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

Philip A. Haile  Han Hong
Dept. of Economics  Dept. of Economics
University of Wisconsin-Madison  Princeton University
Matthew Shum
Dept. of Economics
Johns Hopkins University

July 19, 2002

Preliminary and Incomplete

Abstract
We develop tests for the presence of common values at first-price sealed-bid auctions. Our tests are nonparametric, require observation only of the bids submitted at each auction, and are based on the fact that the “winner’s curse” is present only in common value auctions. We build on recent work of Guerre, Perrigne, and Vuong (2000) and Li, Perrigne, and Vuong (2000), who devise methods for using observed bids to estimate each bidder’s conditional expectation of the value of winning the auction. We point out that under the common values hypothesis these expectations are stochastically decreasing in the number of bidders, whereas with private values the distributions of these expectations are identical across all numbers of bidders. This forms the basis of our tests. We consider both exogenous and endogenous variation in the number of bidders. We develop the necessary asymptotic distribution theory for our tests, evaluate them in Monte Carlo experiments, and illustrate their application using data from auctions of offshore drilling rights.

Keywords: first-price auctions, common values, private values, nonparametric testing, winner’s curse, stochastic dominance

*We thank Don Andrews, Steve Berry, Tang Li, Harry Paarsch, Isabelle Perrigne, Rob Porter, seminar participants at the Tow conference in Iowa, the World Congress of the Econometric Society in Seattle, and various US universities for insightful comments. We thank Hai Che and Grigory Kosenok for exceptional research assistance. We thank the National Science Foundation and SSHRC (Canada) for financial support. Send e-mail correspondence to (respectively): phaile@ssc.wisc.edu, doubleh@princeton.edu, mshum@jhu.edu.
Introduction

At least since the influential work of Hendricks and Porter (1988), studies of auction data have played a particularly important role in demonstrating the empirical relevance of economic models of strategic interaction between agents with asymmetric information. However, a fundamental issue remains unresolved: how to choose between private and common value models of bidders’ information. The distinction is fundamental in the theoretical literature on auctions,\(^1\) with important implications for bidding strategies, and the design of selling mechanisms. For example, it is well-known (cf. Bulow and Klemperer (2002)) that expected seller revenues can be decreasing in the number of bidders in a common values setting (some evidence of this phenomenon for government procurement auctions is presented in Hong and Shum (2002)), so that the distinction between private and common values is of some policy importance.

This question was in fact the motivation for the first work on structural estimation of auction models (Paarsch (1992)). More generally, models in which agents’ private information leads to adverse selection (of which a common values auction model is just one example) have played a prominent role in the theoretical literature, yet the prevalence and significance of this type of informational asymmetry is not well established. Because the institutional rules of first-price auctions are particularly well captured by tractable theoretical models, data from these auctions provide an unusually promising opportunity to test for adverse selection.

We develop nonparametric tests for distinguishing between the common value (CV) and private value (PV) paradigms based on observed bids at first-price sealed-bid auctions. Several previous approaches have relied on parametric assumptions about the distribution functions governing bidders’ private information (e.g. Paarsch (1992), Sareen (1999)). Such tests necessarily confound evaluation of the economic hypotheses of interest with evaluation of parametric distributional assumptions. Other authors (e.g., Paarsch (1992), Gilley and Karels (1981)) have proposed examination of variation in bid levels with the number of bidders as a test for common values. However, Pinkse and Tan (2000) have recently shown that such reduced-form tests generally cannot distinguish common from private values models in first-price auctions: bids themselves may increase or decrease in the number of bidders in both PV and CV models. We overcome both of these limitations by taking a nonparametric structural approach, exploiting the relationships between observable bids and bidders’ latent expectations implied by theory. This testing approach builds on nonparametric estimation methods recently developed by Guerre, Perrigne, and Vuong (2000) (hereafter GPV) and extended by Li, Perrigne, and Vuong (2002) (hereafter LPV) and Hendricks, Pinkse, and Porter (2001) (hereafter HPP).

The importance of such tests for empirical research on auctions is emphasized by recent results (Athey and Haile (2002)) showing that CV models are identified only under strong conditions on the underlying information structure or on the types of data available. In fact Laffont and Vuong (1996) have argued that any common value model is observationally equivalent to some private

---

\(^1\)See, e.g., Milgrom and Weber (1982) for discussion of the distinctions between these models.
values model, suggesting that testing is impossible. However, they did not consider the possibility of binding reserve prices, variation in the numbers of bidders, or observation of ex-post profits, any of which could aid in distinguishing between the private and common value paradigms. Our tests exploit variation in the number of bidders and are based on detecting the effects of the winner’s curse on equilibrium bidding.

The winner’s curse is an adverse selection phenomenon arising in CV but not PV auctions. Loosely, winning a common values auction reveals to the winner that he was more optimistic about the object’s value than were any of his opponents. This “bad news” (Milgrom (1981)) becomes worse as the number of opponents faced by a bidder increases—being most optimistic among a large group of bidders implies (on average) even greater over-optimism about the value of the object than does being most optimistic among a small group of bidders. A rational bidder adjusts his expectation of the value of winning (and, therefore, his bid) to incorporate the information that his winning would reveal. 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 expectation of the value of the object conditional on winning. Relying on this distinction, our testing approach is based on detection of the adjustments rational bidders make in order to avoid the winner’s curse. Note that even in a PV model variation in the level of competition affects the aggressiveness of bidding; however, economic theory enables us to separate this competitive response from responses (if any) to the winner’s curse.

A number of approaches could be taken to test for this winner’s curse effect. We consider a class of multivariate one-sided hypothesis tests based on quantiles or first moments of the distributions of bidders’ expected values under the null and alternative hypotheses. We are not the first to use restrictions on equilibrium bidding implied by the winner’s curse to distinguish between the CV and PV hypotheses in first-price auctions. HPP proposed several tests but require either binding reserve prices or data on the ex post value of the objects sold. Our tests require observation only of the bids—the only data available from most first-price auctions. Paarsch (1991) and Bajari and Hortacsu (1999) have considered tests based on the winner’s curse in the context of second-price and English auctions, where a simple regression approach can be employed. As will be clear below, in a first-price auction the analysis is substantially complicated by competitive responses to changes in the level of competition. Athey and Haile (2002) have proposed a testing approach similar to that developed here. They focus on cases in which the researcher has access to only a subset of the bids submitted in an auction, but do not develop statistical tests. The formal tests developed in this paper could easily be extended to enable implementation the testing approach they propose.

The remainder of the paper is organized as follows. The next section summarizes the underlying affiliated values auction model, the method for inferring bidders’ expectations of their valuations from observed bids, and the main principle of our tests. In section 3 we provide the details of our tests, deriving the necessary asymptotic distributions. In section 4 we discuss unobserved heterogeneity and endogenous participation, showing how our testing approach can be extended to cases in...
which the number of bidders is correlated with unobserved factors affecting bidder valuations, using instrumental variables. In Section 5 we report the results of Monte Carlo experiments undertaken to evaluate the performance of our tests. Section 6 then presents an application to auctions of offshore drilling rights. We conclude in section 7.

2 Model and Testing Principle

The underlying theoretical framework is Milgrom and Weber’s (1982) symmetric affiliated values model. Throughout we denote random variables in upper case and their realizations in lower case. An auction has \( M \leq \bar{n} \) risk-neutral bidders. Each bidder \( i \) has a valuation \( U_i \) for the object and receives a private signal \( X_i \) of this valuation. Bidders’ valuations and signals have the joint distribution \( \tilde{F}(U_1, \ldots, U_n, X_1, \ldots, X_n) \) which is assumed to have a positive joint density. We make the following standard assumptions:

**Assumption 1** *(Symmetry)* Bidders are symmetric, in the sense that \( \tilde{F}(U_1, \ldots, U_n, X_1, \ldots, X_n) \) is exchangeable with respect to the indices 1, \ldots, \( n \).

**Assumption 2** *(Affiliation)* The random variables \( U_1, \ldots, U_n, X_1, \ldots, X_n \) are affiliated.\(^2\)

Initially, we also assume that the number of bidders is not correlated with bidder valuations:

**Assumption 3** For each \( n < \bar{n} \), \( \tilde{F}(U_1, \ldots, U_n, X_1, \ldots, X_n) \) is the same as the marginal distribution of \( U_1, \ldots, U_n, X_1, \ldots, X_n \) implied by \( \tilde{F}(U_1, \ldots, U_{\bar{n}}, X_1, \ldots, X_{\bar{n}}) \).\(^3\)

This is an assumption that variation in the number of bidders across auctions is exogenous, not driven by unobserved object heterogeneity across auctions.\(^4\) While we initially consider only exogenous variation in \( n \), below we will consider endogenous participation and show how Assumption 3 can be avoided when instrumental variables are available.

Given our assumptions, in an \( n \)-bidder auction there exists a unique symmetric Bayesian Nash equilibrium in which each bidder employs a strictly increasing strategies \( s_n(\cdot) \). As shown by Milgrom and Weber (1982), the first-order condition characterizing the equilibrium bid function \( s_n(\cdot) \) is

\[
v(x_i, x_i, n) = s_n(x_i) + \frac{s_n'(x_i) F_n(x_i | x_i)}{f_n(x_i | x_i)}
\]

where

\[
v(x, y, n) = E \left[ U_i | X_i = x, \max_{j \neq i} X_j = y \right]
\]

\(^2\)See Milgrom and Weber (1982) for a discussion of affiliation.

\(^3\)Therefore \( \tilde{F}(U_1, \ldots, U_n, X_1, \ldots, X_n) \) is considered fixed and the realization \( n \) serves to select the marginal distribution of the appropriate elements of \( (U_1, \ldots, U_n, X_1, \ldots, X_n) \). This assumption is not made by Milgrom and Weber (1982) since they consider fixed \( n \). However, standard models motivating the affiliated values setting satisfy this assumption. The assumption rules out structures like \( U_i = \sum_j X_j \), although such examples have little economic appeal.

\(^4\)Athey and Haile (2002) rely on this assumption discuss situations in which this will hold.
and \( F_n(\cdot|x_i) \) is the conditional distribution of the maximum signal among \( i \)'s opponents (with \( f_n(\cdot|x_i) \) the corresponding density).

Our testing approach is based on the fact that the conditional expectation \( v(x_i, x_i, n) \) is decreasing in the number of bidders whenever valuations contain a common value element. To show this, we first formally define common values.

**Definition 1** Within the symmetric affiliated values model, bidders have common values iff for all \( i \) and some realizations of \( X_i, E[U_i|X_1, \ldots, X_n] \) varies with \( X_j \) for \( j \neq i \).

This definition of common values incorporates all affiliated values models in which bidders' valuations contain a common component, not just the special case of pure common values, where the value of the object is unknown but identical for all bidders. A model within the Milgrom-Weber framework that is not a common values model is a private values model.\(^5\)

**Theorem 1** Bidders have common values if and only if \( v(x, x, n) \) is nonincreasing in \( n \) for all \( x \) and strictly decreasing in \( n \) for some \( x \).

**Proof:** Given symmetry, focus on bidder 1 without loss of generality. Without common values, \( E[U_1|X_1, \ldots, X_n] = E[U_1|X_1] \), which does not depend on \( n \). Hence dependence of \( v(x, x, n) \) on \( n \) for some \( x \) implies common values. To see that common values implies \( v(x, x, n) \) decreasing in \( n \), observe that

\[
v(x, x, n) \equiv E[U_1|X_1 = x, X_2 = x, \ldots, X_n = x, X_{n-1} = x, \ldots, X_1 = x] \\
= E_{X_n \leq x} E[U_1|X_1 = x, X_2 = x, \ldots, X_n = x, X_{n-1} = x, \ldots, X_1 = x] \\
\leq E \left[ U_1|X_1 = x, X_2 = x, \ldots, X_{n-1} \leq x \right] \\
\equiv v(x, x; n-1)
\]

with the inequality resulting from the common value assumption and the fact that with symmetry and affiliation, \( E[U_1|X_1, \ldots, X_n] \) cannot decrease in any \( X_j, j \neq 1 \). Furthermore, under common values the inequality must be strict for some \( x \).

Informally, in a common values auction, when a bidder \( i \) with signal \( x \) conditions on the event that \( \max_{j \neq i} X_j = x \), this reduces his expectation of the object’s value to him relative to his expectation when conditioning only on his own signal \( X_i = x \). Even if each signal \( X_i \) is an unbiased estimate of \( U_i \), the highest of \( n \) such signals (the signal of the winner) is an upwardly biased estimate in a CV

\(^5\)Our taxonomy is not the only one used in the literature. Some authors reserve the term “common values” for the special case of pure common values and use “interdependent values” or “affiliated values” for the class 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 or correlated. Note in particular that correlation of bidders’ information is neither necessary nor sufficient for common values.
auction. This bias is increasing in \( n \) so the conditional expectation \( v(x, x, n) \), which corrects for this bias, will be decreasing in \( n \).

### 2.1 Structural Interpretation of Observed Bids

We assume the researcher observes the bids \( b_1, \ldots, b_n \) from \( T_n \) \( n \)-bidder auctions. For simplicity we abstract from variation in the characteristics of objects sold at different auctions. As shown by GPV, standard nonparametric techniques can be applied to condition on auction-specific covariates. We assume throughout that each auction is independent of all others.\(^6\)

The observed data do not directly reveal the distribution of signals, but rather the distribution of bids. The key insight of GPV is that, in equilibrium, the joint distribution of bidder signals is related to the joint distribution of bids through the relations

\[
F_n(y|x) = G_n(s_n(y)|s_n(x))
\]

\[
f_n(y|x) \times \frac{1}{s'_{n}(y)} = g_n(s_n(y)|s_n(x))
\]

where \( G_n(\cdot|s_n(x)) \) is the conditional distribution (assuming all bidders follow \( s_n(\cdot) \)) of the highest bid submitted by \( i \)'s competitors. Given the symmetry and monotonicity of \( s_n(\cdot) \) that underlie (2), this will be the bid of the bidder whose signal is highest signal among \( i \)'s \( n-1 \) competitors.

Since in equilibrium \( b_i = s_n(x_i) \), the differential equation (1) can be rewritten

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

The left side of this equation gives bidder \( i \)'s expected value of the object conditional on winning the auction and on the highest signal among his opponents being the same as his own. While these “values” are not observed directly, both \( G_n(\cdot|\cdot) \) and \( g_n(\cdot|\cdot) \) are nonparametrically identified from the observed bids.

Let \( b_{it} \) denote the bid made by bidder \( i \) at auction \( t \) and let \( b^*_it \) represent the highest bid among \( i \)'s rivals in auction \( t \). GPV and LPV suggest nonparametric kernel estimates of the conditional bid distribution and density, viz.,

\[
\hat{G}_n(b; b) \equiv \hat{G}_n(b|b) \hat{g}_n(b) = \frac{1}{T_n \times h_G \times n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} K \left( \frac{b - b_{it}}{h_G} \right) 1 \left( b^*_it < b \right),
\]

\[
\hat{g}_n(b; b) \equiv \hat{g}_n(b|b) \hat{g}_n(b) = \frac{1}{T_n \times h_g^2 \times n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} K \left( \frac{b - b_{it}}{h_g} \right) K \left( \frac{b - b^*_it}{h_g} \right).
\]

Here \( h_G \) and \( h_g \) are bandwidths, \( K(\cdot) \) is a kernel and \( \hat{g}_n(\cdot) \) is an estimate of the (marginal) density \( g_n(\cdot) \) of the bids. \( \hat{G}_n(b; b) \) and \( \hat{g}_n(b; b) \) are nonparametric estimates of

\[
G_n(b; b) = \frac{\partial}{\partial m} \Pr(B_{it} \leq b, B^*_it \leq m)|_{m=b}
\]

---

\(^6\)This is a standard assumption, but one that serves to qualify many empirical studies of auctions, where data are taken from auctions in which bidders compete repeatedly over time.
and
\[ g_n(b; b) = \frac{\partial^2}{\partial m \partial b} \Pr(B_{it} \leq b, B_{it}^* \leq m)|_{m=b}. \]

For simplicity we will refer to these nonparametric estimates as, respectively, \( \hat{G}_n \) and \( \hat{g}_n \). Note that the subscript \( n \) emphasizes the fact that different estimates are required for auctions with different number of bidders.

By evaluating \( \hat{G}_n \) and \( \hat{g}_n \) at each of the observed bids, we can construct a pseudo-sample of the underlying random variable \( V_{it} \equiv v(X_{it}, X_{it}, n) \) using (3):
\[ \hat{v}_{it} \equiv b_{it} + \frac{\hat{G}_n(b_{it}; b_{it})}{\hat{g}_n(b_{it}; b_{it})} = \hat{\xi}(b_{it}; n). \]

This possibility was first articulated for independent private values models by Laffont and Vuong (1993) and GPV, and has been extended to affiliated values models by LPV and HPP. Following HPP and LPV, we refer to the estimated \( \hat{v}_{it} \) as a “pseudo-value.”

2.2 Main Principle of the Test

Each pseudo-value \( \hat{v}_{it} \) is an estimate of \( v(x_{it}, x_{it}, n) \). In PV models, \( v(x_{it}, x_{it}, n) = x_{it} \). In the general affiliated values model. However, if we do not make parametric functional form assumptions, the function \( v(\cdot, \cdot, n) \) is not known and we can not recover the private signals \( x_{it} \). However, we can nonparametrically estimate the distribution of pseudo-values across different \( n \), enabling us to exploit the result (from Theorem 1) that \( v(x_{it}, x_{it}, n) \) is decreasing in \( n \) in a CV setting, but constant across all \( n \) in a PV setting. Let \( F_{v,n} \) denote the distribution of \( V_{it} \) in \( n \)-bidder auctions. Under the PV hypothesis, the distributions \( F_{v,n} \) must be identical for \( n = 2, \ldots, \bar{n} \), while under the CV alternative, these distributions should be increasing in \( n \). This is formalized in the following corollary to Theorem 1.

**Corollary 1** Under the private values hypothesis
\[ F_{v,2}(v) = F_{v,3}(v) = \ldots = F_{v,n}(v) \quad \forall v. \]

Under the common values hypothesis
\[ F_{v,2}(v) \leq F_{v,3}(v) \leq \ldots \leq F_{v,n}(v) \quad \forall v \]

with each inequality strict for some \( v \).

**Proof:** The first claim is immediate. \( F_{v,n}(v^*) = \Pr(v(X_{it}, X_{it}, n) \leq v^*) = \Pr(X_{it} \leq (v^*, n)) \), where \( x(v, n) \equiv \sup \{x|v(x, x, n) \leq v) \}. Under the common value assumption, \( x(v, n) \) is nondecreasing in \( n \) for all \( v \), and strictly increasing for some \( v \). This implies that \( F_{v,n} \) first-order stochastically dominates \( F_{v,n+1} \).
3 Tests for Stochastic Dominance

Given Corollary 1, a test for stochastic dominance applied to estimates of the distributions $F_{v,n}$ provides a test for common values. If the “values” $v_{it} = v(x_{it}, x_{it}', n)$ were directly observed, a wide variety of tests for stochastic dominance from the statistics and econometrics literature could be used (see, e.g., McFadden (1989), Barrett and Donald (2001), Davidson and Duclos (2000) and Anderson (1996)). The empirical distribution function is commonly used to form test statistics. Supposing that $V_{i}$ were directly observed, let $F_{v,n}(y)$ denote the empirical distribution of $V_{i}$ for all auctions with $n$ bidders:

$$
F_{v,n}(y) = \frac{1}{T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} 1 (v_{it} \leq y)
$$

Some of these tests are consistent against all possible deviations from the null hypothesis that $F_{v,n} \equiv F_{v,m} \forall n, m$. These include the Kolmogorov-Smirnov test of McFadden (1989), which uses the sup norm in constructing the test statistics $\sup_{n \leq m} \sup_y \left( F_{v,n}(y) - F_{v,m}(y) \right)$, and related Von-Mises type statistics $\sum_{n \leq m} \int \left( F_{v,n}(y) - F_{v,m}(y) \right)^2 w(y) dy$, for some weighting function $w(y)$. Another consistent test is the rank test of stochastic dominance (Hajek, Sidak, and Sen (1999)), which uses the statistics $R = \frac{1}{m \sum_{i=1}^{m} \sum_{s=1}^{T_m} \sum_{j=1}^{m} 1 (y_{it} - y_{jt}) - \frac{1}{2}}$. For $n < m$, $R < 0$ suggests evidence against the null hypothesis of no stochastic dominance. Some tests can only detect the deviation from $H_0$ in given directions. For example Anderson (1996) compared $F_{v,n}(y)$ and $F_{v,m}(y)$ at a fixed grid of points of $y$. Many of these tests involve multivariate inequality constraints. Wolak (1989) proposed to formulate multivariate inequality tests based on the likelihood ratio principle using finite sample or asymptotic normality of the test statistics. Andrews (1998) noted that likelihood ratio test does not have optimal power properties for multivariate inequality hypothesis and proposed a class of weighted power tests.

Our testing problem has the complication that the realizations $v_{it} = v(x_{it}, x_{it}', n)$ are not directly observed but estimated from the first step construction (5). Hence, the empirical distributions we can construct are

$$
\hat{F}_{v,n}(y) = \frac{1}{T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} 1 (\hat{v}_{it} \leq y)
$$

Although we can formulate consistent tests based on the above principles, deriving the approximate large sample distribution for inference purposes is difficult. To facilitate implementation in empirical applications, we focus on testing the implications of stochastic dominance for a given set of functional of each $F_{v,n}$. In particular, let each of $\gamma_n$, $n = 1, \ldots, \bar{n}$ denote a vector of functionals of $F_{v,n}$. The one-sided tests we consider distinguish between hypotheses are all of the type

$$
H_0 \ (PV): \ \gamma_1 = \gamma_2 = \cdots = \gamma_N
$$

$$
H_1 \ (CV): \ \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_N, \ \text{with at least one inequality strict}
$$

\footnote{See Linton, Massoumi, and Whang (2002) for a recent example of stochastic dominance tests using estimated regression residuals.}
or, letting $\delta_{m,n} \equiv \gamma_m - \gamma_n$ and $\delta \equiv (\delta_{1,2}, \delta_{2,3}, \ldots, \delta_{N-1,N})'$,

$$H_0 \ (PV) : \delta = 0$$
$$H_1 \ (CV) : \delta \geq 0, \text{ with at least one inequality strict.} \tag{6}$$

We consider two types of vectors $\delta$ that allow tractable practical implementation. The first consists of quantiles of $F_{v,n}$ at a fixed grid of points. The second set involves first moments of each $F_{v,n}$.

While each of these options results in a small erosion in the power of the test, they offer practically feasible test statistics. In the next two sections, we show that for both the quantiles and first moments, estimates of each $\gamma_l$ (or the difference vector $\delta$) are available and have an approximate multivariate normal distribution in large samples. The results in these sections require the following assumptions on $G_n(b; b)$ and $g_n(b; b)$:

**Assumption 4**
1. $G_n(b; b)$ is $R + 1$ times differentiable in the first argument and $R$ times differentiable in the second argument. $g_n(b; b)$ is $R$ times differentiable in both arguments. The derivatives are bounded and continuous.
2. $K(u)$ has bounded support. $\int K(u) \, du = 1$ and $\int u^r K(u) \, du = 0$ for all $r < R$.
3. $h_G = h_g = h$. As $T_n \to \infty$, $h \to 0$, $T_n h^2 / \log T_n \to \infty$, $T_n h^{2+2R} \to 0$.

**3.1 Quantiles of $F_{v,n}$**

Let $b_{r,n}$ be the $\tau$th quantile of the bid distribution at $n$-bidder auctions, i.e.,

$$b_{r,n} = G_n^{-1}(\tau) = \inf \{ b : G_n(b) \geq \tau \}$$

where $G_n(\cdot)$ is the empirical distribution of the $n \times T_n$ observed bids in auctions with $n$ bidders. The pseudo-value for the $\tau$-quantile bidder in $n$-bidder auctions can be estimated by

$$\hat{v}_{r,n} = \hat{b}_{r,n} + \frac{G_n(b_{r,n}; b_{r,n})}{g_n(b_{r,n}; b_{r,n})}.$$

**Theorem 2** Suppose Assumption 4 holds. Then as $T_n \to \infty$

1. $b_{r,n} - s_n \left( F^{-1}_r(\tau) \right) = O_p \left( \frac{1}{\sqrt{T_n}} \right)$. 

---

*The main difficulties in deriving distributional results for tests that are consistent against all deviations from the null are complicated trimming arguments used to handle the boundary of the support of the pseudo value distribution. Trimming arguments are difficult because of the presence of local nonparametric kernel density estimate $\hat{g}_n(b_{r1}; b_{r1})$ in the denominator of (5). The trimming problem might be alleviated if we instead consider global nonparametric estimated of the conditional inverse hazard function $G_n(b_{r1}; b_{r1})$ using sieve based methods (see for example Chen and Shen (1998)).*
2. For each fixed $b$ such that $g_n(b; b) \equiv g_n (b|b) g_n (b) > 0$,
\[ \sqrt{nT_n h^2} \left[ \tilde{\xi}(b; n) - v \left( s_n^{-1} (b), s_n^{-1} (b), n \right) \right] = \sqrt{nT_n h^2} \left( \frac{\tilde{G}_n (b|b)}{\tilde{g}_n (b|b)} - \frac{G_n (b|b)}{g_n (b|b)} \right) \]
\[ \overset{d}{\longrightarrow} N \left( 0, \frac{1}{n} \frac{G_n (b|b)^2}{g_n (b|b)^2} \left[ \int K(x)^2 K(y)^2 dxdy \right] \right) \]

3. For distinct quantile values $\tau_1, \ldots, \tau_L$, assume
\[ g_n \left( s_n \left( F_x^{-1} (\tau_l) \right) | s_n \left( F_x^{-1} (\tau_l) \right) \right) g_n \left( s_n \left( F_x^{-1} (\tau_l) \right) \right) > 0, \quad l = 1, \ldots, L. \]
Then the $L$-dimensional vector of elements $\sqrt{nT_n h^2} \left( \tilde{b}_{\tau_l; n} - v \left( F_x^{-1} (\tau_l), F_x^{-1} (\tau_l), n \right) \right)$ converges in distribution to $Z \sim N(0, \Omega)$, where $\Omega$ is a diagonal matrix with $l$-th element
\[ \Omega_l = \frac{1}{n} \frac{G_n \left( s_n (x_l) \right) | s_n (x_l) \right)^2}{g_n \left( s_n (x_l) \right)^4} \left[ \int K(x)^2 K(y)^2 dxdy \right] \]
where $x_l \equiv F_x^{-1} (\tau_l)$.

The proof is given in the appendix. Note that $G_n (b|b)$ is estimated more precisely than the $g_n (b|b)$ for all bandwidth sequence $h$. For simplicity we have chosen $h_G = h_y$, in which case the sampling variance will be dominated by that from estimation of $g_n (b|b)$.

Since the sample quantile converges at rate $\sqrt{nT_n}$ to the population quantile, the sampling variance of $\hat{v}_{\tau,n} - v(x, x, n)$ is governed by the slow pointwise nonparametric convergence rate of $\hat{g}_n (\cdot; \cdot)$. As shown in GPV, for fixed $b$, $\hat{g}_n (b; b)$ converges at rate $\sqrt{nT_n h_n^2}$ to $g_n (b; b)$ due to the estimation of a univariate conditional density. Theorem 2 below describes the limiting behavior of each $\hat{v}_{\tau,n}$.

The above theorem directly implies the following corollary, which considers a vector of quantiles of $F_{x,n}$ at a fixed value of the quantile $\tau$, but across different values of $n$, the number of bidders:

**Corollary 2** Define the $(n-2)$-dimensional vector $\hat{\delta}_l = \begin{pmatrix} \hat{\delta}_{2,3,l} & \hat{\delta}_{3,4,l} & \cdots & \hat{\delta}_{n-1,n,l} \end{pmatrix}^T$, where $\hat{\delta}_{m,n,l} \equiv \hat{v}_{m,n} - \hat{v}_{n,n}$ denotes the unnormalized difference. Under the null PV hypothesis, the $N$-2-dimensional vector of $\hat{\delta}_l$ converges in distribution to $Z \sim N(0, \Sigma_l)$, where
\[ \Sigma_l = \left[ \begin{array}{cccc} \omega_{2,2} & \omega_{2,3} & \cdots & 0 \\ \omega_{3,2} & \omega_{3,3} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_{n-1,n} \\ \end{array} \right] \]
\[ \omega_{m,l} = \frac{1}{mT_n h_n^2 g_n^2} \left( s_m \left( F_x^{-1} (\tau_l) \right) | s_m \left( F_x^{-1} (\tau_l) \right) \right) g_n (b) \left[ \int K^2 (x) K^2 (y) dxdy \right]. \]

---

9 We have assumed undersmoothing rather than optimal smoothing to avoid estimating the asymptotic bias term for inference purpose. An alternative is to choose the $h_G$ and $h_y$ sequences differently. In fact, if we have chosen $h_G$ and $h_y$ close to their optimal range, the sampling variance will still be dominated by that of $\hat{g}_n (b; b)$ and the result of the theorem will not change. On the other hand we have chosen $h_G \approx h_y^2$ so that $G_n (b; b)$ and $\hat{g}_n (b; b)$ share the same magnitude of variance, then the convergence rate for $G_n (b; b)$ will be far from optimal.
3.2 First moments of $F_{x,n}$

We can estimate the average pseudovalue $E_x[v(x, x, n)] = \int v dF_{x,n}$ for the set of $n$-bidder auctions using the sample average of the pseudo-values calculated across all auctions with $n$ bidders:

$$\hat{\mu}_n = \frac{1}{n \times T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \hat{v}_{it}. \quad (7)$$

One difficulty of implementing the first moment approach is the issue of the boundary, where kernel density estimates such as those appearing in the denominator of (5). Unlike partial mean nonparametric regression problems (cf. Newey (1994)) where fixed exogenous trimming is possible, care has to be taken here to preserve consistency of the test statistics with trimming at the boundary. In particular, we must avoid trimming off the part of the distribution in which the difference between the null and alternative hypotheses will appear.\(^\text{10}\) To overcome this problem, we use a fixed trimming scheme by equalizing quantiles over the distribution of bids with different number of bidders. This fixed trimming scheme produces consistent test statistics against almost all common value alternatives except those that demonstrate the effects of the winner’s curse only at the boundaries of the support of signals. Such situations maybe considered rare.

To be precise, for small $\tau > 0$ let $\hat{b}_{\tau,n}$ and $\hat{b}_{1-\tau,n}$ denote the $\tau$th and $1-\tau$th quantile of the bid distribution with $n$ bidders. The **quantile-trimmed first moment** is defined as

$$\bar{\mu}_{n,\tau} = \frac{1}{n \times T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \hat{v}_{it} 1 \left( \hat{b}_{\tau,n} \leq b_{it} \leq \hat{b}_{1-\tau,n} \right)$$

Strictly speaking, for fixed $\tau$ a test based on the quantile-trimmed first moment tests the modified hypotheses

$$H_0 : E[v(x, x, 1) 1 (x_\tau \leq x \leq x_{1-\tau})] = \cdots = E[v(x, x, N) 1 (x_\tau \leq x \leq x_{1-\tau})]$$

$$H_1 : E[v(X, X, 2) 1 (x_\tau \leq X \leq x_{1-\tau})] > \cdots > E[v(X, X, n) 1 (x_\tau \leq X \leq x_{1-\tau})].$$

For small $\tau$, however, little power is lost by the fixed trimming method.\(^\text{11}\) The next theorem shows the consistency and asymptotic distribution of $\hat{\mu}_{n,\tau}$.

**Theorem 3** Suppose Assumption 4 holds, and in addition $(\log T)^2 (Th^3) \to \infty$ and $T_n h^{1+2R} \to 0$, then (i) (Consistency) $\bar{\mu}_{n,\tau} \xrightarrow{p} E[v(X, X, n) 1 (x_\tau \leq X \leq x_{1-\tau})]$. (ii) (Asymptotic distribution)\(^\text{10}\)

$$\sqrt{T_n h} (\hat{\mu}_{n,\tau} - E[v(X, X, n) 1 (x_\tau \leq X \leq x_{1-\tau})]) \xrightarrow{d} N(0, \omega_n)$$

where

$$\omega_n = \left[ \int \int K(v) K(u+v) dv \right] \frac{1}{n} \int_{F^{-1}_{\tau}(\tau)} \frac{G_n(b; b)^2}{g_n(b; b)^2} db. \quad (9)$$

and the integration is over the support of the kernel function $K(\cdot)$.

\(^{10}\)One might attempt to mimic the stochastic trimming used in, for example, Lavergne and Vuong (1996) and Lewbel (1998). However, this approach usually requires smoothness conditions on $g_n(b; b)$ to reduce the order bias. However, since the pseudo-value involves the inverse of $g_n(b; b)$, the asymptotic variance may not be finite when $g_n(b; b)$ is smooth.

\(^{11}\)In principle, $\tau$ can be made to shrink to zero as the sample grows.
We show that the convergence rate of our estimate of each $\hat{\mu}_n$ is $\sqrt{T_n h}$, which is slower than the parametric rate $\sqrt{T_n}$ but faster than the $\sqrt{T_n h^2}$ rate of the quantile-based tests described above. Intuitively, the intermediate $\sqrt{T_n h}$ rate of convergence arises because $\hat{g}_n(b; b)$ is an estimated bivariate density function, but in constructing the estimate $\hat{\mu}_n$ we average along the one-dimensional 45° line ($b_{ij}, b^*_{ij} = b_{ij}$) (cf. Newey (1994)).

3.3 Testing approaches

A likelihood ratio (LR) test (e.g., Bartholomew (1959), Wolak (1989)) or the weighted power test of Andrews (1998) are possible alternative procedures to formulate test statistics based on the asymptotic normal distribution of $(\hat{\delta} - \delta)$. Since we do not have a good a priori choice of the weighting function for Andrews’ weighted power test, we have chosen to use the LR test. In fact, 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).

In this section we describe a multivariate LR first moment test for the hypotheses (8), based on Bartholomew (1959). An analogous test can be constructed using the quantiles of $F_{v,n}$, which we omit in the interest of brevity.

Let $\omega_n$ denote the asymptotic variance given in (10) for each value of $n = n, \ldots, \bar{n}$ and define

$$a_n = \frac{T_n h}{\omega_n}.$$ 

Denote the asymptotic covariance matrix of the vector $(\hat{\mu}_{n,\tau} \ldots \hat{\mu}_{n,\tau})'$ by

$$\Sigma = \begin{bmatrix} \frac{1}{a_n^2} & 0 & 0 & 0 \\ 0 & \frac{1}{a_{n+1}^2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{a_{\bar{n}}^2} \end{bmatrix}.$$ 

The restricted maximum-likelihood estimate of the mean pseudovalue under the null hypothesis $H_0 : \mu_{n,\tau} = \cdots = \mu_{n,\tau}$ is given by

$$\bar{\mu} = \frac{\sum_{n=n}^{\bar{n}} a_n \hat{\mu}_{n,\tau}}{\sum_{n=2}^{\bar{n}} a_n}.$$ 

To test against the alternative $H_1 : \mu_{n,\tau} \geq \mu_{n+1,\tau} \geq \cdots \geq \mu_{n,\bar{\tau}}$,

12While the test based on averaged pseudo-values converges faster than that based on fixed number of quantiles, the improvement of the convergence rate is not proportional since the conditions on bandwidth for the partial mean case is more stringent than the pointwise estimate. However there are still improvements even after taking this into account.

13Complete details of the quantile-based tests are available from the authors upon request.
let $\mu_{n}^{*}, \ldots, \mu_{\bar{n}}^{*}$ denote the solution to

$$
\min_{\mu_{n}^{*}, \ldots, \mu_{\bar{n}}^{*}} \sum_{n=1}^{\bar{n}} a_{n} (\hat{\mu}_{n, \tau} - \mu_{n})^{2} \quad s.t. \quad \mu_{1} \geq \mu_{2} \geq \cdots \geq \mu_{\bar{n}}.
$$

This solution can be found using the well-known “pool adjacent violaters” algorithm, using the weights $a_{n}$. Now define the test statistic

$$
\bar{\chi}^2 = \sum_{n=1}^{\bar{n}} a_{n} (\mu_{n, \tau}^{*} - \bar{\mu})^{2}.
$$

Let $K$ denote the number of “distinct values” in $\mu_{n}^{*}, \ldots, \mu_{\bar{n}}^{*}$. Note that $K \in \{1, \ldots, \bar{n} - n + 1\}$.

Under the null hypothesis, the LR statistic $\bar{\chi}^2$ has an asymptotic distribution equal to a mixture of $\chi^2$ random variables:

**Corollary 3**

$$
P (\chi^2 \geq c) = \sum_{k=1}^{n-2} P (\chi^2_k \geq c) w(k+1; \Sigma)
$$

where $\chi^2_k$ denotes a standard chi-square distribution with $k$ degrees of freedom. The weight $w(k+1; \Sigma)$ is the probability that a vector from the multivariate $N(0, \Sigma)$ distribution has $k+1$ “distinct values”, as defined in the manner above.

The proof is given in Bartholomew (1959), Section 3. This is the form of the tests which we employ in the Monte Carlo experiments described below, and also for the small application that we consider.

### 4 Unobserved Heterogeneity and Endogenous Participation

In many applications one may be concerned that the number of bidders is both determined endogenously and correlated with unobserved factors affecting bidder valuations, potentially invalidating our testing approach. There are at least two types of concerns here. The first arises when there is a binding reserve price and we observe only the number of participating bidders, not the number of potential bidders. HPP have proposed a testing approach applicable in such situations. Given this prior work and the fact that many important auctions lack binding reserve prices, we focus on testing in the complementary cases. The second concern is that, even when we observe the number of potential bidders, factors determining this number are correlated with unobserved factors affecting the distribution of bidder valuations. Here we show how our approach can be extended to provide a valid test in this case using an instrumental variable.\(^{15}\)

\(^{14}\)For example, HPP are unable to apply their test to the OCS auctions they study, due to the extremely low frequency with which the reserve price binds. Other authors, e.g., Baldwin, Marshall, and Richard (1997), Haile (2001), and Campo, Perrigne, and Vuong (1998) have argued that reserve prices in USFS timber auctions are nonbinding.

\(^{15}\)A third possibility is that the number of potential bidders is observed but, due to the presence of a binding reserve price, bids are not observed from all potential bidders. Athey and Haile (2002) propose a testing approach applicable...
Assume that bidders observe, in addition to their own signals $X_i$ and the number of bidders $N$, the realization of an auction-specific random variable $W$, which is also observable to the econometrician. Dropping Assumption 3, we now assume $W$ is correlated with the number of bidders $N$; however, whatever the realization of $N$, conditioning on $W$ has no effect on the joint distribution of signals and valuations:

**Assumption 5** $\Pr(N \leq n|W = w)$ is decreasing in $w$.

**Assumption 6** For all $n$, $\bar{F}(X_1, \ldots, X_n, U_1, \ldots, U_n|W) = \bar{F}(X_1, \ldots, X_n, U_1, \ldots, U_n)$.

An example satisfying these conditions is the simple linear model in which

\[
\begin{align*}
    u_i &= u + \epsilon_1^i \\
    x_i &= v_i + \epsilon_2^i \\
    n &= \phi(w) + \epsilon^3
\end{align*}
\]

where $\phi(\cdot)$ is a (possibly unknown) strictly increasing function; the error terms are all mean zero and independent of $w$ and $u$; $\epsilon_2$ and $\epsilon_1^2$ are independent; $\epsilon_1$ and $\epsilon_3$ are independent; and $\epsilon_1^i$ and $\epsilon_3$ are correlated. Assuming $U$ is not degenerate, one obtains a private values model if the variance of $\epsilon_2$ is degenerate and a common values model otherwise. One can see the nature of the endogeneity problem in this example. Under the PV hypothesis,

\[
\Pr(v(X_i, X_i; n) \leq v) = \Pr(X_i \leq v|N = n) = \Pr(U + \epsilon_1^i \leq v|\phi(W) + \epsilon_3 = n) \neq \Pr(U + \epsilon_1^i \leq v|\phi(W) + \epsilon_3 = n + 1)
\]

i.e., under the PV null, the distribution $F_{v,n}$ varies with $n$. Likewise, under the CV alternative, the stochastic dominance relation of Corollary 1 need not hold.

Define

\[
v(x, x; n, w) \equiv E \left[ U_i \mid X_i = x, \max_{j \neq i} X_j = x, N = n, W = w \right].
\]

Let $s_n(\cdot; w)$ denote the symmetric equilibrium bid function conditional on the realizations of $N$ and $W$. The first-order condition characterizing equilibrium bidding in an auction with $N = n$ and $W = w$ is now

\[
v \left( s_n^{-1}(b; w), s_n^{-1}(b; w); n, w \right) = b + \frac{G_n(b, b; w)}{g_n(b, b; w)}
\]

in such situations. As noted above, the formal tests of stochastic dominance developed here could be extended to their testing approach as well. Yet another possibility, when the reserve price binds, is that bidders know both the number of potential bidders and the number of bidders with sufficiently high signals that they will bid (or at least a signal of the latter). This case is one of several models studied by Shneyerov (2002). In most applications, however, information about the number of participating bidder will not be available to bidders before they bid.

\[\text{Possible instruments might include the number of sawmills near a tract of timber (Haile, 2001) or the length of an internet auction.}\]
The dependence of the distribution \( G_n(b, b; w) \) on \( w \) is due to the fact that, given \( N = n \), the realization of \( W \) provides information about the realization of unobservable factors affecting the number of bidders, and these unobservable factors may be correlated with the joint distribution of bidder signals and valuations.

Following the same type of estimation procedure described above (but now conditioning on the realization of \( W_t \) in addition to that of \( N_t \)) we can recover a pseudo-sample of realizations of the random variable \( v(X_{it}, X_{it}; N_t, W_t) \). Define

\[
\tilde{v}(x, w) \equiv E\left[ v(X_{it}, X_{it}; N_t, W_t) | X_{it} = x, W_t = w \right].
\]

Then the sample of pseudo-values \( v(x_{it}, x_{it}; n_t, w_t) \) is a sample of the realizations of

\[
\tilde{v}(X_{it}, W_{it}) + \eta_t
\]

where \( \eta_t \) has mean zero conditional on \((X_t, W_t)\) by construction. Assumption 6 implies that \( \tilde{v}(x, w) \) is

\[
\tilde{v}(x, w) = \sum_{k=1}^{n} \Pr(N_t = k | W_t = w) E\left[ U_{it} | X_{it} = \max_{j \in \{1, \ldots, k\}} X_{jt} = x \right].
\]

As argued previously, \( E[V_{it} | X_{it} = \max_{j \in \{1, \ldots, k\}} X_{jt} = x \] is invariant to \( k \) in a private values model but strictly decreasing in \( k \) in a common values model. Hence, Assumption 5 implies that \( \tilde{v}(x, w) \) is

\[
\text{invariant to } w \text{ in a private values model but decreasing in } w \text{ in a common values model. Therefore,}
\]

our first-moment test can be applied using exogenous variation in \( W_t \) instead of exogenous variation in \( N_t \). In particular, we can examine whether the mean pseudo-value is constant or decreasing in \( W_t \).

5 Monte Carlo Simulations

Here we summarize the results of Monte Carlo experiments performed to evaluate our test. To examine both size and power, we consider data generated by two private values models and two common values models:

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

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

(CV1) common values, i.i.d. signals \( x_i \sim u[0, 1] \), \( u_i = \alpha x_i + (1 - \alpha) \sum_{j \neq i} x_j n^{-1} \);\(^{17}\)

(CV2) pure common value \( u_i = u \sim u[0, 1] \), conditionally independent signals, \( x_i \sim u[0, u] \).\(^{18}\)

\(^{17}\)For the case \( \alpha = \frac{1}{2} \) considered below, one can show that \( v(x, x; n) = \frac{3n-2}{4n-1} x \), leading to the equilibrium bid function \( s_n(x) = \frac{3n-2}{4n} x \). In this example we see that although \( v(x, x; n) \) is strictly decreasing in \( n \), \( s_n(x) \) strictly increases in \( n \).

\(^{18}\)The symmetric equilibrium bid function for this model, given in Matthews (1984), is \( s_n(x) = \int_0^x \tilde{v}(t, n) \left( \frac{n-1}{x} \right) (\frac{t}{x})^{n-2} dt \) where \( \tilde{v}(t, n) = \int_x^t c g(c|n) dc \) and \( g(c|x, n) = \frac{c^{n-1}}{\Gamma(n) x^n} \).
Tables 1 and 2 contain the results from Monte Carlo experiments utilizing the multivariate (Bartholomew) version of the mean test. We have chosen to focus on the means-based tests in this section and in the subsequent application, due to its faster rate of convergence relative to the quantile-based tests.

The last two rows in Table 1 indicate that in the PV1 and PV2 paradigms, where the PV null holds, there is a tendency to over-reject, which becomes more apparent as we consider more bidders. For example, if we consider a test with significance level 0.10, under the PV1 paradigm our results indicate that we would reject 20.5% of the time when the range of bidders is 2–4, and 39% of the time when the range of bidders increases to 2–5. The asymptotic rejection probability should, in both cases, be 10%. However, the results also indicate that our tests have good power properties, since uniformly across the CV1 and CV2 paradigms, as well as across different ranges of #bidders, the p-values are very high, which would lead to overwhelming rejections for tests with any reasonable significance level.

One reason for the over-rejections under the null may be that the asymptotic approximations of the variances of the average pseudovalues are not good for the rather modest sample sizes that we consider. To address this possibility, we consider results from tests based upon bootstrap estimates of these variances. The results, reported in Table 2, indicate that the tendency towards over-rejection is attenuated when we estimate the variances via bootstrap: again, for a test with significance level 0.10, we now reject 13.5% of the time when the range of #bidders is 2–4, and 18% of the time when

---

**Table 1: Monte Carlo Results: Bartholomew LR Mean Test**

For each experiment, calculate standard error by asymptotic formula. 200 replications of each experiment.

<table>
<thead>
<tr>
<th>Range of #bidders:</th>
<th>PV1</th>
<th>CV1</th>
<th>PV2</th>
<th>CV2</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>
| %(%pval < 10%)  
 a             | 0.205 | 0.39 | 1.0 | 1.0 |
| %(%pval < 5%)  
 b             | 0.11 | 0.29 | 1.0 | 1.0 |

*Under private value null, this should be 0.10.
*Under private value null, this should be 0.05.

**Table 2: Monte Carlo Results: Bartholomew LR Mean Test**

For each experiment, calculate standard error by bootstrap. 200 replications of each experiment.

<table>
<thead>
<tr>
<th>Range of bidders:</th>
<th>PV1</th>
<th>CV1</th>
<th>PV2</th>
<th>CV2</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>
| %(%pval < 10%)  
 a | 0.135 | 0.18 | 1.0 | 1.0 |
| %(%pval < 5%)  
 b | 0.095 | 0.115 | 1.0 | 1.0 |

*Under private value null, this should be 0.10.
*Under private value null, this should be 0.05.
the range of #bidders is 2–5. The power properties remain very good.

We believe that these results are encouraging, and we proceed to an application of our tests to bids submitted in the USFS timber auctions.

6 Application: USFS Timber Auctions

We consider a small application of our proposed test to detect for the presence of common values in timber auctions run by the United States Forest Service (USFS). In each sale, contracts for timber harvesting on federal land are sold by first-price sealed bid auction. Detailed descriptions of the auctions can be found in Baldwin, Marshall, and Richard (1997), Haile (2001), or Athey and Levin (2000). We discuss only a few key features that are particularly relevant to our application.

First, despite considerable attention to these auctions in the empirical auction literature, there has been disagreement about whether these should be viewed as common or private value auctions. At first blush, the common values model might seem the obvious choice, since there is always considerable common uncertainty over the volume and quality of timber available on a particular tract. Indeed, in what are known as lump sum sales bidders bear all of this uncertainty, since the winning bidder pays a fixed price to the government regardless of the realized volume of harvested timber. More important, bidders typically have private information in lump sum sales, since they usually will send their own “cruisers” to assess the value of the tract before the auction.

However, in many Forest Service timber sales, bidders are insured against much of this uncertainty. In what is known as a “scaled sale,” the bidding is over unit prices for each species on the tract. Payment (other than a deposit) is not collected from the winning bidder until the timber is harvested and is then based on the actual volumes of each timber species. This policy has been used (often along with price indexing, so that bids are real prices per unit) specifically to avoid placing risk on the bidders. As a result, much of the common uncertainty is insured away. Furthermore, bidders are much less likely to send their own cruisers to estimate the tract value in the case of a scaled sale (National Resources Management Corporation (1997)), leaving little room for private information regarding any remaining common uncertainty. However, because bidders are specialized sawmills producing different timber products and, often, using different technologies, they have private information about their own sales and inventories of end products, contracts for future sales, and inventories of uncut timber from private timber sales. This has led many authors (e.g., Baldwin, Marshall, and Richard (1997) and Haile (2001)) to assume a private values model. However, this is not without controversy. Athey and Levin (2000) and Baldwin (1995) argue for a common values model even for scaled sales.

19 The forest service also conducts English auctions, although we do not consider these here.
21 Other studies assuming common values models for Forest Service timber auctions include Chatterjee and Harrison (1988), Lederer (1994), and Leffler, Rucker, and Munn (1994).
Based on these arguments, we consider two sets of auctions. The first is a sample of lumpsum sales, which we have argued seem particularly likely to involve common value components. The second sample contains only scaled sales. The auctions considered were took place between 1982 and 1990 in Forest Service regions 1, 5, and 6. The restriction to sales after 1981 is made due to policy changes in 1981 that reduced the significance of subcontracting as a factor affecting bidder valuations, since resale opportunities can confound the empirical implications of the winner’s curse (Haile (2001)). For the same reason, we restrict attention to sales with no more than 18 months between the auction and the harvest deadline. For consistency, we consider only sales in which the Forest Service provided ex ante estimate of the tract values using the predominant method of this time period, known as the “residual value method.” We also exclude salvage sales, sales set aside for small bidders, and sales of contracts requiring the winner to construct roads.

6.1 Test results

The results are reported in Tables 3, 4, and 5. The message from the top panels of Tables 3 and 4 are unambiguous: the data appear to support the importance of private values in the unit-price sub-sample, and common values in the lump-sum sales sub-sample. The $p$-values are uniformly higher for the unit-price auctions than the lump-sum auctions. Furthermore, the results are qualitatively quite similar across both the tests using the asymptotic expression for the variance, and the tests using a bootstrapped variance.

The bottom panels of Tables 3 and 4 present results from tests where we attempt to control for observed heterogeneity. In a preliminary step, we regressed the logged bids on a set of exogenous covariates which proxy for the public information known to all bidders at the time of bidding, and subsequently we performed the test using the residuals from this regression (exponentiated to ensure non-negative support). The covariates used in these regression were: (i) year and region dummies; (ii) the appraised value of the tract; (iii) volume; (iv) estimated manufacturing cost; (v) estimated harvest cost; (v) estimated selling value; (vi) species concentration index; (vi) 6-month inventory; and (vii) the contract term (i.e., the months the winner has to cut the timber). The results are clearly less conclusive than before: now, the $p$-values are uniformly large across both the lump-sum as well as unit-price subsamples.

Finally, in Table 5, we present results from a version of the test which attempts to control for endogenous participation using instrumental variables, in the way described in Section 4 above. The instrument used is a measure of the number of sawmills in the county of the sale, as well as adjacently surrounding counties. The results from these tests echo the results from the earlier Tables 3 and 4, and unambiguously support the PV hypothesis for the unit-prices sub-sample, but reject in favour of the CV hypothesis for the lump-sum sales subsample.

---

22Region 1 covers Montana, Northern Idaho, North Dakota, and northwestern South Dakota. Region 5 includes national forests in California. Our sample of region 6 auctions include tracts in Washington and Oregon.
Table 3: Bartholomew LR Mean Test for USFS Timber Auctions: Unit Prices Sub-sample  

*Using Bid Levels*

Test statistic constructed using asymptotic variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>1.6587</td>
</tr>
<tr>
<td><em>p</em>-value</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.3360</td>
</tr>
</tbody>
</table>

Test statistic constructed using bootstrap variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0015</td>
</tr>
<tr>
<td><em>p</em>-value</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.8787</td>
</tr>
</tbody>
</table>

*Using exp(residuals) from regression of log(bid) on exogenous covariates*

Test statistic constructed using asymptotic variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0060</td>
<td>0.0000</td>
</tr>
<tr>
<td><em>p</em>-value</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.6638</td>
<td>0.8050</td>
</tr>
</tbody>
</table>

Test statistic constructed using bootstrap variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0174</td>
<td>0.0045</td>
</tr>
<tr>
<td><em>p</em>-value</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.7301</td>
<td>0.8017</td>
</tr>
</tbody>
</table>
Table 4: Bartholomew LR Mean Test for USFS Timber Auctions: Lumpsum Sales Sub-sample

*Using bid levels*

Test statistic constructed using asymptotic variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>7.3058</td>
<td>42.7702</td>
<td>38.3413</td>
<td>35.2750</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0079</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Test statistic constructed using bootstrap variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>4.0473</td>
<td>11.0494</td>
<td>12.1821</td>
<td>10.8472</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0468</td>
<td>0.0020</td>
<td>0.0018</td>
<td>0.0047</td>
</tr>
</tbody>
</table>

*Using exp(residuals) from regression of log(bid) on exogenous covariates*

Test statistic constructed using asymptotic variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>0.0000</td>
<td>1.4461</td>
<td>0.1185</td>
<td>0.0000</td>
</tr>
<tr>
<td>p-value</td>
<td>1.0000</td>
<td>0.2452</td>
<td>0.6639</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Test statistic constructed using bootstrap variances for average pseudovalues

<table>
<thead>
<tr>
<th>Range of bidders</th>
<th>2–4</th>
<th>2–5</th>
<th>2–6</th>
<th>2–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test stat</td>
<td>0.0040</td>
<td>0.0601</td>
<td>0.0321</td>
<td>0.7410</td>
</tr>
<tr>
<td>p-value</td>
<td>0.5806</td>
<td>0.6076</td>
<td>0.7672</td>
<td>0.5360</td>
</tr>
</tbody>
</table>
Table 5: Multivariate Bartholomew LR Mean Test for USFS Timber Auctions: IV Approach

Using IV (number of sawmills in county of sale, and all adjacent counties) to control for possible endogenous participation.

Test statistic constructed using bootstrap variances for average pseudovalues.

<table>
<thead>
<tr>
<th>Range of IV* →</th>
<th>2–3</th>
<th>2–4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–2</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(1.000)</td>
<td>(1.000)</td>
<td></td>
</tr>
<tr>
<td>1–3</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(1.000)</td>
<td>(1.000)</td>
<td></td>
</tr>
<tr>
<td>1–4</td>
<td>1.058</td>
<td>2.050</td>
</tr>
<tr>
<td>(0.329)</td>
<td>(0.181)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Range of bidders→</th>
<th>0 1 2 3 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2 39 13 8 8</td>
<td>74</td>
</tr>
<tr>
<td>3</td>
<td>3 12 5 16 11</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>1 16 7 14 5</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>0 3 8 9 5</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>0 8 4 7 6</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>1 2 4 5 4</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>0 3 1 4 1</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>0 2 3 2 1</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>0 0 2 1 0</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>0 0 0 1 1</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>0 2 2 1 0</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>7 87 49 68 42</td>
<td>268</td>
</tr>
</tbody>
</table>

*: IV is discretized as int [#{sawmills}/50], where int denotes the closest integer.

Cross-tabs of #bidders, vs, value of instrument: Lump-sum sales sub-sample

<table>
<thead>
<tr>
<th>numbid</th>
<th>0 1 2 3 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3 24 11 20 17</td>
<td>86</td>
</tr>
<tr>
<td>3</td>
<td>3 10 7 10 11</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>5 16 8 13 6</td>
<td>64</td>
</tr>
<tr>
<td>5</td>
<td>0 14 6 6 5</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>1 10 2 5 5</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td>1 5 4 5 3</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>0 2 0 5 1</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>0 2 2 3 1</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>1 1 2 3 1</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>0 1 0 0 0</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>0 0 2 2 1</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>14 85 44 72 51</td>
<td>360</td>
</tr>
</tbody>
</table>

Cross-tabs of #bidders, vs, value of instrument: Unit Prices sub-sample
7 Conclusions and Extensions

We have provided a method to test for common values in symmetric first-price sealed bid auctions. We have also pointed out the applicability of tools developed in GPV, not just for structural estimation of auction models, but also for testing the PV hypothesis against the CV alternative.

While we have assumed symmetric bidders, this is not necessary. It is possible to extend our methodology to detect common value elements with ex ante asymmetric bidders, as long as at least one bidder participates in auctions with different numbers of competitors. Two modifications of the approach above are required. The first is that we must focus on one bidder at a time rather than treating them symmetrically. In particular, consider, without loss of generality, bidder 1. A test for the presence of common values for bidder 1 can be based on (say) the pseudo-value corresponding to $b_{1\tau}$, the empirical $\tau$-th quantile of the bids submitted by bidder 1 in $n$-bidder auctions:

$$\hat{v}_{\tau,n,1} \equiv \left( b_{1\tau} + \hat{G}_1 (b_{1\tau}) \right) \hat{g}_1 (b_{1\tau})$$

where $\hat{G}_1 (b)$ and $\hat{G}_1 (b)$ are nonparametric estimates analogous to those in equation (4) above. Under the PV hypothesis, the population analog, $v_{\tau,n,1}$, is constant across $n$ for all $\tau \in (0,1)$.

Under the CV alternative, in order to obtain the stochastic ordering

$$v_{\tau,2,1} \geq v_{\tau,3,1} \geq \cdots \geq v_{\tau,n,1}$$

we require the second modification: in considering auctions with $n = 2, 3, \ldots$, we construct a sequence of sets of opponents faces by bidder 1, e.g., \{bidder 2\}, \{bidder 2, bidder 3\}, \{bidder 2, bidder 3, bidder 4\}, etc. This structure ensures that the severity of the winner’s curse faced by bidder 1 is greater in auctions with larger number of participants, even though opponents are not perfect substitutes for each other. While constructing such a sequence for $n$ large would typically require a great deal of data, doing so for $n \in 2, 3$ (where the change in the severity of the winner’s curse is typically largest) will often be feasible.

References


A Proof of Theorem 2

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 notations that $G_n \equiv G_n (b; b)$, $g_n \equiv g_n (b; b)$, $\hat{G}_n \equiv \hat{G}_n (b; b) = \frac{1}{nT_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} 1 (b_{it} < b) K \left( \left( \frac{b-b_{it}}{b} \right) \right)$ and $\hat{g}_n \equiv \hat{g}_n (b; b) = \frac{1}{nT_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} K \left( \left( \frac{b-b_{it}}{b} \right) \right) K \left( \left( \frac{b-b_{it}}{b} \right) \right)$.

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

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

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

$$= \frac{\hat{G}_n - E\hat{G}_n}{g_n} + \frac{E\hat{G}_n - G_n}{g_n} (\hat{g}_n - E\hat{g}_n) - \frac{G_n}{g_n} (E\hat{g}_n - g_n) + o (\hat{G}_n - G_n) + o (\hat{g}_n - g_n)$$

Standard bias calculation for kernel estimation shows that by Assumption 4, both

$$|E\hat{G}_n - G_n| \leq \left| \int (G_n (b; b + uh) - G_n (b; b)) K (u) \, du \right| \leq C h^R \int |u|^R K (u) \, du = o \left( \frac{1}{\sqrt{T_n h^2}} \right)$$

and

$$|E\hat{g}_n - g_n| \leq \left| \int \int (g_n (b + uh; b + vh) - g_n (b; b)) K (u) K (v) \, dudv \right| \leq Ch^R = o \left( \frac{1}{\sqrt{T_n h^2}} \right).$$
Next it will be shown that
\[ \sqrt{T_n h^2} (\hat{g}_n - E\hat{g}_n) \xrightarrow{d} N \left( 0, \frac{1}{n} \left( \int \int K^2(x) K^2(y) \, dx \, dy \right) g_n(b; b) \right) \] (11)

For this purpose it suffices to show that
\[ \lim_{T_n \to \infty} \text{Var} \left( \sqrt{T_n h^2} (\hat{g}_n(b; b) - E\hat{g}_n(b; b)) \right) = \frac{1}{n} \left( \int \int K^2(x) K^2(y) \, dx \, dy \right) g_n(b; b) \]

This is verified by the following calculation.
\[ \text{Var} \left( \frac{1}{\sqrt{T_n h^2}} \cdot n \sum_{i=1}^{T_n} \left[ K \left( \frac{b_{it} - b}{h} \right) K \left( \frac{\hat{b}_{it} - b}{h} \right) \right] \right) = T_n \left( \frac{1}{T_n h^2} \text{Var} \left( \sum_{i=1}^{n} \left[ K \left( \frac{b_{it} - b}{h} \right) K \left( \frac{\hat{b}_{it} - b}{h} \right) \right] \right) \right) = \frac{1}{nh^2} \left\{ \text{Var} \left[ K \left( \frac{b_{it} - b}{h} \right) K \left( \frac{\hat{b}_{it} - b}{h} \right) \right] + (n - 1) \text{Cov} \left[ K \left( \frac{b_{it} - b}{h} \right), K \left( \frac{\hat{b}_{it} - b}{h} \right) \right] \right\} \]

It is a standard result that
\[ E \left( K \left( \frac{b_{it} - b}{h} \right), K \left( \frac{\hat{b}_{it} - b}{h} \right) \right) = O(h^2) \]
and it can be verified that
\[ E \left[ K \left( \frac{b_{it} - b}{h} \right) K \left( \frac{\hat{b}_{it} - b}{h} \right) K \left( \frac{\hat{b}_{it} - b}{h} \right) \right] = O(h^4) . \]

Therefore we can write
\[ \text{Var} \left( \sqrt{T_n h^2} (\hat{g}_n(b; b) - E\hat{g}_n(b; b)) \right) = \frac{1}{n} E \left[ K^2 \left( \frac{b_{it} - b}{h} \right) K^2 \left( \frac{\hat{b}_{it} - b}{h} \right) \right] + O(h^4) \]
\[ = \frac{1}{n} \int \int \frac{1}{h^2} K^2 \left( \frac{u - b}{h} \right) K^2 \left( \frac{v - b}{h} \right) g_n(u, v) \, du \, dv + O(h^4) \]
\[ = \frac{1}{n} \left( \int \int K^2(x) K^2(y) \, dx \, dy \right) g_n(b; b) + o(1) \]
where the last equality uses the substitutions \( x = (u - b) / h \) and \( y = (v - b) / h \). Finally the same variance calculation shows that
\[ \text{Var} \left( \sqrt{T_n h^2} \left( \hat{G}_n - E\hat{G}_n \right) \right) \to 0. \]

Hence the proof for part 2 is complete.

3. Since the sample quantiles of the bid distribution converge at rate \( T_n \) to the population quantile, which is faster than the convergence rate for the estimated pseudo-values, for large \( T_n \) the sampling error in the \( \tau \)th quantile of the bid distribution does not affect the large sample property of the nonparametric test, for each \( \tau \in \{ \tau_1, \ldots, \tau_l \} \):
\[ (\hat{b}_{\tau_1, n}, \ldots, \hat{b}_{\tau_l, n}) - \hat{\nu} (\bar{x}_{\tau}, n) = O_p \left( \frac{1}{\sqrt{T_n}} \right) = o_p \left( \frac{1}{T_n h^2} \right) , \]
(12)
A formal proof would proceed using uniform convergence of the kernel estimates of $G_n(b;b)$ and $g_n(b;b)$ and their derivatives, and stochastic equicontinuity arguments (see for example Andrews (1994) and Pollard (1984)). We omit these rather tedious technical details and proceed by noting that (12) implies that the limiting distribution of

$$
\sqrt{T_n h^2} \left( \xi \left( \hat{b}_{\tau_1}; n \right) - v \left( F_x^{-1}(\tau_1), F_x^{-1}(\tau_2), n \right) \right) \quad \tau = \{\tau_1, \ldots, \tau_L\}
$$

is the same as the limiting distribution of the vector

$$
\sqrt{T_n h^2} \left( \xi \left( \hat{s}_n (x_\tau); n \right) - v (x_\tau, x_\tau, n) \right) \quad \tau = \{\tau_1, \ldots, \tau_L\}
$$

In part 2 we showed that each element of this vector is asymptotically normal with limit variance given by the diagonal element of $\Omega$. It remains to show that the off-diagonal elements of $\Omega$ are 0. For this purpose it suffices to show, using a standard result from kernel estimation, that nonparametric 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. In other words,

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

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

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

This is completed in the following calculation:

$$
\text{Cov} \left[ \frac{1}{\sqrt{T_n h^2 n}} \sum_{i=1}^{T_n} \sum_{i=1}^{n} K \left( \frac{b_{it} - b_{\tau}}{h} \right) K \left( \frac{b_{it}^* - b_\tau}{h} \right), \frac{1}{\sqrt{T_n h^2 n}} \sum_{i=1}^{T_n} \sum_{i=1}^{n} K \left( \frac{b_{it} - b_{\tau'}}{h} \right) K \left( \frac{b_{it}^* - b_{\tau'}}{h} \right) \right] = \frac{1}{n^2 h^2} \text{Cov} \left[ \sum_{i=1}^{n} K \left( \frac{b_{it} - b_{\tau}}{h} \right) K \left( \frac{b_{it}^* - b_\tau}{h} \right), \sum_{i=1}^{n} K \left( \frac{b_{it} - b_{\tau'}}{h} \right) K \left( \frac{b_{it}^* - b_{\tau'}}{h} \right) \right].
$$

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 $i \neq j$,

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

we can further rewrite the covariance function as

$$
\frac{1}{n^2 h^2} \sum_{i=1}^{n} \sum_{j=1}^{n} E \left[ K \left( \frac{b_{it} - b_\tau}{h} \right) K \left( \frac{b_{it}^* - b_\tau}{h} \right) K \left( \frac{b_{jt} - b_{\tau'}}{h} \right) K \left( \frac{b_{jt}^* - b_{\tau'}}{h} \right) \right] + O \left( h^2 \right)
$$

$$
= \frac{1}{n^2 h^2} \sum_{i=1}^{n} E \left[ K \left( \frac{b_{it} - b_\tau}{h} \right) K \left( \frac{b_{it}^* - b_\tau}{h} \right) K \left( \frac{b_{it} - b_{\tau'}}{h} \right) K \left( \frac{b_{it}^* - b_{\tau'}}{h} \right) \right] + O \left( h^2 \right)
$$

$$
= \frac{1}{n} \int \int K(x) K(y) \left( x + \frac{b - b_{\tau'}}{h} \right) K \left( y + \frac{b - b_{\tau'}}{h} \right) g_n(b + x h, b + y h) dx dy + O \left( h^2 \right) \to 0.
$$
B Proof of Theorem 3

Assumption 4 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)).

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

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

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 hand side of part (ii) as

$$\sqrt{T_n h} \left(\tilde{\mu}_{n,\tau} - E\left[\xi(b;n) 1(b_{n,\tau} \leq b \leq b_{n,\tau})\right]\right)$$

$$= \sqrt{T_n h} \left(\frac{1}{T_n n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \left(\frac{G_n(b_{i,t};b_{i,t})}{g_n(b_{i,t};b_{i,t})} - \frac{G_n(b_{i,t};b_{i,t})}{g_n(b_{i,t};b_{i,t})}\right) 1\left(b_{n,\tau} \leq b_{i,t} \leq b_{n,\tau}\right) - 1\left(b_{n,\tau} \leq b_{i,t} \leq b_{n,\tau}\right)\right)$$

where

$$\tilde{\mu}_{n,\tau} = \sqrt{T_n h} \left(\frac{1}{T_n n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \left(\frac{G_n(b_{i,t};b_{i,t})}{g_n(b_{i,t};b_{i,t})} - \frac{G_n(b_{i,t};b_{i,t})}{g_n(b_{i,t};b_{i,t})}\right) 1\left(b_{n,\tau} \leq b_{i,t} \leq b_{n,\tau}\right) - 1\left(b_{n,\tau} \leq b_{i,t} \leq b_{n,\tau}\right)\right)$$

The function in the summand of $\tilde{\mu}_{n,\tau}$ satisfies stochastic equicontinuity conditions (a type I function of Andrews (1994)), hence using the parametric convergence rates of $b_{n}$ and $b_{n,\tau}$,

$$\tilde{\mu}_{n,\tau} = \sqrt{T_n h} \left(E_n \xi(b;n) 1\left(b_{n,\tau} \leq b \leq b_{n,\tau}\right) - E_n \xi(b;n) 1\left(b_{n,\tau} \leq b \leq b_{n,\tau}\right)\right) + O_p(1)$$

$$= C \sqrt{T_n h} \left(O\left(b_{n,\tau} - b_{n,\tau}\right) + O\left(b_{n,\tau} - b_{n,\tau}\right)\right) + O_p(1) = \sqrt{T_n h} O_p\left(\frac{1}{\sqrt{T_n}}\right) + O_p(1) = O_p(1)$$
Similarly, the function in the summand of \( \mu_{n,r}^4 \) also satisfies stochastic equicontinuity conditions (product of type I and type III functions in Andrews (1994)), and hence

\[
\hat{\mu}_{n,r}^4 = \sqrt{T_n h} \hat{E}_b \left( \frac{\hat{G}_n (b; b)}{g_n (b; b)} - \frac{G_n (b; b)}{g_n (b; b)} \right) \left( 1 \left( b_r \leq b \leq \hat{b}_{1 - r} \right) - 1 \left( b_{\gamma,n} \leq b \leq b_{1 - r,n} \right) \right) + o_p(1)
\]

\[
= o_p \left( \sup_{b \in [b_{\gamma,n}, b_{1 - r,n}]} \left| \frac{\hat{G}_n (b; b)}{g_n (b; b)} - \frac{G_n (b; b)}{g_n (b; b)} \right| \right) \sqrt{T_n h} \left( O \left( \hat{b}_{r,n} - b_{r,n} \right) + O \left( b_{1 - r,n} - b_{1 - r,n} \right) \right) + o_p(1)
\]

\[
= o_p(1) \sqrt{T_n h} O_p \left( \frac{1}{\sqrt{T_n h}} \right) + o_p(1) = o_p(1)
\]

Combining the above results of rates of convergence, we have shown that

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

\( \hat{\mu}_{n,r}^2 \) can be further decomposed using a second order Taylor expansion:

\[
\hat{\mu}_{n,r}^2 = \hat{\mu}_{n,r}^5 + \hat{\mu}_{n,r}^6 + \hat{\mu}_{n,r}^7
\]

where

\[
\hat{\mu}_{n,r}^5 = \sqrt{T_n h} \frac{1}{n T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \frac{1}{g_n (b_{it}; b_{it})} \left( \hat{G}_n (b_{it}; b_{it}) - G_n (b_{it}; b_{it}) \right) 1 (b_{r,n} \leq b_{it} \leq b_{1 - r,n})
\]

\[
\hat{\mu}_{n,r}^6 = -\sqrt{T_n h} \frac{1}{n T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \frac{G_n (b_{it}; b_{it})}{g_n (b_{it}; b_{it})^2} (\hat{g}_n (b_{it}; b_{it}) - g_n (b_{it}; b_{it})) 1 (b_{r,n} \leq b_{it} \leq b_{1 - r,n})
\]

\[
\hat{\mu}_{n,r}^7 = \sqrt{T_n h} \frac{1}{n T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \hat{h}^1_{n} (b_{it}) \left( \hat{G}_n (b_{it}; b_{it}) - G_n (b_{it}; b_{it}) \right)^2 1 (b_{r,n} \leq b_{it} \leq b_{1 - r,n})
\]

\[
+ \sqrt{T_n h} \frac{1}{n T_n} \sum_{t=1}^{T_n} \sum_{i=1}^{n} \hat{h}^2_{n} (b_{it}) (\hat{g}_n (b_{it}; b_{it}) - g_n (b_{it}; b_{it}))^2 1 (b_{r,n} \leq b_{it} \leq b_{1 - r,n})
\]

where \( \hat{h}^1_{n} (\cdot) \) and \( \hat{h}^2_{n} (\cdot) \) are the second derivatives with respect to \( G_n (\cdot) \) and \( g_n (\cdot) \) evaluated at some mean values between \( \hat{G}_n (\cdot) \) and \( G_n (\cdot) \) and between \( \hat{g}_n (\cdot) \) and \( g_n (\cdot) \). We first bound \( \hat{\mu}_{n,r}^7 \) by the uniform convergence rates of \( G_n (\cdot) \) and \( \hat{g}_n (\cdot) \):

\[
\left| \hat{\mu}_{n,r}^7 \right| \leq C \sqrt{T_n h} \left( O_p \left( \frac{\log T}{T_n h} + h^{2R} \right) + O_p \left( \frac{\log T}{T_n h^2} + h^{2R} \right) \right)
\]

\[
= O_p \left( \frac{\log T}{\sqrt{T_n h}} + \frac{\log T}{\sqrt{T_n h^3}} + \sqrt{T_n h^{1+4R}} \right) = o_p(1)
\]
Consider
\[ \hat{\mu}_{n,t}^8 = -\sqrt{T_n h} \frac{1}{n T_n} \sum_{t=1}^{T_n} \frac{G_n(b_{i,t};b_{i,t})}{g_n(b_{i,t};b_{i,t})} (\hat{g}_n(b_{i,t};b_{i,t}) - E[\hat{g}_n(b_{i,t};b_{i,t})] 1(b_{r,n} \leq b_{i,t} \leq b_{1-r,n})) \]

\[ - \sqrt{T_n h} \frac{1}{n T_n} \sum_{t=1}^{T_n} \frac{G_n(b_{i,t};b_{i,t})}{g_n(b_{i,t};b_{i,t})} (E[\hat{g}_n(b_{i,t};b_{i,t})] - g_n(b_{i,t};b_{i,t})) 1(b_{r,n} \leq b_{i,t} \leq b_{1-r,n}) \]

\[ = - \sqrt{T_n h} \frac{1}{n T_n} \sum_{t=1}^{T_n} \frac{G_n(b_{i,t};b_{i,t})}{g_n(b_{i,t};b_{i,t})} (\hat{g}_n(b_{i,t};b_{i,t}) - E[\hat{g}_n(b_{i,t};b_{i,t})] 1(b_{r,n} \leq b_{i,t} \leq b_{1-r,n})) + o_p(1) \]

\[ \equiv \hat{\mu}_{n,r}^8 + o_p(1) \]

because by assumption the bias in the second term is of the order
\[ \sqrt{T_n h} O \left( h^R \right) = O \left( \sqrt{T_n h^{1+2R}} \right) = o(1) \]

Next we show that
\[ \hat{\mu}_{n,r}^8 \overset{d}{\to} N \left( 0, \Omega = \left[ \int \left( \int K(v) K(u+v) dv \right)^2 du \right] \left[ \frac{1}{n} \int F_{n,1}^{-1}(1-r) \frac{G_n(b; b)^2}{g_n(b; b)^2} g_n(b)^2 db \right] \right) \]

This follows from U-statistics limit variance calculation. We can write
\[ \hat{\mu}_{n,r}^8 = \sqrt{T_n h} \frac{1}{n^2 T_n^2} \sum_{t=1}^{T_n} \sum_{s=1}^{T_n} m(w_t, w_s) \]

where
\[ m(w_t, w_s) = \sum_{i=1}^{n} \sum_{j=1}^{n} G_n(b_{i,t};b_{i,s}) \left[ \frac{1}{h^2} K \left( \frac{b_{i,j} - b_{i,t}}{h} \right) K \left( \frac{b_{i,j} - b_{i,s}}{h} \right) - \frac{1}{h^2} K \left( \frac{b_{i,j} - b_{i,t}}{h} \right) K \left( \frac{b_{i,j} - b_{i,s}}{h} \right) \right] 1(b_{r,n} \leq b_{i,t} \leq b_{1-r,n}) \]

Using lemma 8.4 of Newey and McFadden (1994), we can routinely verify that
\[ \sqrt{T_n h} E \left[ m(w_t, w_t) \right] = O_p \left( \frac{1}{T_n h} \right) = O_p \left( \frac{1}{\sqrt{T_n h}} \right) = o_p(1) \]

and
\[ \sqrt{T_n h} \frac{E m(w_t, w_s)^2}{T_n} = O_p \left( \frac{1}{T_n h} \right) = O_p \left( \frac{1}{\sqrt{T_n h^2}} \right) = o_p(1) \]

It then follows from lemma 8.4 of Newey and McFadden (1994) that
\[ \hat{\mu}_{n,r}^8 = \sqrt{T_n h} \frac{1}{n^2 T_n} \sum_{t=1}^{T_n} \left[ E( m(w_t, w_s) | w_t) + E( m(w_t, w_s) | w_s) \right] + o_p(1) \]

The first term is asymptotically negligible, since
\[ \sqrt{T_n h} \frac{1}{n^2 T_n} \sum_{t=1}^{T_n} E( m(w_t, w_s) | w_t) \]

\[ = \sqrt{T_n h} \frac{1}{n^2 T_n} \sum_{t=1}^{T_n} \left[ g_n(b_{i,t};b_{i,t}) 1(b_{r,n} \leq b_{i,t} \leq b_{1-r,n}) - E[ g_n(b_{i,t};b_{i,t}) 1(b_{r,n} \leq b_{i,t} \leq b_{1-r,n})] \right] + O \left( h^R \right) \]

\[ = \sqrt{T_n h} O_p \left( \frac{1}{\sqrt{T_n}} \right) + O \left( \sqrt{T_n h^{1+2R}} \right) = o_p(1) \]
It remains only to verify by straightforward though somewhat tedious calculation that

\[
\text{Var} \left( \frac{1}{n^2 \bar{T}_n} \frac{1}{T_n} \sum_{i=1}^{T_n} E \left( m(w_i, w_s) | w_s \right) \right)
\]

\[
= h \text{Var} \left( \frac{1}{n} \sum_{j=1}^{n} \int_{b_{j-1}}^{b_j} G_n(b; b) \frac{1}{h^2} K \left( \frac{b_j - b}{h} \right) K \left( \frac{b_j - b}{h} \right) g_1(b) db \right)
\]

\[
= h \frac{1}{n} \text{Var} \left( \int_{b_{j-1}}^{b_j} G_n(b; b) \frac{1}{h^2} K \left( \frac{b_j - b}{h} \right) K \left( \frac{b_j - b}{h} \right) g_1(b) db \right) + o(1)
\]

\[
= h \frac{1}{n} E \left( \int_{b_{j-1}}^{b_j} G_n(b; b) \frac{1}{h^2} K \left( \frac{b_j - b}{h} \right) K \left( \frac{b_j - b}{h} \right) g_1(b) db \right)^2 + o(1)
\]

\[
\rightarrow \Omega \equiv \left[ \int \left( \int K(u) K(u + v) dv \right)^2 du \right] \left[ \frac{1}{n} \int \frac{G_{n-1}(1-\tau)}{G_n(b; b) g_n^2(b)} db \right]
\]

Finally, we note that if we apply the same calculation for \( \hat{\mu}_6 \) to \( \hat{\mu}_5 \), then we can verify that

\[
E \left[ \hat{\mu}_5 \right] = o(1) \quad \text{and} \quad \text{Var} \left( \hat{\mu}_5 \right) = o(1)
\]

which then implies that

\[
\hat{\mu}_5 \xrightarrow{p} 0
\]

The proof is now completed by putting these terms together. □