Nonparametric Tests for Common Values In First-Price Sealed-Bid Auctions*

Philip A. Haile† Han Hong‡ Matthew Shum§

November 13, 2006

*****INCOMPLETE MID-REVISION DRAFT*****

Abstract

We develop and apply tests to discriminate between private values and common values models using data from first-price sealed-bid auctions. Although the distinction depends on the nature of bidders' latent private information, distinguishing between them is possible using equilibrium conditions and the facts that the "winner's curse" is (a) present only in common values auctions and (b) more severe when bidders face more competition. We develop nonparametric and semi-parametric procedures that allow for observable and unobservable auction-specific heterogeneity, as well as endogenous bidder participation. Variations of our tests enable evaluation of other hypotheses, including our maintained assumption of equilibrium bidding in the affiliated values auction paradigm. Empirically, we study two types of timber auctions held by the United States Forest Service.

Keywords: first-price auctions, common values, interdependent values, private values, winner's curse, endogenous participation, timber auctions

---

*We thank Don Andrews, Steve Berry, Ali Hortaçsu, Yuichi Kitamura, Tong Li, Costas Meghir, Harry Paarsch, Isabelle Perrigne, Rob Porter, Quang Vuong, three anonymous referees, seminar participants at the Tow conference in Iowa, SITE, the World Congress of the Econometric Society in Seattle, and several universities for insightful comments. We thank Hai Che, Ying Fan, Kun Huang, Grigory Kosenok, and Gustavo Soares for capable research assistance. We are grateful to the NSF, SSHRC, Sloan Foundation, and Vilas Trust for financial support.

†Department of Economics and Cowles Foundation, Yale University, and NBER. P.O. Box 208264, New Haven, CT 06520, `philip.haile@yale.edu`

‡Department of Economics, Duke University, P.O. Box 90097, Durham, NC 27708, `hanhong@duke.edu`

§Department of Economics, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218, `mshum@jhu.edu`

1 Introduction

In a common values (or "interdependent values") auction, information about factors affecting all bidders' valuations is dispersed. So a bidder would update his assessment of the value of winning if he learned the private information of an opponent. In a private values auction, opponents' private information would be of interest to a bidder only for strategic reasons—learning an opponent's assessment of the good would not affect his beliefs about his own valuation. This distinction is fundamental in the auction literature and has significant implications for empirical work and policy. Despite considerable attention to the econometrics of auctions, the question of how to empirically distinguish between the two paradigms remains largely unanswered. Laffont and Vuong (1996) have suggested that doing so is impossible.

Here we propose and apply an approach for discriminating between the common values (CV) and private values (PV) paradigms using observed bids at first-price sealed-bid auctions. Our approach is nonparametric in the sense that identification of the true paradigm relies only on the maintained assumptions of the standard theoretical framework, not on particular functional form or parametric distributional assumptions. In principle this testing strategy could be applied using parametric or nonparametric statistical methods. Here we develop and apply nonparametric and semi-parametric approaches.

The distinction between private and common values is important for theory, empirical work, and policy. Private value models are simpler to work with, although their relevance is sometimes questioned. On the other hand, without a formal testing procedure, we cannot be sure whether the nuances of equilibrium bidding in common values models are important or merely interesting. From an empirical perspective, the interpretation of bid data depends on which type of model applies. In fact, with the types of data typically available from auctions, identification holds for a large class of private values models but fails for many common values models (see, e.g., Laffont and Vuong (1996), Athey and Haile (2002), and Athey and Haile (2006b)). A test for common values could, therefore, be a valuable diagnostic tool for researchers hoping to use demand estimates from auctions to guide the design of markets. As another example, it is widely believed that sealed-bid auctions are less susceptible to collusion than ascending auctions (see, e.g., Athey, Levin, and Seira (2004)). Working against this advantage in a CV auction is the fact that, with competitive bidding, an ascending auction will yield higher expected revenues (Milgrom and Weber (1982)). A test for common values could determine whether there is a tradeoff to be concerned with in choosing the auction format.

Although suggestions are sometimes offered regarding when one might expect a private or common values model to apply, the distinction concerns the nature of private information—something about which there is little guidance from economic theory and little empirical evidence from which to generalize. One result is that different researchers studying similar auctions have made different choices between the paradigms.[1] Given the implications of estimating a misspecified model, a more formal approach for determining the appropriate paradigm based on observed behavior would be valuable in many applications.

From a broader perspective, models of adverse selection (a common values auction being just one example) have played a prominent role in the theoretical economics literature, yet the prevalence and magnitude of this type of informational asymmetry is not well established empirically. Because a first-price auction is a market institution particularly well captured by a tractable theoretical model, data from these auctions offer a promising opportunity to test for adverse selection using structure obtained from an economic model. By studying different types of timber auction contracts, we have an opportunity in our empirical work to explore some standard intuitions regarding when adverse selection is likely to arise, perhaps shedding light on what to expect in other types of markets as well.

Our testing approach is based on detecting bidders' rational responses to changes in the severity of the *winner's curse*, which exists only in a CV auction. Loosely, winning a CV auction reveals to the winner that he was more optimistic about the object's value than his opponents were. This "bad news" (Milgrom (1981)) becomes "worse" as the number of opponents increases—having the most optimistic signal among many bidders implies (on average) greater over-optimism than does being most optimistic among a few bidders. A rational bidder anticipates this and adjusts his expectation of the value of winning accordingly. In a PV auction, by contrast, the value a bidder places on the object does not depend on his opponents' information, so the number of bidders does not affect his expected utility from winning.

This idea is simple but involves a comparative static prediction about an unobserved expectation, not the observed bids. Variation in the level of competition affects the aggressiveness of *bids* even in a PV auction. However, economic theory enables us to separate this competitive response from responses to the winner's curse in equilibrium.[2] Here we rely heavily on insights from the recent literature on nonparametric identification and estimation of first-price auction models (e.g., Guerre, Perrigne, and Vuong (2000), Li, Perrigne,

[1] For example, for mineral rights auctions, contrast Hendricks, Pinkse, and Porter (2003) and Li, Perrigne, and Vuong (2000). For timber auctions, contrast Baldwin (1995) or Athey and Levin (2001) with Baldwin, Marshall, and Richard (1997), Athey, Levin, and Seira (2004) , Haile (2001), or Haile and Tamer (2003).

[2] Hong and Shum (2002) explore this separation in a parametric model to examine the overall effect of additional competition on revenue.

and Vuong (2000), Li, Perrigne, and Vuong (2002), Hendricks, Pinkse, and Porter (2003)). While our approach requires on a maintained hypothesis of equilibrium behavior, a first-price auction may be an environment particularly amenable to such an assumption. We also show how this maintained hypothesis can be tested. Our formal tests compare distributions of bidders' expected valuations (actually, particular conditional expectations, defined below) in auctions with varying numbers of bidders. We propose likelihood ratio tests that can account for the error introduced by estimation of bidders' conditional expected valuations.

Three issues, each of broader interest, arise in developing a testing procedure that can be applied widely: (i) observed auction heterogeneity, (ii) endogenous bidder participation, and (iii) unobserved heterogeneity. As is typical for the literature, our data contain a rich set of auction covariates that are observed by all bidders prior to their bidding decisions. We initially suppress these covariates and focus on the heterogeneity across bidders' beliefs about their valuations arising from private information. As usual, this can be interpreted conditioning on one value of the covariates, and it is well known that standard nonparametric smoothing techniques can be used in practice (see, e.g., Guerre, Perrigne, and Vuong (2000)). However, this will often be impractical, and we propose a simple alternative "homogenization" technique that can be applied under a separability assumption.

Our testing approach can be applied (in most cases directly or with minor modification) to a variety of models of endogenous participation, including models with binding reserve prices, entry fees, costly information acquisition, or unobserved heterogeneity that affects both participation and valuations (see Haile, Hong and Shum (2003) for details). We give particular attention here to the last of these, which is also the most difficult case. Unobserved heterogeneity by itself is one of the most serious challenges in the recent literature on first-price auctions, since identification of equilibrium bid functions relies on the econometrician's ability to fix bidders' beliefs about the distributions of competing bids they face. It is not clear that this is possible when bidders condition on information unavailable to the econometrician. Standard approaches (e.g.,Guerre, Perrigne, and Vuong (2000), Li, Perrigne, and Vuong (2000, 2002)) fail in such cases. We present two methods that enable identification and testing in such environments. These approaches are related to instrumental variables and control function approaches for triangular systems (e.g., Newey, Powell and Vella (1999), Imbens and Newey (2003)). Both approaches rely on an exclusion restriction (an instrument for participation) and a strong monotonicity assumption. Both require estimation of a participation model in the presence of two-sided truncation.

In our empirical analysis we examine two types of timber harvesting contracts sold by the United States Forest Service (USFS). In "scaled sales," which are often modeled with private values, bids are made in real prices per unit, with payments by the winner determined *ex post* by the product of these prices and the volume of timber actually harvested. In “lumpsum” sales, bids are once-and-for-all offers for a tract, raising the importance of bidders’ shared uncertainty over timber volumes and prices of wood products. Intuition might suggest that lumpsum sales are more likely to have a significant common value component, although there has been no consensus on this view in the literature.

Turning to the prior literature, interest in discriminating between private and common values goes back at least to Gilley and Karels (1981) and was the motivation behind Paarsch’s (1992) pioneering work on estimation of auction models. Paarsch (1992) proposed tests of particular parametric models of independent private values and pure common values models (see also Sareen (1999)). By exploiting recent developments in the literature (e.g., Guerre, Perrigne, and Vuong (2000)), we are able to avoid reliance on parametric restrictions for identification and to test a null hypothesis that includes all private value models within the standard symmetric affiliated values framework (Milgrom and Weber (1982)) against an alternative including all CV models (i.e., the complement) in that framework.

We are also not the first to explore variation in the number of bidders as a way of testing for common values. Gilley and Karels (1981) suggested examining the sign of the correlation between bids and the number of bidders as a test for common values. However, Pinkse and Tan (2004) have recently shown that this approach generally cannot distinguish CV from PV models in first-price auctions: in equilibrium, bids can increase or decrease in the number of bidders under either paradigm. For second-price sealed-bid auctions or English auctions, Paarsch (1991) and Bajari and Hortacsu (2003) have considered testing for the winner’s curse using a regression of bids on the number of bidders. Unfortunately, second-price sealed-bid auctions are rare in practice, and the applicability of this approach to English auctions is limited by the fact that the winning bidder’s willingness to pay is never revealed, creating a missing data problem (see, e.g., Athey and Haile (2006b) or Athey and Haile (2006a)), and further by ambiguity regarding the appropriate interpretation of losing bids (e.g., Bikhchandani, Haile, and Riley (2002), Haile and Tamer (2003)). Athey and Haile’s (2002) study of identification in auction models includes sufficient conditions for discriminating between common and private values using variation in the number of bidders. They focus on cases in which only a subset of the bids is observable, consider only exogenous participation, and do not develop statistical tests.

Finally, Hendricks, Pinkse and Porter (2003, footnote 2) have suggested a testing approach applicable when there is a binding reserve price, based on whether the support of the equilibrium bid distribution extends to the reserve price. Although our approach allows a binding reserve price, this is not required—an important advantage in our application and many others, including the mineral rights auctions studied by Hendricks, Pinkse, and Porter (2003).

The remainder of the paper is organized as follows. The next section describes the underlying model, the method for inferring bidders' expectations of their valuations from observed bids, and the main principle of our testing approach. For clarity, we focus initially on the baseline case in which participation is exogenous and there is no auction heterogeneity. In section 3 we provide the details of our formal tests and develop the necessary asymptotic theory. In section 4 we present our "homogenization" approach for incorporating auction-specific covariates. Section 5 takes on the more challenging problems of unobserved heterogeneity and endogenous participation. In section 6 we summarize the results of Monte Carlo experiments illustrating the performance of our tests. Section 7 then presents the empirical application. We conclude in section 8.

2 Model and Testing Principle

The underlying theoretical framework is Milgrom and Weber's (1982) general affiliated values model. Throughout we denote random variables in upper case and their realizations in lower case. We use boldface to denote vectors. An auction has $N \in \mathcal{N} = \{\underline{n} \ldots \bar{n}\}$ risk-neutral bidders, with $\underline{n} \geq 2$. We let $\mathcal{I}$ denote auction-specific public information, known to all bidders—e.g., characteristics of a tract of timber to be harvested. For now we fix $\mathcal{I}$ and suppress it in the notation. Each bidder i has a valuation $U_i \in (\underline{u}, \overline{u})$ for the object. His information about U_i consists of a private signal $X_i \in (\underline{x}, \overline{x})$. We let $\mathbf{X}_{-i}$ denote the vector of signals of i's opponents. Valuations and signals have joint distribution $\tilde{F}_n(U_1, \ldots, U_n, X_1, \ldots, X_n)$, which is assumed to have a positive joint density on $(\underline{u}, \overline{u})^n \times (\underline{x}, \overline{x})^n$. Because signals have purely an information role, we may assume each X_i has a uniform marginal distribution without loss of generality. We make the following standard assumptions (see Milgrom and Weber (1982)).

Assumption 1 *(Symmetry)* $\tilde{F}_n(U_1, \ldots, U_n, X_1, \ldots, X_n)$ *is exchangeable with respect to the indices* $1, \ldots, n$.

Assumption 2 *(Affiliation)* $U_1, \ldots, U_n, X_1, \ldots, X_n$ *are affiliated.*

Assumption 3 *(Nondegeneracy)* $E[U_i | X_i = x, \mathbf{X}_{-i} = \mathbf{x}_{-i}]$ *is strictly increasing in* x $\forall \mathbf{x}_{-i}$.

Initially, we also assume that the number of bidders is independent of valuations or signals:

Assumption 4 *(Exogenous Participation) For each $n < \bar{n}$ and all $(u_1, \ldots, u_n, x_1, \ldots, x_n)$, $\tilde{F}_n(u_1, \ldots, u_n, x_1, \ldots, x_n) = \tilde{F}_{\bar{n}}(u_1, \ldots, u_n, \infty, \ldots, \infty, x_1, \ldots, x_n, \infty, \ldots, \infty)$.*

Although we refer to this as an "exogenous participation" assumption, it is in fact consistent with many models of "endogenous participation" considered in the literature, including Levin and Smith's (1994) model of costly signal acquisition and the variants in Athey, Levin, and Seira (2004), Li and Zheng (2005) and Krasnokutskaya and Seim (2004). In these models, participation decisions are the results of mixed (or purified mixed) strategies, yielding the independence in Assumption 4. We drop this assumption in section 5.

A seller conducts a first-price sealed-bid auction for a single object; i.e., sealed bids are collected from all bidders, and the object is sold to the high bidder at a price equal to his own bid.[3] For ease of exposition (and consistent with our application) we assume there is no binding reserve price (see Haile, Hong and Shum (2003) for the extension, however). Under Assumptions 1–3, in an n-bidder auction there exists a unique symmetric Bayesian Nash equilibrium in which each bidder employs a strictly increasing bidding strategy $\beta(\cdot; n)$. As shown by Milgrom and Weber (1982) this strategy is characterized by the differential equation

$$v(x, x, n) = \beta(x; n) + \frac{\beta'(x; n) F_n(x|x)}{f_n(x|x)} \quad \forall x \tag{1}$$

where

$$v(x, x', n) \equiv E\left[U_i | X_i = x, \max_{j \neq i} X_j = x'\right], \tag{2}$$

$F_n(\cdot|x)$ is the distribution of the maximum signal among a given bidder's opponents conditional on his own signal being x, and $f_n(\cdot|x)$ is the corresponding conditional density.

The conditional expectation $v(x, x, n)$ plays a key role in our work. It gives a bidder's expectation of his valuation conditional on his signal and on his equilibrium bid being pivotal. For simplicity we will refer to this as the bidder's *conditional expected valuation.* Our testing approach is based on the fact that this expectation is decreasing in n whenever valuations contain a common value element. To show this, we first formally define private and common values.[4]

[3] We describe the auction as one in which bidders compete to buy. The translation to the procurement setting, where bidders compete to sell, is straightforward.

[4] Affiliation implies that $E[U_i|X_1, \ldots, X_n]$ is nondecreasing in all X_j, and symmetry implies that when the expectation strictly increases in some $X_j, j \neq i$, it strictly increases in all $X_j, j \neq i$. For simplicity our definition of common values excludes cases in which the winner's curse arises for some realizations of signals but not others. Without this, the results below hold with weak inequalities replacing some strict inequalities. Up to this simplification, our PV and CV definitions partition Milgrom and Weber's (1982) affiliated values framework.

Definition 1 *Bidders have* **private values** *iff* $E[U_i|X_1, \ldots, X_n] = E[U_i|X_i]$*; bidders have* **common values** *iff* $E[U_i|X_1, \ldots, X_n]$ *strictly increases in* X_j *for* $j \neq i$*.*

The definition of common values incorporates a wide range of models with a common value component, not just the special case of *pure common values*, where the value of the object is unknown but identical for all bidders.[5] Note as well that existence of factors affecting all valuations does not imply common values if information about these factors is symmetric.

The following lemma gives a key result enabling discrimination between PV and CV models: an exogenous increase in the number of competing bidders intensifies the winner's cure (in common values model), causing each bidder's conditional expected valuation to decline.[6]

Lemma 1 *Under Assumptions 1–4,* $v(x, x, n)$ *is invariant to* n *for all* x *in a PV model but strictly decreasing in* n *for all* x *in a CV model.*

Proof: Given symmetry, we focus on bidder 1 without loss of generality. With private values, $E[U_1|X_1, \ldots, X_n] = E[U_1|X_1]$, which does not depend on n. With common values

$$\begin{aligned} v(x, x, n) &\equiv E\left[U_1|X_1 = X_2 = x, X_3 \le x, \ldots, X_{n-1} \le x, X_n \le x\right] \\ &< E\left[U_1|X_1 = X_2 = x, X_3 \le x, \ldots, X_{n-1} \le x, X_n < \infty\right] \\ &= E\left[U_1|X_1 = X_2 = x, X_3 \le x, \ldots, X_{n-1} \le x\right] \equiv v(x, x, n-1) \end{aligned}$$

with the inequality following from the definition of common values. □

Informally, in equilibrium a rational bidder adjusts his expectation of his valuation downward to reflect the fact that he wins only when his own signal is higher than those of all opponents. This adjustment is larger when n is larger because each low signal implies bad news about his valuation.

[5] Our terminology corresponds to that used by, e.g., Klemperer (1999) and Athey and Haile (2002), although it is not the only one used in the literature. Some authors reserve the term "common values" for the special case we call pure common values and use the term "interdependent values" (e.g., Krishna (2002)) or the less accurate "affiliated values" for the class of models we call common values. Additional confusion sometimes arises because the partition of the Milgrom-Weber framework into CV and PV models is only one of two partitions that might be of interest, the other being defined by whether bidders' signals are independent. Note that dependence of bidders' signals is neither necessary nor sufficient for common values.

[6] The idea behind this result is well known. To our knowledge the first correct formal statements appeared simultaneously in the early drafts of Athey and Haile (2002) and Haile, Hong, and Shum (2003).

2.1 Interpretation of Observed Bids

We assume that for each n the researcher observes the bids $B_1, \ldots, B_n$ from T_n n-bidder auctions. We let $T = \sum_n T_n$ and assume that for all n, $\frac{T_n}{T} \to \rho_n \in (0, 1)$ as $T \to \infty$. Below we will add the auction index $t \in \{1, \ldots, T\}$ to the notation defined above as necessary. For simplicity we initially assume an identical object is sold at each auction, delaying explicit incorporation of auction-specific heterogeneity to sections 4 and 5. We assume throughout that each auction is independent of all others.[7]

As pointed out by Guerre, Perrigne, and Vuong (2000), the strict monotonicity of $\beta(\cdot; n)$ implies that in equilibrium the joint distribution of bidder signals is related to the joint distribution of bids through the relations

$$\begin{aligned} F_n(y|x) &= G_n(\beta(y;n)|\beta(x;n)) \\ f_n(y|x) &= g_n(\beta(y;n)|\beta(x;n))\, \beta'(y;n) \end{aligned} \tag{3}$$

where $G_n(\cdot|\beta(x;n))$ is the equilibrium distribution of the highest bid among i's competitors conditional on i's equilibrium bid being $\beta(x;n)$, and $g_n(\cdot|\beta(x;n))$ is the corresponding conditional density. Because $b_i = \beta(x_i;n)$ in equilibrium, the differential equation (1) can then be rewritten

$$v(x_i, x_i, n) = b_i + \frac{G_n(b_i|b_i)}{g_n(b_i|b_i)} \equiv \xi(b_i; n). \tag{4}$$

Although the realizations of $v(X_i, X_i; n)$ are not observed directly, the joint distribution of bids is. Hence, the ratio $\frac{G_n(\cdot|\cdot)}{g_n(\cdot|\cdot)}$ is nonparametrically identified. Because $x_i = \beta^{-1}(b_i; n)$, equation (4) implies that each $v\left(\beta^{-1}(b_i;n), \beta^{-1}(b_i;n), n\right)$ is identified as well. This need not be sufficient to identify the model (i.e., to identify $\tilde{F}_n(\cdot)$); however, identification of the distribution of $v(X_i, X_i, n)$ for each n will be sufficient for our purpose.

To address estimation, let B_{it} denote the bid made by bidder i at auction t, and let B^*_{it} represent the highest bid among i's opponents. We follow Guerre, Perrigne, and Vuong (2000) and Li, Perrigne, and Vuong (2000, 2002) and use nonparametric estimates of the form

$$\begin{aligned} \hat{G}_n(b;b) &= \frac{1}{T_n \times h_G \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} K\left(\frac{b - b_{it}}{h_G}\right) \mathbf{1}\left(b^*_{it} < b, n_t = n\right) \\ \hat{g}_n(b;b) &= \frac{1}{T_n \times h_g^2 \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbf{1}(n_t = n) K\left(\frac{b - b_{it}}{h_g}\right) K\left(\frac{b - b^*_{it}}{h_g}\right). \end{aligned} \tag{5}$$

[7] This is a standard assumption, but a strong one that serves to qualify almost all empirical studies of bidding, where data are taken from auctions in which bidders compete repeatedly over time. An important exception is Jofre-Bonet and Pesendorfer (2003).

Here h_G and h_g are bandwidths and $K(\cdot)$ is a kernel. $\hat{G}_n(b;b)$ and $\hat{g}_n(b;b)$ are nonparametric estimates of

$$G_n(b;b) \equiv G_n(b|b)g_n(b) = \frac{\partial}{\partial b}\Pr(B_{it}^* \leq m, B_{it} \leq b)|_{m=b}$$

and

$$g_n(b;b) \equiv g_n(b|b)g_n(b) = \frac{\partial^2}{\partial m \partial b}\Pr(B_{it}^* \leq m, B_{it} \leq b)|_{m=b}$$

respectively, where $g_n(\cdot)$ is the marginal density of bids in equilibrium. Because

$$\frac{G_n(b;b)}{g_n(b;b)} = \frac{G_n(b|b)}{g_n(b|b)} \tag{6}$$

$\frac{\hat{G}_n(b;b)}{\hat{g}_n(b;b)}$ is a consistent estimator of $\frac{G_n(b|b)}{g_n(b|b)}$. Hence, by evaluating $\hat{G}_n(\cdot,\cdot)$ and $\hat{g}_n(\cdot,\cdot)$ at each observed bid, we can construct a pseudo-sample of consistent estimates of the realizations of each $V_{it} \equiv v(X_{it}, X_{it}, n)$ using (4):

$$\hat{v}_{it} \equiv \hat{\xi}(b_{it}; n_t) = b_{it} + \frac{\hat{G}_n(b_{it}; b_{it})}{\hat{g}_n(b_{it}; b_{it})}. \tag{7}$$

This possibility was first articulated for the independent private values model by Laffont and Vuong (1993) and Guerre, Perrigne, and Vuong (2000), and has been extended to affiliated values models by Li, Perrigne, and Vuong (2000, 2002) and Hendricks, Pinkse, and Porter (2003).

2.2 Main Principle of the Test

Each $\hat{v}_{it}$ obtained from (7) is an estimate of $v(x_{it}, x_{it}, n_t)$. Let $F_{v,n}(\cdot)$ denote the distribution of the random variable $V_{it} = v(X_{it}, X_{it}, n)$. Because $F_{v,n}(v) = \Pr\left(v(X_{it}, X_{it}, n) \leq v\right)$, Lemma 1 and Assumption 4 immediately imply that under the PV hypothesis, $F_{v,n}(\cdot)$ must be the same for all n, while under the CV alternative, $F_{v,n}(v)$ must strictly increase in n for all v.

Theorem 1 *Under the private values hypothesis*

$$F_{v,\underline{n}}(v) = F_{v,\underline{n}+1}(v) = \ldots = F_{v,\bar{n}}(v) \quad \forall v. \tag{8}$$

Under the common values hypothesis

$$F_{v,\underline{n}}(v) < F_{v,\underline{n}+1}(v) < \ldots < F_{v,\bar{n}}(v) \quad \forall v. \tag{9}$$

3 Tests

3.1 Test for Common Values

If each $v_{it} = v\left(x_{it}, x_{it}, n\right)$ were directly observed, a wide variety of tests for stochastic dominance from the statistics and econometrics literature could be used directly (e.g, McFadden (1989), Anderson (1996), Davidson and Duclos (2000), Barrett and Donald (2003), Linton, Massoumi, and Whang (2002)). We must rely on estimates, $\hat{v}_{it}$, and any testing procedure must account for the estimation error this introduces. In addition, the smoothing involved in the estimation introduces finite sample dependence between estimates of $v(x, x, n)$ and $v(x', x'; n)$ for x' near x', which must also be accounted for. After exploring several possibilities, we have focused on a simple consistent testing approach based on standard multivariate one-sided linear hypothesis tests (see, e.g., Bartholomew (1959),Perlman (1969)).

The form of our null and alternative hypotheses suggests comparing means of the distributions of $F_{v,n}(\cdot)$ across different values of n. By Corollary 1, $E[v(X, X, n)]$ is the same for all n under private values but strictly decreasing in n with common values. A complication, however, is the need to trim boundary values of $\hat{v}_{it}$ (due to the first-stage kernel density estimation) while preserving the sharp predictions of the null and alternative hypotheses. We use a trimming rule that equalizes the truncation of bidder types caused by trimming across all values of n. Let $\hat{G}_n(\cdot)$ denote the empirical distribution of bids in n-bidder auctions. Let $\hat{b}_{\tau,n}$ denote the τth quantile of $\hat{G}_n(\cdot)$ and x_τ the τth quantile of $F_X(\cdot)$. We define the *quantile-trimmed mean* as

$$\mu_n \equiv E\left[v(X, X, n)\ \mathbf{1}\left(x_\tau \leq X \leq x_{1-\tau}\right)\right].$$

With strict monotonicity of equilibrium bid functions, this has sample analog

$$\hat{\mu}_{n,\tau} \equiv \frac{1}{n \times T_n} \sum_{t=1}^{T} \sum_{i=1}^{n} \hat{v}_{it} \mathbf{1}\{\hat{b}_{\tau,n} \leq b_{it} \leq \hat{b}_{1-\tau,n}, n_t = n\}$$

We can then test the hypotheses

$$H_0 : \mu_{\underline{n},\tau} = \cdots = \mu_{\bar{n},\tau} \tag{10}$$

$$H_1 : \mu_{\underline{n},\tau} > \cdots > \mu_{\bar{n},\tau} \tag{11}$$

which are implied by (8) and (9), respectively. Thus, quantile-trimmed means are sufficient to distinguish between private and common values. Let $\tilde{g}_n(\cdot)$ denote the marginal density of equilibrium bids in an n-bidder auction. Theorem 2 below then shows the consistency and asymptotic normality of each $\hat{\mu}_{n,\tau}$ using the following standard assumptions.

Assumption 5 *1. Each $G_n(b;b)$ is $R+1$ times differentiable in its first argument and R times differentiable in its second argument. Each $g_n(b;b)$ is R times differentiable in both arguments. The derivatives are bounded and continuous.*

2. $\int K(\epsilon)\,d\epsilon = 1$ and $\int \epsilon^r K(\epsilon)\,d\epsilon = 0$ for all $r < R$. $\int |\epsilon|^R K(\epsilon)\,d\epsilon < \infty$.

3. $h_G = h_g = h$. As $T \to \infty$, $h \longrightarrow 0$, $Th^2 \big/ \log T \longrightarrow \infty$, $Th^{2+2R} \longrightarrow 0$.

Theorem 2 *Suppose Assumption 5 holds, $\frac{(\log T)^2}{Th^3} \longrightarrow 0$ and $Th^{1+2R} \longrightarrow 0$. Then*
(i) $\hat{\mu}_{n,\tau} \xrightarrow{p} \mu_{n,\tau}$.
(ii) $\sqrt{T_n h}\left(\hat{\mu}_{n,\tau} - \mu_{n,\tau}\right) \xrightarrow{d} N(0, \sigma_n)$ where

$$\sigma_n = \left[\int \left(\int K(\epsilon') K(\epsilon' - \epsilon)\, d\epsilon'\right)^2 d\epsilon\right]\left[\frac{1}{n}\int_{b_{\tau,n}}^{b_{1-\tau,n}} \frac{G_n(b;b)^2}{g_n(b;b)^3}\tilde{g}_n(b)^2\, db\right]. \tag{12}$$

The proof is given in the appendix. An attractive feature of the trimmed mean is its convergence rate of $\sqrt{T_n h}$, which exceeds the rate $\sqrt{T_n h^2}$ obtainable (Guerre, Perrigne, and Vuong (2000)) for the for the estimated conditional valuations themselves. Intuitively, although the term $\hat{g}_n(m;b)$ in (7) is a bivariate kernel density estimator converging at rate $\sqrt{T_n h^2}$, in constructing the estimate $\hat{\mu}_{n,\tau}$ we average along the diagonal $m = b$.

The likelihood ratio (LR) test or the weighted power test of Andrews (1998) provide possible approaches for formulating test statistics based on Theorem 2. Because we do not have a good a prior choice of weighting function for Andrews' weighted power test, we have chosen to use the LR test.[8] [9]

Let σ_n denote the asymptotic variance given in (12) for each value of $n = \underline{n}, \ldots, \bar{n}$ and define $a_n \equiv \frac{T_n h}{\sigma_n}$. Then the asymptotic covariance matrix of the vector $\left(\hat{\mu}_{\underline{n},\tau} \;\ldots\; \hat{\mu}_{\bar{n},\tau}\right)'$ is

$$\Sigma = \begin{bmatrix} \frac{1}{a_{\underline{n}}} & 0 & 0 & 0 \\ 0 & \frac{1}{a_{\underline{n}+1}} & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{a_{\bar{n}}} \end{bmatrix}.$$

The restricted maximum-likelihood estimate of the quantile-trimmed mean under the null hypothesis (10) is then given by

$$\bar{\mu} = \frac{\sum_{n=\underline{n}}^{\bar{n}} a_n\, \hat{\mu}_{n,\tau}}{\sum_{n=\underline{n}}^{\bar{n}} a_n}.$$

[8]Monte Carlo results in Andrews (1998) comparing the LR test to his more general tests for multivariate one-sided hypotheses, which are optimal in terms of a "weighted average power," suggests that the LR tests are "close to being optimal for a wide range of [average power] weighting functions" (pg. 158).

[9]Because our null hypothesis consider consists of a point hypothesis on the "parameter" vector $\delta = (\mu_{\underline{n},\tau} - \mu_{\underline{n}+1,\tau}, \ldots, \mu_{\overline{n}-1,\tau} - \mu_{\overline{n},\tau})$, the difficulties discussed in Wolak (1991) do not arise here.

To test against the alternative (11), let $\mu^*_{\underline{n}}, \ldots, \mu^*_{\bar{n}}$ denote the solution to

$$\min_{\mu_{\underline{n}},\ldots,\mu_{\bar{n}}} \sum_{n=\underline{n}}^{\bar{n}} a_n \left(\hat{\mu}_{n,\tau} - \mu_n\right)^2 \quad s.t. \quad \mu_{\underline{n}} \geq \mu_{\underline{n}+1} \geq \cdots \geq \mu_{\bar{n}}. \tag{13}$$

Following Bartholomew (1959), define the test statistic

$$\bar{\chi}^2 = \sum_{n=\underline{n}}^{\bar{n}} a_n \left(\mu^*_{n,\tau} - \bar{\mu}\right)^2.$$

Under the null hypothesis, $\bar{\chi}^2$ is asymptotically distributed as a mixture of Chi-square random variables (Bartholomew (1959)):

$$\Pr\left(\bar{\chi}^2 \geq c\right) = \sum_{k=2}^{\bar{n}-\underline{n}+1} \Pr\left(\chi^2_{k-1} \geq c\right)\omega(k;\Sigma) \qquad \forall c > 0$$

where χ^2_j denotes a standard Chi-square distribution with j degrees of freedom, and each mixing weight $\omega(k;\Sigma)$ is the probability that the solution to (13) has exactly k distinct values when the vector $\{\hat{\mu}_{\underline{n},\tau}, \ldots, \hat{\mu}_{\bar{n},\tau}\}$ has a multivariate $N(0,\Sigma)$ distribution.

In practice the weights $\omega(k;\Sigma)$ can be obtained by simulation from the $MVN(0,\hat{\Sigma})$ distribution, where $\hat{\Sigma}$ is a diagonal matrix with elements obtained from sample analogs of (12). An alternative to using equation (12), is to estimate each element of Σ using bootstrapped distributions of the quantile trimmed means.[10]

3.2 Specification Tests

Our testing approach relies on the interpretation of bids through the assumption of equilibrium bidding. This assumption is testable using a very similar approach. In equilibrium of the affiliated values model, the union of the null and alternative hypotheses in (8) and (9) must hold:[11]

$$F_{v,\underline{n}}(v) \leq F_{v,\underline{n}+1}(v) \leq \ldots \leq F_{v,\bar{n}}(v) \quad \forall v. \tag{14}$$

The alternative hypothesis (misspecification) can be written

$$F_{v,m}(v) > F_{v,n}(v) \quad \exists\ v\ \exists\ \underline{n} \leq m < n \leq \bar{n}. \tag{15}$$

[10]Bootstrapping the distribution of the test statistic itself is problematic, because resampling under the null requires resampling *bids* under a restriction on the distribution of conditional expected valuations.

[11]Guerre, Perrigne, and Vuong (2000) have previously pointed out another testable restriction—monotonicity of the inverse bid function (i.e. the right-hand side of (4)). We will examine both restrictions in our application.

The least favorable null hypothesis for any element of the alternative is the point hypothesis

$$F_{v,\underline{n}}(v) = F_{v,\underline{n}+1}(v) = \ldots = F_{v,\bar{n}}(v) \quad \forall v. \tag{14'}$$

The hypotheses (14') and (15) are similar to the null and alternative for our test for common values. One difference is the reversed directionality of the alternative. A more important difference is that the alternative (15) is weaker than the strict first-order stochastic dominance of (9). Because of this, comparisons of means could fail to detect evidence of (15) even in large sample. A variation of our LR test, using a vector of quantiles, can do much better.[12] Let $v_{\tau,n}$ denote the τth quantile of the distribution of $v(X_i, X_i; n)$. For a given pair (m, n) with $m < n$, we will examine the hypotheses

$$H_0^s : v_{\tau_l,m} - v_{\tau_l,n} = 0, \ \forall l = 1, \ldots, L \tag{16}$$

$$H_1^s : v_{\tau_l,m} - v_{\tau_l,n} \leq 0, \ \exists l = 1, \ldots, L. \tag{17}$$

Equation (4) and monotonicity of the equilibrium bid function imply that $v_{\tau,n}$ can be estimated by

$$\hat{v}_{\tau,n} = \hat{b}_{\tau,n} + \frac{\hat{G}_n(\hat{b}_{\tau,n}; \hat{b}_{\tau,n})}{\hat{g}_n(\hat{b}_{\tau,n}; \hat{b}_{\tau,n})}.$$

Theorem 3 describes the limiting behavior of each $\hat{v}_{\tau,n}$. The proof is given in the appendix.

[12] A Kolmogorov-Smirnov test, examining the supremum distance between $F_{v,m}(\cdot)$ and $F_{v,n}(\cdot)$ is a natural candidate here and would provide a consistent specification test. In Haile, Hong and Shum (2003) we explored a modified Kolmogorov-Smirnov test relying on subsampling that initially appeared promising. Simulations, however, indicated that the test performed poorly in practice, especially in samples of modest size. Note that as the number of quantiles considered is grows, the distinction in practice between our quantile test and a Kolmogorov-Smirnov test becomes smaller, since a finite grid of values is typically used for the latter as well.

Theorem 3 *Assume (i) $K(\mu)$ and $|\mu K(\mu)|$ are bounded; (ii) $\int \mu K(\mu)\, d\mu = 0$; (iii) $\int \mu^2 K(\mu)\, d\mu < \infty$; (iv) $\lim T_n h^2 = \infty$ and $\lim T_n h^6 = 0$. Then as $T_n \longrightarrow \infty$ for each n,*
(i) $\hat{b}_{\tau,n} - b_{\tau,n} = O_p\left(\frac{1}{\sqrt{T_n}}\right)$.
(ii) for all b such that $\tilde{g}_n(b) > 0$

$$\sqrt{T_n h^2}\left[\hat{\xi}(b;n) - v\left(\beta^{-1}(b;n), \beta^{-1}(b;n), n\right)\right] = \sqrt{T_n h^2}\left(\frac{\hat{G}_n(b;b)}{\hat{g}_n(b;b)} - \frac{G_n(b|b)}{g_n(b|b)}\right)$$

$$\xrightarrow{d} N\left(0, \frac{1}{n}\frac{G_n(b|b)^2}{g_n(b|b)^3\, \tilde{g}_n(b)}\left[\int\int K(e)^2 K(e')^2\, de\, de'\right]\right).$$

(iii) For distinct values $\tau_1, \ldots, \tau_L$ in $(0,1)$, the L-dimensional vector with elements $\sqrt{T_n h^2}\left(\hat{\xi}\left(\hat{b}_{\tau_l,n}; n\right) - v\left(x_{\tau_l}, x_{\tau_l}, n\right)\right)$ converges in distribution to $N(0,\Omega)$, where Ω is a diagonal matrix with lth diagonal element

$$\Omega_l = \frac{1}{n}\frac{G_n(b_{\tau_l,n})|b_{\tau_l,n})^2}{g_n(b_{\tau_l,n})|b_{\tau_l,n})^3 \tilde{g}_n(b_{\tau_l,n})}\left[\int\int K(e)^2 K(e')^2\, de\, de'\right].$$

This immediately gives the following result.

Corollary 1 *Assume the same conditions on the kernel function $K(\cdot)$ and the bandwidth sequence. Under the null hypothesis $H_0 : v(x,x,n) = v(x,x,m)\,\forall x$, the L-dimensional vector $\hat{d}_{m,n} = \{\hat{d}_{m,n,l}, \ldots, \hat{d}_{m,n,L}\}$, where*

$$\hat{d}_{m,n,l} \equiv \frac{\sqrt{m T_m h^2}\left(\hat{v}(x_{\tau_l}, x_{\tau_l}, m) - \hat{v}(x_{\tau_l}, x_{\tau_l}, n)\right)}{\sigma_{m,l} + \sqrt{\frac{m T_m}{n T_n}}\sigma_{n,l}}$$

converges in distribution to $\mathcal{Z} \sim N(0,I)$, where for $j = n, m$:

$$\sigma^2_{j,l} = \frac{G_j\left(b_{\tau_l,j}|b_{\tau_l,j}\right)^2}{g_j\left(b_{\tau_l,j}|b_{\tau_l,j}\right)^3 \tilde{g}_j\left(b_{\tau_l,j}\right)}\left[\int\int K^2(x)\, K^2(y)\, dx dy\right].$$

For $m < n$ the null hypothesis (14') can be rewritten $H_0^s : d_{m,n,l} = 0, \forall l = 1, \ldots, L$ and the alternative as $H_1^s : d_{m,n,l} < 0, \exists l \in \{1, \ldots, L\}$. Following Perlman (1969) (see also Wolak (1989))a LR test statistic can then be defined as

$$EI = \max_d \hat{d}'_{m,n}\hat{d}_{m,n} - \left(\hat{d}_{m,n} - d\right)'\left(\hat{d}_{m,n} - d\right) \quad \text{subject to} \quad d \leq 0.$$

This is equivalent to L separate problems

$$\max_{d_l} \hat{d}_{m,n,l}^2 - \left(\hat{d}_{m,n,l} - d_l\right)^2 \quad \text{subject to} \quad d_l \leq 0$$

with solution given by $d_l = \hat{d}_{m,n,l}$ if $\hat{d}_{m,n,l} \leq 0$, and $d_l = 0$ if $\hat{d}_{m,n,l} > 0$. Therefore, the LR test statistic equals the sum of the square of the negative elements of $\hat{d}_{m,n,l}, l = 1, \ldots, L$:

$$EI = \sum_{l=1}^{L} \hat{d}_{m,n,l}^2 1\{\hat{d}_{m,n,l} \leq 0\}$$

The p-values are simple to obtain and do not require simulation from the asymptotic distribution. For a critical value c,

$$P\left(EI \geq c\right) = \sum_{l=0}^{L} P\left(\chi_l^2 \geq c\right) \binom{L}{l} \left(\frac{1}{2}\right)^L.$$

A limitation of this approach to specification is the restriction to comparison of two distributions at a time. Combining several nonoverlapping pairwise tests is straightforward given the independence of $\hat{F}_{\hat{v},n}, \hat{F}_{\hat{v},m}$ for $m \neq n$. Alternatively, the test based on quantile trimmed means may be used for the specification test (simply reversing the directionality of the test for common values). On the other hand, in many applications the moderate sample sizes will make it desirable to compare only 2 distribution functions, pooling estimates of $v(x_{it}, x_{it}, n)$ for "small n" together and comparing to those for "large" n. This will be the case in our application below.

3.3 Sequential Testing

A natural procedure in practice is to test increasingly restrictive hypotheses: first testing the maintained assumption of equilibrium bidding, as just described. If this test fails to reject, one can use the test of section 3.1 to discriminate further between private and common values.

In general, with such a sequential procedure, the distribution of the second test statistic will reflect the conditioning event that the first test was passed. Under the null hypothesis of equal distributions, sampling error will occasionally lead to large deviations from equality in the direction opposite that implied by common values. Such deviations will not be taken as evidence against the null in a one-sided test of private values, but could lead to rejection in the first-stage specification test. Passing the first stage-test rules out large deviations of this type.

The correction required will typically be quite small since passing the first test only rules out differences between $\hat{F}_{\hat{v},n}$ and $\hat{F}_{\hat{v},m}$ in the tail of the null distribution that is opposite the tail of interest for the second test.[13] Here, in fact, we can proceed with no correction, since the two test statistics are asymptotically independent.

Theorem 4 *As $T \to \infty$, $\forall c_1, c_2 \geq 0$, $\Pr(\bar{\chi}^2 > c_2 | EI < c_1) \to \Pr(\bar{\chi}^2 > c_2)$.*

Proof. To be added.

4 Observable Auction Heterogeneity

So far we have assumed that data were available from auctions of *ex ante* identical goods. This is rarely the case. Bidders typically observe some public information $\mathcal{I}$ about each auction before bidding. This information affects not only bidders' valuations, but also their beliefs about competing bids. To make this explicit, we can rewrite the first-order condition (4) as

$$v(x_i, x_i, n, \mathcal{I}) = b_i + \frac{G_n(b_i|b_i;\mathcal{I})}{g_n(b_i|b_i;\mathcal{I})} \equiv \xi(b_i; n, \mathcal{I}) \tag{18}$$

where $G_n(b|b;\mathcal{I}) = \Pr(\max_{j \neq i} B_j \leq b | B_i = b, N = n, \mathcal{I})$ etc.

One example of public information is auction-specific observables. In USFS timber auctions, as in many other applications, we observe auction-specific characteristics $\mathbf{Y} \in \mathcal{Y}$ that are also observable to bidders before the auction. As is well known (see, e.g., Guerre, Perrigne, and Vuong (2000)), standard smoothing techniques can be used to estimate $\frac{G_n(b|b;\mathbf{y})}{g_n(b|b;\mathbf{y})}$, replacing $\mathcal{I}$ with $\mathbf{y}$ in (18). With many covariates, however, this will require larger data sets than those typically available.

An alternative is to place some structure on how covariates affect valuations. Consider a function $\Gamma : \mathcal{Y} \to \mathbb{R}$ and, for simplicity, assume there is some $\mathbf{y}_0 \in \mathcal{Y}$ such that[14]

$$E[\Gamma(\mathbf{Y})] = \Gamma(\mathbf{y}_0).$$

Now consider the additively separable structure

$$v(x, x, n, \mathbf{y}) = v(x, x, n, \mathbf{y}_0) + \Gamma(\mathbf{y}) \tag{19}$$

[13] For intuition, imagine a sequence of two 5% t-tests, each with a null hypothesis that a parameter μ equals zero. In the first test the alternative is $\mu < 0$, while in the second it is $\mu > 0$. If the first test fails to reject, the 5% critical value for the second test will be at the 0.9525 quantile of the t distribution instead of the 0.95 quantile.

[14] The covariate value $\mathbf{y}_0$ need not be unique and, in fact, need not exist at all. This is only convenient notation for a normalization.

with $\mathbf{Y}$ independent of $X_1, \ldots, X_n$.[15] Additive separability is a strong assumption, although one used frequently in practice. It will be more natural when valuations are normalized, e.g. by an engineer's estimate of a project's value or by the "size" of the project (e.g., miles of road to be paved or board-feet of timber to be harvested). If the covariates enter multiplicatively rather than additively, an approach analogous to that proposed here applies. This separable structure is particularly useful because it is preserved by equilibrium bidding.

Lemma 2 *Assume the separable structure (19) with* $\mathbf{Y}$ *independent of* $\mathbf{X}$*. Then the equilibrium bid function, conditional on* $\mathbf{Y} = \mathbf{y}$*, has the additively separable form* $\beta(x; n, \mathbf{y}) = \beta(x; n, \mathbf{y}_0) + \Gamma(\mathbf{y})$.

The proof follows the standard equilibrium derivation for a first-price auction (only the boundary condition for the differential equation (1) changes) and is therefore omitted. An important implication of this result is that we can account for observable heterogeneity in a two-stage procedure that avoids the need to condition on (smooth over) $\mathbf{Y}$ when estimating distributions and densities of bids. Letting $\alpha(n) = E[\beta(X; n, \mathbf{y}_0)]$ we can write the equilibrium bidding strategy as

$$\beta(x; n, \mathbf{y}) = \alpha(n) + \Gamma(\mathbf{y}_0) + \tilde{\Gamma}(\mathbf{y}) + \tilde{\beta}(x; n, \mathbf{y}_0)$$

where $\tilde{\Gamma}(\mathbf{y}) = \Gamma(\mathbf{y}) - \Gamma(\mathbf{y}_0)$ and the stochastic term $\tilde{\beta}(x; n, \mathbf{y}_0)$ has mean zero conditional on $(n, \mathbf{y})$ by independence of $\mathbf{Y}$ and X. Now observe that

$$b_{it}^0 \equiv \alpha(n_t) + \Gamma(\mathbf{y}_0) + \tilde{\beta}(x_{it}; n_t, \mathbf{y}_0) \tag{20}$$

is the bid that bidder i *would have submitted* in equilibrium if, instead of auction t, he were in a "generic" n_t-bidder auction—i.e., an auction with $\tilde{\Gamma}(\mathbf{y}) = 0$. Hence, given estimates $\widehat{\tilde{\Gamma}(\mathbf{y})}$ of each $\tilde{\Gamma}(\mathbf{y})$, we can construct estimates $\hat{b}_{it}^0 = b_{it} - \widehat{\tilde{\Gamma}(\mathbf{y})}$ of each b_{it}^0. Our tests can then be applied using these "homogenized" bids *as if* they were from a sample of auctions of identical goods. The asymptotic distributions of our test statistics are not affected by the homogenization procedure as long as $\widehat{\tilde{\Gamma}(\mathbf{y})}$ converges at a faster rate than the nonparametric estimates of the conditional expected valuations. This is guaranteed, for example, if $\tilde{\Gamma}(\cdot)$ is parametrically specified. Note that the function $\tilde{\Gamma}(\cdot)$ is estimated using all bids in the sample rather than separately for each value of n. This can make it possible to incorporate a large set of covariates and to employ a flexible specification of $\tilde{\Gamma}(\cdot)$.

[15] By (2), this follows if valuations U_i are themselves additively separable in $\mathbf{Y}$ and a stochastic component $\tilde{U}_i$, with $\tilde{U}_1, \ldots, \tilde{U}_n, X_1, \ldots, X_n$ independent of $\mathbf{Y}$.

5 Unobserved Heterogeneity and Endogenous Participation

As discussed already, the results above allow standard models endogenous participation due to information acquisition costs. They can also be extended in fairly straightforward ways to models with binding reserve prices, entry fees, or bid preparation costs as well (see Haile, Hong, and Shum (2003) for details). In this section we consider an important and challenging case in which participation is determined in part by unobserved factors that also affect the distribution of bidders' valuations.

Two problems arise in such an environment. The first is a problem of an endogenous "treatment": if auctions with many bidders tend to be those in which the good has a particularly high (or low) value, this correlation will confound attempts to determine whether variation in n leads to variation in the severity of the winner's curse. Second, unobserved heterogeneity threatens the identifiability of bidders' values from first-order conditions that underlies our testing approach. This is easy to see in the first-order condition (18): unless the econometrician can condition on all elements of $\mathcal{I}$ affecting bidders' beliefs, the right-hand-side cannot be estimated. This alone is a difficult problem that arises in any application of the first-order condition approach to first-price auctions where unobserved heterogeneity is present. Here we will describe additional structure under which both problems can be overcome. Our solution to the second problem builds on an idea in Campo, Perrigne and Vuong (2003) and may be useful many other applications.[16]

Conditional on auction characteristics $(\mathbf{Y}, \mathbf{Z}, W)$, let

$$N = \phi(\mathbf{Y}, \mathbf{Z}, W). \tag{21}$$

Here the vectors $\mathbf{Y}$ and $\mathbf{Z}$ are observable to bidders and to the researcher, and W is an index reflecting factors observable only to bidders and which may be correlated with their valuations. Aside from the restriction to a scalar representation of unobservables, this is a very general model of an environment in which participation is affected by both observable and unobservable factors. We make the following additional assumptions.

Assumption 6 $\mathbf{Y}$ *is independent of* W*.*

[16] Building on the result in Lemma 2, Krasnokutskaya (2003) has recently shown that methods from the literature on measurement error can be used to identify a private values model in which unobserved heterogeneity enters multiplicatively (or additively) and is independent of the idiosyncratic components (themselves independently distributed) of bidders' values. Hence, whereas we impose structure on the participation process, she puts restrictions on the underlying structure of bidder's private information. Athey, Levin, and Seira (2004) extend her approach using different functional form assumptions, again for an independent private values framework. Li and Zheng (2005) have explored an alternative parametric model of participation and bidding with unobserved heterogeneity in independent private values auctions. None of these allows participation decisions to depend on unobservables affecting bidders' valuations.

Assumption 7 *Conditional on* $\mathbf{Y}$, $\mathbf{Z}$ *is independent of* $(U_1, \ldots, U_n, X_1, \ldots, X_n, W)$.

Assumption 8 ϕ *is strictly increasing in* W, *and non-constant in* $\mathbf{Z}$.

Assumption 9 *For all* $(\mathbf{y}, \mathbf{z})$, *the support of* $\phi(\mathbf{y}, \mathbf{z}, W)$ *is a finite set of contiguous integers.*

Assumption 6 merely makes explicit that we consider the case of a single endogenous auction characteristic. Assumption 7 is an exclusion restriction providing an instrument for N. In Assumption 8 the requirement that ϕ be non-constant in $\mathbf{z}$ is the usual condition that the instrument have predictive power for the endogenous variable. Strict monotonicity of $\phi(\cdot)$ in the unobservable W is a strong restriction that is key to resolving the identifiability problem described above. This assumption requires that W be discrete, since N is discrete. This is a strong assumption, although in practice it is common to assume a discrete support for unobserved heterogeneity. In our application we will consider up to 9 points of support. The key implication of strict monotonicity is that $(N, \mathbf{Y}, \mathbf{Z})$ are joint sufficient statistics for W: fixing the values of the observables N, $\mathbf{Y}$ and $\mathbf{Z}$ indirectly holds fixed the unobservable W as well.[17] Assumption 9 seems fairly innocuous, primarily ruling out gaps in the support of $N|\mathbf{z}$.

Existence of an instrument $\mathbf{Z}$ is critical to our approach. In practice $\mathbf{Z}$ might be the number of potential bidders, proxies for this number, or observable auction characteristics that partially determine the set of firms that might bid—e.g., those with the appropriate specializations for the project—without affecting valuations (conditional on $\mathbf{Y}$). Possible examples from the literature include the number of firms requesting plans for a construction project (Krasnokutskaya and Seim (2004)), numbers of local firms in the industry (Haile (2001); Cantillon and Pesendorfer (2003)), the number of local firms eligible to bid (Paarsch (1997)), or the number of firms ever to bid in a similar auction (Hendricks, Pinkse, and Porter (2003)). In our application below, we use the numbers of local logging firms and sawmills as instruments. Along with observable characeristics $\mathbf{Y}$ of a tract, these provide a prediction of the number of participating bidders. Deviations from this prediction then arise due to the unobservable.

[17] Similar ideas have been used in a number of other contexts (e.g., Olley and Pakes (1996)). In the auction context, Campo, Perrigne, and Vuong (2003) have suggested this kind of structure to obtain identification of a private values first-price auction model with unobserved auction-specific heterogeneity. They assumed sufficiency of $(N, \mathbf{Z})$ for W directly. We model participation more explicitly and require an instrument because we also need to use this structure to develop an approach for testing. We will see below that our structure allows us to address some important practical considerations that do not arise in Campo, Perrigne and Vuong's application, since it involved no covariates and no instruments.

Let $\{w_k\}_{k=1}^{K}$denote the support of W, where $K \leq \infty$. Without loss of generality, let $w_{k+1} > w_k\ \forall k$. Let $p_k \equiv \Pr(W = w_k)$. The following result provides a useful representation of (21) in terms of a conditional mean or quantile.

Lemma 3 *Let*
(i) $\zeta(\mathbf{y}, \mathbf{z}) \equiv \max\{n \in \mathbb{N} : E[N|\mathbf{y}, \mathbf{z}] \geq n\}$ *or*
(ii) $\zeta(\mathbf{y}, \mathbf{z}) \equiv \max\{n \in \mathbb{N} : \Pr(N \leq n|\mathbf{y}, \mathbf{z}) \leq \alpha\}$ *for some given* $\alpha \in (0, 1)$.
Then under Assumptions 6–9 we may write $N = \zeta(\mathbf{Y}, \mathbf{Z}) + W$ *without loss of generality.*

Proof: See Appendix C

Below we develop two testing approaches that exploit this result and, therefore, require estimation of the participation model

$$N = \zeta(\mathbf{Y}, \mathbf{Z}) + W. \tag{22}$$

Complicating this estimation is the fact that in our application, as in many others, there is reason to be concerned about sample selection on the unobservable. For example, consider a tract of timber for which $\zeta(\mathbf{y}, \mathbf{z}) = 2$ and W takes its lowest possible value. This tract might attract no bidders (i.e., $\phi(\mathbf{y}, \mathbf{z}, w) \leq 0)$). Given the cost of organizing an auction, such a tract would not be offered for sale if the Forest Service observed W, leading to a truncated sample. Even if such an auction were offered and recorded in the data set, the first-order condition (4) does not apply (since this requires $n \geq 2$), forcing us to exclude it from the analysis of bids.

More generally, among the population of tracts with values of $(\mathbf{Y}, \mathbf{Z})$ suggesting a small number of bidders, those offered for sale and in which the first-order condition (4) holds will be only those from auctions with sufficiently favorable realizations of W; i.e., those for which $\zeta(\mathbf{Y}, \mathbf{Z}) + W \geq \underline{n} = 2$. Similar selection can arise for tracts with $(\mathbf{Y}, \mathbf{Z})$ suggesting a large number of bidders: one will typically observe too few auctions with the largest observed values of n to include them in a nonparametric analysis that must condition on n. This systematically truncates auctions with large values of W. In our application, for example, we exclude auctions with $n > \overline{n} = 7$. The result is a truncated sample.

As usual, truncation can lead to a violation of independence of W and $(\mathbf{Y}, \mathbf{Z})$ in the sample.[18]. Here this introduces two problems. One is the confounding effects of sample selection on attempts to assess the winner's curse through exogenous variation in $\mathbf{Z}$. We address this below. The second is the identification and estimation of the participation

[18] Assumptions 6 and 7 should be interpreted as assumptions on the latent random variables of the economic model, prior to any sample selection.

model (22) itself. This model is slightly nonstandard, due to the discreteness;[19] however, in Appendix D we provide conditions ensuring nonparametric identification of $\zeta(\cdot)$ under two-sided truncation.[20] There we also propose maximum likelihood estimation of the model using a parametric specification of $\zeta(\cdot)$.

5.1 Instrumental Variables Approach

5.1.1 Estimation

We will account for the auction-specific observables $\mathbf{Y}$ as in section 4, retaining the assumption $X \perp \mathbf{Y}$.[21] By Lemma 2, under the separable structure

$$\begin{aligned} v(x_i, x_i; n, \mathbf{y}, w) &\equiv E\left[U_i | X_i = \max_{j \neq i} X_j = x, \phi(\mathbf{Y}, \mathbf{Z}, W) = n, \mathbf{Y} = \mathbf{y}, W = w\right] \\ &= v(x_i, x_i; n, \mathbf{y}_0, w) + \Gamma(\mathbf{y}) \end{aligned}$$

equilibrium bidding takes the form

$$\begin{aligned} \beta(x; n, \mathbf{y}, w) &= \beta(x; n, \mathbf{y}_0, w) + \Gamma(\mathbf{y}) \\ &= \alpha(n, w) + \Gamma(\mathbf{y}_0) + \tilde{\Gamma}(\mathbf{y}) + \tilde{\beta}(x; n, \mathbf{y}_0, w) \end{aligned} \tag{23}$$

where $E\left[\tilde{\beta}(X; n, \mathbf{y}_0, w) | n, w, \mathbf{y}\right] = 0$

Since by Lemma 3, $W = N - \zeta(\mathbf{Y}, \mathbf{Z})$, (23) can be rewritten

$$\beta(x; n, \mathbf{y}, w) = \bar{\alpha}(n, \zeta(\mathbf{y}, \mathbf{z})) + \tilde{\Gamma}(\mathbf{y}) + \tilde{\beta}_0(x; n, \zeta(\mathbf{y}, \mathbf{z})). \tag{24}$$

The regression equation (24) can then be estimated using separate intercepts for each $(n, \zeta(\mathbf{y}, \mathbf{z}))$ combination. If the term $\tilde{\Gamma}(\mathbf{y})$ is then dropped, one obtains the "homogenized" bid

$$\beta(x_i; n, \mathbf{y}_0, w) = \bar{\alpha}(n, \zeta(\mathbf{y}, \mathbf{z})) + \tilde{\beta}_0(x_i; n, \zeta(\mathbf{y}, \mathbf{z}))$$

[19] In some cases one might be able to estimate the participation model using a censored rather than truncated regression model—for example if (a) one believes that the seller does not observe W before deciding which tracts to offer and (b) $\mathbf{Y}$ and $\mathbf{Z}$ are observed for all offered tracts, including those attracting no bidders.

[20] If one believes that the true sample selection process by the seller is less sharp than two-sided truncation at some known $\underline{n}$ and $\overline{n}$ (e.g., if some auctions attracting fewer than 2 bidders are offered anyway, perhaps due to noisy observation of W by the seller) the researcher can impose the truncated sampling structure directly by dropping auctions with n outside the appropriate interval.

[21] Auction covariates $\mathbf{Y}$ are not necessary, nor is the homogenization approach the only possible means of incorporating them. However, we discuss them explicitly here since the use of a "homogenized" sample that nonetheless allows different values of $\zeta(\mathbf{y}, \mathbf{z})$ is otherwise not completely transparent, and we will follow this approach in our application.

that bidder i would have made in a hypothetical auction in with $\mathbf{Y} = \mathbf{y}_0$, but with W nonetheless taking on the value $n - \zeta(\mathbf{y}, \mathbf{z})$ that it did in the actual auction. As in section 4, consistent estimates of these homogenized bids are easily constructed using the estimates of equation (24).

Now let B_i^0 denote the homogenized equilibrium bid of bidder i, characterized by first-order condition (the analog of (4))

$$v(x_i, x_i; n, \mathbf{y}_0, w) = b_i^0 + \frac{\Pr\left(\max_{j\neq i} B_j^0 \leq b_i^0 | B_i^0 = b_i^0, N = n, \mathbf{Y} = \mathbf{y}_0, W = w\right)}{\frac{\partial}{\partial m} \Pr\left(\max_{j\neq i} B_j^0 \leq m | B_i^0 = b_i^0, N = n, \mathbf{Y} = \mathbf{y}_0, W = w\right)\Big|_{m=b_i^0}}. \tag{25}$$

Let $\zeta = \zeta(\mathbf{y}, \mathbf{z})$ and define

$$G_{n,\zeta}(m|b) = \Pr(\max_{j\neq i} B_j^0 \leq m | B_i^0 = b, N = n, \mathbf{Y} = \mathbf{y}_0, W = n - \zeta)$$

with $g_{n,\zeta}(b_i|b_i)$ denoting the corresponding conditional density. Then we can rewrite (25) as

$$v(x_i, x_i; \mathbf{y}_0, n, w) = b_i + \frac{G_{n,\zeta}(b_i|b_i)}{g_{n,\zeta}(b_i|b_i)}. \tag{26}$$

Since each $\zeta = \zeta(\mathbf{y}, \mathbf{z})$ is identified, the right-hand side now consists of observables, resolving the problem of identification of bidders' conditional expected valuations in the presence of unobserved heterogeneity.[22]

Given a consistent estimate $\hat{\zeta}$ of each $\zeta(\mathbf{y}, \mathbf{z})$ with a sufficient rate of convergence, consistent estimators of bidders' conditional expected valuations can be obtained by conditioning on both n and $\hat{\zeta}(\mathbf{y}, \mathbf{z})$ to estimate the bidder markdowns $\frac{G_{n,\zeta}(b_i^0|b_i^0)}{g_{n,\zeta}(b_i^0|b_i^0)}$, using the obvious modifications of the kernel-based estimators described in section refstructrual. Here we see some of the practical value of Lemma 3: nonparametric conditioning on (smoothing over) the full vector of instruments $\mathbf{z}$, will often be impractical; however, $\zeta(\mathbf{y}, \mathbf{z})$ is a discrete scalar, and conditioning n and $\zeta(\mathbf{y}, \mathbf{z})$ is sufficient to fix W.

5.1.2 Testing

Lemma 3 suggests that we can test for common values by comparing estimates of $v(x_i, x_i; n, \mathbf{y}, w)$ in auctions with different $\zeta(\mathbf{y}, \mathbf{z})$. This is correct, although we must make an adjustment

[22] If the relationship between N and W were only weakly monotone, conditioning on $(n, \mathbf{y}, \mathbf{z})$ would be equivalent to conditioning on n and the event $\{W \in \mathcal{W}\}$, for some set $\mathcal{W}$. In some applications this might be sufficient to enable the use of the first-order condition (26) as a useful approximation.

for the sample selection discussed above. Intuitively, if small-$\zeta(\mathbf{y},\mathbf{z})$ auctions in the sample tend to have large realizations of W, variation in $\zeta(\mathbf{y},\mathbf{z})$ is not exogenous. This problem will be eliminated, however, if the truncation of W can be equalized across the sample.[23]

To do this, note that by Lemma 3 a consistent estimator $\hat{\zeta}(\mathbf{y},\mathbf{z})$ of $\zeta(\mathbf{y},\mathbf{z})$ immediately gives the consistent estimator

$$\hat{w}_t = n_t - \hat{\zeta}(\mathbf{y}_t, \mathbf{z}_t) \tag{27}$$

of each w_t. Now suppose that we want to compare estimates of $v(x_i, x_i; n, \mathbf{y}, w)$ across J values of $\zeta(\mathbf{y},\mathbf{z})$. Let

$$\{\underline{w}, \ldots, \overline{w}\} = \bigcap_{j=1}^{J} \bigcup_{t:\hat{\zeta}(\mathbf{z}_t)=\zeta_j} \hat{w}_t$$

denote the common support of $\hat{W}_t$ across all values of $\zeta(\mathbf{z})$. To equalize truncation over the sample we consider only auctions with w falling in this common support; i.e., we include auction t only if

$$\hat{w}_t \in \{\underline{w}, \ldots, \overline{w}\}.$$

For example, if $\text{supp}\{N|\zeta(\mathbf{y},\mathbf{z})\} = \zeta(\mathbf{y},\mathbf{z}) \pm 4$, comparing auctions with $\zeta(\mathbf{y},\mathbf{z}) = 3$ to those with $\zeta(\mathbf{y},\mathbf{z}) = 5$ will be valid if one restricts the first set of auctions to include only 2-, 3-, and 4-bidder auctions (i.e., auctions with $w \in \{-1, 0, +1\}$, and the second set to only 4-, 5-, and 6-bidder auctions. After this correction, the distribution of $W|\zeta(\mathbf{y},\mathbf{z})$ is again identical for all $\mathbf{y},\mathbf{z}$ in the sample, as in the population. With this *equalized truncation* property, we can prove the following theorem.

Theorem 5 *Suppose*
(i) $\zeta(\mathbf{Y},\mathbf{Z}) \equiv \max\{n \in \mathcal{N} : E[N|\mathbf{Y},\mathbf{Z}] \geq n\}$ *or*
(ii) $\zeta(\mathbf{Y},\mathbf{Z}) \equiv \max\{n \in \mathcal{N} : \Pr(N \leq n|\mathbf{Y},\mathbf{Z}) \leq \alpha\}$ *for some* $\alpha \in (0,1)$.
Then under Assumptions 6–9 and equalized truncation, $\Pr(v(X, X; \mathbf{y}, N, W)) \leq v|\zeta(\mathbf{y},\mathbf{Z}) = \zeta)$ *is (i) invariant to* ζ *for all* $\mathbf{y}$ *under PV, and (ii) strictly increasing in* ζ *for all* $\mathbf{y}$ *under CV.*

Proof. With equalized truncation and Assumptions 6 and 7, $\Pr(W = w_k|\mathbf{Y} = \mathbf{y}, \zeta(\mathbf{y},\mathbf{z}) = \zeta) =$

[23] This idea is also that underlying Powell's (1986) approach to truncated regression. The estimation problem in Powell (1986) is closer to that we face in estimating the participation model itself, although our model is discrete and we do not require a symmetry assumption.

p_k for $k = 1, \ldots, K$. Letting $F(\cdot|w)$ denote the marginal distribution of X_i given $W = w$,

$$
\begin{aligned}
&\Pr\left(v\left(X, X; \mathbf{y}, N, W\right) \leq v | \zeta\left(\mathbf{y}, \mathbf{Z}\right) = \zeta\right) \\
&= \sum_{k=1}^{K} p_k \int_{\underline{x}}^{\overline{x}} 1\left\{v\left(x, x; \mathbf{y}, \zeta + w_k, w_k\right) \leq v\right\} \, dF(x|w_k) \\
&= \sum_{k=1}^{K} p_k \int_{\underline{x}}^{\overline{x}} 1\left\{E\left[U_1 | X_1 = \max_{j \in \{2, \ldots, \zeta + w_k\}} X_j = x, \mathbf{Y} = \mathbf{y}, W = w_k\right] \leq v\right\} dF(x|w_k). \quad (28)
\end{aligned}
$$

By Assumption 7, $\zeta\left(\mathbf{y}, \mathbf{Z}\right) \perp \left(U_1, \ldots, U_n, X_1, \ldots, X_n\right)$. Therefore, by the argument proving Lemma 1, the expectation in the integrand of (28) is strictly decreasing in ζ under common values, but invariant to ζ under private values. $\square$

This result implies that our testing approach will remain valid under the assumptions above with three modifications. First, following homogenization of bids to control for auction-specific observables, nonparametric estimation of bidders' conditional expected valuations should be performed using the first-order condition (26)—in particular, by conditioning on both $\zeta(\mathbf{y}, \mathbf{z})$ (i.e., on an estimate thereof) and n. Second, some auctions must be excluded to equalize truncation, as described above.[24] Finally, the estimates $v\left(x_{it}, x_{it}; n_t, y_0, w_t\right)$, which may be rewritten $v\left(x_{it}, x_{it}; n_t, y_0, \zeta_t\right)$, must then be pooled into categories based on ζ_t (the empirical analog of summing over k in (28)) so that estimates of

$$F_{v,\zeta}\left(v\right) \equiv \Pr\left(v\left(X, X; n, \zeta\right) \leq v | \zeta\right)$$

can be compared at different values of ζ to test the null of equal distributions against the alternative of first-order stochastic dominance.

Some limitations of this approach should be mentioned. Equalizing the sample truncation can eliminate a considerable number of auctions from consideration — about *** percent of the auctions in our application. Further, even ignoring this the IV approach can require fairly large data sets, since estimating $g_{n,\zeta}(m|b)$ involves bivariate kernel density estimation on a sample that has already been split into "bins" by $\left(n, \zeta\left(\mathbf{y}, \mathbf{z}\right)\right)$. We next consider an alternative that makes somewhat weaker demands on the data, relying on additional separability and independence assumptions.

[24]Since the kernel densities are estimated separately for each value of w (indirectly, by conditioning on n and $\zeta(\mathbf{y}, \mathbf{w})$) one need not equalize the truncation before estimating the conditional expected valuations; however, for the same reason, this makes no difference in the end. Hence, the equalization can be performed prior to the kernel estimation to reduce computation. Conditioning on w (i.e., on an estimate) directly is also possible. Depending on the dimensionality of W relative to that of $\zeta(\mathbf{y}, \mathbf{w})$ this may more may not be advantageous in practice. However, see section 5.1.3 below.

5.1.3 Control Function Method

An alternative to the IV method is to use (27) to control directly for W, as in control function approaches for other triangular systems (e.g., Newey, Powell, and Vella (1999), Imbens and Newey (2002)). While similar to the IV approach, this can offer a significant practical advantage if one is willing to strengthen the assumptions above by requiring that valuations be additively separable in W.[25] In particular, letting

$$v(x,x;n,\mathbf{y},w) \equiv E[U_i|X_i = \max_{j\neq i} X_j = x, N=n, \mathbf{Y}=\mathbf{y}, W=w]$$

assume the separable structure

$$v(x,x;n,\mathbf{y},w) = v\left(x,x;n,\mathbf{y}_0\right) + \Gamma\left(\mathbf{y},w\right) \tag{29}$$

with $X \perp (\mathbf{Y}, W)$. Such an assumption, which imposes the same structure on the auction-specific unobservable that we made earlier on the auction-specific observables, will make it possible to eliminate the effects of the unobservable heterogeneity within the "homogenization" step. This will reduce the dimensionality of the nonparametric estimation of bidders' conditional expected valuations and also avoid the need for truncation equalization by directly controlling for the effects of W.

By Lemma 2, under (29) equilibrium bids take the same separable form

$$\begin{aligned} \beta\left(x;n,\mathbf{y},w\right) &= \beta\left(x;n,\mathbf{y}_0,w_0\right) + \Gamma(\mathbf{y},w) \\ &= \alpha\left(n\right) + \tilde{\Gamma}\left(\mathbf{y},w\right) + \tilde{\beta}\left(x;n,\mathbf{y}_0,w_0\right) \end{aligned} \tag{30}$$

with $E[\tilde{\beta}\left(X;n,\mathbf{y}_0,w_0\right)|\mathbf{Y},W] = 0$ and $(\mathbf{y}_0,\mathbf{w}_0)$ defined by $\tilde{\Gamma}(\mathbf{y}_0,\mathbf{w}_0) = 0$. If W were observable, its effects could therefore be controlled for in a fully flexible way by including dummy variables for each value of w when estimating the homogenization regression (30). This would be equivalent to treating w as one of the observables $\mathbf{y}$. Although W is not observable, we have already seen that it can be consistently estimated from the participation equation. With a sufficient rate of convergence (obtained by the MLE we propose in Apendix D) these estimates $\hat{w}_t$ may be substituted for the unobserved w_t and treated as data. Following previous arguments, estimation of the regression equation (30) then enables construction of consistent estimators of the homogenized bids

$$\beta\left(x;n,\mathbf{y}_0,w_0\right) = \alpha\left(n\right) + \tilde{\beta}\left(x;n,\mathbf{y}_0,w_0\right)$$

which are now free of the effects of both $\mathbf{Y}$ and W.

[25] Again, multiplicative separability would also do.

This effectively returns us to the baseline case of exogenous participation: biders' conditional expected valuations can be estimated using estimates of the markdowns $\frac{G_n(b|b)}{g_n(b|b)}$ that do not require conditioning on w or $\zeta(\mathbf{y},\mathbf{z})$. Further, our tests for common values can be based on the variation in n itself rather than an instrument, using the original baseline approach.

While additional separability and independence assumptions are required, these might be viewed as a mild additions to the same assumptions on the auction observables $\mathbf{Y}$. The practical advantage of this approach can be significant.

6 Monte Carlo Simulations

Here we summarize the results of Monte Carlo experiments performed to evaluate our testing approaches. We examine the performance of our tests on data generated by two PV models and two CV models:

(PV1) independent private values, $x_i \sim \mathrm{u}[0,1]$;

(PV2) independent private values, $\ln x_i \sim N(0,1)$;

(CV1) common values, i.i.d. signals $x_i \sim \mathrm{u}[0,1]$, $u_i = \frac{x_i}{2} + \frac{\sum_{j\neq i} x_j}{2(n-1)}$;[26]

(CV2) pure common values, $u_i = u \sim \mathrm{u}[0,1]$, conditionally independent signals x_i uniform on $[0,u]$.[27]

Before reporting the results, we turn to Figure 1. Here we illustrate the empirical distributions of estimated conditional expected valuations obtained by applying the first-stage nonparametric estimators using one simulated data set from each of the four models. We do this for $n = 2,\ldots,5$, with $T_n = 200$. For the PV models, the estimated distributions are very close to each other. For the CV models these distributions clearly suggest the first-order stochastic dominance relation implied by the winner's curse. Note that in both model CV1 and model CV2, the effect of a change in n on the distribution appears to be largest when n is small. This is the case in many CV models and is quite intuitive: the difference between $E[U_1|X_1 = \max_{j\in\{2,\ldots,n\}} X_j = x]$ and $E[U_1|X_1 = \max_{j\in\{2,\ldots,n+1\}} X_j = x]$ typically shrinks as n grows. This is important because most auction data sets contain relatively few

[26] Here $v(x,x,n) = \frac{3n-2}{4(n-1)}x$, leading to the equilibrium bid function $\beta(x;n) = \frac{3n-2}{4n}x$. It is easy to see that although $v(x,x,n)$ is strictly decreasing in n, $\beta(x;n)$ strictly increases in n.

[27] The symmetric equilibrium bid function for this model is given in Matthews (1984).

observations for n large but many observations for n small—exactly where the effects of the winner's curse are most pronounced.

We perform monte carlo simulations that compare the finite sample performance of a LR test based on quantile-trimmed means that is discussed in details in Haile, Hong, and Shum (2003) and the generalized Kolmogorov-Smirnov test. Tables 1 and 2 report the performances of the LR test based on quantile-trimmed means, with table 1 using an analytic asymptotic variance and table 2 using a bootstrap asymptotic variance for the test statistic. Table 3 reports the performance of the generalized Kolmogorov-Smirnov test statistic. Each of tables 1 to 3 is based on 200 replications of each experiment. In addition to the LR test based on quantile-trimmed means and the generalized Kolmogorov-Smirnov test, Haile, Hong, and Shum (2003) also studied a quantile test that is based on comparing a finite number of quantiles of the distribution across different number of bidders. It is much easier to derive the asymptotic distribution of the quantile test statistic. However, both the asymptotic and finite sample powers of the quantile test are much lower than those of the LR test based on quantile-trimmed means and the generalized Kolmogorov-Smirnov statistic. Therefore we do not report the monte carlo results for the quantile test here.

We first consider the LR test based on quantile-trimmed means. Tables 1 and 2 summarize the test results, using tests with nominal size 5% and 10%. The last two rows in Table 1 indicate that in the PV models there is a tendency to over-reject when sample analogs of the asymptotic variance covariance matrix are used to construct the mixing weights in Corollary 2 of Haile, Hong, and Shum (2003). For example, for tests with nominal size 10% and data generated by the PV1 model, we reject 20.5% of the time when the range of bidders is 2–4, and 39% of the time when the range of bidders is 2–5. The tests do appear to have good power properties, rejecting the CV models in 70 to 100 percent of the replications. However, the over-rejection under the null is a concern.

One possible reason for the over-rejections is that the asymptotic approximations of the variances of the mean estimated conditional expected valuations derived in Theorem 3 of Haile, Hong, and Shum (2003) may be poor at the modest sample sizes we consider. We have considered an alternative of using bootstrap estimates of the elements of Σ.[28] We use a block bootstrap procedure that repeatedly selects an auction from the original sample at random and includes all bids from that auction in the bootstrap sample, thereby preserving any dependence between bids within each auction. The results, reported in Table 2, indicate that the tendency towards over-rejection is attenuated when we estimate these variances

[28]Note that we are not bootstrapping the distribution of the test statistic given in Corollary 2 of Haile, Hong, and Shum (2003), only the component Σ. Bootstrapping the distribution of the test statistic would require resampling bids under the null hypothesis on the latent conditional expected valuations $v(x, x, n)$.

with the bootstrap. For a test with nominal size 10%, we now reject no more than 14% of the time when the range of n is 2–4, and 18% of the time when the range of n is 2–5. With a 5% nominal size, our rejection rates range between 4% and 12%. The power properties remain very good. These results are encouraging and suggest use of the bootstrap in practice.

... to be completed ...

7 Application to U.S. Forest Service Timber Auctions

7.1 Data and Background

We apply our tests to auctions held by the United States Forest Service (USFS). In each sale, a contract for timber harvesting on Federal land was sold by first-price sealed bid auction. Detailed descriptions of the contracts being sold and the auctions themselves can be found in, e.g., Baldwin (1995), Baldwin, Marshall, and Richard (1997), Athey and Levin (2001), Haile (2001), or Haile and Tamer (2003). Here we discuss a few key features that are particularly relevant to our analysis.

We will separately consider two types of USFS auctions, for which the significance of common value elements may be different. The first type is known as a *lumpsum* sale. As the name suggests, here each bidder offers a total bid for an entire tract of standing timber. The winning bidder pays his bid regardless of the volume actually realized at the time of harvest. Bidders, therefore, may face considerable common uncertainty over the volume of timber on the tract. More significant, individual bidders often conduct their own "cruises" of the tract before the auction, creating a natural source of the private information essential to the CV model. Before each sale, however, the Forest Service conducts its own cruise of the tract to provide bidders with estimates of (among other things) timber volumes by species, harvesting costs, costs of manufacturing end products from the timber, and selling prices of these end products. This creates a great deal of common knowledge information about the tract.[29] Whether scope remains for significant private information regarding determinants of tract value common to all bidders is uncertain. It should also be pointed out that the fact that bidders cruise does not by itself imply common values. Information acquired in a private cruise could concern only factors idiosyncratic to each bidder, consistent with a private values model. Information from a private cruise regarding shared determinants of valuations — e.g., volume, quality, distribution of volume across species — would lead to

[29] Indeed, USFS cruises for lumpsum sales are, by design, more thorough than those for "scaled sales," discussed below. Below we treat the two samples separately, allowing for this difference in the covariates.

common values only if bidders get different signals, e.g., due to sampling techniques used by the private "cruisers."

The second type of auction is known as a "scaled sale." Here, bids are made on a per unit (thousand board-feet of timber) basis. The winner is selected based on these unit prices and the *ex ante* estimates of timber volumes obtained from the Forest Service cruise. However, payments to the Forest Service are based on the winning bidder's unit prices and the *actual* volumes, measured by a third party at the time of harvest. As a result, the importance of common uncertainty regarding tract values may be reduced. In fact, bidders are less likely to send their own cruisers to assess the tract value for a scaled sale (National Resources Management Corporation (1997)). This may leave less scope for private information regarding any shared determinants of bidders' valuations and, therefore, less scope for common values. Bidders may still have private information of an idiosyncratic nature, e.g., regarding their own sales and inventories of end products, contracts for future sales, or inventories of uncut timber from private timber sales. This has led several authors (e.g., Baldwin, Marshall, and Richard (1997), Haile (2001), Haile and Tamer (2003)) to assume private values models for scaled sales.[30] However, this is not without controversy. Baldwin (1995) and Athey and Levin (2001), for example, argue for a common values model even for scaled sales. Athey, Levin, and Seira (2004) assume private values for both scaled and lumpsum sales.[31]

The auctions in our samples took place between 1982 and 1990 in Forest Service regions 1 and 5. Region 1 covers Montana, eastern Washington, Northern Idaho, North Dakota, and northwestern South Dakota. The Region 5 data consist of sales in California. The restriction to sales after 1981 is made due to policy changes in 1981 that (among other things) reduced the significance of subcontracting as a factor affecting bidder valuations, because resale opportunities can alter bidding in ways that confound the empirical implications of the winner's curse (cf. Bikhchandani and Huang (1989), Haile (1999), and Haile (2001)). For the same reason, we restrict attention to sales with no more than 12 months between the auction and the harvest deadline.[32] For consistency, we consider only sales in which the Forest Service provided *ex ante* estimates of the tract values (based on the cruise) using the predominant method of this time period, known as the "residual value method" (cf.

[30]Other studies assuming private values at timber auctions (USFS and others) include Cummins (1994), Elyakime, Laffont, Loisel, and Vuong (1994), Hansen (1985), Hansen (1986), Johnson (1979), Paarsch (1991), and Paarsch (1997).

[31]Other studies assuming common values models for timber auctions include Chatterjee and Harrison (1988), Lederer (1994), and Leffler, Rucker, and Munn (2000).

[32]This is the same rule used by Haile and Tamer (2003) and the opposite of that used by Haile (2001) to focus on sales with significant resale opportunities.

Baldwin, Marshall, and Richard (1997)). We exclude salvage sales, sales set aside for small firms, and sales of contracts requiring the winner to construct roads.

Table 3 describes the resulting sample sizes for auctions with each number of bidders $n = 2, 3, \ldots, 12$. There are fairly few auctions with more than four bidders, particularly in the sample of lumpsum sales. However, the unit of observation, both for estimation of the pseudo-values and for estimation of the distribution of pseudo-values, is a bid. Our data set contains 75 or more bids for auctions of up to seven bidders in both samples.

Our data include all bids[33] for each auction, as well as a large number of auction-specific observables. These include the year of the sale, the appraised value of the tract, the acreage of the tract, the length (in months) of the contract term, the volume of timber sold by the USFS in the same region over the previous six months, and USFS estimates of the volume of timber on the tract, harvesting costs, costs of manufacturing end products, selling value of the end products, and an index of the concentration of the timber volume across species (cf. Haile (2001)). All dollar values are in constant 1983 dollars per thousand board-feet of timber. Table 4 provides summary statistics.

7.2 Results

to be completed

8 Final Remarks

Our tests are not without limitations that should be kept in mind when interpreting our empirical results and applying our tests elsewhere. While we have allowed a rich class of models in our underlying framework, we have maintained the assumption of equilibrium competitive bidding in a static game, ruling out collusion and dynamic factors that might influence bidding decisions. Our techniques for dealing with endogenous participation and auction heterogeneity have required additional assumptions and finite sample approximations. Finally, while our test for common values is consistent, in some applications the effects of the winner's curse in the USFS auctions may be sufficiently small that they are difficult to detect. In some cases a failure to reject the PV null may suggest only that any CV elements are fairly small relative to other sources of variation in the data.

[33] In practice separate prices are bid for each identified species on the tract. Following, e.g., Baldwin, Marshall, and Richard (1997), Haile (2001), and Haile and Tamer (2003), we consider only the total bid of each bidder, which is also the statistic used to determine the auction winner. See Athey and Levin (2001) for an analysis of the distribution of bids across species.

In addition to providing an approach for formal testing, comparing distributions of estimated conditional expected valuations as n varies provides one natural way for quantifying the *magnitude* of any deviation from a private values model. For example, our estimates can be used to describe how much bidders adjust their expectations of the value of winning in response to an exogenous increase in competition (on average, or at various quantiles, etc.). This can be done at any quantile of the distribution of signals, for example. Although this provides a one measure of the severity of the winner's curse, in some applications one would like to address questions like how far wrong a particular policy prescription would go if a private values model were incorrectly assumed, or how much revenue might be gained if the common value component could be eliminated, e.g., by public revelation of all shared determinants of bidders valuations. Answering such questions will generally require identification of the model. Thus far identification results for common values models are mostly negative except with strong functional form assumptions once the PV hypothesis is dropped (see, e.g., Li, Perrigne and Vuong (2000) and Athey and Haile (2002, 2006)). Indeed, the lack of nonparametric identification of CV models is one motivation for developing formal tests for common values.

Appendix

A Proof of Theorem 2

First note that Assumption 5 directly implies the following uniform rates of convergence for $\hat{G}_n(b;b)$ and $\hat{g}_n(b;b)$ (see Horowitz (1998) and Guerre, Perrigne, and Vuong (2000)).

$$\sup_{b\in\mathbb{R}}\left|\tilde{G}_n(b;b)\right| \equiv \sup_{b\in\mathbb{R}}\left|\hat{G}_n(b;b)-G_n(b;b)\right| = O_p\left(\sqrt{\frac{\log T}{Th}}\right)+O\left(h^R\right)$$
$$\sup_{b\in R}\left|\tilde{g}_n(b;b)\right| \equiv \sup_{b\in R}\left|\hat{g}_n(b;b)-g_n(b;b)\right| = O_p\left(\sqrt{\frac{\log T}{Th^2}}\right)+O\left(h^R\right).$$

Since part (i) is an immediate consequence of part (ii), we proceed to prove part (ii) directly. Letting $\xi(b;n)=v\left(s^{-1}(b),s^{-1}(b),n\right)$, we can decompose the left side of part (ii) as

$$\begin{aligned}
&\sqrt{T_n h}\left(\hat{\mu}_{n,\tau}-E\left[\xi(b;n)\mathbf{1}\left(b_{\tau,n}\le b\le b_{1-\tau,n}\right)\right]\right)\\
=&\sqrt{T_n h}\left(\frac{1}{T_n n}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathcal{I}_t^n\left(b_{it}+\frac{\hat{G}_n(b_{it};b_{it})}{\hat{g}_n(b_{it};b_{it})}\right)\mathbf{1}\left(\hat{b}_{\tau,n}\le b_{it}\le\hat{b}_{1-\tau,n}\right)-E\left[\xi(b;n)\mathbf{1}\left(b_{\tau,n}\le b\le b_{1-\tau,n}\right)\right]\right)\\
=&\hat{\mu}^1_{n,\tau}+\hat{\mu}^2_{n,\tau}+\hat{\mu}^3_{n,\tau}+\hat{\mu}^4_{n,\tau}
\end{aligned}$$

where we have again let $\mathcal{I}_t^n=\mathbf{1}(n_t=n)$, and

$$\begin{aligned}
\hat{\mu}^1_{n,\tau}=&\sqrt{T_n h}\frac{1}{T_n n}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\frac{\hat{G}_n(b_{it};b_{it})}{\hat{g}_n(b_{it};b_{it})}-\frac{G_n(b_{it};b_{it})}{g_n(b_{it};b_{it})}\right)\left(\mathbf{1}\left(\hat{b}_{\tau,n}\le b_{it}\le\hat{b}_{1-\tau,n}\right)-\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\right)\mathcal{I}_t^n\\
\hat{\mu}^2_{n,\tau}=&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\left[\frac{\hat{G}_n(b_{it};b_{it})}{\hat{g}_n(b_{it};b_{it})}-\frac{G_n(b_{it};b_{it})}{g_n(b_{it};b_{it})}\right]\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\mathcal{I}_t^n\\
\hat{\mu}^3_{n,\tau}=&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(b_{it}+\frac{G_n(b_{it};b_{it})}{g_n(b_{it};b_{it})}\right)\left(\mathbf{1}\left(\hat{b}_{\tau,n}\le b_{it}\le\hat{b}_{1-\tau,n}\right)-\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\right)\mathcal{I}_t^n\\
=&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\xi(b_{it};n)\left(\mathbf{1}\left(\hat{b}_{\tau,n}\le b_{it}\le\hat{b}_{1-\tau,n}\right)-\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\right)\mathcal{I}_t^n\\
\hat{\mu}^4_{n,\tau}=&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\left(b_{it}+\frac{G_n(b_{it};b_{it})}{g_n(b_{it};b_{it})}\right)\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)-E\left[\xi(b;n)\mathbf{1}\left(b_{\tau,n}\le b\le b_{1-\tau,n}\right)\right]\right)\mathcal{I}_t^n\\
=&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\xi(b_{it};n)\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)-E\left[\xi(b;n)\mathbf{1}\left(b_{\tau,n}\le b\le b_{1-\tau,n}\right)\right]\right)\mathcal{I}_t^n.
\end{aligned}$$

We consider the properties of each of these terms in turn. For $\hat{\mu}^4_{n,\tau}$, the law of large numbers gives

$$\hat{\mu}^4_{n,\tau}=\sqrt{h}\ O_p(1)=o_p(1).$$

The function in the summand of $\hat{\mu}^3_{n,\tau}$ satisfies stochastic equicontinuity conditions (a type I function of Andrews (1994)). Hence using the parametric convergence rates of $\hat{b}_\tau$ and $\hat{b}_{1-\tau}$,

$$\begin{aligned}\hat{\mu}^3_{n,\tau} =&\sqrt{T_n h}\left(E_b \xi\left(b;n\right)\mathbf{1}\left(\hat{b}_{\tau,n}\leq b\leq \hat{b}_{1-\tau,n}\right)-E_b\left[\xi\left(b;n\right)\mathbf{1}\left(b_{\tau,n}\leq b\leq b_{1-\tau,n}\right)\right]\right)+o_p\left(1\right)\\ =&C\sqrt{T_n h}\left(O\left(\hat{b}_{\tau,n}-b_{\tau,n}\right)+O\left(\hat{b}_{1-\tau,n}-b_{1-\tau,n}\right)\right)+o_p\left(1\right)=\sqrt{T_n h}O_p\left(\frac{1}{\sqrt{T_n}}\right)+o_p\left(1\right)=o_p\left(1\right).\end{aligned}$$

Similarly, the function in the summand of $\hat{\mu}^1_{n,\tau}$ also satisfies stochastic equicontinuity conditions (product of type I and type III functions in Andrews (1994)), and hence

$$\begin{aligned}\hat{\mu}^1_{n,\tau} =&\sqrt{T_n h}E_b\left(\frac{\hat{G}_n\left(b;b\right)}{\hat{g}_n\left(b;b\right)}-\frac{G_n\left(b;b\right)}{g_n\left(b;b\right)}\right)\left(\mathbf{1}\left(\hat{b}_\tau\leq b\leq\hat{b}_{1-\tau}\right)-\mathbf{1}\left(b_{\tau,n}\leq b\leq b_{1-\tau,n}\right)\right)+o_p\left(1\right)\\ =&O_p\left(\sup_{b\in\left[b_\tau-\tilde{\epsilon},b_{1-\tau}+\tilde{\epsilon}\right]}\left|\frac{\hat{G}_n\left(b;b\right)}{\hat{g}_n\left(b;b\right)}-\frac{G_n\left(b;b\right)}{g_n\left(b;b\right)}\right|\right)\sqrt{T_n h}\left(O\left(\hat{b}_{\tau,n}-b_{\tau,n}\right)+O\left(\hat{b}_{1-\tau,n}-b_{1-\tau,n}\right)\right)+o_p\left(1\right)\\ =&o_p\left(1\right)\sqrt{T_n h}O_p\left(\frac{1}{\sqrt{T_n}}\right)+o_p\left(1\right)=o_p\left(1\right).\end{aligned}$$

Combining the above results, we have thus far shown that

$$\sqrt{T_n h}\left(\hat{\mu}_{n,\tau}-E\left[\xi\left(b;n\right)\mathbf{1}\left(b_{\tau,n}\leq b\leq b_{1-\tau,n}\right)\right]\right)=\hat{\mu}^2_{n,\tau}+o_p\left(1\right).$$

The term $\hat{\mu}^2_{n,\tau}$ can be further decomposed using a second order Taylor expansion:

$$\hat{\mu}^2_{n,\tau}=\hat{\mu}^5_{n,\tau}+\hat{\mu}^6_{n,\tau}+\hat{\mu}^7_{n,\tau}$$

where

$$\begin{aligned}\hat{\mu}^5_{n,\tau} =&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{1}{g_n\left(b_{it};b_{it}\right)}\left(\hat{G}_n\left(b_{it};b_{it}\right)-G_n\left(b_{it};b_{it}\right)\right)\mathbf{1}\left(b_{\tau,n}\leq b_{it}\leq b_{1-\tau,n}\right)\mathcal{I}^n_t\\ \hat{\mu}^6_{n,\tau} =&-\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{G_n\left(b_{it};b_{it}\right)}{g_n\left(b_{it};b_{it}\right)^2}\left(\hat{g}_n\left(b_{it};b_{it}\right)-g_n\left(b_{it};b_{it}\right)\right)\mathbf{1}\left(b_{\tau,n}\leq b_{it}\leq b_{1-\tau,n}\right)\mathcal{I}^n_t\\ \hat{\mu}^7_{n,\tau} =&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\underline{\mathrm{h}}^1_n\left(b_{it}\right)\left(\hat{G}_n\left(b_{it};b_{it}\right)-G_n\left(b_{it};b_{it}\right)\right)^2\mathbf{1}\left(b_{\tau,n}\leq b_{it}\leq b_{1-\tau,n}\right)\mathcal{I}^n_t\\ &+\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\underline{\mathrm{h}}^2_n\left(b_{it}\right)\left(\hat{g}_n\left(b_{it};b_{it}\right)-g_n\left(b_{it};b_{it}\right)\right)^2\mathbf{1}\left(b_{\tau,n}\leq b_{it}\leq b_{1-\tau,n}\right)\mathcal{I}^n_t.\end{aligned}$$

Here the functions $\underline{\mathrm{h}}^1_n\left(\cdot\right)$ and $\underline{\mathrm{h}}^2_n\left(\cdot\right)$ denote the second derivatives with respect to $G_n\left(\cdot\right)$ and $g_n\left(\cdot\right)$ evaluated at some mean values between $\hat{G}_n\left(\cdot\right)$ and $G_n\left(\cdot\right)$ and between $\hat{g}_n\left(\cdot\right)$ and $g_n\left(\cdot\right)$. We first bound $\hat{\mu}^7_{n,\tau}$ using the uniform convergence rates of $\hat{G}_n\left(\cdot\right)$ and $\hat{g}_n\left(\cdot\right)$:

$$\begin{aligned}\left|\mu^7_{n,\tau}\right| \leq&C\sqrt{T_n h}\left(O_p\left(\frac{\log T}{T_n h}+h^{2R}\right)+O_p\left(\frac{\log T}{T_n h^2}+h^{2R}\right)\right)\\ =&O_p\left(\frac{\log T}{\sqrt{T_n h}}+\frac{\log T}{\sqrt{T_n h^3}}+\sqrt{T_n h^{1+4R}}\right)=o_p\left(1\right).\end{aligned}$$

Now consider

$$
\begin{aligned}
\hat{\mu}^6_{n,\tau} &= -\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{G_n\left(b_{it};b_{it}\right)}{g_n\left(b_{it};b_{it}\right)^2}\left(\hat{g}_n\left(b_{it};b_{it}\right)-E\left[\hat{g}_n\left(b_{it};b_{it}\right)\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\right]\right)\mathcal{I}^n_t \\
&\quad -\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{G_n\left(b_{it};b_{it}\right)}{g_n\left(b_{it};b_{it}\right)^2}\left(E\left[\hat{g}_n\left(b_{it};b_{it}\right)\right]-g_n\left(b_{it};b_{it}\right)\right)\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\mathcal{I}^n_t \\
&= -\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{G_n\left(b_{it};b_{it}\right)}{g_n\left(b_{it};b_{it}\right)^2}\left(\hat{g}_n\left(b_{it};b_{it}\right)-E\left[\hat{g}_n\left(b_{it};b_{it}\right)\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\right]\right)\mathcal{I}^n_t+o_p\left(1\right) \\
&\equiv \hat{\mu}^8_{n,\tau}+o_p\left(1\right)
\end{aligned}
$$

because by assumption the bias in the second term on the right-hand side of the first line is of order

$$\sqrt{T_n h}O\left(h^R\right)=O\left(\sqrt{T_n h^{1+2R}}\right)=o\left(1\right).$$

Next we show that

$$\hat{\mu}^8_{n,\tau}\xrightarrow{d}N\left(0,\Omega=\left[\int\left(\int K\left(\epsilon'\right)K\left(\epsilon'-\epsilon\right)d\epsilon'\right)^2 d\epsilon\right]\left[\frac{1}{n}\int_{F_b^{-1}(\tau)}^{F_b^{-1}(1-\tau)}\frac{G_n\left(b;b\right)^2}{g_n\left(b;b\right)^3}g_n\left(b\right)^2 db\right]\right).$$

This follows from a limit variance calculation for U-statistics. Letting $\mathbf{b}_t$ represent the vector of all bids at auction t, we can write

$$\hat{\mu}^8_{n,\tau}=\sqrt{T_n h}\frac{1}{n^2T_n^2}\sum_{t=1}^{T}\sum_{s=1}^{T}m\left(\mathbf{b}_t,\mathbf{b}_s\right)\mathcal{I}^n_t\mathcal{I}^n_s$$

where

$$
\begin{aligned}
m\left(\mathbf{b}_t,\mathbf{b}_s\right)=&\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{G_n\left(b_{it};b_{it}\right)}{g_n^2\left(b_{it};b_{it}\right)}\left[\frac{1}{h^2}K\left(\frac{b_{sj}-b_{ti}}{h}\right)K\left(\frac{b^*_{sj}-b_{ti}}{h}\right)\right.\\
&\left.-E\frac{1}{h^2}K\left(\frac{b_{sj}-b_{ti}}{h}\right)K\left(\frac{b^*_{sj}-b_{ti}}{h}\right)\right]\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right).
\end{aligned}
$$

Using lemma 8.4 of Newey and McFadden (1994), we can verify that

$$
\begin{aligned}
\sqrt{T_n h}\frac{E|m\left(\mathbf{b}_t,\mathbf{b}_t\right)|}{T_n}&=O_p\left(\sqrt{T_n h}\frac{1}{T_n h}\right)=O_p\left(\frac{1}{\sqrt{T_n h}}\right)=o_p\left(1\right),\quad\text{and}\\
\sqrt{T_n h}\frac{\sqrt{Em\left(\mathbf{b}_t,\mathbf{b}_s\right)^2}}{T_n}&=O_p\left(\sqrt{T_n h}\frac{1}{T_n\sqrt{h^3}}\right)=O_p\left(\frac{1}{\sqrt{T_n h^2}}\right)=o_p\left(1\right).
\end{aligned}
$$

It then follows from Lemma 8.4 of Newey and McFadden (1994) that

$$\hat{\mu}^8_{n,\tau}=\sqrt{T_n h}\frac{1}{n^2T_n}\left[\sum_{t=1}^{T}E\left(m\left(\mathbf{b}_t,\mathbf{b}_s\right)|\mathbf{b}_t\right)\mathcal{I}^n_t+\sum_{s=1}^{T}E\left(m\left(\mathbf{b}_t,\mathbf{b}_s\right)|\mathbf{b}_s\right)\mathcal{I}^n_s\right]+o_p\left(1\right).$$

The first term is asymptotically negligible, because

$$
\begin{aligned}
&\sqrt{T_n h}\frac{1}{n^2 T_n}\sum_{t=1}^{T} E\left(m\left(\mathbf{b}_t,\mathbf{b}_s\right)|\mathbf{b}_t\right)\mathcal{I}_t^n \\
=&\sqrt{T_n h}\frac{1}{nT_n}\sum_{t=1}^{T}\left[g_n\left(b_{it};b_{it}\right)\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)-E\left[g_n\left(b_{it};b_{it}\right)\mathbf{1}\left(b_{\tau,n}\le b_{it}\le b_{1-\tau,n}\right)\right]\mathcal{I}_t^n+O\left(h^R\right)\right] \\
=&\sqrt{T_n h}O_p\left(\frac{1}{\sqrt{T_n}}\right)+O\left(\sqrt{T_n h^{1+2R}}\right)=o_p\left(1\right).
\end{aligned}
$$

It remains only to verify by straightforward though somewhat tedious calculation that

$$
\begin{aligned}
&Var\left(\sqrt{T_n h}\frac{1}{n^2 T_n}\sum_{s=1}^{T}E\left(m\left(\mathbf{b}_t,\mathbf{b}_s\right)|\mathbf{b}_s\right)I_s^n\right) \\
=&hVar\left(\frac{1}{n}\sum_{j=1}^{n}\int_{b_\tau}^{b_{1-\tau}}\frac{G_n\left(b;b\right)}{g_n^2\left(b;b\right)}\frac{1}{h^2}K\left(\frac{b_{sj}-b}{h}\right)K\left(\frac{b_{sj}^*-b}{h}\right)g_n\left(b\right)db\right) \\
=&h\frac{1}{n}Var\left(\int_{b_\tau}^{b_{1-\tau}}\frac{G_n\left(b;b\right)}{g_n^2\left(b;b\right)}\frac{1}{h^2}K\left(\frac{b_{sj}-b}{h}\right)K\left(\frac{b_{sj}^*-b}{h}\right)g_n\left(b\right)db\right)+o\left(1\right) \\
=&h\frac{1}{n}E\left(\int_{b_\tau}^{b_{1-\tau}}\frac{G_n\left(b;b\right)}{g_n^2\left(b;b\right)}\frac{1}{h^2}K\left(\frac{b_{sj}-b}{h}\right)K\left(\frac{b_{sj}^*-b}{h}\right)g_n\left(b\right)db\right)^2+o\left(1\right) \\
\longrightarrow\ &\Omega\equiv\left[\int\left(\int K\left(\epsilon'\right)K\left(\epsilon'-\epsilon\right)d\epsilon'\right)^2 d\epsilon\right]\left[\frac{1}{n}\int_{G_n^{-1}(\tau)}^{G_n^{-1}(1-\tau)}\frac{G_n^2\left(b;b\right)}{g_n^3\left(b;b\right)}g_n^2\left(b\right)db\right].
\end{aligned}
$$

Finally, we note that if we apply the calculations performed for $\hat{\mu}_{n,\tau}^6$ to $\hat{\mu}_{n,\tau}^5$, we see that

$$
E\left(\hat{\mu}_{n,\tau}^5\right)=o\left(1\right)\qquad\text{and } Var\left(\hat{\mu}_{n,\tau}^5\right)=o\left(1\right)
$$

which then implies that $\hat{\mu}_{n,\tau}^5\xrightarrow{p}0$. The proof is now completed by putting these terms together.□

B Proof of Theorem 3

1. This is a standard result on the $\sqrt{T_n}$-convergence of sample to population quantiles (cf. van der Vaart (1999), Corollary 21.5).
2. For simplicity we introduce the notation $\mathcal{I}_t^n=\mathbf{1}\left(n_t=n\right)$, $G_n\equiv G_n\left(b;b\right)$, $g_n\equiv g_n\left(b;b\right)$, $\hat{G}_n\equiv\hat{G}_n\left(b;b\right)=\frac{1}{nT_nh}\sum_{t=1}^{T}\mathcal{I}_t^n\sum_{i=1}^{n}\mathbf{1}\left(b_{it}^*<b\right)K\left(\frac{b-b_{it}}{h}\right)$ and

$$
\hat{g}_n\equiv\hat{g}_n\left(b;b\right)=\frac{1}{nT_nh^2}\sum_{t=1}^{T}\mathcal{I}_t^n\sum_{i=1}^{n}K\left(\frac{b-b_{it}}{h}\right)K\left(\frac{b-b_{it}^*}{h}\right).
$$

Then we can use a first-order Taylor expansion to write

$$\hat{v}\left(s^{-1}(b), s^{-1}(b), n\right) - v\left(s^{-1}(b), s^{-1}(b), n\right) = \frac{\hat{G}_n}{\hat{g}_n} - \frac{G_n}{g_n}$$

$$= \frac{\hat{G}_n - G_n}{g_n} - \frac{G_n}{g_n^2}(\hat{g}_n - g_n) + o\left(\hat{G}_n - G_n\right) + o(\hat{g}_n - g_n)$$

$$= \frac{\hat{G}_n - E\hat{G}_n}{g_n} + \frac{E\hat{G}_n - G_n}{g_n} - \frac{G_n}{g_n^2}(\hat{g}_n - E\hat{g}_n) - \frac{G_n}{g_n^2}(E\hat{g}_n - g_n) + o\left(\hat{G}_n - G_n\right) + o(\hat{g}_n - g_n).$$

Standard bias calculations for kernel estimators yield, by Assumption 5,

$$|E\hat{G}_n - G_n| \le |\int (G_n(b; b + h\epsilon) - G_n(b; b)) K(\epsilon)\, d\epsilon| \le Ch^R \int |\epsilon|^R K(\epsilon)\, d\epsilon = o\left(\frac{1}{\sqrt{Th^2}}\right)$$

and

$$|E\hat{g}_n - g_n| \le \left| \int\int (g_n(b + h\epsilon; b + h\epsilon') - g_n(b; b)) K(\epsilon) K(\epsilon')\, d\epsilon\, d\epsilon' \right| \le C' h^R = o\left(\frac{1}{\sqrt{Th^2}}\right)$$

where C and C' are constants. Next it will be shown that

$$\sqrt{T_n h^2}(\hat{g}_n - E\hat{g}_n) \xrightarrow{d} N\left(0, \frac{1}{n}\left(\int\int K(e)^2 K(e')^2\, de\, de'\right) g_n(b; b)\right).$$

For this purpose it suffices to show that

$$\lim_{T_n \to \infty} Var\left(\sqrt{T_n h^2}(\hat{g}_n(b; b) - E\hat{g}_n(b; b))\right) = \frac{1}{n}\left(\int\int K(e)^2 K(e')^2\, de\, de'\right) g_n(b; b).$$

This is verified by the following calculation:

$$\begin{aligned}
& Var\left(\frac{1}{\sqrt{T_n h^2} \cdot n} \sum_{t=1}^{T} \sum_{i=1}^{n} \left[K\left(\frac{b_{it} - b}{h}\right) K\left(\frac{b^*_{it} - b}{h}\right)\right] \mathcal{I}^n_t\right) \\
=& T_n \left(\frac{1}{T_n n^2 h^2} Var\left(\sum_{i=1}^{n} \left[K\left(\frac{b_{it} - b}{h}\right) K\left(\frac{b^*_{it} - b}{h}\right)\right]\right)\right) \\
=& \frac{1}{nh^2}\left\{Var\left[K\left(\frac{b_{it} - b}{h}\right) K\left(\frac{b^*_{it} - b}{h}\right)\right]\right. \\
& \left. + (n-1)\, Cov\left[K\left(\frac{b_{it} - b}{h}\right) K\left(\frac{b^*_{it} - b}{h}\right), K\left(\frac{b_{jt} - b}{h}\right) K\left(\frac{b^*_{jt} - b}{h}\right)\right]_{j \ne i}.\right\}
\end{aligned}$$

It is a standard result that

$$E\left(K\left(\frac{b_{it} - b}{h}\right) K\left(\frac{b^*_{it} - b}{h}\right)\right) = O\left(h^2\right)$$

and it can be verified that for $j \ne i$

$$E\left[K\left(\frac{b_{it} - b}{h}\right) K\left(\frac{b^*_{it} - b}{h}\right) K\left(\frac{b_{jt} - b}{h}\right) K\left(\frac{b^*_{jt} - b}{h}\right)\right] = O\left(h^4\right).$$

Therefore we can write

$$Var\left(\sqrt{T_n h^2}\left(\hat{g}_n\left(b;b\right)-E\hat{g}_n\left(b;b\right)\right)\right)=\frac{1}{n\,h^2}E\left[K\left(\frac{b_{it}-b}{h}\right)^2K\left(\frac{b_{it}^*-b}{h}\right)^2\right]+O\left(h^4\right)$$
$$=\frac{1}{n}\int\int\frac{1}{h^2}K\left(\frac{\epsilon-b}{h}\right)^2K\left(\frac{\epsilon'-b}{h}\right)^2g_n\left(\epsilon,\epsilon'\right)d\epsilon d\epsilon'+O\left(h^4\right)$$
$$=\frac{1}{n}\left(\int\int K\left(e\right)^2K\left(e'\right)^2de\,de'\right)g_n\left(b;b\right)+o\left(1\right)$$

where the last equality uses the substitutions $e=\left(\epsilon-b\right)/h$ and $e'=\left(\epsilon'-b\right)/h$. Finally the same type of variance calculation shows that

$$Var\left(\sqrt{T_n h^2}\left(\hat{G}_n-E\hat{G}_n\right)\right)\longrightarrow 0.$$

Hence the proof for part 2 is complete.

3. Because the sample quantiles of the bid distribution converge at rate $\sqrt{T_n}$ to the population quantile, which is faster than the convergence rate for the estimated expected valuations, for large T_n the sampling error in the τth quantile of the bid distribution does not affect the large sample properties of the estimated quantiles of the distribution of estimated conditional expected valuations. Hence, for each $\tau\in\{\tau_1,\ldots,\tau_l\}$

$$\left(\hat{v}\left(s_n^{-1}\left(\hat{b}_{\tau_l,n}\right),s_n^{-1}\left(\hat{b}_{\tau_l,n}\right),n\right)-\hat{v}\left(x_\tau,x_\tau,n\right)\right)=O_p\left(\frac{1}{\sqrt{T_n}}\right)=o_p\left(\frac{1}{T_nh^2}\right).\tag{31}$$

This implies that the limiting distribution of the vector with elements

$$\sqrt{T_nh^2}\left(\hat{\xi}\left(\hat{b}_{\tau_l,n};n\right)-v\left(F_x^{-1}\left(\tau\right),F_x^{-1}\left(\tau\right),n\right)\right)\qquad\tau=\{\tau_1,\ldots,\tau_L\}$$

is the same as that of the vector with elements

$$\sqrt{Th^2}\left(\hat{\xi}\left(s_n\left(x_\tau\right);n\right)-v\left(x_\tau,x_\tau,n\right)\right)\qquad\tau=\{\tau_1,\ldots,\tau_L\}.$$

In part 2 we showed that each element of this vector is asymptotically normal with limit variance given by the corresponding diagonal element of Ω. It remains to show that the off-diagonal elements are 0. For this purpose it suffices to show, using the standard result that kernel estimates at two distinct points (here, two quantiles $b_\tau\equiv s(x_\tau)$ and $b_{\tau'}\equiv s(x_{\tau'})$) are asymptotically independent, i.e., that

$$\lim_{T_n\to\infty}Cov\left(\sqrt{T_nh^2}\left(\hat{\xi}\left(b_\tau;n\right)-v\left(x_\tau,x_\tau,n\right)\right),\sqrt{T_nh^2}\left(\hat{\xi}\left(b_{\tau'};n\right)-v\left(x_{\tau'},x_{\tau'},n\right)\right)\right)=0.$$

Using the bias calculation and convergence rates derived in part 2, it suffices for this purpose to show that

$$\lim_{T_n\to\infty}Cov\left(\sqrt{T_nh^2}\left(\hat{g}_n\left(b_\tau;b_\tau\right)-Eg_n\left(b_\tau;b_\tau\right)\right),\sqrt{T_nh^2}\left(\hat{g}_n\left(b_{\tau'};b_{\tau'}\right)-Eg_n\left(b_{\tau'};b_{\tau'}\right)\right)\right)=0$$

To show this, first observe that the left-hand side can be written

$$Cov\left[\frac{1}{\sqrt{T_nh^2}n}\sum_{t=1}^{T}\sum_{i=1}^{n}K\left(\frac{b_{it}-b_\tau}{h}\right)K\left(\frac{b_{it}^*-b_\tau}{h}\right)\mathcal{I}_t^n,\frac{1}{\sqrt{T_nh^2}n}\sum_{t=1}^{T}\sum_{i=1}^{n}K\left(\frac{b_{it}-b_{\tau'}}{h}\right)K\left(\frac{b_{it}^*-b_{\tau'}}{h}\right)\mathcal{I}_t^n\right]$$
$$=\frac{1}{n^2h^2}Cov\left[\sum_{i=1}^{n}K\left(\frac{b_{it}-b_\tau}{h}\right)K\left(\frac{b_{it}^*-b_\tau}{h}\right),\sum_{i=1}^{n}K\left(\frac{b_{it}-b_{\tau'}}{h}\right)K\left(\frac{b_{it}^*-b_{\tau'}}{h}\right)\right].$$

Using the fact that for each i

$$E\left[K\left(\frac{b_{it}-b_\tau}{h}\right)K\left(\frac{b^*_{it}-b_\tau}{h}\right)\right]=O\left(h^2\right)$$

and for each $i\neq j$

$$E\left[K\left(\frac{b_{it}-b_\tau}{h}\right)K\left(\frac{b^*_{it}-b_\tau}{h}\right)K\left(\frac{b_{jt}-b_{\tau'}}{h}\right)K\left(\frac{b^*_{jt}-b_{\tau'}}{h}\right)\right]=O\left(h^4\right)$$

we can further rewrite the covariance function as

$$\begin{aligned}
&\frac{1}{n^2h^2}\sum_{i=1}^n\sum_{j=1}^n E\left[K\left(\frac{b_{it}-b_\tau}{h}\right)K\left(\frac{b^*_{it}-b_\tau}{h}\right)K\left(\frac{b_{jt}-b_{\tau'}}{h}\right)K\left(\frac{b^*_{jt}-b_{\tau'}}{h}\right)\right]+O\left(h^2\right)\\
=&\frac{1}{n^2h^2}\sum_{i=1}^n E\left[K\left(\frac{b_{it}-b_\tau}{h}\right)K\left(\frac{b^*_{it}-b_\tau}{h}\right)K\left(\frac{b_{it}-b_{\tau'}}{h}\right)K\left(\frac{b^*_{it}-b_{\tau'}}{h}\right)\right]+O\left(h^2\right)\\
=&\frac{1}{n}\int\int K(e)\,K(e')\,K\left(e+\frac{b_\tau-b_{\tau'}}{h}\right)K\left(e'+\frac{b_\tau-b_{\tau'}}{h}\right)g_n\left(b_\tau+he,b_\tau+he'\right)de\,de'+O\left(h^2\right)\to 0.
\end{aligned}$$

C Proof of Lemma 3

Given any function $\zeta(\cdot)$ we can write

$$N=\phi(\mathbf{Y},\mathbf{Z},W)=\zeta\left(\mathbf{Y},\mathbf{Z}\right)+\Omega\left(\mathbf{Y},\mathbf{Z},W\right) \tag{32}$$

for some function $\Omega(\cdot)$. Because here $\zeta(\mathbf{y},\mathbf{z})$ is an integer by definition, strict monotonicity of $\phi(\cdot)$ in W along with Assumption 8 imply that, for any $(\mathbf{y},\mathbf{z})$, the random variable $\Omega(\mathbf{y},\mathbf{z},W)$ has support given by a set of K contiguous integers

$$\{\Omega(\mathbf{y},\mathbf{z},w_1),\ldots,\Omega(\mathbf{y},\mathbf{z},w_K)\}$$

with $\Pr(\Omega(\mathbf{y},\mathbf{z},w_k))=p_k\forall\mathbf{z},k$ by Assumption 6.

Case (i). Since $\mathrm{int}(\zeta(\mathbf{y},\mathbf{z})+E[\Omega(\mathbf{y},\mathbf{z},W)])=\zeta(\mathbf{y},\mathbf{z})$ by definition and $\zeta(\mathbf{y},\mathbf{z})$ is an integer, we must have $\mathrm{int}(E[\Omega(\mathbf{y},\mathbf{z},W)])=0$. For every value of $(\mathbf{y},\mathbf{z})$ there must then be some $\delta(\mathbf{y},\mathbf{z})\in\{1,\ldots,K\}$ such that $\Omega\left(\mathbf{y},\mathbf{z},w_{\delta(\mathbf{y},\mathbf{z})}\right)=0$, with

$$\Omega(\mathbf{y},\mathbf{z},w_k)=k-\delta(\mathbf{y},\mathbf{z})\qquad k=1,\ldots,K. \tag{33}$$

Since $\delta(\mathbf{y},\mathbf{z})$ is defined by $\mathrm{int}\left(\sum_{k=1}^K p_k\left(k-\delta(\mathbf{y},\mathbf{z})\right)\right)=0$, we must have $\delta(\mathbf{y},\mathbf{z})=\delta=\mathrm{int}\left(\sum_{k=1}^K p_k k\right)$ for all $(\mathbf{y},\mathbf{z})$, so that

$$\Omega(\mathbf{y},\mathbf{z},w_k)=k-\delta\qquad\forall(\mathbf{y},\mathbf{z}),k=1,\ldots,K. \tag{34}$$

Then by (32) and (34), we can normalize the unobservable W by letting

$$w_k=k-\delta\qquad k=1,\ldots,K \tag{35}$$

and writing

$$N=\zeta(\mathbf{Y},\mathbf{Z})+W.$$

Case (ii): The argument is similar. Since $\zeta(\mathbf{y},\mathbf{z})$ is the α-quantile of $N|(\mathbf{y},\mathbf{z})$, the α-quantile of $\Omega(\mathbf{y},\mathbf{z},W)$ must be zero. Since $\zeta(\mathbf{y},\mathbf{z})$ is integer valued, we must have $\Omega\left(\mathbf{y},\mathbf{z},w_{\delta(\mathbf{y},\mathbf{z})}\right)=0$ for some $\delta(\mathbf{y},\mathbf{z})\in\{1,\ldots,K\}$, where now, for all $(\mathbf{y},\mathbf{z})$, $\delta(\mathbf{y},\mathbf{z})=\delta=\max\left\{d:\sum_{k=1}^d p_k\leq\alpha\right\}\ \forall(\mathbf{y},\mathbf{z})$. Then, $\Omega(\mathbf{y},\mathbf{z},w_k)=k-\delta$ and, as above, we can normalize the unobservable W according to (35) and writing $N=\zeta(\mathbf{Y},\mathbf{Z})+W$. $\square$

D Participation Model: Identification and Estimation

Here we have no need to distinguish between $\mathbf{Y}$, and $\mathbf{Z}$, so we will use $\mathbf{Z}$ to denote $(\mathbf{Y}, \mathbf{Z})$. For clarity, it will also be useful to make explicit distinction between the latent random variables $(N, \mathbf{Z}, W)$ and their observable counterparts $(\tilde{N}, \tilde{\mathbf{Z}}, \tilde{W})$, which reflect truncation of N from below at $\underline{n}$ and from above at $\overline{n}$. For every observation, of course, $\tilde{n} = n$, $\tilde{\mathbf{z}} = \mathbf{z}$, and $\tilde{w} = w$; however, the observable random variables have a different probability law from those in the population, making it important to distinguish between them when discussing identification.[34] Let $\tilde{\mathcal{Z}}$ denote the support of $\tilde{\mathbf{Z}}$.

D.1 Identification

Let $P_{\mathbf{Z}}$ denote the probability measure on $\mathbf{Z}$.

Assumption 10 $\{w_1, \ldots, w_K\} \subseteq \{\bar{w}_1, \ldots, \bar{w}_m\} \subset \mathbb{N}$, *with* $K + 1 < 2(\overline{n} - \underline{n})$.

Here we are not assuming that the support of W is known, nor even that K is known—only that K isn't too large relative to the observed variation in n. Let $\bar{p}_1, \ldots, \bar{p}_m$ be the true probabilities on $\bar{w}_1, \ldots, \bar{w}_m$; note that some may be zero, while those that are nonzero are equal to p_j for some j. Let $\gamma = \min\{K - 1, \overline{n} - \underline{n} - 1\}$.

Assumption 11 *(i)* $\exists \mathbf{z}'$ *and* $\mathbf{z}'' \in \mathbb{R}^L$ *such that (a) for some* $\epsilon \geq 0$, $P_{\mathbf{Z}}(\mathcal{B}) > 0$ *for an* ϵ*-ball* $\mathcal{B} = \mathcal{B}(\mathbf{z}')$ *around* $\mathbf{z}'$ *and an* ϵ*-ball* $\mathcal{B} = \mathcal{B}(\mathbf{z}'')$ *around* $\mathbf{z}''$, *(b) max supp*$\{N|\mathbf{z} \in \mathcal{B}(\mathbf{z}')\} = \overline{n} - 1$, *and (c) min supp* $\{N|\mathbf{z} \in \mathcal{B}(\mathbf{z}'')\} = \underline{n} + 1$; *(ii) there exist no distinct* i *and* $i' \in \{1, \ldots, m\}$ *such that* $\frac{p_{i+1}}{p_i} = \frac{p_{i'+1}}{p_{i'}} > 0$.

Part (i) ensures identifiability of a positive measure of values of $\mathbf{Z}$ where at most one tail of the distribution of $\mathbf{W}$ is truncated, enabling observation of the other tail. Part (ii) rules out, e.g., a geometric distribution for W, but is satisfied when W has any unimodal probability mass function that is everywhere nonlinear (i.e., no w such that $\frac{\Pr(W=w+1)}{\Pr(W=w)} = \frac{\Pr(W=w+2)}{\Pr(W=w+1)} > 0$). This and Assumption 10 ensure that we can "paste together" the two tail distributions to form the full distribution of W.

Lemma 4 *Under Assumptions 10 and 11, observation of* $(\tilde{N}, \tilde{\mathbf{Z}})$ *identifies (i)* $\zeta(\cdot)$ *on* $\tilde{\mathcal{Z}}$ *and (ii)* $\bar{p}_1, \ldots, \bar{p}_m$.

Proof. From $\tilde{N}|\tilde{\mathbf{Z}}$ at $\tilde{\mathbf{Z}} = \mathbf{z}'$ we observe the strictly positive $\bar{p}_{j'-\gamma}, \bar{p}_{j'-\gamma+1}, \ldots, \bar{p}_{j'}$ up to a scaling factor equal to $\Pr(N \in \{\underline{n}, \ldots, \overline{n}\}|\mathbf{z}')$, where j' is such that $\bar{w}_{j'} = w_K$. Similarly, from $\tilde{N}|\tilde{\mathbf{Z}}$ at $\tilde{\mathbf{Z}} = \mathbf{z}''$ we observe $\bar{p}_{j''}, \bar{p}_{j''+1}, \ldots, \bar{p}_{j''+\gamma}$ up to a (different) scaling factor, where $\bar{w}_{j''} = w_1$. There will be overlap in the indices of these probabilities if

$$j' - \gamma \leq j'' + \gamma. \tag{36}$$

Since $j' - j'' = K - 1$, (36) is immediate if $\gamma = K - 1$. And if $\gamma = \overline{n} - \underline{n} - 1$, the assumption $K + 1 < 2(\overline{n} - \underline{n})$ implies (36). Hence, there is at least one point in common in the supports of $W|\mathbf{z}'$

[34] In principle, one could introduce such notation for all random variables in the model and carry this notation throughout. This would significantly complicate the exposition and, except in this section, suppressing this should not lead to confusion.

and $W|\mathbf{z}''$; further, since the ratios p_i/p_j are preserved by truncation, part (ii) of Assumption 11 ensures that the number of points of overlap is uniquely determined. This implies that the support of W and all ratios p_i/p_j are identified. Since $\sum_{i=j'}^{j''} p_i = 1$, all p_i are then identified. Knowledge of all p_i immediately implies that the distribution of $N|\mathbf{z}$ (which reveals $\zeta(\mathbf{z})$) can be recovered for all $\mathbf{z} \in \tilde{\mathcal{Z}}$ from the distribution of $\tilde{N}|\tilde{\mathbf{Z}}$. □

D.2 Estimation

Suppose $\zeta(\mathbf{z}) = \text{int}(\mathbf{z}\theta)$ for θ in some compact set Θ. No further assumption is made on W, which therefore has a multinomial distribution defined by probabilities $\bar{p}_1, \ldots, \bar{p}_m$ on $\bar{w}_1, \ldots, \bar{w}_m$. Imposing $\bar{p}_m = 1 - \sum_{k<m} \bar{p}_k$ and either the mean-zero or quantile-zero restriction on W, the likelihood function for the observations of $n, \mathbf{z}$ is given by

$$\mathcal{L}(\theta, p_1, \ldots, p_{m-1}) = \prod_{t=1}^{T} \frac{\sum_{k=1}^{m} p_k \ \mathbf{1}\{\text{int}(\mathbf{z}_t\theta) + w_k = n_t\}}{\sum_{k=1}^{m} p_k \ \mathbf{1}\{\text{int}(\mathbf{z}_t\theta) + w_k \in \{\underline{n}, \ldots, \overline{n}\}\}}$$

Note that although $\zeta(\mathbf{z})$ is nonparametrically identified under the assumptions described above, the parameter θ need not be point identified. However, set identification of θ is gauranteed by Lemma 4 and is sufficient for our purposes; in particular, any $\hat{\theta}$ maximizing the likelihood will lead to the same estimate of $\zeta(\mathbf{z})$. Convergence of this estimator is at rate $\sqrt{T}$.[35]

[35] The log-likelihood function need not be smooth or strictly concave. We use repeated application of the Nelder-Mead simplex algorithm, starting at different locations in the parameter space. Simulations confirm that the estimator performs well in practice.

References

ANDERSON, G. (1996): "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica*, 64, 1183–1194.

ANDREWS, D. (1994): "Empirical Process Methods in Econometrics," in *Handbook of Econometrics, Vol. 4*, ed. by R. Engle, and D. McFadden. North Holland.

——— (1998): "Hypothesis Testing With a Restricted Parameter Space," *Journal of Econometrics*, 86, 155–199.

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

ATHEY, S., AND P. HAILE (2006a): "Empirical Models of Auctions," in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, ed. by R. Blundell, W. Newey, and T. Persson, vol. 2. Cambridge University Press.

——— (2006b): "Nonparametric Approaches to Auctions," in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. Leamer, vol. 6. Elsevier.

ATHEY, S., AND J. LEVIN (2001): "Information and Competition in U.S. Forest Serve Timber Auctions," *Journal of Political Economy*, 109, 375–417.

ATHEY, S., J. LEVIN, AND E. SEIRA (2004): "Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions," Working Paper (Stanford University).

AYER, M., D. BRUNK, G. EWING, W. REID, AND E. SILVERMAN (1955): "An Empirical Distribution Function for Sampling with Incomplete Information," *Annals of Mathematical Statistics*, 26, 641–647.

BAJARI, P., AND A. HORTACSU (2003): "Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions," *RAND Journal of Economics*, 34, 329–355.

BALDWIN, L. (1995): "Risk Aversion in Forest Service Timber Auctions," working paper, RAND Corporation.

BALDWIN, L., R. MARSHALL, AND J.-F. RICHARD (1997): "Bidder Collusion at Forest Service Timber Sales," *Journal of Political Economy*, 105, 657–699.

BARRETT, G., AND S. DONALD (2003): "Consistent Tests for Stochastic Dominance," *Econometrica*, 71, 71–104.

BARTHOLOMEW, D. (1959): "A Test of Homogeneity for Ordered Alternatives," *Biometrika*, 46, 36–48.

BIKHCHANDANI, S., P. HAILE, AND J. RILEY (2002): "Symmetric Separating Equilibria in English Auctions," *Games and Economic Behavior*, 38, 19–27.

BIKHCHANDANI, S., AND C. HUANG (1989): "Auctions with Resale Markets: An Exploratory Model of Treasury Bill Markets," *Review of Financial Studies*, 2, 311–339.

Campo, S., I. M. Perrigne, and Q. Vuong (2003): "Asymmetry in First-Price Auctions with Affiliated Private Values," *Journal of Applied Econometrics*, 18, 197–207.

Cantillon, E., and M. Pesendorfer (2003): "Combination Bidding in Multi-Unit Auctions," Working Paper (London School of Economics and Political Science).

Chatterjee, K., and T. Harrison (1988): "The Value of Information in Competitive Bidding," *European Journal of Operational Research*, 36, 322–333.

Cummins, J. (1994): "Investment Under Uncertainty: Estimates from Panel Data on Pacific Northwest Forest Products Firms," working paper, Columbia University.

Davidson, R., and J.-Y. Duclos (2000): "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 68, 1435–1464.

Elyakime, B., J. Laffont, P. Loisel, and Q. Vuong (1994): "First-Price Sealed-Bid Auctions with Secret Reserve Prices," *Annales d'Economie et Statistiques*, 34, 115–141.

Gilley, O., and G. Karels (1981): "The Competitive Effect in Bonus Bidding: New Evidence," *Bell Journal of Economics*, 12, 637–648.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525–74.

Haile, P. (1999): "Auctions with Resale," mimeo, University of Wisconsin-Madison.

——— (2001): "Auctions with Resale Markets: An Application to U.S. Forest Service Timber Sales," *American Economic Review*, 92, 399–427.

Haile, P., H. Hong, and M. Shum (2003): "Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions," NBER Working Paper #10105.

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

Hansen, R. (1985): "Empirical Testing of Auction Theory," *American Economic Review, Papers and Proceedings*, 75, 156–159.

——— (1986): "Sealed-Bid Versus Open Auctions: The Evidence," *Economic Inquiry*, 24, 125–143.

Hendricks, K., J. Pinkse, and R. Porter (2003): "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common-Value Auctions," *Review of Economic Studies*, 70, 115–146.

Hong, H., and M. Shum (2002): "Increasing Competition and the Winner's Curse: Evidence from Procurement," *Review of Economic Studies*, 69, 871–898.

Horowitz, J. (1998): "Bootstrap Methods for Median Regression Models," *Econometrica*, 66, 1327–1352.

Imbens, G. W., and W. K. Newey (2002): "Identification and Estimation of Triangular Simultaneous Equations," Working Paper (M.I.T.).

Jofre-Bonet, M., and M. Pesendorfer (2003): “Bidding Behavior in a Repeated Procurement Auction,” *Econometrica*, 71, 1443–1490.

Johnson, R. (1979): “Oral Auction Versus Sealed Bids: An Empirical Investigation,” *Natural Resources Journal*, 19, 315–335.

Klemperer, P. (1999): “Auction Theory: A Guide to the Literature,” *Journal of Economic Surveys*, 13, 227–286.

Krasnokutskaya, E. (2003): “Identification and Estimation of Auction Models under Unobserved Auction Heterogeneity,” working paper, Yale University.

Krasnokutskaya, E., and K. Seim (2004): “Bid Preference Programs and Participation in Highway Procurement,” Working Paper (University of Pennsylvania).

Krishna, V. (2002): *Auction Theory*. Academic Press.

Laffont, J. J., and Q. Vuong (1993): “Structural Econometric Analysis of Descending Auctions,” *European Economic Review*, 37, 329–341.

——— (1996): “Structural Analysis of Auction Data,” *American Economic Review, Papers and Proceedings*, 86, 414–420.

Lederer, P. (1994): “Predicting the Winner’s Curse,” *Decision Sciences*, 25, 79–101.

Leffler, K., R. Rucker, and I. Munn (2000): “Transaction Costs and the Collection of Information: Presale Measurement on Private Timber Sales,” *Journal of Law Economics and Organizations*, 16, 166–188.

Li, T., I. Perrigne, and Q. Vuong (2000): “Conditionally Independent Private Information in OCS Wildcat Auctions,” *Journal of Econometrics*, 98, 129–161.

——— (2002): “Structural Estimation of the Affiliated Private Value Auction Model,” *RAND Journal of Economics*, 33, 171–193.

Li, T., and X. Zheng (2005): “Procurement Auctions with Entry and an Uncertain Number of Actual Bidders: Theory, Structural Inference, and an Application,” Working Paper (Indiana University).

Linton, O., E. Massoumi, and Y. Whang (2002): “Consistent Testing for Stochastic Dominance: A Subsampling Approach,” mimeo., LSE.

Matthews, S. (1984): “Information Acquisition in Discriminatory Auctions,” in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom. North-Holland.

McFadden, D. (1989): “Testing for Stochastic Dominance,” in *Studies in the Economics of Uncertainty*, ed. by T. B. Fomby, and T. K. Seo. Springer-Verlag.

Milgrom, P. (1981): “Good News and Bad News: Representation Theorems and Applications,” *The Bell Journal of Economics*, 13, 380–391.

Milgrom, P., and R. Weber (1982): “A Theory of Auctions and Competitive Bidding,” *Econometrica*, 50, 1089–1122.

National Resources Management Corporation (1997): "A Nationwide Study Comparing Tree Measurement and Scaled Sale Methods for Selling United States Forest Service Timber," Report to the U.S. Forest Service, Department of Agriculture.

Newey, W., and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics, Vol. 4*, ed. by R. Engle, and D. McFadden. North Holland.

Newey, W. K., J. L. Powell, and F. Vella (1999): "Nonparametric Estimation of Triangular Simultaneous Equation Models," *Econometrica*, 67, 565–603.

Olley, S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Paarsch, H. (1991): "Empirical Models of Auctions and an Application to British Columbian Timber," University of Western Ontario, Department of Economics Technical Report 91-19.

——— (1997): "Deriving an Estimate of the Optimal Reserve Price: An Application to British Columbian Timber Sales," *Journal of Econometrics*, 78, 333–357.

Perlman, M. D. (1969): "One-Sided Testing Problems in Multivariate Analysis," *Annals of Mathematical Statistics*, 40, 549–567.

Pinkse, J., and G. Tan (2004): "The Affiliation Effect in First-Price Auctions," Penn State University and USC, forthcoming Econometrica.

Powell, J. L. (1986): "Symmetrically Trimmed Least Squares Esimation for Tobit Models," *Econometrica*, 54, 1435–1460.

Sareen, S. (1999): "Posterior Odds Comparison of a Symmetric Low-Price, Sealed-Bid Auction Within the Common Value and the Independent Private Value Paradigms," *Journal of Applied Econometrics*, 14, 651–676.

van der Vaart, A. (1999): *Asymptotic Statistics*. Cambridge University Press.

Wolak, F. (1989): "Testing Inequality Constraints in Linear Econometric Models," *Journal of Econometrics*, 41, 205–235.

——— (1991): "The Local Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models," *Econometrica*, 59, 981–995.

Table 1: Monte Carlo Results

	PV1		CV1		PV2		CV2	
Range of n:	2–4	2–5	2–4	2–5	2–4	3–5	3–5	3–6
T_n	200	200	200	200	200	200	200	200
share of p-values < 10%	0.21	0.39	1.00	1.00	0.12	0.27	0.94	0.99
share of p-values < 5%	0.11	0.29	1.00	1.00	0.05	0.18	0.91	0.99

Table 2: Monte Carlo Results
Bootstrap Estimation of Σ

	PV1		CV1		PV2		CV2	
Range of n:	2–4	2–5	2–4	3–6	2–4	3–5	3–5	3–6
T_n	200	200	200	200	200	200	200	200
share of p-values < 10%	0.14	0.18	1.00	1.00	0.13	0.21	0.80	0.91
share of p-values < 5%	0.10	0.12	1.00	1.00	0.04	0.11	0.70	0.83

Table 3: Data Configuration
USFS Timber Auctions

	Scaled Sales		Lumpsum Sales	
	number of auctions	number of bids	number of auctions	number of bids
$n = 2$	63	126	54	108
$n = 3$	39	117	40	120
$n = 4$	42	168	33	132
$n = 5$	33	165	16	80
$n = 6$	23	138	18	108
$n = 7$	14	98	11	77
$n = 8$	4	32	6	48
$n = 9$	9	81	7	63
$n = 10$	11	110	3	30
$n = 11$	1	11	0	0
$n = 12$	4	48	3	36
TOTAL	243	1094	191	802

Table 4: Summary Statistics
USFS Timber Auctions

	Scaled Sales		Lumpsum Sales	
	mean	std dev	mean	std dev
number of bidders	4.50	2.47	4.20	2.30
winning bid	80.50	51.49	77.53	46.57
appraised value	36.12	32.56	36.10	29.08
estimated volume	609.89	640.50	390.04	555.86
est. manuf cost	141.51	45.79	153.46	43.08
est. harvest cost	120.57	29.55	118.36	24.92
est. selling value	312.04	75.85	335.74	96.88
species concentration	0.5267	0.5003	0.5497	0.4988
6-month inventory	334161	120445	389821	139625
contract term	7.31	3.27	6.39	3.63
acres	697.78	2925.45	266.82	615.28
region 5 dummy	0.8519		0.6806	