Identification and Testing in Ascending Auctions with Unobserved Heterogeneity

Andrés Aradillas-López, Amit Gandhi, Daniel Quint*

July 22, 2010

Abstract

This paper empirically studies the consequences of unobserved heterogeneity on auction design. Unobserved heterogeneity in the objects for sale induces correlation among bidders' valuations, which violates the standard modeling assumption of independence. We show that standard data from ascending auctions partially identifies a model of correlated private values, in a way that is useful for mechanism design purposes. In particular, if larger auctions are ones where bidders on average have higher valuations (a condition which follows naturally from models of endogenous entry), we get an upper bound on the seller's expected profit at a given reserve price. If bidders' valuations are independent of auction size, a stronger assumption, then we also get a lower bound. We then show that the stronger identifying assumption of independence implies nonparametrically testable restrictions, and we develop a precise asymptotic test of these restrictions. We apply both our identification and testing results to data from the United States Forest Service. We fail to reject independence between valuations and auction size in the data. Furthermore, we find that unobserved heterogeneity has significant implications for USFS reserve price policy, suggesting substantially lower reserve prices than standard analysis and rationalizing changes made to the selling mechanism in the 1990s.

*The authors thank Steven Durlauf, Jeremy Fox, Phil Haile, Bruce Hansen, Ali Hortaçsu, Jack Porter, Art Shneyerov, Chris Taber, seminar participants at Chicago, Chicago Booth, Concordia, Duke, Michigan, Wisconsin and Yale, and conference attendees at IIOC 2009, SITE 2009, IIOC 2010, and Cowles Summer Conference 2010 for helpful comments. Aradillas-López acknowledges the support of the NSF through grant SES-0922062.

1 Introduction

Applying the insights from auction theory to real world settings requires knowledge of the primitives that define the game being played by bidders. If bidders at an auction have private values, the key primitive of interest is the latent distribution of these private values in the underlying population. The distribution of bidder valuations represents the demand curve facing the seller; once it is known, a seller can determine the optimal auction design parameters (such as the optimal reserve price), forecast expected revenue, and so on. Thus, recovering this distribution is a crucial step in the structural analysis of auction data.

Beginning with Paarsch (1997) and Laffont, Ossard, and Vuong (1995), a large literature has developed on the identification and estimation of the distribution of private values using auction data. (For a comprehensive synthesis, see Paarsch and Hong (2006).) In both first-price and ascending auctions, the key challenge is that the mapping from valuations to bids is non-trivial. In first-price auctions, this is because equilibrium bids are a complex, non-linear function of valuations. Guerre, Perrigne, and Vuong (2000), however, show that this function can nonetheless be inverted to nonparametrically identify the distribution of valuations. In ascending auctions, bids are more transparently related to valuations, as bidders have a dominant strategy: to keep bidding until the price reaches their willingness-to-pay, then stop. But in an ascending auction, a losing bidder may not bid all the way up to his valuation if multiple other bidders are still active; and since the auction ends when all but one bidder is unwilling to go higher, we never observe how much the winner would have been willing to pay. Nonetheless, Athey and Haile (2002) and Haile and Tamer (2003) show that under certain assumptions, the mapping from valuations to bids can still be inverted to nonparametrically identify or bound the underlying distribution of valuations.

A critical assumption in the above literature is that bidders' valuations are statistically *independent*. Implicitly, independence requires that differences in bidders' willingness-to-pay reflect purely idiosyncratic differences in costs or preferences. The problem with this assumption is that real-world auctions are often heterogeneous – rarely is the exact same good auctioned over and over. In order for independence of valuations to be consistent with the presence of heterogenous goods for sale, one must assume that this heterogeneity can be "conditioned out"; that is, that conditional on a vector of observed covariates $\mathbf{X}$ characterizing the object for sale, bidder values are *i.i.d.* draws from a conditional population distribution $F_V(\cdot|\mathbf{X})$, which can then be identified and estimated by the above approaches.

Independence is a strong assumption, however, both empirically and theoretically. Empirically, it rules out any *unobserved* heterogeneity shifting the demand curve across auctions. That is, if there exist any unobserved factors θ that affect demand in addition to the observables $\mathbf{X}$, then these unobservables would cause bidders' valuations to be correlated when we condition only on $\mathbf{X}$. And theoretically, correlated valuations have very different implications for mechanism design than independent valuations. For example, a strategic reserve price – a reserve price higher than the seller's own valuation for the unsold good – is always beneficial to a seller when valuations are independent, but need not be when valuations are correlated.

The extent of correlation among valuations, and its exact implications for mechanism design, are fundamentally empirical questions. In this paper, we address this problem by showing identification of a model of correlated private values. Our model is a generalization of the traditional IPV (independent private values) framework; in particular, we allow a vector of unobservable characteristics $\theta \in \Theta$ to affect the distribution of demand. We assume bidders have symmetric, *conditionally independent* private values – private values which are *i.i.d.* draws from a distribution which depends on both observable *and unobservable* covariates $(\mathbf{X}, \theta)$. Since the unobservables cannot be conditioned out, from the seller's point of view, they induce correlation among bidders' valuations.

We focus attention on one of the dominant auction formats used in practice, the English auction, and show that enough features of demand are identified to inform the reserve price policy question facing the seller. We apply our identification strategy to the sales of logging rights conducted by the United States Forest Service (USFS), and find that correlation has significant implications for the seller's problem. In particular, we find that the reserve policy that would appear optimal under IPV is too aggressive – when correlation is accounted for, these reserve prices yield much lower profits than analysis under IPV would suggest, and may be worse than not using a strategic reserve at all. We also find that the policy change that actually took place in the USFS during the 1990s – the switch from the "residual method" of timber appraisal to the "transaction evidence appraisal" method – was sufficiently conservative to yield profits even in light of the correlation we identify.

While our paper focuses on English auctions, two recent papers by Krasnokutskaya (2009) and Hu, McAdams, and Shum (2009) (building on work by Li, Perrigne, and Vuong (2000)) have addressed the issue of unobserved heterogeneity in first-price auctions.[1] By restricting unobserved heterogeneity to a scalar unobservable affecting private values in a parametric or monotonic way, and by imposing regularity conditions on the distribution of the unobservable scalar, these authors have shown that it is possible to use the statistics of measurement error to recover the distribution of the unobservable, and hence the joint distribution of private values among bidders in a first-price auction. The key to this identification strategy is that in a first-price auction, there exist multiple

[1] Guerre, Perrigne, and Vuong (2009) and Haile, Hong, and Shum (2003) also make some attempt to control for unobserved heterogeneity in first-price auctions, under stronger assumptions.

informative bids that are linked to the underlying valuations of bidders through the equilibrium of the game. In ascending auctions, on the other hand, as we discuss further below, only one bidder's valuation in each auction can be tightly pinned down by equilibrium bidding. This makes it difficult to detect and measure correlation among values using the joint distribution of bids;[2] the issue of unobserved heterogeneity in ascending auctions thus remains very much an open question.

Unlike the approach taken for first-price auctions, we place no restrictions on the dimensionality or distribution of the unobserved auction characteristics. Not surprisingly, it will be impossible to identify the full joint distribution of private values without restricting the nature of the unobservable. However, we show that we can use variation in transaction prices across auctions with different numbers of bidders to partially identify our model of correlated demand; and the features of demand that we do identify are sufficient to inform the mechanism design problem. Whereas the information revealed by auctions of different sizes is over-identifying under the assumption of independence, it becomes just identifying in an environment with correlation.

In order to use auctions of different sizes as part of a single identification strategy, we require a theory of how the valuations of bidders might change across auctions of different sizes. We can proceed under either of two identifying assumptions. The weaker assumption is that bidder valuations are stochastically higher in auctions with more bidders – roughly, the number of bidders is positively correlated with each bidder's valuation, or said another way, more people participate in auctions for more valuable goods. Under this assumption, variation in transaction prices across auctions of different sizes leads to an upper bound on the seller's expected profit under different reserve prices in auctions of a fixed size. The stronger assumption we can make is that the number of bidders in each auction is independent of their valuations. Under independence, we get both upper and lower bounds on expected profit, leading to tighter bounds on the profit-maximizing reserve price.

The assumption that valuations are stochastically increasing in auction size is a natural one, and (as we discuss later) is consistent under fairly general conditions with standard models of endogenous participation. Independence between valuations and auction size is a stronger assumption, although as we discuss below, it has been exploited in several recent papers on first-price auctions; one of the contributions of our paper is to show that this assumption is nonparametrically testable against the alternative that valuations are stochastically increasing with auction size. In particular, we show that independence implies a testable restriction on bid data, which we show would typically

[2] In contrast with sealed-bid auctions, correlation among bids across ascending auctions need not indicate correlation among values; even in an IPV setting, the winning bidder in each auction only has to bid high if another bidder bids high as well.

be violated if valuations were instead stochastically increasing. We introduce a test statistic that is applicable to this inequality restriction, and show that it has a known asymptotic distribution that forms the basis of an asymptotically valid test. The test generalizes tests of stochastic dominance between distributions, which were recently explored by Lee, Linton, and Whang (2009), in a way that allows the distributions to be transformed nonlinearly as required by auction theory, and can be modified to apply to the equality and inequality tests suggested by Athey and Haile (2002) and Haile and Tamer (2003). When we apply this test to the USFS timber auction data, we fail to reject the hypothesis of independence between valuations and auction size, which is consistent with the fact that auction size is also strikingly independent of observable auction covariates in this data.

The rest of the paper proceeds as follows. In section 2, we introduce our model of conditionally independent private values, and show how identification is obtained using our two possible identifying assumptions. In section 3, we show that independence between valuations and auction size implies a nonparametrically testable restriction in the CIPV framework. The restriction generalizes the equality test proposed by Athey and Haile (2002) for IPV ascending auctions, to an inequality test applicable to auctions with unobserved heterogeneity. We show that this test has power against the most likely violation, positive correlation between the number of bidders and auction primitives favoring higher valuations. We then introduce nonparametric test statistics for the inequality restrictions implied by the model, and derive their asymptotic properties. In section 4, we apply this test to data from U.S. Forest Service timber auctions, and show that the data fails to reject independence between valuations and auction size. We then apply our identification strategy to the same data, and demonstrate the counterfactual expected revenue and optimal reserve implications, contrasting them with the implications of "standard" analysis under the assumption of independent values; our analysis supports the relatively cautious approach to reserve price policy that the Forest Service historically has taken. Section 5 concludes. Appendix A demonstrates the implications of two standard models of auctions with endogenous entry for our identification strategy and test. (Under fairly general conditions, both models lead to valuations which are stochastically increasing in auction size but not independent of auction size, and lead to violations of our test.) Proofs and examples omitted from the text are in Appendix B; step-by-step proofs of the asymptotic properties of the test statistics are in a separate Technical Supplement.

2 Identification

2.1 Model of Correlated Demand

We use standard notation for private value auctions. N denotes the number of bidders in an auction, with n and n' denoting generic values it can take.[3] Each bidder i has a valuation V_i which is his or her private information, and gets payoff $V_i - P$ from winning the auction (where P is the price paid) and 0 from losing. If there are n bidders at the auction, let the joint distribution of their valuations $\mathbf{V} = (V_1, \ldots, V_n)$ be given by $F^n_{\mathbf{V}}(v_1, \ldots, v_n \mid \mathbf{X})$, where $\mathbf{X}$ is a vector of observables that characterize demand (such as characteristics of the good being auctioned or local market conditions).

Under the standard assumption of independent private values, bidders' valuations are modeled as *i.i.d.* random variables, and hence

$$F^n_{\mathbf{V}}(v_1, \ldots, v_n \mid \mathbf{X}) = \prod_{i=1}^{n} F(v_i \mid \mathbf{X})$$

We generalize the independent private value framework and allow $F^n_{\mathbf{V}}$ to be correlated even after controlling for $\mathbf{X}$. A natural source of such correlation is unobserved heterogeneity (from the perspective of the seller); thus, in addition to the observable covariates $\mathbf{X}$, we assume there is a vector of unobserved characteristics $\theta \in \Theta$ that affect demand at the auction. Bidder valuations for the object for sale in an auction with the complete vector of characteristics $(\mathbf{X}, \theta)$ are *i.i.d.* draws from a probability distribution $F(\,\cdot\,|\mathbf{X}, \theta)$. Since $\theta \in \Theta$ is a demand shifter that is unobserved by the seller, the seller effectively faces the correlated demand environment given by

$$F^n_{\mathbf{V}}(v_1, \ldots, v_n \mid \mathbf{X}) = \int_{\theta \in \Theta} \prod_{i=1}^{n} F(v_i \mid \mathbf{X}, \theta)\, dG(\theta \mid \mathbf{X}, n) \tag{1}$$

where $G(\,\cdot\,|\mathbf{X}, n)$ is the distribution of θ (conditional on $\mathbf{X}$ and n). Thus, heterogeneity in the unobservable θ is the mechanism that induces correlation in bidders' valuations, while any variation not caused by θ is idiosyncratic and independent across bidders. Since bidders' valuations are assumed to be symmetrically distributed and independent *conditional* on $(\mathbf{X}, \theta)$, we will refer to this model as CIPV (conditionally independent private values), in contrast with the standard model of IPV.

For the remainder of this section and the next, we will suppress the dependence on $\mathbf{X}$ for ease of exposition, but our analysis is still implicitly being done conditional on $\mathbf{X}$. For an auction with n bidders, knowledge of the joint distribution $F^n_{\mathbf{V}}(v_1, \ldots, v_n)$ allows the seller to anticipate the

[3] In principle, this should refer to the number of *potential* bidders, that is, the number of bidders who learned their valuations and considered participating. This is often assumed to be observed in the data; in our application, as we discuss later, there are specific institutional features that make this a reasonable assumption.

revenue effects of changing auction design parameters. In the context of an ascending auction, the main auction design parameter of interest is the reserve price, a central question being whether there is any value in setting a strategic reserve price (i.e., a reserve price above the seller's own value), and if so what is the optimal amount of "markup" over opportunity cost that should be used to set the reserve.

A key result in auction theory is the so-called "exclusion principle" (see Krishna (2002)): under independent private values, it is always in the seller's interest to potentially exclude bidders by setting a reserve price strictly above his own valuation. The optimal markup for the seller is given by

$$\arg\max_{r>0}(r - v_0)(1 - F_V(r))$$

where F_V is the marginal distribution of a single bidder's valuations and v_0 the seller's valuation for the unsold good. If bidder values are correlated, however, neither result holds. In particular, the gains from using a strategic reserve price, and the degree of the optimal markup, are likely much smaller in the correlated demand environment as compared to IPV.[4] This is because the optimal reserve price trades off the decreased likelihood of a sale (due to the possibility that all bidders have valuations below the reserve) against the increased revenue in the event the reserve binds (just one bidder has a valuation above the reserve). Correlation among bidder valuations makes the former more likely – when bidder values are correlated, they are more likely to all be low simultaneously – while decreasing the value of the latter – when bidder values are correlated, the top two are likely to be closer together, so the increase in revenue is smaller than under independence. (This intuition is explained further in the next section.) The actual degree of correlation among values will dictate the extent to which IPV over-estimates the gains from a reserve price, which is the main empirical problem we now confront.

2.2 The Identification Problem

Let $V_{k:n}$ denote the k^{th} lowest bidder valuation in an auction with n bidders, and $F_{k:n}$ its probability distribution. Haile and Tamer (2003) show that under the assumption that bidders have private values and play undominated strategies, the top two bids in an ascending auction lead to sharp bounds on $V_{n-1:n}$. This is because $V_{n-1:n}$ must be no less than the second-highest bid, or the second-highest bidder would have risked overpaying for the object; and no more than the winning bid plus one bid increment – the lowest bid still available when the auction ended – or the second-

[4] Quint (2008) shows this for valuations which are symmetric and *affiliated*: the expected gain from any reserve price $r > v_0$, and the profit-maximizing reserve, are bounded above by what they would be in the IPV model consistent with the same observables.

highest bidder would have preferred to raise his bid to that level rather than let the auction end. Thus, observations of the top two bids lead to pointwise bounds on the distribution $F_{n-1:n}$, even in an "incomplete" model where exact equilibrium strategies (which bidder bids at each stage) are not specified.[5] However, bid data and the restrictions implied by equilibrium bidding behavior provide very little direct information about the distribution of the highest valuation $V_{n:n}$, or about the *joint* distribution among values.[6]. We will therefore assume that what is directly identifiable in the data are only the distributions $F_{n-1:n}$ for various n.

As a simplification, rather than using the bounds on $F_{n-1:n}$ established in Haile and Tamer (2003), we will proceed under the stronger assumption that $V_{n-1:n}$ is exactly equal to transaction price, so that $F_{n-1:n}$ is point-identified from the data. This would be true to within the minimum bid increment under the assumptions of Haile and Tamer (2003) if there are no "jump bids" at the end of an auction; in our data, the winning bid is on average just 2% higher than the second-highest, so this is a reasonable approximation. Our results extend without difficulty when only pointwise upper and lower bounds on $F_{n-1:n}$ are known rather than its exact value.

In an IPV setting, the marginal distribution $F_{n-1:n}$ for any one n is sufficient to identify the parent distribution F_V. In a correlated demand environment, however, the marginal distribution $F_{n-1:n}$ does not uniquely pin down the joint distribution $F^n_{\mathbf{V}}$, and is insufficient for mechanism design purposes. What we now show, however, is that the seller's profit problem in an auction of size n does not depend on the entire joint distribution $F^n_{\mathbf{V}}$, but only on the marginal distributions of the top *two* valuations, $F_{n-1:n}$ and $F_{n:n}$.[7] The fact that we do not need to identify the joint distribution of these two order statistics, only their marginals, to describe the seller's revenue curve is a key simplification that will enable us to empirically study the model presented above.

Lemma 1 *In an ascending auction with n bidders and reserve price r, the seller's expected profit and each bidder's ex-ante expected surplus depend only on $F_{n-1:n}$ and $F_{n:n}$.*

Proof. Again, we assume that transaction price equals $V_{n-1:n}$, provided it exceeds the reserve price; if not, then the winner pays r. We show in Appendix B.1 that expected profit $\pi_n(r)$, and each

[5]The data can also be used to provide upper and lower bounds on $F_{k:n}$ for $k < n-1$, but these bounds are generally much wider and less informative.

[6]Unlike in first-price auctions with Bayes-Nash bidding behavior, correlation among values in an ascending auctions cannot be inferred from correlation among bids.

[7]In a different but analogous setting, Athey and Haile (2007) point out that identification of the *joint* distribution of $(V_{n-1:n}, V_{n:n})$ is sufficient for "evaluation of rent extraction by the seller, the effects of introducing a reserve price, and the outcomes under a number of alternative selling mechanisms."

bidder's ex-ante expected surplus $u_n(r)$, can be written as

$$\pi_n(r) = (F_{n-1:n}(r) - F_{n:n}(r))(r - v_0) + \int_r^{+\infty} (v - v_0) dF_{n-1:n}(v) \quad (2)$$

$$u_n(r) = \frac{1}{n} (E\{\max\{V_{n:n}, r\}\} - E\{\max\{V_{n-1:n}, r\}\}) \quad (3)$$

where v_0 is again the seller's valuation for the unsold good.[8] □

The expression (2) helps to clarify why the gains from using a reserve price are smaller in a correlated demand environment. A reserve price is only beneficial when it binds, i.e., when it is greater than $V_{n-1:n}$ but less than $V_{n:n}$. The probability of this event is captured by the first term $(F_{n-1:n}(r) - F_{n:n}(r))$ in (2). As Lemma 2 below will confirm, the distance between the top two bidders at an auction, as measured by this difference, is smaller when values are correlated as compared to when values are independent. For the purposes of identifying the revenue implications of alternative reserve price policies, $(F_{n-1:n}(r) - F_{n:n}(r))$ is thus a sufficient measure of correlation among values; so from the seller's point of view, all payoff-relevant information in $F^n_{\mathbf{V}}$, is contained in the marginal distributions $F_{n-1:n}$ and $F_{n:n}$. Once these distributions are known, Lemma 1 gives us $\pi_n(r)$, and we can solve $\max_r \pi_n(r)$ to find the optimal reserve price for an auction with n bidders.[9]

As we have already discussed, we assume $F_{n-1:n}$ is directly revealed by the auction data; so what remains to be identified is $F_{n:n}$. Of course, $V_{n:n}$ is the key order statistic that is censored in an ascending auction - we never observe what the winning bidder would have been willing to pay since the auction ends before his reservation value is reached. What we will show, however, is that we can use the distribution of transaction prices across different auction sizes to identify bounds on $F_{n:n}$, which will lead via (2) to bounds on π_n. To better understand our identification strategy, we will first show that if we only have data from auctions of a fixed size n, the bounds we can identify on $F_{n:n}$ are too wide to be empirically useful. This then motivates our approach of using auctions of different sizes to identify correlation in an empirically useful way.

[8] Under two further assumptions – that bidders observe θ in addition to their own valuations, and that they play Bayesian Nash equilibrium strategies – (2) and (3) holds exactly for *any* "standard" auction format, that is, any mechanism satisfying the usual conditions for revenue equivalence. If bidders know θ, then conditioning on any realization of θ, they are in a symmetric IPV setting; so for a given θ, any mechanism under which the object is always allocated to the bidder with the highest valuation (provided it is greater than r), and a bidder with value less than r gets expected payoff 0, would be revenue-equivalent to the second-price sealed-bid auction, which satisfies (2) and (3) exactly. Averaging over θ, then, would complete the proof.

[9] Under independent private values, the optimal r does not depend on n, but with correlated values, it does. Levin and Smith (1996) give conditions (Prop. 5) under which it is strictly decreasing. A seller who must select a reserve price without knowing n would solve $\max_r \sum_n p(n)\pi_n(r)$, where $p(n)$ is the probability of n bidders participating; as we discuss later, in our particular application, it is reasonable to assume the seller knows n when selecting a reserve price.

To show the bounds on $F_{n:n}$ that can be obtained from auctions of a fixed size n, define a function $\psi_{n-1:n} : [0,1] \to [0,1]$ by

$$\psi_{n-1:n}(s) \quad = \quad ns^{n-1} - (n-1)s^n \tag{4}$$

The function $\psi_{n-1:n}$ has the property that if $F(\cdot)$ is any distribution function, the second-highest of n independent draws from F has distribution $\psi_{n-1:n}(F(\cdot))$. It is straightforward to show that $\psi_{n-1:n}$ is strictly increasing, and therefore invertible.

Lemma 2 *Under CIPV,* $F_{n:n}(v) \geq \left(\psi^{-1}_{n-1:n}(F_{n-1:n}(v))\right)^n$ *and* $F_{n:n}(v) \leq F_{n-1:n}(v)$. *If no information is available other than* $F_{n-1:n}$*, both these bounds are sharp.*[10]

Proof. Since $V_{n:n} \geq V_{n-1:n}$, $F_{n:n}(v) \leq F_{n-1:n}(v)$. Next, let $\psi_{n:n}(s) = s^n$. The key step of the proof is to note that $\psi_{n:n} \circ \psi^{-1}_{n-1:n}$ is convex. Differentiating,

$$\left(\psi_{n:n} \circ \psi^{-1}_{n-1:n}\right)'(s) \;=\; \frac{\psi'_{n:n}(t)}{\psi'_{n-1:n}(t)} \;=\; \frac{nt^{n-1}}{n(n-1)t^{n-2}(1-t)} \;=\; \frac{t}{(n-1)(1-t)}$$

where $t = \psi^{-1}_{n-1:n}(s)$. $\frac{t}{(n-1)(1-t)}$ is increasing in t, so $\left(\psi_{n:n} \circ \psi^{-1}_{n-1:n}\right)'$ is increasing, establishing that $\psi_{n:n} \circ \psi^{-1}_{n-1:n}$ is convex. For a given θ, the CDF of the greatest of n independent draws from $F(v|\theta)$ is $(F(v|\theta))^n$; taking the expectation over θ implies $F_{n:n}(v) = E_\theta\,(F(v|\theta))^n$, and applying Jensen's inequality gives

$$\begin{aligned} F_{n:n}(v) &= E_\theta\left\{\psi_{n:n}\left(F(v|\theta)\right)\right\} \\ &= E_\theta\left\{\left(\psi_{n:n} \circ \psi^{-1}_{n-1:n}\right)\left(\psi_{n-1:n}\left(F(v|\theta)\right)\right)\right\} \\ &\geq \left(\psi_{n:n} \circ \psi^{-1}_{n-1:n}\right)\left(E_\theta\left\{\psi_{n-1:n}\left(F(v|\theta)\right)\right\}\right) \\ &= \psi_{n:n} \circ \psi^{-1}_{n-1:n}\left(F_{n-1:n}(v)\right) \end{aligned}$$

As for sharpness, if values are *i.i.d.* $\sim \psi^{-1}_{n-1:n}(F_{n-1:n}(\cdot))$, the lower bound is achieved; and if bidder values are perfectly correlated (e.g., if $\theta \sim F_{n-1:n}$ and $V_i = \theta$), the upper bound is achieved. □

In essence, this result says we learn nothing about how correlated values are if we only observe the distribution of transaction prices in auctions of a fixed size n – the lower bound is independence, and the upper bound perfect correlation. But these have very different implications for the optimal reserve price. In order to tighten the bounds on $F_{n:n}$, we need more information.

2.3 Identification Strategy

The additional information we will use is variation in transaction prices across auctions of different sizes. Intuitively, the speed at which transaction prices increase with N contains information about

[10] Quint (2008) shows these same bounds hold for symmetric, affiliated private values, whether or not they are conditionally independent.

how correlated bidder values are. (The more highly correlated they are, the less impact an additional bidder has on the distribution of the highest ones.) However, to interpret this information, we will need to consider how individual bidders' valuations might differ across auctions of different sizes.

To see how identification will be achieved, consider the following thought experiment. Start with an $n+1$-bidder auction, and randomly remove one of the bidders. With probability $\frac{n}{n+1}$, the bidder with the highest valuation was not removed, and so the highest of the n remaining is the highest of the original $n+1$. With probability $\frac{1}{n+1}$, the bidder with the highest valuation was removed, in which case the highest of the remaining n chosen is the second-highest of the original $n+1$. If we let $F_{n:n}^{n+1}$ denote the marginal distribution of the highest of n bidders randomly chosen from an $n+1$-bidder auction, then[11]

$$F_{n:n}^{n+1}(v) = \frac{1}{n+1} F_{n:n+1}(v) + \frac{n}{n+1} F_{n+1:n+1}(v) \tag{5}$$

However, the distribution of $V_{n:n}$, which we want to identify, is not $F_{n:n}^{n+1}$ but rather $F_{n:n}$. In order to make useful inferences from (5), we need to be able to make comparisons between bidders in auctions of different sizes. Next, we introduce two possible assumptions about how to make such comparisons, either of which will lead to at least partial identification of our model.

Definition 1 *Valuations are* **stochastically increasing in** N *if for any* n, $F_{n:n}^{n+1} \succsim_{FOSD} F_{n:n}$.

That is, valuations are stochastically increasing in N if auctions with more bidders tend to have bidders with weakly higher valuations – for example, three bidders chosen at random from a four-bidder auction have a stochastically greater highest valuation than the three bidders in a three-bidder auction. This is arguably a natural assumption, as we would intuitively expect auctions for more valuable objects to attract more bidders. For example, we show in Appendix A that under fairly general conditions, the well-known endogenous entry models of Levin and Smith (1994) and Samuelson (1985) both yield valuations which are stochastically increasing in N.

When valuations are stochastically increasing, we get partial identification in the form of one-sided bounds.

Lemma 3 *If valuations are stochastically increasing in* N*, then*

$$F_{n:n}(v) \quad \geq \quad \sum_{m=n+1}^{\bar{n}} \left(\frac{1}{n-1} \prod_{i=n}^{m-1} \frac{i-1}{i+1} \right) F_{m-1:m}(v) \quad + \quad \frac{n}{\bar{n}} F_{\bar{n}:\bar{n}}(v) \tag{6}$$

for any $\bar{n} > n$ *and any* v.

[11] This relationship among order statistics can be viewed as a generalization of Eq. 9 in Athey and Haile (2002).

Proof. A full proof is given in Appendix B.2. By assumption, $F_{n:n}^{n+1} \succsim_{FOSD} F_{n:n}$, which implies $F_{n:n}(v) \geq F_{n:n}^{n+1}(v) = \frac{1}{n+1} F_{n:n+1}(v) + \frac{n}{n+1} F_{n+1:n+1}(v)$; iterating gives (6).[12] $\square$

In the limit as $\bar{n}$ goes to infinity, the coefficient on the trailing term $F_{\bar{n}:\bar{n}}$ vanishes and we have a lower bound on $F_{n:n}$, which leads in turn to an upper bound on the seller's profit π_n. Under a stronger identifying assumption, we can get point identification in the limit.

Definition 2 *Valuations are* **independent of** N *if for any* n, $F_{n:n}^{n+1} = F_{n:n}$.

Under the CIPV model defined in (1), this can be interpreted simply as the distribution of θ, $G(\cdot|N)$, being invariant across N, that is, θ being independent of N. Independence leads to the following result:

Lemma 4 *If valuations are independent of* N*, then*

$$F_{n:n}(v) \quad = \quad \sum_{m=n+1}^{\bar{n}} \left(\frac{1}{n-1} \prod_{i=n}^{m-1} \frac{i-1}{i+1} \right) F_{m-1:m}(v) \quad + \quad \frac{n}{\bar{n}} F_{\bar{n}:\bar{n}}(v) \tag{7}$$

for any $\bar{n} > n$ *and any* v.

Proof. Again, see Appendix B.2 for a full proof. If $F_{n:n} = F_{n:n}^{n+1}$, (5) becomes $F_{n:n}(v) = \frac{1}{n+1} F_{n:n+1}(v) + \frac{n}{n+1} F_{n+1:n+1}(v)$; iterating gives (7). $\square$

As $\bar{n}$ grows, the coefficient on the trailing term in (7) vanishes; so in the limit where every $\{F_{m-1:m}\}_{m=n+1}^{\infty}$ is learned, $F_{n:n}$ is point-identified.[13]

In this paper, we remain agnostic as to the actual process determining bidder participation; we simply show how different assumptions about the relationship between valuations and auction size lead to identification, and test whether the stronger of these assumptions is supported by the data. (Even if N is the endogenous outcome of an entry game such as the one described by Levin and Smith (1994), it could still be independent of valuations if potential bidders had the same information as the seller prior to entry – that is, if bidders only observed θ *after* deciding whether to participate in

[12] If valuations were stochastically *decreasing* in N, the right-hand side of (6) would be an *upper* bound on $F_{n:n}$, and the bounds $\overline{F}_{n:n}$, $\underline{\pi}_n$, and $\underline{u}_n$ defined below for Theorem 1 would hold. We do not think this case is likely or particularly interesting.

[13] Note that Lemmas 1, 3, and 4 do not depend on the additional structure imposed by CIPV, and hold for arbitrary joint distributions of private values. We maintain the assumption of CIPV throughout this paper, which yields both bounds on the trailing term $\frac{n}{\bar{n}} F_{\bar{n}:\bar{n}}$ and a testable restriction implied by valuations being independent of N. Theorem 1 below would still hold if values were symmetric and affiliated but not conditionally independent; under a more general model, a weaker lower bound on $F_{\bar{n}:\bar{n}}$ could be used to get analogous results.

an auction.) We do not explicitly test whether valuations are stochastically increasing in N; instead, we see it as the natural alternative hypothesis to independence.

There are two remaining complications that we must address before we can translate Lemmas 3 and 4 into an identification of the seller's profit. The first complication is that the maximal auction size $\bar{n}$ is not unbounded in actual data, but is finite. In this case, the coefficient on the trailing term $F_{\bar{n}:\bar{n}}$ in equations (6) and (7) does not go to zero; so the term $F_{\bar{n}:\bar{n}}$, which is not observed in the data, cannot be ignored. The second complication is that although we have proceeded under the assumption that each distribution $F_{m-1:m}$ is learned exactly, it is natural to ask what would happen if $F_{m-1:m}$ could only be pointwise bounded as in Haile and Tamer (2003). We now show that both complications can be handled in a consistent way. Intuitively, using lower bounds for the terms $\{F_{m-1:m}\}_{m=n}^{\bar{n}}$ and $F_{\bar{n}:\bar{n}}$ gives a lower bound on $F_{n:n}$ under the assumption that valuations are stochastically increasing in N. Likewise, using upper bounds for the terms $\{F_{m-1:m}\}_{m=n}^{\bar{n}}$ and $F_{\bar{n}:\bar{n}}$ then allows us to add an upper bound on $F_{n:n}$ under the stronger assumption that valuations are independent of N.

To see precisely how this works, fix n and $\bar{n} > n$, and suppose that there are upper and lower bounds $\left\{\overline{F}_{m-1:m}, \underline{F}_{m-1:m}\right\}_{m=n}^{\bar{n}}$, such that $F_{m-1:m}(v) \in \left[\underline{F}_{m-1:m}(v), \overline{F}_{m-1:m}(v)\right]$ for every v and every $m \in \{n, n+1, \ldots, \bar{n}\}$.[14] Plugging these bounds into the expressions in Lemmas 3 and 4, and using Lemma 2 to bound the trailing terms, defines bounds on $F_{n:n}$,

$$\begin{aligned} \overline{F}_{n:n}(v) &\equiv \textstyle\sum_{m=n+1}^{\bar{n}} \alpha_m^n \overline{F}_{m-1:m}(v) + \frac{n}{\bar{n}} \overline{F}_{\bar{n}-1:\bar{n}}(v) \\ \underline{F}_{n:n}(v) &\equiv \textstyle\sum_{m=n+1}^{\bar{n}} \alpha_m^n \underline{F}_{m-1:m}(v) + \frac{n}{\bar{n}} \left(\psi_{\bar{n}-1:\bar{n}}^{-1}\left(\underline{F}_{\bar{n}-1:\bar{n}}(v)\right)\right)^{\bar{n}} \end{aligned}$$

where $\alpha_m^n \equiv \frac{1}{n-1} \prod_{i=n}^{m-1} \frac{i-1}{i+1}$ for $m > n$. We will define bounds on π_n and u_n by plugging the bounds on $F_{n-1:n}$ and $F_{n:n}$ into (2) and (3), giving

$$\begin{aligned} \underline{\pi}_n(r) &\equiv \left(\overline{F}_{n-1:n}(r) - \overline{F}_{n:n}(r)\right)(r - v_0) + \textstyle\int_r^{\infty} (v - v_0) d\overline{F}_{n-1:n}(v) \\ \overline{\pi}_n(r) &\equiv \left(\underline{F}_{n-1:n}(r) - \underline{F}_{n:n}(r)\right)(r - v_0) + \textstyle\int_r^{\infty} (v - v_0) d\underline{F}_{n-1:n}(v) \end{aligned}$$

and

$$\begin{aligned} \underline{u}_n(r) &\equiv \tfrac{1}{n}\left(\textstyle\int_0^{\infty} \max\{r, v\} d\overline{F}_{n:n}(v) - \int_0^{\infty} \max\{r, v\} d\underline{F}_{n-1:n}(v)\right) \\ \overline{u}_n(r) &\equiv \tfrac{1}{n}\left(\textstyle\int_0^{\infty} \max\{r, v\} d\underline{F}_{n:n}(v) - \int_0^{\infty} \max\{r, v\} d\overline{F}_{n-1:n}(v)\right) \end{aligned}$$

Finally, we will use $\overline{\pi}_n$ and $\underline{\pi}_n$ to bound the optimal reserve price; define

$$\begin{aligned} \underline{r}_n &\equiv \min\left\{r \geq v_0 \ : \ \overline{\pi}_n(r) \geq \max_{r'} \underline{\pi}_n(r')\right\} \\ \overline{r}_n &\equiv \max\left\{r \geq v_0 \ : \ \overline{\pi}_n(r) \geq \max_{r'} \underline{\pi}_n(r')\right\} \\ \widetilde{r}_n &\equiv \max\left\{r \geq v_0 \ : \ \overline{\pi}_n(r) \geq \underline{\pi}_n(v_0)\right\} \end{aligned}$$

[14] Under the model of Haile and Tamer (2003), $\overline{F}_{n-1:n}(v) = G_{n-1:n}(v)$ and $\underline{F}_{n-1:n}(v) = G_{n:n}(v - \delta)$, where $G_{n:n}$ and $G_{n-1:n}$ are the distributions of the highest and second-highest bids and δ the minimum bid increment.

and let $r_n^* = \arg\max_r \pi_n(r)$ be the true profit-maximizing reserve price.

Theorem 1 *If $F_{m-1:m}(v) \in \left[\underline{F}_{m-1:m}(v), \overline{F}_{m-1:m}(v)\right]$ for every $m \in \{n, n+1, \ldots, \bar{n}\}$ and every v, then for any $r \geq v_0$,*

1. *If valuations are independent of N, then $F_{n:n}(r) \in \left[\underline{F}_{n:n}(r), \overline{F}_{n:n}(r)\right]$, $\pi_n(r) \in [\underline{\pi}_n(r), \overline{\pi}_n(r)]$, $r_n^* \in [\underline{r}_n, \overline{r}_n]$, and $u_n(r) \in [\underline{u}_n(r), \overline{u}_n(r)]$*

2. *If valuations are stochastically increasing in N, then $F_{n:n}(r) \geq \underline{F}_{n:n}(r)$, $\pi_n(r) \leq \overline{\pi}_n(r)$, $r_n^* \leq \widetilde{r}_n$, and $u_n(r) \leq \overline{u}_n(r)$*

Proof. $\overline{F}_{n:n}$ and $\underline{F}_{n:n}$ follow immediately from Lemmas 2, 3, and 4.[15] For bounds on π_n, Quint (2008) shows that (2) (the expression for $\pi_n(r)$) is stochastically increasing in both $V_{n-1:n}$ and $V_{n:n}$ (in the first-order stochastic dominance sense); so plugging the lower bounds $\underline{F}_{n-1:n}$ and $\underline{F}_{n:n}$ into (2) gives the upper bound $\overline{\pi}_n(r)$, and (when valuations are independent of N) the upper bounds $\overline{F}_{n-1:n}$ and $\overline{F}_{n:n}$ give the lower bound $\underline{\pi}_n(r)$. By (3), u_n is the expected value of a function which is increasing in $V_{n:n}$ and decreasing in $V_{n-1:n}$, so expected bidder surplus is stochastically increasing in $V_{n:n}$ and decreasing in $V_{n-1:n}$; plugging $\overline{F}_{n:n}$ and $\underline{F}_{n-1:n}$ into (3) gives the lower bound $\underline{u}_n(r)$, and $\underline{F}_{n:n}$ and $\overline{F}_{n-1:n}$ yield the upper bound $\overline{u}_n$.

When valuations are independent of N, both bounds $\overline{\pi}_n$ and $\underline{\pi}_n$ hold; as in Haile and Tamer (2003), we can use them to bound the profit-maximizing reserve price for an auction of a given size. Letting $r_* = \arg\max_{r'} \underline{\pi}_n(r')$,

$$\overline{\pi}_n(r_n^*) \quad \geq \quad \pi_n(r_n^*) \quad \geq \quad \pi_n(r_*) \quad \geq \quad \underline{\pi}_n(r_*) \quad = \quad \max_{r'} \underline{\pi}_n(r')$$

and so $r_n^* \in \{r : \overline{\pi}_n(r) \geq \max_{r'} \underline{\pi}_n(r')\}$. When valuations are stochastically increasing in N, the upper bound $\overline{\pi}_n$ still holds. Although the lower bound $\underline{\pi}_n$ need not hold generally, $\underline{\pi}_n(v_0)$ depends only on $\overline{F}_{n-1:n}$, not $\overline{F}_{n:n}$, and therefore $\pi_n(v_0) \geq \underline{\pi}_n(v_0)$ even when valuations are stochastically increasing; $\overline{\pi}_n(r_n^*) \geq \pi_n(r_n^*) \geq \pi_n(v_0) \geq \underline{\pi}_n(v_0)$ gives the bound $\widetilde{r}_n$. □

When $\{F_{m-1:m}\}$ are learned exactly, this is simply the special case $\underline{F}_{m-1:m} = \overline{F}_{m-1:m} = F_{m-1:m}$. In this case, when valuations are independent of N, all uncertainty about $F_{n:n}$ is due to uncertainty about the trailing term $\frac{n}{\bar{n}} F_{\bar{n}:\bar{n}}$. This means that bounds on $F_{n:n}$, and therefore bounds on π_n and r_n^*, will be tighter when n is lower – which is exactly when setting the correct reserve price is most important.[16]

[15] The bounds placed on $F_{\bar{n}:\bar{n}}$ by Lemma 2 are the tightest available using only $F_{\bar{n}-1:\bar{n}}$, but could potentially be tightened using $\{F_{n-1:n}\}_{n \neq \bar{n}}$; thus, the bounds on $F_{n:n}$, π_n, r_n^*, and u_n are not necessarily sharp.

[16] As n grows, the probabilities that a given reserve price r either binds or precludes a sale both go to 0.

To summarize, then, we have identified two-sided bounds on expected profit when valuations are independent of N, and shown that the upper bound is still valid when valuations are stochastically increasing. (As it happens, in our application, the upper bound alone is enough to reach economically interesting conclusions.) The next question, of course, is whether the stronger identifying assumption, and hence the lower bound on expected profit, is empirically valid. To answer this, we will show that the assumption that valuations are independent of N has nonparametrically testable implications, and construct a test statistic which we argue has power against the alternative that valuations are stochastically increasing.

3 Testing Independence Between Valuations and N

Independence of bidder valuations and N has been used as an identifying assumption in several papers on first-price auctions with *independent* signals: by Guerre, Perrigne, and Vuong (2009) to identify the coefficient of risk aversion; by Haile, Hong, and Shum (2003) to test between common and private values; and by Gillen (2009) to identify the distribution of behavioral types in a "level-k" model. We, on the other hand, work in a setting where private values are generally believed to hold, and consider ascending auctions, in which strategic sophistication and risk preferences play no role; thus, we are able to use this assumption to identify a complexity that these other papers must assume away, correlation among values. In particular, the results of Section 2 established that with conditionally independent private values, when valuations are independent of N, variation in auction size allows for two-sided bounds on expected revenue (as a function of reserve price) and as a result the optimal reserve. When valuations are stochastically increasing in auction size N, only one side of these bounds holds.

We would therefore like a way to determine whether our stronger identifying assumption, independence of valuations and N, is empirically valid in a particular data set. While this assumption of independence has been used several times, it has never been formally tested. In our context, independence between valuations and N imposes nonparametric restrictions on observable distributions, which we argue would likely be violated if the assumption failed. We can therefore use these restrictions to design a nonparametric test of this assumption.

3.1 Benchmark: Testing under IPV

Our test for independence between valuations and N under our model is most easily understood if we first introduce an analogous test within the *independent* private values framework. Recall that $\psi_{n-1:n}$ was defined above by $\psi_{n-1:n}(s) = ns^{n-1} - (n-1)s^n$, and that the CDF of the second-highest of n independent draws from a probability distribution $F(\cdot)$ is $\psi_{n-1:n}(F(\cdot))$. In the symmetric IPV model, valuations of participating bidders are independent draws from some probability distribution F_V, so $F_{n-1:n}(v) = \psi_{n-1:n}(F_V(v))$, or $\psi^{-1}_{n-1:n}(F_{n-1:n}(v)) = F_V(v)$.

Since we assume $F_{n-1:n}$ is identified from bid data for each n, a test of the symmetric IPV model with valuations independent of N is therefore whether

$$\psi^{-1}_{n-1:n}(F_{n-1:n}(v)) \quad = \quad \psi^{-1}_{n'-1:n'}(F_{n'-1:n'}(v)) \tag{8}$$

for each pair (n', n). (This is the test considered in Theorem 1 of Athey and Haile (2002).)

It is useful to consider more carefully the underlying economics of this test. If valuations are independent of auction size, then as n increases, the distribution of $V_{n-1:n}$ shifts toward higher values (increases in a first-order stochastic dominance sense), simply because $V_{n-1:n}$ is the second-highest of a larger sample; we refer to this effect as the *competition effect*. The significance of (8) is that the IPV model exactly determines the magnitude of the competition effect, that is, how "quickly" transaction prices rise as n increases: for $n > n'$, $F_{n-1:n}(v) < F_{n'-1:n'}(v)$, but $F_{n-1:n}(v) = \psi_{n-1:n} \circ \psi^{-1}_{n'-1:n'}(F_{n'-1:n'}(v))$.

3.2 Testing under CIPV

We now generalize the test of independence between valuations and auction size to the case where valuations are CIPV instead of IPV. We first observe that under CIPV, the test (8) is no longer valid. To see why not, consider the extreme case of perfectly correlated values, where θ is one-dimensional and each bidder's valuation in an auction with characteristics θ is simply θ. In this case, $F_{n-1:n}(v)$ for $n \geq 2$ is simply the ex-ante distribution of θ, and therefore does not vary with n. That is, the competition effect has vanished, and hence the test (8) will over-predict its size: for $n > n'$, $F_{n-1:n}(v) \leq F_{n'-1:n'}(v)$, but now $F_{n-1:n}(v) > \psi_{n-1:n} \circ \psi^{-1}_{n'-1:n'}(F_{n'-1:n'}(v))$ instead of equality; $F_{n-1:n}$ has fallen less than IPV would have predicted.

We now show that the direction of this inequality is in fact a general feature of CIPV – if valuations are independent of auction size, then conditional independence always slows down the size of the competition effect relative to IPV.

Theorem 2 *Under symmetric, conditionally independent private values, if valuations are independent of N, then $F_{n-1:n}(v)$ is decreasing in n and $\psi^{-1}_{n-1:n}(F_{n-1:n}(v))$ is increasing in n for all v.*

Proof. For the first result, note that $\psi_{n-1:n}(s)$ is decreasing in n.[17] At a given realization of θ, the distribution of $V_{n-1:n}$ is $\psi_{n-1:n}(F(\cdot|\theta))$; so the unconditional distribution $F_{n-1:n}(v) = E_\theta\{\psi_{n-1:n}(F(v|\theta))\}$ is decreasing in n as well.

For the second result, it suffices to show $\psi^{-1}_{n:n+1}(F_{n:n+1}(v)) \geq \psi^{-1}_{n-1:n}(F_{n-1:n}(v))$. First, we show $\psi_{n-1:n} \circ \psi^{-1}_{n:n+1} : [0,1] \to [0,1]$ concave: differentiating,

$$\left(\psi_{n-1:n} \circ \psi^{-1}_{n:n+1}\right)'(s) = \frac{\psi'_{n-1:n}(t)}{\psi'_{n:n+1}(t)} = \frac{n(n-1)t^{n-2}(1-t)}{(n+1)nt^{n-1}(1-t)} = \frac{n-1}{n+1} \cdot \frac{1}{t}$$

where $t = \psi^{-1}_{n:n+1}(s)$. So $\left(\psi_{n-1:n} \circ \psi^{-1}_{n:n+1}\right)'(s) = \frac{n-1}{n+1} / \psi^{-1}_{n:n+1}(s)$, which is decreasing in s. Jensen's inequality then yields

$$\begin{aligned}
\left(\psi_{n-1:n} \circ \psi^{-1}_{n:n+1}\right)(F_{n:n+1}(v)) &= \left(\psi_{n-1:n} \circ \psi^{-1}_{n:n+1}\right) E_\theta\{\psi_{n:n+1}(F(v|\theta))\} \\
&\geq E_\theta\left\{\left(\psi_{n-1:n} \circ \psi^{-1}_{n:n+1}\right)(\psi_{n:n+1}(F(v|\theta)))\right\} \\
&= E_\theta\{\psi_{n-1:n}(F(v|\theta))\} \\
&= F_{n-1:n}(v)
\end{aligned}$$

so $\psi^{-1}_{n:n+1}(F_{n:n+1}(v)) \geq \psi^{-1}_{n-1:n}(F_{n-1:n}(v))$. □

Again, since we assume $F_{n-1:n}$ for each n can be inferred from bid data, Theorem 2 gives us two testable inequality restrictions implied by independence between valuations and N under the model of symmetric, conditionally independent private values: specifically, that

$$n > n' \quad \longrightarrow \quad F_{n-1:n}(v) \leq F_{n'-1:n'}(v) \tag{9}$$

$$n > n' \quad \longrightarrow \quad \psi^{-1}_{n-1:n}(F_{n-1:n}(v)) \geq \psi^{-1}_{n'-1:n'}(F_{n'-1:n'}(v)) \tag{10}$$

for all v.[18]

[17] For $n > 2$, algebra shows $\psi_{n-1:n}(s) - \psi_{n-2:n-1}(s) = -(n-1)s^{n-2}(1-s)^2 \leq 0$.

[18] Under the assumptions of Haile and Tamer (2003), with $G_{n-1:n}$ and $G_{n:n}$ the (observable) distributions of the second-highest and highest bids in n-bidder auctions, (9) and (10) would become $G_{n:n}(v-\delta) \leq G_{n'-1:n'}(v)$ and $\psi^{-1}_{n-1:n}(G_{n-1:n}(v)) \geq \psi^{-1}_{n'-1:n'}(G_{n':n'}(v-\delta))$, where δ is the minimum bid increment. With some modification, the tests we describe below could be applied to these inequalities instead.

Also note that both proposed tests are sharp in one sense: under an IPV data-generating process, (10) would hold everywhere with equality, and with perfectly correlated values which are independent of N, (9) would hold everywhere with equality. However, if either inequality holds strictly for some (n, n'), it could potentially be tightened for other pairs. For example, if $F_{n-4:n-3}(v) > F_{n-3:n-2}(v) > F_{n-2:n-1}(v)$, (9) could be tightened to $F_{n-1:n}(v) \leq F_{n-2:n-1}(v) - \frac{(n-1)(n-3)}{(n-2)^2} \frac{\left(F_{n-3:n-2}(v) - F_{n-2:n-1}(v)\right)^2}{F_{n-4:n-3}(v) - F_{n-3:n-2}(v)}$ – although due to the difference term in the denominator, this stronger version of (9) is highly unstable when applied to a reasonably-sized dataset.

3.3 Power Against Violations of Independence

Since we cannot directly test for independence, only for a consequence of it, it remains to argue why we expect this test to have power; that is, why dependence between valuations and N would likely lead to a violation of either (9) or (10). The logic is as follows.

The results of the last section state that if valuations are independent of auction size, then we should see a smaller competition effect in the data under CIPV as compared to that predicted by IPV. A violation of the test (10) would mean that the size of the competition effect in the data was observed to be *larger* than that predicted under IPV. What would cause transaction prices to increase with auction size at a rate faster than predicted by IPV? This would likely occur if larger auctions tended to coincide with higher individual valuations. But this is exactly our weaker assumption that valuations are stochastically increasing with auction size. Hence our test (10) of independence between valuations and auction size under CIPV is likely to have power against our weaker assumption that valuations are stochastically increasing with auction size.

To formalize this intuition, let $E_{\theta|n}$ denote an expectation taken over the distribution of θ among auctions with exactly n bidders. Independence requires θ to be independent of N, so this expectation would be the same for every n. Below, we show that if $\theta \not\perp N$ and auctions with more bidders have characteristics which put more weight (on average) in the upper tail of bidders' values, this will always lead to a violation of (10).

Theorem 3 *Suppose that for each θ, $F(\cdot|\theta)$ is continuous and twice differentiable, with derivative $f(\cdot|\theta)$, and has the same bounded support $[\underline{v}, \overline{v}]$. If $n > n'$ implies*

$$E_{\theta|n} \left(f(\overline{v}|\theta) \right)^2 > E_{\theta|n'} \left(f(\overline{v}|\theta) \right)^2$$

then (10) *will be violated for v sufficiently close to $\overline{v}$.*

The proof is in Appendix B.3. Under the same smoothness and bounded-support assumptions, we can also show that if valuations are stochastically increasing in N, $E_{\theta|n} f(\overline{v}|\theta)$ must be weakly increasing in n. While this latter condition is not enough to guarantee a violation of (10) under Theorem 3, the conditions are obviously very similar – both indicate a positive relationship between auction size and high bidder valuations. Thus, economic forces that give rise to valuations which are strictly stochastically increasing in N, will also likely lead to valuations satisfying the conditions for Theorem 3. For example, in Appendix A, we show general conditions under which the endogenous entry models of Levin and Smith (1994) and Samuelson (1985) would both lead to violations of (10) via Theorem 3. Thus, informally, we see (10) as a test of independence between valuations and N, against the alternative that valuations are stochastically increasing.

A *negative* relationship between auction size and valuations – valuations which are stochastically *decreasing* in auction size – would not lead to a violation of (10); but if it were a strong enough effect, it would lead to a violation of (9).[19] Thus, together, we feel (9) and (10) give a reasonable test of the combined assumptions of conditionally independent private values which are independent of auction size.

3.4 Econometric Test

Next, we introduce formal econometric tests for the inequalities (9) and (10). We construct a test statistic for each inequality which is calculated from observed auction data; asymptotically, each one will diverge to $+\infty$ if the corresponding inequality is violated anywhere; converge to 0 in probability if it holds everywhere as a strict inequality; and converge to a normal distribution with mean 0 and bounded variance if it holds everywhere, and holds with equality on a positive measure of v. We choose to test (9) and (10) individually (rather than jointly) because knowing which – if either – is rejected will be suggestive of whether larger auctions are associated with higher or lower individual valuations.[20]

Lee, Linton, and Whang (2009) have recently described a test for stochastic monotonicity; their test follows a Kolmogorov-Smirnov (K-S) type of approach, based on the supremum of a properly normalized statistic. This approach could be used for our test of (9), a comparison of two distribution functions; however, it is not designed to test (10), which involves a *nonlinear* transformation of one of the distributions. By its nature, it seems unavoidable that any test of (10) must rely on an explicit estimation of a distribution function. A K-S approach in this setting would have the drawback of leading to a nonpivotal asymptotic distribution (see Remark 2.1 in Lee, Linton, and Whang (2009)). Consequently, critical values for the resulting test would have to rely on simulation or resampling. Rather than relying on this approach, our test is based on the *expected value* of a properly-designed

[19] In settings where independence is taken for granted, a violation of (9) (a stochastic decrease in transaction price as N increases) is interpreted as evidence of common values, since bidders shade their bids more as N increases due to a stronger winner's curse. (See, for example, Hong and Shum (2002), and the discussion in Bajari and Hortacsu (2003).) Here, we take private values as given, and argue that such a left-shift must therefore indicate negative correlation between valuations and N. While counterintuitive, with private values, this could occur naturally if more valuable objects also tend to be valued more consistently by different buyers. This could be the case, for example, if the most valuable objects tended to be purchased by professional dealers, who were fairly unanimous in their assessments, while lower-value objects appealed to individual collectors, whose tastes were more idiosyncratic. Still, we feel intuitively that a positive relationship between valuations and N is the more likely type of dependence; thus, we expect (10) to be the more meaningful test.

[20] These two possibilities have very different implications for the validity of our bounds; contrast Lemma 3 (and part 2 of Theorem 1) with footnote 12.

functional. The key advantage of our approach is that the critical values come from a known distribution. The asymptotic analysis of our test statistic is also simpler than the one that would follow from a K-S approach.

Our test is based on functionals whose expected values are zero if and only if (9) and (10) are satisfied with probability 1. Similar functionals were also used by Khan and Tamer (2009) in a setting of censored regression. As in their case, the sample analogs of the functionals we use take the form of U-statistics. The rejection rules that we recommend are based on the asymptotic properties of these U-statistics. The presence of U-statistics in econometrics is extensive, both for estimation and inference. Examples of M-estimators where the objective function is written explicitly as a U-statistic include the maximum rank correlation estimator studied by Han (1987) and Sherman (1993), as well as the estimation procedures in Dominguez and Lobato (2004) and Khan and Tamer (2009). U-statistics have also been used to construct consistent specification tests. Some examples include Fan and Li (1996), Zheng (1998), Chen and Fan (1999). Even though our goal is to devise a specification test (as opposed to the estimation of an unknown parameter), Khan and Tamer (2009) is perhaps the closest to the spirit of our test, which relies on transforming functional (e.g, conditional moment) inequalities into moment equalities. The goal in Khan and Tamer (2009) is to estimate a finite-dimensional parameter that is point-identified by *conditional moment inequalities* that must hold with probability one. They propose an M-estimation procedure based on a U-statistic objective function whose probability limit is uniquely minimized at the true parameter value. The general idea of transforming moment inequalities into equalities has also been recently studied by Andrews and Shi (2009) and Kim (2009), who propose inferential methods for finite-dimensional parameters partially identified by conditional moment inequalities. The approach in Andrews and Shi (2009) relies on the use of properly chosen “instruments”, while Kim (2009) employs U-statistics of the type used in Khan and Tamer (2009).

For clarity, we first present the test without conditioning on any auction-specific covariates. After that, we will extend the test to be run conditional on observable covariates; in our application, we run both the unconditional and conditional tests.

Assume we have observations from L ascending auctions, indexed by $i \in \{1, 2, \ldots, L\}$. For each auction i, we observe the winning bid (transaction price), W_i, and the number of bidders N_i. We assume these observations $Z_i = (W_i, N_i)$ are *i.i.d.* draws from some ex-ante distribution of auctions. Let

$$p_N(n) = \Pr(N = n), \quad F_{W|N}(w|n) = \Pr(W \leq w|N = n), \quad \text{and} \quad F_W(w) = \Pr(W \leq w)$$

be the unconditional distribution of N, the conditional distribution of W given N, and the unconditional distribution of W under the true data-generating process. Let $\mathcal{S}_Z$ denote the joint support of (W, N), and $\mathcal{S}_W$ and $\mathcal{S}_N$ the marginal supports of W and N, respectively. We assume that $F_{W|N}(\cdot|n)$ is continuous for each n in $\mathcal{S}_N$.

The function $\psi_{n-1:n}(\cdot)$ was defined in (4). For any pair $n, n' \in \mathcal{S}_N$ and any s and w, define

$$\begin{aligned}
\Omega(s, n, n') &= \psi_{n-1:n} \circ \psi^{-1}_{n'-1:n'}(s) \\
\Delta_{W|N}(w, n, n') &= F_{W|N}(w|n) - F_{W|N}(w|n') \\
\Phi_{W|N}(w, n, n') &= \Omega\left(F_{W|N}(w|n'), n, n'\right) - F_{W|N}(w|n)
\end{aligned}$$

so (9) and (10) are exactly the conditions that $\Delta_{W|N}(w, n, n') \leq 0$ and $\Phi_{W|N}(w, n, n') \leq 0$ for $n > n'$.

We will test (9) and (10) over a pre-specified subset of $\mathcal{S}_W$ and $\mathcal{S}_N$. Let $\mathscr{N} \subseteq \mathcal{S}_N$ be a compact, pre-specified set. For each $n, n' \in \mathscr{N}$ with $n \neq n'$, let $\mathscr{W}_{n,n'} \subseteq \mathcal{S}_W$ be any compact subset on which both $F_{W|N}(\cdot|n)$ and $F_{W|N}(\cdot|n')$ are strictly bounded away from 0 and 1. We will test whether $\Delta_{W|N}(w, n, n') \leq 0$ and $\Phi_{W|N}(w, n, n') \leq 0$ for almost all $w \in \mathscr{W}_{n,n'}$, for all $n, n' \in \mathscr{N}$ with $n > n'$.

Take any distinct $i, j \in \{1, \ldots, L\}$. For any $n, n' \in \mathscr{N}$, consider the two functions

$$\begin{aligned}
T^F_{n,n'}(Z_i, Z_j) &= \Big(\mathbb{1}\{W_i \leq W_j\} - F_{W|N}(W_j|n')\Big) \cdot \mathbb{1}\{N_i = n\} \\
&\quad \cdot \mathbb{1}\big\{\Delta_{W|N}(W_j, n, n') \geq 0\big\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\} \\
T^\Omega_{n,n'}(Z_i, Z_j) &= \Big(\Omega\big(F_{W|N}(W_j|n'), n, n'\big) - \mathbb{1}\{W_i \leq W_j\}\Big) \cdot \mathbb{1}\{N_i = n\} \\
&\quad \cdot \mathbb{1}\big\{\Phi_{W|N}(W_j, n, n') \geq 0\big\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\}
\end{aligned} \tag{11}$$

and let

$$\mu^F_{n,n'} \equiv E\Big[T^F_{n,n'}(Z_i, Z_j)\Big] \qquad \text{and} \qquad \mu^\Omega_{n,n'} \equiv E\Big[T^\Omega_{n,n'}(Z_i, Z_j)\Big] \tag{12}$$

Theorem 4 *If Z_i and Z_j are i.i.d., then for any $n, n' \in \mathscr{N}$, $\mu^F_{n,n'} \geq 0$ and $\mu^\Omega_{n,n'} \geq 0$. Further,*

(i) $\mu^F_{n,n'} = 0$ if and only if $F_{W|N}(w|n) \leq F_{W|N}(w|n')$ for almost all $w \in \mathscr{W}_{n,n'}$

(ii) $\mu^\Omega_{n,n'} = 0$ if and only if $\psi^{-1}_{n-1:n}\left(F_{W|N}(w|n)\right) \geq \psi^{-1}_{n'-1:n'}\left(F_{W|N}(w|n')\right)$ for almost all $w \in \mathscr{W}_{n,n'}$

Proof. To prove part (ii), fix (N_i, Z_j), and define

$$\begin{aligned}
\overline{T}^\Omega_{n,n'}(N_i, Z_j) &\equiv E_{W_i|N_i,Z_j}\left[T^\Omega_{n,n'}(Z_i, Z_j)\right] \\
&= E_{W_i|N_i}\Big[\Big(\Omega\big(F_{W|N}(W_j|n'), n, n'\big) - \mathbb{1}\{W_i \leq W_j\}\Big) \cdot \mathbb{1}\{N_i = n\} \cdot \mathbb{1}\big\{\Phi_{W|N}(W_j, n, n') \geq 0\big\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\}\Big] \\
&= \Big(\Omega\big(F_{W|N}(W_j|n'), n, n'\big) - F_{W|N}(W_j|N_i)\Big) \cdot \mathbb{1}\{N_i = n\} \cdot \mathbb{1}\big\{\Phi_{W|N}(W_j, n, n') \geq 0\big\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\} \\
&= \mathbb{1}\{N_i = n\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\} \cdot \max\big\{0, \Phi_{W|N}(W_j, n, n')\big\}
\end{aligned}$$

so $\overline{T}^\Omega_{n,n'}(N_i, Z_j) \geq 0$, and $\overline{T}^\Omega_{n,n'}(N_i, Z_j) > 0$ if and only if $N_i = n$, $W_j \in \mathscr{W}_{n,n'}$, and $\Phi_{W|N}(W_j, n, n') > 0$. By iterated expectations,

$$\mu^\Omega_{n,n'} = p_{\scriptscriptstyle N}(n) \cdot E\Big[\max\big\{0\ ,\ \Phi_{W|N}(W_j, n, n')\big\} \cdot \mathbb{1}\big\{W_j \in \mathscr{W}_{n,n'}\big\}\Big]$$

which is therefore weakly positive, and strictly positive if and only if $\Phi_{W|N}(W_j, n, n') > 0$ occurs with positive probability on $\mathscr{W}_{n,n'}$. The analogous result holds for $\mu^F_{n,n'}$ and $\Delta_{W|N}(W_j, n, n')$. $\square$

Thus, we have established the equivalence between a set of *inequality* constraints and a moment *equality* condition. $T^{\Omega}_{n,n'}(Z_i, Z_j)$ is a function whose expected value is 0 if (10) holds almost everywhere in $\mathscr{W}_{n,n'}$ and is strictly positive otherwise; and likewise $T^F_{n,n'}(Z_i, Z_j)$ with (9). We can use Theorem 4 to test (9) and (10) in our testing range by focusing on all pairs $n > n'$ in $\mathscr{N}$. By Theorem 4, the sums $\sum_{n,n' \in \mathscr{N}, n > n'} T^{\Omega}_{n,n'}(Z_i, Z_j)$ and $\sum_{n,n' \in \mathscr{N}, n > n'} T^F_{n,n'}(Z_i, Z_j)$ will have expected value 0 if (9) and (10) hold almost everywhere in $\mathscr{W}_{n,n'}$ for each $n > n'$ in $\mathscr{N}$, and each will have strictly positive expected value if the corresponding inequality is violated with positive probability. In broad terms, then, we will calculate sample averages of these sums over the pairs (Z_i, Z_j) in our data, and test whether the resulting averages are significantly different from 0. Sample analog estimators would take the form of (second order) U-statistics. U-statistics arise as generalizations of sample averages. They were introduced by Hoeffding (1948) and Halmos (1946); for a detailed overview of their general properties, see Chapter 5 in Serfling (1980). In many instances (but not always), statistics of this class have an asymptotically normal behavior.

However, we cannot directly calculate $T^{\Omega}_{n,n'}(Z_i, Z_j)$ and $T^F_{n,n'}(Z_i, Z_j)$, since the distributions $F_{W|N}$, and therefore the functions $\Delta_{W|N}$ and $\Phi_{W|N}$, are not known; we therefore replace them with nonparametric estimates. To facilitate the asymptotic analysis of our test statistics, we exclude observations i and j from the estimates of $F_{W|N}$ that are used in $T^{\Omega}_{n,n'}(Z_i, Z_j)$ and $T^F_{n,n'}(Z_i, Z_j)$. That is, for each $i \neq j$, we define

$$\widehat{R}^{-i,j}_{W|N}(w|n) \equiv \tfrac{1}{L-2} \textstyle\sum_{\ell \neq i,j} \mathbb{1}\{W_\ell \leq w\} \cdot \mathbb{1}\{N_\ell = n\}$$

$$\widehat{p}^{\,-i,j}_{N}(n) \equiv \tfrac{1}{L-2} \textstyle\sum_{\ell \neq i,j} \mathbb{1}\{N_\ell = n\}$$

$$\widehat{F}^{-i,j}_{W|N}(w|n) \equiv \widehat{R}^{-i,j}_{W|N}(w|n) \,/\, \widehat{p}^{\,-i,j}_{N}(n)$$

(and $\widehat{F}^{-i,j}_{W|N}(w|n) = 0$ if $\widehat{p}^{\,-i,j}_{N}(n) = 0$), and calculate corresponding estimates for $\Delta_{W|N}$ and $\Phi_{W|N}$

$$\widehat{\Delta}^{-i,j}_{W|N}(w, n, n') = \widehat{F}^{-i,j}_{W|N}(w|n) - \widehat{F}^{-i,j}_{W|N}(w|n')$$

$$\widehat{\Phi}^{-i,j}_{W|N}(w, n, n') = \Omega\big(\widehat{F}^{-i,j}_{W|N}(w|n'), n, n'\big) - \widehat{F}^{-i,j}_{W|N}(w|n)$$

In addition to replacing $F_{W|N}$, $\Delta_{W|N}$, and $\Phi_{W|N}$ with estimates, we make one other modification to (11). Our null hypothesis allows for both (9) and (10) to hold strictly or with equality. To deal with the possibility that either holds with equality with positive probability, we introduce a nonzero bandwidth b_L, which disappears at a rate slower than $\sqrt{L}$, so that we can characterize exponential bounds for the probability that $\widehat{\Delta}^{-i,j}_{W|N}(W_j, n, n') < -b_L$ and $\Delta_{W|N}(W_j, n, n') \geq 0$ (or that $\widehat{\Phi}^{-i,j}_{W|N}(W_j, n, n') < -b_L$ and $\Phi_{W|N}(W_j, n, n') \geq 0$). Thus, our estimators are of the form[21]

$$
\begin{aligned}
\widehat{T}^F_{n,n'}(Z_i, Z_j) &= \Big(\mathbb{1}\{W_i \leq W_j\} - \widehat{F}^{-i,j}_{W|N}(W_j|n')\Big) \cdot \mathbb{1}\{N_i = n\} \\
&\quad \cdot \mathbb{1}\big\{\widehat{\Delta}^{-i,j}_{W|N}(W_j, n, n') \geq -b_L\big\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\} \\
\widehat{T}^\Omega_{n,n'}(Z_i, Z_j) &= \Big(\Omega\big(\widehat{F}^{-i,j}_{W|N}(W_j|n'), n, n'\big) - \mathbb{1}\{W_i \leq W_j\}\Big) \cdot \mathbb{1}\{N_i = n\} \\
&\quad \cdot \mathbb{1}\big\{\widehat{\Phi}^{-i,j}_{W|N}(W_j, n, n') \geq -b_L\big\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\}
\end{aligned}
\tag{13}
$$

Our test statistics are then the sample averages of $\widehat{T}^F_{n,n'}$ and $\widehat{T}^\Omega_{n,n'}$, summed over the different (n, n') combinations where $n > n'$:

$$
\begin{aligned}
U^{\widehat{T}^F}_{L^{(2)}} &= \frac{1}{L(L-1)} \sum\sum_{i,j \in \{1,\ldots,L\}, i \neq j} \left[\sum\sum_{n,n' \in \mathscr{N}, n > n'} \widehat{T}^F_{n,n'}(Z_i, Z_j) \right] \\
U^{\widehat{T}^\Omega}_{L^{(2)}} &= \frac{1}{L(L-1)} \sum\sum_{i,j \in \{1,\ldots,L\}, i \neq j} \left[\sum\sum_{n,n' \in \mathscr{N}, n > n'} \widehat{T}^\Omega_{n,n'}(Z_i, Z_j) \right]
\end{aligned}
\tag{14}
$$

Appendix B.4 provides regularity conditions involving the distributions $F_{W|N}$ under which we can fully characterize the asymptotic distribution of these test statistics. The results are given in Theorem 5; the proof is contained in a separate Technical Supplement. In brief, we first show that under these regularity assumptions, we can express our test statistics as $\sqrt{L} \cdot U^{\widehat{T}^F}_{L^{(2)}} = \sqrt{L} \cdot U^{T^F}_{L^{(2)}} + \sqrt{L} \cdot U^{\widehat{G}^F}_{L^{(3)}} + o_p(1)$ and $\sqrt{L} \cdot U^{\widehat{T}^\Omega}_{L^{(2)}} = \sqrt{L} \cdot U^{T^\Omega}_{L^{(2)}} + \sqrt{L} \cdot U^{\widehat{G}^\Omega}_{L^{(3)}} + o_p(1)$, where $U^{T^F}_{L^{(2)}}$ and $U^{T^\Omega}_{L^{(2)}}$ are the second order U-statistics that would result if we replaced $\widehat{T}^F_{n,n'}$ and $\widehat{T}^\Omega_{n,n'}$ with $T^F_{n,n'}$ and $T^\Omega_{n,n'}$ in (14). $U^{\widehat{G}^F}_{L^{(3)}}$ and $U^{\widehat{G}^\Omega}_{L^{(3)}}$ are third-order U-statistics whose asymptotic presence reflects our use of nonparametric estimators in the construction of $U^{\widehat{T}^F}_{L^{(2)}}$ and $U^{\widehat{T}^\Omega}_{L^{(2)}}$. We then characterize the Hoeffding decomposition (see Hoeffding (1961) and Serfling (1980), Chapter 5) of each of these U-statistics. From here, we obtain linear representations of the form

$$
\begin{aligned}
\sqrt{L} \cdot U^{\widehat{T}^F}_{L^{(2)}} &= \sqrt{L} \cdot \left[\sum\sum_{n,n' \in \mathscr{N}, n > n'} \mu^F_{n,n'} \right] + \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \eta^F_{\mathscr{N}}(Z_i) + o_p(1) \\
\sqrt{L} \cdot U^{\widehat{T}^\Omega}_{L^{(2)}} &= \sqrt{L} \cdot \left[\sum\sum_{n,n' \in \mathscr{N}, n > n'} \mu^\Omega_{n,n'} \right] + \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \eta^\Omega_{\mathscr{N}}(Z_i) + o_p(1)
\end{aligned}
\tag{15}
$$

with $E[\eta^F_{\mathscr{N}}(Z_i)] = E[\eta^\Omega_{\mathscr{N}}(Z_i)] = 0$. The function $\eta^F_{\mathscr{N}}(Z_i)$ arises from the conditional expectations which appear as leading terms of the Hoeffding decompositions of the U-statistics $U^{T^F}_{L^{(2)}}$ and $U^{\widehat{G}^F}_{L^{(3)}}$.

[21] Kim (2009) used a bandwidth sequence serving the same purpose as ours. Our asymptotic results are valid for *any* bandwidth sequence satisfying the convergence rate requirements discussed in Appendix B.4. Choosing bandwidths for optimal finite-sample properties is an open question, and one we hope to address in future work.

The remaining terms of these decompositions are of order $o_p(L^{-1/2})$. The function $\eta^{\Omega}_{\mathscr{N}}(Z_i)$ arises analogously. We discuss the structure of these functions in Appendix B.4 and in the separate Technical Supplement. The following asymptotic results, then, follow directly from Theorem 5:[22]

Summary of Asymptotic Results. *Let $\widehat{\Sigma}_{F_{W|N}}$ be a consistent estimator of Var$\left(\eta^F_{\mathscr{N}}(Z_i)\right)$.*[23] *Let $c_1 > 0$ be any positive number, and $\widehat{D}_{F_{W|N}} = \widehat{\Sigma}^{1/2}_{F_{W|N}} + c_1$. Under appropriate regularity conditions (described in Appendix B.4), as $L \to \infty$,*

1a. $\sqrt{L} \cdot U^{\widehat{T}^F}_{L(2)} \Big/ \widehat{D}_{F_{W|N}} \xrightarrow{p} +\infty$ *if* (9) *is violated with positive probability*

1b. $\sqrt{L} \cdot U^{\widehat{T}^F}_{L(2)} \Big/ \widehat{D}_{F_{W|N}} \xrightarrow{d} N(0, \sigma^2)$ *with* $\sigma^2 \in (0,1)$ *if* (9) *holds almost everywhere, and holds with equality with positive probability*

1c. $\sqrt{L} \cdot U^{\widehat{T}^F}_{L(2)} \Big/ \widehat{D}_{F_{W|N}} \xrightarrow{p} 0$ *if* (9) *holds strictly almost everywhere*

Similarly, let $\widehat{\Sigma}_{\Omega_{W|N}}$ be a consistent estimator of Var$\left(\eta^{\Omega}_{\mathscr{N}}(Z_i)\right)$, $c_2 > 0$ any positive number, and $\widehat{D}_{\Omega_{W|N}} = \widehat{\Sigma}^{1/2}_{\Omega_{W|N}} + c_2$; under appropriate regularity conditions, as $L \to \infty$,

2a. $\sqrt{L} \cdot U^{\widehat{T}^\Omega}_{L(2)} \Big/ \widehat{D}_{\Omega_{W|N}} \xrightarrow{p} +\infty$ *if* (10) *is violated with positive probability*

2b. $\sqrt{L} \cdot U^{\widehat{T}^\Omega}_{L(2)} \Big/ \widehat{D}_{\Omega_{W|N}} \xrightarrow{d} N(0, \sigma^2)$ *with* $\sigma^2 \in (0,1)$ *if* (10) *holds almost everywhere, and holds with equality with positive probability*

2c. $\sqrt{L} \cdot U^{\widehat{T}^\Omega}_{L(2)} \Big/ \widehat{D}_{\Omega_{W|N}} \xrightarrow{p} 0$ *if* (10) *holds strictly almost everywhere*

Based on these results, we can apply rejection rules based on the standard normal distribution. Consider the null hypothesis H^F_0 :*(9) is satisfied at almost all $w \in \mathscr{W}_{n,n'}$, for every $n, n' \in \mathscr{N}$ with* $n > n'$. Let $\alpha \in (0,1)$ denote a pre-specified significance level and let z_{α} satisfy $\Pr(\mathscr{Z} \geq z_\alpha) = \alpha$, where $\mathscr{Z} \sim \mathcal{N}(0,1)$. Consider the rule

$$\textit{Reject } H^F_0 \textit{ if } \frac{\sqrt{L} \cdot U^{\widehat{T}^F}_{L(2)}}{\widehat{D}_{F_{W|N}}} \geq z_{\alpha} \tag{16A}$$

This decision rule has the following asymptotic properties:

$$\begin{array}{lcl} \lim_{L \longrightarrow \infty}\Big\{\Pr\big[\ H^F_0 \text{ is rejected when it is true}\big]\Big\} & \leq & \alpha \\ \lim_{L \longrightarrow \infty}\Big\{\Pr\big[\ H^F_0 \text{ is rejected when it is false }\big]\Big\} & = & 1 \end{array} \tag{17}$$

[22] Note that "positive probability" refers to non-zero probability over $\mathscr{W}_{n,n'}$ for *some* $n, n' \in \mathscr{N}$ with $n > n'$, and "almost everywhere" refers to probability 1 over $\mathscr{W}_{n,n'}$ for *every* $n, n' \in \mathscr{N}$ with $n > n'$.

[23] Analytic expressions leading to consistent estimators for Var $\left(\eta^F_{\mathscr{N}}(Z_i)\right)$ and Var $\left(\eta^{\Omega}_{\mathscr{N}}(Z_i)\right)$ are described in (31) and (32), in Appendix B.4.

The equivalent result holds for H_0^Ω :*(10) is satisfied at almost all* $w \in \mathscr{W}_{n,n'}$*, for every* $n, n' \in \mathscr{N}$ *with* $n > n'$, and a decision rule of the form

$$\textit{Reject } H_0^\Omega \textit{ if } \frac{\sqrt{L} \cdot U_{L(2)}^{\widehat{T}^\Omega}}{\widehat{D}_{\Omega_{W|N}}} \geq z_\alpha \tag{16B}$$

Rejecting (9) or (10) would imply the rejection not only of CIPV, but also of IPV.[24] On the other hand, failing to reject (9) and (10) could be attributed to the data being consistent with IPV. To check whether this is the case, we can test whether the reverse inequality in (10) holds: namely, whether

$$\psi_{n-1:n}^{-1}\left(F_{n-1:n}(v)\right) \quad \leq \quad \psi_{n'-1:n'}^{-1}\left(F_{n'-1:n'}(v)\right) \tag{10'}$$

for every $n, n' \in \mathscr{N}$ with $n > n'$ and almost all $v \in \mathscr{W}_{n,n'}$. If the data supports (10), then rejecting (10') implies that the inequalities in (10) are strict with positive probability, which rules out IPV as the true model. A test of (10') would replace $T_{n,n'}^\Omega(Z_i, Z_j)$ with

$$\begin{aligned} T_{n,n'}^{-\Omega}(Z_i, Z_j) \quad = \quad & \Big(\mathbb{1}\{W_i \leq W_j\} - \Omega\big(F_{W|N}(W_j|n'), n, n'\big)\Big) \cdot \mathbb{1}\{N_i = n\} \\ & \cdot \mathbb{1}\big\{-\Phi_{W|N}(W_j, n, n') \geq 0\big\} \cdot \mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\} \end{aligned} \tag{11'}$$

The estimator $\widehat{T}_{n,n'}^{-\Omega}(Z_i, Z_j)$ is constructed accordingly. The resulting U-statistic will have asymptotic properties analogous to the other test statistics, under similar regularity conditions – see Appendix B.4.4 for details.

3.5 Monte Carlo Simulations

To illuminate its small-sample properties, we perform our test of (10) on simulated data. For three different data-generating processes, we generated $1,000$ data sets of sizes $L = 200$, $L = 400$ and $L = 800$, and calculated the test-statistic for each simulated data set. The three data-generating processes were:

Valuations dependent on N: we used an example of the entry game from Levin and Smith (1994) to jointly generate N and W. (This example is solved in Appendix B.3; theoretical graphs of $\psi_{n-1:n}^{-1}\left(F_{n-1:n}(v)\right)$, showing a violation of (10), are shown in Figure 4.)

Valuations independent of N *(A)*: we used the same data-generating process for valuations as the previous case, but chose N independently of θ; this led to (10) holding *strictly* everywhere.

[24] That is, it would reject the notion that valuations are independent of N *and* independent across bidders.

Valuations independent of N (B): we used a different data-generating process for valuations, chosen so that (10) would hold *with equality* with positive probability.

The details of the simulations – the exact data-generating processes, the bandwidths b_L used, and so on – are described in Appendix A.

Table 1 gives the empirical rejection rates observed from applying the test above to the simulated data, with a target significance level of 5%. These results exactly reflect the asymptotic predictions described above. As L grows, the probability of rejecting independence between valuations and N goes to 1 when there is dependence. The probably of rejecting independence goes to 0 when the true data-generating process has independence and (10) holds as a strict inequality with probability one; and if (10) holds with equality with nonzero probability, the empirical rejection probabilities are close to the target significance level of 5%. (These results are particularly encouraging since in our application, the sample size is over 2,000.)

Table 1: Empirical rejection rates at target significance level of 5%

	$L = 200$	$L = 400$	$L = 800$
Valuations dependent on N	58.4%	82.4%	98.3%
Valuations independent of N (A)	3.1%	1.2%	1.1%
Valuations independent of N (B)	7.0%	4.7%	5.2%

3.6 Testing Conditional on Observable Covariates

In Appendix B.5, we show how the test above can be modified to condition nonparametrically on observable covariates. We test whether the conditional analogs of (9) and (10) hold at each realization of $\mathbf{X}$ – that is, whether

$$n > n' \quad \longrightarrow \quad F_{n-1:n}(v|\mathbf{X}) \;\leq\; F_{n'-1:n'}(v|\mathbf{X}) \tag{18}$$

$$n > n' \quad \longrightarrow \quad \psi^{-1}_{n-1:n}\left(F_{n-1:n}(v|\mathbf{X})\right) \;\geq\; \psi^{-1}_{n'-1:n'}\left(F_{n'-1:n'}(v|\mathbf{X})\right) \tag{19}$$

for almost all v and $\mathbf{X}$ in a pre-specified testing range $\mathscr{C}_{n,n'}$ (as in the unconditional case). For clarity, in Appendix B.5 we present the test conditional on a one-dimensional covariate X which is continuously distributed; as we discuss below, we will apply the test conditional only on appraisal value. The test extends naturally to cases where $\mathbf{X}$ is multi-dimensional and includes both discrete and continuous covariates. As is usually the case in nonparametric procedures, preserving the asymptotic features of our test when $\mathbf{X}$ includes multidimensional continuously distributed covariates comes at a computational cost brought about by the need to use so-called "bias-reducing kernels" of higher order. The degree of smoothness required from the various functionals involved also increases more than proportionately with the dimension of the continuously distributed elements in $\mathbf{X}$.

Our conditional tests are based on *third-order* U-statistics (regardless of the dimension of $\mathbf{X}$). A consistent test of (18) and (19) requires matching $\mathbf{X}$ across pairs of observations in our sample. If $\mathbf{X}$ contains continuously distributed covariates, the probability of having two (or more) identical realizations of $\mathbf{X}$ in our sample is zero. Since we focus on the case where $\mathbf{X}$ is continuous, in our proposed test this matching is achieved through the use of kernel-based weights and a bandwidth sequence converging to zero. With the proper kernel and bandwidth, our methodology ensures the type of asymptotic matching of $\mathbf{X}$ needed to test (18) and (19) consistently. Weighting methods with these types of asymptotic properties are referred to as *pairwise differencing* and have been studied, for example, in Honoré and Powell (1994), Honoré and Powell (2005), Aradillas-Lopez, Honoré, and Powell (2007) and Hong and Shum (2009). Under appropriate regularity conditions (given in Appendix B.6), the conditional test has asymptotic properties exactly analogous to the unconditional test; these properties are summarized in Theorem 7.

4 Application to USFS Timber Auctions

4.1 Description of Data

We apply our test to data from timber auctions run by the Unites States Forest Service. A number of other papers have considered Forest Service auctions. Baldwin, Marshall, and Richard (1997) provide much institutional background. Their focus is to test for collusion. Haile (2001) considers the effects of resale on valuations. Haile, Hong, and Shum (2003) develop and apply a test for common values against private values, assuming away (for the most part) unobserved heterogeneity. Lu and Perrigne (2008) use the USFS data to estimate risk aversion among bidders, using the fact that the service conducts both first price sealed bid auctions and open ascending auctions. Finally, Athey and Levin (2001), Athey, Levin, and Seira (2008), and Haile and Tamer (2003) analyze the data to empirically study mechanism design issues.[25] All papers besides a few elements of Haile, Hong, and Shum (2003) do not consider the effects of unobserved heterogeneity in drawing inferences

[25] One of the key insights of Athey, Levin, and Seira (2008) is that there are ex-ante asymmetries between two types of bidders: millers and loggers. This does not preclude the use of our model. Suppose bidders of each type have private values drawn from the distributions $F_m(v_i \mid \mathbf{X}, \theta)$ and $F_l(v_i \mid \mathbf{X}, \theta)$, respectively, and that a fraction $q(\mathbf{X}, \theta)$ of bidders are expected to be millers. As long as each bidder is imagined to be randomly (independently) either a miller or a logger, each bidder's valuation can still be seen as an independent draw from $F(v_i \mid \mathbf{X}, \theta) = q(\mathbf{X}, \theta)F_m(v_i \mid \mathbf{X}, \theta) + (1 - q(\mathbf{X}, \theta))F_l(v_i \mid \mathbf{X}, \theta)$. Since beliefs about other bidders' valuations do not affect bidding in a private-value

from the data, which is our current concern. Thus we hope this application to be fruitful, as we analyze widely-studied auctions in which the role of heterogeneity has not been thoroughly addressed.

These auctions proceed as follows. The Forest Service first conducts a "cruise" of the tract and publishes detailed information on the tract for potential bidders. It also announces an appraisal value for the tract, which serves as a reserve price for the qualifying round of bids: bidders must submit sealed bids of at least that reserve price to be eligible to participate. The oral round of the auction then begins at the highest of these sealed bids.

Data on these auctions comes to us from Phil Haile, who has posted timber auction data from 1978 to 1996. In addition to each bidder's highest bid in the oral round of the auction, we have data on the size of the tract, the estimated volumes of the various timber species, estimated costs of harvesting and delivering the timber to mills, and estimated revenue from such sales, in addition to other measures. Potential bidders can also conduct their own cruises (though this is rarely done for the subset of auctions we consider). It is generally recognized that the reserve prices set by the Forest Service are too low and do not bind.[26] A particular empirical question of interest for us is whether the Forest Service should be setting a higher reserve to extract greater rents, which has been a subject of ongoing debate within the timber service.

We use the cleaning conventions of Haile and Tamer (2003) to select auctions which are most likely to satisfy the assumption of private values. We focus on sales in Region 6 (which encompasses mostly Oregon) and select sales whose contracts expire within a year, to shut down the effect of resale possibilities on valuations. We focus on scaled sales, where bids are per unit of timber actually harvested, and therefore common-value uncertainty about the total amount of timber should not affect valuations. We consider only auctions held between 1982 and 1990, as the reserve price policy within this period was stable. What is left is a sample of 2,181 auctions with N between 2 and 12, in which the private values assumption is thought best to hold.[27] A unique feature of the data is that we have clean observations on the number of bidders in the room when the bidding begins, which is crucial to both our test and our identification strategy. We drop those observations with $N = 12$, because auctions with more than twelve bidders are listed as $N = 12$; this leaves us with an actual sample size of $L = 2,036$.

---

ascending auction, our model still applies, although our stronger identifying assumption now requires that conditional on observables, q is independent of N.

[26] 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).

[27] Like Haile and Tamer (2003), we drop auctions with one bidder from the analysis, because without binding reserve prices, they give us no information about the distribution of valuations. $F_{1:1}$ can still be identified from $\{F_{m-1:m}\}_{m\geq 2}$ via (6) or (7), and $\max_r (r - v_0)(1 - F_1(r))$ can then be solved to find the optimal reserve price when $N = 1$.

An interesting feature of the data is that the observable auction characteristics broadly appear to be independent of the number of bidders. As we discuss below, the government's appraisal value is the most significant covariate in explaining transaction price; but it has almost no relation to N. (The distribution of appraisal values is nearly the same across different values of N; if we regress $\log N$ on the log of appraisal value, we get a statistically insignificant relationship and an R^2 of 0.001.) Other auction-specific covariates also do not appear to be meaningfully related to N. Since bidders do not appear to be selectively participating in auctions based on their observable characteristics, it might be reasonable to think that they are also not selectively participating based on unobservable characteristics, and so independence between valuations and N seems plausible. Next, we move on to testing this possibility.

4.2 Conditioning on Observable Covariates

Since our methods explicitly allow for unmodeled heterogeneity, they can in principle be applied either unconditionally or conditional on observed covariates. If we apply our test of independence between valuations and auction size unconditionally, it becomes a test of whether valuations are *i.i.d.* draws from a distribution $F(\cdot|\mathbf{X}, \theta)$, with "independence" now meaning that N is independent of $(\mathbf{X}, \theta)$. If this is true, then conditional on a realization of $\mathbf{X}$, N is still independent of θ; and so if our "unconditional" model is valid, there is no need to worry about it becoming invalid when we condition on any subset of the covariates.

However, conditioning on observable covariates is potentially appealing for several reasons. First, if N is independent of θ but not of $\mathbf{X}$, then our hypothesis of independence (and thus our identification strategy) is invalid when we do not condition, but valid when we condition on $\mathbf{X}$. Second, we pointed out above that positive correlation between N and valuations would speed up the rate at which $F_{n-1:n}$ shifts to the right as n increases, while correlation among valuations slows it down; thus, heterogeneity can in a sense "conceal" some degree of dependence between valuations and N. By conditioning on $\mathbf{X}$, we remove some of the heterogeneity that could do this, giving us a more powerful test. So conditioning on available covariates makes our identifying assumption more likely to be valid, and more likely to be rejected if it is not. It would also be useful to know whether our model is in some sense "necessary" – that is, whether there truly is *unobservable* heterogeneity inducing correlation, or whether the data could be well described by the standard (IPV) model when observable covariates were controlled for. And finally, when we move on from testing to applying our identification results and analyzing counterfactuals, conditioning on observables will give us tighter results.

As discussed above, our data includes a rich set of auction-specific covariates, including the Forest Service's appraisal value for each tract. These appraisal values were constructed using the "residual" method, which takes the estimated revenue from the downstream sales of the lumber derived from the timber and subtracts off the estimated costs of cutting, transporting, and processing the timber, and an additional amount designed to serve as a "reasonable" profit for lumber supply. Thus, by design, the appraisal value is meant to contain all profit-relevant information contained in the auction-tract covariates; it therefore might plausibly play the role of a sufficient statistic for all observable auction heterogeneity.

For this reason, several papers on timber auctions use the empirical strategy of conditioning only on appraisal value. Campo, Guerre, Perrigne, and Vuong (2002), analyzing first-price auctions with risk-averse bidders and independent private values, condition only on appraisal value, writing, "When regressing the logarithm of bids per mbf [thousand board-feet] on a complete set of variables characterizing the auctioned lot... only two variables are strongly significant, namely the number of bidders and the appraisal value.... Thus, the appraisal value seems to be the best candidate to capture the heteregoneity across auctioned objects." Perrigne (2003) similarly notes, "Because the appraisal value is the most accurate variable capturing the different species and their calibers, we ignore other characteristics in the empirical analysis." Lu and Perrigne (2008) find both appraisal value and volume of timber to have explantory power; but we focus only on scaled sales, where volume does not play a signifcant role.

We therefore will perform both our conditional test and counterfactual analyses conditional only on appraisal value, which we label X. We think of appraisal value as being an "almost-sufficient statistic" for observable covariates. Our methods do not depend on appraisal value actually capturing all observable heterogeneity, as whatever observable heterogeneity remains is simply picked up as part of the unobserved θ; but at the same time, following the literature, we believe appraisal value to capture *most* of the observable variation, and therefore think of θ as, for the most part, reflecting truly *unobserved* heterogeneity.

4.3 Application of the Test

Recall that Theorem 2 gives two strings of inequalities which must hold if valuations are independent of N:

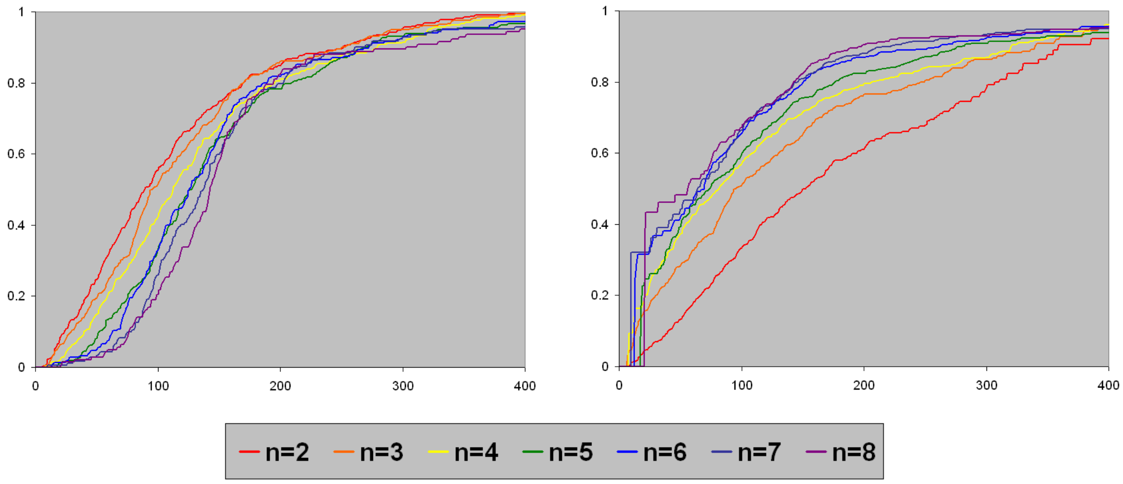
$$\begin{array}{ccccccc} F_{1:2}(v) & \geq & F_{2:3}(v) & \geq & F_{3:4}(v) & \geq & \cdots \\ \psi_{1:2}^{-1}(F_{1:2}(v)) & \leq & \psi_{2:3}^{-1}(F_{2:3}(v)) & \leq & \psi_{3:4}^{-1}(F_{3:4}(v)) & \leq & \cdots \end{array} \tag{20}$$

As discussed above, we assume that the transaction price in each auction in the data was exactly equal to the second-highest valuation $V_{n-1:n}$. As a first check, then, we can simply construct empirical analogs of the distribution functions $F_{n-1:n}(v)$ directly from bid data, plot $\widehat{F}_{n-1:n}(v)$ and $\psi_{n-1:n}^{-1}\left(\widehat{F}_{n-1:n}(v)\right)$ against v for various n, and check whether these distributions satisfy (20). Even without formalizing this into a proper test, this should give some intuition for whether independence between valuations and auction size is a plausible hypothesis given our data.

For a first pass, we let θ capture all auction heterogeneity, that is, we make no attempt to control for observable covariates. Figure 1 shows $\widehat{F}_{n-1:n}(v)$ and $\psi_{n-1:n}^{-1}\left(\widehat{F}_{n-1:n}(v)\right)$ as n varies from 2 to 8. (For visual ease, colors go in rainbow order – red, orange, yellow, green, blue, indigo, violet – as n increases from 2 to 8.) Visual inspection shows that these curves do indeed shift nearly monotonically in the predicted direction as n changes. Thus, without controlling for any observed heterogeneity, the bid data seems to be consistent with our model and the assumption of independence between valuations and auction size.

Figure 1: The "eyeball test" for independence between valuations and auction size

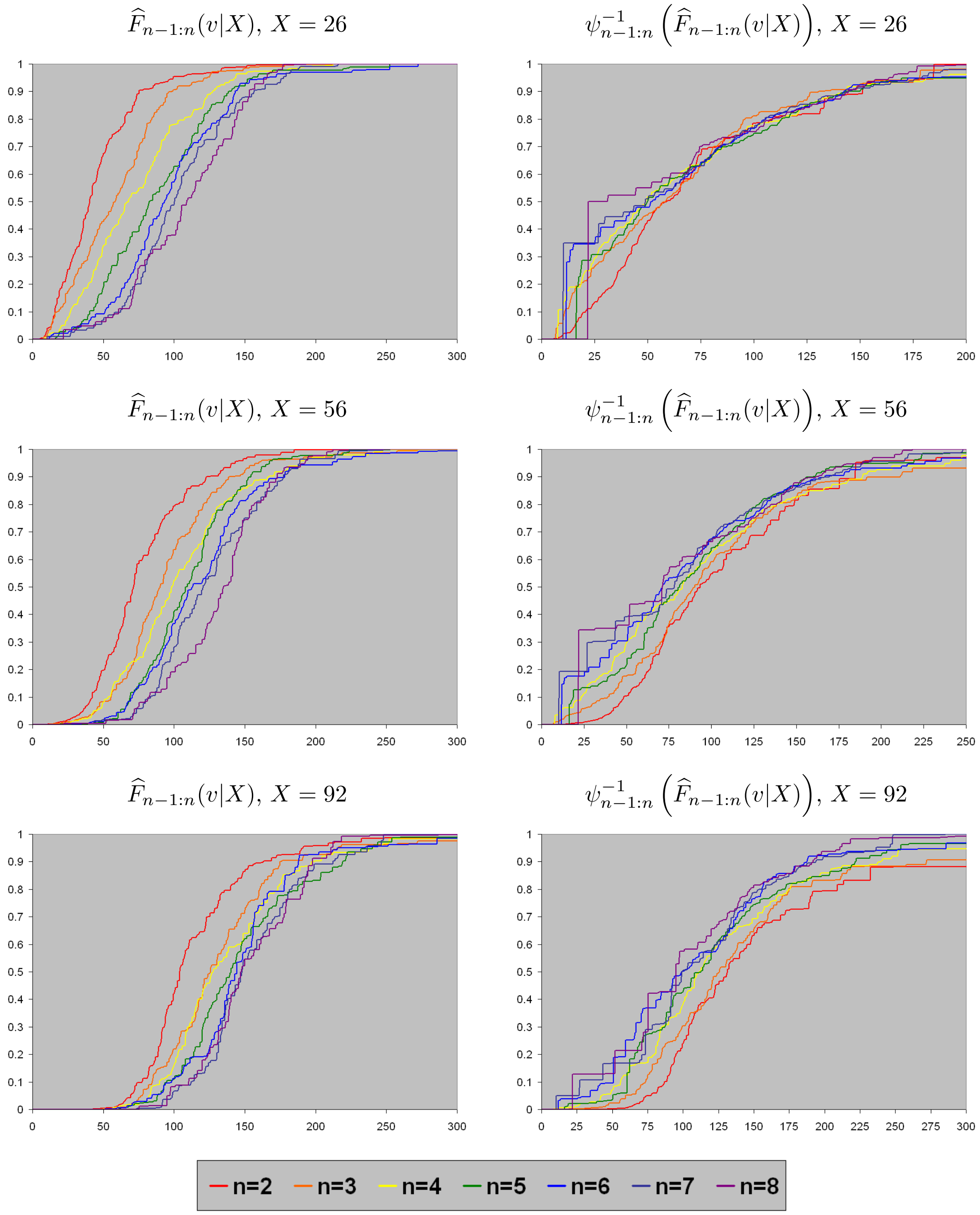
$\widehat{F}_{n-1:n}(v)$ (should shift **right** as n increases) | $\psi_{n-1:n}^{-1}\left(\widehat{F}_{n-1:n}(v)\right)$ (should shift **left** as n increases)



Next, we perform the same visual test conditional on appraisal value – that is, based on nonparametric estimates of the distributions of transaction prices for each N, conditional on appraisal value X. Letting $(W_i, N_i, X_i)_{i=1}^{L}$ denote the transaction price, number of bidders, and appraisal value observed in the i^{th} auction in the data, we construct a Nadaraya-Watson kernel estimate for each distribution $F_{n-1:n}(v|X)$ as

$$\widehat{F}_{n-1:n}(v|X) = \frac{\sum_{i=1}^{L} \mathbb{1}\{N_i = n\} K_b(X - X_i) \mathbb{1}\{W_i \leq v\}}{\sum_{i=1}^{L} \mathbb{1}\{N_i = n\} K_b(X - X_i)}$$

Figure 2: The same "eyeball test," using estimates of $F_{n-1:n}$ conditional on appraisal value



where K_b is the Gaussian kernel with bandwidth b. We use a bandwidth based on the rule of thumb used by Haile and Tamer (2003) for the same data. We can then apply the same visual test as before, to these conditional CDFs. Figure 2 shows graphs of $\widehat{F}_{n-1:n}(\cdot|X)$ and $\psi_{n-1:n}^{-1}\left(\widehat{F}_{n-1:n}(\cdot|X)\right)$ for three representative appraisal values, 26, 56, and 92 – the 25^{th}, 50^{th}, and 75^{th} percentiles of appraisal values in the data. Our test is again the claim that in each row, the curves in the first column shift to the right as n increases, while the curves in the second column shift to the left.

Examining Figure 2, there are two main features to notice. The first is that by controlling for the observable auction heterogeneity through appraisal value, there is less heterogeneity in the residual θ to slow down the effect of n on transaction prices; the curves in the right column are therefore much closer together than in the unconditional case. Second, the data seems to fit the hypothesis of independence of valuations and N fairly well. All three plots of $\widehat{F}_{n-1:n}(v|X)$ show a clear shift to the right as n increases, although there is some crossing of curves in a few spots. As for the plots of $\psi_{n-1:n}^{-1}\left(\widehat{F}_{n-1:n}(v|X)\right)$, for $X = 56$ and 92, these are also for the most part ordered correctly (shifting to the left as n increases); at $X = 26$, for v above about 75, the curves are virtually indistinguishable from each other. These "eyeball" tests suggest that either unconditionally or conditional on appraisal value, our data appear to be broadly consistent with our model and independence of valuations and N.

Next, we apply the formal test of (9) and (10) introduced above, to see whether the occasional apparent violations of these inequalities are explained by sampling variability or refute the model. In addition, the assumption of IPV implies the added restriction that (10) holds everywhere with equality, which appears not to hold in Figures 1 and 2; we naturally would like to ask whether the curves represent real econometric rejections of the equality test proposed in Athey and Haile (2002). Table 2 shows the results of all three unconditional tests.[28]

Table 2: Unconditional Test Results

Eq. (9)	Eq. (10)	IPV†
-2.0209	-4.4134	19.2590

(†) Test-statistic for (10'), using (11').

[28] We applied the test on the range $\mathscr{N} = \{2, 3, 4, \ldots, 11\}$ and $\mathscr{W}_{n,n'} = \{w : 0.02 \leq F_{W|N}(w|m) \leq 0.98 \text{ for } m = n, n'\}$, and used $\mathbb{1}\{0.02 \leq \widehat{F}_{W|N}^{-i,j}(W_j|n) \leq 0.98\} \cdot \mathbb{1}\{0.02 \leq \widehat{F}_{W|N}^{-i,j}(W_j|n') \leq 0.98\}$ to estimate $\mathbb{1}\{W_j \in \mathscr{W}_{n,n'}\}$ in the calculation of $\widehat{T}_{n,n'}^{F}(Z_i, Z_j)$, $\widehat{T}_{n,n'}^{\Omega}(Z_i, Z_j)$, and $\widehat{T}_{n,n'}^{-\Omega}(Z_i, Z_j)$. For all three tests, we used $b_L = 0.001$ and the additive constants $c = 10^{-6}$. For estimating variances, we used the analytic expressions given in (31) and (32) of Appendix B.4.3 and (37) and (38) in Appendix B.4.4. The sample size is $L = 2036$.

These results strongly suggest a data-generating process satisfying (9) and (10). (Their magnitude could be the result of both holding as strict inequalities with probability one, in which case the variance estimators used in the denominator might be quite small. Asymptotically, the variance estimators would be dominated by $c_1 = c_2 = 10^{-6}$, causing both test statistics to converge to 0 in probability; but at our sample size, this may not yet be happening. Nonetheless, we see strong support for (9) and (10); since both test statistics are negative, we could not reject either (9) or (10) regardless of the normalizing variance.)

Our results also reflect strong evidence against (10'). Thus, if we do not condition on any covariates, our data are consistent with conditionally independent private values which are independent of N, but not consistent with independent private values, confirming the visual intuition in Section 4.3. All of our qualitative results were found to be robust to moderate changes in b_L which led, in turn, to moderate fluctuations in the values of all three statistics, but the qualitative rejection results remained intact for a 5% significance level.

Finally, we apply the formal test conditional on appraisal value. The details of the implementation are given in Appendix B.6.5. Table 3 shows the results of the three conditional tests:

Table 3: Conditional Test Results

Eq. (18)	Eq. (19)	IPV
-1.2648	-1.3407	11.8961

As in the unconditional case, these show that, conditional on appraisal value, our data is consistent with conditionally independent private values which are independent of N, but not consistent with independent private values. These conclusions were again robust to moderate changes in our bandwidths, and in the support of the kernels used.

4.4 Application of the Identification Strategy

Next, we apply the identification results from Section 2 to the timber auction data. The main goal is to better understand the consequences of a strategic reserve policy when there is unobserved heterogeneity affecting demand. As discussed above, we condition only on appraisal value, which we believe accounts for *most* of the heterogeneity contained in observable covariates, and allow remaining observables to be picked up in θ.

In the counterfactuals we consider, we exploit the fact that the Forest Service holds timber auctions in two rounds. In the first round, bidders must submit bids of at least the reserve price (equal to the appraisal value) in order to qualify for the oral auction round of bidding. The oral auction then starts at the highest of the sealed bids. The first round of bidding thus reveals to both the seller and the other bidders the number of potential bidders in the ascending auction. The seller can thus use the number of potential bidders to design an alternative reserve price policy for the oral round of bidding, which would mean beginning the oral auction at some strategic level above the highest of the sealed bids.

The number of bidders who show up for the sealed bidding round is not observed in our data. However, as discussed above (footnote 26), it is widely known that the posted reserve price rarely if ever deters a potential bidder from participating in an auction; as a result, we can use the the number of bidders participating in the ascending auction (which is observed in the data) as a reasonable measure of the number of potential bidders in that auction.

Since both the eyeball test and formal econometric test above failed to reject the assumption of independence between valuations and auction size, we will present both the upper and lower bounds on seller's profit for the counterfactuals we consider in this section. If valuations are instead assumed to be stochastically increasing in N, the upper bounds $\overline{\pi}_n$ on expected profit are still valid, but the lower bounds $\underline{\pi}_n$ no longer apply.

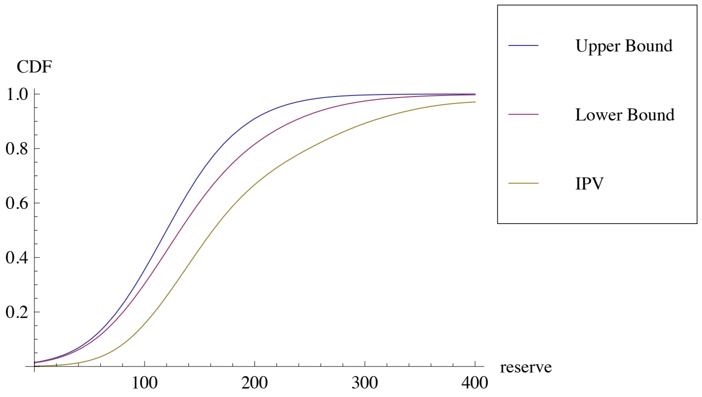
We begin by visually demonstrating the effect of correlation on the distribution $F_{n:n}$ of the highest valuation. As a representative example, we consider a timber tract with the median appraisal value in the data ($X = 55.9$) and $N = 3$ bidders. The first graph in Figure 3 shows three different ways to calculate $F_{n:n}$. "Upper bound" and "lower bound" refer to the bounds $\overline{F}_{n:n}$ and $\underline{F}_{n:n}$ defined in section 2. (These are calculated using nonparametric estimates of the distributions $F_{m-1:m}$ conditional on appraisal value.) "IPV" is the standard calculation of $F_{n:n}$ under the assumption that, conditional on appraisal value, bidder valuations are independent; this calculation is $F_{n:n}(v|X) = \left(\psi_{n-1:n}^{-1}\left(F_{n-1:n}(v|X)\right)\right)^n$.[29]

Examining Figure 3, we see that even though we do not achieve point-identification of $F_{3:3}$, our bounds are nonetheless informative. Even the lower bound $\underline{F}_{n:n}$ lies above the distribution implied by IPV, giving confirmation that there is indeed correlation among valuations in the data. This lower bound $\underline{F}_{n:n}$ remains valid when the assumption of independence between valuations and N is replaced with the weaker assumption of stochastic monotonicity; so the observation that $F_{3:3}$ lies above the IPV calculation still holds under our weaker identifying assumption. Also, relative to IPV, our upper and lower bounds are reasonably close together, so under independence between valuations and auction size, the true value of $F_{3:3}$ is fairly narrowly pinned down.

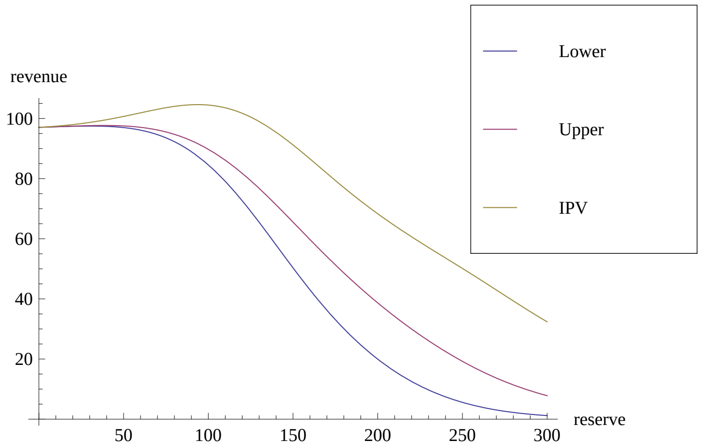
[29] In all three calculations of $F_{n:n}$, and in our calculations of bounds on π_n below, we smooth the nonparametric estimates of the conditional CDFs $F_{m-1:m}(\cdot|X)$ using the procedures suggested by Hansen (2004).

Figure 3: Analyzing a "Typical" Auction with $n = 3$ and $X = 56$

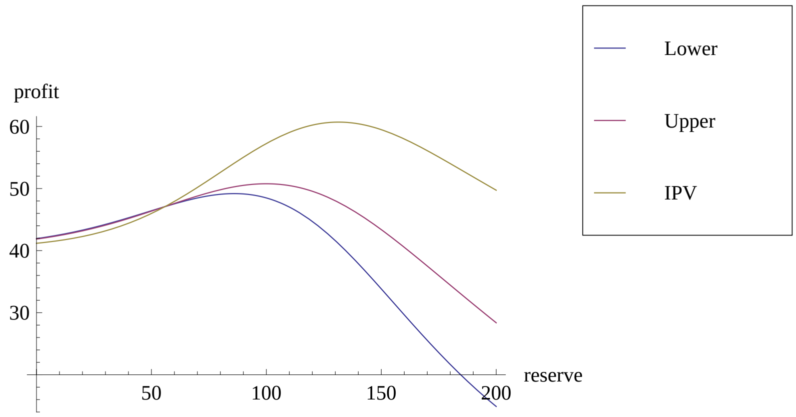
$F_{n:n}(v|X)$, calculated three ways



Expected Revenue



Expected Profit, with $v_0 = X = 56$



Next, we examine the effect that these different estimates of $F_{n:n}$ have on the seller's expected profit. Starting with non-parametric estimates of $F_{2:3}$, and then calculating bounds $\overline{F}_{3:3}$ and $\underline{F}_{3:3}$ (all conditional on the median appraisal value $X = 56$), Theorem 1 defines upper and lower bounds on the seller's expected profit at various reserve prices. The only missing piece is the seller's own valuation for the good, denoted v_0. We begin with the assumption that $v_0 = 0$, i.e., that the government has no private value for the unsold timber tract. In this case, expected profit to the seller is identical to expected revenue. The second graph in Figure 3 shows the bounds on expected revenue when we account for correlation among bidder valuations, alongside the expected revenue implied by the assumption that bidder valuations are independent (IPV).

As Figure 3 shows, the standard IPV model suggests there is a small but noticeable increase in revenue from using a reserve price around 100 (nearly double the appraisal value). Our bounds on expected revenue, accounting for unobserved heterogeneity, suggest there is no such increase, and that expected revenue begins to decrease in reserve price much sooner. As discussed above, a reserve price is only beneficial when just one bidder's value exceeds the reserve. When bidder values are positively correlated, then when $n - 1$ bidders have values below r, the last bidder is more likely to have a low value as well, as compared to the case where bidders are independent; so the benefit of a reserve price is smaller when bidder have correlated valuations as compared to IPV. These effects of correlation are exactly what is being captured in Figure 3, by the fact that both our upper and lower bounds on expected revenue lie below the IPV curve and peak at a lower reserve price.

As the seller's own valuation v_0 increases, the value to setting a reserve price increases (since the cost of not selling is reduced). We now make the more natural assumption that the value to the seller is equal to the appraisal value. (Since appraisal value is meant as a conservative estimate of the profitability of logging the land, this seems appropriate.) The third graph in Figure 3 shows expected profit under this assumption, calculated under IPV and under our bounds. We think of a "nonstrategic reserve price" as a reserve price equal to the seller's valuation, and a "strategic reserve price" as a reserve higher than that. Both standard IPV analysis and our analysis show that there are gains to using a strategic reserve; but the magnitude of the gains, as well as the reserve that maximizes them, are once again much smaller under our analysis than under IPV.

To better understand the implications of the the last graph in Figure 3, Table 4 shows, under each of the three estimates above, the reserve price r^* that maximizes expected profit; the ratio of that optimal reserve to the seller's valuation; the expected profit under a nonstrategic reserve price; the expected profit at the optimal reserve price; and the increase in profits that this represents. As discussed above, our analysis shows that there is indeed a gain to using a strategic reserve price, but that this gain is much smaller than the gain suggested by IPV. From Theorem 1, when valuations are

Table 4: Optimal Reserve and Effect on Profits, Calculated Three Ways

	r^*	$\frac{r^*}{v_0}$	$\pi_n(r=v_0)$	$\pi_n(r=r^*)$	$\%\Delta$
IPV	131.4	2.35	47.1	60.7	22.4%
Upper	99.9	1.79	47.1	50.8	7.2%
Lower	86.0	1.54	47.1	49.2	4.2%

independent of N, bounds on the optimal reserve price are given by $\overline{\pi}_n(r_n^*) \geq \max_{r'} \underline{\pi}_n(r') = 49.2$; this constrains r_n^* to lie within the interval $[73.7, 122.8]$. While this is a wide range, it is still informative, as the optimal reserve is sure to be below the value of 131.4 implied by IPV. Further, at the IPV-implied optimal reserve of 131.4, the upper and lower bound on expected profit accounting for unobserved heterogeneity are 47.8 and 41.1, respectively, implying anything from a negligible gain to a non-negligible loss relative to a non-strategic reserve. We emphasize once again that when valuations are stochastically increasing in N, the upper bounds on expected revenue and profit still hold. This means that the qualitative takeaway – that both the optimal reserve price and the increase in profits are substantially lower than IPV would suggest – still holds when our stronger identifying assumption fails.

While these results illustrate the effects of reserve on one particular auction, they are fairly representative. Maintaining the assumption that the seller's value v_0 in each auction is equal to the appraisal value, we can consider the aggregate effect of different reserve price policies applied to all the auctions in the data set. We consider two main counterfactual exercises. In the first exercise, we consider the policy of setting each auction's reserve price such that the probability of no sale (i.e., no bidder having a value above the reserve) is 15%. The 15% number comes from the mandate of the Forest Service to sell at least 85 percent of timber tracts. Relative to the reserve prices that appear optimal under IPV, this is a more conservative policy, implying a smaller markup over cost; thus, when we consider profit maximization based on IPV analysis, the 15% mandate is a binding constraint. As pointed out by Haile and Tamer (2003), the IPV model predicts that the government could nearly double the reserve price in most auctions without threatening the 15% no-sale benchmark; we find that when we condition only on appraisal value and raise each reserve to the level predicting a 15% chance of no-sale under IPV, the average ratio of reserve price to appraisal value is 3.1 (and the median is 2).

However, when unobserved heterogeneity is accounted for, the probability of no-sale at these new reserve prices is substantially higher, since correlation implies a less optimistic distribution $F_{n:n}$ (and the probability of no-sale at a reserve price r is simply $F_{n:n}(r)$). As Table 5 shows, averaging across the auctions in our data, the upper and lower bounds on the probability of no-sale are 36.7% and 24.2%, respectively. That is, at the reserve prices which IPV predicts give a 15% chance of no-sale, we calculate the actual chance of no-sale (averaged across auctions) to instead be between 24.2% and 36.7%. Table 5 also shows the impact this reserve price policy would have on expected profits π_n, again assuming that $v_0 = X$. Bids and valuations are all stated in dollars per unit of timber, but our data also contains the volume of timber in each tract, so we can calculate the gain or loss per auction in dollars as well. The major take-away message is that when we allow for bidders' valuations being correlated, the gains to profits from the 15 percent reserve price policy are at best about half what they would appear to be under IPV, and at worst significantly negative.

Table 5: Change in Expected Profit from 15 Percent "No Sale" Policy

	$\underline{\pi}_n$	$\overline{\pi}_n$	IPV
Average probability of no sale	36.7%	24.2%	15.0%
Average change in expected profit (relative to $r = v_0 = X$)	−3.9%	7.4%	16.0%
Average change in expected profit, dollars	−$9,269	$3,447	$7,896

While this "15 percent" policy is not one that was seriously considered by the USFS, there has been a long debate about the appropriate way to set reserves for timber auctions. In the 1990s, the USFS made an effort to move toward a market-based method of appraising timber, the Transaction Evidence Appraisal (TEA) method, which was a change from the "residual method" that was operational during the period of our data. Athey, Crampton, and Ingraham (2003) give more detail on the TEA method. The idea was that rather than relying on accounting information on cost and revenue, timber could be appraised using data on transaction prices from past sales of similar tracts. The actual appraisal value (and reserve price) would be the predicted value of the high bid in a given auction, discounted by some "rollback" factor to encourage competition. USFS policy requires rollback rates to be no more than 30% of the predicted transaction price; Athey, Crampton, and Ingraham (2003) report discount rates actually employed in various locations ranging from 5% to 30%, with the vast majority between 10% and 30%. We can now examine the sensibility of this policy. We mimic the TEA approach by regressing transaction price on appraisal value in our data, and then applying a discount rate to that predicted value. Table 6 shows the average effect (across the different auctions in our data set) of the change in expected profits from moving from a non-strategic reserve price to this transactions-based approach with various discount rates.

Table 6: Change in Expected Profit from TEA Policy at Various Discount ("Rollback") Rates

Discount	Average $\frac{r}{v_0}$	$\underline{\pi}_n$	$\overline{\pi}_n$	IPV	$\underline{\pi}_n$	$\overline{\pi}_n$	IPV
0%	2.50	−0.8%	7.3%	17.1%	−\$3,076	\$3,774	\$7,427
10%	2.25	1.7%	6.7%	13.7%	−\$898	\$3,386	\$5,940
20%	2.00	2.5%	5.3%	9.8%	\$252	\$2,654	\$4,274
30%	1.75	2.0%	3.4%	5.9%	\$567	\$1,737	\$2,645

As Table 6 shows, with a discount rate of 20% or more, the transactions-based approach unambiguously increases expected profits relative to a nonstrategic reserve policy, and mitigates the downside risk of the more aggressive "15% policy" discussed earlier. Even with a discount rate of 10%, the lowest discount rate commonly used, the "worst-case" reduction in expected profits is not that significant, and the "best-case" profits are substantially higher. Thus, the TEA policy is fairly robust: even allowing for unobserved heterogeneity in the data and the possibility of correlated values, the TEA policy, as typically implemented, appears to be profit-enhancing.

5 Conclusion

In this paper, we have introduced new empirical methods for identification and counterfactual analysis in ascending auctions, which are robust to correlation in values induced by unobserved heterogeneity. Applying them to timber data, we find that analysis accounting for unobserved heterogeneity rationalizes a much more cautious reserve price policy than traditional analysis, at least when appraisal value is taken to be a sufficient statistic for observable covariates. We have made use of two potential identifying assumptions: the stronger of these, independence between valuations and auction size, may not be valid for all applications, but it seems to fit well in the timber auction setting, and it lends itself to formal testing; and the qualitative conclusions of our analysis are the same under the weaker assumption that valuations are stochastically increasing in auction size.

More generally, we believe the methods presented above to be applicable in three types of situations:

1. Bidder valuations are correlated, but this correlation cannot be linked to any observable auction-specific covariates

2. Covariates known to shift bidders' valuations are missing from the data

3. Conditioning nonparametrically on all relevant covariates is unappealing due to sample size and the "curse of dimensionality"[30]

Depending on the inferential setting, any one of these issues could be a significant concern for the researcher. While the first two are self-explanatory and are likely to characterize many empirical applications, let us elaborate briefly on the third. Even if bidder values are indeed independent conditional on all relevant observables, use of the IPV framework constrains the analyst to be fully nonparametric and to condition on "everything," since any misspecification or omitted variable would lead to correlated residuals and invalidate the theoretical model. Our techniques, on the other hand, allow for more flexibility.

To illustrate this flexibility, consider the following avenue for future research. Suppose $N \perp (\mathbf{X}, \theta)$ and $V_i \overset{\text{iid}}{\sim} F(\cdot\,|\mathbf{X}, \theta)$; let $g : \text{supp}(\mathbf{X}) \to \Re^k$ be *any* function of $\mathbf{X}$ on which we might wish to condition our analysis.[31] Conditional on a realization of $g(\mathbf{X})$, V_i are still *i.i.d.* draws from $F(\cdot\,|\mathbf{X}, \theta)$, just with a different (posterior) distribution of $(\mathbf{X}, \theta)$ for each value of $g(\mathbf{X})$.[32] Regardless of the choice of g, our model (and identifying assumption) still hold, and we can get valid estimates of expected revenue and do valid counterfactuals from estimates $\widehat{F}_{n-1:n}(\cdot\,|g(\mathbf{X}))$ – even if g does not accurately model the effect of $\mathbf{X}$ on V_i in the data-generating process and is therefore "misspecified."[33]

Applied researchers often face a dilemma. Asymptotically, including as many observable covariates $\mathbf{X}$ as possible will help control unmodelled heterogeneity across auctions; but nonparametric methods with a high-dimensional $\mathbf{X}$ may yield estimators for $\widehat{F}_{n-1:n}(\cdot\,|\mathbf{X})$ that are imprecise in small, or even moderately large, samples. Such imprecision would permeate to the resulting estimates of $\overline{\pi}_n$ and $\underline{\pi}_n$ and, consequently, to any policy conclusions derived from them. To counteract these finite-sample issues, the researcher may want to condition only on a subset of elements in $\mathbf{X}$ or, more generally, on a lower-dimensional transformation $g(\mathbf{X})$. The methods and results in this paper allow the researcher to conduct robust estimation and inference whose validity does not hinge on whether the chosen function g entirely controls the unmodelled heterogeneity across auctions.

[30]This curse of dimensionality results from the number of *continuously distributed* elements in $\mathbf{X}$.

[31]If g is a constant, we are conditioning on nothing, and letting all of $(\mathbf{X}, \theta)$ be picked up as unobserved heterogeneity. If $k = \dim(\mathbf{X})$ and g is the identity map, we are conditioning on all of $\mathbf{X}$. If $k < \dim(\mathbf{X})$ and $g(\mathbf{X}) = (X_1, X_2, \ldots, X_k)$, then (as in this paper) we are conditioning on a subset of $\mathbf{X}$. If $k = 1$, g is a one-dimensional index constructed from $\mathbf{X}$. And so on. For any given choice of $g(\cdot)$, the methodology presented here allows us to test whether values are independent of N conditional on $g(\cdot)$.

[32]Specifically, if $\rho(\mathbf{X}, \theta)$ is the prior distribution of $(\mathbf{X}, \theta)$, then $\rho(\mathbf{X}, \theta | g(\mathbf{X}) = \overline{g}) = \frac{\rho(\mathbf{X}, \theta) \mathbb{1}\{g(\mathbf{X}) = \overline{g}\}}{\int_{\mathbf{X} \times \Theta} \rho(\mathbf{X}', \theta') \mathbb{1}\{g(\mathbf{X}') = \overline{g}\} d\mathbf{X}' d\theta'}$.

[33]Another alternative would be to model valuations parametrically, for example, as $V_i = \beta \mathbf{X} + \epsilon_i$. Even if the true effect of $\mathbf{X}$ on V_i is not linear, the residuals $\epsilon_i = V_i - \beta \mathbf{X}$ are still *i.i.d.* conditional on $(\mathbf{X}, \theta)$; so our techniques would still pin down the distribution of $\epsilon_{n:n}$ from the distributions of $\epsilon_{m-1:m}$, giving us the "average" expected revenue across all sets of covariates $\mathbf{X}$ with the same value of $\beta \mathbf{X}$.

A Appendix – Implications of Standard Entry Models

In this appendix, we consider two well-known models of auctions with endogenous entry. We give conditions under which both of them generate valuations which are stochastically increasing in N but not independent of N, and conditions under which both are guaranteed to lead to violations of the test (10). We then solve numerical examples of both, and show that the violations of (10) are substantial; and use one of these examples as the basis for Monte Carlo simulations of our test statistic.

In the main body of the paper, we allow θ to have arbitrary dimension and arbitrary effect on the distribution of valuations $F(\,\cdot\,|\theta)$. Throughout this appendix, however, we will restrict θ to be one-dimensional and assume $F(\,\cdot\,|\theta)$ is stochastically increasing in θ, that is, $\theta \subseteq \Re$, and $\theta > \theta'$ implies $F(\,\cdot\,|\theta) \succ_{FOSD} F(\,\cdot\,|\theta')$.

Models

The first model we consider is that of Levin and Smith (1994), modified to condition on θ. There are $\overline{n}$ identical potential bidders, with identical entry costs c. Bidders observe the realization of θ but not their valuations, and then decide simultaneously whether to enter. Those who enter pay their entry costs, learn their valuations, and participate in an ascending auction. (Since entry decisions are made prior to learning valuations, Levin and Smith interpret c as the cost of actually discovering one's valuation of the object.) Those who do not enter get payoff 0.

The Levin-Smith model has a unique symmetric equilibrium. Let $u_n(\theta)$ denote the expected payoff to each bidder from participating in an n-bidder auction given a realization of θ. We focus on the nondegenerate case where $u_1(\theta) > c > u_{\overline{n}}(\theta)$, so the symmetric equilibrium is in mixed strategies: for each realization of θ, potential bidders each enter with probability q_θ, where q_θ solves

$$c \;=\; \sum_{n=0}^{\overline{n}-1} \left[\binom{\overline{n}-1}{n} (q_\theta)^n (1-q_\theta)^{\overline{n}-1-n} \right] u_{n+1}(\theta)$$

(The term in square brackets is the probability that exactly n of a bidder's opponents enter, so the right-hand side is a bidder's expected payoff after entering.)

The second model we consider is that of Samuelson (1985). Bidders again have homogeneous entry costs c, but learn both θ and their valuation before deciding whether to pay c and participate in the auction. In this model, for each realized θ, potential bidders play a cutoff strategy, entering if and only if their valuation is at least $v^*(\theta)$, where $v^*(\theta)$ solves $c = (v-r)\,(F(v|\theta))^{\overline{n}-1}$.

If $\theta > \theta'$ implies $F(\,\cdot\,|\theta) \succ_{FOSD} F(\,\cdot\,|\theta')$, it is straightforward to show that $v^*(\theta)$ is increasing in θ and $F(v^*(\theta)|\theta)$ is decreasing in θ – for higher θ, bidders require a higher valuation to enter, but still enter with greater probability. Note that conditional on entering, bidders have valuations drawn from the truncated distribution $F(v|\theta, v > v^*(\theta)) = \frac{F(v|\theta) - F(v^*(\theta)|\theta)}{1 - F(v^*(\theta)|\theta)}$.

Results

First, we show conditions under which either entry model is consistent with our weaker identifying assumption:

Proposition 1 *Suppose $\theta \subseteq \Re$, and $\theta > \theta'$ implies $F(\,\cdot\,|\theta) \succ_{FOSD} F(\,\cdot\,|\theta')$.*

1. *If q_θ (the expected participation rate) is increasing in θ, the Levin-Smith entry model leads to valuations which are stochastically increasing in N*
2. *If the truncated distribution $F(v|\theta, v > v^*(\theta))$ is stochastically increasing in θ, the Samuelson entry model leads to valuations which are stochastically increasing in N*

Proof. We first show that in either model, the conditional distribution of θ is stochastically increasing in N, that is, $G(\,\cdot\,|n) \succ_{FOSD} G(\,\cdot\,|n')$ for $n > n'$. We show this for discrete θ; the proof is the same for continuous θ, with integrals replacing sums. For the Levin-Smith case, letting $P(\theta)$ denote the prior probability distribution of θ,

$$\begin{aligned} G(\tilde{\theta}|n) &= \frac{\sum_{\theta \leq \tilde{\theta}} P(\theta)_{\overline{n}}C_n q_\theta^n (1-q_\theta)^{\overline{n}-n}}{\sum_{\theta \in \Theta} P(\theta)_{\overline{n}}C_n q_\theta^n (1-q_\theta)^{\overline{n}-n}} \\ &= \frac{\sum_{\theta \leq \tilde{\theta}} P(\theta)_{\overline{n}}C_n q_\theta^n (1-q_\theta)^{\overline{n}-n}}{\sum_{\theta \leq \tilde{\theta}} P(\theta)_{\overline{n}}C_n q_\theta^n (1-q_\theta)^{\overline{n}-n} + \sum_{\theta > \tilde{\theta}} P(\theta)_{\overline{n}}C_n q_\theta^n (1-q_\theta)^{\overline{n}-n}} \\ &= R\left(\frac{\sum_{\theta > \tilde{\theta}} P(\theta) q_\theta^n (1-q_\theta)^{\overline{n}-n}}{\sum_{\theta \leq \tilde{\theta}} P(\theta) q_\theta^n (1-q_\theta)^{\overline{n}-n}}\right) \end{aligned}$$

where $R(x) = \frac{1}{1+x}$. As n increases, each term in the numerator gets multiplied by $\frac{q_\theta}{1-q_\theta} > \frac{q_{\tilde{\theta}}}{1-q_{\tilde{\theta}}}$, while each term in the denominator gets multiplied by $\frac{q_\theta}{1-q_\theta} \leq \frac{q_{\tilde{\theta}}}{1-q_{\tilde{\theta}}}$, so the argument of R increases; since R is decreasing, this means $G(\tilde{\theta}|n)$ is decreasing in n, meaning $G(\,\cdot\,|n) \succ_{FOSD} G(\,\cdot\,|n')$ for $n > n'$. For the Samuelson model, the same logic holds, but with $1 - F(v^*(\theta)|\theta)$ (each potential bidder's probability of entry) replacing q_θ.

For the Levin-Smith case, since $F(\,\cdot\,|\theta)$ is stochastically increasing in θ,

$$F_{n:n}^{n+1}(v) \;=\; \int_{\theta \in \Theta} (F(v|\theta))^n \, dG(\theta|n+1) \;\leq\; \int_{\theta \in \Theta} (F(v|\theta))^n \, dG(\theta|n) \;=\; F_{n:n}(v)$$

because $G(\,\cdot\,|n+1) \succ_{FOSD} G(\,\cdot\,|n)$ and $(F(v|\theta))^n$ is decreasing in θ. Under the Samuelson model, the same logic holds, but with the truncated distribution $F(v|\theta, v > v^*(\theta))$ replacing $F(v|\theta)$. □

(Note that q_θ increasing in θ is by no means guaranteed in the Levin-Smith model. If, for example, higher θ were associated with higher average valuations but lower variance, $u_n(\theta)$ could be decreasing in θ for $n > 1$, in which case q_θ could be decreasing for some levels of entry costs. A sufficient condition for q_θ to be increasing would be $u_n(\theta)$ increasing in θ for each n, which would tend to hold, for example, if higher θ implied both higher mean and higher variance.)

Next, we show that under similar conditions, either entry model will generate a violation of (10), meaning our test of independence between valuations and N will have power against the dependence introduced by either of these models:

Proposition 2 *Suppose $\theta \subseteq \Re$, and $\theta > \theta'$ implies $F(\cdot|\theta) \succ_{FOSD} F(\cdot|\theta')$. Suppose also that $F(\cdot|\theta)$ is continuous and twice differentiable, with derivative $f(\cdot|\theta)$, and has the same bounded support $[\underline{v}, \overline{v}]$ for all θ.*

1. *If q_θ and $f(\overline{v}|\theta)$ are both increasing in θ, the Levin-Smith entry model leads to a violation of* (10).

2. *If $f(\overline{v}|\theta, v > v^*(\theta)) = \frac{f(\overline{v}|\theta)}{1-F(v^*(\theta)|\theta)}$ is increasing in θ, the Samuelson entry model leads to a violation of* (10).

Proof. We've already showed that under either model, $G(\cdot|n)$ is stochastically increasing in n. If $f(\overline{v}|\theta)$ is increasing in θ, then, $E_{\theta|n}\left(f(\overline{v}|\theta)\right)^2$ is increasing in n, so the Levin-Smith model generates a violation by Theorem 3. The same logic holds for the Samuelson model, using the density conditional on entry in place of the unconditional density $f(\overline{v}|\theta)$. □

Numerical Examples

Next, we give numerical examples of these two entry models. Let θ take two values, H and L, with equal probabilities, and $F(\cdot|\theta)$ be log-normal distributions: when $\theta = H$, $\ln(V_i) \sim N(2.5, 0.5)$, and when $\theta = L$, $\ln(V_i) \sim N(2.0, 0.5)$. V_i has mean 13.8 and median 12.2 when $\theta = H$, and mean 8.4 and median 7.4 when $\theta = L$.

Consider the Levin-Smith entry game, with 12 potential bidders and entry costs of \$1.50. This causes bidders to each enter with probability $q_H = 0.3884$ when $\theta = H$, and probability $q_L = 0.2597$ when $\theta = L$. Given q_θ, we can then calculate the distribution of N, conditional on θ:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
$\Pr(N = n \mid \theta = H)$	0.3%	2.1%	7.3%	15.4%	22.1%	22.4%	16.6%	9.0%	3.6%	1.0%	0.2%	0.0%	0.0%
$\Pr(N = n \mid \theta = L)$	2.7%	11.4%	22.0%	25.7%	20.3%	11.4%	4.7%	1.4%	0.3%	0.0%	0.0%	0.0%	0.0%

And likewise, assuming the prior probability of $\theta = H$ is $\frac{1}{2}$, the conditional probability distribution of θ given a value of N:

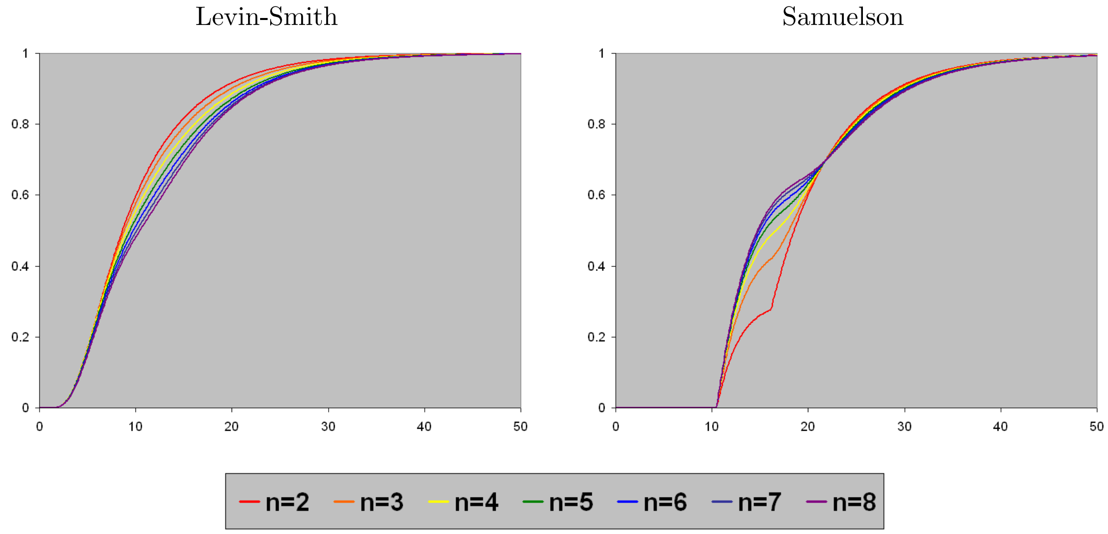
n	0	1	2	3	4	5	6	7	8	9	10	11	12
$\Pr(\theta = H \mid N = n)$	9%	15%	25%	37%	52%	66%	78%	87%	92%	95%	97%	99%	99%

Next, consider the Samuelson entry game, with 8 potential bidders and the same entry costs of \$1.50. This leads to cutoff values $v^*(H) = \$16.12$ and $v^*(L) = \$10.48$; each bidder enters with probability $1 - F(v^*(H)|H) = 0.288$ when $\theta = H$, and probability $1 - F(v^*(L)|L) = 0.242$ when $\theta = L$. Again, we can calculate the distribution of N conditional on θ, and the distribution of θ conditional on N:

n	0	1	2	3	4	5	6	7	8
$\Pr(N = n \mid \theta = H)$	6.6%	21.4%	30.3%	24.5%	12.3%	4.0%	0.8%	0.1%	0.0%
$\Pr(N = n \mid \theta = L)$	10.8%	27.8%	31.1%	19.9%	8.0%	2.0%	0.3%	0.0%	0.0%
$\Pr(\theta = H \mid N = n)$	38%	44%	49%	55%	61%	66%	71%	76%	80%

Figure 4 shows $\psi^{-1}_{n-1:n}(F_{n-1:n}(v))$ for $n = 2, 3, 4, 5, 6, 7, 8$. If $\psi^{-1}_{n-1:n}(F_{n-1:n}(v))$ is increasing in n everywhere, we would fail to reject (10). If it is decreasing in n at any v, we would reject independence of valuations and N given enough data. Both of these entry models would clearly lead to rejection given enough data.

Figure 4: $\psi^{-1}_{n-1:n}(F_{n-1:n}(v))$ under two entry models. (Recall that we reject independence if $\psi^{-1}_{n-1:n}(F_{n-1:n}(v))$ is decreasing in n for some v.)



Monte Carlo Simulations

For three different data-generating processes (described below), we generated 1000 data sets of sizes $L = 200$, $L = 400$ and $L = 800$ and calculated the test-statistic of (10) for each simulated data set. In every instance we constructed the statistic exactly as described in the unconditional-test results for the timber data application (see footnote 28). We used $b_L = 0.001$ for $L = 800$, $b_L = 0.0012$ for $L = 400$ and $b_L = 0.0015$ for $L = 200$. These satisfy the convergence rate conditions[34] of b_L in Theorem 5. The data-generating processes were:

1. The above example of the entry game from Levin and Smith (1994), discarding observations with $N < 2$

2. Valuations independent of N (A) – the distributions of θ and $V_i|\theta$ were the same as in case 1, but N was chosen independently of θ, to match the distribution of N in case 1. (So for example, in case 1, $\Pr(N = 5|N \geq 2) = 22.9\%$ when $\theta = H$ and 13.3% when $\theta = L$; so in this case, each observation had probability $\frac{22.9\%+13.3\%}{2} = 18.1\%$ of having $N = 5$, independent of θ.) Under this data-generating process, (10) holds as a strict inequality with probability 1.

3. Valuations independent of N (B) – the distributions of θ and N were the same as in the first two cases, but the distributions $F(\cdot|\theta)$ were as follows:

$$F(v|L) = \Phi\left(\frac{\ln(v) - 2}{3/4}\right) \qquad and \qquad F(v|H) = \begin{cases} \Phi\left(\frac{\ln(v)-2.5}{1/3}\right) & if \quad \ln(v) < 2.9 \\ \Phi\left(\frac{\ln(v)-2}{3/4}\right) & if \quad \ln(v) \geq 2.9 \end{cases}$$

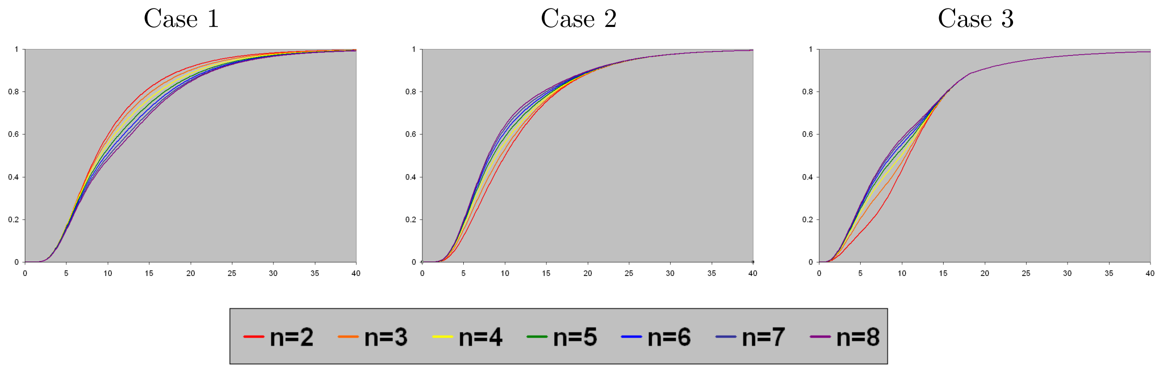
   where Φ is the CDF of the standard normal distribution. This led to (10) holding with equality at $v \geq \exp(2.9)$, which occurs for about 11.5% of valuations, and therefore about 8% of transaction prices in our simulated data.

Figure 5 shows graphs of $\psi_{n-1:n}^{-1}(F_{n-1:n}(v))$ for various values of n for each of these three data-generating processes. The three examples represent the three relevant asymptotic scenarios: (10) is violated in case 1, holds strictly everywhere in case 2, and holds everywhere but sometimes with equality in case 3.

Our rejection rule used a 5% target significance level. The results of the simulations are given in the text, in section 3.5.

[34]Using the same bandwidth b_L for the three sample sizes would also have been a valid and meaningful exercise, since our focus here is on the finite-sample properties of our test.

Figure 5: $\psi_{n-1:n}^{-1}\left(F_{n-1:n}(v)\right)$ under the three scenarios used in Monte Carlo simulations



B Appendix – Omitted Proofs

B.1 Proof of Lemma 1

By assumption, the transaction price is $\max\{r, V_{n-1:n}\}$, provided $V_{n:n} \geq r$. We can therefore write expected profits as

$$\pi_n(r) \quad = \quad E_{V_{n-1:n},V_{n:n}}\left\{\mathbf{1}_{V_{n:n}\geq r\geq V_{n-1:n}}\left(r-v_0\right)+\mathbf{1}_{V_{n:n}\geq V_{n-1:n}>r}\left(V_{n-1:n}-v_0\right)\right\}$$

Since the first event happens with probability $F_{n-1:n}(r)-F_{n:n}(r)$, and the second happens whenever $V_{n-1:n} > r$, we can rewrite this as

$$\pi_n(r) \quad = \quad \left(F_{n-1:n}(r)-F_{n:n}(r)\right)\left(r-v_0\right)+\int_r^{+\infty}(v-v_0)dF_{n-1:n}(v)$$

As for bidder surplus, each bidder has ex-ante probability $\frac{1}{n}$ of having the highest value, which earns a surplus of 0 when $V_{n:n} \leq r$ and $V_{n:n} - \max\{V_{n-1:n}, r\}$ when $V_{n:n} > r$ (where $V_{0:1}$ is understood to be 0). So

$$\begin{aligned}
u_n(r) &= \tfrac{1}{n}E_{V_{n-1:n},V_{n:n}}\left\{\mathbf{1}_{V_{n:n}>r}\left(V_{n:n}-\max\{V_{n-1:n},r\}\right)\right\} \\
&= \tfrac{1}{n}E\left\{\mathbf{1}_{r\geq V_{n:n}\geq V_{n-1:n}}\left(r-r\right)+\mathbf{1}_{V_{n:n}>r\geq V_{n-1:n}}\left(V_{n:n}-r\right)+\mathbf{1}_{V_{n:n}\geq V_{n-1:n}>r}\left(V_{n:n}-V_{n-1:n}\right)\right\} \\
&= \tfrac{1}{n}E\left\{\mathbf{1}_{V_{n:n}\leq r}r+\mathbf{1}_{V_{n:n}>r}V_{n:n}-\mathbf{1}_{V_{n-1:n}\leq r}r-\mathbf{1}_{V_{n-1:n}>r}V_{n-1:n}\right\} \\
&= \tfrac{1}{n}\left(E\left\{\max\{V_{n:n},r\}\right\}-E\left\{\max\{V_{n-1:n},r\}\right\}\right)
\end{aligned}$$

B.2 Proof of Lemmas 3 and 4

We begin with Lemma 4. As noted in the text, valuations independent of N implies $F_{n:n}(v) = \frac{1}{n+1}F_{n:n+1}(v) + \frac{n}{n+1}F_{n+1:n+1}(v)$. Fixing n, we will use induction on $\bar{n}$ to show (7),

$$F_{n:n}(v) \quad = \quad \frac{1}{n-1}\sum_{m=n+1}^{\bar{n}}\left(\prod_{i=n}^{m-1}\frac{i-1}{i+1}\right)F_{m-1:m}(v) \quad + \quad \frac{n}{\bar{n}}F_{\bar{n}:\bar{n}}(v)$$

for any $\bar{n} > n$. For the base case, $\bar{n} = n+1$, the right-hand side is $\frac{1}{n+1}F_{n:n+1}(v) + \frac{n}{n+1}F_{n+1:n+1}(v)$, which as noted is equal to $F_{n:n}(v)$. For the inductive step, if (7) holds for $\bar{n} = K$, then

$$\begin{aligned}
&\tfrac{1}{n-1}\textstyle\sum_{m=n+1}^{K+1}\left(\prod_{i=n}^{m-1}\frac{i-1}{i+1}\right)F_{m-1:m}(v) + \frac{n}{K+1}F_{K+1:K+1}(v) \\
&\quad = \tfrac{1}{n-1}\textstyle\sum_{m=n+1}^{K}\left(\prod_{i=n}^{m-1}\frac{i-1}{i+1}\right)F_{m-1:m}(v) + \frac{1}{n-1}\left(\prod_{i=n}^{K}\frac{i-1}{i+1}\right)F_{K:K+1}(v) + \frac{n}{K+1}F_{K+1:K+1}(v) \\
&\quad = F_{n:n}(v) - \tfrac{n}{K}F_{K:K}(v) + \tfrac{1}{n-1}\tfrac{(n-1)n}{K(K+1)}F_{K:K+1}(v) + \tfrac{n}{K+1}F_{K+1:K+1}(v) \\
&\quad = F_{n:n}(v) + \tfrac{n}{K}\left(-F_{K:K}(v) + \tfrac{1}{K+1}F_{K:K+1}(v) + \tfrac{K}{K+1}F_{K+1:K+1}(v)\right)
\end{aligned}$$

which is again equal to $F_{n:n}(v)$ since by (5), the terms in parentheses sum to 0; so (7) holds for $\bar{n} = K+1$. To show (6) holds when valuations are stochastically increasing in N, the steps are identical, only starting with the weak inequality $F_{n:n}(v) \geq \frac{1}{n+1}F_{n:n+1}(v) + \frac{n}{n+1}F_{n+1:n+1}(v)$.

B.3 Proof of Theorem 3

The proof is basically the Taylor expansion of $\psi_{n-1:n}^{-1} \circ F_{n-1:n}$ around $\overline{v}$, which gives

$$\psi_{n-1:n}^{-1}\left(F_{n-1:n}(v)\right) \quad = \quad 1 - (\overline{v} - v)\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2} + O\left((\overline{v} - v)^2\right) \tag{21}$$

Calculation is complicated by the fact that $\left(\psi_{n-1:n}^{-1}\right)'$ is infinite at 1. $\psi_{n-1:n}^{-1} \circ F_{n-1:n}$ is still differentiable at $\overline{v}$, because $F_{n-1:n}$ is very flat near $\overline{v}$; we just need to be very careful in how we calculate things.

Differentiating gives

$$\begin{aligned}
\left(\psi_{n-1:n}^{-1} \circ F_{n-1:n}\right)'(v) \quad &= \quad \left(\psi_{n-1:n}^{-1}\right)'\left(F_{n-1:n}(v)\right) \times F_{n-1:n}'(v) \\
&= \quad \tfrac{1}{\psi_{n-1:n}'\left(\psi_{n-1:n}^{-1}(F_{n-1:n}(v))\right)} \times F_{n-1:n}'(v) \\
&= \quad \tfrac{1}{\psi_{n-1:n}'\left(\psi_{n-1:n}^{-1}\left(E_{\theta|n}F_{n-1:n}(v|\theta)\right)\right)} \times \left(E_{\theta|n}F_{n-1:n}(v|\theta)\right)' \\
&= \quad \tfrac{1}{\psi_{n-1:n}'\left(\psi_{n-1:n}^{-1}\left(E_{\theta|n}\psi_{n-1:n}(F(v|\theta))\right)\right)} \times \left(E_{\theta|n}\psi_{n-1:n}\left(F(v|\theta)\right)\right)' \\
&= \quad \tfrac{1}{\psi_{n-1:n}'\left(\psi_{n-1:n}^{-1}\left(E_{\theta|n}\psi_{n-1:n}(F(v|\theta))\right)\right)} \times E_{\theta|n}\psi_{n-1:n}'\left(F(v|\theta)\right)f(v|\theta)
\end{aligned}$$

so letting $\epsilon \equiv \overline{v} - v$,

$$\left(\psi_{n-1:n}^{-1} \circ F_{n-1:n}\right)'(\overline{v} - \epsilon) \quad = \quad \frac{E_{\theta|n}\{\psi_{n-1:n}'\left(F(\overline{v} - \epsilon|\theta)\right)f(\overline{v} - \epsilon|\theta)\}}{\psi_{n-1:n}'\left(\psi_{n-1:n}^{-1}\left(E_{\theta|n}\{\psi_{n-1:n}\left(F(\overline{v} - \epsilon|\theta)\right)\}\right)\right)}$$

To calculate the denominator, we first fix θ, and suppress the dependence of $F(\cdot|\theta)$ and $f(\cdot|\theta)$ on θ. The Taylor expansion of $F(v|\theta)$ around $\overline{v}$ gives $F(\overline{v}-\epsilon) = 1 - f(\overline{v})\epsilon + \frac{1}{2}f'(\overline{v})\epsilon^2 + O(\epsilon^3)$; $\psi_{n-1:n}(s)$ can be written as $ns^{n-1}(1-s)+s^n$, so

$$\begin{aligned}
\psi_{n-1:n}(F(\overline{v}-\epsilon)) &= n\left(1 - f(\overline{v})\epsilon + \tfrac{1}{2}f'(\overline{v})\epsilon^2 + O(\epsilon^3)\right)^{n-1}\left(f(\overline{v})\epsilon - \tfrac{1}{2}f'(\overline{v})\epsilon^2 + O(\epsilon^3)\right) \\
&\quad + \left(1 - f(\overline{v})\epsilon + \tfrac{1}{2}f'(\overline{v})\epsilon^2 + O(\epsilon^3)\right)^n \\
&= nf(\overline{v})\epsilon - \tfrac{n}{2}f'(\overline{v})\epsilon^2 - n(n-1)(f(\overline{v}))^2\epsilon^2 \\
&\quad +1 - nf(\overline{v})\epsilon + \tfrac{n}{2}f'(\overline{v})\epsilon^2 + \tfrac{n(n-1)}{2}(f(\overline{v}))^2\epsilon^2 + O(\epsilon^3) \\
&= 1 - \tfrac{n(n-1)}{2}(f(\overline{v}))^2\epsilon^2 + O(\epsilon^3) \\
E_{\theta|n}\left\{\psi_{n-1:n}\left(F(\overline{v}-\epsilon)\right)\right\} &= 1 - \tfrac{n(n-1)}{2}\epsilon^2 E_{\theta|n}\{(f(\overline{v}|\theta))^2\} + O(\epsilon^3)
\end{aligned}$$

Next, we calculate $y = \psi_{n-1:n}^{-1}\left(E_{\theta|n}\{\psi_{n-1:n}\left(F(\overline{v}-\epsilon)\right)\}\right)$. For ϵ small, the argument of $\psi_{n-1:n}^{-1}$ will be close to 1, so y will be close to 1; let $x = 1-y$, and solve

$$\begin{aligned}
\psi_{n-1:n}(1-x) &= 1 - \tfrac{n(n-1)}{2}\epsilon^2 E_{\theta|n}\{(f(\overline{v}|\theta))^2\} + O(\epsilon^3) \\
n(1-x)^{n-1}x + (1-x)^n &= 1 - \tfrac{n(n-1)}{2}\epsilon^2 E_{\theta|n}\{(f(\overline{v}|\theta))^2\} + O(\epsilon^3) \\
nx - n(n-1)x^2 + 1 - nx + \tfrac{n(n-1)}{2}x^2 + O(x^3) &= 1 - \tfrac{n(n-1)}{2}\epsilon^2 E_{\theta|n}\{(f(\overline{v}|\theta))^2\} + O(\epsilon^3) \\
1 - \tfrac{n(n-1)}{2}x^2 + O(x^3) &= 1 - \tfrac{n(n-1)}{2}\epsilon^2 E_{\theta|n}\{(f(\overline{v}|\theta))^2\} + O(\epsilon^3) \\
x^2 + O(x^3) &= \epsilon^2 E_{\theta|n}\{(f(\overline{v}|\theta))^2\} + O(\epsilon^3) \\
x &= \epsilon\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2} + O(\epsilon^2)
\end{aligned}$$

(For the last equality, we're solving $x^2 + O(x^3) = z^2 + O(z^3)$. If $x = z + \beta z^\alpha + \ldots$ with $\alpha < 2$, then the leading term of $x^2 - z^2$ would be of order $z^{1+\alpha}$, giving a contradiction.)

Finally, taking $\psi'_{n-1:n}(1-x)$ to complete the denominator gives

$$\begin{aligned}
\psi'_{n-1:n}(s) &= n(n-1)s^{n-2}(1-s) \\
\psi'_{n-1:n}(1-x) &= n(n-1)(1-x)^{n-2}x \\
&= n(n-1)(1-\epsilon\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2} + O(\epsilon^2))^{n-2}\left(\epsilon\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2} + O(\epsilon^2)\right) \\
&= n(n-1)\epsilon\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2} + O(\epsilon^2)
\end{aligned}$$

and so the denominator of $\left(\psi_{n-1:n}^{-1} \circ F_{n-1:n}\right)'(\overline{v}-\epsilon)$ is

$$\psi'_{n-1:n}\left(\psi_{n-1:n}^{-1}\left(E_{\theta|n}\{\psi_{n-1:n}\left(F(\overline{v}-\epsilon|\theta)\right)\}\right)\right) = n(n-1)\epsilon\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2} + O(\epsilon^2)$$

To calculate the numerator, $E_{\theta|n}\left\{\psi'_{n-1:n}\left(F(\overline{v}-\epsilon)\right)f(\overline{v}-\epsilon)\right\}$, we again fix θ and suppress the dependence of F and f on θ. $F(\overline{v}-\epsilon)=1-f(\overline{v})\epsilon+O(\epsilon^2)$ and $\psi'_{n-1:n}(s)=n(n-1)s^{n-2}(1-s)$, so

$$\begin{aligned}
\psi'_{n-1:n}(F(\overline{v}-\epsilon)) &= n(n-1)\left(1-f(\overline{v})\epsilon+O(\epsilon^2)\right)^{n-2}\left(f(\overline{v})\epsilon+O(\epsilon^2)\right) \\
&= n(n-1)f(\overline{v})\epsilon+O(\epsilon^2) \\
\psi'_{n-1:n}(F(\overline{v}-\epsilon))f(\overline{v}-\epsilon) &= \left(n(n-1)f(\overline{v})\epsilon+O(\epsilon^2)\right)\left(f(\overline{v})-f'(\overline{v})\epsilon+O(\epsilon^2)\right) \\
&= n(n-1)(f(\overline{v}))^2\epsilon+O(\epsilon^2) \\
E_{\theta|n}\left\{\psi'_{n-1:n}(F(\overline{v}-\epsilon|\theta))f(\overline{v}-\epsilon|\theta)\right\} &= n(n-1)\epsilon E_{\theta|n}\{(f(\overline{v}|\theta))^2\}+O(\epsilon^2)
\end{aligned}$$

Putting it together, then,

$$\begin{aligned}
\left(\psi^{-1}_{n-1:n}\circ F_{n-1:n}\right)'(\overline{v}-\epsilon) &= \frac{E_{\theta|n}\left\{\psi'_{n-1:n}\left(F(\overline{v}-\epsilon|\theta)\right)f(\overline{v}-\epsilon|\theta)\right\}}{\psi'_{n-1:n}\left(\psi^{-1}_{n-1:n}\left(E_{\theta|n}\{\psi_{n-1:n}\left(F(\overline{v}-\epsilon|\theta)\right)\}\right)\right)} \\
&= \frac{n(n-1)\epsilon E_{\theta|n}\{(f(\overline{v}|\theta))^2\}+O(\epsilon^2)}{n(n-1)\epsilon\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2}+O(\epsilon^2)} \\
&= \frac{E_{\theta|n}(f(\overline{v}|\theta))^2}{\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2}}+O(\epsilon) \\
&= \sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2}+O(\epsilon)
\end{aligned}$$

so

$$\begin{aligned}
\psi^{-1}_{n-1:n}\left(F_{n-1:n}(v)\right) &= \psi^{-1}_{n-1:n}\left(F_{n-1:n}(\overline{v})\right)-\int_v^{\overline{v}}\left(\psi^{-1}_{n-1:n}\circ F_{n-1:n}(s)\right)'(s)ds \\
&= 1-\int_v^{\overline{v}}\left[\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2}+O\left(\overline{v}-v\right)\right]ds \\
&= 1-(\overline{v}-v)\sqrt{E_{\theta|n}(f(\overline{v}|\theta))^2}+O\left((\overline{v}-v)^2\right)
\end{aligned}$$

Thus, if $E_{\theta|n}\left(f(\overline{v}|\theta)\right)^2>E_{\theta|n'}\left(f(\overline{v}|\theta)\right)^2$, then $\psi^{-1}_{n-1:n}\left(F_{n-1:n}(v)\right)<\psi^{-1}_{n'-1:n'}\left(F_{n'-1:n'}(v)\right)$ for v close to $\overline{v}$, a violation of (10) if $n>n'$. □

B.4 Asymptotic Properties of the Unconditional Test Statistics in (14)

Here we describe and discuss the main asymptotic results and the corresponding assumptions for the statistics described in (14). Detailed proofs of all our results can be found in the Technical Supplement.

B.4.1 Assumptions

We begin by formalizing the basic distribution assumptions, all of which are compatible with the model presented in Section 2.

Assumption T1

We observe an i.i.d. sample $(W_i,N_i)_{i=1}^L\equiv(Z_i)_{i=1}^L$ of winning bids (transaction prices) W and auction sizes N. The distribution $F_{W|N}(w|n)$ is continuous in w for each $n\in\mathcal{S}_N$. The set $\mathcal{N}\subseteq\mathcal{S}_N$

is compact and $\min_{n\in\mathscr{N}} p_N(n) \equiv \underline{p}_N > 0$. *For each* n, n' *in* $\mathscr{N}$, *the set* $\mathscr{W}_{n,n'}$ *is compact and both* $F_{W|N}(\cdot|n)$ *and* $F_{W|N}(\cdot|n')$ *are strictly bounded away from* 0 *and* 1 *everywhere in* $\mathscr{W}_{n,n'}$.

The reason for restricting the test to a region $\mathscr{W}_{n,n'}$ on which $F_{W|N}$ is bounded away from 0 and 1 is as follows. Define

$$\nabla_1 \Omega(s,n,n') = \frac{\partial}{\partial s}\Omega(s,n,n')$$

Following the logic from the proof of Theorem 2, $\nabla_1\Omega(s,n,n') = \frac{n(n-1)t^{n-1}(1-t)}{n'(n'-1)t^{n'-2}(1-t)}$, where $t = \psi^{-1}_{n'-1:n'}(s)$. If $s \in (0,1)$, then $t \in (0,1)$, and so $\nabla_1\Omega(s,n,n')$ is well-defined, non-zero, and finite; so under Assumption T1, $\nabla_1\Omega\big(F_{W|N}(\cdot|n'),n,n'\big)$ is well-defined and bounded everywhere in $\mathscr{W}_{n,n'}$.

Take any (w,n,n'), where $n,n' \in \mathcal{S}_N^2$. We will define

$$\begin{aligned}
\varphi^F(Z,w,n) &= \frac{\mathbb{1}\{W \le w\} - F_{W|N}(w|n)}{p_N(n)} \cdot \mathbb{1}\{N=n\},\\
\varphi^\Delta(Z,w,n,n') &= \varphi^F(Z,w,n) - \varphi^F(Z,w,n'),\\
\varphi^\Omega(Z,w,n,n') &= \nabla_1\Omega\big(F_{W|N}(w|n'),n,n'\big)\cdot\varphi^F(Z,w,n'),\\
\varphi^\Phi(Z,w,n,n') &= \varphi^\Omega(Z,w,n,n') - \varphi^F(Z,w,n).
\end{aligned} \tag{22}$$

Note that for any $n \in \mathcal{S}_N$ and any w, we have $E\big[\varphi^F(Z,w,n)\big] = 0$ and therefore $E\big[\varphi^\Delta(Z,w,n,n')\big] = E\big[\varphi^\Omega(Z,w,n,n')\big] = E\big[\varphi^\Phi(Z,w,n,n')\big] = 0$. Take any $i \neq j$. From the conditions of Assumption (T1), the linear representations for our nonparametric estimators are given by

$$\begin{aligned}
\widehat{F}^{-i,j}_{W|N}(w|n) &= F_{W|N}(w|n) + \frac{1}{L-2}\sum_{k\neq i,j}\varphi^F(Z_k,w,n) + \varsigma^{-i,j}_L(w,n),\\
\widehat{\Delta}^{-i,j}_{W|N}(w,n,n') &= \Delta_{W|N}(w,n,n') + \frac{1}{L-2}\sum_{k\neq i,j}\varphi^\Delta(Z_k,w,n,n') + \xi^{-i,j}_L(w,n,n'),
\end{aligned} \tag{23A}$$

and

$$\begin{aligned}
\Omega\big(\widehat{F}^{-i,j}_{W|N}(w|n'),n,n'\big) &= \Omega\big(F_{W|N}(w|n'),n,n'\big) + \frac{1}{L-2}\sum_{k\neq i,j}\varphi^\Omega(Z_k,w,n') + \widetilde{\varsigma}^{-i,j}_L(w,n,n'),\\
\widehat{\Phi}^{-i,j}_{W|N}(w,n,n') &= \Phi_{W|N}(w,n,n') + \frac{1}{L-2}\sum_{k\neq i,j}\varphi^\Phi(Z_k,w,n,n') + \widetilde{\xi}^{-i,j}_L(w,n,n').
\end{aligned} \tag{23B}$$

Using results from empirical process theory (see, e.g., Pakes and Pollard (1989), Andrews (1994) or Part 2 in van der Vaart and Wellner (1996)), compactness and the remaining features of $\mathscr{N}$ and $\mathscr{W}_{n,n'}$ described in Assumption (T1), and the classes of functions involved yield

$$\begin{aligned}
&\sup_{\substack{w\in\mathscr{W}_{n,n'}\\(n,n')\in\mathscr{N}}}\big|\varsigma^{-i,j}_L(w,n,n')\big| = O_p\Big(\frac{1}{L}\Big), \quad \sup_{\substack{w\in\mathscr{W}_{n,n'}\\(n,n')\in\mathscr{N}}}\big|\widetilde{\varsigma}^{-i,j}_L(w,n,n')\big| = O_p\Big(\frac{1}{L}\Big),\\
&\sup_{\substack{w\in\mathscr{W}_{n,n'}\\(n,n')\in\mathscr{N}}}\big|\xi^{-i,j}_L(w,n,n')\big| = O_p\Big(\frac{1}{L}\Big), \quad \sup_{\substack{w\in\mathscr{W}_{n,n'}\\(n,n')\in\mathscr{N}}}\big|\widetilde{\xi}^{-i,j}_L(w,n,n')\big| = O_p\Big(\frac{1}{L}\Big).
\end{aligned} \tag{24}$$

Thus, $\xi^{-i,j}_L$ and $\widetilde{\xi}^{-i,j}_L$ are to be interpreted as the remainder terms of the linear asymptotic representation of the nonparametric estimators $\widehat{\Delta}^{-i,j}_{W|N}$ and $\widehat{\Phi}^{-i,j}_{W|N}$ respectively. From the conditions in Assumption (T1), these terms vanish in probability at a rate faster than $L^{-1/2}$. Additional properties of these remainders will be important throughout the proof of our main result. In order to rule out irregular cases, we will assume that the underlying data-generating process is such that the following additional conditions are satisfied. First, let $\widehat{\Delta}^{-i,j,k}_{W|N}(w,n,n')$ and $\widehat{\Phi}^{-i,j,k}_{W|N}(w,n,n')$ denote nonparametric estimators of $\Delta_{W|N}(w,n,n')$ and $\Phi_{W|N}(w,n,n')$ obtained after dropping the i,j,k observations in the sample. Their corresponding linear representations are straightforward generalizations of (23A) and (23B). The corresponding remainder terms would now be denoted by $\xi^{-i,j,k}_{L}(w,n,n')$ and $\widetilde{\xi}^{-i,j,k}_{L}(w,n,n')$. We assume the following.

Assumption T2

(i) Take $i \neq j$ and let ℓ denote either i or j. There exists a $\tau > 0$ and a deterministic sequence $\overline{H}_L = O(1)$ such that, for any $n, n' \in \mathscr{N}$,

$$\Pr\Big(-s \leq \Delta_{W|N}(W_\ell,n,n') < 0 \;\Big|\; \xi^{-i,j}_{L}(W_\ell,n,n'), W_\ell \in \mathscr{W}_{n,n'}\Big) \leq \overline{H}_L \cdot |s| \quad \forall\, 0 < s \leq \tau.$$

(ii) Take any n,n' in $\mathscr{N}$ and any $w \in \mathscr{W}_{n,n'}$. For $k \neq i,j$ let

$$\gamma^{-i,j}_{L}(Z_k,w,n,n') = \xi^{-i,j}_{L}(w,n,n') - \xi^{-i,j,k}_{L}(w,n,n').$$

Take any n,n' in $\mathscr{N}$. There exists a deterministic sequence $\overline{J}_L = O(1)$ such that for any $t \equiv (w,n,n')$ where $w \in \mathscr{W}_{n,n'}$

$$\left| E\Big[\varphi^{\Delta}(Z_k,t) \;\Big|\; \xi^{-i,j,k}_{L}(t) + \beta\cdot\gamma^{-i,j}_{L}(Z_k,t)\Big] - E\Big[\varphi^{\Delta}(Z_k,t) \;\Big|\; \xi^{-i,j,k}_{L}(t) + \beta'\cdot\gamma^{-i,j}_{L}(Z_k,t)\Big] \right|$$
$$\leq \overline{J}_L \cdot |\beta - \beta'| \cdot \left|\gamma^{-i,j}_{L}(Z_k,t)\right| \quad \forall\,(\beta,\beta') \in [0,1).$$

Assumption (T2.i) essentially requires that, for $\ell = i,j$, conditional on $\xi^{-i,j}_{L}(W_\ell,n,n')$ and $W_\ell \in \mathscr{W}_{n,n'}$, the density of $\Delta_{W|N}(W,n,n')$ be bounded in a semi open interval of the form $[-\tau,0)$. This condition will help establish results of the following form for each n,n' in $\mathscr{N}$,

$$\Pr\Big(-|\xi^{-i,j}_{L}(W_\ell,n,n')| \leq \Delta_{W|N}(W_\ell,n,n') < 0 \;\Big|\; \xi^{-i,j}_{L}(W_\ell,n,n'),\, W_\ell \in \mathscr{W}_{n,n'}\Big) = O\Big(|\xi^{-i,j}_{L}(W_\ell,n,n')|\Big).$$

We wish to stress that, since Assumption (T2.i) deals exclusively with semi-open intervals of the form $[-\tau,0)$, it does not preclude $\Delta_{W|N}(W,n,n')$ from having a point mass at zero (which would be the case if (9) is binding as equality with positive probability). Combined with the conditions leading to (24), the Lipschitz-type restriction in Assumption (T2.ii) will help ensure that

$$\sup_{\substack{w \in \mathscr{W}_{n,n'} \\ n,n' \in \mathscr{N}}} \left| E\big[\varphi^{\Delta}(Z_k,w,n,n') \mid \xi^{-i,j}_{L}(w,n,n')\big]\right| = O_p\left(\frac{1}{L}\right)$$

for $k \neq i,j$. This result will help yield an exponential probability bound which helps establish our main result.

Assumption (T2) is relevant for the inequalities in (9); the following are the analogous conditions for (10).

Assumption T2'

(i) Take $i \neq j$ and let ℓ denote either i or j. There exists a $\tau > 0$ and a deterministic sequence $\overline{H}_L = O(1)$ such that, for any $n, n' \in \mathscr{N}$,

$$\Pr\Big(-s \leq \Phi_{W|N}(W_\ell, n, n') < 0 \;\Big|\; \widetilde{\xi}_L^{-i,j}(W_\ell, n, n'), W_\ell \in \mathscr{W}_{n,n'}\Big) \leq \overline{H}_L \cdot |s| \quad \forall\, 0 < s \leq \tau.$$

(ii) Take any n, n' in $\mathscr{N}$ and any $w \in \mathscr{W}_{n,n'}$. For $k \neq i, j$ let

$$\widetilde{\gamma}_L^{-i,j}(Z_k, w, n, n') = \widetilde{\xi}_L^{-i,j}(w, n, n') - \widetilde{\xi}_L^{-i,j,k}(w, n, n').$$

Take any n, n' in $\mathscr{N}$. There exists a deterministic sequence $\overline{J}_L = O(1)$ such that for any $t \equiv (w, n, n')$ where $w \in \mathscr{W}_{n,n'}$

$$\left| E\Big[\varphi^{\Phi}(Z_k, t) \;\Big|\; \widetilde{\xi}_L^{-i,j,k}(t) + \beta \cdot \widetilde{\gamma}_L^{-i,j}(Z_k, t)\Big] - E\Big[\varphi^{\Phi}(Z_k, t) \;\Big|\; \widetilde{\xi}_L^{-i,j,k}(t) + \beta' \cdot \widetilde{\gamma}_L^{-i,j}(Z_k, t)\Big] \right|$$
$$\leq \overline{J}_L \cdot |\beta - \beta'| \cdot \left|\widetilde{\gamma}_L^{-i,j}(Z_k, t)\right| \quad \forall\, (\beta, \beta') \in [0, 1).$$

Assumption (T2') will serve purposes analogous to those of Assumption (T2).

Assumption T3

b_L is a positive sequence that satisfies $\sqrt{L} \cdot b_L \longrightarrow \infty$ and $\sqrt{L} \cdot b_L^2 \longrightarrow 0$.

Along with our previous assumptions, using a nonzero bandwidth b_L that converges to zero at a rate slower than $L^{-1/2}$ will help yield exponential probability bounds while allowing for either $\Delta_{W|N}(W, n, n')$ or $\Phi_{W|N}(W, n, n')$ to have positive probability mass at zero (i.e., allowing for either (9) or (10) to be binding as equalities with positive probability). Qualitatively speaking, we require b_L to converge to zero faster than $L^{-1/4}$ because the rate at which objects of the following type disappear in probability will be relevant during the proof of our main result,

$$\Big(|\xi_L^{-i,j}(W_\ell, n, n')| + b_L\Big) \cdot \Pr\Big(-|\xi_L^{-i,j}(W_\ell, n, n')| - b_L \leq \Delta_{W|N}(W_\ell, n, n') < 0 \;\Big|\; \xi_L^{-i,j}(W_\ell, n, n'),\, W_\ell \in \mathscr{W}_{n,n'}\Big),$$
$$\Big(|\widetilde{\xi}_L^{-i,j}(W_\ell, n, n')| + b_L\Big) \cdot \Pr\Big(-|\widetilde{\xi}_L^{-i,j}(W_\ell, n, n')| - b_L \leq \Phi_{W|N}(W_\ell, n, n') < 0 \;\Big|\; \widetilde{\xi}_L^{-i,j}(W_\ell, n, n'),\, W_\ell \in \mathscr{W}_{n,n'}\Big).$$

Combined with the conditions leading to (24) and with Assumptions (T2.i) and (T2'.i), the above objects will disappear at a rate of b_L^2. Assumption (T3) then ensures that this convergence rate is faster than $L^{-1/2}$.

B.4.2 Main result

For a given n, n' in $\mathscr{N}$ and a given w, let

$$
\begin{aligned}
&\Gamma^F(w,n,n') = \\
&\quad E_W\left[\left(\mathbb{1}\{w \le W\} - F_{W|N}(W|n')\right) \cdot \mathbb{1}\left\{\Delta_{W|N}(W,n,n') \ge 0\right\} \cdot \mathbb{1}\left\{W \in \mathscr{W}_{n,n'}\right\}\right], \\
&\Gamma^\Phi(w,n,n') = \\
&\quad E_W\left[\left(\Omega\left(F_{W|N}(W|n'),n,n'\right) - \mathbb{1}\{w \le W\}\right) \cdot \mathbb{1}\left\{\Phi_{W|N}(W,n,n') \ge 0\right\} \cdot \mathbb{1}\left\{W \in \mathscr{W}_{n,n'}\right\}\right], \\
&\Gamma^\Omega(w,n,n') = \\
&\quad E_W\left[\nabla_1\Omega\left(F_{W|N}(W|n'),n,n'\right) \cdot \left(\mathbb{1}\{w \le W\} - F_{W|N}(W|n')\right) \cdot \mathbb{1}\left\{\Phi_{W|N}(W,n,n') \ge 0\right\} \cdot \mathbb{1}\left\{W \in \mathscr{W}_{n,n'}\right\}\right]
\end{aligned} \tag{25}
$$

These functionals will help us characterize the asymptotic distribution of our unconditional test-statistics. Let

$$
U_{L^{(2)}}^{\widehat{T}^F_{n,n'}} = \frac{1}{L(L-1)} \sum\sum_{i,j\in\{1,\ldots,L\}, i\neq j} \widehat{T}^F_{n,n'}(Z_i,Z_j) \quad U_{L^{(2)}}^{\widehat{T}^\Omega_{n,n'}} = \frac{1}{L(L-1)} \sum\sum_{i,j\in\{1,\ldots,L\}, i\neq j} \widehat{T}^\Omega_{n,n'}(Z_i,Z_j).
$$

By (14) we have

$$
U_{L^{(2)}}^{\widehat{T}^F} = \sum\sum_{n,n'\in\mathscr{N}, n>n'} U_{L^{(2)}}^{\widehat{T}^F_{n,n'}}, \quad \text{and} \quad U_{L^{(2)}}^{\widehat{T}^\Omega} = \sum\sum_{n,n'\in\mathscr{N}, n>n'} U_{L^{(2)}}^{\widehat{T}^\Omega_{n,n'}}.
$$

We establish the asymptotic properties of our statistics $U_{L^{(2)}}^{\widehat{T}^F}$ and $U_{L^{(2)}}^{\widehat{T}^\Omega}$ by characterizing those of $U_{L^{(2)}}^{\widehat{T}^F_{n,n'}}$ and $U_{L^{(2)}}^{\widehat{T}^\Omega_{n,n'}}$ for each n, n'.

Theorem 5 *Let $\mu^F_{n,n'}$ and $\mu^\Omega_{n,n'}$ be as defined in* (12)*.*

(i) If Assumptions (T1), (T2) and (T3) are satisfied, then

$$
\sqrt{L} \cdot U_{L^{(2)}}^{\widehat{T}^F_{n,n'}} = \sqrt{L} \cdot \mu^F_{n,n'} + \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \eta^F_{n,n'}(Z_i) + o_p(1), \quad \textit{where} \quad E\left[\eta^F_{n,n'}(Z_i)\right] = 0.
$$

From (14)*, it follows by construction that*

$$
\sqrt{L} \cdot U_{L^{(2)}}^{\widehat{T}^F} = \sqrt{L} \cdot \left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}} \mu^F_{n,n'}\right] + \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \eta^F_{\mathscr{N}}(Z_i) + o_p(1), \quad \textit{where} \quad \eta^F_{\mathscr{N}}(Z_i) = \left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}} \eta^F_{n,n'}(Z_i)\right]. \tag{26}
$$

For each n, n' in $\mathscr{N}$, if (9) *is satisfied with probability one in $\mathscr{W}_{n,n'}$, the function $\eta^F_{n,n'}$ reduces to*

$$
\eta^F_{n,n'}(Z_i) = \mathbb{1}\{N_i = n\} \cdot \Gamma^F(W_i,n,n') - \frac{p_N(n)}{p_N(n')} \cdot \mathbb{1}\{N_i = n'\} \cdot \Gamma^F(W_i,n,n'). \tag{27}
$$

(ii) If Assumptions (T1), (T2') and (T3) are satisfied, then

$$
\sqrt{L} \cdot U_{L^{(2)}}^{\widehat{T}^\Omega_{n,n'}} = \sqrt{L} \cdot \mu^\Omega_{n,n'} + \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \eta^\Omega_{n,n'}(Z_i) + o_p(1), \quad \textit{where} \quad E\left[\eta^\Omega_{n,n'}(Z_i)\right] = 0.
$$

From (14), *it follows by construction that*

$$\sqrt{L}\cdot U_{L(2)}^{\widehat{T}^{\Omega}} = \sqrt{L}\cdot\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\mu_{n,n'}^{\Omega}\right] + \frac{1}{\sqrt{L}}\sum_{i=1}^{L}\eta_{\mathscr{N}}^{\Omega}(Z_i) + o_p(1), \quad where \quad \eta_{\mathscr{N}}^{\Omega}(Z_i) \quad = \left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\eta_{n,n'}^{\Omega}(Z_i)\right]. \tag{28}$$

For each n, n' *in* $\mathscr{N}$, *if* (10) *is satisfied with probability one in* $\mathscr{W}_{n,n'}$, *the function* $\eta_{n,n'}^{\Omega}$ *reduces to*

$$\eta_{n,n'}^{\Omega}(Z_i) = \mathbb{1}\{N_i = n\}\cdot\Gamma^{\Phi}(W_i,n,n') + \frac{p_N(n)}{p_N(n')}\cdot\mathbb{1}\{N_i = n'\}\cdot\Gamma^{\Omega}(W_i,n,n'). \tag{29}$$

Note that the functions $\eta_{n,n'}^{F}(Z_i)$ and $\eta_{n,n'}^{\Omega}(Z_i)$ have mean zero regardless of whether (9) or (10) are satisfied. These functions are the leading terms in the Hoeffding decompositions (see Hoeffding (1961) and Serfling (1980), Chapter 5) of the relevant U-statistics involved. Their general expressions can be found in Equations (T-42) and (T-46) of the Technical Supplement. Theorem 5 characterizes these expressions for the case where (9) and (10) are satisfied w.p.1. in $\mathscr{W}_{n,n'}$. In this instance, using (25) and iterated expectations, we can easily show that $E\big[\mathbb{1}\{N_i = n\}\cdot\Gamma^{F}(W_i,n,n')\big] = \mu_{n,n'}^{F} = 0$ and $E\big[\mathbb{1}\{N_i = n'\}\cdot\Gamma^{F}(W_i,n,n')\big] = 0$, which yields $E\big[\eta_{n,n'}^{F}(Z_i)\big] = \mu_{n,n'}^{F} = 0$. Likewise, iterated expectations yields $E\big[\mathbb{1}\{N_i = n\}\cdot\Gamma^{\Phi}(W_i,n,n')\big] = \mu_{n,n'}^{\Omega} = 0$ and $E\big[\mathbb{1}\{N_i = n'\}\cdot\Gamma^{\Omega}(W_i,n,n')\big] = 0$, yielding $E\big[\eta_{n,n'}^{\Omega}(Z_i)\big] = \mu_{n,n'}^{\Omega} = 0$.

B.4.3 A Rejection Rule based on Theorem 5

Based on the statement of Theorem 5 and the expressions in (27) and (29), we see that if (9) is satisfied almost everywhere in $\mathscr{W}_{n,n'}$ for each $n > n'$ in $\mathscr{N}$ as a *strict* inequality, then $\eta_{\mathscr{N}}^{F}(Z_i) = 0$ w.p.1, and therefore $\mathrm{Var}\big(\eta_{\mathscr{N}}^{F}(Z_i)\big) = 0$. Otherwise, said variance is strictly positive. The same is true about (10) and $\mathrm{Var}\big(\eta_{\mathscr{N}}^{\Omega}(Z_i)\big)$. Combining (26) with Theorems 4 and 5, $\sqrt{L}\cdot U_{L(2)}^{\widehat{T}^{F}}$ will diverge to $+\infty$ with probability 1 if the inequalities in (9) are violated with positive probability on $\mathscr{W}_{n,n'}$ for any $n, n' \in \mathscr{N}$ $(n > n')$. Conversely, if (9) is satisfied as strict inequality almost everywhere in each $\mathscr{W}_{n,n'}$, then $\sqrt{L}\cdot U_{L(2)}^{\widehat{T}^{F}}$ will vanish in probability. Finally, if (9) is satisfied almost everywhere in each $\mathscr{W}_{n,n'}$ and binding with nonzero probability in some $\mathscr{W}_{n,n'}$, the statistic $\sqrt{L}\cdot U_{L(2)}^{\widehat{T}^{F}}$ will be asymptotically normally distributed with mean zero and variance $\mathrm{Var}\big(\eta_{\mathscr{N}}^{F}(Z_i)\big) > 0$. The same results hold for (10) and $\sqrt{L}\cdot U_{L(2)}^{\widehat{T}^{\Omega}}$.

Let $\widehat{\eta}_{n,n'}^{F}(Z_i)$ and $\widehat{\eta}_{n,n'}^{\Omega}(Z_i)$ be estimators of the expressions given in (27) and (29). It follows from Theorem 5 that, for any pair of arbitrary constants $c_1 > 0$ and $c_2 > 0$, the rejection rules described in (16A) and (16B) based on

$$\frac{\sqrt{L}\cdot U_{L(2)}^{\widehat{T}^{F}}}{\sqrt{\widehat{\mathrm{Var}}\left[\sum\sum_{n>n'}\widehat{\eta}_{n,n'}^{F}(Z_i)\right] + c_1}} \quad \text{and} \quad \frac{\sqrt{L}\cdot U_{L(2)}^{\widehat{T}^{\Omega}}}{\sqrt{\widehat{\mathrm{Var}}\left[\sum\sum_{n>n'}\widehat{\eta}_{n,n'}^{\Omega}(Z_i)\right] + c_2}}, \tag{30}$$

would satisfy the asymptotic properties described in (17) for our testing range. Let φ^F and φ^Ω be as defined in (22). Define

$$\begin{aligned} \widehat{G}^F_{n,n'}(Z_i,Z_j,Z_k) &= \varphi^F(Z_i,W_k,n')\cdot \mathbb{1}\{N_j=n\}\cdot \mathbb{1}\big\{\widehat{\Delta}^{-i,j,k}_{W|N}(W_k,n,n')\geq -b_L\big\}\cdot \mathbb{1}\big\{W_k\in \mathscr{W}_{n,n'}\big\},\\ \widehat{G}^\Omega_{n,n'}(Z_i,Z_j,Z_k) &= \varphi^\Omega(Z_i,W_k,n,n')\cdot \mathbb{1}\{N_j=n\}\cdot \mathbb{1}\big\{\widehat{\Phi}^{-i,j,k}_{W|N}(W_k,n,n')\geq -b_L\big\}\cdot \mathbb{1}\big\{W_k\in \mathscr{W}_{n,n'}\big\}. \end{aligned} \tag{31}$$

We can estimate $\eta^F_{n,n'}$ and $\eta^\Omega_{n,n'}$ as described in (27) and (29) by using

$$\begin{aligned} \widehat{\eta}^F_{n,n'}(Z_i) &= \frac{1}{L-1}\sum_{j\neq i}\widehat{T}^F_{n,n'}(Z_i,Z_j) - \frac{1}{(L-1)(L-2)}\sum_{j\neq i}\sum_{\substack{k\neq j\\ k\neq i}}\widehat{G}^F_{n,n'}(Z_i,Z_j,Z_k),\\ \widehat{\eta}^\Omega_{n,n'}(Z_i) &= \frac{1}{L-1}\sum_{j\neq i}\widehat{T}^\Omega_{n,n'}(Z_i,Z_j) + \frac{1}{(L-1)(L-2)}\sum_{j\neq i}\sum_{\substack{k\neq j\\ k\neq i}}\widehat{G}^\Omega_{n,n'}(Z_i,Z_j,Z_k) \end{aligned} \tag{32}$$

The resulting estimated variances to be used in (30), $\widehat{\text{Var}}\Big[\sum_{n>n'}\sum\widehat{\eta}^F_{n,n'}(Z_i)\Big]$ and $\widehat{\text{Var}}\Big[\sum_{n>n'}\sum\widehat{\eta}^\Omega_{n,n'}(Z_i)\Big]$, are consistent under the conditions of Theorem 5.

B.4.4 A Test for the Reverse Inequality in (10) and Its Asymptotic Properties

We go back to the test of (10'), which was used in Section 3.4 to discriminate between CIPV and IPV. (10') is simply the reverse inequality in (10). As we pointed out there, if the data supports (10), then rejecting (10') implies that the inequalities in (10) are strict with positive probability, which rules out IPV as the true model. Let $T^{-\Omega}_{n,n'}(Z_i,Z_j)$ be as defined in (11'). A test of (10') replaces $T^\Omega_{n,n'}(Z_i,Z_j)$ with

$$\begin{aligned} \widehat{T}^{-\Omega}_{n,n'}(Z_i,Z_j) = \Big(&\mathbb{1}\{W_i\leq W_j\} - \Omega\big(\widehat{F}^{-i,j}_{W|N}(W_j|n'),n,n'\big)\Big)\cdot \mathbb{1}\{N_i=n\}\\ &\cdot \mathbb{1}\big\{-\widehat{\Phi}^{-i,j}_{W|N}(W_j,n,n')\geq -b_L\big\}\cdot \mathbb{1}\{W_j\in\mathscr{W}_{n,n'}\} \end{aligned}$$

Let

$$\begin{aligned} U^{T^{-\Omega}}_{L^{(2)}} &= \frac{1}{L(L-1)}\sum\sum_{i,j\in\{1,\dots,L\},i\neq j}\left[\sum\sum_{n,n'\in\mathscr{N},n>n'} T^{-\Omega}_{n,n'}(Z_i,Z_j)\right]\\ U^{\widehat{T}^{-\Omega}}_{L^{(2)}} &= \frac{1}{L(L-1)}\sum\sum_{i,j\in\{1,\dots,L\},i\neq j}\left[\sum\sum_{n,n'\in\mathscr{N},n>n'} \widehat{T}^{-\Omega}_{n,n'}(Z_i,Z_j)\right] \end{aligned} \tag{33}$$

For a given n,n' in $\mathscr{N}$ and a given w, let

$$\begin{aligned} &\Gamma^{-\Phi}(w,n,n') =\\ &E_W\left[\big(\mathbb{1}\{w\leq W\} - \Omega\big(F_{W|N}(W|n'),n,n'\big)\big)\cdot \mathbb{1}\big\{-\Phi_{W|N}(W,n,n')\geq 0\big\}\cdot \mathbb{1}\big\{W\in\mathscr{W}_{n,n'}\big\}\right],\\ &\Gamma^{-\Omega}(w,n,n') =\\ &E_W\left[\nabla_1\Omega\big(F_{W|N}(W|n'),n,n'\big)\cdot \big(\mathbb{1}\{w\leq W\} - F_{W|N}(W|n')\big)\cdot \mathbb{1}\big\{-\Phi_{W|N}(W,n,n')\geq 0\big\}\cdot \mathbb{1}\big\{W\in\mathscr{W}_{n,n'}\big\}\right]. \end{aligned}$$

Let

$$\mu^{-\Omega}_{n,n'} = p_N(n)\cdot E\Big[\max\big\{0\ ,\ -\Phi_{W|N}(W_j,n,n')\big\}\cdot \mathbb{1}\big\{W_j\in\mathscr{W}_{n,n'}\big\}\Big]$$

We have $E\big[T^{-\Omega}_{n,n'}(Z_i,Z_j)\big]=\mu^{-\Omega}_{n,n'}$; note that $\mu^{-\Omega}_{n,n'}\geq 0$ by construction. Having $\mu^{-\Omega}_{n,n'}>0$ for some $n>n'$ in $\mathscr{N}$ would indicate that (10) holds as a *strict* inequality with positive probability in $\mathscr{W}_{n,n'}$, leading us to reject the notion that, unconditionally, the data generating process is consistent with IPV. Denote

$$U_{L^{(2)}}^{\widehat{T}^{-\Omega}_{n,n'}}=\frac{1}{L(L-1)}\sum\sum_{i,j\in\{1,\dots,L\},i\neq j}\widehat{T}^{-\Omega}_{n,n'}(Z_i,Z_j).$$

Under the type of conditions leading to part (ii) of Theorem 5, we can show that

$$\sqrt{L}\cdot U_{L^{(2)}}^{\widehat{T}^{-\Omega}_{n,n'}}=\sqrt{L}\cdot\mu^{-\Omega}_{n,n'}+\frac{1}{\sqrt{L}}\sum_{i=1}^{L}\eta^{-\Omega}_{n,n'}(Z_i)+o_p(1),\quad\text{where}\quad E\big[\eta^{-\Omega}_{n,n'}(Z_i)\big]=0.$$

It follows by construction from (33) that

$$\sqrt{L}\cdot U_{L^{(2)}}^{\widehat{T}^{-\Omega}}=\sqrt{L}\cdot\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\mu^{-\Omega}_{n,n'}\right]+\frac{1}{\sqrt{L}}\sum_{i=1}^{L}\eta^{-\Omega}_{\mathscr{N}}(Z_i)+o_p(1),\quad\text{where}\quad\eta^{-\Omega}_{\mathscr{N}}(Z_i)\quad=\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\eta^{-\Omega}_{n,n'}(Z_i)\right].\quad(34)$$

For each n,n' in $\mathscr{N}$, if the reverse inequality in (10) is satisfied with probability one in $\mathscr{W}_{n,n'}$, the function $\eta^{-\Omega}_{n,n'}$ reduces to

$$\eta^{-\Omega}_{n,n'}(Z_i)=\mathbb{1}\{N_i=n\}\cdot\Gamma^{-\Phi}(W_i,n,n')-\frac{p_N(n)}{p_N(n')}\cdot\mathbb{1}\{N_i=n'\}\cdot\Gamma^{-\Omega}(W_i,n,n').\quad(35)$$

Let $\widehat{\eta}^{-\Omega}_{n,n'}$ denote an estimator of the function described in (35) and let $c>0$ denote a pre-specified constant. A rejection rule like the ones described previously could be used, based on

$$\frac{\sqrt{L}\cdot U_{L^{(2)}}^{\widehat{T}^{-\Omega}}}{\sqrt{\widehat{\mathrm{Var}}\left[\sum\sum_{n>n'}\widehat{\eta}^{-\Omega}_{n,n'}(Z_i)\right]+c}}\quad(36)$$

Let φ^{Ω} be as defined in (22) and denote

$$\widehat{G}^{-\Omega}_{n,n'}(Z_i,Z_j,Z_k)=\varphi^{\Omega}(Z_i,W_k,n,n')\cdot\mathbb{1}\{N_j=n\}\cdot\mathbb{1}\big\{-\widehat{\Phi}^{-i,j,k}_{W|N}(W_k,n,n')\geq -b_L\big\}\cdot\mathbb{1}\big\{W_k\in\mathscr{W}_{n,n'}\big\}.\quad(37)$$

We can estimate the expression in (35) with

$$\widehat{\eta}^{-\Omega}_{n,n'}(Z_i)=\frac{1}{L-1}\sum_{j\neq i}\widehat{T}^{-\Omega}_{n,n'}(Z_i,Z_j)-\frac{1}{(L-1)(L-2)}\sum_{j\neq i}\sum_{\substack{k\neq j\\ k\neq i}}\widehat{G}^{-\Omega}_{n,n'}(Z_i,Z_j,Z_k).\quad(38)$$

From here we estimate $\widehat{\mathrm{Var}}\left[\sum\sum_{n>n'}\widehat{\eta}^{-\Omega}_{n,n'}(Z_i)\right]$ to be plugged into (36).

B.5 Testing Conditional on Observable Covariates

Next, we modify the test to condition on observable covariates. Let X be a vector of covariates, and let $Y\equiv(N,X)$ and $Z\equiv(W,Y)$. We maintain the assumption of having an *i.i.d.* sample $(Z_i)_{i=1}^{L}$. Define

$$\begin{aligned}F_{W|Y}(w|n,x)&\equiv\Pr(W\leq w|N=n,X=x)\\ \Delta_{W|Y}(w,n,n',x)&\equiv F_{W|Y}(w|n,x)-F_{W|Y}(w|n',x)\\ \Phi_{W|Y}(w,n,n',x)&=\Omega\big(F_{W|Y}(w|n',x),n,n'\big)-F_{W|Y}(w|n,x)\end{aligned}$$

We will test whether conditional analogs of (9) and (10) hold at each realization of X – that is, whether $n > n'$ implies

$$F_{W|Y}(w|n,x) \quad \leq \quad F_{W|Y}(w|n',x) \tag{39}$$

and

$$\psi^{-1}_{n-1:n}\left(F_{W|Y}(w|n,x)\right) \quad \geq \quad \psi^{-1}_{n'-1:n'}\left(F_{W|Y}(w|n',x)\right) \tag{40}$$

(or, equivalently, $\Delta_{W|Y}(w,n,n',x) \leq 0$ and $\Phi_{W|Y}(w,n,n'x) \leq 0$ almost everywhere). For clarity, we present the case where X is one-dimensional and continuously distributed, as is the case in our application. Later we discuss extending the test to cases where X is multi-dimensional and includes both discrete and continuous covariates.

As before, let $\mathscr{N} \subseteq \mathcal{S}_N$ be a fixed subset with $p_{_N}(n) > 0$ for all $n \in \mathscr{N}$, and let $\mathscr{C}_{n,n'} \subseteq \mathcal{S}_X \times \mathcal{S}_W$ be a pre-specified range of values of (x,w) on which we will test (39) and (40). (Again, $\mathscr{C}_{n,n'}$ should be chosen such that $F_{W|Y}(\cdot|n,\cdot)$ and $F_{W|Y}(\cdot|n',\cdot)$ are both bounded away from 0 and 1 on $\mathscr{C}_{n,n'}$.) Let $h_{_L} \longrightarrow 0$ be a nonnegative bandwidth sequence converging to zero, and let $K : \mathbb{R} \longrightarrow \mathbb{R}$ be a bias-reducing kernel of order M.[35] We will impose conditions on M below. Define

$$\begin{aligned}
S^F_{L_{n,n'}}(Z_i,Z_j,Z_k) &= \Big(\mathbb{1}\{W_i \leq W_k\} - F_{W|Y}(W_k|n',X_j)\Big) \cdot \mathbb{1}\{N_i = n\} \\
&\quad \cdot \mathbb{1}\big\{\Delta_{W|Y}(W_k,n,n',X_j) \geq 0\big\} \cdot \mathbb{1}\big\{(X_j,W_k) \in \mathscr{C}_{n,n'}\big\} \cdot \tfrac{1}{h_L} K\left(\tfrac{X_i - X_j}{h_L}\right) \\
S^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k) &= \Big(\Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big) - \mathbb{1}\{W_i \leq W_k\}\Big) \cdot \mathbb{1}\{N_i = n\} \\
&\quad \cdot \mathbb{1}\big\{\Phi_{W|Y}(W_k,n,n',X_j) \geq 0\big\}\mathbb{1}\big\{(X_j,W_k) \in \mathscr{C}_{n,n'}\big\} \cdot \tfrac{1}{h_L} K\left(\tfrac{X_i - X_j}{h_L}\right)
\end{aligned} \tag{41}$$

Both functions are constructed such that only triples (Z_i,Z_j,Z_k) for which $X_i - X_j$ lies in a vanishing neighborhood around zero will matter asymptotically. The following result describes a set of sufficient conditions under which $\lim_{L\to\infty} E\big[S^F_{L_{n,n'}}(Z_i,Z_j,Z_k)\big] = 0$ and $\lim_{L\to\infty} E\big[S^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k)\big] = 0$ if and only if $\Delta_{W|Y}$ and $\Phi_{W|Y}$ are negative almost everywhere within the test range.

Theorem 6 *Let $f_{X|N}(\cdot|n)$ denote the density of X conditional on $N = n$. Suppose that the support of $f_{X|N}(\cdot|n)$ is the same for all $n \in \mathscr{N}$. Further, suppose that for any $n \in \mathscr{N}$, $f_{X|N}(x|n)$ and $F_{W|Y}(w|n,x)$ are both M times differentiable with respect to x, with bounded derivatives at almost all x for which $(x,w) \in \mathscr{C}_{n,n'}$ for some w. Then for $i \neq j \neq k$,*

$$\begin{aligned}
E\big[S^F_{L_{n,n'}}(Z_i,Z_j,Z_k)\big] &= \gamma^F_{n,n'} + O(h^M_L) \\
E\big[S^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k)\big] &= \gamma^\Omega_{n,n'} + O(h^M_L)
\end{aligned}$$

where

$$\gamma^F_{n,n'} = p_{_N}(n) \cdot E\Big[\max\big\{0,\Delta_{W|Y}(W_k,n,n',X_j)\big\} \cdot f_{X|N}(X_j|n) \cdot \mathbb{1}\big\{(X_j,W_k) \in \mathscr{C}_{n,n'}\big\}\Big]$$

$$\gamma^\Omega_{n,n'} = p_{_N}(n) \cdot E\Big[\max\big\{0,\Phi_{W|Y}(W_k,n,n',X_j)\big\} \cdot f_{X|N}(X_j|n) \cdot \mathbb{1}\big\{(X_j,W_k) \in \mathscr{C}_{n,n'}\big\}\Big]$$

[35] That is, K satisfies $K(s) = K(-s)$, $\int_{-\infty}^{\infty} K(s)ds = 1$, $\int_{-\infty}^{\infty} s^r K(s)ds = 0$ for all $r = 1,\ldots,M-1$, and $\int_{-\infty}^{\infty} s^M K(s)ds < \infty$.

Proof. We focus on $E\big[S^{\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)\big]$; the proof for $E\big[S^{F}_{L_{n,n'}}(Z_i,Z_j,Z_k)\big]$ follows identical steps. As before, define

$$\begin{aligned}
\overline{S}^{\Omega}_{L_{n,n'}}(N_i,Z_j,Z_k) &= E_{W_i,X_i|N_i}\left[S^{\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)\right] = E_{X_i|N_i}\left[E_{W_i|N_i,X_i}\left[S^{\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)\right]\right] \\
&= E_{X_i|N_i}\Bigg[\Big(\Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big)-F_{W|Y}(W_k|N_i,X_i)\Big)\cdot \mathbb{1}\{N_i=n\} \\
&\quad \cdot \mathbb{1}\big\{\Phi_{W|Y}(W_k,n,n',X_j)\geq 0\big\}\cdot \mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\}\cdot \tfrac{1}{h_L}K\Big(\tfrac{X_i-X_j}{h_L}\Big)\Bigg] \\
&= E_{X_i|N_i}\Bigg[\Big(\Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big)-F_{W|Y}(W_k|N_i,X_i)\Big)\cdot \tfrac{1}{h_L}K\Big(\tfrac{X_i-X_j}{h_L}\Big)\Bigg] \\
&\quad \cdot \mathbb{1}\{N_i=n\}\cdot \mathbb{1}\big\{\Phi_{W|Y}(W_k,n,n',X_j)\geq 0\big\}\cdot \mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\} \\
&= \Big[\Big(\Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big)-F_{W|Y}(W_k|n,X_j)\Big)f_{X|N}(X_j|n)+O(h_L^M)\Big] \\
&\quad \cdot \mathbb{1}\{N_i=n\}\cdot \mathbb{1}\big\{\Phi_{W|Y}(W_k,n,n',X_j)\geq 0\big\}\cdot \mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\} \\
&= \mathbb{1}\{N_i=n\}\cdot \mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\} \\
&\quad \cdot \max\{0,\Phi_{W|Y}(W_k,n,n',X_j)\}\cdot f_{X|N}(X_j|n)+O(h_L^M)
\end{aligned}$$

The second-to-last inequality comes from an M^{th}-order Taylor approximation, which given our smoothness assumption yields

$$\begin{aligned}
&E_{X_i|N_i}\Bigg[\Big(\Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big)-F_{W|Y}(W_k|N_i,X_i)\Big)\cdot \tfrac{1}{h_L}K\Big(\tfrac{X_i-X_j}{h_L}\Big)\Bigg] \\
&\quad = \int\Big(\Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big)-F_{W|Y}(W_k|N_i,x)\Big)\tfrac{1}{h_L}K\Big(\tfrac{x-X_j}{h_L}\Big)f_{X|N}(x|N_i)dx \\
&\quad = \Big(\Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big)-F_{W|Y}(W_k|N_i,X_j)\Big)f_{X|N}(X_j|N_i)+O(h_L^M)
\end{aligned}$$

along with the fact that the entire expression is zero unless $N_i=n$. Taking the expectation of $\overline{S}^{\Omega}_{L_{n,n'}}$ over (N_i,Z_j,Z_k) then proves the result. □

As with the unconditional test, we modify (41) by replacing $F_{W|Y}$ with a nonparametric estimate calculated on all the observations but i,j,k and introducing a positive but vanishing bandwidth b_L. Let $\widetilde{K}$ and $\widetilde{h}_L$ be an additional kernel function and bandwidth sequence, and define

$$\widehat{R}^{-i,j,k}_{W|Y}(w|n,x) = \tfrac{1}{(L-3)\widetilde{h}_L}\textstyle\sum_{\ell\neq i,j,k}\mathbb{1}\{W_\ell\leq w\}\mathbb{1}\{N_\ell=n\}\widetilde{K}\Big(\tfrac{X_\ell-x}{\widetilde{h}_L}\Big)$$

$$\widehat{p}^{-i,j,k}_{Y}(x,n) = \tfrac{1}{(L-3)\widetilde{h}_L}\textstyle\sum_{\ell\neq i,j,k}\mathbb{1}\{N_\ell=n\}\widetilde{K}\Big(\tfrac{X_\ell-x}{\widetilde{h}_L}\Big)$$

$$\widehat{F}^{-i,j,k}_{W|Y}(w|n,x) = \begin{cases}\widehat{R}^{-i,j,k}_{W|Y}(w|n,x)/\widehat{p}^{-i,j,k}_{Y}(x,n) & if \quad \widehat{p}^{-i,j,k}_{Y}(x,n)\neq 0\\ 0 & otherwise\end{cases}$$

and

$$\widehat{\Delta}^{-i,j,k}_{W|Y}(w,n,n',x) = \widehat{F}^{-i,j,k}_{W|Y}(w|n,x)-\widehat{F}^{-i,j,k}_{W|Y}(w|n',x)$$

$$\widehat{\Phi}^{-i,j,k}_{W|Y}(w,n,n',x) = \Omega\big(\widehat{F}^{-i,j,k}_{W|Y}(w|n',x),n,n'\big)-\widehat{F}^{-i,j,k}_{W|Y}(w|n,x)$$

Define

$$\begin{aligned}
\widehat{S}^F_{L_{n,n'}}(Z_i,Z_j,Z_k) &= \Big(\mathbb{1}\{W_i \le W_k\} - \widehat{F}^{-i,j,k}_{W|Y}(W_k|n',X_j)\Big)\cdot \mathbb{1}\{N_i=n\} \\
&\quad \cdot \mathbb{1}\big\{\widehat{\Delta}^{-i,j,k}_{W|Y}(W_k,n,n',X_j) \ge -b_L\big\}\cdot \mathbb{1}\big\{(X_j,W_k)\in \mathscr{C}_{n,n'}\big\}\cdot \tfrac{1}{h_L}K\left(\tfrac{X_i-X_j}{h_L}\right) \\
\widehat{S}^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k) &= \Big(\Omega\big(\widehat{F}^{-i,j,k}_{W|Y}(W_k|n',X_j),n,n'\big) - \mathbb{1}\{W_i\le W_k\}\Big)\cdot \mathbb{1}\{N_i=n\} \\
&\quad \cdot \mathbb{1}\big\{\widehat{\Phi}^{-i,j,k}_{W|Y}(W_k,n,n',X_j)\ge -b_L\big\}\mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\}\cdot \tfrac{1}{h_L}K\left(\tfrac{X_i-X_j}{h_L}\right)
\end{aligned} \tag{42}$$

Our conditional test statistics are

$$\begin{aligned}
U^{\widehat{S}^F_L}_{L^{(3)}} &= \frac{1}{L(L-1)(L-2)}\sum\sum\sum_{\substack{i,j,k\in\{1,\dots,L\}\\ i\ne j\ne k}}\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\widehat{S}^F_{L_{n,n'}}(Z_i,Z_j,Z_k)\right] \\
U^{\widehat{S}^\Omega_L}_{L^{(3)}} &= \frac{1}{L(L-1)(L-2)}\sum\sum\sum_{\substack{i,j,k\in\{1,\dots,L\}\\ i\ne j\ne k}}\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\widehat{S}^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k)\right]
\end{aligned} \tag{43}$$

Appendix B.6 gives conditions under which $\sqrt{L}\cdot U^{\widehat{S}^F_L}_{L^{(3)}}$ and $\sqrt{L}\cdot U^{\widehat{S}^\Omega_L}_{L^{(3)}}$ have asymptotic behavior analogous to the unconditional test statistics; the formal result is given as Theorem 7.

B.6 Asymptotic Properties of the Conditional Test Statistics in (43)

Here we describe and discuss the main asymptotic results and the corresponding assumptions for the statistics described in (43). Detailed proofs of all our results can be found in the Technical Supplement.

B.6.1 Assumptions

Again, we begin with the distributional assumptions, all of which are compatible with the model presented in the text.

Assumption C1

(i) We observe an i.i.d. sample $(W_i,N_i,X_i)_{i=1}^L \equiv (Z_i)_{i=1}^L$. *Let* $f_{X|N}(\cdot|n)$ *denote the density of* X *conditional on* $N=n$. *Then* $f_{X|N}(x|n)$ *is continuous in* x *for each* $n\in\mathcal{S}_N$ *and its support is the same for every* $n\in\mathscr{N}$. *The distribution* $F_{W|Y}(w|n,x)$ *is continuous in* w *for almost every* $(n,x)\in\mathcal{S}_N\times\mathcal{S}_X$. *The sets* $\mathscr{N}\subseteq\mathcal{S}_N$ *and* $\mathscr{C}_{n,n'}$ *are compact (the latter being true for each* n,n' *in* $\mathscr{N}$*). For each* n,n' *in* $\mathscr{N}$ *and almost everywhere in* $\mathscr{C}_{n,n'}$, $F_{W|Y}(w|n,x)$ *is strictly bounded away from* 0 *and* 1.

(ii) Let $p_Y(x,n)=f_{X|N}(x|n)\cdot p_N(n)$. *Then inf* $\big\{p_Y(x,n) : n\in\mathscr{N}$ *and* $(w,x)\in\mathscr{C}_{n,n'}$ *for some* w *and* $n'\in\mathscr{N}\big\}\equiv \underline{p}_Y > 0$. *For each* $n,n'\in\mathscr{N}$ *and almost everywhere in* $\mathcal{C}_{n,n'}$, *the functions* $f_{X|N}(x|n)$ *and* $F_{W|Y}(w|n,x)$ *are* M *times differentiable with respect to* x *with bounded derivatives.*

As before, part (i) guarantees that $\nabla_1\Omega\big(F_{W|Y}(w|n',x),n,n'\big)$ is well-defined and strictly bounded in our testing range. Let

$$\begin{aligned}
\Upsilon^F_L(Z,w,n,x) &= \frac{\mathbb{1}\{W\leq w\}-F_{W|Y}(w|n,x)}{p_Y(n,x)}\cdot \mathbb{1}\{N=n\}\cdot\frac{1}{\widetilde{h}_L}\widetilde{K}\left(\frac{X-x}{\widetilde{h}_L}\right),\\
\Upsilon^\Delta_L(Z,w,n,n',x) &= \Upsilon^F_L(Z,w,n,x)-\Upsilon^F_L(Z,w,n',x),\\
\Upsilon^\Omega_L(Z,w,n,n',x) &= \nabla_1\Omega\big(F_{W|Y}(w|n',x),n,n'\big)\cdot\Upsilon^F_L(Z,w,n',x),\\
\Upsilon^\Phi_L(Z,w,n,n',x) &= \Upsilon^\Omega_L(Z,w,n,n',x)-\Upsilon^F_L(Z,w,n,x).
\end{aligned}\tag{44}$$

Let M be as described in Assumption (C1). Suppose the kernel function $\widetilde{K}$ is Lipschitz-continuous, bias-reducing of order M with compact support, and suppose the bandwidth sequence $\widetilde{h}_L$ satisfies $L^{1/2}\cdot\widetilde{h}^M_L\longrightarrow 0$. An $M^{th}-$order approximation yields, for all n,n' in $\mathscr{N}$ and $(w,x)\in\mathscr{C}_{n,n'}$,

$$\begin{aligned}
&E\big[\Upsilon^F_L(Z,w,n,x)\big]=O(h^M_L)=o(L^{-1/2}),\quad E\big[\Upsilon^\Delta_L(Z,w,n,n',x)\big]=O(h^M_L)=o(L^{-1/2}),\\
&E\big[\Upsilon^\Omega_L(Z,w,n,n',x)\big]=O(h^M_L)=o(L^{-1/2}),\quad E\big[\Upsilon^\Phi_L(Z,w,n,n',x)\big]=O(h^M_L)=o(L^{-1/2}).
\end{aligned}$$

Furthermore, suppose $L^{1/2}\cdot\widetilde{h}_L\longrightarrow\infty$ (we will impose stronger bandwidth convergence conditions below in Assumption (C4)). Then, given the class of functions involved and the conditions described in Assumption (C1), we can show (see e.g, Lemma 3 in Collomb and Hardle (1986), or Theorem 1' in Lewbel (1997) and the references cited there) that the linear representations for our nonparametric estimators are given by

$$\begin{aligned}
\widehat{F}^{-i,j,k}_{W|Y}(w|n,x) &= F_{W|Y}(w|n,x)+\frac{1}{L-3}\sum_{\ell\neq i,j,k}\Upsilon^F_L(Z_\ell,w,n,x)+\varsigma^{-i,j,k}_L(w,n,x),\\
\widehat{\Delta}^{-i,j,k}_{W|Y}(w,n,n',x) &= \Delta_{W|Y}(w,n,n',x)+\frac{1}{L-3}\sum_{\ell\neq i,j,k}\Upsilon^\Delta_L(Z_\ell,w,n,n',x)+\xi^{-i,j,k}_L(w,n,n',x),
\end{aligned}\tag{45A}$$

and

$$\begin{aligned}
\Omega\big(\widehat{F}^{-i,j,k}_{W|Y}(w|n',x),n,n'\big) &= \Omega\big(F_{W|Y}(w|n',x),n,n'\big)+\frac{1}{L-3}\sum_{\ell\neq i,j,k}\Upsilon^\Omega_L(Z_\ell,w,n,n',x)+\widetilde{\varsigma}^{-i,j,k}_L(w,n,n',x),\\
\widehat{\Phi}^{-i,j,k}_{W|Y}(w,n,n',x) &= \Phi_{W|Y}(w,n,n',x)+\frac{1}{L-3}\sum_{\ell\neq i,j,k}\Upsilon^\Phi_L(Z_\ell,w,n,n',x)+\widetilde{\xi}^{-i,j,k}_L(w,n,n',x),
\end{aligned}\tag{45B}$$

where, for *any* $\delta>0$ we have

$$\begin{aligned}
&\sup_{\substack{(w,x)\in\mathscr{C}_{n,n'}\\(n,n')\in\mathscr{N}}}\left|\varsigma^{-i,j,k}_L(w,n,n',x)\right|=O_p\left(\frac{1}{L^{1-\delta}\cdot\widetilde{h}_L}\right),\quad \sup_{\substack{(w,x)\in\mathscr{C}_{n,n'}\\(n,n')\in\mathscr{N}}}\left|\widetilde{\varsigma}^{-i,j,k}_L(w,n,n',x)\right|=O_p\left(\frac{1}{L^{1-\delta}\cdot\widetilde{h}_L}\right),\\
&\sup_{\substack{(w,x)\in\mathscr{C}_{n,n'}\\(n,n')\in\mathscr{N}}}\left|\xi^{-i,j,k}_L(w,n,n',x)\right|=O_p\left(\frac{1}{L^{1-\delta}\cdot\widetilde{h}_L}\right),\quad \sup_{\substack{(w,x)\in\mathscr{C}_{n,n'}\\(n,n')\in\mathscr{N}}}\left|\widetilde{\xi}^{-i,j,k}_L(w,n,n',x)\right|=O_p\left(\frac{1}{L^{1-\delta}\cdot\widetilde{h}_L}\right).
\end{aligned}\tag{46}$$

Equations (45A)-(45B) and (46) are analogous to (23A)-(23B) and (24), respectively. Analogously to the unconditional-test case, our main result in this section relies on regularity conditions of the remainder terms $\xi^{-i,j,k}_L$ and $\widetilde{\xi}^{-i,j,k}_L$. Accordingly, we will impose conditions that serve an analogous purpose to those described in Assumptions (T2) and (T2').

Assumption C2

(i) Take a distinct triple i,j,k in $1,\dots,L$ and let $\ell \neq \ell'$, where each denotes either i, j or k. There exists a $\tau > 0$ and a deterministic sequence $\overline{H}_L = O(1)$ such that, for any $n,n' \in \mathscr{N}$,

$$\Pr\Big(-s \leq \Delta_{W|Y}(W_\ell, n, n', X_{\ell'}) < 0 \;\Big|\; \xi_L^{-i,j,k}(W_\ell, n, n', X_{\ell'}),\, (W_\ell, X_{\ell'}) \in \mathscr{C}_{n,n'}\Big) \leq \overline{H}_L \cdot |s| \quad \forall\, 0 < s \leq \tau.$$

(ii) Take any n,n' in $\mathscr{N}$ and any $(w,x) \in \mathscr{C}_{n,n'}$. For $\ell \neq i,j,k$ let

$$\gamma_L^{-i,j,k}(Z_\ell, w, n, n', x) = \xi_L^{-i,j,k}(w,n,n',x) - \xi_L^{-i,j,k,\ell}(w,n,n',x).$$

Take any n,n' in $\mathscr{N}$. There exists a deterministic sequence $\overline{J}_L = O(1)$ such that for any $t \equiv (w,n,n',x)$ where $(w,x) \in \mathscr{C}_{n,n'}$,

$$\left| E\Big[\Upsilon_L^{\Delta}(Z_\ell, t) \;\Big|\; \xi_L^{-i,j,k,\ell}(t) + \beta \cdot \gamma_L^{-i,j,k}(Z_\ell, t)\Big] - E\Big[\Upsilon_L^{\Delta}(Z_\ell, t) \;\Big|\; \xi_L^{-i,j,k,\ell}(t) + \beta' \cdot \gamma_L^{-i,j,k}(Z_\ell, t)\Big] \right|$$
$$\leq \overline{J}_L \cdot |\beta - \beta'| \cdot \Big|\gamma_L^{-i,j,k}(Z_\ell, t)\Big| \quad \forall\, (\beta, \beta') \in [0,1).$$

(iii) There exists V_L such that, for any $i \neq j \neq k$,

$$\sup_{\substack{(n,n') \in \mathscr{N} \\ (w,x) \in \mathscr{C}_{n,n'}}} \Big|\xi_L^{-i,j,k}(w,n,n',x)\Big| \leq V_L, \quad \text{and} \quad \overline{\lim_{L \longrightarrow \infty}} E\big[V_L^4\big] < \infty.$$

Similarly to Assumption (T2), imposing (C2) in this context will help ensure that, for $\ell, \ell' \in i,j,k$ and $\ell \neq \ell'$,

$$\Pr\Big(-\big|\xi_L^{-i,j,k}(W_\ell, n, n', X_{\ell'})\big| \leq \Delta_{W|Y}(W_\ell, n, n', X_{\ell'}) < 0 \;\Big|\; \xi_L^{-i,j,k}(W_\ell, n, n', X_{\ell'}),\, W_\ell \in \mathscr{W}_{n,n'}\Big)$$
$$= O\Big(\big|\xi_L^{-i,j,k}(W_\ell, n, n', X_{\ell'})\big|\Big),$$

and

$$\sup_{\substack{(w,x) \in \mathscr{C}_{n,n'} \\ n,n' \in \mathscr{N}}} \Big| E\big[\Upsilon_L^{\Delta}(Z_\ell, w, n, n', x) \mid \xi_L^{-i,j,k}(w,n,n',x)\big]\Big| = O_p\left(\frac{1}{L^{1-\delta} \cdot \widetilde{h}_L}\right) \quad \forall\, \delta > 0$$

Assumption (C2.iii) includes an existence-of-moments condition that was redundant in Assumption (T2) because, in the latter case, the remainder term $\xi_L^{-i,j}$ was bounded w.p.1. by construction of the nonparametric estimators. This is no longer the case in our current context due to the use of bias-reducing kernels. We extend Assumption (T2') as follows.

Assumption C2'

(i) Take a distinct triple i,j,k in $1,\dots,L$ and let $\ell \neq \ell'$, where each denotes either i, j or k. There exists a $\tau > 0$ and a deterministic sequence $\overline{H}_L = O(1)$ such that, for any $n,n' \in \mathscr{N}$,

$$\Pr\Big(-s \leq \Phi_{W|Y}(W_\ell, n, n', X_{\ell'}) < 0 \;\Big|\; \widetilde{\xi}_L^{-i,j,k}(W_\ell, n, n', X_{\ell'}),\, (W_\ell, X_{\ell'}) \in \mathscr{C}_{n,n'}\Big) \leq \overline{H}_L \cdot |s| \quad \forall\, 0 < s \leq \tau.$$

(ii) Take any n, n' in $\mathscr{N}$ and any $(w,x) \in \mathscr{C}_{n,n'}$. For $\ell \neq i,j,k$ let

$$\widetilde{\gamma}_{L}^{-i,j,k}(Z_\ell, w, n, n', x) = \widetilde{\xi}_{L}^{-i,j,k}(w,n,n',x) - \widetilde{\xi}_{L}^{-i,j,k,\ell}(w,n,n',x).$$

Take any n, n' in $\mathscr{N}$. There exists a deterministic sequence $\overline{J}_{L} = O(1)$ such that for any $t \equiv (w,n,n',x)$ where $(w,x) \in \mathscr{C}_{n,n'}$,

$$\left| E\Big[\Upsilon_{L}^{\Phi}(Z_\ell,t) \;\Big|\; \widetilde{\xi}_{L}^{-i,j,k,\ell}(t) + \beta \cdot \widetilde{\gamma}_{L}^{-i,j,k}(Z_\ell,t)\Big] - E\Big[\Upsilon_{L}^{\Phi}(Z_\ell,t) \;\Big|\; \widetilde{\xi}_{L}^{-i,j,k,\ell}(t) + \beta' \cdot \widetilde{\gamma}_{L}^{-i,j,k}(Z_\ell,t)\Big] \right|$$
$$\leq \overline{J}_{L} \cdot \left|\beta - \beta'\right| \cdot \left|\widetilde{\gamma}_{L}^{-i,j,k}(Z_\ell,t)\right| \quad \forall\ (\beta,\beta') \in [0,1).$$

(iii) There exists V_{L} such that, for any $i \neq j \neq k$,

$$\sup_{\substack{(n,n') \in \mathscr{N} \\ (w,x) \in \mathscr{C}_{n,n'}}} \left|\widetilde{\xi}_{L}^{-i,j,k}(w,n,n',x)\right| \leq V_{L}, \quad \text{and} \quad \overline{\lim_{L \longrightarrow \infty}} E\big[V_{L}^4\big] < \infty.$$

The results that will follow from Assumption (C2') are analogous to those discussed above for the case of (C2). In addition to these regularity conditions about the linear representation of our nonparametric estimators, we will require additional smoothness assumptions, described next.

Assumption C3

For a given n, n' in $\mathscr{N}$ and a given (w,x), let

$$\lambda^F(w,n,n',x) =$$
$$E_W\left[\left(\mathbb{1}\{w \leq W\} - F_{W|Y}(W|n',x)\right) \cdot \mathbb{1}\left\{\Delta_{W|Y}(W,n,n',x) \geq 0\right\} \cdot \mathbb{1}\left\{(x,W) \in \mathscr{C}_{n,n'}\right\}\right],$$
$$\lambda^\Phi(w,n,n',x) =$$
$$E_W\left[\left(\Omega\big(F_{W|Y}(W|n',x),n,n'\big) - \mathbb{1}\{w \leq W\}\right) \cdot \mathbb{1}\left\{\Phi_{W|Y}(W,n,n',x) \geq 0\right\} \cdot \mathbb{1}\left\{(x,W) \in \mathscr{C}_{n,n'}\right\}\right],$$
$$\lambda^\Omega(w,n,n',x) =$$
$$E_W\left[\nabla_1\Omega\big(F_{W|Y}(W|n',x),n,n'\big) \cdot \left(\mathbb{1}\{w \leq W\} - F_{W|Y}(W|n',x)\right) \cdot \mathbb{1}\left\{\Phi_{W|Y}(W,n,n',x) \geq 0\right\} \cdot \mathbb{1}\left\{(x,W) \in \mathscr{C}_{n,n'}\right\}\right] \tag{47}$$

Let M be as described in Assumption (C1). Then, for each n, n' in $\mathscr{N}$ and almost everywhere in $\mathscr{C}_{n,n'}$, the functions $\lambda^F(w,n,n',x)$, $\lambda^\Phi(w,n,n',x)$ and $\lambda^\Omega(w,n,n',x)$ are M times differentiable with respect to x with bounded derivatives.

The functionals described in (C3) will appear in the leading terms of the Hoeffding decompositions of the relevant U-statistics involved in our constructions. As a result of the smoothness requirements in (C3) and (C3') and the bandwidth conditions to be described below, these leading terms will be asymptotically normally distributed at the parametric rate of $\sqrt{L}$.

Assumption C4

Let M be the constant described in Assumptions (C1) and (C3). Both $K(\cdot)$ and $\widetilde{K}(\cdot)$ are Lipschitz-continuous, bounded and symmetric (around zero) bias-reducing kernels of order M. Each has compact support of the form $[-a,a]$ (the support of K can differ from that of $\widetilde{K}$), and they are both bounded by some constant $\overline{K}$. The sequences b_L, h_L and $\widetilde{h}_L$ satisfy

$$L^{\frac{1}{2}}\cdot b_L\cdot\widetilde{h}_L\cdot h_L\longrightarrow\infty,\quad \frac{L^{\frac{1}{2}}\cdot b_L^2}{\widetilde{h}_L^{\frac{1}{2}}\cdot h_L^2}\longrightarrow 0,\quad \left(\frac{\widetilde{h}_L^M}{h_L}\right)\cdot L^{1/2}\longrightarrow 0\quad\text{and}\quad\left(\frac{h_L^M}{\widetilde{h}_L}\right)\cdot L^{1/2}\longrightarrow 0.$$

Values of M compatible with Assumption (C4) can be as low as $M=8$. In this case ($M=8$), our bandwidth convergence conditions would be satisfied, for instance, if we let $\epsilon=0.0001$ and we set $b_L\propto L^{-c_1}$, $h_L\propto L^{-c_2}$ and $\widetilde{h}_L\propto L^{-c_3}$, with $c_1=\frac{1}{2}\cdot\left(\frac{3}{4}-\frac{3}{8(M-1)}-\frac{3}{4}(M-1)\cdot\epsilon\right)\approx 0.3479$, $c_2=\frac{1}{2(M-1)}+\epsilon\approx 0.0715$, $c_3=\frac{1}{2(M-1)}+(M-1)\cdot\epsilon\approx 0.0721$. It follows that the various functionals described in Assumptions (C1) and (C3) would have to be $M=8$ times differentiable, but we remind the reader that these smoothness restrictions can be violated at any subset of our testing range with measure zero.

B.6.2 Main Result

Let

$$U_{L^{(3)}}^{\widehat{S}^F_{L_{n,n'}}}=\frac{1}{L(L-1)(L-2)}\sum\sum\sum_{\substack{i,j,k\in\{1,\ldots,L\}\\ i\neq j\neq k}}\widehat{S}^F_{L_{n,n'}}(Z_i,Z_j,Z_k),\quad U_{L^{(3)}}^{\widehat{S}^\Omega_{L_{n,n'}}}=\frac{1}{L(L-1)(L-2)}\sum\sum\sum_{\substack{i,j,k\in\{1,\ldots,L\}\\ i\neq j\neq k}}\widehat{S}^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k)$$

By (43) we have

$$U_{L^{(3)}}^{\widehat{S}^F_L}=\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}U_{L^{(3)}}^{\widehat{S}^F_{L_{n,n'}}}\quad\text{and}\quad U_{L^{(3)}}^{\widehat{S}^\Omega_L}=\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}U_{L^{(3)}}^{\widehat{S}^\Omega_{L_{n,n'}}}.$$

We study the asymptotic properties of the statistics $U_{L^{(3)}}^{\widehat{S}^F_L}$ and $U_{L^{(3)}}^{\widehat{S}^\Omega_L}$ by characterizing those of $U_{L^{(3)}}^{\widehat{S}^F_{L_{n,n'}}}$ and $U_{L^{(3)}}^{\widehat{S}^\Omega_{L_{n,n'}}}$.

Theorem 7 *Let $\gamma^F_{n,n'}$ and $\gamma^\Omega_{n,n'}$ be as defined in Theorem 6.*

(i) If Assumptions (C1), (C2), (C3) and (C4) are satisfied, then

$$\sqrt{L}\cdot U_{L^{(3)}}^{\widehat{S}^F_{L_{n,n'}}}=\sqrt{L}\cdot\gamma^F_{n,n'}+\frac{1}{\sqrt{L}}\sum_{i=1}^{L}\phi^F_{n,n'}(Z_i)+o_p(1),\quad\textit{where}\quad E\left[\phi^F_{n,n'}(Z_i)\right]=0.$$

From (43), *it follows by construction that*

$$\sqrt{L}\cdot U_{L^{(3)}}^{\widehat{S}^F_L}=\sqrt{L}\cdot\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\gamma^F_{n,n'}\right]+\frac{1}{\sqrt{L}}\sum_{i=1}^{L}\phi^F_{\mathscr{N}}(Z_i)+o_p(1),\quad\textit{where}\quad\phi^F_{\mathscr{N}}(Z_i)=\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\phi^F_{n,n'}(Z_i)\right].\tag{48}$$

For each n, n' in $\mathscr{N}$, if (39) is satisfied with probability one in $\mathscr{C}_{n,n'}$, the function $\phi^F_{n,n'}$ reduces to

$$\phi^F_{n,n'}(Z_i) = \mathbb{1}\{N_i = n\}\cdot\lambda^F(W_i, n, n', X_i)\cdot f_X(X_i) - \frac{p_Y(n, X_i)}{p_Y(n', X_i)}\cdot\mathbb{1}\{N_i = n'\}\cdot\lambda^F(W_i, n, n', X_i)\cdot f_X(X_i) \quad (49)$$

(ii) If Assumptions (C1), (C2'), (C3) and (C4) are satisfied, then

$$\sqrt{L}\cdot U^{\widehat{S}^{\Omega}_{L_{n,n'}}}_{L^{(3)}} = \sqrt{L}\cdot\gamma^{\Omega}_{n,n'} + \frac{1}{\sqrt{L}}\sum_{i=1}^{L}\phi^{\Omega}_{n,n'}(Z_i) + o_p(1), \quad \textit{where} \quad E\big[\phi^{\Omega}_{n,n'}(Z_i)\big] = 0.$$

From (43)*, it follows by construction that*

$$\sqrt{L}\cdot U^{\widehat{S}^{\Omega}_L}_{L^{(3)}} = \sqrt{L}\cdot\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\gamma^{\Omega}_{n,n'}\right] + \frac{1}{\sqrt{L}}\sum_{i=1}^{L}\phi^{\Omega}_{\mathscr{N}}(Z_i) + o_p(1), \quad \textit{where} \quad \phi^{\Omega}_{\mathscr{N}}(Z_i) \quad = \left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\phi^{\Omega}_{n,n'}(Z_i)\right]. \quad (50)$$

For each n, n' in $\mathscr{N}$, if (40) *is satisfied with probability one in $\mathscr{C}_{n,n'}$, the function $\phi^{\Omega}_{n,n'}$ reduces to*

$$\phi^{\Omega}_{n,n'}(Z_i) = \mathbb{1}\{N_i = n\}\cdot\lambda^{\Phi}(W_i, n, n', X_i)\cdot f_X(X_i) + \frac{p_Y(n, X_i)}{p_Y(n', X_i)}\cdot\mathbb{1}\{N_i = n'\}\cdot\lambda^{\Omega}(W_i, n, n', X_i)\cdot f_X(X_i) \quad (51)$$

The functions $\phi^F_{n,n'}(Z_i)$ and $\phi^{\Omega}_{n,n'}(Z_i)$ have mean-zero regardless of whether (39) or (40) are satisfied. Once again, these functions appear as the leading terms of the various Hoeffding decompositions involved. Their general structure is provided in Equations (T-72) and (T-79) of the Technical Supplement. Theorem 7 characterizes these expressions for the case where (39) and (40) are satisfied w.p.1. in $\mathscr{C}_{n,n'}$. In this instance, using the definitions in (47) and iterated expectations, it is easy to verify that $E\big[\mathbb{1}\{N_i = n\}\cdot\lambda^F(W_i, n, n', X_i)\cdot f_X(X_i)\big] = \gamma^F_{n,n'} = 0$ and $E\Big[\frac{p_Y(n,X_i)}{p_Y(n',X_i)}\cdot\mathbb{1}\{N_i = n'\}\cdot\lambda^F(W_i, n, n', X_i)\cdot f_X(X_i)\Big] = 0$, which yields $E\big[\phi^F_{n,n'}(Z_i)\big] = \gamma^F_{n,n'} = 0$. Likewise, iterated expectations yields $E\big[\mathbb{1}\{N_i = n\}\cdot\lambda^{\Phi}(W_i, n, n', X_i)\cdot f_X(X_i)\big] = \gamma^{\Omega}_{n,n'} = 0$ and $E\Big[\frac{p_Y(n,X_i)}{p_Y(n',X_i)}\cdot\mathbb{1}\{N_i = n'\}\cdot\lambda^{\Omega}(W_i, n, n', X_i)\cdot f_X(X_i)\Big] = 0$, and therefore $E\big[\phi^{\Omega}_{n,n'}(Z_i)\big] = \gamma^{\Omega}_{n,n'} = 0$.

B.6.3 A Rejection Rule based on Theorem 7

The implications of Theorem 7 are reminiscent those of the unconditional-test case. Combined with Theorem 6, we have that $\sqrt{L}\cdot U^{\widehat{S}^F_L}_{L^{(3)}}$ diverges w.p.1 to $+\infty$ if (39) is violated with positive probability in $\mathscr{C}_{n,n'}$; it vanishes in probability to zero if (39) is satisfied as a strict inequality w.p.1. in $\mathscr{C}_{n,n'}$, and it converges in distribution to a normal random variable with mean zero and variance $\mathrm{Var}\big(\phi^F_{n,n'}(Z_i)\big) > 0$ if (39) is satisfied w.p.1. in each $\mathscr{C}_{n,n'}$ *and* is binding as an equality with nonzero probability in some $\mathscr{C}_{n,n'}$. There is an equivalent relationship between $\sqrt{L}\cdot U^{\widehat{S}^{\Omega}_L}_{L^{(3)}}$ and (40).

Let $\widehat{\phi}^F_{n,n'}(Z_i)$ and $\widehat{\phi}^{\Omega}_{n,n'}(Z_i)$ be estimators of the expressions given in (49) and (51). It follows from Theorem 7 that, for any pair of arbitrary constants $c_1 > 0$ and $c_2 > 0$, rejection rules like those described in (16A) and (16B) based on

$$\frac{\sqrt{L}\cdot U_{L^{(3)}}^{\widehat{S}^F}}{\sqrt{\widehat{\text{Var}}\left[\sum\limits_{n>n'}\sum \widehat{\phi}^F_{n,n'}(Z_i)\right]+c_1}} \quad \text{and} \quad \frac{\sqrt{L}\cdot U_{L^{(3)}}^{\widehat{S}^\Omega}}{\sqrt{\widehat{\text{Var}}\left[\sum\limits_{n>n'}\sum \widehat{\phi}^\Omega_{n,n'}(Z_i)\right]+c_2}}, \tag{52}$$

would satisfy the asymptotic properties described in (17) for our testing range. Let Υ^F_L and Υ^Ω_L be as defined in (44). Let

$$\widehat{G}^F_{L_{n,n'}}(Z_i,Z_j,Z_k,Z_\ell) =$$
$$\Upsilon^F_L(Z_i,W_\ell,n',X_k)\cdot \mathbb{1}\{N_j=n\}\cdot \mathbb{1}\left\{\widehat{\Delta}^{-i,j,k,\ell}_{W|Y}(W_\ell,n,n',X_k)\geq -b_L\right\}\cdot \mathbb{1}\left\{(X_k,W_\ell)\in \mathscr{C}_{n,n'}\right\}\cdot \frac{1}{h_L}K\left(\frac{X_j-X_k}{h_L}\right),$$
$$\widehat{G}^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k,Z_\ell) =$$
$$\Upsilon^\Omega_L(Z_i,W_\ell,n,n',X_k)\cdot \mathbb{1}\{N_j=n\}\cdot \mathbb{1}\left\{\widehat{\Phi}^{-i,j,k,\ell}_{W|Y}(W_\ell,n,n',X_k)\geq -b_L\right\}\cdot \mathbb{1}\left\{(X_k,W_\ell)\in \mathscr{C}_{n,n'}\right\}\cdot \frac{1}{h_L}K\left(\frac{X_j-X_k}{h_L}\right).$$

We can estimate $\phi^F_{n,n'}$ and $\phi^\Omega_{n,n'}$ as described in (49) and (51) by using

$$\begin{aligned}
\widehat{\phi}^F_{n,n'}(Z_i) &= \frac{1}{(L-1)(L-2)}\sum_{j\neq i}\sum_{\substack{k\neq j\\k\neq i}}\widehat{S}^F_{L_{n,n'}}(Z_i,Z_j,Z_k) - \frac{1}{(L-1)(L-2)(L-3)}\sum_{j\neq i}\sum_{\substack{k\neq j\\k\neq i}}\sum_{\substack{\ell\neq k\\ \ell\neq j\\ \ell\neq i}}\widehat{G}^F_{L_{n,n'}}(Z_i,Z_j,Z_k,Z_\ell)\\
\widehat{\phi}^\Omega_{n,n'}(Z_i) &= \frac{1}{(L-1)(L-2)}\sum_{j\neq i}\sum_{\substack{k\neq j\\k\neq i}}\widehat{S}^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k) + \frac{1}{(L-1)(L-2)(L-3)}\sum_{j\neq i}\sum_{\substack{k\neq j\\k\neq i}}\sum_{\substack{\ell\neq k\\ \ell\neq j\\ \ell\neq i}}\widehat{G}^\Omega_{L_{n,n'}}(Z_i,Z_j,Z_k,Z_\ell)
\end{aligned} \tag{53}$$

The resulting estimated variances to be used in (52), $\widehat{\text{Var}}\left[\sum\limits_{n>n'}\sum \widehat{\phi}^F_{n,n'}(Z_i)\right]$ and $\widehat{\text{Var}}\left[\sum\limits_{n>n'}\sum \widehat{\phi}^\Omega_{n,n'}(Z_i)\right]$, are consistent under the conditions of Theorem 7.

Remark 1 (Extension to multivariate X). Our methodology can be extended to the multivariate $\mathbf{X}$ case, with both discrete and continuous covariates. Our nonparametric estimators would use indicator functions for the discrete components of $\mathbf{X}$, and a multivariate (e.g, multiplicative) bias-reducing kernel for the continuous elements. The value of M needed to preserve $\sqrt{L}-$consistency and asymptotic normality of the resulting test-statistic would increase with the number of continuously distributed elements in $\mathbf{X}$. To be precise, let d denote the number of continuous elements in $\mathbf{X}$. We can show that the bandwidth convergence conditions analogous to Assumption (C4) that correspond to the multidimensional case would be satisfied if $M > \frac{1}{4}\cdot\left(8+5\cdot d+\sqrt{(8+5\cdot d)^2+120\cdot d}\right)$. Thus, for instance if $d=2$ we need $M\geq 11$, if $d=3$, $M\geq 14$, if $d=4$, $M\geq 16$, and so on. The use of bias-reducing kernels of these (and higher) orders is highly feasible computationally. Furthermore, as we pointed out in the one-dimensional case studied above, even though higher values of M would in turn imply a higher degree of smoothness for the relevant distributions and functionals involved, this feature can be violated at any subset of our testing range that has measure zero.

B.6.4 A Test for the Reverse Inequality in (40) and Its Asymptotic Properties

Failing to reject either (39) or (40) could be attributed to the data being consistent with IPV conditional on X. As we did in the unconditional case (see Appendix B.4.4), to check whether this is the case we can test whether the reverse inequality in (40) holds: namely, whether

$$\psi_{n-1:n}^{-1}\left(F_{W|Y}(w|n,x)\right) \leq \psi_{n'-1:n'}^{-1}\left(F_{W|Y}(w|n',x)\right) \tag{40'}$$

for every $n, n' \in \mathscr{N}$ with $n > n'$ and almost all $(w,x) \in \mathscr{C}_{n,n'}$. If the data supports (40), then rejecting (40') implies that the inequalities in (40) are strict with positive probability, which rules out IPV (conditional on X) as the true model. A test of (40') would replace $S^{\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)$ with

$$\begin{aligned} S^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k) &= \Big(\mathbb{1}\{W_i \leq W_k\} - \Omega\big(F_{W|Y}(W_k|n',X_j),n,n'\big)\Big)\cdot \mathbb{1}\{N_i = n\} \\ &\quad \cdot \mathbb{1}\big\{-\Phi_{W|Y}(W_k,n,n',X_j) \geq 0\big\}\mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\}\cdot \frac{1}{h_L}K\Big(\frac{X_i - X_j}{h_L}\Big). \end{aligned}$$

We estimate this function with

$$\begin{aligned} \widehat{S}^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k) &= \Big(\mathbb{1}\{W_i \leq W_k\} - \Omega\big(\widehat{F}^{-i,j,k}_{W|Y}(W_k|n',X_j),n,n'\big)\Big)\cdot \mathbb{1}\{N_i = n\} \\ &\quad \cdot \mathbb{1}\big\{-\widehat{\Phi}^{-i,j,k}_{W|Y}(W_k,n,n',X_j) \geq -b_L\big\}\mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\}\cdot \frac{1}{h_L}K\Big(\frac{X_i - X_j}{h_L}\Big) \end{aligned}$$

Let

$$\begin{aligned} U^{S^{-\Omega}_L}_{L^{(3)}} &= \frac{1}{L(L-1)(L-2)}\sum\sum\sum_{\substack{i,j,k\in\{1,\ldots,L\}\\ i\neq j\neq k}}\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}} S^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)\right], \\ U^{\widehat{S}^{-\Omega}_L}_{L^{(3)}} &= \frac{1}{L(L-1)(L-2)}\sum\sum\sum_{\substack{i,j,k\in\{1,\ldots,L\}\\ i\neq j\neq k}}\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}} \widehat{S}^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)\right] \end{aligned} \tag{54}$$

For a given n, n' in $\mathscr{N}$ and a given (w,x), let

$$\lambda^{-\Phi}(w,n,n',x) =$$
$$E_W\left[\big(\mathbb{1}\{w \leq W\} - \Omega\big(F_{W|Y}(W|n',x),n,n'\big)\big)\cdot \mathbb{1}\big\{-\Phi_{W|Y}(W,n,n',x)\geq 0\big\}\cdot \mathbb{1}\big\{(x,W)\in\mathscr{C}_{n,n'}\big\}\right],$$
$$\lambda^{-\Omega}(w,n,n',x) =$$
$$E_W\left[\nabla_1\Omega\big(F_{W|Y}(W|n',x),n,n'\big)\cdot\big(\mathbb{1}\{w\leq W\} - F_{W|Y}(W|n',x)\big)\cdot \mathbb{1}\big\{-\Phi_{W|Y}(W,n,n',x)\geq 0\big\}\cdot \mathbb{1}\big\{(x,W)\in\mathscr{C}_{n,n'}\big\}\right]$$

Let

$$\gamma^{-\Omega}_{n,n'} = p_N(n)\cdot E\Big[\max\big\{0, -\Phi_{W|Y}(W_k,n,n',X_j)\big\}\cdot f_{X|N}(X_j|n)\cdot \mathbb{1}\big\{(X_j,W_k)\in\mathscr{C}_{n,n'}\big\}\Big]$$

Under the type of smoothness and bandwidth-convergence conditions of Theorem 7, we have $E\big[\widehat{S}^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)\big] = \gamma^{-\Omega}_{n,n'} + o(L^{-1/2})$. Note that $\gamma^{-\Omega}_{n,n'} \geq 0$. Having $\gamma^{-\Omega}_{n,n'} > 0$ for some $n > n'$ in $\mathscr{N}$ would indicate that (40) holds as a *strict* inequality with positive probability in $\mathscr{C}_{n,n'}$, leading us to reject the notion that, conditional on X, the data generating process is consistent with IPV. Let

$$U^{\widehat{S}^{-\Omega}_{L_{n,n'}}}_{L^{(3)}} = \frac{1}{L(L-1)(L-2)}\sum\sum\sum_{\substack{i,j,k\in\{1,\ldots,L\}\\ i\neq j\neq k}}\widehat{S}^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k).$$

Under the type of conditions leading to part (ii) of Theorem 7, we can show that

$$\sqrt{L}\cdot U_{L^{(3)}}^{\widehat{S}^{-\Omega}_{L_{n,n'}}} = \sqrt{L}\cdot\gamma^{-\Omega}_{n,n'} + \frac{1}{\sqrt{L}}\sum_{i=1}^{L}\phi^{-\Omega}_{n,n'}(Z_i) + o_p(1), \quad \text{where} \quad E\big[\phi^{-\Omega}_{n,n'}(Z_i)\big] = 0.$$

By construction in (54), it follows that

$$\sqrt{L}\cdot U_{L^{(3)}}^{\widehat{S}^{-\Omega}_{L}} = \sqrt{L}\cdot\left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\gamma^{-\Omega}_{n,n'}\right] + \frac{1}{\sqrt{L}}\sum_{i=1}^{L}\phi^{-\Omega}_{\mathscr{N}}(Z_i) + o_p(1), \quad \text{where} \quad \phi^{-\Omega}_{\mathscr{N}}(Z_i) \;=\; \left[\sum\sum_{\substack{n,n'\in\mathscr{N}\\ n>n'}}\phi^{-\Omega}_{n,n'}(Z_i)\right]. \tag{55}$$

For each n, n' in $\mathscr{N}$, if the reverse inequality in (40) holds w.p.1. in $\mathscr{C}_{n,n'}$, the function $\phi^{-\Omega}_{n,n'}$ reduces to

$$\phi^{-\Omega}_{n,n'}(Z_i) = \mathbb{1}\{N_i = n\}\cdot\lambda^{-\Phi}(W_i,n,n',X_i)\cdot f_X(X_i) - \frac{p_Y(n,X_i)}{p_Y(n',X_i)}\cdot\mathbb{1}\{N_i = n'\}\cdot\lambda^{-\Omega}(W_i,n,n',X_i)\cdot f_X(X_i) \tag{56}$$

Let Υ^{Ω}_{L} be as defined in (44). Define

$$\widehat{G}^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k,Z_\ell) =$$

$$\Upsilon^{\Omega}_{L}(Z_i,W_\ell,n,n',X_k)\cdot\mathbb{1}\{N_j = n\}\cdot\mathbb{1}\big\{-\widehat{\Phi}^{-i,j,k,\ell}_{W|Y}(W_\ell,n,n',X_k)\geq -b_L\big\}\cdot\mathbb{1}\big\{(X_k,W_\ell)\in\mathscr{C}_{n,n'}\big\}\cdot\frac{1}{h_L}K\left(\frac{X_j - X_k}{h_L}\right).$$

We can estimate $\phi^{-\Omega}_{n,n'}$ as described in (56) by using

$$\widehat{\phi}^{-\Omega}_{n,n'}(Z_i) = \frac{1}{(L-1)(L-2)}\sum_{j\neq i}\sum_{\substack{k\neq j\\ k\neq i}}\widehat{S}^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k) - \frac{1}{(L-1)(L-2)(L-3)}\sum_{j\neq i}\sum_{\substack{k\neq j\\ k\neq i}}\sum_{\substack{\ell\neq k\\ \ell\neq j\\ \ell\neq i}}\widehat{G}^{-\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k,Z_\ell) \tag{57}$$

B.6.5 Application of the Conditional Test

We apply our test, conditional on X = appraisal value, to our timber data. We use $\mathscr{N} = \{2,3,\ldots,11\}$ and $\mathscr{C}_{n,n'} = \big\{(x,w) : 0.02 \leq F_{W|Y}(w|m,x) \leq 0.98 \text{ for } m = n,n'\big\}$ as our testing range. We use $\mathbb{1}\big\{0.02 \leq \widehat{F}^{-i,j,k}_{W|Y}(W_k|n,X_j) \leq 0.98\big\}\cdot\mathbb{1}\big\{0.02 \leq \widehat{F}^{-i,j,k}_{W|Y}(W_k|n',X_j) \leq 0.98\big\}$ to estimate $\mathbb{1}\{(X_j,W_k)\in\mathscr{C}_{n,n'}\}$ in the construction of $\widehat{S}^{F}_{L_{n,n'}}(Z_i,Z_j,Z_k)$ and $\widehat{S}^{\Omega}_{L_{n,n'}}(Z_i,Z_j,Z_k)$. We employed polynomial kernels with bounded support of form $K(z) = \widetilde{K}(z) = \big(a_0 + \sum_{j=1}^{7}a_j\cdot z^{2j}\big)\cdot\mathbb{1}\big\{-C_k \leq z \leq C_k\big\}$. The a_j's are chosen to satisfy the conditions $\int_{-C_k}^{C_k}K(z)dz = 1$ and $\int_{-C_k}^{C_k}z^{2j}\cdot K(z)dz = 0$ for $j = 1,\ldots,6$. Since our kernel is symmetric around zero, it guarantees $\int_{-C_k}^{C_k}z^{r}\cdot K(z)dz = 0$ for any odd r. In our results we used $C_k = 8$, which yields a support large enough to include most of the probability mass, e.g., of a standard normal kernel. Using polynomial kernels helped simplify the computation of our test-statistics significantly. The bandwidths we used are of the type $b_L = 0.015\cdot L^{-0.344}$, $h_L = d_1\cdot\widehat{\sigma}_X\cdot L^{-0.073}$, and $\widetilde{h}_L = d_2\cdot\widehat{\sigma}_X\cdot L^{-0.073}$, where $\widehat{\sigma}_X$ denotes the sample standard deviation of X (appraisal value). The expression for b_L was chosen so that $b_L \approx 0.001$ given our sample size. Combined with the exponents used for the remaining bandwidths, the conditions in Assumption (C4) are satisfied for $M = 16$. The constants of proportionality d_1 and d_2 were fixed at 2 in our implementation.

The test-statistics for (39) and (40) were constructed using the analytic expressions for variances described in Equations (52)-(53) in Appendix B.6.3, with $c_1 = c_2 = 10^{-6}$. We also implement the IPV test described in Appendix B.6.4, using the analytic expressions for the variance given in Equations (56)-(57) with $c = 10^{-6}$. As we can see in those equations, constructing the analytic variance estimators requires the use of a fourth-order U-statistic. Given our sample size, this would amount to a sum with approximately 2000^4 terms. In order to make the task of estimating this variance computationally feasible, we chose 20 random subsamples of size $L = 500$ and computed the corresponding estimate of the variance for each subsample. Our final estimate was computed as the average of these 20 estimates.

References

Andrews, D. (1994). Empirical process methods in econometrics. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics, Vol. 4*. North Holland.

Andrews, D. and X. Shi (2009). Inference based on conditional moment inequality models. Working Paper. Yale University.

Aradillas-Lopez, A., B. Honoré, and J. Powell (2007). Pairwise difference estimation with nonparametric control variables. *International Economic Review 48*(4), 1119–1158.

Athey, S., P. Crampton, and A. Ingraham (2003). Upset Pricing in Auction Markets: An Overview. White Paper, Market Design Inc., on behalf of British Columbia Ministry of Forests.

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 (2008). Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions. NBER Working Paper W14590.

Bajari, P. and A. Hortacsu (2003). The winner's curse, reserve prices, and endogenous entry: Empirical insights from eBay auctions. *The RAND Journal of Economics 34*(2), 329–355.

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.

Campo, S., E. Guerre, I. Perrigne, and Q. Vuong (2002). Semiparametric Estimation of First-Price Auctions with Risk Averse Bidders. Working Paper.

Chen, X. and Y. Fan (1999). Consistent hypothesis testing in semiparametric and nonparametric models for econometric time series. *Journal of Econometrics 91*(2), 373–401.

Collomb, G. and W. Hardle (1986). Strong uniform convergence rates in robust nonparametric time series analysis and prediction : Kernel regression from dependent observations. *Stochastic Processes and their Applications 23*, 77–89.

Dominguez, M. and I. Lobato (2004). Consitent estimation of models defined by conditional moment restrictions. *Econometrica 72*(5), 1601–1615.

Fan, Y. and Q. Li (1996). Consistent model specification tests: Omitted variables and semiparametric functional forms. *Econometrica 64*(4), 865–890.

Froeb, L. and P. McAfee (1988). Deterring Bid Rigging in Forest Service Timber Auctions. US Department of Justice, mimeo.

Gillen, B. (2009). Identification and Estimation of Level-k Auctions. Working Paper.

Guerre, E., I. Perrigne, and Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica 68*(3), 525–574.

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., 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.

Halmos, P. (1946). The theory of unbiased estimation. *Annals of Mathematical Statistics 17*, 34–43.

Han, A. (1987). Non-parametric analysis of a generalized regression model the maximum rank correlation estimator. *Journal of Econometrics 35*(2-3), 303–316.

Hansen, B. (2004). Nonparametric estimation of smooth conditional distributions. Working Paper.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics 19*, 293–325.

Hoeffding, W. (1961). The strong law of large numbers for u-statistics. Inst. Statist. Mimeo Series.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association 1*(2), 13–30.

Hong, H. and M. Shum (2002). Imcreasing competition and the winner's curse: Evidence from procurement. *Review of Economic Studies 69*(4), 871–898.

Hong, H. and M. Shum (2009). Pairwise difference estimator of a dynamic optimization model. *Review of Economic Studies, forthcoming*.

Honoré, B. and J. Powell (1994). Pairwise difference estimators of censored and truncated regression models. *Journal of Econometrics 64*(1-2), 241–278.

Honoré, B. and J. Powell (2005). Pairwise difference estimation in nonlinear models. In D. Andrews and J. Stock (Eds.), *Identification and Inference in Econometric Models. Essays in Honor of Thomas Rothenberg*, pp. 520–553. Cambridge University Press.

Hu, Y., D. McAdams, and M. Shum (2009). Nonparametric identification of auction models with non-separable unobserved heterogeneity. Working Paper.

Khan, S. and E. Tamer (2009). Inference on endogenously censored regression models using conditional moment inequalities. *Journal of Econometrics 152*(2), 104–119.

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

Krasnokutskaya, E. (2009). Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity. Forthcoming, *Review of Economic Studies.*

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

Laffont, J.-J., H. Ossard, and Q. Vuong (1995). Econometrics of first-price auctions. *Econometrica 63*(4), 953–980.

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

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

Levin, D. and J. Smith (1996). Optimal reservation prices in auctions. *The Economic Journal 106*(438), 1271–1283.

Lewbel, A. (1997). Semiparametric estimation of location and other discrete choice moments. *Econometric Theory 13*(1), 32–51.

Li, T., I. Perrigne, and Q. Vuong (2000). Conditionally independent private information in OCS wildcat auctions. *Journal of Econometrics 98*(1), 129–161.

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.

Paarsch, H. (1997). Deriving an estimate of the optimal reserve price: an application to British Columbian timber sales. *Journal of Econometrics 78*(2), 333–357.

Paarsch, H. and H. Hong (2006). *An Introduction to the Structural Econometrics of Auction Data*. The MIT Press.

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

Perrigne, I. (2003). Random Reserve Prices and Risk Aversion in Timber Sale Auctions. Working Paper.

Quint, D. (2008). Unobserved correlation in private-value ascending auctions. *Economics Letters 100*(3), 432–434.

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. (1993). The limiting distribution of the maximum rank correlation estimator. *Econometrica 61*(1), 123–137.

van der Vaart, A. and J. Wellner (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag.

Zheng, J. (1998). Consistent specification testing for conditional symmetry. *Econometric Theory 14*(1), 139–149.