in this paper use naive estimates that replace the bid distribution conditional on variables observed to the bidder, $G_{B_1}(B_i|U)$, with the bid distribution conditional on variables observed to the researcher. If the researcher observes the same bidder in each auction, or observes all bids and ignores their dependence structure, this will mean replacing $G_{B_1}(B_i|U)$ with $G_{B_1}(B_i)$ (and the same for the pdf). If the researcher observes only the highest bid, this will mean replacing $G_{B_1}(B_i|U)$ with $\tilde{G}(t)$. Depending on which is used, the bounds derived below can be calculated with data on a single bidder or the highest bidder. If only this data is available, a data generating process with no unobserved heterogeneity cannot be ruled out, so the bounds are sharp. I first present bounds on the marginal value distribution using data on a single bidder.

2.1 Bounds on the Marginal Value Distribution

The following theorem shows that attempting to recover the value distribution using the marginal distribution of the bid of a single bidder gives an upper bound for the mean of the marginal value distribution. The mean of the marginal value distribution gives the expected surplus of a counterfactual mechanism that allocates the object randomly (see Section 2.4). It can also be used to obtain the social surplus from a nonrival technology that is a perfect substitute for the good being auctioned. For example, in an auction for emissions permits, $nE(V_i)$ would give the expected total surplus to all participants of a technology that allowed the participants to produce just as efficiently without polluting.

Theorem 1. *In the symmetric IPV model with unobserved heterogeneity considered above,*

$$E(B_i) \leq E(V_i) = E(B_i) + E\left(\frac{G_{B_1}(B_i|U)}{(n-1)g_{B_1}(B_i|U)}\right) \leq E(B_i) + E\left(\frac{G_{B_1}(B_i)}{(n-1)g_{B_1}(B_i)}\right)$$

If the researcher observes only B_i for a single $i \in \{1, \dots, n\}$ and $b \mapsto b + \frac{G_{B_1}(b)}{(n-1)g_{B_1}(b)}$ is strictly increasing, these bounds are sharp.

Proof. The mean of the value distribution conditional on U is

$$E(V_i|U) = E(B_i|U) + E\left(\frac{G_{B_1}(B_i|U)}{(n-1)g_{B_1}(B_i|U)}\middle| U\right) = E(B_i|U) + \int_{\underline{b}(U)}^{\bar{b}(U)} \frac{1}{n-1} G_{B_1}(b|U)\, db.$$

Taking expectations over U and changing the order of integration in the second term gives

(letting P_U be the probability distribution of U)

$$
\begin{aligned}
E(V_i) &= E(B_i) + \frac{1}{n-1}\iint I(\underline{b}(u) \le b \le \overline{b}(u)) G_{B_1}(b|u)\, dP_U(u)\, db \\
&\le E(B_i) + \frac{1}{n-1}\int I(\underline{b} \le b \le \overline{b}) \int G_{B_1}(b|u)\, dP_U(u)\, db \\
&= E(B_i) + \frac{1}{n-1}\int I(\underline{b} \le b \le \overline{b}) G_{B_1}(b)\, db = E(B_i) + E\left(\frac{G_{B_1}(B_i)}{(n-1)g_{B_1}(B_i)}\right).
\end{aligned}
$$

Here, the inequality follows because, since $G_{B_1}(b|u)$ is nonnegative, increasing the area of integration in the inner integral from $\{\underline{b}(u) \le b \le \overline{b}(u)\}$ to $\{\underline{b} \le b \le \overline{b}\}$ can only increase the value of the integral. Sharpness follows since the lower bound is achieved when V_i is degenerate given U, and the upper bound is achieved when there is no unobserved heterogeneity. $\square$

As noted by Athey and Haile (2007), the expression for the mean of value distribution can be written in terms of the mean of the bid distribution and the upper support of the bid distribution. This holds for the misspecified estimate considered here:

$$
\begin{aligned}
&E(B_i) + E\left(\frac{G_{B_1}(B_i)}{(n-1)g_{B_1}(B_i)}\right) = E(B_i) + \int_{\underline{b}}^{\overline{b}} \frac{1}{n-1} G_{B_1}(b)\, db = E(B_i) + \frac{1}{n-1}\left(\overline{b} - E(B_i)\right) \\
&= \frac{n-2}{n-1}E(B_i) + \frac{1}{n-1}\overline{b}.
\end{aligned}
$$

Since bidders never bid above valuations in a symmetric Nash equilibrium of this game, the above theorem along with this characterization of the misspecified estimate of the mean of the value distribution gives $\left[E(B_i), \frac{n-2}{n-1}E(B_i) + \frac{1}{n-1}\overline{b}\right]$ as bounds on the mean of the value distribution. The lower bound is reached if $V_i = U$ for all i, and the upper bound is reached if U is a constant, so, as long as the marginal bid distribution can be rationalized by a model with no unobserved heterogeneity, this bound is the tightest interval that can be obtained with data on a single participant's bids and the number of bidders. The requirement in the theorem that the estimated bid function be nondecreasing is a technical condition that ensures that the data can be rationalized by a model with no unobserved heterogeneity when only one bid is observed. The formulation in the above display also shows that the upper bound will be obtained even if there is nontrivial unobserved heterogeneity as long as the unobserved heterogeneity does not shift the upper support point of the bid distribution. Section A.5 provides an example of a data generating process where this is the case.

If only the winning bid and the number of bidders are observed, the marginal distribution of bids is not identified, so Theorem 1 cannot be used directly. However, note that

$$G_{B_1}(t) \geq G_{B_{(n)}}(t) = E\left(G_{B_{(n)}}(t|U)\right) = E\left(G_{B_1}(t|U)^n\right) \geq E\left(G_{B_1}(t|U)\right)^n = G_{B_1}(t)^n \qquad (1)$$

(see, for example, Shaked (1977)) so the marginal bid distribution can be bounded by

$$G_{B_{(n)}}(t) \leq G_{B_1}(t) \leq G_{B_{(n)}}(t)^{\frac{1}{n}} \equiv \tilde{G}(t). \qquad (2)$$

Where $\tilde{G}(t)$ is the misspecified marginal bid distribution defined above that uses the assumption of no unobserved heterogeneity . These bounds on the marginal bid distributions can be combined with Theorem 1 to compute bounds for the mean of the marginal value distribution by using $\tilde{G}(t)$ to compute the lower bound and $G_{B_{(n)}}$ to compute the upper bound. However, these bounds are likely to be crude since the lower bound obtains the marginal bid distribution assuming U is degenerate and obtains the mean of the value distribution from this distribution assuming V_i is degenerate given U, and the reverse holds for the upper bound. Thus, these bounds will only be reached when the value distribution is degenerate.

2.2 Bounds on the Distribution of the Highest Value

When only the winning bid is observed, a similar result holds for the mean of the greatest order statistic of the joint distribution of values. The mean of the distribution of the greatest value gives the expected surplus under of the first price auction in a symmetric setting considered here, or any other mechanism that leads to efficient allocation, and can be used to examine how much surplus is lost under alternative mechanisms (as in Section 2.4), or in weighing the welfare of auction participants against potential externalities of the good being auctioned (for example, in a policy analysis examining whether certain forests should be auctioned for timber or preserved for environmental reasons).

As above, let $\tilde{G}(b) = G_{B_{(n)}}(b)^{\frac{1}{n}}$ be the estimate of the bid distribution obtained from the distribution of the highest bid assuming independence and ignoring unobserved heterogeneity. The mean of the greatest order statistic of the value distribution assuming U is degenerate can be estimated as $E(B_{(n)}) + E\left(\frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right)$. When there is unobserved heterogeneity, this provides an upper bound.

Theorem 2. *In the symmetric IPV model with unobserved heterogeneity,*

$$E(B_{(n)}) \leq E(V_{(n)}) \leq E(B_{(n)}) + E\left(\frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right).$$

If the researcher observes only $B_{(n)}$ and $b \mapsto b + \frac{\tilde{G}(b)}{(n-1)\tilde{g}(b)}$ is strictly increasing, these bounds are sharp.

Proof. The pdf of the first order statistic of the bid distribution given U is

$$g_{B_{(n)}}(t|u) = \frac{d}{dt} G_{B_1}(t|u)^n = n G_{B_1}(t|u)^{n-1} g_{B_1}(t|u).$$

Combining this with the bid inversion formula for the affiliated values model gives

$$\begin{aligned}
E(V_{(n)}) &= E(B_{(n)}) + E\left(\frac{G_{B_1}(B_{(n)}|U)}{(n-1)g_{B_1}(B_{(n)}|U)}\right) \\
&= E(B_{(n)}) + \iint_{\underline{b}(u)}^{\bar{b}(u)} \frac{G_{B_1}(b|u)}{(n-1)g_{B_1}(b|u)} n G_{B_1}(b|u)^{n-1} g_{B_1}(b|u)\, db\, dP_u(u) \\
&= E(B_{(n)}) + \iint_{\underline{b}(u)}^{\bar{b}(u)} \frac{n}{n-1} G_{B_1}(b|u)^n\, db\, dP_u(u) \le E(B_{(n)}) + \int_{\underline{b}}^{\bar{b}} \frac{n}{n-1} \int G_{B_1}(b|u)^n\, dP_u(u)\, db \\
&= E(B_{(n)}) + \int_{\underline{b}}^{\bar{b}} \frac{n}{n-1} \int G_{B_{(n)}}(b|u)\, dP_u(u)\, db = E(B_{(n)}) + \frac{n}{n-1} \int_{\underline{b}}^{\bar{b}} G_{B_{(n)}}(b)\, db
\end{aligned}$$

Noting that $\tilde{g}(t) = \frac{d}{dt} G_{B_{(n)}}(t)^{\frac{1}{n}} = \frac{1}{n} G_{B_{(n)}}(t)^{\frac{1-n}{n}} g_{B_{(n)}}(t)$ gives the estimated "bid shade" term under a misspecified assumption of no unobserved heterogeneity as

$$E\left(\frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right) = \int_{\underline{b}}^{\bar{b}} \frac{n G_{B_{(n)}}(b)^{\frac{1}{n}}}{(n-1) G_{B_{(n)}}(b)^{\frac{1-n}{n}} g_{B_{(n)}}(b)} g_{B_{(n)}}(b)\, db = \frac{n}{n-1} \int_{\underline{b}}^{\bar{b}} G_{B_{(n)}}(b)\, db.$$

Plugging this into the previous display gives the result. Sharpness follows since the lower bound is achieved when V_i is degenerate given U, and the upper bound is achieved when there is no unobserved heterogeneity. □

Using the expression for the "bid shade" term from the above proof, the expectation of the highest value estimated under the assumption of independence and no unobserved heterogeneity can be written as

$$E(B_{(n)}) + \frac{n}{n-1} \int_{\underline{b}}^{\bar{b}} G_{B_{(n)}}(b)\, db. = E(B_{(n)}) + \frac{n}{n-1}\left(\bar{b} - E(B_{(n)})\right) = \frac{n}{n-1}\bar{b} - \frac{1}{n-1} E(B_{(n)}).$$

This gives $\left[E(B_{(n)}), \frac{n}{n-1}\bar{b} - \frac{1}{n-1}E(B_{(n)})\right]$ as bounds for the expectation of the valuation of the winning bidder. As with the bounds in Theorem 1, the upper bound will be reached when there is no unobserved heterogeneity (or nontrivial unobserved heterogeneity that does not shift the upper support of the bid distribution), and the lower bound will be reached when the value distribution is degenerate conditional on U, so no smaller interval can be obtained.

If bounds on the mean of the value of the highest bidder are desired and data is only available on one bidder, the bounds in equation 1 can be used to bound the distribution of the highest bid. Then, the upper bound in Theorem 2 can be calculated integrating over the upper bound for the distribution of the highest bid (that is, the lower bound for the cdf).

2.3 Entry Costs and Expected Profits

The bounds on the mean of the distribution of the value of the highest bidder lead to bounds on the expected profits for a given bidder entering the auction. By symmetry, each bidder has probability $1/n$ of drawing the highest value and winning the auction. The ex post profit from winning is equal to $V_{(n)} - B_{(n)}$. Thus, the expected profit $\Pi(n, U)$ from entering an auction with n total bidders for a given value of U is $\Pi(n, U) = \frac{1}{n}E(V_{(n)} - B_{(n)}|U)$. The bound on $E(V_{(n)})$ using data on the highest bid derived above leads immediately to a bound on profits averaged over auctions with different realizations of the unobserved heterogeneity U. This is stated in the following theorem.

Theorem 3. *In the symmetric IPV model with unobserved heterogeneity,*

$$0 \leq E\left(\Pi(n, U)\right) = \frac{1}{n}E(V_{(n)} - B_{(n)}) \leq \frac{1}{n}E\left(\frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right).$$

If the researcher observes only $B_{(n)}$ and $b \mapsto b + \frac{\tilde{G}(b)}{(n-1)\tilde{g}(b)}$ is strictly increasing, these bounds are sharp.

Using the same integration identity as before, this bound can be written as $\frac{1}{n}E\left(\frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right) = \frac{1}{n-1}(\bar{b} - E(B_{(n)}))$, which gives $[0, \frac{1}{n-1}(\bar{b} - E(B_{(n)}))]$ as bounds for expected bidder profit. As with the other bounds, the upper bound is achieved under no unobserved heterogeneity (or nontrivial unobserved heterogeneity that does not shift the upper support of the bid distribution), and the lower bound is achieved when values are a deterministic function of U, subject to the caveat, stated in the theorem, that the estimated bid function must be nondecreasing. As before, the upper bound will be obtained even with nontrivial unobserved heterogeneity as long as it does not shift the support of the bid distribution.

$\Pi(n, U)$ can be interpreted as the expected value of entering an auction where the unobserved heterogeneity is given by U. $E(\Pi(n, U))$ can be interpreted as the long run value of participating in a large number of auctions similar to the ones in the sample. In most entry models, these can be related to the cost of preparing a bid using zero profit conditions, although the specifics of this will depend on the information structure of the model and the equilibrium being played (pure or mixed strategy, etc.). As an example, consider a pure strategy equilibrium in a model in which n_p potential entrants first decide whether to pay a cost K to bid in an auction. In the second stage, each potential entrant i that decided to bid observes the number of bidders n (which is nonstochastic anyway because of the pure strategy assumption), the unobserved (to the econometrician) heterogeneity U, and her value V_i. In a pure strategy equilibrium, the number of entrants n will be the largest value such that the expected second stage profit $E(\Pi(n, U))$ is greater than the entry cost K. Thus, the upper bound for $E(\Pi(n, U))$ derived above will be an upper bound for K. The upper bound on $E(\Pi(n, U))$ and on K can be used to analyze the effect of counterfactual entry fees and subsidies. For example, in this setting, assuming K is nonnegative, the upper bound on $E(\Pi(n, U))$ gives an upper bound on the entry fee that could be charged without decreasing the number of bidders.

2.4 Efficiency Loss from Random Assignment

Efficiency comparisons of different auction formats are of interest in many applications. Often, a researcher has access to data from a series of auctions (in this case, first price sealed bid auctions), and would like to know how changing the selling mechanism would affect things like revenue and allocative efficiency. In this section, I derive an upper bound on the inefficiency from replacing a first price sealed bid auction with a particular alternative mechanism - random assignment - under unobserved heterogeneity.

Lotteries are a common mechanism for governments assigning contracts and property rights. If potential buyers have similar values for the good being assigned, or transaction costs for reselling the good are small, the loss in efficiency from random assignment will be small. In fact, economists are often surprised to find goods that do not seem to fit this description being assigned randomly as well, even when the same good is auctioned in other settings. Although highway construction contracts are often assigned by first price sealed bid auction in the U.S., Decarolis (2010) lists numerous examples of highway construction contracts being assigned though mechanisms that lead to random assignment in equilibrium (see Decarolis, 2010, for conditions under which such equilibria exist and additional conditions under which they are unique). As Decarolis points out, civil engineers and others in policy debates have proposed replacing many of these highway procurement auctions with mechanisms that lead to random assignment in equilibrium play.

It should also be noted that the random assignment considered in this section need not come from a lottery with a uniform distribution over all bidders. Any arbitrary assignment that leads to the long run average value of the object to the winner being $E(V_i)$ will have the same total surplus, even if the allocation is different. For example, this will be the case if firms collude by choosing the winner in some arbitrary way that does not depend on realizations of valuations.

Even in settings where random assignment is not a likely policy outcome, a lottery is a useful benchmark for other efficiency comparisons. Suppose, for example, that we find that subsidizing certain bidders would have a small effect on allocative efficiency relative to the average selling price. One explanation for this is that the subsidy program does a particularly good job of favoring these bidders only when they are close to having the highest value anyway. Another explanation is that values do not vary much between bidders, so that most mechanisms will have similar efficiency properties. Finding that random assignment has similar efficiency properties to both mechanisms favors the latter explanation.

The total surplus in a symmetric first price sealed bid auction (or any mechanism that assigns the good to the agent with the highest valuation) is $V_{(n)}$. Any mechanism that leads to random assignment has the same expected total surplus as a mechanism that fixes some arbitrary bidder i and always awards the object to that bidder, giving ex post surplus of V_i. Thus, the expected efficiency loss from replacing auctions described by U with lotteries is $E(V_{(n)} - V_i|U)$. The expected efficiency loss from replacing all auctions with random assignment is $E(V_{(n)} - V_i)$.

The following theorem shows that estimating the efficiency loss from replacing all auctions with lotteries using data on the highest bid and ignoring unobserved heterogeneity gives an upper bound on $E(V_{(n)} - V_i)$. With data on all bids, a tighter upper bound can be obtained by using the marginal bid distribution to estimate $E(V_i)$.

Theorem 4. *In the symmetric IPV model with unobserved heterogeneity,*

$$\begin{aligned} 0 \leq E(V_{(n)} - V_i) &\leq E\left[\left(B_{(n)} + \frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right) - \left(B_i + \frac{G_{B_1}(B_i)}{(n-1)g_{B_1}(B_i)}\right)\right] \\ &\leq E\left(B_{(n)} + \frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right) - \int \left(b + \frac{\tilde{G}(b)}{(n-1)\tilde{g}(b)}\right) d\tilde{G}(b) \end{aligned}$$

If the researcher observes only $B_{(n)}$ and $b \mapsto b + \frac{\tilde{G}(b)}{(n-1)\tilde{g}(b)}$ is strictly increasing, the bound on *the last line is sharp.*

Proof. Using the integration by parts identities mentioned above, we have

$$E(V_{(n)}|U) - E(V_i|U) = \left(\frac{n}{n-1}\bar{b}(U) - \frac{1}{n-1}E(B_{(n)}|U)\right) - \left(\frac{1}{n-1}\bar{b}(U) + \frac{n-2}{n-1}E(B_i|U)\right)$$
$$= \frac{n-1}{n-1}\bar{b}(U) - \frac{1}{n-1}E(B_{(n)}|U) - \frac{n-2}{n-1}E(B_i|U). \qquad (3)$$

Using the law of iterated expectations and $\bar{b} \geq E\bar{b}(U)$ gives

$$E(V_{(n)}) - E(V_i) \leq \bar{b} - \frac{1}{n-1}E(B_{(n)}) - \frac{n-2}{n-1}E(B_i).$$

Again using integration by parts, the right hand side of the above display is equal to $E\left[\left(B_{(n)} + \frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right) - \left(B_i + \frac{G_{B_1}(B_i)}{(n-1)g_{B_1}(B_i)}\right)\right]$, giving the first inequality.

For the second inequality, note that, by equation 2, $G_{B_1}(b)$ first order stochastically dominates $\tilde{G}(b)$, so that $\int b d\tilde{G}(b) \leq \int b dG_{B_1}(b)$ and

$$\bar{b} - \frac{1}{n-1}E(B_{(n)}) - \frac{n-2}{n-1}E(B_i) \leq \bar{b} - \frac{1}{n-1}E(B_{(n)}) - \frac{n-2}{n-1}\int b\, d\tilde{G}(b).$$

Again using integration by parts and cancellation of densities in the denominator, the right hand side of the above display is equal to $E\left(B_{(n)} + \frac{\tilde{G}(B_{(n)})}{(n-1)\tilde{g}(B_{(n)})}\right) - \int\left(b + \frac{\tilde{G}(b)}{(n-1)\tilde{g}(b)}\right) d\tilde{G}(b)$, giving the second inequality.

Sharpness follows because the bound on the last line is attained under independent bids. □

As mentioned in the proof, these bounds can be written using integration by parts identities as

$$E(V_{(n)} - V_i) \leq \bar{b} - \frac{1}{n-1}E(B_{(n)}) - \frac{n-2}{n-1}E(B_i) \leq \bar{b} - \frac{1}{n-1}E(B_{(n)}) - \frac{n-2}{n-1}\int b\, d\tilde{G}(b). \qquad (4)$$

The last bound is sharp if only data on the winning bid is available, since it is attained if there is no unobserved heterogeneity (or nontrivial unobserved heterogeneity that does not shift the upper support of the bid distribution). All of the quantities in the final expression of this display can be estimated using data on the winning bid. For future reference, let $\mu \equiv \int b\, d\tilde{G}(b)$ and let $\tau \equiv \bar{b} - \frac{1}{n-1}E(B_{(n)}) - \frac{n-2}{n-1}\mu$ be the upper bound obtained using data on the winning bid. While methods for inference on $\bar{b}$ and $E(B_{(n)})$ are well established, estimation of $\int b\, d\tilde{G}(b)$ appears to be an unsolved problem. As it turns out, estimation of this quantity using sample analogues can lead to nonstandard asymptotics and rates of convergence that are slower than the square root of the number of auctions observed, depending on the data generating process and number of bidders. I turn to these issues in the next section.

3 Estimation

The bounds in the previous section are functions of $\bar{b}$, $E(B_{(n)})$ and $\int b\, d\tilde{G}(b)$. Given a sample $B_{(n),1}, \dots, B_{(n),T}$, of winning bids from T auctions, $\bar{b}$ and $E(B_{(n)})$ can be estimated using $\max_{1\le t\le T} B_{(n),t}$ and $1/T \sum_{t=1}^{T} B_{(n),t}$ respectively, and the asymptotic distributions of these estimators are well known. However, estimation of $\int b\, d\tilde{G}(b)$ has not been treated. In this section, I propose estimators for the bounds in the previous sections, and show that the sample analogue estimator $\int b\, d\tilde{G}(b)$ can have a nonstandard rate of convergence and asymptotic distribution depending on n and the shape of the distribution of the highest bid.

Let $\hat{G}_{B_{(n)}}(b) = 1/T \sum_{t=1}^{T} I(B_{(n),t} \le b)$ be the empirical distribution from the sample of winning bids. An obvious candidate for an estimator for $\mu \equiv \int b\, d\tilde{G}(b) = \int b\, dG_{B_{(n)}}^{1/n}(b)$ is the sample analogue $\hat{\mu}_T \equiv \int b\, d\hat{G}_{B_{(n)}}^{1/n}(b)$. This estimator can be written as an L-statistic (see Chapter 22 of van der Vaart, 2000). Letting $B_{(n),(t)}$ be the tth smallest of the sample of winning bids,

$$\int b\, d\hat{G}_{B_{(n)}}^{1/n}(b) = \sum_{t=1}^{T} B_{(n),(t)} \left[\left(\frac{t}{T}\right)^{1/n} - \left(\frac{t-1}{T}\right)^{1/n} \right]. \tag{5}$$

The rate of convergence and asymptotic distribution of $\hat{\mu}_T$ depend on the behavior of $G_{B_{(n)}}$ near the lower end of its support $\underline{b}$. I make the following assumption.

Assumption 1. *For some γ, $\underline{b}$ and $h_0 > 0$, $h(b) \equiv G_{B_{(n)}}(b)/(b - \underline{b})^{\gamma}$ is differentiable with a bounded derivative and* $\lim_{b\downarrow \underline{b}} h(b) = h_0$.

To allow for greater generality, this assumption places conditions directly on the marginal distribution of the winning bid. Section A.4 in the appendix gives sufficient conditions on the distributions of values and unobserved heterogeneity for this assumption to hold. If there is no unobserved heterogeneity and the value distribution F_V behaves like $(b - \underline{b})^{\beta}$ near $\underline{b}$ for some β (for example, $\beta = 1$ if V_1 is uniform), Assumption 1 will hold with $\gamma = n\beta$. However, Assumption 1 allows for the more general case where $B_{(n)}$ is not the greatest order statistic from n independent draws of some distribution.

To see why $\hat{\mu}_T$ can converge at a slower than $\sqrt{T}$ rate, note that the first term in the sum in (5) is $B_{(n),(1)}T^{-1/n}$. By standard results for extreme order statistics, $B_{(n),(1)}$ will converge to $\underline{b}$ at a $T^{1/\gamma}$ rate. Thus, the first term will converge at a $1/n + 1/\gamma$ rate. If $1/n + 1/\gamma > 1/2$, this will be faster than the $\sqrt{T}$ rate of convergence of the intermediate order statistics, so that extreme order statistics will not affect the asymptotic distribution. However, if $1/n + 1/\gamma < 1/2$, the extreme order statistics will converge more slowly than the intermediate order statistics, the rate of convergence will be slower than $\sqrt{T}$ and there will not be a normal limiting distribution. In the intermediate case where $1/n+1/\gamma = 1/2$, it turns out that the rate of convergence is $\sqrt{T}/\sqrt{\log T}$ and the limiting distribution is normal. Intuitively, the issue is that only observing the winning bid does not give enough information about the lower parts of the distribution. In order to estimate the mean of the bid distribution, the smaller draws of the winning bid have to be weighted heavily, which can lead to slower than $\sqrt{T}$ rates of convergence.

In the next theorem, I describe the asymptotic behavior of of $\hat{\mu}_T$ under the various types of tail behavior covered by Assumption 1. The proof, given in the appendix, uses results for L-statistics in Mason and Shorack (1992). Importantly, that paper allows for the types of increasing weights at the tails that lead to nonstandard rates of convergence, as is the case here.

Theorem 5. *Under Assumption 1, we have the following.*

If $1/n + 1/\gamma > 1/2$, *then*

$$\sqrt{T}(\hat{\mu}_T - \mu) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[K(G_{B_{(n)}}(B_{(n),t})) - EK(G_{B_{(n)}}(B_{(n),t}))] + o_P(1) \xrightarrow{d} N(0, \sigma^2)$$

where K *and* σ^2 *are defined below in the appendix. If* $1/n + 1/\gamma = 1/2$, *then*

$$\frac{\sqrt{T}}{\sqrt{\log T}}(\hat{\mu}_T - \mu) \xrightarrow{d} N(0, [n\gamma h_0^{1/\gamma}(1 - 1/n - 1/\gamma)]^{-2}).$$

If $1/n + 1/\gamma < 1/2$, *then* $\sqrt{T}(\hat{\mu}_T - B_T)/A_T$ *will not converge to a nondegenerate normal distribution for any sequences* B_T *and* A_T. *In this case,*

$$T^{1/n+1/\gamma}(\hat{\mu}_T - \mu) = \mathcal{O}_P(1).$$

A useful benchmark case is when the pdf g_{B_1} of B_i converges to a strictly positive constant at the lower support point $\underline{b}$, and there is no unobserved heterogeneity. In this case, $G_{B_1}(b)$ behaves like $b - \underline{b}$ near $\underline{b}$, so $G_{B_{(n)}}(b) = G_{B_1}(b)^n$ behaves like $(b - \underline{b})^n$ near $\underline{b}$, and Assumption 1 holds with $\gamma = n$. Thus, applied to this case, Theorem 5 says that the estimated mean will converge to a normal distribution at a $\sqrt{T}$ rate only if $1/n + 1/n > 1/2$ iff. $n < 4$. If $n = 4$, the estimate will converge at a $\sqrt{T}/\sqrt{\log T}$ rate to a normal distribution. If $n > 4$, the estimation error will be of order $T^{-2/n}$, and the limit will not be a nondegenerate normal distribution.

Theorem 5 applies to estimating the mean of any distribution with repeated observations from the greatest order statistic, regardless of whether the order statistics are bids in a set of first price auctions. Similar methods will also apply to estimating the mean or other functionals of a distribution from repeated observations of another order statistic, such as the second largest value in an ascending auction, but I leave the details of this for future research, since observations of the winning bid in the first price auctions treated in this paper correspond to the largest order statistic.

The asymptotic distributions derived in the first two cases can be used for inference once the variance is estimated. One possibility that applies immediately and only uses the fact that these asymptotic distributions are atomless is subsampling. In the second case, where $1/n + 1/\gamma = 1/2$, inference could also be based on estimating h_0 and, if γ is treated as unknown, testing for $1/n + 1/\gamma = 1/2$ as well. In a recent paper, Hill and Shneyerov (2010) treat inference on tail behavior of bid distributions for a different purpose, and some of the results used in that paper would be useful for this as well. Another possibility would be to use the results for bootstrapping L-statistics from Shorack (1997), or to use a sample analogue estimator of the asymptotic variance. Some of these methods require knowledge of γ in order to determine which case in Theorem 5 applies. This could be done using the methods for inference on the tail behavior of distributions used by Hill and Shneyerov (2010) and the papers they cite, or by directly estimating the rate of convergence as described in Chapter 8 of Politis, Romano, and Wolf (1999) (although the latter may run into problems in the intermediate case where $1/n + 1/\gamma = 1/2$, since the methods described in that book for inference under unknown rates of convergence require that the rate be given by a power of the sample size). I perform a small monte carlo examining other methods of estimating the asymptotic distribution in Section A.3 of the appendix, but leave the rest of these ideas for future research. I use subsampling with an assumed $\sqrt{T}$ rate of convergence in the application in the next section. The $\sqrt{T}$ rate of convergence corresponds to the case where

$1/n + 1/\gamma > 1/2$, which seems plausible in the application since I concentrate on a set of auctions with $n = 3$.

In the last case, where $1/n + 1/\gamma < 1/2$, the results in Mason and Shorack (1992) only give the rate of convergence and state that any limiting distribution must be in a certain class of sums of functionals of Poisson processes and normal random variables. In this case, Theorem 5 can be used for conservative inference. Another possibility would be to truncate an increasing number of the smallest winning bids and use results from Mason and Shorack (1990) for truncated L-statistics. The resulting asymptotic distribution will be the sum of an asymptotically normal variance term and a bias term that converges at a slower rate. To do exact, rather than conservative, inference using this strategy, the bias term would have to be estimated using smoothness assumptions on the tail of the bid distribution. This would have the additional advantage of improving the rate of convergence by extrapolating tail behavior.

While estimating or bounding the mean of the marginal bid distribution may be interesting in its own right in some applications, one of the primary motivations for estimating it in this paper is in estimating efficiency losses from random allocation using the formula in (4). For the case where $1/n + 1/\gamma > 1/2$, Theorem 5 gives an influence function representation that can be used to derive the asymptotic distributions of statistics that are functions of $\hat{\mu}_T$ and other estimators (for the case where $1/n + 1/\gamma = 1/2$, the results of Mason and Shorack (1992) also give an influence function representation, but I leave this out because the slower rate of convergence means that functions of this and other statistics will typically have an asymptotic distribution that depends only on the asymptotic distribution of $\hat{\mu}_T$ in this case). The next theorem describes the asymptotic behavior of the sample analogue estimator of the bound $\tau \equiv \overline{b} - \frac{1}{n-1}E(B_{(n)}) - \mu$ in (4). It places an additional assumption on the behavior of the distribution of the winning bid near its upper support point to deal with the estimator of the upper support point $\overline{b}$.

Theorem 6. *Suppose that Assumption 1 holds and that* $(1 - G_{B_{(n)}}(b))/(\overline{b} - b)^2 \to \infty$ *as* $b \to \overline{b}$. *Define* $\hat{\tau}_T = \max_{0 \le t \le T} B_{(n),(t)} - \frac{1}{n-1}\frac{1}{T}\sum_{t=1}^{T} B_{(n),(t)} - \frac{n-2}{n-1}\hat{\mu}_T$ *to be the sample analogue estimator of* $\tau = \overline{b} - \frac{1}{n-1}E(B_{(n)}) - \frac{n-2}{n-1}\int b\, d\tilde{G}(b)$.

If $1/n + 1/\gamma > 1/2$, *then*

$$\sqrt{T}(\hat{\tau}_T - \tau) \overset{d}{\to} N(0, var(\frac{1}{n-1}B_{(n)} + \frac{n-2}{n-1}K(G_{B_{(n)}}(B_{(n),t}))))$$

where K is defined in the appendix. If $1/n + 1/\gamma = 1/2$, then

$$\frac{\sqrt{T}}{\sqrt{\log T}}(\hat{\tau}_T - \tau) \xrightarrow{d} N(0, [n\gamma h_0^{1/\gamma}(1 - 1/n - 1/\gamma)]^{-2}).$$

If $1/n + 1/\gamma < 1/2$, then $\sqrt{T}(\hat{\tau}_T - B_T)/A_T$ will not converge to a nondegenerate normal distribution for any sequences B_T and A_T. In this case,

$$T^{1/n+1/\gamma}(\hat{\tau}_T - \tau) = \mathcal{O}_P(1).$$

The theorem follows immediately from Theorem 5 and the fact that the estimator of the upper support point $\bar{b}$ converges at a faster than $\sqrt{T}$ rate under these conditions. If the upper tail of the distribution of the winning bid is such that $\max_{1\leq t\leq T} B_{(n),t}$ converges at a slower rate than $\hat{\mu}_T$, the extreme value limit of $\max_{1\leq t\leq T} B_{(n),t}$ will dominate.

4 Application

To illustrate the estimation of these bounds, I provide an application to bounding the loss in surplus from replacing a subset of highway procurement auctions in Michigan with random allocation. Random allocation is a useful benchmark for comparing outcomes under other counterfactual mechanisms in many applications. If the loss in efficiency from random allocation is small, we would be more surprised to find large efficiency effects for, say, a bid subsidy scheme than if it is large. Highway procurement auctions provide a particularly interesting application because mechanisms that lead to random allocation are common in highway procurement, as are proposals to replace first price auctions with these mechanisms. See Decarolis (2010) for a treatment of some of these mechanisms, and examples of where they have been implemented. Since these are procurement auctions in which the lowest bid wins, the results stated above will hold with $\underline{b}$ replaced by $\bar{b}$, $B_{(n)}$ replaced by $B_{(1)}$, etc.

The formula in (4) (or the version of it that reverses signs for bids from a procurement auction) is equal to or bounds the expected difference between the lowest cost and a randomly chosen bidder under the conditions given above. However, there are some caveats to interpreting this as the surplus loss from random or arbitrary allocation. If contracts were allocated randomly and subcontracting were allowed, then, depending on negotiation costs, subcontracting could lead to a more efficient allocation so that the bounds in this paper would overstate efficiency losses even more. Another reason that random allocation

	Mean	Std. Dev.	Min	Max
Eng. Est.	906.4	1,318.7	30.6	13,392.2
$\frac{B_{(1)}}{\text{Eng. Est.}}$	0.9351	0.1357	0.6131	1.4713
$\frac{B_{(2)}-B_{(1)}}{\text{Eng. Est.}}$	0.0795	0.0757	0.0007	0.5440

Table 1: Summary Statistics for 3 Bidder Sample (215 Auctions)

might be more efficient relative to a first price auction than the estimates here suggest is if the costs to a firm of taking a contract do not represent the social costs of the contract. Decarolis (2010) argues that this is the case in the Italian highway procurement auctions he studies because firms tend to have lower costs when they are more willing to default on the contract if it turns out to be more expensive than expected. On the other hand, the estimates of the surplus loss from random allocation used here might understate these losses if such a mechanism leads less efficient firms to participate, since these estimates hold the set of participants fixed.

One of the advantages of the bound given in (4) is that it can be estimated using only repeated observations of the winning bid. In addition to requiring additional assumptions on the form of the unobserved heterogeneity, methods for taking into account unobserved heterogeneity such as those proposed by Krasnokutskaya (2009) require repeated observations from at least two bidders. These methods require that these observations be from the same two bidders or, in the symmetric case, two random bidders, rather than, say, the top two bids. Since data sets with repeated observations from the same strict subset of bidders are rare, this usually means that all bids must be observed. To compare the bounds in this paper to other methods that require additional assumptions and data, I use a data set with observations of all bids and calculate the Krasnokutskaya (2009) estimates for comparison.

The data set is the same as the one used by Einav and Esponda (2008) and Somaini (2011) and contains observations on procurement auctions for highway construction projects in Michigan run by the Michigan Department of Transportation. Contracts are awarded through a first price sealed bid auction in which the lowest bidder is awarded the project, subject to some rules governing eligibility. I focus on a subset of auctions between 2001 and 2004 that mainly involve maintenance. The data set contains all bids for each auction along with an engineer's estimate of the cost of the project. I focus on the subset of these auctions with exactly three bidders. Table 1 shows summary statistics for these auctions. Engineer's estimates are in thousands of dollars.

A word is in order about the extent to which the assumptions used in deriving the bounds in this paper are likely to hold for this data set. The bounds require that values be symmetric and independent conditional on the unobserved heterogeneity, U, observed to all agents. While these assumptions cannot be tested with data on only the winning bid, with data on multiple bids or additional data on observed variables that shift values or costs, they have some testable implications. In this data set, bids are positively correlated within an auction, which is consistent with this model, but also with models of correlation not caused by unobserved heterogeneity. As for the symmetry assumption, the recent work of Somaini (2011) provides evidence that bids are negatively correlated with distance from a firm's location in this data set, suggesting that there is some asymmetry in the data (although the effect of distance does not appear to be too large relative to overall variation in bids). Under an exclusion restriction involving the distance of other firms, Somaini (2011) also finds evidence of common values ruled out by the assumptions of this paper. The results in this section are intended to be illustrative and, to the extent that these assumptions do not hold for this data set, they should be interpreted with appropriate caution.

In calculating the bounds in this paper and the Krasnokutskaya (2009) estimates, I incorporate observed heterogeneity in cost distributions by modeling cost distributions as fractions of the engineer's estimate. Specifically, letting C_i be bidder i's cost and A the engineer's estimate, I assume that $C_i = A \cdot \tilde{C}_i$ where $(\tilde{C}_i, U)$ (U being the unobserved heterogeneity) is independent of A. Under these assumptions, bidder i's bid B_i will satisfy $B_i = A \cdot \tilde{B}_i$ where $\tilde{B}_i$ is bidder i's bid in an auction where $A = 1$ and bidder i has cost $\tilde{C}_i$. Thus, we can divide all bids by the engineer's estimate and the bounds will still hold with the average efficiency loss from holding a lottery instead of a first price auction reinterpreted as the average surplus loss from random allocation as a proportion of the engineer's estimate. I also control for observed heterogeneity by focusing on auctions with a fixed number of bidders. The assumptions describing how cost distributions relate to engineer's estimates could be done away with by conditioning on the engineer's estimate directly and using auctions with similar engineer's estimates, but, because of the relatively small size of the data set after conditioning on the number of bidders as well, I take the approach described above. See Sections 5.1 and 5.2 for discussions of other ways of incorporating observed heterogeneity, and Section A.2 in the appendix for alternate estimates that incorporate observed heterogeneity in engineer's estimates in a more nonparametric way.

Under the assumption that distribution of the highest bid behaves like the distribution of the least order statistic from a draw of three random variables with density bounded away from zero near the upper support point, Theorem 5 states that the estimator will converge at a $\sqrt{T}$ rate ($n = \gamma = 3$, so $1/n + 1/\gamma > 1/2$). I assume that γ in Assumption 1 is such that $1/n + 1/\gamma > 1/2$ so that convergence is at a $\sqrt{T}$ rate, and perform inference by subsampling at this rate. I use $\lceil T^{3/4} \rceil = \lceil 215^{3/4} \rceil = 57$ as the subsample size, and fix the lower endpoint estimate of $\underline{b}$ in the subsamples, since it converges more quickly than the other components of $\hat{\tau}_T$ under these conditions. I use $1,000$ subsample replications. I report confidence regions based on other methods of estimating the asymptotic distribution in Section A.3 of the appendix. In cases where the researcher is uncomfortable making assumptions on the tail of the bid distribution, the methods used here could be extended to use an estimate of γ or estimate the rate of convergence directly, as discussed in Section 3. The upper endpoint of the reported 95% confidence region inverts a set of one sided level .05 subsampling based tests on the estimated upper bound of the identified set. Since the lower endpoint of the identified set is known, any confidence region that takes the form of an interval with zero as the lower endpoint will contain the entire identified set with probability .95 iff. it contains any fixed point in the identified set with probability .95. That is, letting τ be the true upper bound for the efficiency loss from random allocation and $\hat{c}_n$ the upper endpoint of a confidence region, $P([0, \tau] \subseteq [0, \hat{c}_n]) \geq .95$ iff. $P(\tau' \in [0, \hat{c}_n]) \geq .95$ for all $\tau' \in [0, \tau]$ (see Imbens and Manski, 2004, for a discussion of the distinction between these two criteria for confidence regions for set identified parameters). The subsampling confidence interval reported here will satisfy these criteria asymptotically under the appropriate assumptions.

The cost distribution estimators proposed by Krasnokutskaya (2009) require the additional assumption that the unobserved heterogeneity takes a multiplicatively separable form. In particular, we require that, for $\tilde{C}_i$ defined above, $\tilde{C}_i = U \cdot K_i$ where K_i is independent of the unobserved heterogeneity U. As Krasnokutskaya describes, this allows for the application of deconvolution methods from the measurement error literature. I report estimates of the efficiency loss from random allocation that use these estimates of the bid distribution along with the bounds described above. I estimate the surplus loss from random allocation by using the formula

$$E(C_i - C_{(1)}|U) = \frac{1}{n-1} E(B_{(1)}|U) + \frac{n-2}{n-1} E(B_i|U) - \underline{b}(U)$$

(this is just the procurement version of (3)) and the law of iterated expectations to get

$$E(C_i - C_{(1)}) = \frac{1}{n-1} E(B_{(1)}) + \frac{n-2}{n-1} E(B_i) - E(\underline{b}(U)).$$

The estimates use this formula with $E(B_{(1)})$ and $E(B_i)$ replaced by their sample analogues,

	$EC_i - EC_{(1)}$
Deconvolution method (Krasnokutskaya, 2009)	0.1285
Ignoring heterogeneity	0.4166
95% CI	$[0, 0.0.4306]$

Table 2: Estimation Results for 3 Bidder Sample (215 Auctions)

and $E(\underline{b}(U))$ replaced by the expectation of $\underline{b}(U)$ when the joint distribution of U and B is given by the deconvolution estimates described by Krasnokutskaya. In taking the Fourier transform of the estimated characteristic functions, I set the truncation parameter governing the limits of the integral (T in Krasnokutskaya's notation) equal to 15.

Estimation results are in Table 2. The deconvolution method of Krasnokutskaya (2009) gives a point estimate of about 12% of the engineer's estimate as the average surplus loss from replacing these auctions with random allocation. The point estimate for the upper bound using only the winning bid is around 42%, with the upper endpoint of the 95% confidence region of around 43% of the engineer's estimate. The bounds using only data on the winning bid are conservative (assuming the Krasnokutskaya (2009) assumptions hold), but still informative. It should be noted that the deconvolution estimates are point estimates rather than upper endpoints of confidence regions. Since only rates of convergence and not nondegenerate asymptotic distributions are available for these estimators, it is not clear how to perform inference other than adding an arbitrary constant after scaling by something slower than the rate of convergence. Krasnokutskaya and others have used the bootstrap, but it is not clear if it is valid in this setting, especially since the application in this paper uses the formula in (3), which involves support points of the estimated distribution. It is well known that the bootstrap fails for estimating support points even in simpler settings.

The fact that the bounds using winning bid data are somewhat conservative in this application is not surprising giving the large amount of unobserved heterogeneity in this data set. The correlation of log bids in this data set is about .6, so that (with the assumption of independent values conditional on unobserved heterogeneity made throughout this paper) about 60% of the variation in log bids is due to unobserved heterogeneity. With only data on the winning bid, it is impossible to rule out bids and values having a much smaller correlation, which would mean more efficiency loss from random allocation.

While this application focuses on efficiency loss from random allocation, it is interesting to compare the bounds and point estimates for the expected profit from entering an auction, $E(B_{(1)} - C_{(1)})/n$, as well. Interdependence between the bids in this sample also leads to the bounds for the expected profit from entering an auction that use only one bid being somewhat conservative relative to those computed using all bids and the assumptions of Krasnokutskaya (2009). The point estimate for the upper bound for the expected profit from entering an auction as a fraction of the engineer's estimate using the winning bid is 0.1610, with a 95% CI of $[0, 0.1686]$, while the Krasnokutskaya (2009) estimate is 0.0404 (neither of these estimates are reported in these tables).

5 Extensions and Limitations

This section discusses some possible extensions of the results in this paper, as well as some limitations of the approach taken in this paper.

5.1 Incorporating Observed Heterogeneity

In practice, a researcher will often have access to some, but not all, of the variables that are common knowledge to auction participants when they prepare bids. In other words, it is appropriate to model the environment as an IPV auction taking place conditional on random variables (X, U), where X is observed to the researcher and the bidders, and U is observed to bidders only. The observed covariate X may contain engineer's estimates of the value of the object (as in the application in Section 4) or other observable measures of quality.

In this case, all of the population bounds derived in Section 2 will hold, with all probability distributions interpreted as being conditional on X in addition to any other conditioning variables. If X is discrete, the estimation results of Section 3 will also apply immediately to estimates of the bounds conditional on each value that X takes. The methods for estimation and inference introduced in that section can simply be applied separately to the observations corresponding to each distinct value of X.

If X is continuously distributed (or takes on a large number of values relative to the sample size), estimation becomes more difficult. One possibility is to estimate bounds conditional on X by splitting the sample into discrete bins corresponding to different values of X, or by computing the estimates in Section 5 with estimates of the mean and cdf of the winning bid replaced by kernel estimates of these objects conditional on X taking some value x. Such estimates will likely be consistent for the corresponding bounds conditional on X if the bandwidth is taken to zero or the number of bins is taken to infinity, and, if the bandwidth decreases quickly enough, the asymptotic distribution will be normal or nonstandard depending on the tail behavior of the conditional bid distribution. Confirming these conjectures would require extending the results of Mason and Shorack (1992) to the analogous nonparametric kernel estimates, a topic that I leave for future research.

One advantage of an approach based on binning the data according to X is that these estimates still give valid bounds even if the bins do not change with the sample size. This holds because the bounds in Section 2 apply with variation of X within a bin interpreted as part of the unobserved heterogeneity U. If we restrict the sample to auctions t with $\|X_t - x\| \leq h$ for some h that does not change with the number of auctions T, the bounds in Section 2 will hold for quantities conditional on $\|X_t - x\| \leq h$, with (X, U) taking the place of U. Applying the estimators of Section 3 to auctions with $\|X_t - x\| \leq h$ will give consistent estimates of upper bounds for $E(V_{(n)} - V_i | \|X_t - x\| \leq h)$. These estimates can be averaged over nonoverlapping bins to get upper bounds for $E(V_{(n)} - V_i)$ that are tighter than bounds that ignore X completely. Section A.2 reports estimates for the application in Section 4 that incorporate observed heterogeneity in engineer's estimates in this way (the estimates in Section A.2 use a stronger separability and independence assumption to incorporate the engineer's estimate).

5.2 Mapping the Bounds to Parametric Models

Another approach is to model observed heterogeneity parametrically. One way of doing this is to assume some separable form for the observed heterogeneity, such as

$$V_{i,t} = \exp(X_t'\beta) \cdot \tilde{V}_{i,t} \tag{6}$$

where

$$\{\tilde{V}_{i,t}\}_{i=1}^{n} \text{ are independent conditional on } U_i \text{, and } X_t \text{ is independent of } \{\tilde{V}_{i,t}\}_{i=1}^{n}. \tag{7}$$

This is essentially the approach taken in the application in Section 4, but with β known a priori. More generally, β can be estimated up to the constant term from winning bid data using the regression

$$\log B_{(n),t} = X_t'\beta + \log \tilde{B}_{(n),t},$$

which follows from separability of the bid function, and the bounds can be estimated using $\hat{\tilde{B}}_{(n),t} = B_{(n),t} / \exp(X_t'\hat{\beta})$. Note that, despite the general nature of the unobserved heterogeneity in this setup, the separability and independence assumption allows the regression parameter β to be point identified up to the constant term.

More generally, one may be interested in inference on a regression parameter after imposing a separability assumption like (6), but with a weaker mean independence assumption on $\tilde{V}_{i,t}$. After a normalization, this leads to a model of the form

$$E(V_{i,t}|X_t) = \exp(X_t'\beta).$$

Without the full independence assumption (7), the bid function may not be separable in X, so that β will no longer be point identified in general. However, sharp bounds on β can be obtained from data on a single bidder by mapping the bounds of Section 2.1 to this equation, leading to the restrictions

$$E(B_{i,t}|X_t) \leq \exp(X_t'\beta) \leq \frac{n-2}{n-1}E(B_{i,t}|X_t) + \frac{1}{n-1}\bar{b}(X_t) \tag{8}$$

where, by some abuse of notation, $\bar{b}(X_t)$ is the upper support point of $B_{i,t}$ given X_t. The inequality restrictions (8) define an interval regression that is a slight extension of one of the models considered by Manski and Tamer (2002). Inference on β in this setting can be done using any of several recently proposed methods (see, among others, Andrews and Shi, 2009; Armstrong, 2011; Chernozhukov, Lee, and Rosen, 2009; Lee, Song, and Whang, 2011). With data on the winning bid, the results in Section 2.2 could be used to estimate a similar regression model where the conditional mean of the highest value $E(V_{(n),t}|X_t)$ is modeled directly, but this seems less natural.

If the researcher imposes a full parametric model on the distibution of (V, U) given X_i, the bounds in this paper can still be applied, but they will not lead to sharp bounds on the parameter. In this case, the additional information given by the parametric model may even be enough to point identify the parameter.

5.3 Multiple Bid Observations

Most of the bounds derived in this paper are sharp if only the winning bid (or a single bidder) is observed. If more than one bid is observed, the bounds can, of course, still be computed. The bounds may even be tightened by using bounds on other order statistics derived in a similar way to those derived in this paper using the winning bid. However, these bounds will typically not be sharp, since they ignore the correlation structure of multiple bids in each auction (although some of them will still be sharp as long as the joint distribution of observed bids can be rationalized by unobserved heterogeneity that does not shift the upper support point of bids, such as the example in Section A.5 of the appendix). With enough data and some additional assumptions, identification results such as those derived by Krasnokutskaya (2009) or Hu, McAdams and Shum (2011) can be used. This leaves open the questions of (1) how much can be learned with data that satisfies the requirements of one of these papers without imposing the additional assumptions of these papers and (2) how much can be learned with or without these additional assumptions if more than one bid is observed, but the data requirements of these papers are not met. It seems likely that the answers to these questions will require methods substantially different from those developed in this paper and elsewhere in the literature.

It should be noted that, while question (1) is important only to the extent that the conditions of papers such as Krasnokutskaya (2009) and Hu, McAdams and Shum (2011) are questionable in applications, there are many cases where one has data on more than one bidder, but the data requirements of these papers are not met. One common case is data sets where only the top two bids are reported and the number of bidders is known. Even if one imposes the separability assumptions of Krasnokutskaya (2009), the approach in that paper, which requires repeated observations of the same two bidders, cannot be used. On the other hand, using the bounds in the present paper on such a data set throws away information about the correlation of bids given by multiple bid observations. Deriving sharper bounds or identification results for these situations is an important question for future research.

5.4 Limitations

This paper assumes throughout that, conditional on the unobserved heterogeneity, values are drawn independently from the same distribution. In addition, this paper derives bounds on averages of draws from the value distribution, rather than bounding the entire distribution. With only data on the winning bid or a single bidder, it is difficult to imagine getting informative bounds after replacing symmetry with a weaker condition. Without some structure on how value distributions are related, little can be done to recover or bound valuations of one bidder with the bids of another. It also appears that using naive estimates that ignore unobserved heterogeneity to bound the the entire value distribution, rather than the quantities based on averages of order statistics considered in this paper, is not possible under this level of generality. Section A.5 of the Appendix provides a counterexample.

Thus, using these or similarly strong symmetry assumptions and giving up on bounding the entire value distribution appears to be necessary when dealing with data on a single bid and unobserved heterogeneity of such a general form. It should be recognized that this approach limits the questions that can be asked and answered. Ascending auctions and first price auctions lead to the same outcomes and average payoffs under these assumptions, so one cannot use this framework to compare revenue or efficiency of these auction formats empirically. Allowing asymmetry or affiliation conditional on the unobserved heterogeneity would make such revenue comparisons interesting, but this would require more data or additional assumptions. In addition, while sufficient to bound total surplus under random allocation or bidder profits, the bounds in this paper cannot be used to bound optimal reserve prices. This paper shows that much can be learned with a relatively small amount of data (only the winning bid or a single bidder) even with unobserved heterogeneity of a very general form, but certain questions require better data or more stringent assumptions.

6 Conclusion

Unobserved heterogeneity leads to inconsistent estimates of the value distribution in empirical studies of auctions if it is not taken into account. Recently, several authors have proposed methods for dealing with unobserved heterogeneity in auctions, but these require additional data and assumptions. This paper provides bounds for several economic primitives of interest in auction studies using only data on the winning bid or bids from a single bidder. These bounds also give the direction of bias when unobserved heterogeneity is not taken into account. This shows for which economic primitives intuition about the bias from ignoring unobserved heterogeneity holds in general.

This paper also shows how to estimate and perform inference on these bounds. Depending on the number of bidders and the shape of the bid distribution, statistics that depend on the mean of the bid distribution, such as the expected surplus loss from replacing a first price auction with a lottery, can in some cases only be estimated at a slower rate when only data on the winning bid is available than with data on a single bidder or all bids. I provide conditions under which these estimates are still asymptotically normal and converge at a rate proportional to the number of auctions observed, and when the rate of convergence and asymptotic distribution is different. These results are likely to be of independent interest, since they apply more broadly to estimation of statistics of a distribution from repeated observations of a single order statistic.

An application to bounding the long run surplus loss from replacing a subset of highway procurement auctions in Michigan with random or arbitrary allocation illustrates the results. While the bounds are conservative compared to point estimates that use the methods of Krasnokutskaya (2009), they are still informative, while requiring less data and fewer assumptions.

A Appendix

A.1 Proof of Theorem 5

In this section of the appendix, I prove Theorem 5, which treats the estimation of $\int bd\,\tilde{G}(b)$, the mean of the bid distribution using repeated observations of the largest bid $B_{(n)}$. The results in this section do not require that $B_{(n)}$ be the largest order statistic of n independent observations, and apply to estimation of $\int bd\,\left(G_{B_{(n)}}(b)\right)^{1/n}$ with samples from the distribution $G_{B_{(n)}}$ for any cdf $G_{B_{(n)}}$. I state some of these results using an influence function representation of the estimated mean of the bid distribution so that they can be used in Theorem 6 along with the other terms in Equation 4 to estimate the loss in surplus from replacing a first price auction with a lottery. Throughout this section of the appendix, I assume that Assumption 1 holds for some γ.

The statistic of interest can be written in L-statistic form as

$$\mu \equiv \int bd\,\left(G_{B_{(n)}}(b)\right)^{1/n} = \int_0^1 G^{-1}_{B_{(n)}}(u^n)\,du = \int_0^1 G^{-1}_{B_{(n)}}(u)\,d(u^{1/n}) = \int_0^1 G^{-1}_{B_{(n)}}(u)\frac{1}{n}u^{1/n-1}\,du.$$

Similarly, for the sample analogue,

$$\begin{aligned}\hat{\mu}_T &\equiv \int bd\,\left(\hat{G}_{B_{(n)}}(b)\right)^{1/n} = \int_0^1 \hat{G}^{-1}_{B_{(n)}}(u)\frac{1}{n}u^{1/n-1}\,du = \sum_{t=1}^{T} B_{(n),(t)} \int_{(t-1)/T}^{t/T} \frac{1}{n}u^{1/n-1}\,du \\ &= \sum_{t=1}^{T} B_{(n),(t)} \left[\left(\frac{t}{T}\right)^{1/n} - \left(\frac{t-1}{T}\right)^{1/n}\right].\end{aligned}$$

This is an L-statistic with, in the notation of Mason and Shorack (1992), $J(t) = \frac{1}{n}t^{1/n-1}$ and $g(t) = G^{-1}_{B_{(n)}}(t)$. I derive the asymptotic distribution of this statistic by verifying the conditions of that paper.

The following lemma giving the tail behavior of $\frac{d}{dt}G^{-1}_{B_{(n)}}(t)$ will be useful.

Lemma 1. $k(t) \equiv \left[\frac{d}{dt}G^{-1}_{B_{(n)}}(t)\right]/t^{1/\gamma-1}$ *satisfies* $\lim_{t\downarrow 0} k(t) = h_0^{-1/\gamma}/\gamma$.

Proof. By the inverse function theorem,

$$\frac{d}{dt}G_{B_{(n)}}^{-1}(t) = \left[\frac{d}{du}G_{B_{(n)}}(u)\right]^{-1}\Bigg|_{u=G_{B_{(n)}}^{-1}(t)} = \left[\gamma(u-\underline{b})^{\gamma-1}h(u) + (u-\underline{b})^{\gamma}h'(u)\right]^{-1}\Bigg|_{u=G_{B_{(n)}}^{-1}(t)}$$

$$= \left\{\gamma(u-\underline{b})^{\gamma-1}[h(u) + (u-\underline{b})h'(u)/\gamma]\right\}^{-1}\Bigg|_{u=G_{B_{(n)}}^{-1}(t)}.$$

Under Assumption 1, we will have $G_{B_{(n)}}^{-1}(t) = \underline{b} + t^{1/\gamma}r(t)$ where $r(t) \to h_0^{-1/\gamma}$ as $t \to 0$. Plugging this into the above display, we get

$$\left\{\gamma(t^{1/\gamma}r(t))^{\gamma-1}[h(\underline{b}+t^{1/\gamma}r(t)) + t^{1/\gamma}r(t)h'(\underline{b}+t^{1/\gamma}r(t))/\gamma]\right\}^{-1}.$$

Dividing by $t^{1/\gamma-1}$ gives

$$\left\{\gamma(r(t))^{\gamma-1}[h(\underline{b}+t^{1/\gamma}r(t)) + t^{1/\gamma}r(t)h'(\underline{b}+t^{1/\gamma}r(t))/\gamma]\right\}^{-1},$$

which converges to $(\gamma h_0^{-(\gamma-1)/\gamma}\cdot h_o)^{-1} = h_0^{-1/\gamma}/\gamma$. □

In the notation of Mason and Shorack (1992), $K(t)$ is given by

$$K(t) = \int_c^t J(u)\,dg(u) = \int_c^t \frac{1}{n}u^{1/n-1}\,dG_{B_{(n)}}^{-1}(u) = \int_c^t \frac{1}{n}u^{1/n+1/\gamma-2}k(u)\,du.$$

Here, c is any fixed constant greater than zero, and $\int_c^t$ is defined to be $-\int_t^c$ if $t < c$. The next lemma describes the tail behavior of $K(t)$.

Lemma 2. *If* $1/n+1/\gamma < 1$, *then, defining* $q(t) = K(t)/t^{1/n+1/\gamma-1}$, $q(t) \to q_0 \equiv -[n\gamma h_0^{1/\gamma}(1-1/n-1/\gamma)]^{-1}$ *as* $t \to 0$.

Proof. For any $\varepsilon > 0$, $K(t)$ is sandwiched between

$$\int_c^t \frac{1}{n}\left(\frac{h_0^{-1/\gamma}}{\gamma} - \varepsilon\right)u^{1/n+1/\gamma-2}\,du = \frac{1}{n}\left(\frac{h_0^{-1/\gamma}}{\gamma} - \varepsilon\right)(1/n+1/\gamma-1)^{-1}(t^{1/n+1/\gamma-1} - c^{1/n+1/\gamma-1})$$

and the corresponding expression with ε added rather than subtracted for small enough t. Since $t^{1/n+1/\gamma-1}$ increases without bound as $t \to 0$, for small enough t, we can get rid of the power of c at the cost of another ε to get

$$\frac{1}{n}\left(\frac{h_0^{-1/\gamma}}{\gamma}+2\varepsilon\right)(1/n+1/\gamma-1)^{-1}t^{1/n+1/\gamma-1}$$
$$\leq K(t)\leq\frac{1}{n}\left(\frac{h_0^{-1/\gamma}}{\gamma}-2\varepsilon\right)(1/n+1/\gamma-1)^{-1}t^{1/n+1/\gamma-1}.$$

The result follows since ε can be arbitrarily small. □

Again following the notation of Mason and Shorack (1992), let $K_{ab}(t)$ be defined for any $a, b \in (0,1)$ as $K(a)$, $K(t)$, or $K(b)$ for $t \leq a$, $a < t < b$, and $b \leq t$ respectively. Define

$$\sigma^2[a,b)=\int_0^1 K_{ab}^2(t)\,dt-\left(\int_0^1 K_{ab}(t)\,dt\right)^2$$

and

$$\mu(a,b)=\int_a^b J(t)g(t)\,dt=\int_a^b G_{B_{(n)}}^{-1}(t)\frac{1}{n}t^{1/n-1}\,dt=\int_a^b(\underline{b}+t^{1/\gamma}r(t))\frac{1}{n}t^{1/n-1}\,dt.$$

Define $\sigma^2(a)=\sigma^2[a,1-a)$ and $\mu(a)=\mu[a,1-a)$. The limiting behavior of $\hat{\mu}_T$ depends on a variance term and a bias term. For the variance term, we need to study the limiting behavior of $\sigma^2[a,1)$ as $a\to 0$. For the bias term, we need to study the limiting behavior of $\mu(a,1)-\mu(0,a)$ as $a\to 0$. I treat these issues in the following lemmas.

Lemma 3. *If $1/n+1/\gamma > 1/2$, then $\sigma^2[0,1)$ is finite. If $1/n+1/\gamma<1/2$, then $\sigma^2[a,1)/a^{2(1/n+1/\gamma-1/2)}\to q_0^2\left(1+\frac{1}{1-2/n-2/\gamma}\right)$ as $a\to 0$. If $1/n+1/\gamma=1/2$ then $\sigma^2[a,1)/(-\log a)\to q_0^2$.*

Proof. In all cases, $\left(\int_0^1 K_{a1}(t)\,dt\right)^2$ is bounded independently of a, so the claims will hold for $\sigma^2[a,1)$ iff. they hold with $\sigma^2[a,1)$ replaced by the first term $\int_0^1 K_{a1}^2(t)\,dt$. For any $\varepsilon>0$, there is a b such that $|q(t)^2-q_0^2|\leq\varepsilon$ on $t\leq b$, so that $\int_0^1 K_{a1}^2(t)\,dt$ is bounded from above by

$$a\cdot(q_0^2+\varepsilon)a^{2/n+2/\gamma-2}+(q_0^2+\varepsilon)\int_a^b t^{2/n+2/\gamma-2}\,dt+\int_b^1 t^{2/n+2/\gamma-2}[q(t)]^2\,dt \tag{9}$$

and from below be the same expression with ε subtracted rather than added. The last term is finite and does not depend on a. If $1/n + 1/\gamma < 1/2$, the first two terms add up to

$$\begin{aligned}&(q_0^2+\varepsilon)a^{2/n+2/\gamma-1} + (q_0^2+\varepsilon)\frac{1}{2/n+2/\gamma-1}(b^{2/n+2/\gamma-1} - a^{2/n+2/\gamma-1})\\ &= (q_0^2+\varepsilon)a^{2/n+2/\gamma-1}\left(1+\frac{1}{1-2/n-2/\gamma}\right) + (q_0^2+\varepsilon)\frac{1}{2/n+2/\gamma-1}b^{2/n+2/\gamma-1}.\end{aligned}$$

The last term in this display is finite and does not depend on a, so, for small enough a,

$$\int_0^1 K_{a1}^2(t)\,dt \le (q_0^2+2\varepsilon)a^{2/n+2/\gamma-1}\left(1+\frac{1}{1-2/n-2/\gamma}\right).$$

A similar argument gives the corresponding lower bound, so, since ε is arbitrary,

$$\lim_{a\to 0}\int_0^1 K_{a1}^2(t)\,dt/a^{2/n+2/\gamma-1} = q_0^2\left(1+\frac{1}{1-2/n-2/\gamma}\right).$$

If $1/n+1/\gamma = 1/2$, the last term in (9) is still finite and does not depend on a, while the first two terms are

$$(q_0^2+\varepsilon) + (q_0^2+\varepsilon)(\log b - \log a).$$

For small enough a, this is bounded from above by $-(q_0^2+2\varepsilon)\log a$. This and the corresponding lower bound give the final claim. □

Lemma 4. *As $a \to 0$, $(\mu - \mu(a)) = \underline{b}a^{1/n} + \mathcal{O}(a^{1/n+1/\gamma} + a)$.*

Proof. We have

$$\begin{aligned}\mu - \mu(a) &= \int_0^a (\underline{b} + t^{1/\gamma}r(t))\frac{1}{n}t^{1/n-1}\,dt + \int_{1-a}^1 (\underline{b} + t^{1/\gamma}r(t))\frac{1}{n}t^{1/n-1}\,dt\\ &= \underline{b}a^{1/n} + \int_0^a \frac{1}{n}t^{1/n+1/\gamma-1}r(t)\,dt + \int_{1-a}^1 (\underline{b} + t^{1/\gamma}r(t))\frac{1}{n}t^{1/n-1}\,dt.\end{aligned}$$

The second term is $\mathcal{O}(a^{1/n+1/\gamma})$ and the last term is $\mathcal{O}(a)$. □

Define $\mu_T = \mu(1/T)$ and $\sigma_T^2 = \sigma^2(1/T)$. Also define the trimmed L-statistic $\tilde{\mu}_T = \sum_{t=2}^{T-1} B_{(n),(t)}\left[\left(\frac{t}{T}\right)^{1/n} - \left(\frac{t-1}{T}\right)^{1/n}\right]$. The results in Mason and Shorack (1992) (with $k = m = 1$) apply to

$$\sqrt{T}(\tilde{\mu}_T - \mu_T)/\sigma_T,$$

so we need to show that

$$\sqrt{T}(\hat{\mu}_T - \tilde{\mu}_T - (\mu - \mu_T))/\sigma_T$$

converges to zero or, in the case where $1/n + 1/\gamma < 1/2$, is $\mathcal{O}_P(1)$.

We have

$$\hat{\mu}_T - \tilde{\mu}_T = B_{(n),(1)} \left(\frac{1}{T}\right)^{1/n} + B_{(n),(T)} \left[1 - \left(\frac{T-1}{T}\right)^{1/n}\right].$$

The first term is $\underline{b}T^{-1/n} + \mathcal{O}_P(T^{-1/\gamma-1/n})$ by standard arguments, and the last term converges to zero at a $1/T$ rate, so that $\hat{\mu}_T - \tilde{\mu}_T = \underline{b}T^{-1/n} + \mathcal{O}_P(T^{-1/\gamma-1/n} + T^{-1})$. By the last lemma, $\mu - \mu_T = \underline{b}T^{-1/n} + \mathcal{O}(T^{-1/\gamma-1/n} + T^{-1})$. Thus, the $\underline{b}$ terms cancel and

$$\sqrt{T}(\hat{\mu}_T - \tilde{\mu}_T - (\mu - \mu_T))/\sigma_T = \mathcal{O}_P(T^{1/2-1/n-1/\gamma}/\sigma_T + T^{-1/2}/\sigma_T).$$

If $1/n + 1/\gamma > 1/2$, σ_T is bounded away from zero and $T^{1/2-1/n-1/\gamma}$ and $T^{-1/2}$ go to zero so that this is $o_P(1)$. If $1/n + 1/\gamma = 1/2$, σ_T goes to infinity, $T^{1/2-1/n-1/\gamma}$ is constant, and the last term goes to zero, so that this display is $o_P(1)$. If $1/n + 1/\gamma < 1/2$, σ_T increases like $T^{1/2-1/n-1/\gamma}$ so that the above display is $\mathcal{O}_P(1)$.

Thus, $\hat{\mu}_T - \mu$ is approximated closely enough by $\tilde{\mu}_T - \mu_T$ that we just need to verify the claims of the theorem with $\hat{\mu}_T - \mu$ replaced by $\tilde{\mu}_T - \mu_T$. The first statement of the theorem, where $1/n + 1/\gamma > 1/2$, follows by part (i) of Theorem 1 in Mason and Shorack (1992) since $\sigma^2 < \infty$ in this case. For the second case, where $1/n + 1/\gamma = 1/2$, we have

$$\sqrt{T}(\tilde{\mu}_T - \mu_T)/\sqrt{\log T} = \sqrt{T}(\tilde{\mu}_T - \mu_T)/\sigma_T \cdot (\sigma_T/\sqrt{\log T}).$$

By Lemma 3, $\sigma_T/\sqrt{\log T} \to |q_0|$. $\sqrt{T}(\tilde{\mu}_T - \mu_T)/\sigma_T$ will converge to a $N(0,1)$ distribution by part (ii) of Theorem 1 in Mason and Shorack (1992) as long as condition (1.24) from that paper holds. This is equivalent to having $\sigma(\lambda u)/\sigma(u) \to 1$ as $u \to 0$ for all $0 < \lambda \leq 1$ (condition 1.27 from that paper). To see that this holds, note that

$$\sigma^2(\lambda u)/\sigma(u) = \frac{\sigma^2(\lambda u)/\log(\lambda u)}{\sigma^2(u)/\log u} \frac{\log(\lambda u)}{\log u}.$$

Here, $\sigma(\lambda u)/\log(\lambda u)$ and $\sigma(u)/\log(u)$ both converge to q_0 by Lemma 3, and $\log(\lambda u)/\log u = 1 + (\log \lambda)/(\log u) \to 1$, so the above display converges to 1 as $u \to 0$. Thus, $\sqrt{T}(\tilde{\mu}_T -$

$\mu_T)/\sqrt{\log T} \to N(0, q_0) = N(0, [n\gamma h_0^{1/\gamma}(1 - 1/n - 1/\gamma)]^{-2})$ as claimed.

For the last case, where $1/n + 1/\gamma < 1/2$, I verify the conditions of Theorem 2.1 in Mason and Shorack (1992). We have

$$T^{1/n+1/\gamma}(\tilde{\mu}_T - \mu_T) = \sqrt{T}(\tilde{\mu}_T - \mu_T)/\sigma_T \cdot (\sigma_T/T^{1/2-1/n-1/\gamma}).$$

By Lemma 3, $\sigma_T/T^{1/2-1/n-1/\gamma}$ converges to a constant as $T \to \infty$. As for the other term, $\sqrt{T}(\tilde{\mu}_T - \mu_T)/\sigma_T$ will be $\mathcal{O}_P(1)$ as long as $\limsup |\Phi_{iT}(c)| < \infty$ for all $c > 0$ for all $c > 0$ and $i = 0, 1$ (condition 2.2 in Mason and Shorack (1992)) where Φ_{iT} is defined in (2.1) in that paper. $\Phi_{1T}(c)$ goes to zero since $K(t)$ is bounded for t near 1. As for $\Phi_{0T}(c)$, we have, for large enough T, $\Phi_{0T}(c) = K(c/T)/(\sqrt{T}\sigma_T)$. By Lemma 2, $K(c/T)$ increases like $(c/T)^{1/n+1/\gamma-1}$, and, by Lemma 3, σ_T increases like $T^{1/2-1/n-1/\gamma}$, so the denominator increases like $T^{1-1/n-1/\gamma}$. Thus, $\Phi_{0T}(c)$ is bounded as T increases. The last claim of the theorem now follows by Theorem 2.1 in Mason and Shorack (1992).

A.2 Other Methods of Incorporating Observed Heterogeneity in the Application

The bounds computed in the application in Section 4 incorporate observed heterogeneity in the form of variation in engineer's estimates using somewhat strong independence and separability assumptions. This section reports estimates that control for this observed heterogeneity in a more nonparametric way by averaging over estimates with observations binned by engineer's estimate. The results illustrate how the bounds tighten as one incorporates more of the observed heterogeneity in a data set into the estimates.

The estimates are computed as follows. I order the bids by engineer's estimate and divide the observations into m bins with approximately equal observations (the kth bin contains observations for which the engineer's estimate is strictly greater than its $(k-1)/m$ quantile and less than or equal to its k/m quantile). I then compute the estimate of the upper bound for the efficiency loss from a lottery as the average of the estimates for each bin. These estimates are then divided by the average of the engineer's estimate over all samples, so that they can be interpreted relative to the average scale of the auctions. As discussed in Section 5.1, this gives a valid upper bound on the efficiency loss from a lottery for any fixed m (where heterogeneity in the engineer's estimate within a bin is treated as unobserved), and the bound will be tighter for larger m. I compute 95% confidence regions by bootstrapping, with the upper endpoint held fixed in each bootstrap replication. I use the bootstrap rather than

number of bins	estimate	95% CI
1	2.5762	[0,2.7551]
2	1.6711	[0,1.7901]
5	0.9666	[0,1.0310]
10	0.6327	[0,0.6764]
15	0.4862	[0,0.5253]
20	0.4229	[0,0.4539]

Table 3: Estimation Results for Binned Estimates (3 Bidder Sample)

subsampling because of the small number of observations per bin in some of the estimates.

The results are reported in Table 3. As expected, the upper bounds shrink as more of the observed heterogeneity is controlled for by using a larger number of bins. The estimate that uses only one bin (thereby treating ignoring all observed heterogeneity and treating it as unobserved) is over twice the average engineer's estimate, which is barely informative at all. With 20 bins, which corresponds to 10 or 11 observations per bin, the upper bound is almost as tight as the one obtained under stronger independence and separability assumptions in Section 4.

A.3 Other Methods of Computing Standard Errors

In addition to the subsampling based standard errors used in the application in Section 4, standard errors for the case where a normal distribution is obtained at a root-n rate can be estimated directly using sample analogues. Some calculation shows that the variance of the limiting normal distribution in Theorem 6 can be written as

$$\int\int \left[\frac{1}{n-1}+\frac{n-2}{n(n-1)}G_{B_{(n)}}(u)^{1/n-1}\right]\left[\frac{1}{n-1}+\frac{n-2}{n(n-1)}G_{B_{(n)}}(v)^{1/n-1}\right] \cdot \left[G_{B_{(n)}}(u\wedge v)-G_{B_{(n)}}(u)G_{B_{(n)}}(v)\right]\,dudv$$

Replacing $G_{B_{(n)}}$ with its sample analogue gives

$$\int\int \left[\frac{1}{n-1} + \frac{n-2}{n(n-1)}\hat{G}_{B_{(n)}}(u)^{1/n-1}\right]\left[\frac{1}{n-1} + \frac{n-2}{n(n-1)}\hat{G}_{B_{(n)}}(v)^{1/n-1}\right]$$
$$\cdot \left[\hat{G}_{B_{(n)}}(u \wedge v) - \hat{G}_{B_{(n)}}(u)\hat{G}_{B_{(n)}}(v)\right] dudv$$
$$= \sum_{s=1}^{T-1}\sum_{t=1}^{T-1} \left[\frac{1}{n-1} + \frac{n-2}{n(n-1)}(s/T)^{1/n-1}\right]\left[\frac{1}{n-1} + \frac{n-2}{n(n-1)}(t/T)^{1/n-1}\right]$$
$$\cdot \left[(s \wedge t)/T - (s/T)(t/T)\right](B_{(n),(s+1)} - B_{(n),(s)})(B_{(n),(t+1)} - B_{(n),(t)}). \tag{10}$$

Since the asymptotic variance involves an integral of a function of $\hat{G}_{B_{(n)}}$ that goes to infinity as $\hat{G}_{B_{(n)}}$ goes to zero, consistency of this estimator does not follow immediately from uniform consistency of $\hat{G}_{B_{(n)}}$. Since the consistency of the subsample standard errors reported in the paper follows easily from existing general results, I leave the asymptotic properties of the sample analogue estimator for future research and report monte carlo results for both estimators in this section of the appendix.

I simulate the winning bids from an auction with $n = 3$ bidders from the distribution $G_{B_{(n)}}(t) = (3t/2)^3$. This corresponds to an auction with $F_V(v|U) = v$ and no unobserved heterogeneity, but is consistent with other value functions in the presence of unobserved heterogeneity. The upper bound for $E(V_{(n)} - V_i)$ is $1/4$. I compute $1,000$ monte carlo draws of this data generating process with $T = 215$ observations (the same number as in the application) and compute one sided nominal 95% confidence intervals for the upper bound using the sample analogue estimate (10), subsampling with subsample size 57 (as in the application), and the bootstrap, with the subsample and bootstrap estimates holding the upper endpoint estimate constant in each computation. The monte carlo coverage probabilities are, respectively, 0.8330, 0.7230 and 0.8160. All of these converage probabilities are somewhat below the nominal coverage probability, perhaps because of downward finite sample bias in the estimates.

The 95% confidence region for the efficiency loss from a lottery using the sample analogue estimator (10) of the asymptotic variance for the auctions in the application is $[0, 0.4304]$. Compared to the subsampling based confidence region of $[0, 0.4306]$ reported in Table 2 (last row), the sample analogue estimator of the asymptotic variance actually gives a slightly smaller confidence region than the subsampling estimator for this data set, despite having greater coverage in the monte carlos.

A.4 Primitive Conditions for Assumption 1

Assumption 1 places conditions directly on the distribution of the winning bid $G_{B_{(n)}}$. While this allows for greater generality than primitive conditions on the distribution of values and unobserved heterogeneity, it is more difficult to interpret. This section gives simple sufficient conditions for Assumption 1. If the value distribution $F_V(v|U)$ satisfies Assumption 1 for some γ and U shifts the lower support so that the lower support has a cdf that behaves like $(u-\underline{u})^{\psi}$, Assumption 1 will hold for $G_{B_{(n)}}(b)$ with γ given by $\gamma n+\psi$.

Theorem 7. *Suppose that $F_V(v|U)$ satisfies Assumption 1. If there is no unobserved heterogeneity, $G_{B_{(n)}}(b)$ will satisfy Assumption 1 with γ given by γn. If unobserved heterogeneity takes the form $F_V(v|U) = h(v;u)(v-u)^{\gamma}$ where U has pdf $r(u)(u-\underline{u})^{\psi-1}$ with lower support point $\underline{u}$ and $h(v;u)$ and $r(u)$ are bounded with bounded first derivatives and are bounded away from zero near $v=u$ and near $u=\underline{u}$, then Assumption 1 holds for $G_{B_{(n)}}(b)$ with γ given by $\gamma n+\psi$.*

Proof. The distribution $G_{B_1}(b|U)$ is given by $F_V(\beta^{-1}(b)|U)$, where $\beta(v;U)$ is the equilibrium bid function given by

$$\beta(v;U) = v - \frac{\int_{\underline{v}}^{v} F_V(\tilde{v}|U)^{n-1}\,dv}{F_V(v|U)^{n-1}}.$$

Suppose that the value distribution satisfies Assumption 1 conditional on $U=u$ for some $h(v;u)$. Then

$$\begin{aligned} G_{B_1}(b|U=u) &= F_V(\beta^{-1}(b)|U=u) = h(\beta^{-1}(b;u);u)(\beta^{-1}(b;u)-\underline{b}) \\ &= h(\beta^{-1}(b;u);u)\frac{\beta^{-1}(b;u)-\underline{b}}{b-\underline{b}}(b-\underline{b}). \end{aligned}$$

Let

$$\tilde{h}(b;u) = h(\beta^{-1}(b;u);u)\frac{\beta^{-1}(b;u)-\underline{b}}{b-\underline{b}}.$$

Under these conditions, the bid function can be written as

$$\begin{aligned}
\beta(v) - v &= \frac{\int_{\underline{v}}^{v} h(\tilde{v};u)^{n-1}(\tilde{v}-\underline{v})^{n-1}\,dv}{h(v;u)^{n-1}(v-\underline{v})^{n-1}} \\
&= \frac{[h(\tilde{v};u)^{n-1}(\tilde{v}-\underline{v})^{n}/n]_{\underline{v}}^{v} - \int_{\underline{v}}^{v} h(\tilde{v};u)^{n-2}h'(\tilde{v};u)(\tilde{v}-\underline{v})^{n}(n-1)/n\,dv}{h(v;u)^{n-1}(v-\underline{v})^{n-1}} \\
&= (v-\underline{v})/n - \frac{\int_{\underline{v}}^{v} h(\tilde{v};u)^{n-2}h'(\tilde{v};u)(\tilde{v}-\underline{v})^{n}(n-1)/n\,dv}{h(v;u)^{n-1}(v-\underline{v})^{n-1}} \\
&= (v-\underline{v})/n + k(v;u)(v-\underline{v})^2
\end{aligned}$$

for some function $k(v;u)$ such that k and its derivative with respect to v are bounded by a constant that depends only on the bound on h and its derivatives. It follows that $\beta'(v;u) \to (n-1)/n$, so that $\frac{d}{db}\beta^{-1}(b) = (\beta'(\beta^{-1}(b))^{-1} \to n/(n-1)$, as $v \to \underline{v} = \underline{b}$, where the limit is uniform in u if there is unobserved heterogeneity that satisfies the conditions of the theorem (in which case $\underline{v} = u$). Thus, uniformly in u,

$$\frac{\beta^{-1}(b) - \underline{b}}{b - \underline{b}} = \frac{\frac{d}{db}\beta^{-1}(b^*)(b-\underline{b})}{b-\underline{b}} \to \frac{n}{n-1}$$

as $b \to \underline{b}$, where b^* is between b and $\underline{b}$. It follows from this and taking derivatives of $\tilde{h}(b;u)$ that Assumption 1 holds for $G_{B_1}(b|U=u)$ with h_0 given by $h_0 n/(n-1)$. It follows from this that $G_{B_{(n)}}(b|U=u)$ satisfies this assumption with γ given by $n\gamma$ and h_0 given by $[h_0 n/(n-1)]^n$.

Now, suppose that U has pdf $r(u)(u-\underline{u})^{\psi-1}$ for some ψ and shifts the lower support point of v so that

$$F_V(v|U=u) = h(v;u)(v-u)^{\gamma}$$

Then

$$\begin{aligned}
G_{B_{(n)}}(b) &= \int G_{B_{(n)}}(b|U=u)\,du = \int_{\underline{u}}^{b} \tilde{h}(b;u)^n |b-u|_+^{n\gamma} r(u)(u-\underline{u})^{\psi-1}\,du \\
&= \int_{\tilde{u}=0}^{1} h(b;\tilde{u}(b-\underline{u})+\underline{u})|(1-\tilde{u})(b-\underline{u})|_+^{n\gamma} r(\tilde{u}(b-\underline{u})+\underline{u})[\tilde{u}(b-\underline{u})]^{\psi-1}(b-\underline{u})\,d\tilde{u} \\
&= (b-\underline{u})^{n\gamma+\psi}\int_{\tilde{u}=0}^{1} h(b;\tilde{u}(b-\underline{u})+\underline{u})|1-\tilde{u}|_+^{n\gamma} r(\tilde{u}(b-\underline{u})+\underline{u})\tilde{u}^{\psi-1}\,d\tilde{u}
\end{aligned}$$

where second line uses the change of variables $\tilde{u} = (u - \underline{u})/(b - \underline{u})$. It can be verified that this satisfies Assumption 1 with γ given by $n\gamma + \psi$ by taking limits and derivatives under the integral.

$\square$

A.5 Counterexample

This section provides a counterexample to show two points made in the main text. First, estimates that ignore unobserved heterogeneity do not in general provide bounds for the entire bid distribution. In this example, the value distribution estimated ignoring unobserved heterogeneity does not satisfy any first order stochastic dominance ordering with the true marginal value distribution. Second, the upper bound on bidder profits can be attained even if there is nontrivial unobserved heterogeneity. Thus, there are data generating processes for which this bound is sharp even if all of the bids are observed. In the example below, the bound is attained because the upper endpoint of the support of the bid distribution does not change with U.

Let V be uniform on $(1-U, 1+U)$ conditional on U with 2 bidders and let U be distributed on $(0,1)$ with pdf $f_U(u) = 3u^2$. Some calculation shows that bids are uniform on $(u-1, 1)$ conditional on $U = u$, so the cdf of the highest bid conditional on $U = u$ is $G_{B_{(2)}}(b|U = u) = \left(\frac{b-(1-u)}{u}\right)^2 I(1-u < b < 1)$ and the pdf is $g_{B_{(2)}}(b|U = u) = 2\frac{b-(1-u)}{u^2} I(1-u < b < 1)$. Thus,

$$
\begin{aligned}
g_{B_{(2)}}(b) &= \int_{u=0}^{1} 2\frac{b-(1-u)}{u^2} I(1-u < b < 1) f_U(u)\, du \\
&= \int_{u=0}^{1} 2\frac{b-(1-u)}{u^2} I(1-b < u < 1-b+u) 3u^2\, du = 6\int_{u=1-b}^{1} (u+b-1)\, du \\
&= 6\int_{t=0}^{b} t\, du = 3b^2
\end{aligned}
$$

and

$$
G_{B_{(2)}}(b) = \int_0^b 3t^2\, dt = b^3.
$$

This leads to $\tilde{G}(b) = b^{3/2}$ and $\tilde{g}(b) = \frac{3}{2}b^{1/2}$, so that the estimated bid shade using data on the winning bid is

$$\frac{\tilde{G}(b)}{(2-1)\tilde{g}(b)} = \frac{b^{3/2}}{\frac{3}{2}b^{1/2}} = \frac{2}{3}b.$$

Thus, ignoring unobserved heterogeneity leads to the conclusion that bidders with value less than v are exactly those with $b + 2b/3 < v$, which is equivalent to $b < (3/5)v$. This leads to estimating the cdf of the greatest valuation as

$$\tilde{F}_{V_{(2)}}(v) = G_{B_{(2)}}((3/5)v) = (3/5)^3 v^3$$

for $v \leq 5/3$. Since the estimated distribution of the greatest valuation has no support on above $5/3$, but the true distribution of the greatest valuation takes values in $[5/3, 2]$ with positive probability, the estimated distribution of the greatest valuation cannot first order stochastically dominate the true marginal distribution of the greatest valuation.

References

Andrews, D. W., and X. Shi (2009): "Inference Based on Conditional Moment Inequalities," *Unpublished Manuscript, Yale University, New Haven, CT.*

Andrews, D. W. K., and M. M. A. Schafgans (1998): "Semiparametric Estimation of the Intercept of a Sample Selection Model," *Review of Economic Studies*, 65(3), 497–517.

Aradillas-Lopez, A., A. Gandhi, and D. Quint (2010): "Nonparametric Testing and Identification in Ascending Auctions with Unobserved Heterogeneity," .

Armstrong, T. (2011): "Weighted KS Statistics for Inference on Conditional Moment Inequalities," *Unpublished Manuscript.*

Asker, J. (2008): "A Study of the Internal Organisation of a Bidding Cartel," .

Athey, S., and P. A. Haile (2002): "Identification of Standard Auction Models," *Econometrica*, 70(6), 21072140.

——— (2007): "Chapter 60 Nonparametric Approaches to Auctions," vol. Volume 6, Part 1, p. 38473965. Elsevier.

Athey, S., J. Levin, and E. Seira (2011): "Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions," *Quarterly Journal of Economics*, 126(1).

Chernozhukov, V., S. Lee, and A. M. Rosen (2009): "Intersection bounds: estimation and inference," *Arxiv preprint arXiv:0907.3503*.

de Castro, L. (2010): "Affiliation, Equilibrium Existence and Revenue Ranking of Auctions," Discussion Paper 1530, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

de Castro, L. I. d., and H. J. Paarsch (2010): "Testing affiliation in private-values models of first-price auctions using grid distributions," *The Annals of Applied Statistics*, 4(4), 2073–2098, Zentralblatt MATH identifier: 05910061; Mathematical Reviews number (MathSciNet): MR2829947.

Decarolis, F. (2010): "When the Highest Bidder Loses the Auction: Theory and Evidence from Public Procurement," .

Einav, L., and I. Esponda (2008): "Endogenous Participation and Local Market Power in Highway Procurement," .

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525574.

Haile, P. A., H. Hong, and M. Shum (2003): "Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions," Discussion paper, Cowles Foundation, Yale University.

Hill, J. B., and A. Shneyerov (2010): "Are There Common Values in First-Price Auctions? A Tail-Index Nonparametric Test," Discussion paper.

Imbens, G. W., and C. F. Manski (2004): "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72(6), 1845–1857.

Jun, S. J., J. Pinkse, and Y. Wan (2010): "A consistent nonparametric test of affiliation in auction models," *Journal of Econometrics*, 159(1), 46–54.

Khan, S., and D. Nekipelov (2008): "Semiparametric Efficiency in Irregularly Identified Models," .

KHAN, S., AND E. TAMER (2007): “Irregular identification, support conditions and inverse weight estimation,” .

KRASNOKUTSKAYA, E. (2009): “Identification and Estimation of Auction Models with Unobserved Heterogeneity,” .

LEE, S., K. SONG, AND Y. WHANG (2011): “Testing functional inequalities,” .

LI, T., I. PERRIGNE, AND Q. VUONG (2000): “Conditionally independent private information in OCS wildcat auctions,” *Journal of Econometrics*, 98(1), 129161.

MANSKI, C. F., AND E. TAMER (2002): “Inference on Regressions with Interval Data on a Regressor or Outcome,” *Econometrica*, 70(2), 519–546.

MASON, D. M., AND G. R. SHORACK (1990): “Necessary and sufficient conditions for asymptotic normality of trimmed L-statistics,” *Journal of Statistical Planning and Inference*, 25(2), 111–139.

——— (1992): “Necessary and Sufficient Conditions for Asymptotic Normality of L-Statistics,” *The Annals of Probability*, 20(4), 1779–1804.

MENZEL, K., AND P. MORGANTI (2011): “Estimation of Auctions with Incomplete Bidding Data,” .

MILGROM, P. R., AND R. J. WEBER (1982): “A Theory of Auctions and Competitive Bidding,” *Econometrica*, 50(5), 10891122.

POLITIS, D. N., J. P. ROMANO, AND M. WOLF (1999): *Subsampling*. Springer.

SHAKED, M. (1977): “A Concept of Positive Dependence for Exchangeable Random Variables,” *The Annals of Statistics*, 5(3), 505515.

SHORACK, G. R. (1997): “Uniform CLT, WLLN, LIL and bootstrapping in a data analytic approach to trimmed L-statistics,” *Journal of Statistical Planning and Inference*, 60(1), 1–44.

SOMAINI, P. (2011): “Competition and Interdependent Costs in Highway Procurement,” .

VAN DER VAART, A. W. (2000): *Asymptotic statistics*. Cambridge University Press.