A Simple Test for Moment Inequality Models with an Application to English Auctions *

Andrés Aradillas-López, Amit Gandhi, Daniel Quint

February 16, 2014

Abstract

Testable predictions of many standard economic models take the form of inequality comparisons between transformations of nonparametric conditional moments. Notable among these examples are implications about the statistical interdependence of bidders' valuations in ascending auctions. Motivated by these models, we propose a novel econometric test of these types of restrictions and we describe its properties. Our approach relies on a particular type of one-sided L^p statistic. Unlike existing comparable methods that also rely on one-sided L^p functionals, our approach is not based on so-called least favorable configurations and is therefore less conservative. And unlike methods that require approximating the suprema of a test-statistic, our method remains computationally straightforward to implement even in the presence of a rich set of continuous conditioning covariates. Going back to our motivating example, we introduce a new specification test for standard models of ascending auctions. We apply our econometric test to data from much-studied United States Forest Service timber auctions, and find clear evidence to reject the Independent Private Values model in favor of a model of correlated private values.

Keywords: Ascending auctions, independent private values, conditional moment inequalities.

*A previous version of this paper was titled "Testing Auction Models: Are Private Values Really Independent?" Aradillas-López is grateful to the Human Capital Foundation (www.hcfoundation.ru), and especially Andrey P. Vavilov, for support of the Department of Economics, the Center for the Study of Auctions, Procurements, and Competition Policy (CAPCP, http://capcp.psu.edu/), and the Center for Research in International Financial and Energy Security (CRIFES, http://crifes.psu.edu/) at Penn State University

1 Introduction

Testable implications of economic models often involve restrictions on moments identified in the data. Econometric tests for these restrictions effectively become specification tests for the underlying model. In this paper we study models whose restrictions involve inequalities between nonlinear transformations of conditional moments, and develop a computationally simple econometric methodology to test such restrictions. Our main motivating example are testable implications concerning the interdependence of bidders' valuations in ascending auctions. However, our testing framework is applicable to a variety of empirical models that are prominent in the literature but have not been studied in a unified way, such as stoachastic dominance tests, conditional covariance tests, and tests of other relationships that emerge from models of entry games, and social interaction.

The literature on testing and inferential methods involving some form of conditional moment inequalities has grown in the recent past. Some examples include, among others, Ghosal, Sen, and Vaart (2000), Barrett and Donald (2003), Hall and Yatchew (2005), Lee, Linton, and Whang (2009), Andrews and Shi (2013, 2011), Chernozhukov, Lee, and Rosen (2013), Lee, Song, and Whang (2013), Ponomareva (2010), Kim (2009), Menzel (2011), Armstrong (2011a, 2011b), and Chetverikov (2012). Some of these methods are not designed to test inequalities involving general nonlinear *transformations* of conditional moments and therefore do not include the general problem we study within their scope. Existing methods that can be adapted to handle our problem can be broadly classified into those that rely on one-sided L^p functionals (e.g, Lee, Song, and Whang (2013)) and those that rely on the suprema over a collection of test-statistics (notably, Chernozhukov, Lee, and Rosen (2013)). Our method is a special case of a one-sided L^p functional; however unlike existing tests of this type it does not rely on *least favorable* configurations, making our procedure less conservative. Compared with procedures that rely on finding the suprema over a collection of test-statistics, our method remains computationally straightforward to implement even in the presence of a rich set of continuous conditioning covariates[1]. Our method first expresses the inequality restrictions as an *unconditional mean-zero condition* and then develops a test for this mean-zero condition. In the end, our test statistic involves sample averages; this makes it asymptotically pivotal under certain conditions (which we describe) and therefore straightforward to implement computationally.

[1] A more detailed discussion is included in Section 3.9.

We apply our approach to test the Independent Private Values (IPV) assumption widely used in empirical studies of auctions. Whether the valuations of bidders in an auction are independent or correlated has significant policy implications.[2] Furthermore, in either a first-price or English auction, non-parametric identification of the model, and the choice of an empirical strategy, depend on whether or not values are independent (see, e.g., Athey and Haile (2007)). When values are correlated, policy implications drawn from an IPV-based empirical approach can be misleading; thus, if the tests presented in this paper lead to rejection of the IPV model, an empirical strategy that allows for correlation is necessary.[3] Within the general framework presented in this paper, we derive testable implications of both IPV and correlated private values models of bidding in English auction, and apply these tests to data from USFS timber auctions, which has been widely studied in the literature under the assumption of IPV. As we discuss later, the USFS timber data has several features that argue in favor of the plausibility of a private values model, and a rich vector covariate information available about each auction; conditioning on such detailed covariates is often used to justify the assumption that any remaining private-value differences are independent. We nonetheless find clear evidence to reject independence of values in favor of correlation.

The rest of the paper proceeds as follows. Section 2 describes the structure of our general setup, and gives examples of several economic applications that would fit it. Section 3 describes our econometric testing strategy: we show how testable restrictions from a theoretical model are transformed into unconditional moment equalities, describe our proposed test, and establish its asymptotic properties. We also describe how our approach fits within and contributes to the existing literature. Section 4 is devoted to our motivating economic example: testing standard assumptions used in empirical modeling of English auctions. We establish testable implications of IPV and positively-correlated private values under different models of bidding in English auctions, and show how each model can be expressed as a case of our general econometric setup. We then perform Monte Carlo simulations, using

[2]If private values are independently and symmetrically distributed, a classic result in auction theory is that the optimal selling mechanism takes the form of a standard auction in which the only relevant design parameter is the reserve price. IPV also implies revenue-equivalence of the two most prevalent auction formats, first-price and English auctions; optimality of a reserve price strictly higher than the seller's valuation; and invariance of that optimal reserve price to the number of bidders present.

[3]Identification and estimation in auctions with correlated values has been studied by Li, Perrigne, and Vuong (2002), Krasnokutskaya (2011), and Hu, McAdams, and Shum (2013) in the case of first-price auctions, and Aradillas-López, Gandhi, and Quint (2013) in the case of ascending auctions.

simulated data to study the finite-sample properties of the proposed test; and finally, we apply our tests to actual USFS timber data and present our results. Section 5 concludes. Appendix A is devoted to the econometric proofs and details about the constructions of our proposed tests. Appendix B contains proofs of our auction model results.

2 General setup

Here we describe the generic structure of the type of econometric model we study in this paper. It is designed explicitly with our ascending auction economic motivation in mind, but it also includes other examples as special cases. The setup we consider includes the following components.

2.1 Variables

The variables observed in the data can be classified into the following categories:

(i) **Outcome variables:** Denoted by $Y \in \mathbb{R}^{d_y}$, these are quantities under the control of the decision-makers in the model, such as bids in an auction. Economic theory will characterize how these decisions are made.

(ii) **Conditioning variable of interest:** Denoted by $N \in \mathbb{R}$, this is the key explanatory variable whose relationship with Y is predicted by theory and which we will be testing.

(iii) **Control variables:** Denoted by $X \in \mathbb{R}^{d_x}$, these are other observable characteristics of the environment on which we will be conditioning.

At a high level, we will be testing the relationship between N and Y, while holding X fixed. The variable N could be allowed to be multidimensional, but we focus on the case $N \in \mathbb{R}$ because it corresponds to the main application we study here, where N will denote the number of bidders in an auction and X will denote all other observable details of the auction.

In addition to the variables observed in the data, we will sometimes make use of **index variables**, denoted by $Z \in \mathbb{R}^{d_z}$, which appear when the theoretical predictions being tested are properties which must hold over a range of values. (For example, a test of first-order stochastic dominance can be thought of as a test that the relationship $F_1(z) \leq F_2(z)$ holds over a range of values z.)

2.2 Structural functions

The next component of the models we consider is a *known*, vector-valued function of the variables above. This function will be denoted by $S(y, x, z, n) \in \mathbb{R}^{d_s}$, and its expected value over y, conditional on (x, z, n), by

$$s(x, z, n) = \int S(y, x, z, n) dF_{Y|X,N}(y|x, n) = E_{Y|X,N}\left[S(Y, x, z, n)\,\middle|X = x, N = n\right].$$

(Throughout the paper, F_ξ will refer to the marginal distribution of a random variable ξ, and $F_{\xi|\eta}$ the conditional distribution of ξ given η, with f_ξ and $f_{\xi|\eta}$ the respective densities.) This is a vector of conditional moments, and its construction highlights the different roles played by the control and the index variables: the former (X) affect the distribution of observed outcomes Y, while the latter (Z) enters into s only as a direct argument in S. The choice of the function S will depend on the economic model being tested.

2.3 Transformations

For each pair $(n, n') \in \text{Supp}(N) \times \text{Supp}(N)$, the model produces a finite collection of $Q_{n,n'} \geq 1$ *known* real-valued transformations $\{m^q\}$. These depend on the pair (n, n') in question and on the conditional moments $s(\cdot, n)$ and $s(\cdot, n')$. We will abbreviate

$$R^q(X, Z; n, n') = m^q(s(X, Z, n),\ s(X, Z, n');\ n, n') \in \mathbb{R}.$$

The models we consider have predictions of the type

$$\forall\ n, n' \in \text{Supp}(N),\ Pr\left(R^q(X, Z; n, n') \leq 0\right) = 1 \quad \text{for} \quad q = 1, \ldots Q_{n,n'}. \tag{1}$$

2.4 Examples

Next, we illustrate the use of our framework by showing how it would apply to several standard testing problems. First-order stochastic dominance arises in welfare economics (comparing income distributions), as well as in tests of treatment effects. Second-order stochastic dominance arises in models of portfolio choice and uncertainty, and in tests of differential private information. Conditional covariance relationships arise in many standard models of asymmetric information, as well as in certain models of collusive behavior in auctions. Other tests could be used for each of these particular problems; our point here is to illustrate the elements of our framework. (At the start of the next section, we will explicitly compare our approach to other methods in the literature.)

Example 1: First order stochastic dominance

Suppose Y is real-valued, and that the economic model predicts the first-order stochastic dominance relation $F_{Y|X,N}(\cdot|x,n) \succsim_{FOSD} F_{Y|X,N}(\cdot|x,n')$ for any x whenever $n > n'$. Thus, the model predicts

$$n > n' \implies F_{Y|X,N}(z|x,n) \leq F_{Y|X,N}(z|x,n') \text{ for any } (x,z).$$

This can be written as an instance of our general setup by letting $S(y,x,z,n) = \mathbb{1}\{y \leq z\}$, so that $s(x,z,n) = F_{Y|X,N}(z|x,n)$, and using the single transformation m for each n, n'

$$m\left(s(x,z,n); s(x,z,n'); n, n'\right) = \left(s(x,z,n) - s(x,z,n')\right) \cdot \mathbb{1}\{n > n'\}.$$

Example 2: Second order stochastic dominance

Next, consider testing whether $F_{Y|X,N}(\cdot|x,n) \succsim_{SOSD} F_{Y|X,N}(\cdot|x,n')$ whenever $n > n'$, where SOSD denotes second-order stochastic dominance. This requires that

$$n > n' \implies \int_{-\infty}^{z} F_{Y|X,N}(v|x,n)dv \leq \int_{-\infty}^{z} F_{Y|X,N}(v|x,n')dv \quad \text{for any } (x,z).$$

This too is an instance of our general setup, with $S(y,x,z,n) = \max\{z - y,\, 0\}$. To see why, note that

$$\int_{-\infty}^{z} \mathbb{1}\{Y \leq v\}\, dv = \max\{z - Y,\, 0\},$$

and so for a given (x,z,n),

$$\begin{aligned}
s(x,z,n) &= E_{Y|X,N}\left[S(Y,x,z,n)|X=x, N=n\right] = E_{Y|X,N}\left[\max\{z-Y,\, 0\}\,\middle|X=x, N=n\right] \\
&= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z} \mathbb{1}\{y \leq v\}\, dv\right) f_{Y|X,N}(y|x,n)dy \\
&= \int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} \mathbb{1}\{y \leq v\}\, f_{Y|X,N}(y|x,n)dy\right) dv \\
&= \int_{-\infty}^{z} F_{Y|X,N}(v|x,n)dv.
\end{aligned}$$

Once again, we would have $Q_{n,n'} = 1$ and

$$m\left(s(x,z,n); s(x,z,n'); n, n'\right) = \left(s(x,z,n) - s(x,z,n')\right) \cdot \mathbb{1}\{n > n'\}.$$

Example 3: Covariance Restrictions

Third, let Y_1 and Y_2 denote two real-valued outcome variables, and X a vector of controls. A standard conditional covariance test would examine the sign of $Cov(Y_1, Y_2|X)$. However, we want to allow for the possibility that Y_1, while fully-ordered, might not have a natural scale. For example, we might test for adverse selection by letting Y_1 denote a consumer's choice from several available insurance contracts, and Y_2 be an ex post measure of risk such as whether an accident occurred or the size of the claim. The contracts being considered might be naturally ordered from least to greatest coverage, but not on a natural linear scale. Thus, to be more general, we test whether for any z in the set of contracts and for any x,

$$Cov\left(\mathbb{1}\left\{Y^1 \leq z\right\}, Y_2 \middle| X\right) \quad \geq \quad 0 \tag{2}$$

If $Y_2 = -A$, where A is the binary indicator variable that equals 1 if the agent incurs an accident (requiring them to exercise their insurance contract), then the "positive correlation" test proposed by Chiappori and Salanie (2000) and Chiappori, Jullien, Salanié, and Salanie (2006) for moral hazard and adverse selection in insurance markets takes this form. While there has been much interest in the application of this testable restriction in the literature, almost all empirical work has focused on simple parametric specifications, such as a bivariate probit with linear parameters on X. However, such a restrictive specification can make a negative result (failing to find the predicted positive correlation) difficult to interpret, since it could follow from the "wrong" choice of parametrization; we are not aware of any work that tests the restriction in its full nonparametric form above.

Covariance restrictions also emerge as tests of strategic interaction models, such as entry games. Let Y_1 denote the action of one player i, and $Y_2 = \eta(Y_{-i})$ be a real-valued aggregate measure of the remaining players' actions Y_{-i}. Then a testable implication of strategic interaction models shown by de Paula and Tang (2012) for binary action games, and generalized to an ordered set of actions by Aradillas-López and Gandhi (2011), takes the form of the covariance restriction (2). As a final example of the use of covariance tests, several recent papers have shown that the positive correlation of entry decisions among potential bidders in an auction is a testable implication of various hypotheses about bidding behavior. Li and Zhang (2010) propose such a test of endogenous entry in the standard affiliated signals model for first price auctions; Haile, Hendricks, Porter, and Onuma (2012) and Conley and Decarolis (2013) propose such a covariance test to detect collusive bidding behavior in different institutional settings.

To test this within our framework, note that (2) can be re-expressed as the restriction

$$E\left[\mathbb{1}\left\{Y^1 \leq z\right\} \middle| X = x\right] \cdot E\left[Y_2 \middle| X = x\right] - E\left[\mathbb{1}\left\{Y \leq z\right\} \cdot Y_2 \middle| X = x\right] \leq 0 \text{ for any } (x, z).$$

In this model, there is no variable playing the role of N, so this model fits the special case (1′) discussed below. Let Z denote the index variable chosen for z. The structural functions are

$$S(Y,z) = \Big(\mathbb{1}\{Y_1 \leq z\},\, Y_2,\, \mathbb{1}\{Y_1 \leq z\} \cdot Y_2\Big),$$

and

$$\begin{aligned} s(x,z) &= E\left[S(Y,z)\,\middle|X = x\right] \\ &= \Big(E\left[\mathbb{1}\{Y_1 \leq z\}\,\middle|X = x\right],\, E\left[Y_2\middle|X = x\right],\, E\left[\mathbb{1}\{Y_1 \leq z\} \cdot Y_2\middle|X = x\right]\Big). \end{aligned}$$

The transformation m is then given by

$$m(s(x,z)) = E\left[\mathbb{1}\{Y_1 \leq z\}\,\middle|X\right] \cdot E\left[Y_2\middle|X\right] - E\left[\mathbb{1}\{Y_1 \leq z\} \cdot Y_2\middle|X\right].$$

Example 4: Conditional Moment Inequality Models

A special case of our general framework is when there is no N or Z variable, so $S(y,x,z) = S(y,x)$. In this case, the vector-valued transformations are of the form $S(y,x) \in \mathbb{R}^{d_s}$ and

$$s(x) = \int S(y,x)dF_{Y|X}(y|x) = E_{Y|X}\left[S(Y,x)\,\middle|X = x\right].$$

In this case the model consists of $q = 1,\ldots,Q$ transformations $\{m^q\}$. The arguments of each transformation m^q are now simply $s(x)$, with

$$R^q(x) = m^q(s(x)).$$

These models predict

$$Pr(R^q(X) \leq 0) = 1, \quad q = 1,\ldots Q. \tag{1′}$$

3 An econometric test for our general model

The restrictions in (1) involve inequalities between possibly *nonlinear* transformations of conditional moments. This puts them outside the scope of a number of existing approaches to testing moment inequalities; particularly those that rely on a space of so-called instrument functions (e.g, Andrews and Shi (2011, 2013), Armstrong (2011a, 2011b)). Our test focuses on a special type of one-sided L^p statistic and relies on nonparametric estimators of the conditional moments involved. Unlike other existing tests that also rely on these types of functionals (e.g, Lee, Song, and Whang (2013)), our procedure does not rely on a *least favorable* configuration and is therefore less conservative. Unlike procedures based on approximating the suprema of a test-statistic (e.g, Chernozhukov, Lee, and Rosen (2013)), our test does not become computationally costly to implement when the conditioning variables include a large collection of covariates with rich support. A more detailed discussion of where our test fits and how it contributes to the existing literature is included in Section 3.9 below.

Now we develop a test for the general model described in (1). Our procedure also covers the special case described in (1′).

3.1 Expressing (1) as an unconditional mean-zero condition

We will assume that N has discrete support, but our econometric approach can be extended to cases where N is continuous. X will be allowed to include discrete and/or continuously distributed covariates. We group

$$W \equiv (X, Z) \quad \text{and} \quad U \equiv (Y, X, N, Z)\,.$$

To simplify exposition, we will assume that the support of $W = (X, Z)$ does not vary with N, i.e., for all $n, n' \in \text{Supp}(N)$,

$$\text{Supp}\left(W \middle| N = n\right) = \text{Supp}\left(W \middle| N = n'\right) = \text{Supp}(W) \tag{3}$$

although this assumption can be relaxed. For each $n, n' \in \text{Supp}(N)$ and $q = 1, \ldots, Q_{n,n'}$ consider the unconditional expectation

$$\mathcal{T}^q_{n,n'} = E\left[\max\left\{R^q(W; n, n'),\, 0\right\}\right]$$

and note that this expectation is nonnegative, and it is zero *if and only if* (1) holds for this n, n' and q. Nonnegativity allows us to combine all restrictions into

$$\mathcal{T} = \sum_{n,n' \in \text{Supp}(N)} \sum_{q=1}^{Q_{n,n'}} \mathcal{T}^q_{n,n'}.$$

We then have $\mathcal{T} \geq 0$, with $\mathcal{T} = 0$ if and only if (1) holds.

3.2 Choosing a testing range

Let $\mathcal{N} \subseteq Supp(N)$ denotes the range of values we consider for N. We separate X into its continuous variables X^c and its discrete variables X^d, and assume X^c has an absolutely continuous distribution with respect to the Lebesgue measure; we denote the dimension of X^c as r. Our test will estimate $\mathcal{T}$ by replacing $R^q(w;n,n')$ with a nonparametric estimator $\widehat{R}^q(w;n,n')$. One of the main requirements for our asymptotic theory is that these nonparametric estimators satisfy certain asymptotic properties *uniformly* over the range of values of $(x,z) \equiv w$ considered. For this reason, we will specify a *testing range* $\mathcal{W} \subset Supp(W)$ such that $(x,z) \in \mathcal{W}$ implies $x^c \in \text{int}(\text{Supp}(X^c))$. We will also assume that $\mathcal{W}$ is such that

$$f_{X,N}(x,n) \geq \underline{f} > 0 \quad \forall\ (x,z) \in \mathcal{W},\ n \in \mathcal{N},$$

where $f_{X,N}$ is the joint distribution of X and N. Note that this condition would be automatically satisfied, e.g., if (X,N) have bounded support and their joint density is bounded away from zero everywhere. We will focus on the expectations $\mathcal{T}^q_{n,n'}$ conditional on $W \in \mathcal{W}$. From now on we will re-define

$$\mathcal{T}^q_{n,n'} = E\Big[\max\big\{R^q(W;n,n'),\,0\big\}\cdot \mathbb{1}\left\{W \in \mathcal{W}\right\}\Big] \quad \textit{and} \quad \mathcal{T} = \sum_{n,n'\in\mathcal{N}} \sum_{q=1}^{Q_{n,n'}} \mathcal{T}^q_{n,n'}. \tag{4}$$

3.3 Nonparametric estimators

We maintain that we observe a sample $\{U_i \equiv (Y_i,X_i,N_i,Z_i)\colon\, 1 \leq i \leq L\}$ with $U_i \sim F$. We employ kernel-based estimators constructed using a kernel $K : \mathbb{R}^q \longrightarrow \mathbb{R}$ and a bandwidth sequence $h_L \longrightarrow 0$. For a given $x \equiv \left(x^c,\, x^d\right)$ and $h > 0$ denote

$$\mathcal{H}\left(X_i - x;h\right) = K\left(\frac{X_i^c - x^c}{h}\right)\cdot \mathbb{1}\left\{X_i^d - x^d = 0\right\}.$$

For a given $(x,z) = w$ and n we use

$$\begin{aligned}
\widehat{f}_{X,N}(x,n) &= \frac{1}{L\cdot h_L^r}\sum_{i=1}^{L}\mathcal{H}\left(X_i - x;h_L\right)\cdot \mathbb{1}\left\{N_i = n\right\},\\
\widehat{s}(w,n) &= \frac{1}{L\cdot h_L^r}\sum_{i=1}^{L} S\left(Y_i,w,n\right)\cdot\mathcal{H}\left(X_i - x;h_L\right)\cdot \mathbb{1}\left\{N_i = n\right\}\Big/\widehat{f}_{X,N}(x,n),\\
\widehat{R}^q\left(w;n,n'\right) &= m^q\left(\widehat{s}(w,n),\,\widehat{s}(w,n');\,n,n'\right).
\end{aligned}$$

As we described above, our target testing range depends on $f_{X,N}$. We estimate $\mathcal{T}^q_{n,n'}$ and $\mathcal{T}$ as

$$\begin{aligned}
\widehat{\mathcal{T}}^q_{n,n'} &= \frac{1}{L}\sum_{i=1}^{L}\widehat{R}^q\left(W_i;n,n'\right)\cdot \mathbb{1}\left\{\widehat{R}^q\left(W_i;n,n'\right) \geq -b_L\right\}\cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\},\\
\widehat{\mathcal{T}} &= \sum_{n,n'\in\mathcal{N}}\sum_{q=1}^{Q_{n,n'}}\widehat{\mathcal{T}}^q_{n,n'}.
\end{aligned} \tag{5}$$

$b_L \longrightarrow 0$ is a nonnegative bandwidth sequence, whose inclusion will allow us to deal with the "kink" at zero in the function $\max\{v, 0\}$ and obtain (under assumptions described below) asymptotically pivotal properties for $\widehat{\mathcal{T}}^q_{n,n'}$ and $\widehat{\mathcal{T}}$. Its role is analogous, e.g., to the bandwidth sequences β_n in Jun, Pinkse, and Wan (2010) and τ_n in Kim (2009). Both of these papers also involve testing and inference with moment inequalities.[4]

3.4 Asymptotic properties

We characterize the asymptotic distribution of $\widehat{\mathcal{T}}$ under four types of assumptions:

(i) Smoothness conditions

(ii) A special regularity assumption

(iii) Bias-reducing kernels and bandwidth convergence conditions

(iv) Manageability properties of the empirical processes involved

Assumption 3.1. ***(Smoothness conditions)***
(i) As before we let $w = (x, z)$. For each $(n, n') \in \mathcal{N}^2$ and $F \times F-$almost every $(w_1, w_2) \in \mathcal{W}^2$, the objects

$$E_{Y|X,N}\left[S(Y, w_1, n')\big|X = x_2, N = n\right] \quad \textit{and} \quad f_{X,N}(x_2, n)$$

are M times differentiable with respect to x_2^c, with bounded derivatives. Below, we will describe how large M needs to be.

(ii) For some $-\infty < \underline{s} < \overline{s} < \infty$ and each of the $\ell = 1, \ldots, d_s$ elements in the vector of conditional moments $s(w, n)$, we have $\underline{s} \leq s^\ell(w, n) \leq \overline{s}$ $\forall$ $w \in \mathcal{W}$, $n \in \mathcal{N}$. Whenever these derivatives exist, let

$$\nabla_s m^q\left(s_a, s_b; n, n'\right) = \left(\nabla_{s_a} m^q\left(s_a, s_b; n, n'\right), \nabla_{s_b} m^q\left(s_a, s_b; n, n'\right)\right)'.$$

[4] We could use a bandwidth specific to each of our $q = 1, \ldots, Q_{n,n'}$ restrictions and generalize (5) to

$$\widehat{\mathcal{T}}^q_{n,n'} = \frac{1}{L}\sum_{i=1}^{L} \widehat{R}^q\left(W_i; n, n'\right) \cdot \mathbb{1}\left\{\widehat{R}^q\left(W_i; n, n'\right) \geq -b^q_L(n, n')\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\}. \qquad (5')$$

As long as each of the bandwidths $b^q_L(n, n')$ satisfies the conditions to be described in Assumption 3.3 below, all our asymptotic results would follow through. We focus on the expression given in (5) for expositional purposes.

And define $\nabla_{ss'} m^q\left(s_a, s_b; n, n'\right)$ *accordingly. For some* $\eta > 0$*, the Jacobian* $\nabla_s m^q\left(s_a, s_b; n, n'\right)$ *and Hessian* $\nabla_{ss'} m^q\left(s_a, s_b; n, n'\right)$ *exist for any* $(s_a, s_b) \in \left[\underline{s} - \eta,\ \overline{s} + \eta\right]^{d_s}$ *and any* $(n, n') \in \mathcal{N}^2$*. Furthermore, for some* $D < \infty$*,*

$$\sup_{\substack{(s_a,s_b)\in[\underline{s}-\eta,\ \overline{s}+\eta]^{d_s}\\ n,n'\in\mathcal{N}}} \left\|\nabla_s m^q\left(s_a, s_b; n, n'\right)\right\| \leq D,$$

$$\sup_{\substack{(s_a,s_b)\in[\underline{s}-\eta,\ \overline{s}+\eta]^{d_s}\\ n,n'\in\mathcal{N}}} \left\|\nabla_{ss'} m^q\left(s_a, s_b; n, n'\right)\right\| \leq D.$$

In the auction application we consider in later sections, the functions $\psi_{q:n}^{-1}(s)$ are smooth and differentiable inside any interval in the interior of the unit interval. For this reason, in order to satisfy the conditions in Assumption 3.1 in the auctions models we study here, the testing range $\mathcal{W}$ will have to be such that must be such that

$$0 < \underline{G} \leq G_{q:n}(z|x) \leq \overline{G} < 1 \quad \forall\ (z, x) \equiv w \in \mathcal{W}, n \in \mathcal{N},$$

and the same must be true for $G_{n:n}^{\Delta}(z|x)$.

Next note that any test of (1) must allow for the special case where

$$Pr\left(R^q\left(W; n, n'\right) = 0\right) > 0.$$

That is, $R^q\left(W; n, n'\right)$ may have a point mass at zero.[5] While we allow for this, we assume that $R^q\left(W; n, n'\right)$ has a finite density in an open neighborhood to the left of zero. More precisely we impose the following condition.

Assumption 3.2. ***(A regularity condition)*** *There exist constants* $\overline{b} > 0$ *and* $\overline{A} < \infty$ *such that, for any* $0 < b \leq \overline{b}$ *and each* $n, n' \in \mathcal{N}$ *and* $q = 1, \ldots, Q_{n,n'}$*,*

$$Pr\left(-b \leq R^q\left(W; n, n'\right) < 0\ \middle|W \in \mathcal{W}\right) \leq b \cdot \overline{A}.$$

The smoothness conditions in Assumption 3.1 can lead to $\sqrt{L}-$consistency of $\widehat{\mathcal{T}}$ if we can make the bias in our nonparametric estimators disappear at a fast enough rate. We describe conditions under which this can be achieved in the following assumption.

Assumption 3.3. ***(Kernels and bandwidths)*** *Let* M *be as described in Assumption 3.1. We use a bias-reducing kernel* K *of order* M *with bounded support. The kernel is a function*

[5] This will happen if the inequality being tested holds with equality at a positive measure of points.

of bounded variation, symmetric around zero and satisfies $\sup_{v \in \mathbb{R}^r} \left|K(v)\right| \leq \overline{K} < \infty$. *The bandwidth sequences* b_L *and* h_L *satisfy*

$$L^{1/2} \cdot h_L^r \cdot b_L \longrightarrow \infty$$

and for a small enough $\epsilon_1 > 0$,

$$\frac{b_L \cdot L^{\epsilon_1}}{\sqrt{h_L^r}} \longrightarrow 0 \quad and \quad L^{1/2+\epsilon_1} \cdot b_L^2 \longrightarrow 0.$$

In addition, M *is large enough that*

$$L^{1/2+\epsilon_1} \cdot h_L^M \longrightarrow 0.$$

If our bandwidths are of the form $h_L \propto L^{-\alpha_h}$ and $b_L \propto L^{-\alpha_b}$ it is not hard to verify that the smallest value for which Assumption 3.3 can hold is $M = 2r + 1$. This is the smallest number of bounded derivatives that the functionals in Assumption 3.1 must possess.

Assumption 3.4. ***(Empirical process conditions)***
Let

$$\overline{S}(y) = \sup_{w \in \mathcal{W}, n \in \mathcal{N}} \left\|S(y, w, n)\right\|.$$

Then $E\left[\exp\left\{\overline{S}(Y)^2 \cdot \epsilon\right\}\right] \leq C$ *for some* $\epsilon > 0$ *and* $C < \infty$ *(i.e.,* $\overline{S}(Y)^2$ *possesses a moment generating function). For each* $\ell = 1, \ldots, d_s$ *and each* $n \in \mathcal{N}$*, the class of functions*

$$\mathscr{F}^\ell = \left\{f\colon\ f(y) = S^\ell(y, w, n) \ \ for\ some\ w \in \mathcal{W}\right\}$$

is Euclidean *(see Definition 2.7 in Pakes and Pollard (1989)) with respect to the envelope* $\overline{S}(\cdot)$.

Using the results in Pakes and Pollard (1989) (in particular, Lemmas 2.4 and 2.14), direct inspection shows that each of the classes of functions considered in the examples of Section 4 is in fact Euclidean.

Theorem 1. *If Assumptions 3.1-3.4 hold, then for each $n, n' \in \mathcal{N}$ and $q = 1, \ldots, Q_{n,n'}$,*

$$\widehat{\mathcal{T}}^q_{n,n'} = \mathcal{T}^q_{n,n'} + \frac{1}{L}\sum_{i=1}^{L} \lambda^q_L \left(U_i; n, n'\right) + \xi^q_L(n, n'),$$

where

(i) $E\left[\lambda^q_L\left(U_i; n, n'\right)\right] = 0.$

(ii) If $Pr\left(R^q(W; n, n') < 0 \middle| W \in \mathcal{W}\right) = 1$, *then* $\lambda^q_L\left(U_i; n, n'\right) = 0$ *w.p.1.*

(iii) $|\xi^q_L(n, n')| = O_p\left(L^{-1/2-\epsilon}\right)$ *for some* $\epsilon > 0$.

Proof: In the appendix. □

A precise expression for the "influence function" $\lambda^q_L\left(U_i; n, n'\right)$ is given in the appendix. Let

$$\lambda_L\left(U_i\right) = \sum_{n,n' \in \mathcal{N}} \sum_{q=1}^{Q_{n,n'}} \lambda^q_L\left(U_i; n, n'\right).$$

By Theorem 1,

(i) $E\left[\lambda_L\left(U_i\right)\right] = 0.$

(ii) If $Pr\left(R^q(W; n, n') < 0 \middle| W \in \mathcal{W}\right) = 1 \; \forall \, n, n' \in \mathcal{N}$ and $q = 1, \ldots, Q_{n,n'}$, then $\lambda_L\left(U_i\right) = 0$ w.p.1.

By the representation result in Theorem 1, we have

$$\widehat{\mathcal{T}} = \mathcal{T} + \frac{1}{L}\sum_{i=1}^{L} \lambda_L\left(U_i\right) + \xi_L, \quad \text{where} \quad \xi_L = O_p\left(L^{-1/2-\epsilon}\right) \text{ for some } \epsilon > 0. \tag{6}$$

Let

$$\sigma^2_L = \text{Var}\left(\lambda_L\left(U_i\right)\right).$$

Because $\sigma^2_L = 0$ if our (weak) inequalities are satisfied almost surely as strict inequalities, σ^2_L is the relevant measure for the *slackness* in (1).

3.5 A test statistic

The linear representation in (6) and the specific properties of the influence function $\lambda_L(U_i)$ are the foundation of our test. Let $\kappa_L \longrightarrow 0$ be a nonnegative sequence such that $L^\epsilon \cdot \kappa_L \longrightarrow \infty$ for any $\epsilon > 0$ (e.g., $\kappa_L \propto \log(L)^{-1}$). Define

$$t_L = \frac{\sqrt{L} \cdot \widehat{\mathcal{T}}}{\max\left\{\sigma_L, \, \kappa_L\right\}}$$

We can characterize the asymptotic properties of t_L in three relevant cases.

(i) If (1) is violated with positive probability over our testing range, then

$$t_L = \underbrace{\frac{\sqrt{L} \cdot \mathcal{T}}{\max\{\sigma_L,\, \kappa_L\}}}_{\to +\infty} + \underbrace{\frac{1}{\sqrt{L}} \sum_{i=1}^{L} \frac{\lambda_L(U_i)}{\max\{\sigma_L,\, \kappa_L\}}}_{=O_p(1)} + \underbrace{O_p\left(\frac{L^{-\epsilon}}{\kappa_L}\right)}_{=o_p(1)}$$

(ii) If the restrictions in (1) are satisfied as strict inequalities w.p.1 over our testing range, then

$$t_L = O_p\left(\frac{L^{-\epsilon}}{\kappa_L}\right) = o_p(1).$$

In this case, for any $c > 0$ we have $\lim_{L\to\infty} Pr\,(t_L \leq c) = 1$.

(iii) If the restrictions in (1) are satisfied w.p.1 over our testing range but at least one of them holds with equality with positive probability,

$$t_L = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \frac{\lambda_L(U_i)}{\max\{\sigma_L,\, \kappa_L\}} + \underbrace{O_p\left(\frac{L^{-\epsilon}}{\kappa_L}\right)}_{=o_p(1)}$$

In this case, for any $c > 0$ we have $\lim_{L\to\infty} Pr\,(t_L \leq c) \geq \Phi(c)$, where Φ is the Standard Normal distribution.

Take any $\alpha \in (0,1)$ and let $c_{1-\alpha}$ be the Standard Normal critical value that satisfies $\Phi\,(c_{1-\alpha}) = 1 - \alpha$. From the above results we have

(i) $\lim_{L\to\infty} Pr\,(t_L \leq c_{1-\alpha}) \geq 1 - \alpha$ if (1) is satisfied w.p.1 over our testing range,

(ii) $\lim_{L\to\infty} Pr\,(t_L \leq c_{1-\alpha}) = 0$ otherwise.

σ_L^2 is unknown but is nonparametrically identified and can be estimated as

$$\widehat{\sigma}_L^2 = \frac{1}{L} \sum_{i=1}^{L} \widehat{\lambda}_L^2(U_i).$$

The estimator $\widehat{\lambda}_L(U_i)$ for the influence function is described in the appendix. Under the conditions leading to Theorem 1 we will have $\left|\widehat{\sigma}_L^2 - \sigma_L^2\right| = o_p(1)$. Let

$$\widehat{t}_L = \frac{\sqrt{L} \cdot \widehat{\mathcal{T}}}{\max\{\widehat{\sigma}_L,\, \kappa_L\}}.$$

For a target size $\alpha \in (0,1)$ consider the rejection rule

$$\text{“}\textit{Reject (1) if } \widehat{t}_L > c_{1-\alpha}\text{”} \tag{7}$$

This decision rule would have the following properties,

$$\lim_{L\to\infty} Pr\Big(1 \text{ is rejected when it is true}\Big) \leq \alpha,$$
$$\lim_{L\to\infty} Pr\Big(1 \text{ is rejected when it is violated over our testing range}\Big) = 1.$$

3.6 A study of uniform asymptotic properties

Here we outline the uniform properties of our approach under certain assumptions about the family of distributions that produced our data. We will denote this space as $\mathcal{F}$ and for each $F \in \mathcal{F}$ we will index the various functionals involved in our test by F. Thus, we will denote $\mathcal{T}(F)$, $\lambda_L(U, F)$ and $\sigma_L(F)$. Similarly,

$$R^q\left(X, Z; n, n', F\right) = m^q\left(s(X, Z, n, F), s(X, Z, n', F); n, n'\right).$$

The nonparametric nature of our approach and the various functionals that are relevant can potentially complicate the analysis of the uniform asymptotic properties of our test. The following assumption helps make things tractable.

Assumption 3.5. *Let $\mathcal{F}$ be a family of distributions for U which has common support and satisfies $P_F\left[W \in \mathcal{W}\right] \geq \underline{p} > 0$. In addition,*

(i) Every $F \in \mathcal{F}$ satisfies Assumptions 3.1, 3.2 and 3.4.

(ii) Let

$$\mathcal{F}^*_{\mathcal{W}} = \left\{F \in \mathcal{F}\colon\ P_F\left(R^q\left(X, Z, n, n', F\right) < 0 \middle| W \in \mathcal{W}\right) = 1 \quad \forall\ n, n' \in \mathcal{N},\ q = 1, \ldots, Q_{n,n'}\right\}$$

$\mathcal{F}^*_{\mathcal{W}}$ is therefore the subset of distributions such that the inequalities in (1) *are satisfied as* strict *inequalities $F-a.s$ over our testing range $\mathcal{W}$. There exist $b > 0$ and $\epsilon > 0$ such that*

$$\lim_{L\to\infty} E_F\left[\frac{\lambda_L(U, F)^{2+\epsilon}}{\sigma_L^{2+\epsilon}(F)}\right] \leq b \quad \forall\ F \in \mathcal{F} \setminus \mathcal{F}^*_{\mathcal{W}}$$

Part (i) is meant to ensure that the linear representation in Theorem 1 is valid for each $F \in \mathcal{F}$. Part (ii) is an integrability condition which, intuitively, bounds the information about $\mathcal{T}(F)$ (and $\lambda_L(U, F)$) that is contained in the tails of F. Since one sample of observations yields little information about the tails of the distribution, allowing $\mathcal{T}(F)$ to be sufficiently sensitive to the tails of F could bring about poor (uniform) size properties for our procedure. Assumption 3.5(ii) is analogous to the integrability condition in Romano

(2004, Section 4.2). By Theorem 1, for each $F \in \mathcal{F}^*_{\mathcal{W}}$ we have $\lambda_L(U,F) = 0$ $F-$a.s (and therefore $\sigma_L(F) = 0$). Therefore the ultimate purpose of Assumption 3.5(ii) is to ensure that, for any sequence of distributions $\{F_L\} \in \mathcal{F}$ and any nonnegative sequence $\delta_L \to 0$,

$$\lim_{L\to\infty} E_{F_L}\left[\frac{\lambda_L(U,F_L)^{2+\epsilon}}{\max\left\{\sigma_L^{2+\epsilon}(F_L), \delta_L\right\}}\right] \leq b.$$

Let us generalize our setting and assume we have a triangular array $\{U_i : 1 \leq i \leq L, L \geq 1\}$ that is row-wise iid with distribution F_L where $\{F_L\} \in \mathcal{F}$. To study the properties of $\widehat{t}_L$, let us focus for now on $t_L = \frac{\sqrt{L}\cdot\widehat{\mathcal{T}}}{\max\{\sigma_L(F_L),\kappa_L\}}$, the (unfeasible) statistic normalized by the true standard deviation $\sigma_L(F_L)$. The properties of t_L will be preserved for $\widehat{t}_L$ if $\mathcal{F}$ is equipped with conditions that ensure $\left|\max\left\{\widehat{\sigma}_L(F_L),\kappa_L\right\} - \max\left\{\sigma_L(F_L),\kappa_L\right\}\right| \xrightarrow{p} 0$. Let

$$\widetilde{\mathcal{F}}_{\mathcal{W}} = \left\{F \in \mathcal{F}\colon\ P_F\left(R^q\left(X,Z,n,n',F\right) \leq 0 \middle| W \in \mathcal{W}\right) = 1 \quad \forall\ n,n' \in \mathcal{N},\ q = 1,\ldots,Q_{n,n'}\right\}.$$

$\widetilde{\mathcal{F}}_{\mathcal{W}}$ is the subset of distributions such that the inequalities in (1) are satisfied. That is, the set of distributions for which our null hypothesis is true. Above we denoted $\mathcal{F}^*_{\mathcal{W}}$ as the collection of distributions that satisfy (1) as strict inequalities, therefore $\mathcal{F}^*_{\mathcal{W}} \subseteq \widetilde{\mathcal{F}}_{\mathcal{W}}$. Under the conditions in Assumption 3.5,

$$\lim_{L\to\infty} P_{F_L}\left(t_L > c_{1-\alpha}\right) \leq \alpha \quad \forall\ \{F_L\} \in \widetilde{\mathcal{F}}_{\mathcal{W}}.$$

The proof combines the results in Theorem 1 and the verification of the Lindeberg condition, which would follow from Assumption 3.5 (as in Romano (2004, Lemma 1)). Furthermore if $\widetilde{\mathcal{F}}_{\mathcal{W}} \setminus \mathcal{F}^*_{\mathcal{W}} \neq \emptyset$, then there exists a sequence[6] $\{F_L\} \in \widetilde{\mathcal{F}}_{\mathcal{W}}$ such that

$$\lim_{L\to\infty} P_{F_L}\left(t_L > c_{1-\alpha}\right) = \alpha$$

Combining the above results we would obtain

$$\begin{aligned} &\lim_{L\to\infty} \underset{F\in\widetilde{\mathcal{F}}_{\mathcal{W}}}{sup}\ P_F\left(t_L > c_{1-\alpha}\right) \leq \alpha, \quad \text{with} \\ &\lim_{L\to\infty} \underset{F\in\widetilde{\mathcal{F}}_{\mathcal{W}}}{sup}\ P_F\left(t_L > c_{1-\alpha}\right) = \alpha \quad \text{if} \quad \widetilde{\mathcal{F}}_{\mathcal{W}} \setminus \mathcal{F}^*_{\mathcal{W}} \neq \emptyset. \end{aligned} \tag{8}$$

(8) is relevant for the size properties of our approach. Moving on to the issue of power, there will be two objects of interest. For any sequence of distributions $\{F_L\} \in \mathcal{F}$ let

$$\delta_1 = \lim_{L\to\infty}\left[\frac{\max\left\{\sigma_L(F_L),\kappa_L\right\}}{\sigma_L(F_L)}\right] \quad \text{and} \quad \delta_2 = \lim_{L\to\infty}\left[\frac{\sqrt{L}\mathcal{T}(F_L)\sigma_L(F_L)}{\max\left\{\sigma_L(F_L),\kappa_L\right\}^2}\right],$$

[6]Simply choose a sequence $\{F_L\} \in \widetilde{\mathcal{F}}_{\mathcal{W}} \setminus \mathcal{F}^*_{\mathcal{W}}$.

where δ_1 and δ_2 are allowed to be ∞ (note that $\delta_1 \geq 1$ and $\delta_2 \geq 0$). Under our previous conditions[7],

$$\lim_{L\to\infty} P_{F_L}\left(t_L > c_{1-\alpha}\right) = 1 - \Phi\left(\delta_1 \cdot (c_{1-\alpha} - \delta_2)\right) \tag{9}$$

This expression helps us characterize the power properties of our approach. Since $\delta_1 \geq 1$, we will have asymptotic power of 1 for any sequence $\{F_L\}$ such that $\delta_2 = \infty$. In particular, consider sequences such that $\mathcal{T}(F_L) = \frac{d}{\sqrt{L}}$ for some $d > 0$. Our test will have asymptotic power of 1 against such sequences as long as $\lim\limits_{L\to\infty} \sigma_L(F_L) = 0$ but $\lim\limits_{L\to\infty} \frac{\kappa_L^2}{\sigma_L(F_L)} = 0$ since we would have $\delta_2 = \infty$. That is, as long as $\sigma_L(F_L)$ also converges to zero but at a rate slower than κ_L^2 (and therefore, slower than $\mathcal{T}(F_L)$ given our restrictions regarding the rate of convergence of κ_L). Equation (9) reveals that our procedure will lack power against sequences $\{F_L\} \in \mathcal{F} \setminus \widetilde{\mathcal{F}}$ such that δ_2 is "small" and δ_1 is "large". Specifically, if $\delta_2 < c_{1-\alpha}$, we can find a sufficiently large δ_1 such that $\Phi\left(\delta_1 \cdot (c_{1-\alpha} - \delta_2)\right) \approx 1$ resulting in a loss of power. One immediate case in which this would be ruled out would be if we assumed that

$$\lim_{L\to\infty} \sigma_L(F) \geq \underline{c} > 0 \quad \forall\, F \in \mathcal{F} \setminus \widetilde{\mathcal{F}}.$$

In this case we would have $\delta_1 = 1$ and $\delta_2 \leq \underline{c}^{-1} \cdot \lim\limits_{L\to\infty} \sqrt{L} \cdot \mathcal{T}(F_L)$. However, adding the above assumption would come at the cost of losing the property of asymptotic power of 1 against sequences where $\mathcal{T}(F_L) = \frac{d}{\sqrt{L}}$ for some $d > 0$ (which we described above). For any such sequence we would have $\lim\limits_{L\to\infty} P_{F_L}\left(t_L > c_{1-\alpha}\right) \leq 1 - \Phi\left(c_{1-\alpha} - d/\underline{c}\right)$. Without further restrictions on $\mathcal{F}$, it appears that having asymptotic power of 1 against these types of $\sqrt{L}-$local alternatives comes at the cost of relinquishing power against sequence of alternatives where δ_2 is small. In particular, against sequences where $\delta_2 < c_{1-\alpha}$.

Next we need to take into account the fact that σ_L is estimated in our test. If $\lambda_L(U_i)$ were directly observed, Assumption 3.5(ii) would suffice to ensure that an appropriate LLN for triangular arrays holds (see Romano (2004, Lemma 2)) and we would have

$$\frac{\max\left\{\sum_{i=1}^{L} \lambda_L^2(U_i)/L, \kappa_L\right\}}{\max\left\{\sigma_L^2(F_L), \kappa_L\right\}} \longrightarrow 1.$$

However, $\lambda_L(U_i)$ is not observed and is replaced with a nonparametric estimator in the construction of our test. Thus, to ensure that the size and power properties described in (8)

[7] Recall that $\mathcal{T}(F_L) = 0$ for any $\{F_L\} \in \widetilde{\mathcal{F}}_{\mathcal{W}}$ and therefore $\delta_2 = 0$ for any such sequence. Our result in (8) follows as a special case of (9) since $\delta_1 \geq 1$.

and (9) are preserved for our testing procedure it would suffice to endow $\mathcal{F}$ with conditions that guarantee that

$$\frac{\max\left\{\sum_{i=1}^{L}\widehat{\lambda}_L^2(U_i)/L,\kappa_L\right\}}{\max\left\{\sigma_L^2(F_L),\kappa_L\right\}}\longrightarrow 1.$$

Analyzing each one of the pieces that goes into the construction of $\widehat{\lambda}_L^2(U_i)$ illuminates the type of additional conditions on $\mathcal{F}$ that would suffice to ensure that[8]

$$\underset{u\in Supp(U)\cap\mathcal{W}\times\mathcal{N}}{sup}\left|\widehat{\lambda}_L(u,F_L)-\lambda_L(u,F_L)\right|\xrightarrow{p}0$$

for sequences $\{F_L\}$ in $\mathcal{F}$. From here and Assumption 3.5, the size and power properties described in (8) and (9) would be satisfied by our test.

3.7 Choice of tuning parameters

As is usual in nontrivial nonparametric problems, a definitive theory for choosing tuning parameters appears to be beyond our reach. Clearly, our choice of the sequence b_L will have a direct impact on the finite-sample properties of our method and in particular in the tradeoff between size and power. Values of b_L closer to zero will lead to greater power, but they may also lead to over-rejection rates in finite samples. One possible approach to take both effects into consideration in our choice of b_L could be by using a parametric model as a benchmark. Fitting the reference parametric model and drawing simulations from it we could choose b_L in a way that preserves the target size while emphasizing power against particular alternatives of interest[9]. This could also serve as a guidance for our choice of the remaining tuning parameters involved. Section 4.2 presents a reasonably detailed Monte Carlo analysis for several choices of tuning parameters while applying our test on ascending auctions models.

3.7.1 Tuning parameters and scale invariance

Trivially, if $R^q(W;n,n')\leq 0$, then we will also have $d\cdot R^q(W;n,n')\leq 0$ for any $d\geq 0$. Having a statistic that is invariant to a rescaling of $R^q(\cdot)$ is therefore desirable. Let s_a, s_b

[8]This is stronger than we need since we only require $\frac{\max\left\{\sum_{i=1}^{L}\widehat{\lambda}_L^2(U_i)/L,\kappa_L\right\}}{\max\left\{\sigma_L^2(F_L),\kappa_L\right\}}\longrightarrow 1$, and not necessarily $\frac{\sum_{i=1}^{L}\widehat{\lambda}_L^2(U_i)/L}{\sigma_L^2(F_L)}\longrightarrow 1$.

[9]We can also use an approach like the one suggested in Jun, Pinkse, and Wan (2010, Section 5) as a guidance in our choice for b_L.

be two given values such that $|m^q(s_a, s_b; n, n')| > 0$ and denote $\overline{R}^q_{n,n'} \equiv |m^q(s_a, s_b; n, n')|$, and let $\overline{R} = \sum_{q=1}^{Q_{n,n'}} \sum_{n,n' \in \mathcal{N}} \overline{R}^q_{n,n'}$. We can use a bandwidth b_L specific to each pair (n, n') considered (see footnote 4). Consider bandwidths of the type

$$b^q_L(n, n') = c_b \cdot \max\left\{\Omega^q_{n,n'}, \overline{R}^q_{n,n'}\right\} \cdot L^{-\alpha_b}, \quad \text{where} \quad \Omega^q_{n,n'} = \left\{\mathrm{Var}\left(R^q(W_i; n, n')\right)\right\}^{1/2}$$

where $c_b > 0$ is a constant chosen by the researcher and α_b is selected so as to satisfy the restrictions in Assumption 3.3. We consider $\max\left\{\Omega^q_{n,n'}, \overline{R}^q_{n,n'}\right\}$ because $\Omega^q_{n,n'}$ could be zero if $R^q(W_i; n, n') = 0$ w.p.1 (i.e., the restrictions hold with equality almost surely). In the implementation we would replace $\Omega^q_{n,n'}$ with a nonparametric estimator.

We can proceed similarly with the sequence κ_L. Let $\lambda^*_L(U_i)$ denote the expression for $\lambda_L(U_i)$ that would correspond if $R^q(W_i; n, n') \geq 0$ w.p.1 for each $n, n' \in \mathcal{N}$ and $q = 1, \ldots, Q_{n,n'}$. This would be the case if our inequalities almost surely either held with equality or were violated. Now define $\Sigma^2_L = \mathrm{Var}\left(\lambda^*_L(U_i)\right)$ and let $\kappa > 0$ be a prespecified constant. Using a sequence of the form

$$\kappa_L = c_\kappa \cdot \max\left\{\Sigma_L, \overline{R}\right\} \cdot \log(L)^{-1}$$

would be conducive to scale invariant properties for our test-statistic. In the implementation we would replace Σ_L with a nonparametric estimator. Once again we advocate using $\max\left\{\Sigma_L, \overline{R}\right\}$ because we could have $\Sigma_L = 0$ if all our restrictions hold as strict inequalities almost surely.

3.8 More general L^p-type functionals

Our statistic can be generalized to one of the form

$$\widehat{\mathcal{T}}^q_{n,n'}(p) = \left(\frac{1}{L}\sum_{i=1}^{L}\left(\widehat{R}^q\left(W_i; n, n'\right)\right)^p \cdot \mathbb{1}\left\{\widehat{R}^q\left(W_i; n, n'\right) \geq -b_L\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\}\right)^{\frac{1}{p}},$$

for $p \geq 1$. Under the same type of assumptions we made above we can extend our asymptotic results to this general case. Using $p > 1$ can allow us to use sequences b_L that converge to zero at a faster rate than $p = 1$ (the case we analyzed above).

3.9 Our test relative to existing methods

The literature on inference with conditional moment inequalities has grown in the recent past. One type of approach to this problem relies on "instrument functions". Conditional moment inequality restrictions of the type $E\left[f(Y)|X\right] \leq 0$ translate into restrictions $E\left[f(Y)\cdot g(X)\right] \leq 0$ for any nonnegative measurable function g. Denoting $\mu(g) = E\left[f(Y)\cdot g(X)\right]$ and pre-specifying a space of functions $\mathcal{G}$, the conditional moment inequalities are tested by estimating sample analogs of CvM statistics of the type $\int\limits_{g\in\mathcal{G}} \max\left\{\mu(g),0\right\} dg$ or KS statistics of the type $\underset{g\in\mathcal{G}}{sup} \max\left\{\mu(g),0\right\}$. Notable examples include Andrews and Shi (2013), Armstrong (2011a) and Armstrong (2011b). Using instrument functions allows us to dispense with smoothness assumptions on the conditional moments involved, but by its nature is not suited in general to test inequalities involving nonlinear *transformations* of a collection of conditional moments. Therefore our problem is in general beyond the scope of this approach.

In general, our problem requires some type of direct nonparametric estimation of the functionals in question. Lee, Song, and Whang (2013) present a test for inequalities similar to ours. Like us, they use one-sided L^p-functionals of nonparametric estimators. In contrast with our method, their approach relies on a *least favorable* configuration whereby they standardize their test-statistic based on the least favorable case under the null hypothesis. This is the case where the inequalities are binding as equalities almost surely. In our context this would amount to replacing our test-statistic $\widehat{t}_L$ with

$$\widetilde{t}_L = \frac{\sqrt{L}\cdot\widehat{\mathcal{T}}}{\widehat{\Omega}_L},$$

where $\widehat{\Omega}_L$ is the estimator that results for $\widehat{\sigma}_L$ if we replace $\mathbb{1}\left\{\widehat{R}^q(W_i;n,n') \geq -b_L\right\}$ with 1. By design, the resulting test-statistic would be asymptotically a Standard Normal only in the least favorable case but it would be conservative in other cases. In contrast, our test is asymptotically standard normal also in intermediate cases where the inequalities are satisfied a.s and binding with strictly positive probability (but not with probability one). By being less conservative than with a least-favorable configuration our test would also have more power, particularly against cases where the inequalities are violated with a small probability. There are many economic applications where detecting these types of violations is of particular interest.

Chernozhukov, Lee, and Rosen (2013) (CLR) introduce a novel method for estimation and inference on intersection bounds. Their approach covers many econometric problems of interest and is general enough that it can be adapted and applied to our specific problem. To simplify the notation let $v \equiv (x,z,n,n',q)$, abbreviate $\widehat{R}^q(x,z;n,n') = \widehat{R}(v)$ and let

$\widehat{\theta}(v) \equiv -\widehat{R}(v)$. Denote a prespecified target testing range as

$$\mathcal{V} = \{(x, z, n, n', q) : x \in \mathcal{X}, z \in \mathcal{Z}, n, n' \in \mathcal{N}, q \in 1, \ldots, Q_{n,n'}\}.$$

The approach in CLR in our setting would revolve around a statistic of the form

$$\widehat{\theta} = \underset{v \in \mathcal{V}}{sup} \left[\widehat{\theta}(v) + \widehat{k}(p) \cdot \sigma(v)\right],$$

where $\sigma(v)$ is the standard error of the estimator $\widehat{\theta}(v)$ and $\widehat{k}(p)$ is a critical value chosen based on the properties of the standardized process $Z_L(\cdot)$ given by $Z_L(v) = \frac{\theta(v) - \widehat{\theta}(v)}{\sigma(v)}$. More precisely, $\widehat{k}(p)$ must be an appropriately constructed approximation to the p^{th} quantile of $\underset{v}{sup} Z_L(v)$. CLR present a theory of how to construct $\widehat{k}(p)$, advocating simulation-based methods as the preferable approach. In the context of our problem, implementation can potentially present computational challenges involving the approximation of the suprema involved in its construction when the dimension of X and/or Z is large[10]. In that case doing an exhaustive grid search quickly becomes computationally unfeasible and appropriately chosen grid search methods must be employed. The dimensionality of X or Z present no significant computational challenges in the method we propose, they would just dictate the type of kernel and bandwidths we have to use. Using the methods in CLR, the properties of the set $V_0 = arg \underset{v}{inf} [\theta(v)]$ would play a key role in the details of the resulting estimation rates and local power properties. In our general problem, pinning down the relevant features of V_0 could be challenging. On the other hand, as we described in Section 3.6, the simplicity of our proposed method facilitates the understanding of the power (or lack thereof) of our approach against specific alternatives (see Eq. (9) above). Overall, a general and definitive power comparison between our approach and the test that would result from a CLR approach is not immediately obvious and would be probably better understood on a case by case basis using Monte Carlo experiments. However, the simplicity of our test presents computational advantages when the dimensionality of X or Z is large. This is the case in our empirical application to auctions models in Section 3.6, where X includes six continuously distributed covariates.

[10]In a companion paper, Chernozhukov, Kim, Lee, and Rosen (2013) present a step-by-step computational procedure to implement the inferential approach in CLR, but it allows for multidimensional X only in parametric models. Unfortunately this leaves out our general case of interest.

4 Application: Testing Models of English Auctions

In this section, we demonstrate the usefulness of the testing approach described above by applying it to test certain standard assumptions made in the empirical modeling of English auctions. We first show theoretically how certain testable implications follow from each model, and how these implications fit within our testing framework. We run Monte Carlo simulations to demonstrate the power of the resulting tests, then apply the tests to actual data from United States Forest Service timber auctions.

4.1 Testable Implications of Various Models of English Auctions

4.1.1 General Setup

Each auction in the data is characterized by a set of observable (to the researcher) covariates describing that particular auction, X; a number of bidders, N; and a vector of bids, $\boldsymbol{B} = (B_1, \ldots, B_N)$. The joint distribution of the observables $(X, N, \boldsymbol{B})$ is thus nonparametrically identified by the data. We will maintain the following assumption throughout:

Assumption 4.1. *Bidders have private values, and the joint distribution of these private values is symmetric.*

Thus, we will assume that bidders have private values $\boldsymbol{V} = (V_1, \ldots, V_N)$; let $F(\cdot|x, n)$ denote the joint distribution of these valuations, conditional on $X = x$ and $N = n$.[11] Symmetry imposes the additional restriction that $F(v_1, v_2, \ldots, v_n|x, n) = F(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)}|x, n)$ for $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ any permutation. An independent private values (IPV) model would impose the additional restriction that $F(v_1, \ldots, v_n|x, n) = \prod_{i=1}^{n} F_V(v_i|x, n)$ for some univariate distribution F_V (which may or may not depend on n).

Our primary focus is to understand and test the implications of the IPV assumption in English (ascending) auctions. Unlike the case of first-price auctions (see section 5.5 in Athey and Haile (2007)), bids in ascending auctions may be correlated even if valuations are not. Thus, we cannot test the IPV model in ascending auctions simply by testing for conditional covariance among bids. However, in this section we show it is still possible to derive non-parametric testable implications of ascending auctions using the properties of order statistics.

[11] This is a slight abuse of notation, as the domain of $F : \Re^n_+ \to [0, 1]$ depends on n, but the meaning should be clear.

Fixing $N = n$, let $V_{1:n} \leq V_{2:n} \leq \ldots \leq V_{n:n}$ denote the order statistics of the random vector of valuations $\boldsymbol{V}$, and $F_{k:n}(\cdot \mid x)$ the distribution of $V_{k:n}$ conditional on the realization $(X, N) = (x, n)$. Similarly, let $B_{1:n} \leq \ldots \leq B_{n:n}$ denote the order statistics of the random vector of bids $\boldsymbol{B}$, and $G_{k:n}(\cdot \mid x)$ the distribution of $B_{k:n}$ given $(X, N) = (x, n)$. Since bids are observed, $G_{k:n}(\cdot \mid x)$ are identified from the data.

For $k \leq n$, define a function $\psi_{k:n} : [0,1] \to [0,1]$ by

$$\psi_{k:n}(s) = \frac{n!}{(n-k)!(k-1)!} \int_0^s t^{k-1}(1-t)^{n-k} dt$$

For $t \in (0,1)$, the integrand is positive, so $\psi_{k:n}$ is strictly increasing everywhere and therefore invertible. Athey and Haile (2002) observe that if n independent random variables are drawn from a distribution $H(\cdot)$, the distribution of the k^{th}-lowest is $\psi_{k:n}(H(\cdot))$. Under an IPV model, then, for any k and n, $F_V(v|x,n) = \psi_{k:n}^{-1}(F_{k:n}(v|x))$.

4.1.2 Testing IPV with Fixed N

In an open-outcry English auction, a bidder responds to his opponents' bids; a bidder's valuation therefore does not uniquely determine his bid in an English auction, and inferring valuations from bids is not a straightforward exercise. To address this, Haile and Tamer (2003) introduced an "incomplete model" of bidding in ascending auctions. Rather than impose a unique mapping from primitives to equilibrium outcomes, Haile and Tamer (2003) impose only weak assumptions about bidder behavior, and aim to partially identify an IPV model. They assume bidders never bid higher than their valuations, which implies $B_{k:n} \leq V_{k:n}$; and they assume that bidders never allow the auction to end at a price they could profitably beat, which implies that for $k < n$, $V_{k:n} \leq B_{n:n} + \Delta$, where Δ is the minimum bid increment at the end of the auction. While Haile and Tamer (2003) assume IPV, these bidding assumptions are roughly analogous to bidders playing weakly undominated strategies, and need not depend on bidders' beliefs or the joint distribution of valuations; they are therefore equally natural in any private-values setting. We will use "Haile-and-Tamer bidding" to describe bidding strategies which satisfy these two assumptions, but are otherwise unrestricted.

In order to create a test that will still have power even in such an unstructured model of bidding, we must place some structure on how we might expect independence to be violated, that is, what we see as the alternative hypothesis to IPV. We do this in a theoretically general and non-parametric way:

Assumption 4.2. *For each n and x, the joint distribution $F(\,\cdot\,|x,n)$ is such that for any v and i, the probability $\Pr(V_i < v|X = x, N = n, \|\{j \neq i : V_j < v\}\| = k)$ is nondecreasing in k.*

This formulation of correlated private values was introduced by Aradillas-López, Gandhi, and Quint (2013). There, we show (Lemma 1) that Assumption 4.2 holds under all the standard models of symmetric, correlated private values: specifically, affiliated private values, conditionally-independent private values, and IPV with unobserved heterogeneity.

Haile-and-Tamer bidding implies that $G_{k:n}(v|x) \geq F_{k:n}(v|x)$ and $F_{n-1:n}(v|x) \geq G^{\Delta}_{n:n}(v|x)$, where $G^{\Delta}_{n:n}(\cdot|x)$ is the distribution of $B_{n:n} + \Delta$ (given $X = x$). Even though valuations are not uniquely pinned down, the model is still testable:

Proposition 1. *Under IPV and Haile-and-Tamer bidding, for any (x, n, v) and any $k \leq n-2$,*

$$\psi^{-1}_{k:n}\left(G_{k:n}(v|x)\right) \quad \geq \quad \psi^{-1}_{n-1:n}\left(G^{\Delta}_{n:n}(v|x)\right) \tag{10}$$

On the other hand, if values are not independent, then at any (x, n, v) where Assumption 4.2 holds strictly – that is, where $\Pr\left(V_i < v \mid X = x, N = n, \|\{j \neq i : V_j < v\}\| = k\right)$ is not the same for all k – then for $k \leq n-2$,

$$\psi^{-1}_{k:n}(F_{k:n}(v|x)) \quad < \quad \psi^{-1}_{n-1:n}(F_{n-1:n}(v|x))$$

and (10) will therefore be violated if there is sufficiently little slack in the "Haile and Tamer bounds".

The first part of Proposition 1 was noted by Haile and Tamer (2003) (Remark 2), who point out that it could be used as a test of the IPV model. The second part of the proposition, however, is new, and shows that this test can have power against the standard alternative hypotheses to independence.

In order to apply our testing framework, we now represent the restriction in Proposition 1 as an instance of (1), and in particular, as a case of (1′). The decision variables are $Y = (B_{1:N}, \ldots, B_{N-1:n}, B_{N:N} + \Delta)$. The index variable Z could be any real-valued random variable with the property that $\text{Support}(V) \subseteq \text{Support}(Z)$. For a given $w \equiv (x, z)$, y and n, the structural function S is the vector-valued indicator $S(y, w, n) = \mathbb{1}\{y \leq z\}$, so

$$s\,(w, n) = E_{Y|X,N}\left[S(Y, w, n)\middle|X = x, N = n\right] = \left(G_{1:n}(z|x), \ldots, G_{n-1:n}(z|x), G^{\Delta}_{n:n}(z|x)\right).$$

For each n, the model involves $Q_n = n-2$ transformations $\{m^q\}$, with

$$m^q\left(s\left(w,n\right);\,n\right) = \psi_{n-1:n}^{-1}\left(G_{n:n}^{\Delta}(z|x)\right) - \psi_{q:n}^{-1}\left(G_{q:n}(z|x)\right), \quad q = 1,\ldots,n-2.$$

As Proposition 1 indicates, the power of the test depends on how close $G_{n:n}^{\Delta}$ is to $F_{n-1:n}$ and $G_{k:n}$ to $F_{k:n}$. If the top two bids are close together in most auctions (implying also that Δ is small), then the first inequality will not have much slack: since $G_{n:n}^{\Delta} \leq F_{n-1:n} \leq G_{n-1:n}$, if $G_{n-1:n}$ and $G_{n:n}^{\Delta}$ are close together, $F_{n-1:n}$ must be close to $G_{n:n}^{\Delta}$. Thus, the real concern is whether $G_{k:n}$ is close to $F_{k:n}$ for $k \leq n-2$ – that is, whether losing bidders other than the second-highest bid close to their valuations. Song (2004) considers the possibility that the "top two losers" bid close to their values, even if the others do not; this would be enough for our test to have power. Unfortunately, there is no easy way to check this in the data. And if only the highest losing bidder approaches his value, this test may have little power.

As a result, we consider another approach to testing the IPV model, which relies only on transaction prices (or the winning and highest losing bids) but requires variation in the number of bidders.

4.1.3 Testing IPV using Variation in N

Exploiting variation in N requires an assumption about the nature of this variation. We will assume that variation in the number of bidders is independent of the realization of their valuations. To formalize this condition, let $F_m^n(\cdot \mid x)$ denote the joint distribution of m bidders drawn at random from an auction with n bidders, conditional on $X = x$.[12]

Definition. Values are **independent of** N if $F_m^n(\,\cdot \mid x) = F_m^{n'}(\,\cdot \mid x)$ for all (x, n, n', m).

Under the IPV model, this simply means that the marginal distribution $F_V\left(\,\cdot \mid x, n\right)$ does not depend on n. This assumption has been used in Haile, Hong, and Shum (2003), Guerre, Perrigne, and Vuong (2009), Gillen (2009), and Aradillas-López, Gandhi, and Quint (2013), and has been termed an "exclusion restriction" since N is excluded from the distribution $F_V\left(\,\cdot \mid x\right)$.

This test, and the subsequent ones, are based only on the distribution of the second-highest valuation $V_{N-1:N}$ as N changes. In many applications, including in Aradillas-López, Gandhi, and Quint (2013), this valuation is assumed to be equal to the transaction

[12]Since $F(\,\cdot \mid x, n)$ is symmetric, $F_m^n(v_1, \ldots, v_m \mid x) = F\left(v_1, \ldots, v_m, \infty, \infty, \ldots, \infty \mid x, n\right)$.

price $B_{N:N}$ – as it would be in a "button auction". If bidders only increase their bids by the minimum amount toward the end of the auction, this should be true to within a bid increment under the Haile-and-Tamer bidding assumptions. Here, we present the test under both under the assumption that $B_{n:n} = V_{n-1:n}$, and under the weaker Haile-and-Tamer assumptions; in our application, we use the former test.

Proposition 2. *Assume $B_{n:n} = V_{n-1:n}$ and values are independent of N.*

(a) Under IPV, for any (x, n, n', v),

$$\psi_{n-1:n}^{-1} \left(G_{n:n}(v|x) \right) \quad = \quad \psi_{n'-1:n'}^{-1} \left(G_{n':n'}(v|x) \right) \tag{11}$$

(b) Under Assumption 4.2 (nonnegatively correlated private values), for any (x, n, n', v),

$$n > n' \quad \longrightarrow \quad \psi_{n-1:n}^{-1} \left(G_{n:n}(v|x) \right) \quad \geq \quad \psi_{n'-1:n'}^{-1} \left(G_{n':n'}(v|x) \right) \tag{12}$$

Further, (12) *holds strictly wherever Assumption 4.2 holds strictly, that is, at every point (x, n, n', v) where $n > n'$ and $\Pr(V_i < v \,|\, X = x, N = n, \|\{j \neq i : V_j < v\}\| = k)$ is not the same for all k.*

(11) was proposed by Athey and Haile (2002) as a possible basis for a test of the IPV model. (See also the discussion in Athey and Haile (2007).) The drawback of (11) as a standalone test, however, is that it is really a joint test of two assumptions: IPV *and* the exclusion restriction. That is, a rejection of (11) could follow from a violation either of IPV or of the exclusion restriction. (12), on the other hand, is a new result, and contributes to our testing strategy in two ways. First, it ensures that (11) has power against all the standard models of positively-correlated values when the exclusion restriction holds. More importantly, it provides a testable implication of the exclusion restriction itself which does not depend on independence of values. In Appendix B.5, we show that (12) has power as a test of the exclusion restriction. Specifically, we show fairly general conditions under which a correlated private values model, combined with either of the two standard models of endogenous entry in auctions (those of Levin and Smith (1994) and Samuelson (1985)), would lead to a violation of (12). (This is also illustrated in a numerical example in Section 4.1.7 below.) Thus, if data violates (11) but satisfies (12), this supports the hypothesis that the failure of (11) is caused by a violation of IPV rather than a violation of the exclusion restriction. If this is indeed the case – values are correlated, but independent of the number

of bidders – then both upper and lower bounds are identified for the seller's expected profit and optimal reserve price, using the approach laid out in Aradillas-López, Gandhi, and Quint (2013).[13]

To gain intuition for Proposition 2, consider what happens to the distribution of transaction prices as N increases. As N increases, transaction prices get stochastically higher (the distribution shifts to the right), since the price is set by the second-highest of a bigger group. (Pinkse and Tan (2005) refer to this as the *sampling effect*.) If values are IPV and F_V does not vary with n, Proposition 2 says that this must happen at a particular "speed" – that is, for each v, $F_{n-1:n}(v)$ must fall exactly fast enough so that $\psi_{n-1:n}^{-1}(F_{n-1:n}(v))$ remains constant.

Relative to that benchmark, correlation of values *slows down* the sampling effect – if values are correlated, then each incremental bidder has less impact on transaction price, as bidder values are more prone to be close together. So if values are correlated but independent of N, $F_{n-1:n}(v)$ falls more slowly than under IPV, and $\psi_{n-1:n}^{-1}(F_{n-1:n}(v))$ therefore increases with n.

On the other hand, violations of the exclusion restriction would likely be due to a positive relationship between valuations and N – that is, endogenous participation favoring auctions for more-valuable prizes. This would augment the sampling effect, causing $F_{n-1:n}(v)$ to fall more quickly than under IPV; provided this effect was stronger than the slowing-down due to correlation, it would result in $\psi_{n-1:n}^{-1}(F_{n-1:n}(v))$ decreasing with n. As noted above, we have shown that even in the presence of correlation, the test of (12) has power against a fairly wide class of "typical" violations of the exclusion restriction.

The restriction in (12) is an instance of (1) where the decision variable is $Y = B_{N:N}$ and (as before) the index variable Z could be any real-valued random variable with the property that $\text{Support}(V) \subseteq \text{Support}(Z)$. The structural function is $S(y, w, n) = S(y, z) = \mathbb{1}\{y \leq z\}$ and

$$s(w, n) = E_{Y|X,N}\left[S(Y, w, n)\middle|X = x, N = n\right] = G_{n:n}(z|x).$$

For each n, n' the model involves a single transformation (i.e., $Q_{n,n'} = 1$) given by

$$m\left(s(w, n);\, s(w, n');\, n, n'\right) = \left(\psi_{n'-1:n'}^{-1}\left(G_{n':n'}(z|x)\right) - \psi_{n-1:n}^{-1}\left(G_{n:n}(z|x)\right)\right) \cdot \mathbb{1}\{n > n'\}.$$

[13] In that paper, we also show that the same upper bound on profit, and a weaker upper bound on the optimal reserve price, still hold if the exclusion restriction is violated.

The restriction in (11) involves an equality between transformations of conditional moments. Notice however that we can frame it as the combination of two inequalities: (12) along with its reverse inequality. If either fails, we would reject (11).

4.1.4 Extending Proposition 2 to Haile-and-Tamer bidding behavior

In Appendix B.2, we show that if we drop the assumption $B_{n:n} = V_{n-1:n}$, and instead assume Haile-and-Tamer bidding behavior, then (11) and (12) above become

$$\psi^{-1}_{n-1:n}\left(G_{n-1:n}(v|x)\right) \quad \geq \quad \psi^{-1}_{n'-1:n'}\left(G^{\Delta}_{n':n'}(v|x)\right) \tag{13}$$

and

$$n > n' \quad \longrightarrow \quad \psi^{-1}_{n-1:n}\left(G_{n-1:n}(v|x)\right) \quad \geq \quad \psi^{-1}_{n'-1:n'}\left(G^{\Delta}_{n':n'}(v|x)\right) \tag{14}$$

respectively. Both of these can again be written as instances of (1). The decision variables are $Y = (B_{N-1:N},\, B_{N:N} + \Delta)$, and once again, the index variable Z could be any real-valued random variable with the property that $\text{Support}(V) \subseteq \text{Support}(Z)$. For (13), the structural function S is given by $S(y, w, n) = S(y, z) = \mathbb{1}\{y \leq z\}$ and therefore

$$s\,(w, n) = E_{Y|X,N}\left[S(Y, w, n)\middle| X = x, N = n\right] = \left(G_{n-1:n}(z|x),\, G^{\Delta}_{n:n}(z|x)\right).$$

For each n, n' the model involves a single transformation (i.e., $Q_{n,n'} = 1$) given by

$$m\left(s\,(w, n)\,;\, s\,(w, n')\,;\, n, n'\right) = \psi^{-1}_{n'-1:n'}\left(G^{\Delta}_{n':n'}(z|x)\right) - \psi_{n-1:n}\left(G_{n-1:n}(z|x)\right).$$

For (14), $S(y, w, n) = \mathbb{1}\{y \leq z\}$ (again), and therefore

$$s\,(w, n) = E_{Y|X,N}\left[S(Y, w, n)\middle| X = x, N = n\right] = \left(G_{n-1:n}(z|x),\, G^{\Delta}_{n:n}(z|x)\right).$$

As before, the model entails a single transformation for each n, n' given now by

$$m\left(s\,(w, n)\,;\, s\,(w, n')\,;\, n, n'\right) = \Big(\psi^{-1}_{n'-1:n'}\left(G^{\Delta}_{n':n'}(z|x)\right) - \psi^{-1}_{n-1:n}\left(G_{n-1:n}(z|x)\right)\Big) \cdot \mathbb{1}\{n > n'\}.$$

4.1.5 Testing IPV when the Exclusion Restriction Fails

When the exclusion restriction is rejected, of course, (11) no longer offers a test of IPV. Without any restriction on how $F_V(\,\cdot\,|x, n)$ can vary with n, the IPV model is just-identified from transaction price data, and therefore not testable. However, a natural restriction would be a general positive reslationship between the number of bidders and their valuations. This can be formalized as the following condition:[14]

Assumption 4.3. *If valuations are IPV but the distribution $F_V(\cdot \mid x, n)$ depends on n, then it does so in such a way that (for any x) $n > n'$ implies $F_V(\cdot \mid x, n) \succsim_{FOSD} F_V(\cdot \mid x, n')$.*

Proposition 3. *Assume $B_{n:n} = V_{n-1:n}$. Under IPV and Assumption 4.3,*

$$n > n' \quad \longrightarrow \quad \psi_{n-1:n}^{-1}\left(G_{n:n}(v|x)\right) \quad \leq \quad \psi_{n'-1:n'}^{-1}\left(G_{n':n'}(v|x)\right) \tag{15}$$

for all (x, n, n', v).

Observe that Proposition 3 gives the opposite conclusion as part (b) of Proposition 2; that is, relative to the benchmark of (11), violations of the exclusion restriction work in the opposite direction as correlation among values. When both the exclusion restriction and independence fail, (15) need not always have power as a test of IPV. Nevertheless, a rejection would serve as evidence against IPV.

If we assume Haile-and-Tamer bidding, (15) becomes

$$n > n' \quad \longrightarrow \quad \psi_{n-1:n}^{-1}\left(G_{n:n}^{\Delta}(v|x)\right) \quad \leq \quad \psi_{n'-1:n'}^{-1}\left(G_{n'-1:n'}(v|x)\right) \tag{16}$$

(see Appendix B.2). For testing, both (15) and (16) can be framed as instances of (1). For the case of (15) the decision variable is $Y = B_{N:N}$, and (as before) the index variable Z could be any real-valued random variable with the property that Support(V) $\subseteq$ Support(Z). The structural transformation is $S(y, w, n) = S(y, z) = \mathbb{1}\{y \leq z\}$ and

$$s(w, n) = E_{Y|X,N}\left[S(Y, w, n)\big|X = x, N = n\right] = G_{n:n}(z|x).$$

For each n, n' the model involves a single transformation (i.e., $Q_{n,n'} = 1$) given by

$$m\left(s(w,n)\,;\, s(w,n')\,;\, n, n'\right) = \Big(\psi_{n-1:n}^{-1}\left(G_{n:n}(z|x)\right) - \psi_{n'-1:n'}^{-1}\left(G_{n':n'}(z|x)\right)\Big) \cdot \mathbb{1}\{n > n'\}.$$

For (16) we have $Y = (B_{N-1:N},\, B_{N:N} + \Delta)$, $S(y, w, n) = \mathbb{1}\{y \leq z\}$ and consequently

$$s(w, n) = E_{Y|X,N}\left[S(Y, w, n)\big|X = x, N = n\right] = \left(G_{n-1:n}(z|x),\, G_{n:n}^{\Delta}(z|x)\right).$$

The model consists of a single transformation for each n, n' which is given by

$$m\left(s(w,n)\,;\, s(w,n')\,;\, n, n'\right) = \Big(\psi_{n-1:n}^{-1}\left(G_{n:n}^{\Delta}(z|x)\right) - \psi_{n'-1:n'}^{-1}\left(G_{n'-1:n'}(z|x)\right)\Big) \cdot \mathbb{1}\{n > n'\}.$$

[14]In Aradillas-López, Gandhi, and Quint (2013), we generalize this notion of valuations being "stochastically increasing in N" to settings with correlated values, and show conditions under which it follows from three different models of endogenous entry.

4.1.6 Summarizing the results from auctions models

Table 1 below summarizes the various tests we have derived above:

Table 1: Overview of Observable Implications

Eq.	**Test of**	N	**Bidding assumptions**
10	IPV	Fixed	Haile-and-Tamer bidding
11	IPV $\wedge$ $V \perp N$	Variable	Transaction price $= V_{n-1:n}$
12	$V \perp N$	Variable	Transaction price $= V_{n-1:n}$
15	IPV	Variable	Transaction price $= V_{n-1:n}$
13	IPV $\wedge$ $V \perp N$	Variable	Haile-and-Tamer bidding
14	$V \perp N$	Variable	Haile-and-Tamer bidding
16	IPV	Variable	Haile-and-Tamer bidding

• Rejecting 12 or 15 would also reject 11.

4.1.7 Illustration of Propositions 2 and 3

Equations (11), (12), and (15) are claims that $\psi^{-1}_{n-1:n}(G_{n:n}(v|x))$ is constant, increasing, or decreasing in n, respectively. To illustrate how $\psi^{-1}_{n-1:n}(G_{n:n}(v|x))$ behaves for various types of data-generating processes, we graph its value (as a function of v) for different values of N under four versions of a parametric example. For the example, there are no observable covariates X; bidder values are *i.i.d.* draws from a log-normal distribution, $\log(V_i) \sim N(\mu, \sigma^2)$, with $\sigma^2 = 0.5$ throughout but μ potentially variable. The four cases are as follows:

1. Values are IPV and independent of N: specifically, $\mu = 2.25$ for every N
2. Values are independent of N, but correlated with each other via conditional independence: regardless of N, $\mu = 2.0$ with probability $\frac{1}{2}$ and 2.5 with probability $\frac{1}{2}$. (Variation in μ induces correlation among values.)
3. Values are IPV, but the distribution varies with N: specifically, $\mu = 2 + 0.05N$
4. Values are correlated with each other, and with N. $\mu = 2.5$ or 1.5 with probabilities $\frac{1}{3}$ and $\frac{2}{3}$ respectively, and N is determined endogenously via equilibrium play of the entry game described in Samuelson (1985). There are 12 potential bidders, each of

whom learns μ and his own valuation before deciding whether to pay a cost of \$10 to participate in the auction. Bidders play a symmetric, cutoff-strategy equilibrium, with the cutoff value varying with μ;[15] this induces a positive relationship between N and μ, and therefore between N and valuations.

Figure 1: $\psi^{-1}_{n-1:n}\left(G_{n:n}(v)\right)$ against v under four scenarios

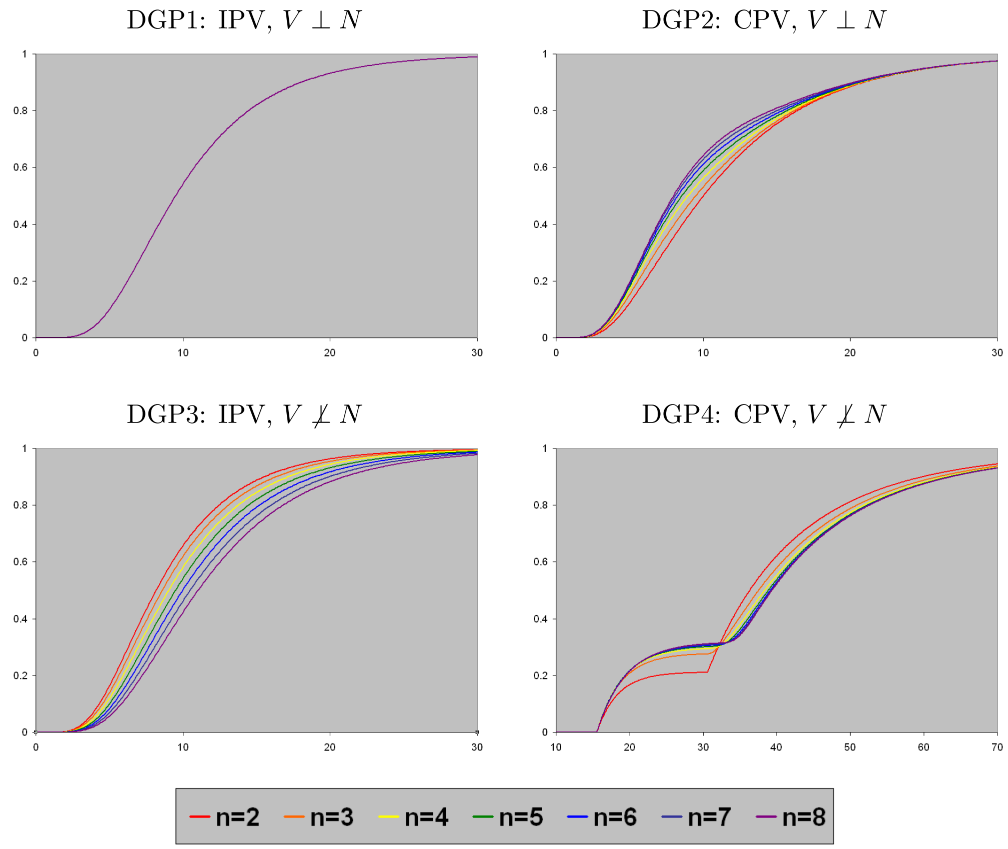

Figure 1 shows plots of $\psi^{-1}_{n-1:n}\left(G_{n:n}(v)\right)$ against v for various values of n for each scenario. Now consider what would happen if we ran our tests, in order, on each of these four data-generating processes. For DGP1, we would fail to reject (11), and conclude (correctly) that the data was consistent with both IPV and the exclusion restriction. For DGP2, we would reject (11) but fail to reject (12), (correctly) rejecting IPV but not the exclusion restriction.

[15] When $\mu = 2.5$, the entry cutoff is 30.57, which is exceeded by 9.7% of bidders; when $\mu = 1.5$, the cutoff is 15.54, which is exceeded by 3.9% of bidders. By Bayes' Law, then, $\Pr(\mu = 2.5 | N = n)$ is increasing in n.

For DGP3, we would reject both (11) and (12), but fail to reject (15), concluding (correctly) that the data was consistent with an IPV model violating the exclusion restriction. Finally, for DGP4, we would reject all three tests, concluding (correctly) that both IPV and the exclusion restriction failed.[16]

4.2 Monte Carlo Experiments

4.2.1 Setup

In order to investigate the finite-sample properties of our econometric methodology, we applied the tests of (12) and (15) above on simulated data. The data is based on modifications of the last three DGPs from Section 4.1.7 to include a single auction-specific covariate X. In all cases, the maximum number of bidders was 12, and valuations satisfy $\log(V_i) \sim \mathcal{N}\left(\mu, \sigma^2\right)$. We fixed $\sigma^2 = 0.5$, let $X \sim \mathcal{N}(0,1)$, and generated μ in ways analogous to DGPs 2, 3 and 4 above in the following way:

(A) Values are independent of N conditional on X but are correlated with each other (even conditional on X) in the following way. Let $\varepsilon \sim \mathcal{N}(0,1)$ such that $\varepsilon \perp X$. If $X + \varepsilon > 0$ then $\mu = 2.0$, otherwise $\mu = 2.5$.

(B) Values are IPV conditional on X, but the distribution of values varies with N in the following way. If $X < 0$ then $\mu = 1.7 + .05 \cdot N$. Otherwise $\mu = 2.3 + .05 \cdot N$. Note that on average we have $\mu = 2 + .05 \cdot N$.

(C) Let $\varepsilon \sim \mathcal{N}(0,1)$ with $\varepsilon \perp X$ and let $c_{\frac{1}{3}}$ denote the $\frac{1}{3}^{rd}$ quantile from the Standard Normal distribution. Then $\mu = 2.5$ if $\frac{X+\varepsilon}{\sqrt{2}} < c_{\frac{1}{3}}$ and $\mu = 1.5$ otherwise. Everything else is as described in DGP4 above, with N being determined endogenously (given μ) by the equilibrium outcome of an entry game.

By design, DGP(A) satisfies (12) almost surely as a *strict* inequality, while it violates (15). The reverse is true for DGP(B). Both (12) and (15) are violated with positive probability in DGP(C), but each of these inequalities is satisfied over some range of x and n. Table 2 summarizes the predicted asymptotic behavior of our econometric test for each one of these designs.

[16] When neither IPV nor the exclusion restriction hold, it is not necessarily the case that equations 12 and 15 will both be violated: a similar example based on a different entry model (that of Levin and Smith (1994)) leads to distributions satisfying equation 15 everywhere.

Table 2: Asymptotic rejection probabilities predicted for our test.

	DGP(A)	**DGP(B)**	**DGP(C)**
$\lim_{L\to\infty} Pr\,(\text{Reject 12})$	0	1	1
$\lim_{L\to\infty} Pr\,(\text{Reject 15})$	1	0	1

These asymptotic predictions are valid at any significance level.

4.2.2 Kernels and bandwidths

We have $r = 1$ (one continuous observable X). The smallest order of the kernel that can satisfy Assumption 3.3 is $M = 2r + 1 = 3$. We employed a bias-reducing kernel of the type

$$K(\psi) = \sum_{\ell=1}^{2} c_\ell \cdot \left(s^2 - \psi^2\right)^{2\ell} \cdot \mathbb{1}\{|\psi| \le s\},$$

with $s = 30$ (as in Section 4.4 of Aradillas-López, Gandhi, and Quint (2013)). The coefficients c_1, c_2 were chosen to ensure that $K(\cdot)$ was bias-reducing of order $M = 4$. Our bandwidth h_L was chosen as

$$h_L = c_h \cdot \widehat{\sigma}(X) \cdot L^{-\alpha_h}.$$

The choices of c_h and α_h are discussed below. To construct b_L and κ_L we followed the general guidelines in Section 3.7.1. Let

$$\widehat{\Omega} = \left\{ \widehat{\text{Var}} \left(\sum_{q=1}^{Q_{n,n'}} \sum_{n,n' \in \mathcal{N}} \widehat{R}^q\left(W; n, n'\right) \right) \right\}, \quad \widehat{\Sigma} = \left\{ \widehat{\text{Var}} \left(\widehat{\lambda}^*_L(U) \right) \right\},$$

where (as described in Section 3.7.1), $\lambda^*_L(U_i)$ is the expression of the influence function $\lambda_L(U_i)$ that would follow if $R^q(W_i; n, n') \ge 0$ w.p.1. for each $n, n' \in \mathcal{N}$ and $q = 1, \ldots, Q_{n,n'}$. From here we used

$$b_L = c_b \cdot \widehat{\Omega} \cdot L^{-\alpha_b}, \quad \kappa_L = c_\kappa \cdot \widehat{\Sigma} \cdot \log(L)^{-1}.$$

The bandwidth convergence restrictions in Assumption 3.3 can be satisfied if we set

$$0 < \epsilon_h \le \frac{1}{4 \cdot r \cdot (2 \cdot r + 1)}, \quad 0 < \epsilon_b < \epsilon_h, \quad \alpha_h = \frac{1}{4 \cdot r} - \epsilon_h, \quad \alpha_b = \frac{1}{4} + \epsilon_b.$$

We chose $\epsilon_h = \frac{9}{10} \cdot \frac{1}{4 \cdot r \cdot (2 \cdot r + 1)}$ and $\epsilon_b = \frac{9}{10} \cdot \epsilon_b$ ($r = 1$ as described previously). This yielded $\alpha_h \approx 0.28$ and $\alpha_b \approx 0.32$. In our experiments we fixed $c_\kappa = 10^{-1}$ and studied the properties of our tests for $c_h \in \{0.25, 0.40\}$ and $c_b \in \left\{10^{-3}, 1\right\}$ for samples of size $L = 1,000,\ 1,250$ and $1,500$.

4.2.3 Testing range

As our range for N we used $\mathcal{N} = \left\{n\colon\ \widehat{Pr}(N = n) \geq .05\right\}$. As our index variable Z we used

$$Z = B_{N:N} \quad \text{(i.e., transaction price)}$$

and the testing range $\mathcal{W}$ was given by

$$\mathcal{W} = \left\{(x, z)\colon\ \widehat{f}_X(x) \geq \widehat{f}_X^{(.005)} \text{ and } 10^{-4} \leq \widehat{G}_{k:n}(z|x) \leq 1 - 10^{-4}\ \forall\ 2 \leq k \leq n,\ n \in \mathcal{N}\right\}$$

where $\widehat{f}_X^{(.005)}$ denotes the $.005^{th}$ quantile of $\widehat{f}_X(\cdot)$.

4.2.4 Results from experiments

Detailed step-by-step details of the construction of our test statistics can be found in Appendix A.2. We generated 500 simulations of each DGP for $c_h \in \{0.25, 0.4\}$ and $c_b \in \left\{10^{-3}, 1\right\}$, fixing $c_\kappa = 10^{-1}$. This value of c_κ made κ_L negligible in every instance of our experiments.[17] The observed rejection frequencies are summarized in Tables 3 and 4.

Overall, the results from our MC experiments fell very much in line with the asymptotic predictions summarized in Table 2. We can summarize our main findings as follows.

- The power properties of our tests in the cases of DGPs (A) and (B) – where one of the inequalities is violated almost everywhere – were remarkably robust to the choices of the tuning parameters. This was also true for the size of the tests (the probability of rejecting each model when it is true) in these cases.

- The choice of tuning parameters was more impactful for DGP (C), where neither inequality is satisfied w.p.1, but each is satisfied with positive probability over some range of the data. Still, even in this case the power properties of our test held reasonably well even in the worst case.

- The performance of the test was more sensitive to the choice of the bandwidth h_L than to b_L. While the effect of changes in b_L was predictable ex-ante (with smaller values of b_L leading to higher rejection probabilities), the effect of h_L was not as straightforward to predict. Our results showed, across the board, that a smaller bandwidth h_L led to more power, without any important tradeoff in size.

[17]This means that by choosing $c_\kappa = 10^{-1}$ we had $\max\{\widehat{\sigma}_L,\, \kappa_L\} = \widehat{\sigma}_L$ in every run of our experiments.

Table 3: Rejection frequencies[†], 500 simulations.

$c_h = 0.4$, $c_b = 10^{-3}$						
	DGP(A)		DGP(B)		DGP(C)	
	Eq. 12	Eq. 15	Eq. 12	Eq. 15	Eq. 12	Eq. 15
$L = 1,000$	1%	76%	99%	0%[b]	51%	45%
$L = 1,250$	1%	85%	99%	0%[c]	57%	60%
$L = 1,500$	1%	92%	100%[a]	0%[d]	67%	70%
$c_h = 0.25$, $c_b = 10^{-3}$						
	DGP(A)		DGP(B)		DGP(C)	
	Eq. 12	Eq. 15	Eq. 12	Eq. 15	Eq. 12	Eq. 15
$L = 1,000$	4%	74%	100%[e]	0%[h]	71%	61%
$L = 1,250$	5%	84%	100%[f]	0%[i]	79%	77%
$L = 1,500$	4%	91%	100%[g]	0%[j]	84%	82%

(†) Critical value used was 1.645, corresponding to a 5% target level.

• Largest values of our test-statistic in cases with 0% rejection were:

(b): 1.11, (c): 1.17, (d): 1.18, (h): 1.37, (i): 1.39, (j): 1.11.

• Smallest values of our test-statistic in cases with 100% rejection were:

(a): 2.37, (e): 1.86, (f): 1.70, (g): 2.21.

- Gains in power did not come at the cost of significant increases in size. At least for the range considered, smaller bandwidths (both for h_L and b_L) consistently led to gains in power which did not come at the expense of important increases in the size of the test.

4.3 Application to USFS Timber Data

4.3.1 Timber Auctions

We apply our tests to data from timber auctions run by the Unites States Forest Service. These are auctions for the right to harvest timber on a tract of public land. Auctions are heterogeneous due to both differences in the tracts themselves (for example, in the type and density of timber present) and differences in the lease contracts (such as the length of the lease).

Table 4: Rejection frequencies[†], 500 simulations.

$c_h = 0.4$, $c_b = 1$						
	DGP(A)		DGP(B)		DGP(C)	
	Eq. 12	Eq. 15	Eq. 12	Eq. 15	Eq. 12	Eq. 15
$L = 1,000$	0.4%	72%	98%	0%^{b}	44%	32%
$L = 1,250$	0.2%	79%	99%	0%^{c}	48%	42%
$L = 1,500$	0.2%	88%	100%^{a}	0%^{d}	59%	52%
$c_h = 0.25$, $c_b = 1$						
	DGP(A)		DGP(B)		DGP(C)	
	Eq. 12	Eq. 15	Eq. 12	Eq. 15	Eq. 12	Eq. 15
$L = 1,000$	0.4%	65%	99%	0%^{f}	60%	44%
$L = 1,250$	0.6%	76%	99%	0%^{g}	70%	59%
$L = 1,500$	0.4%	86%	100%^{e}	0%^{h}	75%	69%

(†) Critical value used was 1.645, corresponding to a 5% target level.

• Largest values of our test-statistic in cases with 0% rejection were:

(b): 0.79, (c): 0.99, (d): 0.71, (f): 1.03, (g): 1.05, (h): 0.68.

• Smallest values of our test-statistic in cases with 100% rejection were:

(a): 2.17, (e): 2.21.

Prior to each auction, the Forest Service conducts a "cruise" of the tract and publishes detailed information on the tract for potential bidders. The timber data therefore includes a rich set of auction covariates, corresponding to the information the bidders had about the tract. It has often been argued that these covariates therefore capture all *systematic* demand shifters for a tract of timber, and that any remaining variation in valuations is likely bidder-specific (such as differences in costs and capacity) and hence independent. We can now explicitly test this claim and evaluate whether independence or correlation better models valuations in the timber data after we control for these covariates.

A number of other papers have studied Forest Service auctions empirically. Nearly all have done so within the framework of independent private values.[18] Two recent papers, however, have found indirect evidence of correlation among valuations. Athey, Levin, and

[18]See, e.g., Baldwin, Marshall, and Richard (1997); Haile (2001); Haile, Hong, and Shum (2003); Lu and Perrigne (2008); Athey and Levin (2001); and Haile and Tamer (2003).

Seira (2011) estimate a model allowing for unobserved heterogeneity on data from first-price auctions, noting that "an extension along these lines appears crucial as... we estimate implausibly high bid margins when we fail to account for within-auction bid correlation."[19] In Aradillas-López, Gandhi, and Quint (2013), we estimate a model allowing for correlated values on English auction data; we find the estimates (of expected profit as a function of reserve price) differ significantly from estimates made under the assumption of independence.

Thus, while independence is a standard assumption in empirical work, both in general and applied to these particular auctions, there is some recent evidence to suggest this assumption might be worrisome; here, we directly examine the validity of the independence assumption on this data.

4.3.2 Data

Data on all USFS timber auctions held between 1978 and 1996 was made available to us by Phil Haile. We focus on the auctions held between 1982 and 1990, as the reserve price policy in place was stable during that period, and the reserve prices used were generally recognized not to be binding, allowing us to infer the number of potential bidders from the number who submitted bids.[20] We use auctions from Region 6 (mostly Oregon), which relative to other regions provides a large sample of English auctions.

We use the same conventions as Haile and Tamer (2003) (which were motivated by the previous literature) to select auctions most likely to satisfy the assumption of private values. In particular, we focus on sales whose contracts expire within a year, to minimize the effect of resale possibilities on valuations. And we focus on scaled sales, where bids are in dollars per unit of timber actually harvested, and therefore common-value uncertainty about the

[19] Another key insight of Athey, Levin, and Seira (2011) is that there are ex-ante asymmetries between two types of bidders: mills and loggers. As we discuss in Aradillas-López, Gandhi, and Quint (2013), our model can accommodate this type of asymmetry, if we imagine each bidder being independently and randomly either a miller or a logger, so that the marginal distribution of each bidder's valuations is the appropriate mixture between a "mill distribution" and a "logger distribution". In first-price auctions, bids depend on beliefs about one's opponents, so the identities of the other bidders would affect bidding; but in ascending auctions, bidding is in dominant strategies and this information has no effect.

[20] Campo, Guerre, Perrigne, and Vuong (2002) write, "It is well known that this reserve price does not act as a screening device to participating," and perform analysis that confirms that "the possible screening effect of the reserve price is negligible" (p. 33). See also Haile (2001), Froeb and McAfee (1988), and Haile and Tamer (2003).

total amount of timber should not affect valuations. Within this selection of auctions, there are both first-price and ascending auctions, allowing us to apply both tests to the same environment.

For each auction in the sample, in addition to the number of bidders and their bids, our data contains detailed covariate information about the auction from the government's cruise report. We control for *six* auction covariates which have been emphasized in the previous literature as being relevant demand shifters: the density of timber (timber volume over acres in the tract, which we label X^1); the government's appraisal value of the timber (which we label X^2); the estimated profit from manufacturing the timber (sales value minus manufacturing cost, X^3); the estimated harvesting cost (per unit of timber, X^4); the species concentration (the HHI (Herfindahl index) computed as a function of the volume of various species present, X^5); the total volume of timber sold in the six months prior to each auction (as a measure of the bidding firms' existing inventory, X^6). Bids and monetary covariates are all measured in 1983 dollars. We let $X = (X^1, X^2, X^3, X^4, X^5, X^6)$ refer to the vector of covariates, and $X_i = (X_i^1, X_i^2, X_i^3, X_i^4, X_i^5, X_i^6)$ the data corresponding to the i^{th} auction. X was treated as a continuously distributed random vector. We drop auctions with $N = 1$ (since there is no second-highest bidder to whose value we can link the transaction price) and $N = 12$ (as this appears to be top-coding for "more than 11"). Thus, our range for N is $\mathcal{N}\{2, 3, \ldots, 11\}$, which leaves us with a sample of $L = 2,034$ auctions.

4.3.3 Kernels and bandwidths used

We used the same types of kernels and bandwidths as in our Monte Carlo experiments in Section 4.2.2 given $r = 6$. We employed a bias-reducing kernel of order $M = 14$ which is essentially an extension of the one used in Section 4.2.2. This was a multiplicative kernel of the type $K(\psi_1, \ldots, \psi_6) = \prod_{\ell=1}^{6} k(\psi_\ell)$, where

$$k(\psi) = \sum_{\ell=1}^{7} c_\ell \cdot \left(s^2 - \psi^2\right)^{2\ell} \cdot \mathbb{1}\{|\psi| \leq s\},$$

with $s = 30$ (as in Section 4.4 of Aradillas-López, Gandhi, and Quint (2013)). The coefficients $c_1, \ldots, c_7$ were chosen to ensure that $k(\cdot)$ was bias-reducing of order $M = 14$. Our bandwidths h_L, b_L and c_L were chosen using the exact same formulas described in Section 4.2.2, with $r = 6$.

4.3.4 Testing range

For $\mathcal{N}$ we used the entire range of values in the data, $\{2, 3, \ldots, 11\}$. As our index variable Z we used

$$Z = B_{N:N} \quad \text{(i.e., transaction price)}$$

and the testing range $\mathcal{W}$ was given by

$$\mathcal{W} = \left\{(x, z)\colon\ \widehat{f}_X(x) \geq \widehat{f}_X^{(.005)} \text{ and } 10^{-4} \leq \widehat{G}_{k:n}(z|x) \leq 1 - 10^{-4}\ \forall\ 2 \leq k \leq n,\ n \in \mathcal{N}\right\}$$

where $\widehat{f}_X^{(.005)}$ denotes the $.005^{th}$ quantile of $\widehat{f}_X(\cdot)$.

4.3.5 Results

The construction of our test statistics followed the step-by-step description in Appendix A.2. Table 5 shows the results for the tests of (10), (12) and (15). The results shown correspond to $c_h = 0.4$, $c_b = 10^{-3}$ and $c_k = 10^{-1}$ (see the bandwidth formulas in Section 4.2.2). Appendix A.3 shows results for alternative values of these tuning parameters. Although the values of our test statistics change (as should be expected), the qualitative findings – in particular, whether each test rejects the model in question – are consistent across the values analyzed.

Table 5: Test Results on Ascending Auction Timber Data

Eq.	**Test of**	N	**Bidding assumptions**	$\widehat{t}$	**Outcome**†
10	IPV	Fixed	Haile-and-Tamer bidding	**13.41**	Reject
12	$V \perp N$	Variable	$B_{n:n} = V_{n-1:n}$	**0.64**	Fail to Reject
15	IPV	Variable	$B_{n:n} = V_{n-1:n}$	**6.75**	Reject

(†) Critical values for rejection are 1.645 for $\alpha = 5\%$ and 2.326 for $\alpha = 1\%$.

These results paint a consistent picture of the timber data. Both testing methods – comparing winning to losing bids in auctions of the same size, and comparing transaction prices across auctions of different sizes – allow us to reject independence of valuations, and instead give strong evidence of positive correlation among valuations.[21] On the other hand,

[21] As we noted previously, rejection of (15) implies automatically that (11) is rejected, too.

we fail to reject a model of correlated values which are independent of N; thus, the exclusion restriction appears plausible in the ascending auction data.

5 Conclusion

In this paper, we considered testing of economic models whose testable implications involve inequality comparisons between nonlinear transformations of nonparametric conditional moments. Our motivating example was specification tests in ascending auctions, but this setup extends to multiple examples of interest. Because many commonly-used models in economics fit this description, it is important to have econometric tools capable of testing these restrictions in a computationally feasible way in the presence of rich covariate data. In this paper, we described an econometric methodology capable of testing this type of restriction in a straightforward way. We studied the asymptotic properties of our procedure, and also highlighted its finite sample features through Monte Carlo experiments. We also made a comparison between our testing procedure and other existing methods and made a case for the contributions of our approach within the existing literature. Our application involved testing for independence in bidders' private values in ascending auctions. Applying our test to data from the United States Forest Service timber auctions, we found clear evidence to reject the IPV model in favor of a model of correlated private values. Because the IPV assumption is at the heart of key auction theory results, this finding has significant policy implications, which are analyzed in detail in Aradillas-López, Gandhi, and Quint (2013).

A Appendix – Econometrics

A.1 Proof of Theorem 1

A.1.1 A useful probability inequality

As in the main body of the paper, $\mathcal{W} \subset Supp(W)$ denotes our testing range for W. Recall that $\mathcal{W} \cap \text{Supp}(X^c) \subset \text{int}(\text{Supp}(X^c))$, and recall we maintained that

$$f_{X,N}(x,n) \geq \underline{f} > 0 \quad \forall\, x \in \mathcal{W},\, n \in \mathcal{N}.$$

Let

$$\mathcal{I} = \{(w,n)\colon\ w \in \mathcal{W},\ n \in \mathcal{N}\}.$$

$\mathcal{I}$ denotes our overall testing range. In this section we will describe conditions that yield a exponential bound for the probability

$$Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\widehat{R}^q(w;n,n') - R^q(w;n,n')\right| \geq b_L\right),$$

where $b_L \longrightarrow 0$ is the bandwidth sequence used in our construction (5). The bound we obtain is given in (A12) and its usefulness will become evident in sections to follow. Invoking Lemma 22 in Nolan and Pollard (1987) and Lemmas 2.4 and 2.14 in Pakes and Pollard (1989) (see also Example 10 there), having a kernel of bounded variation implies that the class of functions

$$\mathscr{G}_K = \left\{g : g(x) = \mathcal{H}\left(x - v; h\right) \quad \text{for some } v \in \mathbb{R}^{dim(X)} \text{ and some } h > 0\right\}$$

is *Euclidean*[22] with respect to the constant envelope $\overline{K}$. For a given $(x,z) \equiv w$ and n, and for each of the $\ell = 1, \ldots, d_s$ elements in the vector-valued function $S(Y_i, w, n)$ denote

$$\begin{aligned}
\widetilde{q}_i^{\ell}(w,n;h) &= S^{\ell}(Y_i,w,n)\cdot\mathcal{H}(X_i - x;h)\cdot\mathbb{1}\{N_i = n\},\\
\widetilde{v}_i^{\ell}(w,n;h) &= \left(\frac{S^{\ell}(Y_i,w,n) - s^{\ell}(w,n)}{f_{X,N}(x,n)}\right)\cdot\mathcal{H}(X_i - x;h)\cdot\mathbb{1}\{N_i = n\},\\
\widehat{\nu}^{\ell}(w,n) &= \frac{1}{L\cdot h_L^r}\sum_{i=1}^{L}\widetilde{v}_i^{\ell}\left(w,n;h_L\right),\\
\widehat{Q}^{\ell}(w,n) &= \frac{1}{L\cdot h_L^r}\sum_{i=1}^{L}\widetilde{q}_i^{\ell}\left(w,n;h_L\right),\\
Q^{\ell}(w,n) &= s^{\ell}(w,n)\cdot f_{X,N}(x,n).
\end{aligned}$$

[22] See Definition 2.7 in Pakes and Pollard (1989).

Using an M^{th} order approximation, the smoothness conditions in Assumption 3.1 imply the existence of a finite constant $\overline{M}$ such that,

$$\begin{aligned}
&\sup_{(x,n)\in\mathcal{I}} \left| E\left[\widehat{f}_{X,N}(x,n)\right] - f_{X,N}(x,n) \right| \leq \overline{M} \cdot h_L^M, \\
&\sup_{(w,n)\in\mathcal{I}} \left| E\left[\widehat{\nu}^{\ell}(w,n)\right] \right| \leq \overline{M} \cdot h_L^M, \qquad\qquad \text{(A1)} \\
&\sup_{(w,n)\in\mathcal{I}} \left| E\left[\widehat{Q}^{\ell}(w,n)\right] - Q^{\ell}(w,n) \right| \leq \overline{M} \cdot h_L^M.
\end{aligned}$$

If the Euclidean properties in Assumption 3.4 hold, Lemma 2.14 in Pakes and Pollard (1989) implies that the processes

$$\left\{\widetilde{q}_i^{\ell}(w,n;h)\colon\, w\in\mathcal{W},\, n\in\mathcal{N},\, h>0,\, 1\leq i\leq L\right\}$$

$$\left\{\widetilde{v}_i^{\ell}(w,n;h)\colon\, w\in\mathcal{W},\, n\in\mathcal{N},\, h>0,\, 1\leq i\leq L\right\}$$

are *manageable* (as described in Definition 7.9 of Pollard (1990)) with respect to the envelopes[23] $\overline{K}\cdot\overline{S}(\cdot)$ and $\left(\overline{K}/\underline{f}\right)\cdot\left(\overline{S}(\cdot)+\max\left\{|\overline{s}|,|\underline{s}|\right\}\right)$ respectively. These envelopes possess a moment generating function by Assumption 3.4. The Euclidean property of the class of functions $\mathscr{G}_K$ described above also implies that the process

$$\left\{\mathcal{H}(X_i - x;h)\cdot\mathbb{1}\left\{N_i=n\right\}\colon\, x\in\mathcal{W},\, h>0,\, 1\leq i\leq L\right\}$$

is manageable with respect to the constant envelope $\overline{K}$. Using the maximal inequality results in Chapter 7 of Pollard (1990) combined with the bias conditions in A1 imply that there exist positive constants A_1, A_2 and A_3 such that for each $\ell=1,\ldots,d_s$ and any $\delta>0$,

$$\begin{aligned}
&Pr\left(\sup_{(x,n)\in\mathcal{I}} \left|\widehat{f}_{X,N}(x,n) - f_{X,N}(x,n)\right| \geq \delta\right) \leq A_1\cdot\exp\left\{-\sqrt{L}\cdot h_L^r\left(A_2\cdot\delta - A_3\cdot h_L^M\right)\right\}, \\
&Pr\left(\sup_{(w,n)\in\mathcal{I}} \left|\widehat{\nu}^{\ell}(w,n)\right| \geq \delta\right) \leq A_1\cdot\exp\left\{-\sqrt{L}\cdot h_L^r\left(A_2\cdot\delta - A_3\cdot h_L^M\right)\right\}, \\
&Pr\left(\sup_{(w,n)\in\mathcal{I}} \left|\widehat{Q}^{\ell}(w,n) - Q^{\ell}(w,n)\right| \geq \delta\right) \leq A_1\cdot\exp\left\{-\sqrt{L}\cdot h_L^r\left(A_2\cdot\delta - A_3\cdot h_L^M\right)\right\}.
\end{aligned} \qquad \text{(A2)}$$

Group

$$\widehat{\mu}^{\ell}(w,n) = \left(\left[\widehat{Q}^{\ell}(w,n) - Q^{\ell}(w,n)\right] \quad \left[\widehat{f}_{X,N}(x,n) - f_{X,N}(x,n)\right]\right)'.$$

[23] If a class is Euclidean it is also necessarily manageable. See page 1033 in Pakes and Pollard (1989).

Using (A2), for any $\delta > 0$ we have

$$Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left\|\widehat{\mu}^{\ell}(w,n)\right\| \geq \delta\right) \leq 2 \cdot A_1 \cdot \exp\left\{-\sqrt{L}\cdot h_L^r\left(A_2\cdot\frac{\delta}{\sqrt{2}} - A_3\cdot h_L^M\right)\right\}. \qquad (A2')$$

For any $(x,z) \equiv w$ and n such that $f_{X,N}(x,n) > 0$, a second order approximation yields

$$\begin{aligned}
s^{\ell}(w,n) - s^{\ell}(w,n) &= \frac{1}{f_{X,N}(x,n)}\cdot\left[\widehat{Q}^{\ell}(w,n) - Q^{\ell}(w,n)\right] - \frac{s^{\ell}(w,n)}{f_{X,N}(x,n)}\cdot\left[\widehat{f}_{X,N}(x,n) - f_{X,N}(x,n)\right] \\
&+ \frac{1}{2}\cdot\widehat{\mu}^{\ell}(w,n)'\begin{pmatrix} 0 & -\frac{1}{\widetilde{f}^2_{X,N}(x,n)} \\ -\frac{1}{\widetilde{f}^2_{X,N}(x,n)} & \frac{2\widetilde{Q}^{\ell}(w,n)}{\widetilde{f}^3_{X,N}(x,n)}\end{pmatrix}\cdot\widehat{\mu}^{\ell}(w,n) \\
&= \widehat{\nu}^{\ell}(w,n) + \frac{1}{2}\cdot\widehat{\mu}^{\ell}(w,n)'\begin{pmatrix} 0 & -\frac{1}{\widetilde{f}^2_{X,N}(x,n)} \\ -\frac{1}{\widetilde{f}^2_{X,N}(x,n)} & \frac{2\widetilde{Q}^{\ell}(w,n)}{\widetilde{f}^3_{X,N}(x,n)}\end{pmatrix}\cdot\widehat{\mu}^{\ell}(w,n)
\end{aligned}$$

where $\left(\widetilde{Q}^{\ell}(w,n),\, \widetilde{f}_{X,N}(x,n)\right)$ belong in the line segment that connects $\left(\widehat{Q}^{\ell}(w,n),\, \widehat{f}_{X,N}(x,n)\right)$ with $\left(Q^{\ell}(w,n),\, f_{X,N}(x,n)\right)$. Denote

$$\xi_L^{\ell}(w,n) = \frac{1}{2}\cdot\widehat{\mu}^{\ell}(w,n)'\begin{pmatrix} 0 & -\frac{1}{\widetilde{f}^2_{X,N}(x,n)} \\ -\frac{1}{\widetilde{f}^2_{X,N}(x,n)} & \frac{2\widetilde{Q}^{\ell}(w,n)}{\widetilde{f}^3_{X,N}(x,n)}\end{pmatrix}\cdot\widehat{\mu}_L^{\ell}(w,n).$$

Then we can express

$$\widehat{s}^{\ell}(w,n) - s^{\ell}(w,n) = \widehat{\nu}^{\ell}(w,n) + \xi_L^{\ell}(w,n). \qquad (A3)$$

By Assumption 3.1 there exists a constant $\overline{Q} < \infty$ such that

$$\underset{(w,n)\in\mathcal{I}}{sup}\left|Q^{\ell}(w,n)\right| \leq \overline{Q}.$$

Define

$$J = \left\|\begin{matrix} 0 & -\frac{1}{\left(\underline{f}/2\right)^2} \\ -\frac{1}{\left(\underline{f}/2\right)^2} & \frac{3\overline{Q}}{\left(\underline{f}/2\right)^3}\end{matrix}\right\|. \qquad (A4)$$

et J be as described in (A4). Combining (A2)-($A2'$), for any $\delta > 0$ we have

$$\begin{aligned}
&Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left|\xi_L^{\ell}(w,n)\right| \geq \delta\right) \leq Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left|\widehat{Q}^{\ell}(w,n) - Q^{\ell}(w,n)\right| \geq \overline{Q}\right) \\
&+ Pr\left(\underset{(x,n)\in\mathcal{I}}{sup}\left|\widehat{f}_{X,N}(x,n) - f_{X,N}(x,n)\right| \geq \underline{f}/2\right) + Pr\left(\underset{(x,n)\in\mathcal{I}}{sup}\left|\widehat{\mu}^{\ell}(w,n)\right| \geq \sqrt{\frac{2\delta}{J}}\right) \qquad (A5) \\
&\leq 4\cdot A_1\cdot\exp\left\{-\sqrt{L}\cdot h_L^r\left(A_2\cdot\min\left\{\sqrt{\frac{\delta}{J}},\, \overline{Q},\, \underline{f}/2\right\} - A_3\cdot h_L^M\right)\right\}.
\end{aligned}$$

Combining (A2), (A3) and (A5), for any $\delta > 0$ we have

$$
\begin{aligned}
&Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left|\widehat{s}^{\ell}(w,n) - s^{\ell}(w,n)\right| \geq \delta\right) \\
&\leq Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left|\widehat{\nu}^{\ell}(w,n)\right| \geq \frac{\delta}{2}\right) + Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left|\xi^{\ell}(w,n)\right| \geq \frac{\delta}{2}\right) \\
&\leq 5 \cdot A_1 \cdot \exp\left\{-\sqrt{L} \cdot h_L^r \left(A_2 \cdot \min\left\{\frac{\delta}{2}, \sqrt{\frac{\delta}{J}}, \overline{Q}, \underline{f}/2\right\} - A_3 \cdot h_L^M\right)\right\}.
\end{aligned}
$$

Our estimator $\widehat{s}(w,n)$ is

$$
\widehat{s}(w,n) = \left(\widehat{s}^1(w,n), \ldots, \widehat{s}^{d_s}(w,n)\right)'.
$$

Let b_L be the vanishing sequence used in our construction. For reasons that will become clear below, we are interested in a bound for $Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n) - s(w,n)\| \geq b_L\right)$. A Bonferroni inequality implies

$$
Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n) - s(w,n)\| \geq b_L\right) \leq \sum_{\ell=1}^{d_s} Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left|\widehat{s}^{\ell}(w,n) - s^{\ell}(w,n)\right| \geq \frac{b_L}{\sqrt{d_s}}\right).
$$

For large enough L we will have $\min\left\{\frac{b_L}{2}, \sqrt{\frac{b_L}{J}}, \overline{Q}, \underline{f}/2\right\} = \frac{b_L}{2}$ and consequently

$$
Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n) - s(w,n)\| \geq b_L\right) \leq B_1 \cdot \exp\left\{-\sqrt{L} \cdot h_L^r \left(B_2 \cdot b_L - A_3 \cdot h_L^M\right)\right\}, \quad \text{(A6)}
$$

where $B_1 \equiv 5 \cdot d_s \cdot A_1$ and $B_2 \equiv \frac{A_2}{2\sqrt{d_s}}$.

We move on to studying the properties of $\widehat{R}^q(w;n,n') - R^q(w;n,n')$. We begin with the following condition. For a given $(x,z) \equiv w$ and n group

$$
\widetilde{v}_i(w,n;h) = \left(\widetilde{v}_i^{\ell}(w,n;h)\right)_{\ell=1}^{d_s} \quad \text{and} \quad \xi_L(w,n) = \left(\xi_L^{\ell}(w,n)\right)_{\ell=1}^{d_s}
$$

and for a pair n, n' let

$$
v_i(w,n,n';h) = \left(\widetilde{v}_i(w,n;h)', \widetilde{v}_i(w,n';h)'\right)' \quad \text{and} \quad \xi_L(w,n,n') = \left(\xi_L(w,n)', \xi_L(w,n')'\right)'.
$$

For a given $(x,z) \equiv w$ and n, n' such that the relevant derivatives exist, define

$$
\begin{aligned}
v_i^q(w,n,n';h) &= \nabla_s m^q\left(s(w,n), s(w,n'); n, n'\right)' v_i(w,n,n';h), \\
\widehat{\nu}^q(w,n,n') &= \frac{1}{L \cdot h_L^r} \sum_{i=1}^{L} v_i^q(w,n,n';h_L).
\end{aligned}
$$

Using an M^{th} order approximation, the smoothness conditions in Assumption 3.1 imply the existence of a finite constant $\overline{M}'$ such that,

$$\sup_{(w,n,n')\in\mathcal{I}} |E[\widehat{\nu}^q(w,n,n')]| \leq \overline{M}' \cdot h_L^M. \tag{$A1'$}$$

Denote

$$\widehat{\mu}(w,n,n') = \Big((\widehat{s}(w,n) - s(w,n))' \quad (\widehat{s}(w,n') - s(w,n'))' \Big)'.$$

By the smoothness conditions in Assumption 3.1, for any $w \in \mathcal{W}$ and $n, n' \in \mathcal{N}$, a second order approximation yields

$$\widehat{R}^q(w;n,n') - R^q(w;n,n') = \widehat{\nu}^q(w,n,n') + \xi_L^{q,1}(w,n,n') + \xi_L^{q,2}(w,n,n'), \tag{A7}$$

where

$$\xi_L^{q,1}(w,n,n') = \nabla_s m^q\left(s(w,n), s(w,n'); n, n'\right)' \xi_L(w,n,n'),$$

$$\xi_L^{q,2}(w,n,n') = \widehat{\mu}(w,n,n')' \nabla_{ss'} m^q\left(\widetilde{s}(w,n), \widetilde{s}(w,n'); n, n'\right) \widehat{\mu}(w,n,n'),$$

where $(\widetilde{s}(w,n), \widetilde{s}(w,n'))$ lie in the line segment connecting $(\widehat{s}(w,n), \widehat{s}(w,n'))$ and $(s(w,n), s(w,n'))$. Let D be as described in Assumption 3.1. The smoothness conditions described there along with the Euclidean properties in Assumption 3.4 imply that the process

$$\{v_i^q(w,n,n';h)\colon\ (w,n,n') \in \mathcal{I},\ h > 0,\ 1 \leq i \leq L\}$$

is manageable with respect to the envelope $\left(D/\underline{f}\right) \cdot \left(\overline{S}(\cdot) + \max\left\{|\underline{s}|, |\overline{s}|\right\}\right)$, which has a moment generating function by Assumption 3.4. As before, this allows us to use the maximal inequality results in Chapter 7 of Pollard (1990) which, combined with the bias conditions in A1', imply that there exist positive constants B_1', B_2' and B_3' such that, for any $\delta > 0$,

$$Pr\left(\sup_{(w,n,n')\in\mathcal{I}} |\widehat{\nu}^q(w,n,n')| \geq \delta\right) \leq B_1' \cdot \exp\left\{-\sqrt{L} \cdot h_L^r \left(B_2' \cdot \delta - B_3' \cdot h_L^M\right)\right\}. \tag{A8}$$

Our ultimate goal here is to bound $Pr\left(\sup_{(w,n,n')\in\mathcal{I}} \left|\widehat{R}^q(w;n,n') - R^q(w;n,n')\right| \geq b_L\right)$. By (A7),

$$Pr\left(\sup_{(w,n,n')\in\mathcal{I}} \left|\widehat{R}^q(w;n,n') - R^q(w;n,n')\right| \geq b_L\right) \leq Pr\left(\sup_{(w,n,n')\in\mathcal{I}} |\widehat{\nu}^q(w,n,n')| \geq \frac{b_L}{3}\right)$$
$$+ Pr\left(\sup_{(w,n,n')\in\mathcal{I}} \left|\xi_L^{q,1}(w,n,n')\right| \geq \frac{b_L}{3}\right) + Pr\left(\sup_{(w,n,n')\in\mathcal{I}} \left|\xi_L^{q,2}(w,n,n')\right| \geq \frac{b_L}{3}\right) \tag{A9}$$

Let D and η be as described in Assumption 3.1. Then

$$
\begin{aligned}
& Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\xi_L^{q,1}(w,n,n')\right| \geq \frac{b_L}{3}\right) \\
& \leq Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n)-s(w,n)\| \geq \eta\right) + Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\|\xi_L(w,n,n')\| \geq \frac{b_L}{3\cdot D}\right) \\
& \leq Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n)-s(w,n)\| \geq \eta\right) + 2\cdot Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\xi_L(w,n)\| \geq \frac{b_L}{\sqrt{18\cdot D^2}}\right)
\end{aligned}
$$

where the last line follows from a Bonferroni inequality and the definition of $\xi_L(w,n,n')$. We also have

$$
\begin{aligned}
& Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\xi_L^{q,2}(w,n,n')\right| \geq \frac{b_L}{3}\right) \\
& \leq Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n)-s(w,n)\| \geq \eta\right) + Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\|\widehat{\mu}(w,n,n')\| \geq \sqrt{\frac{b_L}{3\cdot D}}\right) \\
& \leq Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n)-s(w,n)\| \geq \eta\right) + 2\cdot Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n)-s(w,n)\| \geq \sqrt{\frac{b_L}{6\cdot D}}\right)
\end{aligned}
$$

Combining these we have

$$
\begin{aligned}
& Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\xi_L^{q,1}(w,n,n')\right| \geq \frac{b_L}{3}\right) + Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\xi_L^{q,2}(w,n,n')\right| \geq \frac{b_L}{3}\right) \\
& \leq 2\cdot Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\widehat{s}(w,n)-s(w,n)\| \geq \min\left\{\eta,\ \sqrt{\frac{b_L}{6\cdot D}}\right\}\right) \\
& + 2\cdot Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\xi_L(w,n)\| \geq \frac{b_L}{\sqrt{18\cdot D^2}}\right)
\end{aligned}
$$

Let J be as described in (A4). Using (A5) and a Bonferroni inequality,

$$
\begin{aligned}
& Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\|\xi_L(w,n)\| \geq \frac{b_L}{\sqrt{18\cdot D^2}}\right) \leq \sum_{\ell=1}^{d_s} Pr\left(\underset{(w,n)\in\mathcal{I}}{sup}\left|\xi_L^{\ell}(w,n)\right| \geq \frac{b_L}{\sqrt{18\cdot D^2\cdot d_s}}\right) \\
& \leq 4\cdot A_1\cdot d_s\cdot\exp\left\{-\sqrt{L}\cdot h_L^r\left(A_2\cdot\min\left\{\frac{b_L^{1/2}}{(J^2\cdot 18\cdot D^2\cdot d_s)^{1/4}},\ \overline{Q},\ \underline{f}/2\right\} - A_3\cdot h_L^M\right)\right\}.
\end{aligned}
\tag{A10}
$$

For large enough L we will have $\left(\frac{B_2}{\sqrt{6\cdot D}} + \frac{A_2}{(J^2\cdot 18\cdot D^2\cdot d_s)^{1/4}}\right)\times b_L^{1/2} \leq \min\left\{\overline{Q},\ \underline{f}/2\right\}$. Using (A6) and (A10) we obtain

$$
\begin{aligned}
& Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\xi_L^{q,1}(w,n,n')\right| \geq \frac{b_L}{3}\right) + Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\xi_L^{q,2}(w,n,n')\right| \geq \frac{b_L}{3}\right) \\
& \leq B_1''\cdot\exp\left\{-\sqrt{L}\cdot h_L^r\left(B_2''\cdot b_L^{1/2} - A_3\cdot h_L^M\right)\right\},
\end{aligned}
\tag{A11}
$$

where $B_1'' \equiv 4 \cdot A_1 \cdot d_s + B_1$ and $B_2'' \equiv \min\left\{\frac{B_2}{\sqrt{6 \cdot D}},\, \frac{A_2}{(J^2 \cdot 18 \cdot D^2 \cdot d_s)^{1/4}}\right\}$. Combined with (A7) and (A9), we obtain

$$
\begin{aligned}
&Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\widehat{R}^q(w;n,n') - R^q(w;n,n')\right| \geq b_L\right)\\
&\leq B_1' \cdot \exp\left\{-\sqrt{L}\cdot h_L^r\left(B_2' \cdot \frac{b_L}{3} - B_3' \cdot h_L^M\right)\right\} + B_1'' \cdot \exp\left\{-\sqrt{L}\cdot h_L^r\left(B_2'' \cdot b_L^{1/2} - A_3 \cdot h_L^M\right)\right\}
\end{aligned}
$$

For L large enough we have $\left(\frac{B_2'}{3}\right) \cdot b_L \leq B_2'' \cdot b_L^{1/2}$ and therefore the above bound becomes

$$
Pr\left(\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\widehat{R}^q(w;n,n') - R^q(w;n,n')\right| \geq b_L\right) \quad \leq \overline{K}_1 \cdot \exp\left\{-\sqrt{L}\cdot h_L^r\left(\overline{K}_2 \cdot b_L - \overline{K}_3 \cdot h_L^M\right)\right\}, \tag{A12}
$$

where $\overline{K}_1 \equiv \max\{B_1',\, B_1''\}$, $\overline{K}_2 \equiv \min\left\{\frac{B_2'}{3},\, B_2''\right\}$ and $\overline{K}_3 \equiv \max\{A_3,\, B_3'\}$. Going back to (A7), our results also imply

$$
\begin{aligned}
\widehat{R}^q(w;n,n') - R^q(w;n,n') &= \frac{1}{L \cdot h_L^r}\sum_{i=1}^{L} v_i^q\left(w,n,n';h_L\right) + \xi_L^q(w,n,n'),\\
&\qquad \text{where} \quad \underset{(w,n,n')\in\mathcal{I}}{sup}\left\|\xi_L^q(w,n,n')\right\| = O_p\left(\frac{\log(L)^2}{L \cdot h_L^r}\right).
\end{aligned} \tag{A13}
$$

A.1.2 Proving Theorem 1

Using the main results from our previous section (equations (A12) and (A13)), we show that under our assumptions there is a linear representation for $\widehat{\mathcal{T}}_{n,n'}^q$. Recall from (5) that

$$
\widehat{\mathcal{T}}_{n,n'}^q = \frac{1}{L}\sum_{i=1}^{L}\widehat{R}^q\left(W_i;n,n'\right) \cdot \mathbb{1}\left\{\widehat{R}^q\left(W_i;n,n'\right) \geq -b_L\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\}.
$$

From here, a linear representation will immediately follow for $\widehat{\mathcal{T}}$, since it was defined as

$$
\widehat{\mathcal{T}} = \sum_{n,n'\in\mathcal{N}}\sum_{q=1}^{Q_{n,n'}}\widehat{\mathcal{T}}_{n,n'}^q.
$$

We have

$$
\widehat{\mathcal{T}}_{n,n'}^q = \frac{1}{L}\sum_{i=1}^{L}\widehat{R}^q(W_i;n,n') \cdot \mathbb{1}\left\{R^q(W_i;n,n') \geq 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\} + \zeta_L^q(n,n'),
$$

where

$$|\zeta_L^q(n,n')| \leq \underbrace{\frac{1}{L}\sum_{i=1}^{L}\left|\widehat{R}^q(W_i;n,n')\right| \cdot \mathbb{1}\left\{-2\cdot b_L \leq R^q(W_i;n,n') < 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\}}_{=\left|\zeta_L^{q,1}(n,n')\right|}$$

$$+ \underbrace{\frac{2}{L}\sum_{i=1}^{L}\left|\widehat{R}^q(W_i;n,n')\right| \cdot \mathbb{1}\left\{\left|\widehat{R}^q(W_i;n,n') - R^q(W_i;n,n')\right| \geq b_L\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\}}_{=\left|\zeta_L^{q,2}(n,n')\right|}.$$

We have

$$\begin{aligned}
&\left|\zeta_L^{q,1}(n,n')\right| \\
&\leq \frac{1}{L}\sum_{i=1}^{L}|R^q(W_i;n,n')| \cdot \mathbb{1}\left\{-2\cdot b_L \leq R^q(W_i;n,n') < 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\} \\
&+ \frac{1}{L}\sum_{i=1}^{L}\left|\widehat{R}^q(W_i;n,n') - R^q(W_i;n,n')\right| \cdot \mathbb{1}\left\{-2\cdot b_L \leq R^q(W_i;n,n') < 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\} \\
&\leq 2\cdot b_L \times \frac{1}{L}\sum_{i=1}^{L}\mathbb{1}\left\{-2\cdot b_L \leq R^q(W_i;n,n') < 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\} \\
&+ \underset{(w,n,n')\in\mathcal{I}}{sup}\left|\widehat{R}^q(w;n,n') - R^q(w;n,n')\right| \times \frac{1}{L}\sum_{i=1}^{L}\mathbb{1}\left\{-2\cdot b_L \leq R^q(W_i;n,n') < 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\} \\
&= O_p\left(\frac{\log(L)}{\sqrt{L\cdot h_L^r}}\right) \times \frac{1}{L}\sum_{i=1}^{L}\mathbb{1}\left\{-2\cdot b_L \leq R^q(W_i;n,n') < 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\},
\end{aligned} \tag{A14}$$

where the last line follows from (A13). For $b > 0$ define

$$g_i^{q,1}(b;n,n') = \mathbb{1}\left\{-2\cdot b \leq R^q(W_i;n,n') < 0\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\}.$$

And let

$$\widetilde{g}_i^{q,1}(b;n,n') = g_i^{q,1}(b;n,n') - E\left[g_i^{q,1}(b;n,n')\right], \quad \widehat{\nu}^{q,1}(n,n') = \frac{1}{L}\sum_{i=1}^{L}\widetilde{g}_i^{q,1}(b_L;n,n').$$

Since $\mathcal{N}$ is a finite set, Lemmas 2.4 and 2.14 in Pakes and Pollard (1989) implies that the process

$$\left\{\widetilde{g}_i^{q,1}(b;n,n')\colon\ b\in\mathbb{R},\ n,n'\in\mathcal{N},\ 1\leq i\leq L\right\}$$

is manageable with respect to the envelope 1. Let $\overline{b}$ and $\overline{A}$ be as described in Assumption 3.2. For large enough L we have $2\cdot b_L \leq \overline{b}$, and therefore the regularity condition described in Assumption 3.2 and the aforementioned manageability property yield

$$\underset{n,n'\in\mathcal{N}}{sup}\left|\widehat{\nu}_L^{q,1}(n,n')\right| = O_p\left(\sqrt{\frac{b_L}{L}}\right).$$

When L is large enough that $2\cdot b_L \le \overline{b}$, the regularity condition in Assumption 3.2 implies

$$\frac{1}{L}\sum_{i=1}^{L}\mathbb{1}\left\{-2\cdot b_L \le R^q(W_i;n,n') < 0\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{I}_{n,n'}\right\} = \widehat{\nu}_L^{q,1}(n,n') + \varsigma_L^{q,1}(n,n'),$$

$$\text{where}\quad \underset{n,n'\in\mathcal{N}}{sup}\left|\varsigma_L^{q,1}(n,n')\right| \le 2\cdot\overline{A}\cdot b_L.$$

Therefore,

$$\begin{aligned}\underset{n,n'\in\mathcal{N}}{sup}\left|\frac{1}{L}\sum_{i=1}^{L}\mathbb{1}\left\{-2\cdot b_L \le R^q(W_i;n,n') < 0\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{I}_{n,n'}\right\}\right| &\le O_p\left(\sqrt{\frac{b_L}{L}}\right) + 2\cdot\overline{A}\cdot b_L\\ &= O_p\left(b_L\right),\end{aligned}$$

where the last equality follows from the bandwidth convergence conditions in Assumption 3.3. Going back to (A14), this yields

$$\underset{n,n'\in\mathcal{N}}{sup}\left|\zeta_L^{q,1}(n,n')\right| \le O_p\left(b_L^2\right) + O_p\left(\frac{\log(L)\cdot b_L}{\sqrt{L\cdot h_L^r}}\right) = O_p\left(L^{-1/2-\epsilon}\right)\quad\text{for some } \epsilon>0,$$

where the last equality follows from the bandwidth convergence properties in Assumption 3.3. Using (A12) and (A13),

$$\begin{aligned}\underset{n,n'\in\mathcal{N}}{sup}\left|\zeta_L^{q,2}(n,n')\right| &\le \underset{(w,n,n')\in\mathcal{I}}{sup}\left|\widehat{R}^q(w;n,n')\right|\times\mathbb{1}\left\{\underset{(w,n,n')\in\mathcal{I}}{sup}\left|\widehat{R}^q(w;n,n') - R^q(w;n,n')\right| \ge b_L\right\}\\ &= O_p(1)\times O_p\left(\overline{K}_1^{1/2}\cdot\exp\left\{-\frac{1}{2}\sqrt{L}\cdot h_L^r\left(\overline{K}_2\cdot b_L - \overline{K}_3\cdot h_L^M\right)\right\}\right)\\ &= O_p\left(L^{-1/2-\epsilon}\right)\ \forall\ \epsilon>0.\end{aligned}$$

From here we conclude that

$$\begin{aligned}\widehat{\mathcal{T}}_{n,n'}^q &= \frac{1}{L}\sum_{i=1}^{L}\widehat{R}^q(W_i;n,n')\cdot\mathbb{1}\left\{R^q(W_i;n,n')\ge 0\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{W}\right\} + \zeta_L^q(n,n'),\\ &\quad\text{where}\quad \underset{n,n'\in\mathcal{N}}{sup}\left|\zeta_L^q(n,n')\right| = O_p\left(L^{-1/2-\epsilon}\right)\ \text{for some } \epsilon>0.\end{aligned}\tag{A15}$$

This can be re-expressed as

$$\begin{aligned}\widehat{\mathcal{T}}_{n,n'}^q &= \frac{1}{L}\sum_{i=1}^{L}\max\left\{R^q(W_i;n,n'),\,0\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{W}\right\}\\ &\quad+\frac{1}{L}\sum_{i=1}^{L}\left(\widehat{R}^q(W_i;n,n') - R^q(W_i;n,n')\right)\cdot\mathbb{1}\left\{R^q(W_i;n,n')\ge 0\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{W}\right\} + \zeta_L^q(n,n'),\end{aligned}\tag{A16}$$

We will begin by studying the second term. For each $\ell = 1,\ldots,d_s$ and any two observations i,j in $1,\ldots,L$ denote

$$\begin{aligned}
v^{\ell}\left(U_i,U_j;n,h\right) &= \left(\frac{S^{\ell}(Y_j,W_i,n)-s^{\ell}(W_i,n)}{f_{X,N}(X_i,n)}\right)\cdot\mathcal{H}\left(X_j-X_i;h\right)\cdot\mathbb{1}\left\{N_j=n\right\},\\
v(U_i,U_j;n,h) &= \left(v^{1}\left(U_i,U_j;n,h\right),\ldots,v^{d_s}\left(U_i,U_j;n,h\right)\right)',\\
v\left(U_i,U_j,n,n';h\right) &= \left(v(U_i,U_j;n,h)',\, v(U_i,U_j;n',h)'\right)',
\end{aligned}$$

and let

$$f^{q}\left(U_i,U_j,n,n';h\right) =$$
$$\frac{1}{h^{r}}\cdot\nabla_s m^{q}\left(s(W_i,n),s(W_i,n');n,n'\right)' v\left(U_i,U_j,n,n';h\right)\cdot\mathbb{1}\left\{R^{q}(W_i;n,n')\geq 0\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{W}\right\}$$

Using (A13), we have

$$\begin{aligned}
&\frac{1}{L}\sum_{i=1}^{L}\left(\widehat{R}^{q}(W_i;n,n')-R^{q}(W_i;n,n')\right)\cdot\mathbb{1}\left\{R^{q}(W_i;n,n')\geq 0\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{W}\right\}\\
&=\frac{1}{L^{2}}\sum_{i=1}^{L}\sum_{j=1}^{L}f^{q}\left(U_i,U_j,n,n';h_L\right)+\varrho_L^{q}(n,n'),
\end{aligned}\tag{A17}$$

where

$$\underset{n,n'\in\mathcal{N}}{sup}\left|\varrho_L^{q}(n,n')\right| = O_p\left(\frac{\log(L)^2}{L\cdot h_L^{r}}\right) = O_p\left(L^{-1/2-\epsilon}\right)\quad\text{for some } \epsilon>0.$$

Let us analyze the properties of the first term in the right hand side of (A17) which is a second-order U process (Serfling (1980), Sherman (1994)). Denote

$$\mu_L^{q}(n,n') = E\left[f^{q}\left(U_i,U_j,n,n';h_L\right)\right].$$

Given the smoothness conditions in Assumption 3.1 there exists a constant $\overline{D}$ such that for any $i\neq j$,

$$\underset{n,n'\in\mathcal{N}}{sup}\left|E\left[f^{q}\left(U_i,U_j,n,n';h_L\right)\big|U_i\right]\right|\leq\overline{D}\cdot h_L^{M}.\tag{A18}$$

And a dominated convergence argument and iterated expectations imply

$$\underset{n,n'\in\mathcal{N}}{sup}\left|\mu_L^{q}(n,n')\right|\leq\overline{D}\cdot h_L^{M}.$$

Let

$$\begin{aligned}
\widetilde{f}^{q}\left(U_i,U_j,n,n';h_L\right) &= f^{q}\left(U_i,U_j,n,n';h_L\right)-\mu_L^{q}(n,n'),\\
\widetilde{g}^{q}\left(U_i,U_j,n,n';h_L\right) &= \frac{\widetilde{f}^{q}\left(U_i,U_j,n,n';h_L\right)+\widetilde{f}^{q}\left(U_j,U_i,n,n';h_L\right)}{2},\\
V_L^{q}(n,n') &= \binom{L}{2}^{-1}\sum_{i<j}\widetilde{g}^{q}\left(U_i,U_j,n,n';h_L\right).
\end{aligned}$$

Note first that under our previous assumptions,

$$\underset{n,n'\in\mathcal{N}}{sup}\left|\frac{1}{L^2}\sum_{i=1}^{L} f^q\left(U_i,U_i,n,n';h_L\right)\right| = O_p\left(\frac{1}{L\cdot h_L^r}\right) = o_p\left(L^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon>0.$$

Combined with the vanishing properties of $\mu_L^q(n,n')$, this means that we can express

$$\frac{1}{L^2}\sum_{i=1}^{L}\sum_{j=1}^{L} f^q\left(U_i,U_j,n,n';h_L\right) = \left(\frac{L-1}{L}\right)\cdot V_L^q(n,n') + \vartheta_L(n,n'),$$

$$\text{where} \quad \underset{n,n'\in\mathcal{N}}{sup}\left|\vartheta_L(n,n')\right| = O\left(h_L^M\right) + O_p\left(\frac{1}{L\cdot h_L^r}\right) = o_p\left(L^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon>0.$$

Fix u. Symmetry of $\widetilde{g}^q$ implies $E_U\left[\widetilde{g}^q\left(u,U,n,n';h_L\right)\right] = E_U\left[\widetilde{g}^q\left(U,u,n,n';h_L\right)\right]$. We will denote

$$\theta_L^q\left(u,n,n'\right) = E_U\left[\widetilde{g}^q\left(u,U,n,n';h_L\right)\right].$$

Note that $E\left[\theta_L^q\left(U,n,n'\right)\right]=0$. Let

$$t^q\left(U_i,U_j,n,n';h_L\right) = \widetilde{g}^q\left(U_i,U_j,n,n';h_L\right) - \theta_L^q\left(U_i,n,n'\right) - \theta_L^q\left(U_j,n,n'\right),$$

$$V_L^{q,2}(n,n') = \binom{L}{2}^{-1}\sum_{i<j} t^q\left(U_i,U_j,n,n';h_L\right).$$

The properties of $\mu_L^q(n,n')$ and the Hoeffding decomposition of V_L^q ((Serfling (1980)) imply that

$$V_L^q(n,n') = \frac{2}{L}\sum_{i=1}^{L}\theta_L^q(U_i,n,n') + V_L^{q,2}(n,n') + \widetilde{\tau}_L^q(n,n'), \quad \text{where} \quad \underset{n,n'\in\mathcal{N}}{sup}\left|\widetilde{\tau}_L^q(n,n')\right| = O\left(h_L^M\right).$$

$V_L^{q,2}(n,n')$ is a *degenerate* U-statistic of order 2 and it satisfies $V_L^{q,2}(n,n') = O_p\left(\frac{1}{L\cdot h_L^r}\right)$ (see Serfling (1980), Sherman (1994)). Since $\mathcal{N}$ is finite, this property holds uniformly over $\mathcal{N}$. Let

$$\Delta_L^q\left(u,n,n'\right) = E_U\left[f^q\left(U,u,n,n';h_L\right)\right].$$

Using (A18), our smoothness conditions imply that the last result can be re-expressed as

$$V_L^q(n,n') = \frac{1}{L}\sum_{i=1}^{L}\left(\Delta_L^q\left(U_i,n,n'\right) - E\left[\Delta_L^q\left(U_i,n,n'\right)\right]\right) + \tau_L^q(n,n'),$$

$$\text{where} \quad \underset{n,n'\in\mathcal{N}}{sup}\left|\tau_L^q(n,n')\right| = O_p\left(\frac{1}{L\cdot h_L^r}\right) + O\left(h_L^M\right) = O_p\left(L^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon>0. \tag{A19}$$

Taking these results back to (A16) we obtain the main result in this section. Let

$$\lambda_L^q(U_i; n, n') =$$
$$\left(\max\{R^q(W_i; n, n'), 0\} \cdot \mathbb{1}\{W_i \in \mathcal{W}\} - \mathcal{T}_{n,n'}^q\right) + \left(\Delta_L^q(U_i, n, n') - E\left[\Delta_L^q(U_i, n, n')\right]\right).$$

By construction, $E\left[\lambda_L^q(U_i; n, n')\right] = 0$. A quick inspection also shows that

$$Pr\left(R^q(W_i; n, n') < 0 \middle| W_i \in \mathcal{W}\right) = 1 \implies \lambda_L^q(U_i; n, n') = 0 \ w.p.1.$$

That is, the influence function $\lambda_L^q(U_i; n, n')$ vanishes almost surely if the condition $R^q(W; n, n') \leq 0$ is satisfied as a strict inequality almost everywhere over our testing range. Combining (A16) and (A19), we obtain

$$\widehat{\mathcal{T}}_{n,n'}^q = \mathcal{T}_{n,n'}^q + \frac{1}{L}\sum_{i=1}^{L} \lambda_L^q(U_i; n, n') + \xi_L^q(n, n'), \ \text{ where}$$

(i) $E\left[\lambda_L^q(U_i; n, n')\right] = 0$.

(ii) If $Pr\left(R^q(W; n, n') < 0 \middle| W \in \mathcal{W}\right) = 1$, then $\lambda_L^q(U_i; n, n') = 0$ w.p.1.

(iii) $|\xi_L^q(n, n')| = O_p\left(L^{-1/2-\epsilon}\right)$ for some $\epsilon > 0$.

This proves Theorem 1. $\square$

Let

$$\lambda_L(U_i) = \sum_{q=1}^{Q_{n,n'}} \sum_{n,n' \in \mathcal{N}} \lambda_L^q(U_i; n, n').$$

Note that

(i) $E\left[\lambda_L(U_i)\right] = 0$.

(ii) If $Pr\left(R^q(W; n, n') < 0 \middle| W \in \mathcal{W}\right) = 1$ for each $n, n' \in \mathcal{N}$ and $q = 1, \ldots, Q_{n,n'}$, then $\lambda_L(U_i) = 0$ w.p.1.

By our previous results, it follows that

$$\widehat{\mathcal{T}} = \mathcal{T} + \frac{1}{L}\sum_{i=1}^{L} \lambda_L(U_i) + \xi_L, \quad \text{where} \quad \xi_L = O_p\left(L^{-1/2-\epsilon}\right) \text{ for some } \epsilon > 0.$$

Let

$$\sigma_L^2 = \text{Var}\left(\lambda_L(U_i)\right).$$

Because $\sigma_L^2 = 0$ if our (weak) inequalities are satisfied almost surely as strict inequalities, σ_L^2 is the relevant measure for the slackness in our test.

A.2 Construction of the test statistics

A.2.1 Estimating $\boldsymbol{\sigma_L^2}$

An estimator for σ_L^2 can be constructed by first estimating the influence function $\lambda_L(U_i)$. For $i \neq j$ let

$$\widehat{f}^q(U_j, U_i, n, n'; h_L) =$$
$$\frac{1}{h_L^r} \cdot \nabla_s m^q\left(\widehat{s}(W_j, n), \widehat{s}(W_j, n'); n, n'\right)' \widehat{v}\left(U_j, U_i, n, n'; h\right) \cdot \mathbb{1}\left\{\widehat{R}^q(W_j; n, n') \geq -b_L\right\} \cdot \mathbb{1}\left\{W_j \in \mathcal{W}\right\}$$
$$\times \frac{1}{\widehat{f}_{X,N}(X_j, n)} \cdot \mathcal{H}\left(X_i - X_j; h_L\right) \cdot \mathbb{1}\left\{N_j = N_i\right\}$$

We estimate $\Delta_L^q(U_i, n, n')$ as

$$\widehat{\Delta}_L^q(U_i, n, n') = \frac{1}{L-1} \sum_{j \neq i} \widehat{f}^q\left(U_j, U_i, n, n'; h_L\right).$$

And from here our estimators for $\lambda_L^q(U_i; n, n')$ and $\lambda_L(U_i)$ are

$$\begin{aligned}
\widehat{\lambda}_L^q(U_i; n, n') &= \left(\widehat{R}^q(W_i; n, n') \cdot \mathbb{1}\left\{\widehat{R}^q(W_i; n, n') \geq -b_L\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\} - \widehat{\mathcal{T}}_{n,n'}^q\right) \\
&\quad + \left(\widehat{\Delta}_L^q(U_i, n, n') - \widehat{E}\left[\widehat{\Delta}_L^q(U_i, n, n')\right]\right), \\
\widehat{\lambda}_L(U_i) &= \sum_{q=1}^{Q_{n,n'}} \sum_{n,n' \in \mathcal{N}} \widehat{\lambda}_L^q(U_i; n, n').
\end{aligned}$$

From here we can estimate σ_L^2 as

$$\widehat{\sigma}_L^2 = \frac{1}{L} \sum_{i=1}^{L} \widehat{\lambda}_L^2(U_i).$$

Under Assumptions 3.1-3.4 (the conditions leading to Theorem 1), we will have $\left|\widehat{\sigma}_L^2 - \sigma_L^2\right| = o_p(1)$.

A.2.2 Detailed expressions for the auction models test statistics

Recall that our test statistic has a generic expression of the form

$$\widehat{t}_L = \frac{\sqrt{L} \cdot \widehat{\mathcal{T}}}{\max\left\{\widehat{\sigma}_L, \kappa_L\right\}}$$

Here we describe the precise expressions for $\widehat{\mathcal{T}}$ and $\widehat{\sigma}_L$ for each of the auction models examples. For $s \in (0,1)$ and $1 \leq k \leq n$ denote

$$\nabla \psi_{k:n}^{-1}(s) = \frac{(n-k)! \cdot (k-1)!}{n! \cdot \left[\psi_{k:n}^{-1}(s)\right]^{k-1} \cdot \left(1 - \left[\psi_{k:n}^{-1}(s)\right]\right)^{n-k}}$$

IPV with fixed N

For equation (10) we have $Q_n = n-2$ and for each $q = 1, \dots, n-2$,

$$
\begin{aligned}
&\widehat{R}^q(W_j; n) = \psi^{-1}_{n-1:n}\left(\widehat{G}^{\Delta}_{n:n}(Z_j|X_j)\right) - \psi^{-1}_{q:n}\left(\widehat{G}_{q:n}(Z_j|X_j)\right), \\
&\widehat{f}^q\left(U_j, U_i, n; h_L\right) = \\
&\left\{\frac{1}{h_L^r \cdot \widehat{f}_{X,N}(X_j, n)} \cdot \left[\nabla\psi^{-1}_{n-1:n}\left(\widehat{G}^{\Delta}_{n:n}(Z_j|X_j)\right) \cdot \left(\mathbb{1}\left\{B_{N:N,i} + \Delta \le Z_j\right\} - \widehat{G}^{\Delta}_{n:n}(Z_j|X_j)\right)\right.\right. \\
&\left.-\nabla\psi^{-1}_{q:n}\left(\widehat{G}_{q:n}(Z_j|X_j)\right) \cdot \left(\mathbb{1}\left\{B_{q:N,i} \le Z_j\right\} - \widehat{G}_{q:n}(Z_j|X_j)\right)\right] \cdot \mathbb{1}\left\{N_i = n\right\} \cdot \mathcal{H}\left(X_i - X_j; h_L\right) \\
&\left.\times \mathbb{1}\left\{\widehat{R}^q(W_j; n) \ge -b_L\right\} \cdot \mathbb{1}\left\{W_j \in \mathcal{W}\right\}\right\}
\end{aligned}
$$

From here,

$$
\begin{aligned}
&\widehat{\mathcal{T}}^q_n = \frac{1}{L}\sum_{i=1}^{L} \widehat{R}^q(W_i; n) \cdot \mathbb{1}\left\{\widehat{R}^q(W_i; n) \ge -b_L\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\}, \quad \widehat{\mathcal{T}} = \sum_{q=1}^{n-2}\sum_{n\in\mathcal{N}} \widehat{\mathcal{T}}^q_n, \\
&\widehat{\Delta}^q_L(U_i; n) = \frac{1}{L-1}\sum_{j\neq i} \widehat{f}^q\left(U_j, U_i, n; h_L\right), \\
&\widehat{\lambda}^q_L(U_i; n) = \left(\widehat{R}^q(W_i; n) \cdot \mathbb{1}\left\{\widehat{R}^q(W_i; n) \ge -b_L\right\} \cdot \mathbb{1}\left\{W_i \in \mathcal{W}\right\} - \widehat{\mathcal{T}}^q_n\right) \\
&+ \left(\widehat{\Delta}^q_L(U_i; n) - \widehat{E}\left[\widehat{\Delta}^q_L(U_i; n)\right]\right), \quad \widehat{\lambda}_L(U_i) = \sum_{q=1}^{n-2}\sum_{n\in\mathcal{N}} \widehat{\lambda}^q_L(U_i; n), \quad \widehat{\sigma}^2_L = \frac{1}{L}\sum_{i=1}^{L} \widehat{\lambda}^2_L(U_i),
\end{aligned}
$$

IPV and $V \perp N$

For equation (13) we have $Q_{n,n'} = 1$ and

$$
\begin{aligned}
&\widehat{R}(W_j; n, n') = \psi^{-1}_{n'-1:n'}\left(\widehat{G}^{\Delta}_{n':n'}(Z_j|X_j)\right) - \psi_{n-1:n}\left(\widehat{G}_{n-1:n}(Z_j|X_j)\right), \\
&\left\{\frac{1}{h_L^r} \cdot \left[\nabla\psi^{-1}_{n'-1:n'}\left(\widehat{G}^{\Delta}_{n':n'}(Z_j|X_j)\right) \cdot \frac{\left(\mathbb{1}\left\{B_{N:N,i} + \Delta \le Z_j\right\} - \widehat{G}^{\Delta}_{n':n'}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j, n')} \cdot \mathbb{1}\left\{N_i = n'\right\}\right.\right. \\
&\left.-\nabla\psi^{-1}_{n-1:n}\left(\widehat{G}_{n-1:n}(Z_j|X_j)\right) \cdot \frac{\left(\mathbb{1}\left\{B_{N-1:N,i} \le Z_j\right\} - \widehat{G}_{n-1:n}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j, n)} \cdot \mathbb{1}\left\{N_i = n\right\}\right] \\
&\left.\times \mathcal{H}\left(X_i - X_j; h_L\right) \cdot \mathbb{1}\left\{\widehat{R}(W_j; n, n') \ge -b_L\right\} \cdot \mathbb{1}\left\{W_j \in \mathcal{W}\right\}\right\} \cdot \mathbb{1}\left\{n > n'\right\}.
\end{aligned}
$$

From here,

$$
\begin{aligned}
&\widehat{\mathcal{T}}_{n,n'} = \frac{1}{L}\sum_{i=1}^{L}\widehat{R}(W_i;n,n')\cdot\mathbb{1}\left\{\widehat{R}(W_i;n,n')\geq -b_L\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{W}\right\}, \quad \widehat{\mathcal{T}} = \sum_{n\in\mathcal{N}}\widehat{\mathcal{T}}_{n,n'},\\
&\widehat{\Delta}_L(U_i;n,n') = \frac{1}{L-1}\sum_{j\neq i}\widehat{f}\left(U_j,U_i,n,n';h_L\right),\\
&\widehat{\lambda}_L(U_i;n,n') = \left(\widehat{R}(W_i;n,n')\cdot\mathbb{1}\left\{\widehat{R}(W_i;n,n')\geq -b_L\right\}\cdot\mathbb{1}\left\{W_i\in\mathcal{W}\right\} - \widehat{\mathcal{T}}_{n,n'}\right)\\
&+\left(\widehat{\Delta}_L(U_i;n,n') - \widehat{E}\left[\widehat{\Delta}_L(U_i;n,n')\right]\right), \quad \widehat{\lambda}_L(U_i) = \sum_{n,n'\in\mathcal{N}}\widehat{\lambda}_L(U_i;n,n'),\\
&\widehat{\sigma}_L^2 = \frac{1}{L}\sum_{i=1}^{L}\widehat{\lambda}_L^2(U_i),
\end{aligned} \tag{A20}
$$

Nonnegatively correlated values and $V \perp N$

For equation (12) we have $Q_{n,n'} = 1$ and

$$
\begin{aligned}
&\widehat{R}(W_j;n,n') = \left(\psi^{-1}_{n'-1:n'}\left(\widehat{G}_{n':n'}(Z_j|X_j)\right) - \psi^{-1}_{n-1:n}\left(\widehat{G}_{n:n}(Z_j|X_j)\right)\right)\cdot\mathbb{1}\left\{n>n'\right\},\\
&\widehat{f}\left(U_j,U_i,n,n';h_L\right) =\\
&\left\{\frac{1}{h_L^r}\cdot\left[\nabla\psi^{-1}_{n'-1:n'}\left(\widehat{G}_{n':n'}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N:N,i}\leq Z_j\right\} - \widehat{G}_{n':n'}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n')}\cdot\mathbb{1}\left\{N_i=n'\right\}\right.\right.\\
&\left.-\nabla\psi^{-1}_{n-1:n}\left(\widehat{G}_{n:n}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N:N,i}\leq Z_j\right\} - \widehat{G}_{n:n}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n)}\cdot\mathbb{1}\left\{N_i=n\right\}\right]\\
&\left.\times\mathcal{H}\left(X_i-X_j;h_L\right)\cdot\mathbb{1}\left\{\widehat{R}(W_j;n,n')\geq -b_L\right\}\cdot\mathbb{1}\left\{W_j\in\mathcal{W}\right\}\right\}\cdot\mathbb{1}\left\{n>n'\right\}.
\end{aligned}
$$

For equation (14), we have

$$
\begin{aligned}
&\widehat{R}(W_j;n,n') = \left(\psi^{-1}_{n'-1:n'}\left(\widehat{G}^{\Delta}_{n':n'}(Z_j|X_j)\right) - \psi^{-1}_{n-1:n}\left(\widehat{G}_{n-1:n}(Z_j|X_j)\right)\right)\cdot\mathbb{1}\left\{n>n'\right\},\\
&\widehat{f}\left(U_j,U_i,n,n';h_L\right) =\\
&\left\{\frac{1}{h_L^r}\cdot\left[\nabla\psi^{-1}_{n'-1:n'}\left(\widehat{G}^{\Delta}_{n':n'}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N:N,i}+\Delta\leq Z_j\right\} - \widehat{G}^{\Delta}_{n':n'}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n')}\cdot\mathbb{1}\left\{N_i=n'\right\}\right.\right.\\
&\left.-\nabla\psi^{-1}_{n-1:n}\left(\widehat{G}_{n-1:n}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N-1:N,i}\leq Z_j\right\} - \widehat{G}_{n-1:n}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n)}\cdot\mathbb{1}\left\{N_i=n\right\}\right]\\
&\left.\times\mathcal{H}\left(X_i-X_j;h_L\right)\cdot\mathbb{1}\left\{\widehat{R}(W_j;n,n')\geq -b_L\right\}\cdot\mathbb{1}\left\{W_j\in\mathcal{W}\right\}\right\}\cdot\mathbb{1}\left\{n>n'\right\}.
\end{aligned}
$$

In both cases $\widehat{\mathcal{T}}$ and $\widehat{\sigma}_L$ are constructed following the generic expression (A20).

IPV without independence between V and N

For equation (15) we have $Q_{n,n'} = 1$ and

$$\widehat{R}(W_j;n,n') = \left(\psi^{-1}_{n-1:n}\left(\widehat{G}_{n:n}(Z_j|X_j)\right) - \psi^{-1}_{n'-1:n'}\left(\widehat{G}_{n':n'}(Z_j|X_j)\right)\right)\cdot \mathbb{1}\left\{n > n'\right\},$$

$$\widehat{f}\left(U_j,U_i,n,n';h_L\right) =$$

$$\left\{\frac{1}{h_L^r}\cdot\left[\nabla\psi^{-1}_{n-1:n}\left(\widehat{G}_{n:n}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N:N,i}\le Z_j\right\} - \widehat{G}_{n:n}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n)}\cdot\mathbb{1}\left\{N_i = n\right\}\right.\right.$$

$$\left.-\nabla\psi^{-1}_{n'-1:n'}\left(\widehat{G}_{n':n'}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N:N,i}\le Z_j\right\} - \widehat{G}_{n':n'}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n')}\cdot\mathbb{1}\left\{N_i = n'\right\}\right]$$

$$\left.\times\,\mathcal{H}\left(X_i - X_j;h_L\right)\cdot\mathbb{1}\left\{\widehat{R}(W_j;n,n')\ge -b_L\right\}\cdot\mathbb{1}\left\{W_j\in\mathcal{W}\right\}\right\}\cdot\mathbb{1}\left\{n > n'\right\}$$

In the case of (16),

$$\widehat{R}(W_j;n,n') = \left(\psi^{-1}_{n-1:n}\left(\widehat{G}^{\Delta}_{n:n}(Z_j|X_j)\right) - \psi^{-1}_{n'-1:n'}\left(\widehat{G}_{n'-1:n'}(Z_j|X_j)\right)\right)\cdot \mathbb{1}\left\{n > n'\right\},$$

$$\widehat{f}\left(U_j,U_i,n,n';h_L\right) =$$

$$\left\{\frac{1}{h_L^r}\cdot\left[\nabla\psi^{-1}_{n-1:n}\left(\widehat{G}^{\Delta}_{n:n}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N:N,i}+\Delta\le Z_j\right\} - \widehat{G}^{\Delta}_{n:n}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n)}\cdot\mathbb{1}\left\{N_i = n\right\}\right.\right.$$

$$\left.-\nabla\psi^{-1}_{n'-1:n'}\left(\widehat{G}_{n'-1:n'}(Z_j|X_j)\right)\cdot\frac{\left(\mathbb{1}\left\{B_{N-1:N,i}\le Z_j\right\} - \widehat{G}_{n'-1:n'}(Z_j|X_j)\right)}{\widehat{f}_{X,N}(X_j,n')}\cdot\mathbb{1}\left\{N_i = n'\right\}\right]$$

$$\left.\times\,\mathcal{H}\left(X_i - X_j;h_L\right)\cdot\mathbb{1}\left\{\widehat{R}(W_j;n,n')\ge -b_L\right\}\cdot\mathbb{1}\left\{W_j\in\mathcal{W}\right\}\right\}\cdot\mathbb{1}\left\{n > n'\right\}$$

In both instances we construct $\widehat{\mathcal{T}}$ and $\widehat{\sigma}_L$ using the generic expression (A20).

A.3 Recomputing the tests in Section 4.3.5 with alternative bandwidths

Here we present the results for the test-statistics in Table 5 using alternative values for the constants c_h and c_b used in the construction of h_L and b_L, respectively. As we did in Table 5, c_k was set to 10^{-1} throughout.

Table A.1: Test Results on Auction Timber Data using alternative values of c_h and c_b

$\boldsymbol{c_h = 0.25}$, $\boldsymbol{c_b = 10^{-3}}$						
Eq.	**Test of**	**Auctions**	N	**Bid assump**	$\widehat{t}$	**Outcome**†
10	IPV	English	Fixed	H&T bidding	**1.94**	Reject‡
12	$V \perp N$	English	Variable	$B_{n:n} = V_{n-1:n}$	**0.67**	Fail to Reject
15	IPV	English	Variable	$B_{n:n} = V_{n-1:n}$	**2.52**	Reject
$\boldsymbol{c_h = 0.25}$, $\boldsymbol{c_b = 1}$						
Eq.	**Test of**	**Auctions**	N	**Bid assump**	$\widehat{t}$	**Outcome**†
10	IPV	English	Fixed	H&T bidding	**1.93**	Reject‡
12	$V \perp N$	English	Variable	$B_{n:n} = V_{n-1:n}$	**0.49**	Fail to Reject
15	IPV	English	Variable	$B_{n:n} = V_{n-1:n}$	**2.39**	Reject
$\boldsymbol{c_h = 0.5}$, $\boldsymbol{c_b = 10^{-3}}$						
Eq.	**Test of**	**Auctions**	N	**Bid assump**	$\widehat{t}$	**Outcome**†
10	IPV	English	Fixed	H&T bidding	**16.15**	Reject
12	$V \perp N$	English	Variable	$B_{n:n} = V_{n-1:n}$	**0.71**	Fail to Reject
15	IPV	English	Variable	$B_{n:n} = V_{n-1:n}$	**9.25**	Reject
$\boldsymbol{c_h = 0.5}$, $\boldsymbol{c_b = 1}$						
Eq.	**Test of**	**Auctions**	N	**Bid assump**	$\widehat{t}$	**Outcome**†
10	IPV	English	Fixed	H&T bidding	**15.92**	Reject
12	$V \perp N$	English	Variable	$B_{n:n} = V_{n-1:n}$	**0.10**	Fail to Reject
15	IPV	English	Variable	$B_{n:n} = V_{n-1:n}$	**9.07**	Reject

(†) Critical values for rejection are 1.645 for $\alpha = 5\%$ and 2.326 for $\alpha = 1\%$.

(‡) Denotes rejection at level $\alpha = 5\%$. All other results hold at $\alpha = 1\%$.

B Appendix – Auctions Models

B.1 Proof of Proposition 1

Under the Haile-and-Tamer bidding assumptions, $B_{k:n} \leq V_{k:n}$, which implies $G_{k:n}(v|x) \geq F_{k:n}(v|x)$; and $V_{n-1:n} \leq B_{n:n} + \Delta$, which implies $F_{n-1:n}(v|x) \geq G^{\Delta}_{n:n}(v|x)$. So under IPV,

$$\begin{aligned} \psi^{-1}_{k:n}\left(G_{k:n}(v|x)\right) &\geq \psi^{-1}_{k:n}\left(F_{k:n}(v|x)\right) = F_V(v|x) \\ &= \psi^{-1}_{n-1:n}\left(F_{n-1:n}(v|x)\right) \geq \psi^{-1}_{n-1:n}\left(G^{\Delta}_{n:n}(v|x)\right) \end{aligned}$$

For the second part, fix n, x, and v, and let $\Pr(v|m)$ denote the probability that $V_i < v$, conditional on exactly m of the other $n-1$ valuations being less than v. Suppress the dependence of value distributions on x. Let P_i denote the probability that exactly i valuations are greater than or equal to v, so $P_0 = F_{n:n}(v)$, $P_n = 1 - F_{1:n}(v)$, and $P_i = F_{n-i:n}(v) - F_{n-i+1:n}(v)$ for $1 \leq i < n$. Let $\Pr(m)$ be the probability that $V_1, \ldots, V_m \geq v$ and $V_{m+1}, \ldots, V_{n-1} < v$. By symmetry, $P_{i+1} = {}_nC_{i+1} \Pr(i)(1 - \Pr(v|n-1-i))$ and $P_i = {}_nC_i \Pr(i) \Pr(v|n-1-i)$; so

$$\frac{\frac{1}{{}_nC_{i+1}} P_{i+1}}{\frac{1}{{}_nC_i} P_i} = \frac{\Pr(i)(1 - \Pr(v|n-1-i))}{\Pr(i)\Pr(v|n-1-i)} = \frac{1 - \Pr(v|n-1-i)}{\Pr(v|n-1-i)}$$

By assumption, this is weakly increasing in i, and strictly increasing for some i.

Let $p = \psi^{-1}_{n-1:n}(F_{n-1:n}(v))$, and let $P^I_i = {}_nC_i p^{n-i}(1-p)^i$, and $F^I_{k:n}(v) = \psi_{k:n}(p) = \sum_{i=0}^{n-k} P^I_i$, so P^I_i and $F^I_{k:n}$ are what P_i and $F_{k:n}$ would be if valuations were independent draws from the distribution $F_V(\cdot) = \psi^{-1}_{n-1:n}(F_{n-1:n}(\cdot))$. By construction,

$$\frac{\frac{1}{{}_nC_{i+1}} P^I_{i+1}}{\frac{1}{{}_nC_i} P^I_i} = \frac{1-p}{p}$$

and therefore does not vary with i. Note that $P_0 + P_1 = F_{n-1:n}(v) = F^I_{n-1:n}(v) = P^I_0 + P^I_1$.

Claim 1. $P_0 > P^I_0$.

Proof is by contradiction. Since $P_0 + P_1 = P^I_0 + P^I_1$, if $P_0 \leq P^I_0$, then $P_1 \geq P^I_1$. Then

$$\frac{\frac{1}{{}_nC_2} P_2}{\frac{1}{{}_nC_1} P_1} \geq \frac{\frac{1}{{}_nC_1} P_1}{\frac{1}{{}_nC_0} P_0} \geq \frac{\frac{1}{{}_nC_1} P^I_1}{\frac{1}{{}_nC_0} P^I_0} = \frac{1-p}{p} = \frac{\frac{1}{{}_nC_2} P^I_2}{\frac{1}{{}_nC_1} P^I_1}$$

and so since $P_1 \geq P^I_1$ and $\frac{P_2}{P_1} \geq \frac{P^I_2}{P^I_1}$, then $P_2 \geq P^I_2$. Similarly, since $\frac{P_3}{P_2} \geq \frac{P_2}{P_1} \geq \frac{P^I_2}{P^I_1} = \frac{P^I_3}{P^I_2}$, $P_3 \geq P^I_3$; and likewise, $P_i \geq P^I_i$ for every $i > 3$, with at least one strict inequality due to the requirement that Assumption 4.2 holds strictly. This leads to $\sum_{i=0}^n P_i > \sum_{i=0}^n P^I_i$, which is a contradiction since both must be equal to 1.

Claim 2. $P_2 < P_2^I$.

Since $P_1 < P_1^I$, if $P_2 \geq P_2^I$, then $\frac{P_2}{P_1} > \frac{P_2^I}{P_1^I}$, and so $\frac{P_3}{P_2} > \frac{P_3^I}{P_2^I}$ giving $P_3 > P_3^I$ and so on; this would give $P_0 + P_1 = P_0^I + P_1^I$, $P_2 \geq P_2^I$, and $P_i > P_i^I$ for $i \geq 3$, yielding a contradiction.

(Note that if $\Pr(v|m)$ is only weakly increasing in m, everything up to here applies as weak inequalities and $P_2 \leq P_2^I$, which will be used below in the proof of Proposition 2.)

Claim 3. If $P_k > P_k^I$, then $P_{k'} > P_{k'}^I$ for all $k' > k$.

We know that $P_1 < P_1^I$. Let j denote the smallest $i > 0$ such that $P_i > P_i^I$. This means $P_j > P_j^I$ but $P_{j-1} \leq P_{j-1}^I$, and therefore

$$\frac{\frac{1}{{}_nC_j} P_j}{\frac{1}{{}_nC_{j-1}} P_{j-1}} > \frac{\frac{1}{{}_nC_j} P_j^I}{\frac{1}{{}_nC_{j-1}} P_{j-1}^I} = \frac{1-p}{p}$$

Which means that

$$\frac{\frac{1}{{}_nC_{j+1}} P_{j+1}}{\frac{1}{{}_nC_j} P_j} \geq \frac{\frac{1}{{}_nC_j} P_j}{\frac{1}{{}_nC_{j-1}} P_{j-1}} > \frac{1-p}{p} = \frac{\frac{1}{{}_nC_{j+1}} P_{j+1}^I}{\frac{1}{{}_nC_j} P_j^I}$$

and so $P_j > P_j^I$ and $\frac{P_{j+1}}{P_j} > \frac{P_{j+1}^I}{P_j^I}$, meaning $P_{j+1} > P_{j+1}^I$. Likewise,

$$\frac{\frac{1}{{}_nC_{j+2}} P_{j+2}}{\frac{1}{{}_nC_{j+1}} P_{j+1}} \geq \frac{\frac{1}{{}_nC_{j+1}} P_{j+1}}{\frac{1}{{}_nC_j} P_j} > \frac{1-p}{p} = \frac{\frac{1}{{}_nC_{j+2}} P_{j+2}^I}{\frac{1}{{}_nC_{j+1}} P_{j+1}^I}$$

and so $P_{j+2} > P_{j+2}^I$, and so on, proving the claim.

Claim 4. For $k > 1$, if $F_{n-k:n}(v) \geq F_{n-k:n}^I(v)$ then $P_k > P_k^I$.

By construction, $F_{n-k:n}(v) = \sum_{i=0}^k P_i$ and $F_{n-k:n}^I(v) = \sum_{i=0}^k P_i^I$. We know that $P_0 + P_1 = P_0^I + P_1^I$, and $P_2 < P_2^I$; so if $\sum_{i=0}^k P_k \geq \sum_{i=0}^k P_k^I$, there must be some j ($2 < j \leq k$) such that $P_j > P_j^I$. But then by the previous claim, $P_k > P_k^I$.

Claim 5. For $1 < k < n$, $F_{n-k:n}(v) < F_{n-k:n}^I(v)$.

If $F_{n-k:n}(v) \geq F_{n-k:n}^I(v)$, then by the last claim, $P_k > P_k^I$. But then by the previous claim, $P_{k'} > P_{k'}^I$ for all $k' > k$. So

$$1 = F_{n-k:n}(v) + \sum_{k'>k} P_k > F_{n-k:n}^I(v) + \sum_{k'>k} P_k^I = 1$$

a contradiction. So it must be that $F_{n-k:n}(v) < F_{n-k:n}^I(v)$. But

$$F_{n-k:n}^I(v) = \psi_{n-k:n}(p) = \psi_{n-k:n}\left(\psi_{n-1:n}^{-1}(F_{n-1:n}(v))\right)$$

so the last claim is that $F_{n-k:n}(v) < \psi_{n-k:n}\left(\psi_{n-1:n}^{-1}(F_{n-1:n}(v))\right)$, proving the proposition.

B.2 Generalizing Prop. 2 and 3 to Haile-and-Tamer Bidding

Propositions 2 and 3 have direct analogs which rely on the bidding assumptions of Haile and Tamer rather than the requirement that $G_{n:n} = F_{n-1:n}$:

Proposition B1. *Assume bidding behavior satisfies the Haile-and-Tamer assumptions and valuations are independent of N.*

(a) Under IPV, for any (x, n, n', v),

$$\psi^{-1}_{n-1:n}\left(G_{n-1:n}(v|x)\right) \quad \geq \quad \psi^{-1}_{n'-1:n'}\left(G^{\Delta}_{n':n'}(v|x)\right) \tag{13}$$

(b) Under Assumption 4.2, for any (x, n, n', v),

$$n > n' \quad \longrightarrow \quad \psi^{-1}_{n-1:n}\left(G_{n-1:n}(v|x)\right) \quad \geq \quad \psi^{-1}_{n'-1:n'}\left(G^{\Delta}_{n':n'}(v|x)\right) \tag{14}$$

Proposition B2. *Assume bidding behavior satisfies the Haile-and-Tamer assumptions and Assumption 4.3 holds. Under IPV, for any (x, n, n', v),*

$$n > n' \quad \longrightarrow \quad \psi^{-1}_{n-1:n}\left(G^{\Delta}_{n:n}(v|x)\right) \quad \leq \quad \psi^{-1}_{n'-1:n'}\left(G_{n'-1:n'}(v|x)\right) \tag{16}$$

These are proved side-by-side with Propositions 2 and 3 below.

B.3 Proof of Propositions 2 and B1

Part a. Under IPV and the exclusion restriction,

$$\begin{aligned}\psi^{-1}_{n-1:n}\left(F_{n-1:n}(v|x)\right) &= F_V(v|x,n) \\ &= F_V(v|x,n') \quad = \quad \psi^{-1}_{n'-1:n'}\left(F_{n'-1:n'}(v|x)\right)\end{aligned}$$

If $B_{n:n} = V_{n-1:n}$, then $G_{n:n}(v|x) = F_{n-1:n}(v|x)$ and $G_{n':n'}(v|x) = F_{n'-1:n'}(v|x)$, giving (11). Under Haile-and-Tamer bidding, $G_{n-1:n}(v|x) \geq F_{n-1:n}(v|x)$ and $F_{n'-1:n'}(v|x) \geq G^{\Delta}_{n':n'}(v|x)$, giving (13).

Part b. Fix n, x, and v. As above, let $p = \psi^{-1}_{n-1:n}(F_{n-1:n}(v|x))$, and let P_i be the probability that exactly i (of n) valuations are at least v. Under Assumption 4.2, as noted above in the proof of Proposition 10, $P_2 \leq {}_nC_2 p^{n-2}(1-p)^2$. If valuations are independent of N, plugging $r = n-2$ into equation 9 of Athey and Haile (2002) and rearranging gives

$$\begin{aligned}
F_{n-2:n-1}(v|x) &= F_{n-1:n}(v|x) + \tfrac{2}{n}\left[F_{n-2:n}(v|x) - F_{n-1:n}(v|x)\right] \\
&= np^{n-1} - (n-1)p^n + \tfrac{2}{n}P_2 \\
&\leq np^{n-1} - (n-1)p^n + \tfrac{2}{n}\tfrac{n(n-1)}{2}p^{n-2}(1-p)^2 \\
&= (n-1)p^{n-2} - (n-2)p^{n-1} \\
&= \psi_{n-2:n-1}(p) \\
&= \psi_{n-2:n-1}\left(\psi^{-1}_{n-1:n}\left(F_{n-1:n}(v|x)\right)\right)
\end{aligned}$$

or $\psi^{-1}_{n-2:n-1}(F_{n-2:n-1}(v|x)) \leq \psi^{-1}_{n-1:n}(F_{n-1:n}(v|x))$; if Assumption 4.2 holds strictly at (v, x, n), then (from the proof of Proposition 1 above) $P_2 < {}_nC_2 p^{n-2}(1-p)^2$ and this holds strictly. From there, $G_{n:n} = F_{n-1:n}$ and $G_{n':n'} = F_{n'-1:n'}$ establish (12), and $G_{n-1:n} \geq F_{n-1:n}$ and $F_{n'-1:n'} \geq G^{\Delta}_{n':n'}$ establish (14) under Haile-and-Tamer bidding.

B.4 Proof of Propositions 3 and B2

Let $n > n'$; under Assumption 4.3, $F_V(\,\cdot\,|x,n) \succsim_{FOSD} F_V(\,\cdot\,|x,n')$, so

$$\begin{aligned}
\psi^{-1}_{n-1:n}\left(F_{n-1:n}(v|x)\right) &= F_V(v|x,n) \\
&\leq F_V(v|x,n') \quad = \quad \psi^{-1}_{n'-1:n'}\left(F_{n'-1:n'}(v|x)\right)
\end{aligned}$$

Again, if $B_{n:n} = V_{n-1:n}$, then $G_{n:n} = F_{n-1:n}$ and $G_{n':n'} = F_{n'-1:n'}$, and (15) follows, with strict inequality whenever $F_V(v|x,n) < F_V(v|x,n')$; under Haile-and-Tamer bidding, $G^{\Delta}_{n:n} \leq F_{n-1:n}$ and $G_{n'-1:n'} \geq F_{n'-1:n'}$, and (16) follows.

B.5 Violations of (12) Under Two Entry Models

Here, we show how dependence of valuations on N generated by two standard models of endogenous participation in auctions would lead to rejection of the exclusion restriction due to a violation of (12).

Consider a model of independent private values with unobserved heterogeneity. There is a one-dimensional variable $\theta \in \Re$ which is observed by bidders but not the analyst. Valuations are *i.i.d.* $\sim F_V(\cdot \mid \theta)$, and $\theta > \theta'$ implies $F(\,\cdot\,|\theta) \succsim_{FOSD} F(\,\cdot\,|\theta')$. For each θ, assume $F_V(\cdot|\theta)$ is twice differentiable and has bounded support $[\underline{v}, \overline{v}]$. Let $f_V(\cdot \mid \theta)$ denote the density function.

Now, we apply two standard models of endogeous entry to this environment. In the first model, that of Levin and Smith (1994), there are $\overline{n}$ potential bidders, who each observe θ

but not their own valuations before deciding whether to enter (in which case they incur a cost c and participate in the auction) or not (earning a payoff of 0). Bidders play a different symmetric mixed strategy for each realization of θ, leading to a stochastic N with a different distribution for each θ.

In the second model, that of Samuelson (1985), bidders observe both θ and their own valuation before making their entry decision, and play a different pure-strategy symmetric equilibrium in cutoff strategies for each θ.

Proposition B3. *In the Levin-Smith entry game, if $f_V(\overline{v}|\theta)$ and the equilibrium entry probability are both strictly increasing in θ, then the valuations generated would violate Equation* (12) *over some range of v.*

In the Sameulson entry game, if valuations and θ are related via the Strict Monotone Likelihood Ratio Property, then the valuations generated would violate (12) *over some range of v.*

Proof. The Taylor expansion of $\psi^{-1}_{n-1:n}(F_{n-1:n}(v))$ around $v = \overline{v}$, after a lot of algebra, gives

$$\psi^{-1}_{n-1:n}(F_{n-1:n}(v)) = 1 - (\overline{v} - v)\sqrt{E_{\theta|n}\left(f_V(\overline{v}|\theta)\right)^2} + O\left((\overline{v} - v)^2\right) \tag{B1}$$

Let $n > n'$. If $\theta|N = n \succ_{FOSD} \theta|N = n'$ and $f_V(\overline{v}|\theta)$ is increasing in θ, then $E_{\theta|n}\left(f_V(\overline{v}|\theta)\right)^2 > E_{\theta|n'}\left(f_V(\overline{v}|\theta)\right)^2$, in which case $\psi^{-1}_{n-1:n}(F_{n-1:n}(v)) < \psi^{-1}_{n'-1:n'}(F_{n'-1:n'}(v))$ for v sufficiently close to $\overline{v}$, violating (12).

For the Levin-Smith result, this is all we need. We showed in Aradillas-López, Gandhi, and Quint (2013) that if the entry probability is increasing in θ, then $n > n'$ implies $\theta|N = n \succsim_{FOSD} \theta|N = n'$; the same argument shows this is strict when the entry probability is strictly increasing.

In the Samuelson game, the entry cutoff $v^*(\theta)$ is the solution to $vF_V^{\bar{n}-1}(v|\theta) = c$. Under the strict MLRP, $F_V(v|\theta)$ is strictly decreasing in θ on $(\underline{v}, \overline{v})$, so $v^*(\theta)$ and $1 - F_V(v^*(\theta)|\theta)$ are both strictly increasing in θ; so $\theta|N = n \succ_{FOSD} \theta|N = n'$ when $n > n'$, as above. But now we require the density of valuations at $\overline{v}$ *conditional on entry* to be strictly increasing in θ. This density can be written as

$$\frac{f_V(\overline{v}|\theta)}{1 - F_V(v^*(\theta)|\theta)} \quad = \quad \frac{f_V(\overline{v}|\theta)}{\int_{v^*(\theta)}^{\overline{v}} f_V(v|\theta)dv} \quad = \quad \left(\int_{v^*(\theta)}^{\overline{v}} \frac{f_V(v|\theta)}{f_V(\overline{v}|\theta)} dv\right)^{-1}$$

Since $v^*(\theta)$ is increasing in θ, an increase in θ shrinks the interval $[v^*(\theta), \overline{v}]$ over which the integral is taken; and if v and θ are related by the strict MLRP, since $v < \overline{v}$, $\frac{f_V(v|\theta)}{f_V(\overline{v}|\theta)}$ is strictly decreasing in θ. So $\int_{v^*(\theta)}^{\overline{v}} \frac{f_V(v|\theta)}{f_V(\overline{v}|\theta)} dv$ is strictly decreasing in θ, meaning $\frac{f_V(\overline{v}|\theta)}{1-F_V(v^*(\theta)|\theta)}$ is strictly increasing in θ; so (B1) gives $\psi^{-1}_{n-1:n}(F_{n-1:n}(v)) < \psi^{-1}_{n'-1:n'}(F_{n'-1:n'}(v))$ for v close to $\overline{v}$.

References

Andrews, D. and X. Shi (2011). Nonparametric inference based on conditional moment inequalities. Working Paper. Yale University.

Andrews, D. and X. Shi (2013). Inference based on conditional moment inequality models. *Econometrica 81*(2), 609–666.

Aradillas-López, A. and A. Gandhi (2011). Robust inference in ordered response games with incomplete information: Are firms strategic substitutes? Working Paper. UW-Madison.

Aradillas-López, A., A. Gandhi, and D. Quint (2013). Identification and inference in ascending auctions with correlated private values. *Econometrica 81*(2), 489–534.

Armstrong, T. (2011a). Asymptotically exact inference in conditional moment inequality models. Working paper. Stanford University.

Armstrong, T. (2011b). Weighted KS statistics for inference on conditional moment inequalities. Working paper. Stanford University.

Athey, S. and P. Haile (2002). Identification of standard auction models. *Econometrica 70*(6), 2107–2140.

Athey, S. and P. Haile (2007). Nonparametric approaches to auctions. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, vol. 6A*, Chapter 60, pp. 3847–3965. Elsevier.

Athey, S. and J. Levin (2001). Information and competition in U.S. forest service timber auctions. *Journal of Political Economy 109*(2), 375–417.

Athey, S., J. Levin, and E. Seira (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *Quarterly Journal of Economics 126*(1).

Baldwin, L. H., R. C. Marshall, and J.-F. Richard (1997). Bidder collusion at forest service timber sales. *Journal of Political Economy 105*(4), 657–699.

Barrett, G. and S. Donald (2003). Consistent tests for stochastic dominance. *Econometrica 71*, 71–104.

Campo, S., E. Guerre, I. Perrigne, and Q. Vuong (2002). Semiparametric estimation of first-price auctions with risk averse bidders. *Review of Economic Studies 78*, 112–147.

Chernozhukov, V., W. Kim, S. Lee, and A. Rosen (2013). Implementing intersection bounds in stata. Cemmap Working Papers, CWP38/13. Institute for Fiscal Studies and Department of Economics, UCL.

Chernozhukov, V., S. Lee, and A. Rosen (2013). Intersection bounds: Estimation and inference. *Econometrica 81*(2), 667–738.

Chetverikov, D. (2012). Adaptive test of conditional moment inequalities. Working paper. MIT.

Chiappori, P.-A., B. Jullien, B. Salanié, and F. Salanie (2006). Asymmetric information in insurance: General testable implications. *The RAND Journal of Economics 37*(4), 783–798.

Chiappori, P.-A. and B. Salanie (2000). Testing for asymmetric information in insurance markets. *Journal of political Economy 108*(1), 56–78.

Conley, T. G. and F. Decarolis (2013). Detecting bidders groups in collusive auctions. Working Paper, Boston University.

de Paula, A. and X. Tang (2012). Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica 80*(1), 143–172.

Froeb, L. and P. McAfee (1988). Deterring bid rigging in forest service timber auctions. US Department of Justice, mimeo.

Ghosal, S., A. Sen, and A. V. D. Vaart (2000). Testing monotonicity of regression. *Annals of Statistics 28*, 1054–1082.

Gillen, B. (2009). Identification and estimation of level-k auctions. Working Paper.

Guerre, E., I. Perrigne, and Q. Vuong (2009). Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions. *Econometrica 77*(4), 1193–1227.

Haile, P. (2001). Auctions with resale markets: An application to U.S. forest service timber sales. *American Economic Review 91*(3), 399–427.

Haile, P., K. Hendricks, R. Porter, and T. Onuma (2012). Testing competition in US offshore oil and gas lease bidding. Working Paper.

Haile, P., H. Hong, and M. Shum (2003). Nonparametric tests for common values at first-price sealed-bid auctions. Working Paper.

Haile, P. and E. Tamer (2003). Inference with an incomplete model of English auctions. *Journal of Political Economy 111*(1), 1–51.

Hall, P. and A. Yatchew (2005). Unified approach to testing functional hypotheses in semiparametric contexts. *Journal of Econometrics 127*(2), 225–252.

Hu, Y., D. McAdams, and M. Shum (2013). Identification of first-price auction models with non-separable unobserved heterogeneity. *Journal of Econometrics 174*, 186–193.

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

Kim, K. (2009). Set estimation and inference with models characterized by conditional moment inequalities. Working Paper. University of Minnesota.

Krasnokutskaya, E. (2011). Identification and estimation in procurement auctions under unobserved auction heterogeneity. *Review of Economic Studies 28*.

Lee, S., O. Linton, and Y. Whang (2009). Testing for stochastic monotonicity. *Econometrica 77*(2), 585–602.

Lee, S., K. Song, and Y. Whang (2013). Testing functional inequalities. *Journal of Econometrics 172*(1), 14 – 32.

Levin, D. and J. Smith (1994). Equilibrium in auctions with entry. *The American Economic Review 84*(3), 585–599.

Li, T., I. Perrigne, and Q. Vuong (2002). Structural estimation of the affiliated private value auction model. *RAND Journal of Economics 33*(2), 171–193.

Li, T. and B. Zhang (2010). Testing for affiliation in first-price auctions using entry behavior*. *International Economic Review 51*(3), 837–850.

Lu, J. and I. Perrigne (2008). Estimating risk aversion from ascending and sealed-bid auctions: The case of timber auction data. *Journal of Applied Econometrics 23*, 871–896.

Menzel, K. (2011). Consistent estimation with many moment inequalities. Working Paper. New York University.

Nolan, D. and D. Pollard (1987). U-processes: Rates of convergence. *Annals of Statistics 15*, 780–799.

Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. *Econometrica 57*(5), 1027–1057.

Pinkse, J. and G. Tan (2005). The affiliation effect in first-price auctions. *Econometrica 73*(1), 263–277.

Pollard, D. (1990). *Empirical Processes: Theory and Applications*. Institute of Mathematical Statistics.

Ponomareva, E. (2010). Inference in models defined by conditional moment inequalities with continuous covariates. Working Paper. Northwestern University.

Romano, J. (2004). On non-parametric testing, the uniform behaviour of the t-test, and related problems. *Scandinavian Journal of Statistics 31*, 567 – 584.

Samuelson, W. (1985). Competitive bidding with entry costs. *Economics Letters 17*(1-2), 53–57.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley. New York, NY.

Sherman, R. (1994). Maximal inequalities for degenerate u-processes with applications to optimization estimators. *Annals of Statistics 22*, 439–459.

Song, U. (2004). Nonparametric estimation of an eBay auction model with an unknown number of bidders. Working Paper.