1 Introduction

Applying the insights from auction theory to real world settings requires an empirical knowledge of the primitives that define the game being played by bidders. In the case of private value auctions, a key primitive of interest is the latent distribution of bidder valuations in the underlying population. In ascending price auctions (the most commonly found format), this is in fact the only primitive of interest: the dominance solvability of the game renders the remaining characteristics of the bidders, such as their risk preferences or strategic sophistication, effectively moot. The distribution of private values represents the demand curve facing the seller; once demand is empirically known, a seller can determine the optimal auction design parameters (such as the optimal reserve price), forecast expected revenue, and so on.

In an ideal world, if all auctions in the data are identical, all bidders independently bid their private values, and all bids are observable, then the identification of demand is trivial - the distribution of bids identifies the distribution of private values. Unfortunately, the actual data generating process represented by real-world auctions differs from this ideal in three important ways: (i) auctions in the data may be heterogenous, (ii) not all bidders can be presumed to bid their actual values, and (iii) not all bids are necessarily observed.

The empirical literature on auctions has made remarkable progress on problems (ii) and (iii), but much less on problem (i). With a few notable exceptions (see below), the literature has largely assumed that, conditional on a vector of observed covariates $Z$, all auctions in a particular data set are identical, and bidder valuations can therefore be assumed to be identically and independently distributed. With this assumption, it is possible to identify and estimate the distribution of private values $F_V(\cdot | Z)$ in a way that solves problem (ii) (strategic bidding) and problem (iii) (censored data, such as having data on only the transaction price). Results along these lines were attained for the standard auction formats by [Guerre, Perrigne, and Vuong (2000), Athey and Haile (2007), and Haile and Tamer (2003)].
In this project, we seek to enrich the empirical approach to auctions by confronting issue (i), the heterogeneity of auctions. That is, we will dispense with the usual assumption that all auctions are identical conditional on observable covariates, and allow for the presence of unobserved heterogeneity across the auctions in the data. What we aim to show is that unobserved heterogeneity has significant effects for the analysis of auction data, and in particular for the inference of optimal reserve policy. Unobserved heterogeneity induces correlation of the private values among bidders in an auction, and reserve prices in a correlated environment entail (potentially significant) costs that would be underestimated under the standard assumption of independent private values.

Insofar as the existing literature has addressed the issue of unobserved heterogeneity, it has modeled this heterogeneity as being driven by a scalar unobservable that affects an agent’s private value in a parametric or monotonic way (Krasnokutskaya (2009), Hu, McAdams, and Shum (2009)). By restricting unobserved heterogeneity to have this exact form, 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 joint distribution of private values among bidders in an auction. Moreover, this identification is attained without requiring variation in the number of bidders across auctions in the data.

In contrast, we will take a fundamentally different approach to the problem of unobserved heterogeneity. We do not place any restrictions on the dimensionality of the unobservables or their distribution, and do not make any assumptions about how the unobservables affect agents’ private values. Not surprisingly, it will be impossible to identify the joint distribution of private values in an auction without any restrictions. However, despite the fact that the joint distribution of private values is fundamentally underidentified, we show there exist important features of the joint distribution that are identified by variation across auctions in the number of bidders. Moreover, the features we identify turn out to be sufficient for the purposes of the mechanism design problem.

The crucial assumption we make to achieve identification is that the number of bidders participating in a particular auction is independent of the unobserved characteristics of that auction. (As we discuss below, this same assumption has been exploited in other recent papers on first-price auctions, to identify other sources of variation.) While this assumption of exogenous participation can potentially be defended on a variety of grounds, it remains a strong identifying assumption. In the first part of this paper, we develop an informative test of the assumption of exogenous participation. The test takes the form of an inequality relating the distributions of winning bids across auctions of different sizes, and generalizes the nonparametric tests for independent private values in ascending auctions proposed by Athey and Haile (2002) and Haile and Tamer (2003). In addition, we introduce a test statistic that is applicable to this set of inequalities, and demonstrate its asymptotic normality. We thus construct (for the first time to our knowledge) a precise nonparametric test of the standard assumptions used in auction theory (private values and exogenous entry) that is robust to unobserved heterogeneity.
In the second part of this paper, we use the assumptions tested in the first part (private values and exogenous auction size) to nonparametrically identify key auction primitives which would normally be unidentified in the presence of unobserved heterogeneity. We are able to construct upper and lower bounds on expected revenue and the revenue-maximizing reserve price. We then apply these methods to data from U.S. Forest Service timber auctions. Our preliminary analysis suggests that accounting for unobserved heterogeneity rationalizes a much lower reserve price than analysis under the assumption of independence; this may answer the puzzle found in the existing literature of why reserve prices in timber auctions have historically been set so low.

2 Model

In this section, we introduce our model – symmetric, conditionally independent private values, henceforth CIPV – and give an overview of our results.

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. Each bidder \( i \) in an auction with \( n \) bidders has a valuation \( V_i \) which is his or her private information. Bidder \( i \) gets payoff \( V_i - P \) from winning the auction, where \( P \) is the price paid, and 0 from losing.

Bidders’ valuations are assumed to be symmetrically distributed and conditionally independent. Each auction is characterized by a set of characteristics \( \theta \in \Theta \). Bidder valuations for the object for sale in an auction with characteristics \( \theta \) are i.i.d. draws from a probability distribution \( F(\cdot|\theta) \). Thus, variation in \( \theta \) induces correlation in bidders’ valuations, while any variation not caused by \( \theta \) is idiosyncratic and independent across bidders.

Let \( V_{k:n} \) be the \( k \)th lowest bidder valuation in auctions with \( n \) bidders, and \( F_{k:n} \) its probability distribution. Depending on the modeling assumptions, ascending auctions are typically assumed to reveal, or at least provide bounds on, the second-highest valuation, \( V_{n-1:n} \). For expositional ease, both our test and our identification results will be predicated on the realization of \( V_{n-1:n} \) being exactly revealed by the transaction price in each auction\(^2\). All of our results extend easily to the weaker assumptions of Haile and Tamer (2003), in which bids provide bounds on \( V_{n-1:n} \)\(^3\).

Finally, we define our key assumption, which we will be testing and using for identification:

\(^1\)These can include both characteristics of the object for sale, and of the auction more generally (reliability of a particular seller, shipping costs being charged, etc.)

\(^2\)This would be exactly true in a second-price sealed-bid auction, or in a so-called “button auction,” where bidders hold down a button to remain active and release it to drop out. It is approximately true (up to the minimum bid increment) under the bidding assumptions of Haile and Tamer (2003) in an ascending auction with no “jump bids,” and is a common assumption when modeling ascending auctions. In our data, the winning bid is on average just 2% higher than the second-highest, so this is a fairly good approximation.

\(^3\)Specifically, \( V_{n-1:n} \) is assumed to be bounded below by the highest losing bid, and above by the lowest bid available when the auction ended (the winning bid plus the minimum bid increment). These assumptions lead to upper and lower pointwise
Definition 1 There is exogenous variation in auction size (or exogenous variation in N) if the joint distribution of the valuations of any set of k bidders taken from an auction with n bidders is the same for any n ≥ k.

Really, what this is saying is that the number of bidders in an auction is independent of the valuations of those bidders; however, since values are correlated, we need to state it in a slightly more clunky way. Within our framework of conditionally-independent private values, this simply requires N to be independent of θ, and independent of the realizations of the draws of bidder values from the distribution F(·|θ).

This “exclusion restriction” that the number of bidders is independent of their valuations has been used in several papers on first-price auctions with independent private values: 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, consider ascending auctions, in which strategic sophistication and risk preferences play no role; thus, we are able to use this same exclusion restriction to identify a complexity that these other papers must assume away: unobserved heterogeneity. What we show (Theorems 5 and 6) is that within our model of conditionally independent private values, exogenous variation in auction size allows us to identify the primitives necessary to calculate expected revenue (as a function of reserve price), and therefore revenue-maximizing reserve price.

In addition, while this assumption has been used several times, it has never been formally tested. In our context, exogenous N imposes nonparametric restrictions on bid distributions, which we argue would likely be violated if the assumption failed. (The restrictions take the form of inequalities which generalize the equality test introduced by Athey and Haile (2002) for IPV ascending auctions.) We can therefore use these restrictions to design a nonparametric test of the assumption of exogenous N.

Below, we do three things. First, we derive the restrictions implied by our model, and explain the economics of why they would likely be violated if the exclusion restriction did not hold. Second, we introduce a nonparametric, asymptotically-valid test statistic, and use it to test these restrictions. We apply this test to data from U.S. Forest Service timber auctions, and show that the data fails to reject our model (including exogenous variation in N). Emboldened, we then apply our identification strategy to this same data, to demonstrate the counterfactual expected revenue and optimal reserve implications. These implications are substantially different from the implications that would follow from “standard” analysis under the assumption of independent values.
3 Testable Implications of Exogenous Auction Size

3.1 Benchmark: Testing under IPV

Our test for exogenous variation in auction size within our model is most easily understood if we first introduce an analogous test within the independent private values framework. For \( n > 1 \), define a function \( \psi_{n-1:n} : [0,1] \to [0,1] \) by

\[
\psi_{n-1:n}(s) = ns^{n-1} - (n-1)s^n
\]

This function has the property that given \( n \) independent draws from a probability distribution \( F(\cdot) \), the CDF of the second-highest of the draws is \( \psi_{n-1:n}(F(\cdot)) \). In the symmetric IPV model, valuations of participating bidders are assumed to be independent draws from some probability distribution \( F_V \), so \( F_{n-1:n}(v) = \psi_{n-1:n}(F_V(v)) \), or alternatively \( \psi_{n-1:n}^{-1}(F_{n-1:n}(v)) = F_V(v) \). A test of the symmetric IPV model with exogenous \( N \) is therefore whether

\[
\psi_{n'-1:n'}^{-1}(F_{n'-1:n'}(v)) = \psi_{n-1:n}^{-1}(F_{n-1:n}(v))
\]

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

3.2 Our Test under CIPV

Returning to our model of conditionally independent private values, the assumption of exogenous \( N \) would imply that the valuations of bidders participating in an auction of size \( n \) are independent draws from a distribution \( F(\cdot|\theta) \), where \( \theta \) are the characteristics of the auction (unobserved by the econometrician) and \( \theta \) is independent of \( N \).

To see why the test above is not still valid, consider the extreme case of perfectly correlated values, where \( \theta \) is one-dimensional and each bidder’s valuation in an auction with characteristics \( \theta \) is simply \( \theta \). In this case, \( F_{n-1:n}(v) \) for \( n \geq 2 \) is simply the ex-ante distribution of \( \theta \), and therefore does not vary with \( n \); \( \psi_{n-1:n}^{-1}(F_{n-1:n}(v)) \), then, is strictly increasing in \( n \), and (2) fails.

However, under CIPV, the way in which the test fails is predictable, and leads us to a similarly-motivated inequality test. Under our model, \( F_{n-1:n}(v) \) must still be weakly decreasing in \( n \); but we show below that under CIPV with exogenous \( N \), it decreases more slowly than it would under IPV, and so \( \psi_{n-1:n}^{-1}(F_{n-1:n}(v)) \) must be increasing in \( n \). We prove this result, then argue why this test should have power against the most likely source of endogeneity.

Theorem 1 Under symmetric, conditionally independent private values with exogenous \( N \), for any \( v \), \( F_{n-1:n}(v) \) is decreasing in \( n \) and \( \psi_{n-1:n}^{-1}(F_{n-1:n}(v)) \) is increasing in \( n \).

\footnote{From (1), it is straightforward to show that \( \psi_{n-1:n} \) is strictly increasing and onto, and therefore invertible.}
Proof. For the first result, note that $\psi_{n-1:n}(s)$ is decreasing in $n$. At a given realization of $\theta$, the distribution of $V_{n-1:n}$ is $\psi_{n-1:n}(F(\cdot | \theta))$; so the unconditional distribution of $V_{n-1:n}$
\[ F_{n-1:n}(v) = E_\theta \{ \psi_{n-1:n}(F(v|\theta)) \} \]
is decreasing in $n$ as well.

For the second result, it suffices to show $\psi_{n,n+1}^{-1}(F_{n,n+1}(s)) \geq \psi_{n-1:n}^{-1}(F_{n-1:n}(s))$. The key step of the proof is the observation that $\psi_{n-1:n} \circ \psi_{n,n+1}^{-1} : [0,1] \to [0,1]$ is concave. For any two differentiable functions $f$ and $g$ with $g$ invertible,
\[ \frac{d}{ds} \left( f \circ g^{-1} \right) (s) = f' \left( g^{-1}(s) \right) \cdot (g^{-1})' (s) = \frac{f'(g^{-1}(s))}{g'(g^{-1}(s))}, \]
yielding
\[ \frac{d}{ds} \left( \psi_{n-1:n} \circ \psi_{n,n+1}^{-1} \right) (s) = \frac{\psi_{n-1:n}'(t)}{\psi_{n,n+1}'(t)} = \frac{n(1-t)n^{-1} - t}{n(1-t)n^{-1}} = \frac{n-1}{n+1} \cdot \frac{1}{t} \]
where $t = \psi_{n,n+1}^{-1}(s)$. So $(\psi_{n-1:n} \circ \psi_{n,n+1}^{-1})' (s) = \frac{n-1}{n+1} / \psi_{n,n+1}^{-1}(s)$, which is decreasing in $s$, establishing that $\psi_{n-1:n} \circ \psi_{n,n+1}^{-1}$ is concave. Jensen’s inequality then yields
\[ \left( \psi_{n-1:n} \circ \psi_{n,n+1}^{-1} \right) (F_{n,n+1}(v)) = \left( \psi_{n-1:n} \circ \psi_{n,n+1}^{-1} \right) E_\theta \{ \psi_{n,n+1}(F(v|\theta)) \} \]
\[ \geq E_\theta \{ \left( \psi_{n-1:n} \circ \psi_{n,n+1}^{-1} \right) (\psi_{n,n+1}(F(v|\theta))) \} \]
\[ = E_\theta \{ \psi_{n-1:n} \left( (\psi_{n,n+1}^{-1} \circ \psi_{n,n+1}) (F(v|\theta)) \right) \} \]
\[ = E_\theta \{ \psi_{n-1:n}(F(v|\theta)) \} \]
\[ = F_{n-1:n}(v) \]
so $\psi_{n,n+1}^{-1}(F_{n,n+1}(v)) \geq \psi_{n-1:n}^{-1}(F_{n-1:n}(v))$. \hfill \Box

Since we assume $F_{n-1:n}$ can be inferred from bid data for various $n$, Theorem \ref{thm:proof} gives us a testable restriction implied by exogenous auction size under the assumption of symmetric, conditionally independent private values: specifically, that
\[ n > n' \quad \implies \quad F_{n-1:n}(v) \leq F_{n'-1:n'}(v) \]  \hfill (3)
and
\[ n > n' \quad \implies \quad \psi_{n-1:n}^{-1}(F_{n-1:n}(v)) \geq \psi_{n'-1:n'}^{-1}(F_{n'-1:n'}(v)) \]  \hfill (4)
for any $n,n'$ and any $v$.

3.3 Why (3) and (4) Should Detect Endogeneity of $N$

Since we cannot directly test for exogenous $N$, only for a consequence of it, it remains to argue why we expect this test to have power; that is, why endogeneity of $N$ would likely lead to a violation of either (3) or (4). The logic is as follows.

Under IPV with exogenous $N$, as $n$ increases, the distribution of $V_{n-1:n}$ shifts to the right, so $F_{n-1:n}(v)$ decreases; but the function $\psi_{n-1:n}^{-1}(\cdot)$ increases, and the two effects exactly balance each other.

\footnote{For $n > 2$, it’s just algebra to show that $\psi_{n-1:n}(s) - \psi_{n-2:n-1}(s) = -(n-1)s^{n-2}(1-s)^2 \leq 0$.}
\[ \psi_{n-1,n}^{-1}(F_{n-1,n}(v)) \] is unchanged. Under CIPV with exogenous \( N \), the change in \( \psi_{n-1,n}^{-1}(\cdot) \) is the same as before, but \( \psi_{n-1,n}^{-1}(F_{n-1,n}(v)) \) is now increasing rather than constant; thus, the second part of Theorem 1 therefore says that \( F_{n-1,n}(v) \) must decrease in \( n \) more slowly under CIPV than it would under IPV. Anything that "speeds up" the dependence of \( F_{n-1,n}(v) \) on \( n \), beyond the rate of change under IPV, would lead to a violation of (4), and therefore a failure of the test.

The most plausible source of endogeneity of auction size would be if bidders were more prone to enter auctions for more valuable objects – either auctions with characteristics (\( \theta \)) which make them likely to be more valuable, or auctions for objects that the bidder already knows he values more highly. Such “positive selection” would mean that auctions with larger \( n \) were more likely to have high prices, beyond the natural effect of drawing from a larger pool of bidders; this is exactly the behavior that the test (4) will detect. However, since selection increases the dependence of \( F_{n-1,n} \) on \( n \) but correlation weakens it, the two effects could potentially cancel out and be undetected if selection were weak enough. In Appendix A.1, we present two examples in which any correlation between \( N \) and \( \theta \) leads to a violation of (4) for some values of \( v \).

Negative selection – bidders being more prone to participate in auctions for less-valuable objects – would not lead to a violation of (4); but if it were a strong enough effect, it would lead to a violation of (3). So strong negative selection should also be detected. Weak negative selection, however, could be indistinguishable from correlation in values, since either one would slow the dependence of \( F_{n-1,n} \) without reversing its direction.

(While counterintuitive, negative selection 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 positive selection is the more likely source of endogeneity in \( N \); thus, we expect (4) to be the more meaningful test.)

\[ F_{n-1,n}(v) \leq F_{n-2,n-1}(v) - \frac{(n-1)(n-3)}{(n-2)^2} \frac{(F_{n-3,n-2}(v) - F_{n-2,n-1}(v))^2}{(F_{n-4,n-3}(v) - F_{n-3,n-2}(v))^2} \]

(which we prove in Appendix A.2) to give a stronger test against negative selection. In practice, however, due to the difference term in the denominator, this latter test is highly unstable when applied to a reasonably-sized dataset.

---

\(^6\) In settings where exogenous \( N \) is taken for granted, a violation of (3) (a left-shift in the distribution of winning bids as \( n \) increases) is interpreted as an indication 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 selection.

\(^7\) In theory, we can replace (3) with the stronger result that
4 Applying the Test

4.1 Description of Data

We apply our test to data from timber auctions run by the United States Forest Service. These auctions have received a great deal of attention in the empirical auctions literature; but except for some efforts by Haile, Hong, and Shum (2003), the literature has thus far ignored the issue of unobserved heterogeneity. Thus we hope this application to be fruitful as it is a widely studied auction in which the role of heterogeneity has not been addressed.

The data come to us from Phil Haile who has posted timber auction data from 1978 to 1996. The rough structure of these auctions is that the forestry service announces an auction, and bidders must submit sealed bids that define their opening bid at the auction. Those bidders with a sealed bid above the reserve are eligible to attend. The forestry service conducts a "cruise" of the tract and publishes detailed information on the tract for potential bidders. Potential bidders can also conduct their own cruises. It is commonly understood that the reserve prices set by the Forest Service are too low and do not bind. A particular empirical question of interest for us is whether the Forest Service should be setting a higher reserve to extract greater rents.

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 common-value uncertainty about the total amount of timber should not affect valuations. What is left is a sample of auctions in which the private values assumption is thought best to hold. 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.

A number of other papers have considered these Forest Service auctions. Baldwin, Marshall, and Richard (1997) provide much institutional background. Their focus is to test for collusion. Paarsch (1997) calculates the optimal reserve price for timber auctions in British Columbia using bid data from first-price auctions. Haile (2001) considers the effects of resale on valuations. Haile, Hong, and Shum (2003) consider testing 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. All paper besides a few elements of Haile, Hong, and Shum (2003) do not consider the effects of unobserved heterogeneity in drawing inferences from the data, which is our current concern.

In addition to the bids of each bidder in the room, we also have a set of covariates for each tract which come from the Forest Service’s cruise report.
4.2 A Simple Visual Test

Theorem 1 predicts that two strings of inequalities must hold under conditionally independent private values with exogenous variation in auction size:

\[ F_{1:2}(v) \geq F_{2:3}(v) \geq F_{3:4}(v) \geq \cdots \]

\[ \psi_{1:2}^{-1}(F_{1:2}(v)) \leq \psi_{2:3}^{-1}(F_{2:3}(v)) \leq \psi_{3:4}^{-1}(F_{3:4}(v)) \leq \cdots \]  \hspace{1cm} (5)

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 test, then, we can simply construct empirical analogs of the distribution functions \( F_{n-1:n}(v) \) directly from bid data, plot \( \hat{F}_{n-1:n}(v) \) and \( \psi_{n-1:n}^{-1}(\hat{F}_{n-1:n}(v)) \) against \( v \) for various \( n \), and check whether these distributions satisfy (5). Even without formalizing this into a proper test, this should give some intuition for whether exogenous auction size is a plausible hypothesis given our data, or whether either positive or negative selection seems a likely problem.

For a first pass, we let \( \theta \) capture all auction heterogeneity, that is, we make no attempt to control even for observable covariates. Figure 1 shows \( \hat{F}_{n-1:n}(v) \) and \( \psi_{n-1:n}^{-1}(\hat{F}_{n-1:n}(v)) \) 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 exogenous auction size.

While this certainly seems to suggest the data is consistent with our model and exogenous \( N \), we would still like to develop a more formal test of this. In addition, while the curves in the bottom graph do not appear to be the same, and so the data does not appear to be consistent with the standard model of independent private values when we do not account for covariates, it would be useful to see whether our sample size is large enough to reject the standard model. (If not, failure to reject our own model is not a particularly strong endorsement.)

4.3 Statistical Test

Next, we introduce a formal econometric test for the inequalities (3) and (4). We construct a test statistic which is calculated from observed auction data; asymptotically, our test statistic will diverge to +\( \infty \) if either (3) or (4) is violated anywhere; converge to 0 in probability if both (3) and (4) hold strictly almost everywhere; and converge to a normal distribution with mean 0 and bounded variance if either holds with equality on a positive measure of \( v \).

Our test statistic is based on the realizations of \( N \) and \( V_{N-1:N} \) in each auction in the data. We show that the inequalities (3) and (4) hold for every \( v \) and every pair \( (n, n') \) if and only if a certain pair of functions have mean zero. We characterize these functions and develop a test for these moment restrictions. We continue to assume that the transaction price in each auction exactly reveals the second-highest valuation; the test can easily be modified to work under weaker bidding assumptions.
Figure 1: The “eyeball test” for exogenous auction size (assuming $V_{n-1,n} = \text{transaction price in the data}$)

$\hat{F}_{n-1,n}(v)$

(should shift to the right as $n$ increases)

$\psi_{n-1,n}^{-1}\left(\hat{F}_{n-1,n}(v)\right)$

(should shift to the left as $n$ increases)
For clarity, we first present the test without conditioning on any auction-specific covariates. After that, we extend the test to be run conditional on observable covariates; we will run the test both unconditionally and conditional on appraisal value.

Notation and Preliminaries

We assume we have observations from $L$ ascending auctions, indexed by $i \in \{1, 2, \ldots, L\}$. For each auction, we observe the number of bidders, $N$, and the winning bid (transaction price), $W$, which we assume matches $V_{N-1:N}$.

We assume these observations $Z_i = (W_i, N_i)$ are $i.i.d.$ draws from some ex-ante distribution of auctions $Z = (W, N)$. 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 $S_Z$ denote the joint support of $(W, N)$, and $S_W$ and $S_N$ the marginal supports of $W$ and $N$, respectively. We assume that $S_Z$ is compact and that $F_{W|N}(\cdot|n)$ is continuous.

Test Statistic

Lee, Linton, and Whang (2009) have recently described a test for stochastic monotonicity which could be used to evaluate whether (3) is satisfied almost everywhere in $S_W \times S_N$. However, their procedure is not designed to test (4) which is effectively a comparison between a distribution function and a nonlinear transformation of it. We present a testing procedure whose general design encompasses both (3) and (4), and can be straightforwardly modified to test whether these inequalities hold conditional on a vector of observables $X$ (see below). The invertibility properties of $\psi_{n-1:n}(\cdot)$ will enable us to interpret both (3) and (4) as statements concerning monotonicity properties of certain conditional expectations. Instead of basing our tests on functionals whose suprema over $S_W \times S_N$ is zero if and only if (3) and (4) are satisfied there, we focus on functionals whose expectations over $S_W \times S_N$ are zero in such a case. The use of functionals whose expectation is zero if and only if a conditional moment inequality holds w.p.1 is not new, as it is prominently present, e.g., in Khan and Tamer (2009). 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 function $\psi_{n-1:n}(\cdot)$ was defined in (1). For any pair $n, n' \in S_N$, define

$$\Omega(s, n, n') = \psi_{n-1:n} \circ \psi_{n'-1:n'}^{-1}(s)$$

$$\Delta_{W|N}(w, n, n') = F_{W|N}(w|n) - F_{W|N}(w|n')$$

$$\Phi_{W|N}(w, n, n') = \Omega(F_{W|N}(w|n'), n, n') - F_{W|N}(w|n)$$

8If not, we can restrict ourselves to a compact subset $Z \subset S_Z$ and test whether (3) and (4) are satisfied in $Z$.  

11
Thus, \( \Omega \) appears only once in \( \mathbf{W} \) and let \( \mu \) denote \( \mu_{\mathbf{W}|N}(\mathbf{Z}, \mathbf{Z}', \mathbf{Z}_k) \), then

\[
\mu_{\mathbf{W}|N} = E[T_{\mathbf{W}|N}(\mathbf{Z}, \mathbf{Z}, \mathbf{Z}_k)] \quad \text{and} \quad \mu_{\Omega|N} = E[T_{\Omega|N}(\mathbf{Z}, \mathbf{Z}, \mathbf{Z}_k)]
\]

(Recall that \( Z_{\ell} \) are i.i.d., so \( \mu_{\mathbf{W}|N} \) and \( \mu_{\Omega|N} \) are well-defined independent of \( \{i, j, k\} \).)

**Theorem 2** If \( Z_i, Z_j, Z_k \) are i.i.d., then \( \mu_{\mathbf{W}|N} \geq 0 \) and \( \mu_{\Omega|N} \geq 0 \). Furthermore,

(i) \( \mu_{\mathbf{W}|N} = 0 \) if and only if \( \mathbf{F}_{\mathbf{W}|N}(w|n) \leq \mathbf{F}_{\mathbf{W}|N}(w|n') \) for every \( n > n' \) (\( n, n' \in \mathcal{S}_N \)) and every \( w \), that is, if and only if the true data-generating process satisfies (3) everywhere.

(ii) \( \mu_{\Omega|N} = 0 \) if and only if \( \psi_{n-1,n}(\mathbf{F}_{\mathbf{W}|N}(w|n)) \geq \psi_{n-1,n'}(\mathbf{F}_{\mathbf{W}|N}(w|n')) \) for every \( n > n' \) (\( n, n' \in \mathcal{S}_N \)) and every \( w \), that is, if and only if the true data-generating process satisfies (4) everywhere.

**Proof.** Fixing \((N_i, Z_j, Z_k)\), define \( T_{\mathbf{W}|N}(N_i, Z_j, Z_k) \equiv E[N_i, Z_j, Z_k]T_{\mathbf{W}|N}(N_i, Z_j, Z_k) \). Since \( W_i \) appears only once in \( T_{\mathbf{W}|N} \), in the indicator function \( \mathbb{I}\{W_i \leq W_k\} \), and \( Z_i \perp (Z_j, Z_k) \),

\[
T_{\mathbf{W}|N}(N_i, Z_j, Z_k) = E[N_i, Z_j, Z_k]T_{\mathbf{W}|N}(N_i, Z_j, Z_k)
\]

Thus, \( T_{\mathbf{W}|N} \geq 0 \), and \( T_{\mathbf{W}|N} > 0 \) if and only if \( N_i > N_j \) and \( \Delta_{\mathbf{W}|N}(W_k, N_i, N_j) > 0 \). By iterated expectations, \( \mu_{\mathbf{W}|N} = E_{N_i, Z_j, Z_k}E[N_i, Z_j, Z_k]T_{\mathbf{W}|N}(N_i, Z_j, Z_k) = E_{N_i, Z_j, Z_k}T_{\mathbf{W}|N}(N_i, Z_j, Z_k) \) is therefore weakly positive, and strictly positive if and only if \( N_i > N_j \) and \( \Delta_{\mathbf{W}|N}(W_k, N_i, N_j) > 0 \) with positive probability over \((N_i, N_j, W_k)\).

The same argument holds for \( \mu_{\Omega|N} \): define

\[
T_{\Omega|N}(N_i, Z_j, Z_k) = E_{N_i, Z_j, Z_k}T_{\Omega|N}(N_i, Z_j, Z_k)
\]

Again, \( \mu_{\Omega|N} = E_{N_i, Z_j, Z_k}T_{\Omega|N}(N_i, Z_j, Z_k) \) is weakly positive, and strictly positive if and only if \( N_i > N_j \) and \( \Phi_{\mathbf{W}|N}(W_k, N_i, N_j) > 0 \) with positive probability.

We show in Appendix A.3 that the qualifier “with positive probability” is unnecessary: if for any \( n > n' \) with \( n, n' \in \mathcal{S}_N \) and \( P_N(n), P_N(n') > 0 \), there is any \( w \) such that \( F_{\mathbf{W}|N}(w|n) > F_{\mathbf{W}|N}(w|n') \), then \( \mu_{\mathbf{W}|N} > 0 \); and likewise, if there is any \( w \) such that \( \Phi_{\mathbf{W}|N}(w, n, n') > 0 \), \( \mu_{\Omega|N} > 0 \).
Sample analog estimators for $\mu^{FW|N}$ and $\mu^{\Omega W|N}$ would take the form of (third 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. Let

$$U_{L(3)}^{FW|N} = \frac{1}{L(L-1)(L-2)} \sum_{i=1}^{L} \sum_{j \neq i} \sum_{k \neq i, k \neq j} T^{FW|N}(Z_i, Z_j, Z_k),$$

$$U_{L(3)}^{\Omega W|N} = \frac{1}{L(L-1)(L-2)} \sum_{i=1}^{L} \sum_{j \neq i} \sum_{k \neq i, k \neq j} T^{\Omega W|N}(Z_i, Z_j, Z_k).$$

Note that

$$E[U_{L(3)}^{FW|N}] = \mu^{FW|N} \quad \text{and} \quad E[U_{L(3)}^{\Omega W|N}] = \mu^{\Omega W|N}.$$ 

Thus, a natural test of (3) could be based on $U_{L(3)}^{FW|N}$, and a natural test of (4) on $U_{L(3)}^{\Omega W|N}$. Unfortunately, $T^{FW|N}(Z_i, Z_j, Z_k)$ and $T^{\Omega W|N}(Z_i, Z_j, Z_k)$ cannot be calculated directly, since $F_{W|N}(\cdot | \cdot)$ is not known. Instead, we replace $F_{W|N}(\cdot | \cdot)$ in the calculation of $T^{FW|N}(Z_i, Z_j, Z_k)$ and $T^{\Omega W|N}(Z_i, Z_j, Z_k)$ with a nonparametric estimate based on all the data except observations $i$, $j$, and $k$. Formally, take any pair $(w, n)$ where $n \in S_N$. For each distinct triple $(i, j, k)$ in $(1, \ldots, L)$, we will define

$$\hat{R}_{W|N}^{-i,j,k}(w|n) = \frac{1}{L-3} \sum_{\ell \neq i, j, k} \mathbb{I}\{W_\ell \leq w\} \cdot \mathbb{I}\{N_\ell = n\}$$

$$\hat{P}_{N}^{-i,j,k}(n) = \frac{1}{L-3} \sum_{\ell \neq i, j, k} \mathbb{I}\{N_\ell = n\}$$

$$\hat{\Delta}_{W|N}^{-i,j,k}(w,n,n') = \left\{ \begin{array}{ll} \hat{R}_{W|N}^{-i,j,k}(w|n) / \hat{P}_{N}^{-i,j,k}(n) & \text{if } \hat{P}_{N}^{-i,j,k}(n) \neq 0 \\ 0 & \text{otherwise} \end{array} \right.$$ 

We calculate corresponding estimates for $\Delta_{W|N}$ and $\Phi_{W|N}$:

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

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

For each triple $(i, j, k)$ in $(1, \ldots, L)$, then, we will replace $T^{FW|N}(Z_i, Z_j, Z_k)$ and $T^{\Omega W|N}(Z_i, Z_j, Z_k)$ with

$$\hat{T}^{FW|N}(Z_i, Z_j, Z_k) = \{ \mathbb{I}\{W_i \leq W_k\} - \hat{F}_{W|N}^{-i,j,k}(W_k|N_j) \} \cdot \mathbb{I}\{N_i > N_j\} \cdot \mathbb{I}\{\hat{\Delta}_{W|N}^{-i,j,k}(W_k, N_i, N_j) \geq -b_L\}$$

$$\hat{T}^{\Omega W|N}(Z_i, Z_j, Z_k) = \{ \mathbb{I}(\hat{F}_{W|N}^{-i,j,k}(W_k|N_j), N_i, N_j) - \mathbb{I}\{W_i \leq W_k\} \} \cdot \mathbb{I}\{N_i > N_j\} \cdot \mathbb{I}\{\hat{\Phi}_{W|N}^{-i,j,k}(W_k, N_i, N_j) \geq -b_L\},$$

(6)
where \( b_L \) is a deterministic positive sequence such that \( b_L \to 0 \), \( \sqrt{L} \cdot b_L \to \infty \), and \( \sqrt{L} \cdot b_L^2 \to 0 \) as \( L \to \infty \). Our test statistics will replace \( U_{L(3)}^{F_{W|N}} \) and \( U_{L(3)}^{O_{W|N}} \) with

\[
U_{L(3)}^{F_{W|N}} = \frac{1}{L(L-1)(L-2)} \sum_{i=1}^{L} \sum_{j \neq i} \sum_{k \neq i \atop k \neq j} \widehat{T}_{F_{W|N}}(Z_i, Z_j, Z_k),
\]

\[
U_{L(3)}^{O_{W|N}} = \frac{1}{L(L-1)(L-2)} \sum_{i=1}^{L} \sum_{j \neq i} \sum_{k \neq i \atop k \neq j} \widehat{T}_{O_{W|N}}(Z_i, Z_j, Z_k).
\]

We introduce a nonzero sequence \( b_L \) into our formulation because we allow for the possibility that the inequalities in (3) and (4) may (or may not) hold with equality at some \( \omega \) instead of letting it disappear asymptotically. Dropping the \( \omega \) allows us to characterize exponential bounds for

\[
\Pr(\widehat{\Delta}_{W|N}^{-i,j,k}(W_k, N_i, N_j) < -b_L \text{ and } \Delta_{W|N}(W_k, N_i, N_j) \geq 0 \text{ for } i \neq j \neq k),
\]

and for \( \Pr(\widehat{\Phi}_{W|N}^{-i,j,k}(W_k, N_i, N_j) < -b_L \text{ and } \Phi_{W|N}(W_k, N_i, N_j) \geq 0 \). These bounds would not be valid if we fixed \( b_L \) instead of letting it disappear asymptotically. Dropping the \( i^{th} \), \( j^{th} \), and \( k^{th} \) observations when calculating the estimates \( \widehat{F}_{W|N}^{-i,j,k} \) helps ensure that the overall variances of \( U_{L(3)}^{F_{W|N}} \) and \( U_{L(3)}^{O_{W|N}} \) converge to zero at an appropriate rate.

Theorem 8 (in Appendix A.6) provides regularity conditions for the distribution \( F_{W|N} \) such that \( U_{L(3)}^{F_{W|N}} \) and \( U_{L(3)}^{O_{W|N}} \) satisfy the following asymptotic behavior:

\[
\sqrt{L} \cdot U_{L(3)}^{F_{W|N}} = \sqrt{L} \cdot \mu_{F_{W|N}} + \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \eta_{F_{W|N}}(Z_i) + o_p(1), \text{ where } E[\eta_{F_{W|N}}(Z_i)] = 0,
\]

\[
\sqrt{L} \cdot U_{L(3)}^{O_{W|N}} = \sqrt{L} \cdot \mu_{O_{W|N}} + \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \eta_{O_{W|N}}(Z_i) + o_p(1), \text{ where } E[\eta_{O_{W|N}}(Z_i)] = 0.
\]  \( (7) \)

By Theorem 2, the results in (7) imply that the statistic \( \sqrt{L} \cdot U_{L(3)}^{F_{W|N}} \) (\( \sqrt{L} \cdot U_{L(3)}^{O_{W|N}} \)) will diverge to \(+\infty\) with probability 1 if the inequalities in Equation (3) (Equation (4)) are violated at any \( (n, n', v) \) with \( n, n' \in S_N \). Theorem 8 describes the functions \( \eta_{F_{W|N}} \) and \( \eta_{O_{W|N}} \) precisely; their structure arises from the Hoeffding decomposition of U-statistics (see Hoeffding (1961) and Serfling (1980), Chapter 5). Let \( \Sigma_{F_{W|N}} = \text{Var}(\eta_{F_{W|N}}(Z_i)) \) and \( \Sigma_{O_{W|N}} = \text{Var}(\eta_{O_{W|N}}(Z_i)) \). We show in the appendix that \( \Sigma_{F_{W|N}} > 0 \) (\( \Sigma_{O_{W|N}} > 0 \) if and only if the inequalities in Equation (3) (Equation (4)) are binding with nonzero probability. That is, these variances satisfy \( \Sigma_{F_{W|N}} > 0 \iff E[\mathbb{I}\{N_i > N_j\} \mathbb{I}\{\Delta_{W|N}(W_k, N_i, N_j) = 0\}] > 0 \) and \( \Sigma_{O_{W|N}} > 0 \iff E[\mathbb{I}\{N_i > N_j\} \mathbb{I}\{\Phi_{W|N}(W_k, N_i, N_j) = 0\}] > 0 \) for \( i \neq j \neq k \).

The conditions required for the result in (7) are fully compatible with the assumptions in Section 3.2. They involve regularity conditions on \( F_{W|N} \) reminiscent of those present in many nonparametric estimation and inferential problems. For example, while we allow for the existence of \( n, n' \in S_n \) for which the distribution of
$\Delta_{W|N}(W, n, n')$ has a probability mass at zero, we require that the density of $\Delta_{W|N}(W, n, n')$ be uniformly bounded for all $n, n'$ in an interval of the form $[-s, 0]$ for some $s > 0$. We impose the same type of restriction on $\Phi_{W|N}(W, n, n')$. Section A.6 in the appendix describes in detail the required assumptions, and include a step-by-step proof of the result in (7) (Theorem 8).

Now, suppose (3) is satisfied everywhere by the true data generating process. By Theorem 8, the statistic $\sqrt{L} \cdot U_{L(3)}^{FW|N}$ will be asymptotically normal with mean zero and variance $\Sigma_{FW|N} > 0$ if the inequalities in (3) are binding with nonzero probability; otherwise, we will have $\sqrt{L} \cdot U_{L(3)}^{FW|N} = o_p(1)$. The statistic $\sqrt{L} \cdot U_{L(3)}^{\Omega_{W|N}}$ has analogous asymptotic properties. Based on these results, our testing procedure has the following form. Consider the null hypothesis $H_0^{FW|N}$: is satisfied almost everywhere. Let $\hat{S}_{FW|N}$ denote an estimator whose probability limit ‘$\text{plim}(\hat{S}_{FW|N})$’ satisfies $\text{plim}(\hat{S}_{FW|N}) > 0$ and $\text{plim}(\hat{S}_{FW|N}) \geq \Sigma_{FW|N}^{1/2}$. Let $\alpha \in (0, 1)$ denote a pre-specified significance level and let $z_\alpha$ satisfy $\Pr(\mathcal{F} \geq z_\alpha) = \alpha$, where $\mathcal{F} \sim N(0, 1)$. Consider the following decision rule:

$$\text{Reject } H_0^{FW|N} \text{ if } \frac{\sqrt{L} \cdot U_{L(3)}^{FW|N}}{\hat{S}_{FW|N}} \geq z_\alpha$$

By Theorem 8 this decision rule has the following asymptotic properties:

$$\lim_{L \to \infty} \left\{ \Pr[ H_0^{FW|N} \text{ is falsely rejected } ] \right\} \leq \alpha, \quad \text{and} \quad \lim_{L \to \infty} \left\{ \Pr[ H_0^{FW|N} \text{ is rejected when it is false } ] \right\} = 1.$$

The equivalent result holds for $H_0^{\Omega_{W|N}}$: is satisfied almost everywhere, and a decision rule of the form

$$\text{Reject } H_0^{\Omega_{W|N}} \text{ if } \frac{\sqrt{L} \cdot U_{L(3)}^{\Omega_{W|N}}}{\hat{S}_{\Omega_{W|N}}} \geq z_\alpha$$

where $\hat{S}_{\Omega_{W|N}}$ is such that $\text{plim}(\hat{S}_{\Omega_{W|N}}) > 0$ and $\text{plim}(\hat{S}_{\Omega_{W|N}}) \geq \Sigma_{\Omega_{W|N}}^{1/2}$. We can estimate $\Sigma_{FW|N}$ and $\Sigma_{\Omega_{W|N}}$ by using the analytical expressions for $\text{Var}(\eta_{FW|N}(Z_i))$ and $\text{Var}(\eta_{\Omega_{W|N}}(Z_i))$ (see (24) in the appendix), or by an asymptotically valid resampling method. Equipped with consistent estimators $\hat{S}_{FW|N}$ and $\hat{S}_{\Omega_{W|N}}$, we can implement the tests described above by using $\hat{S}_{FW|N} = \hat{\Sigma}_{FW|N}^{1/2} + \epsilon_1$ and $\hat{S}_{\Omega_{W|N}} = \hat{\Sigma}_{\Omega_{W|N}}^{1/2} + \epsilon_2$ with nonnegative $\epsilon_j$. (The $\epsilon_j$ can be set to zero if the inequalities in (3) or (4) are assumed to be binding with nonzero probability.)

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), and 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).

Testing over a Subset of \( S_W \) and \( S_N \)

We may want to restrict ourselves to test if (3) and (4) are satisfied over a specific subset of values of \((w, n)\). For instance, since \( \frac{\partial}{\partial s} \Omega(s, n, n') \) becomes unbounded at \( s = 0 \) or \( s = 1 \) for any \( n, n' \), the asymptotic properties of \( \sqrt{L} \cdot U_{L(3)}^{FW|N} \) outlined in (4) require that \( F_{W|N}(W_k|N_j) \) be bounded away from 0 and 1 for every \( j \neq k \) in our sample. In general, ensuring this condition asymptotically requires restricting ourselves to a pre-specified range of values in \( S_W \times S_N \). Let \( W \subseteq S_W \) and \( N \subseteq S_N \), and let \( Z = W \times N \). We can test if (3) and (4) are satisfied in \( Z \) by modifying our test-statistics as follows. Let

\[
T_Z^{FW|N}(Z_i, Z_j, Z_k) = T^{FW|N}(Z_i, Z_j, Z_k) \cdot I\{N_i, N_j \in N\} \cdot I\{W_k \in W\},
\]

\[
T_Z^{\Omega W|N}(Z_i, Z_j, Z_k) = T^{\Omega W|N}(Z_i, Z_j, Z_k) \cdot I\{N_i, N_j \in N\} \cdot I\{W_k \in W\}
\]

By inspection, we have \( E[T_Z^{FW|N}(Z_i, Z_j, Z_k)] = 0 \) if and only if (7) is satisfied in \( Z \). Otherwise, \( E[T_Z^{FW|N}(Z_i, Z_j, Z_k)] > 0 \). The same is true for \( E[T_Z^{\Omega W|N}(Z_i, Z_j, Z_k)] \) and (4). Our tests would be based on the corresponding U-statistics produced by \( \hat{T}_Z^{FW|N}(Z_i, Z_j, Z_k) \) and \( \hat{T}_Z^{\Omega W|N}(Z_i, Z_j, Z_k) \). An asymptotic result equivalent to (7) holds.

Results of the Unconditional Test

As described above, our test-statistics in this case are \( \sqrt{L} \cdot U_{L(3)}^{FW|N} \) and \( \sqrt{L} \cdot U_{L(3)}^{\Omega W|N} \). For now, we estimate the variances \( \hat{S}_{FW|N} \) and \( \hat{S}_{\Omega W|N} \) using subsampling. Subsampled estimates of the variance are valid under the assumption that

\[
E[I\{N_i > N_j\} I\{\Delta_{W|N}(W_k, N_i, N_j) \geq 0\}] > 0
\]

\[
E[I\{N_i > N_j\} I\{\Phi_{W|N}(W_k, N_i, N_j) \geq 0\}] > 0
\]

for \( i \neq j \neq k \), that is, that the data generating process is such that (3) and (4) each is either violated, or binds with equality with positive probability. In the latter case (as we stressed previously), (4) implies that our test-statistics will have a nondegenerate asymptotic distribution, enabling us to use subsampling to estimate their variances.

Otherwise, we can calculate the variances using the sample analogs of the analytic variance expressions for the asymptotic variances under the null hypothesis (see Equation (24) in the appendix).

As suggested in the discussion prior to (5), in the construction of our test statistics we stay away from regions in \( S_W \) for which our nonparametric estimate for \( F_{W|N}(w|n) \) is either 0 or 1 for some \( n = 2, \ldots, 12 \) (the

\footnote{Our variance estimates for \( \sqrt{L} \cdot U_{L(3)}^{FW|N} \) and \( \sqrt{L} \cdot U_{L(3)}^{\Omega W|N} \) are based on 500 subsamples of size \( L = 600 \). Unless stated otherwise, this is the case for all our the results presented in this section.}
values observed in our sample). We used \( b_L = 0.001 \) to compute our test statistic\(^{10}\). The resulting values are \( \sqrt{L} \cdot U_{L,3}^{F_{W|N}} / \hat{S}_{F_{W|N}} = -0.2755 \) and \( \sqrt{L} \cdot U_{L,3}^{\hat{F}_{W|N}} / \hat{S}_{\hat{F}_{W|N}} = -0.9898 \). At an asymptotic significance level of 5\%, the corresponding critical value is 1.67. At that level, each test statistic fails to reject the null hypotheses that (3) and (4) are satisfied almost everywhere. This confirms the visual evidence described in Section 4.2.

Also worth considering is whether the data is consistent with the model of independent private values. The assumption of independence is the additional restriction that (4) holds everywhere with equality. The curves in the second graph in Figure 1 do not appear visually to be the same; however, it is not immediately obvious whether we have enough data to statistically reject the hypothesis of IPV, even without conditioning on any covariates.

Rather than testing whether (4) holds with equality, we will test whether it holds weakly with the inequality reversed: that is, we can use the same method developed above to test whether

\[
\psi_{n-1,n'}^{-1}(F_{n-1,n'}(v)) \leq \psi_{n-1,n'}^{-1}(F_{n,n'}(v)) \quad \text{for any } n > n' \text{ and any } v \tag{4}
\]

(This obviously nests equality as a special case, but allows us to proceed without developing a new test from scratch and proving its properties.) To this end, we replace \( T_{W|N}(Z_i, Z_j, Z_k) \) with

\[
T_{W|N}(Z_i, Z_j, Z_k) = \left( \mathbb{1}\{W_i \leq W_k\} - \mathbb{1}\{W_{i,j,k}(W_k|N_j, N_i, L_{1:n})\} \right) \cdot \mathbb{1}\{N_i > N_j\} \cdot \mathbb{1}\{-\hat{F}_{W|N}(W_k, N_i, N_j) \geq -b_L\}
\]

Our test-statistic is now \( \sqrt{L} \cdot U_{L,3}^{\hat{F}_{W|N}} \). This statistic will have asymptotic properties analogous those described in Equation (7) under the same type of regularity conditions as in Theorem 3. Accordingly, this would lead us to the same type of rejection rules as those described for \( \sqrt{L} \cdot U_{L,3}^{F_{W|N}} \) and \( \sqrt{L} \cdot U_{L,3}^{\hat{F}_{W|N}} \). Our results yield \( \sqrt{L} \cdot U_{L,3}^{\hat{F}_{W|N}} / \hat{S}_{\hat{F}_{W|N}} = 19.2359 \), leading us to reject (4). This supports the conjecture that when we do not condition on any covariates, our data is consistent with conditionally independent private values and exogenous \( N \), but not consistent with independent private values, again confirming the visual intuition in Section 4.2.

### 4.4 Testing Conditional of Observable Covariates

Of course, in reality, auction-by-auction heterogeneity will typically be a mix of observed and unobserved variables. If \( X \) denotes the observed covariates and \( \theta \) (as before) the unobserved, our test above was really a test of whether valuations are i.i.d. draws from a distribution \( F(\cdot | X, \theta) \), with “exogenous \( N \)” meaning that \( N \) is independent of \( (X, \theta) \). If this is true, then conditional on a realization of \( X \), \( N \) is still independent of \( \theta \); 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.

---

\(^{10}\) This is consistent with the convergence restrictions for \( b_L \), for example, if we set \( b_L = C_b \cdot L^{-\alpha} \), where \( 1/4 < \alpha < 1/2 \) and \( C_b = (2181^{10}) \cdot 0.001 \).
However, conditioning on observable covariates is potentially appealing for several reasons. First, if \( N \) is independent of \( \theta \) but not of \( X \), then our hypothesis of exogenous \( N \) (and thus our identification strategy) is invalid when we do not condition, but valid when we condition on \( X \). Second, we pointed out above that selective entry speeds up the rate at which \( F_{n-1:n} \) shifts to the right as \( n \) increases, while correlation slows it down; thus, correlation can conceal some degree of endogeneity of \( N \). By conditioning on \( X \), we remove some of the correlation that could conceal positive selection, giving us a more powerful test. So conditioning on available covariates makes our model more likely to be valid, and more likely to be rejected if it is not. And finally, we would like 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 all observable covariates were controlled for.

As we discuss in section 5 like Lu and Perrigne (2008), we find that aside from \( N \), the most important covariate for explaining variation in the winning bid is the appraisal price, given in dollars per thousand board feet. (Lu and Perrigne (2008) find the volume of timber to also have explanatory power, but we focus only on scaled sales, where volume should no longer play a significant role.) Thus, for now, we will consider the question of testing our model while controlling for appraisal value. Later on, we will consider testing while controlling for more covariates.

**Visual Test**

As a first pass, we can replicate our earlier visual test (Figure 1 above) with nonparametric estimates of the distributions of transaction prices for each \( N \), conditional on appraisal value. 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 of the distributions \( F_{n-1:n}(v|X) \) as

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

where \( K_b \) is the Gaussian kernel \( \frac{1}{2\pi b} e^{-s^2/2b^2} \) with bandwidth parameter \( b \). We use a bandwidth based on the rule of thumb suggested by Silverman (1986) (and used by Haile and Tamer (2003) as well). We can then apply the same visual test as before, to these conditional CDFs. Figure 2 shows graphs of \( \hat{F}_{n-1:n}(\cdot|X) \) and \( \psi_{n-1:n}^{-1} \left( \hat{F}_{n-1:n}(\cdot|X) \right) \) for three representative appraisal values, 26, 56, and 92 – the 25th, 50th, and 75th 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, the data seems to fit our prediction fairly well. All three plots of \( \hat{F}_{n-1:n}(v|X) \) 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 \( \psi_{n-1:n}^{-1} \left( \hat{F}_{n-1:n}(v|X) \right) \), 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. Thus, the data again appears to be broadly consistent with our predictions. Again, we
Figure 2: The same “eyeball test,” using nonparametric estimates of \( F_{n-1:n} \) conditional on appraisal value

\[
\hat{F}_{n-1:n}(v|X), \quad X = 26
\]

\[
\psi^{-1}_{n-1:n} \left( \hat{F}_{n-1:n}(v|X) \right), \quad X = 26
\]

\[
\hat{F}_{n-1:n}(v|X), \quad X = 56
\]

\[
\psi^{-1}_{n-1:n} \left( \hat{F}_{n-1:n}(v|X) \right), \quad X = 56
\]

\[
\hat{F}_{n-1:n}(v|X), \quad X = 92
\]

\[
\psi^{-1}_{n-1:n} \left( \hat{F}_{n-1:n}(v|X) \right), \quad X = 92
\]
next move on to a formal econometric test of [3] and [4], to see whether the occasional apparent violations of these predictions are explained by the sample size or refute the model.

**Statistical Test**

Let \( X \) be a vector of observable covariates and denote \( Y \equiv (N, X) \) and \( Z \equiv (W, Y) \). We maintain the assumption of having an i.i.d. sample \((Z_i)_{i=1}^N\). Define

\[
F_{W|Y}(w|n, x) \equiv \Pr(W \leq w|N = n, X = x)
\]

\[
\Delta_{W|Y}(w, n, n', x) \equiv F_{W|Y}(w|n, x) - F_{W|Y}(w|n', x)
\]

\[
\Phi_{W|Y}(w, n, n', x) = \Omega(F_{W|Y}(w|n', x), n, n') - F_{W|Y}(w|n, x)
\]

We can test whether conditional analogs of (3) and (4) hold at each realization of \( X \) – that is, whether \( n > n' \) implies

\[
F_{W|Y}(w|n, x) \leq F_{W|Y}(w|n', x)
\]

(10)

and

\[
\psi^{-1}_{n-1,n'}(F_{W|Y}(w|n, x)) \geq \psi^{-1}_{n-1,n'}(F_{W|Y}(w|n', x))
\]

(11)

(or, equivalently, \( \Delta_{W|Y}(w, n, n', x) \leq 0 \) and \( \Phi_{W|Y}(w, n, n', x) \leq 0 \) for almost all \((n, x), (n', x) \in \mathcal{S}_Y \) and \( w \in \mathcal{S}_W \)).

For illustrative purposes, we focus on the case where \( X \) is real-valued and continuously distributed. Later we will discuss the case where \( X \) is multi-dimensional and includes both discrete and continuous covariates.

Let \( f_{X|N}(.|n) \) denote the conditional density of \( X \) given \( N = n \). For \( i \neq j \neq k \neq \ell \), define

\[
\gamma_{F_{W|Y}} = \mathbb{E} \left[ \mathbb{I} \{ N_i > N_j \} \cdot f_{X|N}(X_i|N_i) \cdot \max \{ 0, \Delta_{W|Y}(W_k, N_i, N_j, X_i) \} \right]
\]

\[
\gamma_{\Omega_{W|Y}} = \mathbb{E} \left[ \mathbb{I} \{ N_i > N_j \} \cdot f_{X|N}(X_i|N_i) \cdot \max \{ 0, \Phi_{W|Y}(W_k, N_i, N_j, X_i) \} \right]
\]

\( \gamma_{F_{W|Y}} \) and \( \gamma_{\Omega_{W|Y}} \) will serve exactly the same purpose as \( \mu_{F_{W|N}} \) and \( \mu_{\Omega_{W|N}} \) did in the unconditional tests. Note that both \( \gamma_{F_{W|Y}} \) and \( \gamma_{\Omega_{W|Y}} \) are weakly positive; and if \( f_{X|N}(.|n) \) has the same support for every \( n \), both will be equal to 0 if and only if \( n > n' \) implies (10) and (11), respectively, hold almost everywhere.

As in the unconditional test, our conditional test will be based on sample analogs of \( \gamma_{F_{W|Y}} \) and \( \gamma_{\Omega_{W|Y}} \). In constructing these sample analogs, the assumption that \( X \) is continuously distributed precludes the use of indicator functions. Instead, we will use a weighting scheme with very particular asymptotic properties.

Let \( h_{\ell} \rightarrow 0 \) denote a nonnegative bandwidth sequence converging to zero and let \( K : \mathbb{R} \rightarrow \mathbb{R} \) denote a bias-reducing kernel\(^{[11]} \).

Let

\[
S_{F_{W|Y}}(Z_i, Z_j, Z_k, Z_\ell) = \left( \mathbb{I} \{ W_i \leq W_k \} - F_{W|Y}(W_k|N_j, X_\ell) \right) \cdot \mathbb{I} \{ N_i > N_j \} \cdot \mathbb{I} \{ \Delta_{W|Y}(W_k, N_i, N_j, X_\ell) \geq 0 \} \cdot \frac{1}{h_{\ell}} K \left( \frac{X_\ell - X_i}{h_{\ell}} \right)
\]

\[
S_{\Omega_{W|Y}}(Z_i, Z_j, Z_k, Z_\ell) = \left( \Omega(F_{W|Y}(W_k|N_j, X_\ell), N_i, N_j) - \mathbb{I} \{ W_i \leq W_k \} \right) \cdot \mathbb{I} \{ N_i > N_j \} \cdot \mathbb{I} \{ \Phi_{W|Y}(W_k, N_i, N_j, X_\ell) \geq 0 \} \cdot \frac{1}{h_{\ell}} K \left( \frac{X_\ell - X_i}{h_{\ell}} \right)
\]

\(^{[11]}\)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 \( \int_{-\infty}^{\infty} s^M K(s)ds < \infty. \)
Both functions are constructed such that only 4-tuples of observations \((Z_i, Z_j, Z_k, Z_{\ell})\) for which \(X_i - X_{\ell}\) belongs in a vanishing neighborhood around zero will matter asymptotically. Weighting methods with these types of asymptotic properties are referred to as *pairwise differencing* and they 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). The following result describes a set of sufficient conditions under which \(\lim_{L \to \infty} E\left[S_L^{F_{W|Y}}(Z_i, Z_j, Z_k, Z_{\ell})\right] = 0\) and \(\lim_{L \to \infty} E\left[S_L^{\Omega_{W|Y}}(Z_i, Z_j, Z_k, Z_{\ell})\right] = 0\) if and only if \(\Delta_{W|Y}\) and \(\Phi_{W|Y}\) are negative almost everywhere.

**Theorem 3** Suppose that the support of \(f_{X|N}(\cdot|n)\) is the same for all \(n \in S_N\). Further, suppose that for any \(n \in S_N\), \(f_{X|N}(x|n)\) and \(F_{W|Y}(w|n, x)\) are both \(M\) times differentiable with respect to \(x\), the latter with bounded derivatives, at almost all \(x \in S_X\). If \(i \neq j \neq k \neq \ell\), then

\[
E\left[S_L^{F_{W|Y}}(Z_i, Z_j, Z_k, Z_{\ell})\right] = \gamma_{F_{W|Y}} + O(h_L^M)
\]

\[
E\left[S_L^{\Omega_{W|Y}}(Z_i, Z_j, Z_k, Z_{\ell})\right] = \gamma_{\Omega_{W|Y}} + O(h_L^M)
\]

**Proof:** We focus on \(E[S_L^{F_{W|Y}}]\), since the proof for \(E[S_L^{\Omega_{W|Y}}]\) follows identical steps. Similar to before, define

\[
S_L^{F_{W|Y}}(N_i, Z_j, Z_k, Z_{\ell}) = E_{W_i, X_i|N_i}\left[S_L^{F_{W|Y}}((W_i, N_i, X_i), Z_j, Z_k, Z_{\ell})\right]
\]

\[
= E_{X_i|N_i}\left[E_{W_i|N_i, X_i}\left[S_L^{F_{W|Y}}((W_i, N_i, X_i), Z_j, Z_k, Z_{\ell})\right]\right]
\]

\[
= E_{X_i|N_i}\left[f_{W|Y}(W_k|N_i, X_i) - F_{W|Y}(W_k|N_i, X_{\ell})\right] f_{X|N}(X_i|N_i) dx
\]

\[
= \Delta_{W|Y}(W_k, N_i, N_j, X_i) f_{X|N}(X_i|N_i) + O(h_L^M)
\]

and so

\[
E\left[S_L^{F_{W|Y}}(Z_i, Z_j, Z_k, Z_{\ell})\right] = E\left[S_L^{F_{W|Y}}(N_i, Z_j, Z_k, Z_{\ell})\right]
\]

\[
= E\left[\Delta_{W|Y}(W_k, N_i, N_j, X_i) f_{X|N}(X_i|N_i) + O(h_L^M)\right] 1\{N_i > N_j\} 1\{\Delta_{W|Y}(W_k, N_i, N_j, X_{\ell}) \geq 0\}
\]

\[
= E\left[\Delta_{W|Y}(W_k, N_i, N_j, X_i) f_{X|N}(X_i|N_i)\right] 1\{N_i > N_j\} 1\{\Delta_{W|Y}(W_k, N_i, N_j, X_{\ell}) \geq 0\} + O(h_L^M)
\]

as claimed. \(\Box\)

Thus, asymptotically, a test of the sample means of \(S_L^{F_{W|Y}}\) and \(S_L^{\Omega_{W|Y}}\) is exactly a test of whether the conditional relationships \(F_{W|Y}(w|n, x) \leq F_{W|Y}(w|n', x)\) and \(\psi_{n-1,n}^{-1} (F_{W|Y}(w|n, x)) \geq \psi_{n-1,n'}^{-1} (F_{W|Y}(w|n', x))\) hold almost everywhere when \(n > n'\).
The rate at which $h_L^{M}$ disappears will depend on how smooth the relevant functionals of $x$ are assumed to be. In particular, constructing test-statistics with asymptotic properties analogous to those we used in the unconditional case will require that $M$ be such that $\sqrt{L} h_L^{M} \to 0$. We describe how our test is constructed next.

Test Statistic

Statistics based directly on $S_L^{W|Y}$ and $S_L^{D|W|Y}$ are not feasible because $F_{W|Y}$ is unknown. However, it is nonparametrically identified under our assumptions, and we can replace it with a nonparametric estimate. Let $\tilde{K}$ and $\tilde{h}_L$ be an additional kernel function and bandwidth sequence, and define

$$
\hat{R}^{-i,j,k,\ell}_{W|Y}(w|n,x) = \frac{1}{(L-4)\tilde{h}_L} \sum_{m \neq i,j,k,\ell} \mathbb{I}\{W_m \leq w\} \mathbb{I}\{N_m = n\} \tilde{K}\left(\frac{X_m - x}{\tilde{h}_L}\right)
$$

$$
\hat{\gamma}^{-i,j,k,\ell}(x,n) = \frac{1}{(L-4)\tilde{h}_L} \sum_{m \neq i,j,k,\ell} \mathbb{I}\{N_m = n\} \tilde{K}\left(\frac{X_m - x}{\tilde{h}_L}\right)
$$

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

and, plugging in these estimates,

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

$$
\tilde{\Phi}_{W|Y}(w,n,n',x) = \Omega\left(\hat{F}^{-i,j,k,\ell}_{W|Y}(w|n',x),n,n'\right) - \hat{F}^{-i,j,k,\ell}_{W|Y}(w|n,x)
$$

In order to obtain an asymptotic result analogous to the unconditional test case, $\tilde{K}$ needs to be bias-reducing of order $M$, and the relative rates of convergence of $b_L$, $h_L$ and $\tilde{h}_L$ need to satisfy very specific conditions (which we will outline below). Let

$$
S_L^{F_{W|Y}}(Z_i,Z_j,Z_k,Z_\ell) = \left(\mathbb{I}\{W_i \leq W_k\} - \hat{F}^{-i,j,k,\ell}_{W|Y}(W_k|N_j,X_\ell)\right) \mathbb{I}\{N_i > N_j\} \mathbb{I}\{\tilde{\Delta}^{-i,j,k,\ell}_{W|Y}(W_k,N_i,N_j,X_\ell) \geq -b_L\} \cdot \frac{1}{\tilde{h}_L} \tilde{K}\left(\frac{X_i - X_j}{\tilde{h}_L}\right)
$$

$$
S_L^{D_{W|Y}}(Z_i,Z_j,Z_k,Z_\ell) = \left(\Omega(\hat{F}^{-i,j,k,\ell}_{W|Y}(W_k|N_j,X_\ell),N_i,N_j) - \mathbb{I}\{W_i \leq W_k\}\right) \mathbb{I}\{N_i > N_j\} \mathbb{I}\{\tilde{\Phi}_{W|Y}(W_k,N_i,N_j,X_\ell) \geq -b_L\} \cdot \frac{1}{\tilde{h}_L} \tilde{K}\left(\frac{X_i - X_j}{\tilde{h}_L}\right)
$$

Our test statistics then take the form

$$
U_{L(4)}^{S_L^{F_{W|Y}}} = \frac{1}{L(L-1)(L-2)(L-3)} \sum_{i=1}^{L} \sum_{j \neq k \neq i} \sum_{k \neq i} S_L^{F_{W|Y}}(Z_i,Z_j,Z_k,Z_\ell),
$$

$$
U_{L(4)}^{S_L^{D_{W|Y}}} = \frac{1}{L(L-1)(L-2)(L-3)} \sum_{i=1}^{L} \sum_{j \neq k \neq i} \sum_{k \neq i} S_L^{D_{W|Y}}(Z_i,Z_j,Z_k,Z_\ell).
$$

(12)
In order for $\sqrt{L} \cdot U_{L(i)}^{f_{W|Y}}$ and $\sqrt{L} \cdot U_{L(i)}^{g_{W|Y}}$ to have an asymptotic behavior analogous to their unconditional counterparts, in addition to the smoothness conditions described in Theorem 3, other functionals of $x$ must also be sufficiently smooth. For instance, for any $(n, n') \in S_{X}$ and almost every $w \in S_{W}$, the following functions must be $M$-times differentiable with respect to $x$ and $x'$ for almost every $(x, x') \in S_{X}$,

$$
\phi(w, n, n', x) = E_{w} \left[ \mathbb{I} \{ w \leq W \} \mathbb{I} \{ \Delta_{W|Y}(W, n, n', x) \geq 0 \} \right] \\
\Psi(n, n', x) = E_{w} \left[ F_{W|Y}(W|n', x) \mathbb{I} \{ \Delta_{W|Y}(W, n, n', x) \geq 0 \} \right] \\
\Upsilon(w, n, x, x') = E_{W|X} \left[ \mathbb{I} \{ N > n \} \mathbb{I} \{ \Delta_{W|Y}(w, N, n, x) \geq 0 \} \mid X = x' \right].
$$

In addition, the relative convergence rates of the bandwidths employed must satisfy

$$L^{\frac{1}{2}} \cdot b_{L} \cdot \tilde{h}_{L} \cdot h_{L} \longrightarrow \infty \quad \text{and} \quad \frac{L^{\frac{1}{2}} \cdot b_{L}^{2}}{\tilde{h}_{L}^{2} \cdot h_{L}^{2}} \longrightarrow 0.
$$

It follows that the sequence $b_{L}$ must converge to zero faster than both $\tilde{h}_{L}$ and $h_{L}$. In addition, $M$ must be large enough that

$$\left( \frac{\tilde{h}_{L}^{M}}{h_{L}^{M}} \right) \cdot L^{1/2} \longrightarrow 0 \quad \text{and} \quad \left( \frac{h_{L}^{M}}{\tilde{h}_{L}^{M}} \right) \cdot L^{1/2} \longrightarrow 0.
$$

We use bandwidth sequences of the form $b_{L} = c_{i} L^{-\alpha_{i}}$, $h_{L} = c_{j} L^{-\alpha_{j}}$, and $\tilde{h}_{L} = c_{j} L^{-\alpha_{j}}$, where $c_{j} > 0$. By inspection, we can show that the smallest integer value of $M$ that satisfies the above restrictions is $M = 16$. If $X$ were to include multiple continuous covariates, the lower bound for $M$ would increase with their dimension.

Appendix A.7 describes precisely the set of conditions under which $\sqrt{L} \cdot U_{L(i)}^{f_{W|Y}}$ and $\sqrt{L} \cdot U_{L(i)}^{g_{W|Y}}$ have asymptotic properties analogous to those outlined in Equation (7) (and described in Theorem 3) for the test without conditioning in $X$. The precise result is stated in Theorem 9. Equipped with this result, we employ rejection rules identical to those described above for the unconditional case.

Our methodology can be extended to the case where $X$ is multivariate, with both discrete and continuous covariates. Our nonparametric estimators would use indicator functions for the discrete components of $X$, and a multivariate (e.g., multiplicative) bias-reducing kernel for the continuous elements. The value of $M$ needed to preserve $\sqrt{L}$-asymptotic normality of $U_{L(i)}^{f_{W|Y}}$ and $U_{L(i)}^{g_{W|Y}}$ would increase with the dimension of continuous components in $X$.

**Results of the Conditional Test**

Coming soon.

---

12Specifically, $b_{L}$ must converge to zero faster than $(\tilde{h}_{L}^{\frac{1}{2}} \cdot h_{L})^{3}$.

13The entire set of bandwidth convergence restrictions would be satisfied for $M = 16$, e.g., if we choose $h_{L} \propto L^{-0.073}$, $\tilde{h}_{L} \propto L^{-0.073}$ and $b_{L} \propto L^{-0.344}$. There is no integer $M < 16$ for which the bandwidth convergence restrictions can be satisfied.
5 Identification

5.1 Theory

Identification of the distribution of bidder valuations is economically important because it allows a seller to forecast expected revenue and to select the optimal (revenue- or profit-maximizing) reserve price. In an IPV setting, as noted above, the marginal distribution of each bidder’s valuations, $F_V$, is identified from $F_{n-1:n}$, so observations of auctions of a single size pin down counterfactual expected revenue at any reserve price, and thus the optimal reserve (which under IPV does not depend on $n$).

When bidder values are correlated, however, the marginal distribution a single bidder’s valuations is no longer identified by bid data from auctions of a single size, nor is it any longer sufficient to characterize expected revenue. Further, correlation in bidder valuations will generally lower the optimal level of reserve price. Intuitively, 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. Formally, Quint (2008) shows that when private values are symmetric and affiliated, both expected revenue at any positive reserve price and the revenue-maximizing reserve price are weakly lower than they would be under an independent private values model consistent with the same distribution $F_{n-1:n}$. However, without further restrictions, neither expected revenue nor optimal reserve is identified, and both potentially vary over a wide range.

However, if auction size varies exogenously, we can use the variation in the distributions $F_{n-1:n}$ across $n$ to identify exactly these objects of interest, along with the ex-ante expected surplus of each bidder. First, we show that all of these quantities require only knowledge of the marginal distributions of order statistics, not the entire (correlated) joint distribution of bidder valuations.

**Theorem 4** In any “standard” auction mechanism with $n$ bidders and reserve price $r$, the seller’s expected profit and each bidder’s ex-ante expected surplus are uniquely determined by $F_{n-1:n}$ and $F_{n:n}$.

**Proof.** By “standard,” we mean, any auction mechanism satisfying the usual conditions for revenue equivalence: 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.

\[14\] In a different but analogous setting, Athey and Halle (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.” Here, however, we do not rely on the joint distribution of $V_{n-1:n}$ and $V_{n:n}$, only the marginal distribution of each.
Consider a second-price sealed-bid auction with \( n \) bidders and reserve price \( r \), with all bidders following their dominant strategy of bidding their valuation. We show in the appendix that expected revenue \( \pi_n(r) \), and each bidder’s ex-ante expected surplus \( u_n(r) \), can be written as

\[
\pi_n(r) = (F_{n-1,n}(r) - F_{n,n}(r)) (r - v_0) + \int_r^{+\infty} (v - v_0) dF_{n-1,n}(v)
\]
\[
u_n(r) = \frac{1}{3} (E \{ \max \{V_{n,n}, r\} \} - E \{ \max \{V_{n-1,n}, r\} \})
\]

where \( v_0 \) is the seller’s valuation for the unsold good (and \( V_{0,1} \) is understood to be 0).

For other auction types, as long as the bidders observe \( \theta \) in addition to their own valuations, then from the bidders’ point of view, conditioning on any realization of \( \theta \), they are in an IPV world; so for a given \( \theta \), any other “standard” auction is revenue-equivalent to the second-price sealed-bid auction. Averaging over \( \theta \), then, this still holds, completing the proof. \( \square \)

Of course, once \( \pi_n(r) \) is known, we can solve \( \max_r \pi_n(r) \) to calculate the optimal reserve price for an auction with a fixed number \( n \) bidders. (Under independent private values, the optimal \( r \) does not depend on \( n \); but with correlated values, it does.) Thus, if he is unable to observe \( \theta \), everything the seller could benefit from knowing, is summarized in the distributions \( F_{n-1:n} \) and \( F_{n:n} \). Next, we show that exogenous variation in auction size allows us to identify \( F_{n:n} \).

**Theorem 5** The distribution \( F_{n:n} \) is uniquely determined by the distributions \( \{F_{m-1:m}\}_{m=n+1}^{\infty} \) if there is exogenous variation in \( N \).

**Proof.** A complete proof is given in the appendix. Exogenous \( N \) leads to the relation discussed in Athey and Haile (2002): for any \( n \):

\[
F_{n:n}(v) = \frac{1}{n+1} F_{n,n+1}(v) + \frac{n}{n+1} F_{n+1,n+1}(v)
\]

Applying this iteratively gives

\[
F_{n:n}(v) = \frac{1}{n-1} \sum_{m=n+1}^{n'} \left( \prod_{i=n}^{m-1} \frac{i-1}{i+1} \right) F_{m-1:m}(v) + \frac{n}{n'} F_{n',n'}(v)
\]

Continuing to iterate, as \( n' \to \infty \), the trailing term \( \frac{n}{n'} F_{n',n'}(v) \) vanishes and \( F_{n:n}(v) \) is pinned down exactly. \( \square \)

As shown in the appendix, for \( j < n-1 \), the distribution \( F_{j:n} \) is also uniquely determined by the distributions \( \{F_{m-1,m}\} \). Thus, if another distribution \( F_{j:n}, j < n-1 \), can be inferred from the data (as from

\[\text{\footnotesize 15}\text{Even without assuming equilibrium play, } F_{1:k}\text{ will hold, at least approximately, for ascending auctions: if bidders do not make jump bids and satisfy the assumptions of Haile and Tamer (2003), then the transaction price in an ascending auction must be within the minimum bid increment of } \max \{r, V_{n-1:n}\}, \text{ and so both } \pi_n \text{ and } u_n \text{ are within a bid increment of the expressions above.}\]

\[\text{\footnotesize 16}\text{To see where this comes from, begin with an auction with } n+1 \text{ bidders, and then choose one at random to remove. With probability } \frac{n}{n+1}, \text{ the bidder with the highest value was removed, in which case the highest of the remaining } n \text{ was the second-highest originally. With probability } \frac{n}{n+1}, \text{ the bidder removed did not have the highest value, so the highest remaining was the highest of the original } n+1.\]
observations of other losing bids besides the highest), the model is overidentified and therefore testable in another way.

We should also note that the proof of Theorem 5 does not require conditionally independent private values, only exogenous \( N \) and private values drawn from an exchangeable (symmetric) distribution.\footnote{De Finetti’s theorem, as formalized by Hewitt and Savage (1955), says that if bidders are drawn at random from an infinite set of potential bidders with a joint distribution of valuations which is exchangeable, then the assumption of conditionally independent values is without loss of generality. Without appealing to infinite bidders, since we place no restrictions on the dimensionality of \( \theta \) or how it affects bidder preferences (other than that it affects them symmetrically), we still feel it is a very general formulation. For a setting like symmetric, affiliated private values, or symmetric but otherwise arbitrarily distributed private values, Theorem 5 would still hold; but the revenue equivalence argument used to prove Theorem 4 would fail, and so it would only hold for ascending auctions (under the maintained assumption that the highest losing bidder bids his valuation), not other auction formats like first-price auctions.}

Theorem 5 does require unbounded variation in auction size to point-identify \( F_{n:n} \). However, with bounded variation, we can still obtain pointwise bounds on \( F_{n:n} \), and therefore on \( \pi_n \) and \( u_n \).

**Theorem 6** Fix \( n \) and \( v \). If \( F_{m-1:m}(v) \in \left[ F_{m-1:m}(v), \overline{F}_{m-1:m}(v) \right] \) for every \( m \in \{ n, n+1, \ldots, n\} \) and auction size varies exogenously, then

\[
F_{n:n}(v) \leq F_{n:n}(v) \equiv \sum_{m=n+1}^{n} \alpha_m F_{m-1:m}(v) + \frac{2}{n} \overline{F}_{n-1:n}(v)
\]

\[
F_{n:n}(v) \geq F_{n:n}(v) \equiv \sum_{m=n+1}^{n} \alpha_m F_{m-1:m}(v) + \frac{2}{n} \underline{F}_{n-1:n}(v)
\]

where \( \alpha_m = \frac{1}{n-1} \prod_{i=n}^{n-1} \frac{1}{i+1} \) for \( m > n \).

**Proof.** Since \( V_{n:n} \geq V_{n-1:n} \), \( F_{n:n}(v) \leq F_{n-1:n}(v) \leq \overline{F}_{n-1:n}(v) \), and so

\[
F_{n:n}(v) = \sum_{m=n+1}^{n} \alpha_m F_{m-1:m}(v) + \frac{2}{n} \overline{F}_{n-1:n}(v)
\]

Next, we will show that \( \underline{F}_{n:n}(v) \geq (\psi^{-1}_{n-1:n}(\overline{F}_{n-1:n}(v)))^{\bar{n}} \); to see this, let \( \psi_{n:n}(s) = s^n \), and note that

\[
(\psi_{n:n} \circ \psi^{-1}_{n-1:n})(t) = \frac{\psi_{n:n}(t)}{\psi_{n-1:n}(t)} = \frac{n^{n-1}}{n(n-1)^{n-2}(1-t)} = \frac{t}{(n-1)(1-t)}
\]

where \( t = \psi^{-1}_{n-1:n}(s) \). \( \frac{t}{(n-1)(1-t)} \) is increasing in \( t \), so \( (\psi_{n:n} \circ \psi^{-1}_{n-1:n})^t \) is increasing, or \( \psi_{n:n} \circ \psi^{-1}_{n-1:n} \) is convex. For a given \( \theta \), the CDF of the greatest of \( n \) independent draws from \( F(v|\theta) \) is \( (F(v|\theta))^n \); so taking the expectation over \( \theta \) implies \( F_{n:n}(v) = E_{\theta} (F(v|\theta))^n \), and applying Jensen’s Inequality gives\footnote{Quint (2008) shows this same lower bound \( F_{n:n}(v) \geq \psi_{n:n} \circ \psi^{-1}_{n-1:n}(F_{n-1:n}(v)) \) for symmetric, affiliated private values, whether or not they are conditionally independent.}
Then
\[ F_{n,n}(v) = \sum_{m=n+1}^{n} a_m F_{m-1,m}(v) + \frac{n}{n} F_{n,n}(v) \]
\[ \geq \sum_{m=n+1}^{n} a_m F_{m-1,m}(v) + \frac{n}{n} (\psi_{n-1,n}^{-1}(F_{n-1,n}(v)))^n \]
\[ \geq \sum_{m=n+1}^{n} a_m F_{m-1,m}(v) + \frac{n}{n} (\psi_{n-1,n}^{-1}(F_{n-1,n}(v)))^n \]
completing the proof. \qed

**Corollary 7** Fix \( n \), and let \( \{ F_{m-1,m}, F_{m-1,m} \}_{m=n}^{n} \) be integrable mappings from \( \mathbb{R}^+ \) to \([0,1] \). If \( F_{m-1,m}(v) \in [F_{m-1,m}(v), F_{m-1,m}(v)] \) for every \( m \in \{n,n+1,\ldots,n\} \) and every \( v \), and auction size varies exogenously, then

1. \( \pi_n(r) \in [\bar{\pi}_n(r), \underline{\pi}_n(r)] \)

2. \( \arg \max_{r} \pi_n(r) \in [\bar{\pi}_n, 1] \), and

3. \( u_n(r) \in [\underline{u}_n(r), \bar{u}_n(r)] \)

where
\[ \bar{\pi}_n(r) = (F_{n-1,n}(r) - F_{n,n}(r)) (r - v_0) + \int_{v_0}^{\infty} (v - v_0) dF_{n-1,n}(v) \]
\[ \underline{\pi}_n(r) = (F_{n-1,n}(r) - F_{n,n}(r)) (r - v_0) + \int_{v_0}^{\infty} (v - v_0) dF_{n-1,n}(v) \]
\[ \bar{\pi}_n = \min \{ r \geq v_0 : \pi_n(r) \geq \max_{r'} \underline{\pi}_n(r') \} \]
\[ \underline{\pi}_n = \max \{ r \geq v_0 : \pi_n(r) \geq \max_{r'} \underline{\pi}_n(r') \} \]
\[ \underline{u}_n(r) = \frac{1}{n} \int_{0}^{\infty} \max\{r,v\} dF_{n,n}(v) - \int_{0}^{\infty} \max\{r,v\} dF_{n-1,n}(v) \]
\[ \bar{u}_n(r) = \frac{1}{n} \int_{0}^{\infty} \max\{r,v\} dF_{n,n}(v) - \int_{0}^{\infty} \max\{r,v\} dF_{n-1,n}(v) \]
and \( F_{n,n} \) and \( F_{n,n} \) are defined as in (16).

**Proof.** For the bounds on \( \pi_n(r) \), Quint (2008) shows that the expression for expected profit (13) is stochastically increasing in both \( V_{n-1,n} \) and \( V_{n,n} \) (in the first-order stochastic dominance sense); so plugging the upper bounds \( F_{n-1,n} \) and \( F_{n,n} \) into (13) gives the lower bound \( \underline{\pi}_n(r) \), and the lower bounds \( E_{n-1,n} \) and \( F_{n,n} \) give the upper bound \( \bar{\pi}_n(r) \). As in Haile and Tamer (2003), we can use \( \pi_n \) and \( \bar{\pi}_n \) to bound the profit-maximizing reserve price for an auction of a given size: letting \( r^*_n = \arg \max_{r'} \bar{\pi}_n(r') \), the true optimal reserve price \( r^*_n \) for an auction with \( n \) bidders must satisfy
\[ \underline{\pi}_n(r^*_n) \geq \pi_n(r^*_n) \geq \bar{\pi}_n(r^*_n) = \max_{r'} \underline{\pi}_n(r') \]
and so \( r^*_n \in \{ r : \pi_n(r) \geq \max_{r'} \underline{\pi}_n(r') \} \).

Finally, by (13), \( u_n \) can be expressed as 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 \( F_{n,n} \) and \( F_{n-1,n} \) into (13) gives the lower bound on \( u_n(r) \), and \( F_{n,n} \) and \( F_{n-1,n} \) yields the upper bound. \qed
It is worth noting that the bounds on $F_{n:n}$, and therefore $\pi_n(r)$ and $r^*_n$, will be tighter when $n$ is lower \textsuperscript{19} – exactly the cases where setting the correct reserve price is most important. To see this, consider the case where the data set is huge but bounded, so that $F_{m-1:m}$ is learned very precisely for $m$ up to an upper bound $M$. By (15), $F_{n:n}(v) = \sum_{m=n+1}^{n} \alpha^m_{n} F_{m-1:m}(v) + \frac{\alpha}{n} F_{\bar{n}:\bar{n}}(v)$, so all uncertainty about $F_{n:n}(v)$ comes from uncertainty about $F_{\bar{n}:\bar{n}}(v)$; but the lower is $n$, the lower is the coefficient $\frac{\alpha}{n}$ on $F_{\bar{n}:\bar{n}}(v)$, and therefore the lower is the uncertainty in $F_{n:n}$.

5.2 Empirics

6 Conclusion

\textsuperscript{19}As $n$ grows, the probabilities that a given reserve price $r$ either binds or precludes a sale both go to 0.
Appendix

A.1 Examples of Endogenous \( N \) Leading to Violations of (4)

Theorem 1 states that if \( N \) is exogenous, \( \psi_{n-1,n}^{-1}(F_{n-1:n}(v)) \) will be increasing in \( n \), our test (4). The logic of our test effectively boils down to: “if \( N \) is endogenous, it will probably lead to a violation of (4); so if the data is satisfies (4), we believe \( N \) to be exogenous.” We argue in the text that if endogeneity of \( N \) comes in the form of positive selection, and this effect is strong enough, then it should lead to a violation of (4). This begs the question, how strong does selection have to be for us to detect it?

To answer this, consider a simple example. Conditional on \( \theta \), bidder values are i.i.d. draws from a truncated normal distribution: take the distribution \( N(\mu_{\theta},\sigma_{\theta}^2) \), truncate it at 0 and 30, and rescale to have total probability 1. Any variability in \( (\mu_{\theta},\sigma_{\theta}^2) \) induces correlation in bidder values; and by making \( \theta \) correlated with \( N \), we can test whether this amount of selection would lead to a violation of (4).

Suppose that \( \theta \) takes two values, \( L \) and \( H \), and that \( \Pr(\theta|N) \) is given by the table below:

\[
\begin{array}{c|cc}
N = 3 & N = 4 \\
\hline
\Pr(\theta = L|N) & p & 1 - p \\
\Pr(\theta = H|N) & 1 - p & p \\
\end{array}
\]

If \( p = \frac{1}{2} \), \( N \) is exogenous, and (4) will hold; if not, \( N \) is endogenous, and we hope that (4) will be violated somewhere, so that our test can detect the endogeneity.

For the first example, suppose \( \theta \) shifts the mean of the distribution: specifically, let \( \mu_H = 20 \), \( \mu_L = 10 \), and \( \sigma_H = \sigma_L = 7.5 \). Since we are mostly concerned with positive selection, we will consider how large \( p \) must be for (4) to be violated. (If \( p < \frac{1}{2} \), (4) will hold, but the selection might be detected by the other test, (3).) Figure 3 below shows \( F_{3,4}(v) - \psi_{3,4} \circ \psi_{2,3}^{-1}(F_{2,3}(v)) \), plotted against \( v \), for various values of \( p \). (4) says that when \( N \) is exogenous, this will be everywhere positive; the top graph, \( p = 0.5 \), confirms this. Our test relies on the hope that if \( p > \frac{1}{2} \), the graph will be negative at least for some values of \( v \).

In fact, even at \( p = 0.51 \), the graph dips below 0 (albeit only slightly) for a significant range of \( v \). As \( p \) increases, the graph gets more negative over a larger range of \( v \), and thus, endogeneity of \( N \) becomes more and more clear; at \( p = 0.54 \), it is negative nearly as much as it’s positive, and at \( p = 0.57 \), it is negative everywhere. For reference, \( p = 0.51 \) implies that the mean valuation of a bidder in a four-bidder auction is 1% higher than the mean valuation of a bidder in a three-bidder auction, \( p = 0.54 \) means it is 4% higher, and \( p = 0.57 \) means it is 7% higher.

For the second example, suppose \( \theta \) shifts \( \sigma^2 \). Specifically, let \( \mu_H = \mu_L = 15 \), but \( \sigma_H = 10 \) and \( \sigma_L = 5 \). In this case, “positive” and “negative” selection do not have as clear meanings, and so we consider \( p \) both greater and less than \( \frac{1}{2} \). Figure 3 again displays \( F_{3,4}(v) - \psi_{3,4} \circ \psi_{2,3}^{-1}(F_{2,3}(v)) \) for various values of \( p \). Again, whenever \( p \neq \frac{1}{2} \), this is negative for some range of \( v \). Thus, our test seems at least theoretically capable of detecting even a small amount of endogeneity in \( N \).
Figure 3: $F_{3,4}(v) - \psi_{3,4} \circ \psi_{2,3}^{-1}(F_{2,3}(v))$ against $v$, for our two examples

Example 1
- $\mu_H = 20$
- $\mu_L = 10$
- $\sigma_H = 7.5$
- $\sigma_L = 7.5$

Example 2
- $\mu_H = 15$
- $\mu_L = 15$
- $\sigma_H = 10$
- $\sigma_L = 5$
A.2 Proof of Footnote 7

The exact claim is that under symmetric, conditionally independent private values with exogenous auction size,

\[ F_{n-1:n}(v) \leq F_{n-2:n-1}(v) = \frac{(n-1)(n-3)(F_{n-3:n-2}(v) - F_{n-2:n-1}(v))^2}{(n-2)^2 F_{n-4:n-3}(v) - F_{n-3:n-2}(v)} \]

Let \( X_n \) denote the statement “\( v_2 > r, v_3 > r, \) and \( v_i \leq r \) for \( i \in \{4, 5, \ldots, n\} \)”. We first show that \( \Pr(v_i \leq r|X_n) \) is increasing in \( n \). To see this, fix \( r \), and let \( G^n \) denote the probability distribution of \( F(r|\theta) \) when the distribution of \( \theta \) is conditioned on the information \( X_n \). (That is, begin with the prior distribution of \( \theta \), apply Bayes’ Law to determine the posterior distribution of \( \theta \) conditional on \( X_n \); \( G^n \) is then the posterior distribution of \( F(r|\theta) \).) Then \( \Pr(v_1 \leq r|X_n) = E_{\theta|X_n} F(r|\theta) = \int p \ dG^n(p) \).

Since \( X_{n+1} \) is just \( X_n \), plus the information that \( v_{n+1} \leq r \), we can generate \( dG^{n+1} \) from \( dG^n \) through Bayes’ Law: abusing notation slightly,

\[ dG^{n+1}(p) = \frac{\Pr(F(r|\theta) = p|X_n) \Pr(v_1 \leq r|F(r|\theta) = p)}{Pr(v_1 \leq r|X_n)} = \frac{p \ dG^n(p)}{\int p' \ dG^n(p')} \]

and so, taking the expectation,

\[ \Pr(v_1 \leq r|X_{n+1}) = \int p \ dG^{n+1}(p) = \frac{\int p^2 \ dG^n(p)}{\int p' \ dG^n(p')} \geq \frac{\Pr(v_1 \leq r|X_n)}{\Pr(v_1 \leq r|X_n)} \]

where the inequality is simply \( E(p^2) \geq (E(p))^2 \) when both expectations are taken with respect to the distribution \( G^n \).

Once we know \( \Pr(v_1 \leq r|X_n) \) is increasing in \( n \), we know that

\[ \frac{\Pr(X_n) \ Pr(v_1 \leq r|X_n)}{\Pr(X_n) \ Pr(v_1 \leq r|X_n)} \]

is decreasing in \( n \). By symmetry,

\[ F_{n-2:n}(r) - F_{n-1:n}(r) = nC_2 \ Pr(v_1 \leq r, v_2 > r, v_3 > r, v_4 \leq 4, \ldots, v_n \leq r) = nC_2 \ Pr(X_n) \ Pr(v_1 \leq r|X_n) \]

\[ F_{n-3:n}(r) - F_{n-2:n}(r) = nC_3 \ Pr(v_1 > r, v_2 > r, v_3 > r, v_4 \leq 4, \ldots, v_n \leq r) = nC_3 \ Pr(X_n) \ Pr(v_1 > r|X_n) \]

so

\[ \frac{1}{nC_2} (F_{n-3:n}(r) - F_{n-2:n}(r)) = \frac{1}{nC_2} (F_{n-2:n}(r) - F_{n-1:n}(r)) \]

is decreasing in \( n \).
Solving \([14]\) for \(F_{j,k}\) gives \(F_{j,k} = \frac{k}{k-j}F_{j:k-1} - \frac{j}{k-j}F_{j+1:k}\) (suppressing the dependence on \(r\)), leading to

\[
F_{n-2:n} - F_{n-1:n} = \frac{n}{2}F_{n-2:n-1} - \frac{n^2}{2}F_{n-1:n} - F_{n-1:n} \\
= \frac{n}{2} (F_{n-2:n-1} - F_{n-1:n})
\]

\[
F_{n-3:n} - F_{n-2:n} = \frac{n}{3}F_{n-3:n-1} - \frac{n^3}{6}F_{n-2:n} - F_{n-2:n} \\
= \frac{n}{3} (F_{n-3:n-1} - F_{n-2:n}) \\
= \frac{n}{3} \left( \frac{n-1}{2}F_{n-3:n-2} - \frac{n^2}{2}F_{n-2:n-1} - \left( \frac{n}{2}F_{n-2:n-1} - \frac{n^2}{2}F_{n-1:n} \right) \right) \\
= \frac{n}{6} \left( (n-1)F_{n-3:n-2} - (2n-3)F_{n-2:n-1} + (n-2)F_{n-1:n} \right) \\
= \frac{n}{6} \left( (n-1)(F_{n-3:n-2} - F_{n-2:n-1}) - (n-2)(F_{n-2:n-1} - F_{n-1:n}) \right)
\]

and so

\[
\frac{1}{nC_3} (F_{n-3:n} - F_{n-2:n}) = \frac{nC_2}{nC_3} \frac{n}{6} \left( (n-1)(F_{n-3:n-2} - F_{n-2:n-1}) - (n-2)(F_{n-2:n-1} - F_{n-1:n}) \right) \\
= \frac{1}{n-2} \left( \frac{n-1}{2} \left( F_{n-2:n-1} - F_{n-1:n} \right) - (n-2) \right) \\
= \frac{(n-1)(F_{n-3:n-2} - F_{n-2:n-1})}{(n-2)(F_{n-2:n-1} - F_{n-1:n})} - 1
\]

must be decreasing in \(n\). Solving

\[
\frac{(n-1)(F_{n-3:n-2} - F_{n-2:n-1})}{(n-2)(F_{n-2:n-1} - F_{n-1:n})} \leq \frac{(n-2)(F_{n-4:n-3} - F_{n-3:n-2})}{(n-3)(F_{n-3:n-2} - F_{n-2:n-1})}
\]

for \(F_{n-1:n}\) gives the upper bound.

### A.3 Finishing the Proof of Theorem 2

We showed in the text that \(\mu^{W:N} > 0\) if and only if the event \(\{N_i > N_j \text{ and } \Delta_{W:N}(W_k, N_i, N_j) > 0\}\) has positive probability under the true data-generating process (and likewise for \(\mu^{\Delta_{W:N}}\) and \(\Phi(\cdot)\)). Here, we show that the qualifier “with positive probability” is unnecessary, at least with regard to \(W_k\). Assume that \(F_N(n) > 0\) for every \(n \in S_N\), or else redefine \(S_N\) to exclude those \(n\) which occur with zero probability. Suppose that for some \(n > n'\), there is some \(w\) such that \(F_{W:N}(w|n) > F_{W:N}(w|n')\); but make no assumptions about \(w\). Define

\[
\epsilon = \frac{1}{2} \left( F_{W:N}(w|n) - F_{W:N}(w|n') \right) \\
w^* = \sup \{ w' : F_{W:N}(w'|n') \leq F_{W:N}(w|n') \} \\
w^{**} = \inf \{ w' : F_{W:N}(w'|n') \geq F_{W:N}(w|n') + \epsilon \}
\]

Since \(F_{W:N}(W_k|N_k) \sim U[0, 1]\), if \(N_k = n'\), then

\[
\Pr(W_k \in (w^*, w^{**})) = \Pr(F_{W:N}(W_k|N_k) \in (F_{W:N}(w|n'), F_{W:N}(w|n') + \epsilon)) = \epsilon
\]
This implies $W_k \geq w^* \geq w$, therefore (since $F_{W|N}(\cdot|n)$ is increasing) $F_{W|N}(W_k|n) \geq F_{W|N}(w|n)$, therefore

$$\Delta_{W|N}(W_k,n,n') \geq F_{W|N}(w|n) - (F_{W|N}(w'|n) + \epsilon) = 2\epsilon - \epsilon = \epsilon$$

So with probability $P_n \cdot (P_n')^2 \cdot \epsilon$, $N_i = n$, $N_j = n'$, and $W_k \in (w^*,w^{**})$, which together imply $\Delta_{W|N}(W_k,N_i,N_j) \geq \epsilon$. Thus, if $F_{W|N}(w|n) > F_{W|N}(w'|n)$ for any $(n,n',w)$ with $n > n'$ and $n,n' \in S_N$, then $F^*(N_i,Z_j,Z_k) \geq \epsilon$ with strictly positive probability over $(N_i,Z_j,Z_k)$, and thus $\mu^F > 0$. The same arguments establish that $\mu^{\Delta_{W|N}} > 0$ if $\Phi(w,n,n') > 0$ for any $w$ and any $n > n'$ with $n,n' \in S_N$.

### A.4 Proof of Theorem 4

In a second-price sealed-bid auction, bidders have a dominant strategy of bidding their valuations. Seller profits are therefore 0 if $V_{n:n} < r$, $r - v_0$ if $V_{n:n} \geq r \geq V_{n-1:n}$, and $V_{n-1:n} - v_0$ if $V_{n:n} \geq V_{n-1:n} > r$ (where $v_0$ is the seller’s cost and $V_{0:1}$ is understood to be 0). We can therefore write expected profits as

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

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

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

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 $V_{0:1}$ under stood to be 0). So

$$u_n(r) = \frac{1}{n} E_{V_{n-1:n},V_{n:n}} \left\{ 1_{V_{n:n} \geq r} (V_{n:n} - \max\{V_{n-1:n}, r\}) \right\}$$

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

$$= \frac{1}{n} E_{V_{n-1:n},V_{n:n}} \left\{ 1_{V_{n:n} \leq r} + 1_{V_{n:n} > r} V_{n:n} + 1_{V_{n-1:n} > r} V_{n-1:n} \right\}$$

$$= \frac{1}{n} \left( E \left\{ \max\{V_{n:n}, r\} \right\} - E \left\{ \max\{V_{n-1:n}, r\} \right\} \right)$$

As argued in the text, conditional on a realization of $\theta$, bidder values are IPV, and so other “standard” auctions are revenue-equivalent to a sealed-bid second-price auction.

### A.5 Proof of Theorem 5

The proof is based on the relation discussed in [Athey and Haile (2002)] and in the text above as [14].

$$F_{n:n}(v) = \frac{1}{n+1} F_{n:n+1}(v) + \frac{n}{n+1} F_{n-1:n+1}(v)$$
To identify \( F_{n,n} \), we fix \( n > 1 \) and use induction on \( n' \) to show \([15]\),

\[
F_{n,n}(v) = \frac{1}{n-1} \sum_{m=n+1}^{n'} \left( \prod_{i=n}^{m-1} i \right) F_{m-1,m}(v) + \frac{n}{n'} F_{n',n'}(v)
\]

for any \( n' > n \). First, the base case, \( n' = n+1 \). 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 by \([14]\) is equal to \( F_{n,n}(v) \).

For the inductive step, if \([15]\) holds for \( n' = K \), then

\[
\frac{1}{n-1} \sum_{m=n+1}^{K+1} \left( \prod_{i=n}^{m-1} i \right) F_{m-1,m}(v) + \frac{n}{K+1} F_{K+1,K+1}(v)
\]

are again equal to \( F_{n,n}(v) \) since by \([14]\), the terms in parentheses sum to 0; so \([15]\) holds for \( n' = K+1 \).

A more general version of \([14]\) is for any \( j < k \), \( F_{j,k-1}(v) = \frac{k-i}{k} F_{j,k}(v) + \frac{i}{k} F_{j+1,k}(v) \), which we can rearrange to

\[
F_{k-i,k}(v) = \frac{k}{i} F_{(k-1)-(i-1):k-1}(v) - \frac{k-i}{i} F_{k-(i-1):k}(v)
\]

As noted in the text, we can use this to pin down \( F_{k-i,k} \) for \( i > 1 \) by induction on \( i \); once \( F_{k'-(i-1):k'}(v) \) is known for every \( k' \), \( F_{k-i,k}(v) \) can be calculated directly. So \( \{F_{n-3,n}\} \) are pinned down from \( \{F_{n-1,n}\}, \{F_{n-2,n}\} \), and so on.

### A.6 Asymptotic properties of \( \hat{U}^{PW|N}_{L(3)} \) and \( \hat{U}^{DW|N}_{L(3)} \)

Rewriting \( \hat{U}^{PW|N}_{L(3)} \) and \( \hat{U}^{DW|N}_{L(3)} \) as Symmetric U-statistics

It is convenient to re-write \( \hat{U}^{PW|N}_{L(3)} \) and \( \hat{U}^{DW|N}_{L(3)} \) as a symmetric U-statistics. For any function \( f \) with three arguments \((x_1, x_2, x_3)\) (where \( x_t \) may be vector-valued), define a function

\[
\mathcal{H}^f_{(3)}(x_1, x_2, x_3) = \frac{1}{3!} \sum_{s_3} f(x_{s_1}, x_{s_2}, x_{s_3})
\]

where \( \sum_{s_3} \) denotes a sum over the \( 3! \) permutations \((s_1, s_2, s_3)\) of \((1, 2, 3)\). Note that \( \mathcal{H}^f_{(3)} \) is symmetric in its arguments. We will use the following notation to denote the third order U-statistic with kernel function \( f \):

\[
\hat{U}^f_{L(3)} = \binom{L}{3}^{-1} \sum_{i<j<k} \mathcal{H}^f_{(3)}(Z_i, Z_j, Z_k)
\]

As usual, \( \sum_{i<j<k} \) denotes the sum over the \( \binom{L}{3} \) combinations of 3 distinct elements \((i, j, k)\) from \((1, \ldots, L)\). It is easy to see that the expressions for \( \hat{U}^{PW|N}_{L(3)}, \hat{U}^{DW|N}_{L(3)}, \hat{U}^{PW|N}_{L(3)} \) and \( \hat{U}^{DW|N}_{L(3)} \) can be rewritten in terms
Fourth-order U-statistics will show up in our asymptotic results. For any function $g$ with four arguments $(x_1, x_2, x_3, x_4)$ we will let

$$
H^g_{(4)}(x_1, x_2, x_3, x_4) = \frac{1}{4!} \sum_{c_4} g(x_{r_1}, x_{r_2}, x_{r_3}, x_{r_4})
$$

where $\sum_{c_4}$ denotes a sum over the $4!$ permutations $(r_1, r_2, r_3, r_4)$ of $(1, 2, 3, 4)$. We will let

$$
U^g_{L(4)} = \binom{L}{4}^{-1} \sum_{1 < j < k < \ell} H^g_{(4)}(Z_i, Z_j, Z_k, Z_\ell).
$$

### Assumptions

We begin by formalizing the basic distribution assumptions, all of which are compatible with the conditions in Section A.6.2

**Assumption T1**

We observe an iid sample $(W_i, N_i)_{i=1}^L \equiv (Z_i)_{i=1}^L$ of winning bids (transaction prices) $W$ and auction participants ‘$N$’ produced by a data generating process that satisfies the assumptions in Section 3.2. The support $\mathcal{S}_Z$ of $Z$ is compact. The distribution $F_{W|N}(w|n)$ is continuous in $w$ for each $n \in \mathcal{S}_N$.

If $\mathcal{S}_Z$ is not compact, we can take any compact subset $\mathcal{Z}$ such that $P_N(n) \equiv \Pr(N = n)$ is bounded away from zero for all $n \in \mathcal{Z}$ and test whether the inequalities in $[3]$ and $[4]$ hold with probability one in $\mathcal{Z}$.

Next we will impose some regularity assumptions concerning the distribution of $(W, n)$, all of which are entirely compatible with our theoretical model. Let us introduce the statistics and functionals involved in our assumptions. Denote

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

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

$$
\begin{align*}
\phi_{W|N}(Z, w, n) &= \frac{\mathbb{I}\{W \leq u\} - F_{W|N}(w|n)}{P_N(n)}, \quad \mathbb{I}\{N = n\},
\phi_{W|N}(Z, w, n, n') &= \phi_{W|N}(Z, w, n) - \phi_{W|N}(Z, w, n'),
\phi_{W|N}(Z, w, n, n') &= \phi_{W|N}(Z, w, n') \cdot \nabla_1 \Omega(F_{W|N}(w|n'), n, n') - \phi_{W|N}(Z, w, n).
\end{align*}
$$

Note that $E[\phi_{W|N}(Z, w, n)|N] = E[\phi_{W|N}(Z, w, n, n')|N] = E[\phi_{W|N}(Z, w, n, n')|N] = 0$. Let

$$
\begin{align*}
\xi_{-i,j,k}^L(w, n, n') &= \Delta_{W|N}(-i,j,k)(w, n, n') - \Delta_{W|N}(w, n, n') - \frac{1}{L-3} \sum_{\ell \neq i,j,k} \phi_{W|N}(Z_\ell, w, n, n'),
\tilde{\xi}_{-i,j,k}^L(w, n, n') &= \tilde{\Delta}_{W|N}(-i,j,k)(w, n, n') - \Phi_{W|N}(w, n, n') - \frac{1}{L-3} \sum_{\ell \neq i,j,k} \phi_{W|N}(Z_\ell, w, n, n').
\end{align*}
$$

The asymptotic behavior of $\xi_{-i,j,k}^L(\cdot)$ corresponds to that of the remainder term in a first order Taylor expansion of $\tilde{\Delta}_{W|N}(-i,j,k)(\cdot)$ around $\Delta_{W|N}(\cdot)$. The same interpretation holds for $\tilde{\xi}_{-i,j,k}^L(\cdot)$ and a first order
approximation of $\hat{\Phi}_{W|N}^{-1,j,k}(\cdot)$ around $\Phi_{W|N}(\cdot)$. We will show below that for any $i \neq j \neq k$,

$$
\sup_{w \in S_N, n,n' \in S_N} |\xi_l^{i,j,k}(w, n, n')| = O_p \left( \frac{1}{L} \right), \quad \sup_{w \in S_N, n,n' \in S_N} |\hat{\xi}_l^{i,j,k}(w, n, n')| = O_p \left( \frac{1}{L} \right).
$$

Assumption T2

The following holds for any $i \neq j \neq k$.

(i) Let $\ell \in (i,j,k)$. There exists a $\tau > 0$ and $\bar{\mathcal{H}} > 0$ such that, for any $n,n' \in S_N^2$,

$$
\Pr \left( -s \leq \Delta_{W|N}(W, n, n') < 0 \bigg| \xi_l^{i,j,k}(W, n, n') \right) \leq |s| \cdot \bar{\mathcal{H}} \quad \forall \ s \leq \tau.
$$

(ii) Take any $(w, n, n')$ such that $n,n' \in S_N^2$. For $\ell \notin (i,j,k)$ let

$$
\xi_l^{i,j,k,\ell}(w, n, n') = \tilde{\Delta}_l^{i,j,k,\ell}(w, n, n') - \Delta_{W|N}(w, n, n') - \frac{1}{L - 3} \sum_{\ell \neq i,j,k} \phi_{\Delta_{W|N}^N}(Z_m, w, n, n'),
$$

$$
\gamma_l^{i,j,k}(Z, w, n, n') = \xi_l^{i,j,k}(w, n, n') - \xi_l^{i,j,k,\ell}(w, n, n').
$$

That is, $\xi_l^{i,j,k,\ell}(\cdot)$ would be the resulting value of $\xi_l^{i,j,k}(\cdot)$ if we drop the $\ell$th observation entirely from its computation. Then, there exists a $\mathcal{J} < \infty$ such that, for all $t \equiv (w, n, n') \in S_W \times S_N^2$,

$$
\left| E \left[ \phi_{\Delta_{W|N}^N}(Z_t, t) \bigg| \xi_l^{i,j,k,\ell}(t) + \beta \gamma_l^{i,j,k}(Z_t) \right] - E \left[ \phi_{\Delta_{W|N}^N}(Z_t, t) \bigg| \xi_l^{i,j,k,\ell}(t) + \beta' \gamma_l^{i,j,k}(Z_t) \right] \right| \\
\leq |\beta - \beta'| \cdot |\gamma_l^{i,j,k}(Z_t)| \cdot \mathcal{J} \quad \forall \ (\beta, \beta') \in [0,1].
$$

Our approach explicitly allows for the case where the inequalities in either (3) or (4) are binding with strictly positive probability. In the context of (3), this means that we allow for the existence of $n,n' \in S_N^2$ such that $\Delta_{W|N}(W, n, n')$ has a point mass at zero. Assumption (T2.i) essentially requires that the density of $\Delta_{W|N}(W, n, n') \mid \xi_l^{i,j,k}(W, n, n')$ be bounded in a semi open interval of the form $[\tau, 0)$. This condition is assumed to hold for all $(n,n') \in S_N^2$. Assumption (T2.ii) will help ensure that

$$
\sup_{w \in S_N, n,n' \in S_N^2} \left| E \left[ \phi_{\Delta_{W|N}^N}(Z_t, w, n, n') \bigg| \xi_l^{i,j,k}(w, n, n') \right] \right| = O_p \left( \frac{1}{L} \right).
$$

This result will turn out to be helpful to prove our main result.

Assumption (T2) is relevant for the inequalities in (4). The following conditions are analogous for (4).

Assumption T2'

The following holds for any $i \neq j \neq k$.

(i) Take any $\ell \in (i,j,k)$. There exists a $\tau > 0$ and $\bar{\mathcal{H}} > 0$ such that, for any $n,n' \in S_N$,

$$
\Pr \left( -s \leq \Phi_{W|N}(W, n, n') < 0 \bigg| \tilde{\xi}_l^{i,j,k}(W, n, n') \right) \leq |s| \cdot \bar{\mathcal{H}} \quad \forall \ s \leq \tau.
$$
(ii) Take any \((w,n,n')\) such that \(n,n'\in S_N\). For each \(\ell \neq (i,j,k)\) let
\[
\tilde{\gamma}_{1-i,j,k}^{\ell}(w,n,n') = \hat{\Phi}_{W|N}^{\ell}(w,n,n') - \Phi_{W|N}(w,n,n') - \frac{1}{L-3} \sum_{m \neq i,j,k,\ell} \varphi_{W|N}(Z_m,w,n,n'),
\]
\[
\tilde{\gamma}_{1-i,j,k}^{\ell}(Z_\ell,w,n,n') = \tilde{\gamma}_{1-i,j,k}^{\ell}(w,n,n') - \tilde{\gamma}_{1-i,j,k}^{\ell}(w,n,n').
\]
That is, \(\tilde{\gamma}_{1-i,j,k}^{\ell}()\) would be the resulting value of \(\tilde{\gamma}_{1-i,j,k}^{\ell}()\) if we drop the \(\ell\)th observation entirely from its computation. Then, there exists a \(\tilde{J} < \infty\) such that, for all \(t \equiv (w,n,n') \in S_W \times S_N \times S_N\),
\[
\left| E\left[ \varphi_{W|N}(Z_t,t) \left| \tilde{\gamma}_{1-i,j,k}^{\ell}(t) + \beta \cdot \tilde{\gamma}_{1-i,j,k}^{\ell}(Z_t,t) \right. \right] - E\left[ \varphi_{W|N}(Z_t,t) \left| \tilde{\gamma}_{1-i,j,k}^{\ell}(t) + \beta' \cdot \tilde{\gamma}_{1-i,j,k}^{\ell}(Z_t,t) \right. \right] \right| \\
\leq |\beta - \beta'| \cdot |\tilde{\gamma}_{1-i,j,k}^{\ell}(Z_t,t)| \cdot \tilde{J} \quad \forall (\beta,\beta') \in [0,1].
\]
Assumption (T2') accomplishes the same results for \([4]\) as Assumption (T2) does for \([3]\).

**Assumption T3**

\(b_L\) is a positive sequence that satisfies \(\sqrt{L} \cdot b_L \to \infty\) and \(\sqrt{L} \cdot b_L^2 \to 0\).

**Main result**

Let
\[
g_{W|N}(Z_i,Z_j,Z_k,Z_\ell) = \varphi_{W|N}(Z_i,W_t,N_k) \cdot \mathbb{1}\{N_i = N_k\} \cdot \mathbb{1}\{N_j > N_k\} \cdot \mathbb{1}\{\Delta_{W|N}(W_t,N_j,N_k) \geq 0\},
\]
\[
g_{D|W|N}(Z_i,Z_j,Z_k,Z_\ell) = \varphi_{W|N}(Z_i,W_t,N_k) \cdot \nabla_1 \Omega(F_{W|N}(W_t|N_k),N_j,N_k) \cdot \mathbb{1}\{N_j > N_k\} \cdot \mathbb{1}\{\Phi_{W|N}(W_t,N_j,N_k) \geq 0\}.
\]
For any \(i \neq j \neq k \neq \ell\) we have \(\overset{20}{=} E[g_{W|N}(Z_i,Z_j,Z_k,Z_\ell)] = E[g_{D|W|N}(Z_i,Z_j,Z_k,Z_\ell)] = 0\).

**Theorem 8** Take \(i \neq j \neq k \neq \ell\).

(i) If Assumptions (T1), (T2) and (T3) hold, then
\[
\sqrt{L} \cdot U_{L(3)}^{\hat{\mu}}_{W|N} = \sqrt{L} \cdot \mu_{W|N} + O_p(1).
\]
Let
\[
\delta_{W|N}(Z_i) = E[g_{W|N}(Z_i,Z_j,Z_k,Z_\ell)|Z_i] - E[g_{W|N}(Z_i,Z_j,Z_k,Z_\ell)|Z_i].
\]
Suppose \([3]\) is satisfied with probability one. Then,
\[
\sqrt{L} \cdot U_{L(3)}^{\hat{\delta}}_{W|N} = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta_{W|N}(Z_i) + o_p(1),
\]
and \(\frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta_{W|N}(Z_i) \xrightarrow{d} {N}(0, \Sigma_{F_{W|N}})\), with \(\Sigma_{F_{W|N}} > 0\) if and only if
\[
E\left[ \mathbb{1}\{N_i > N_j\} \cdot \mathbb{1}\{\Delta_{W|N}(W_k,N_i,N_j) = 0\} \right] > 0.
\]
That is, \(\Sigma_{F_{W|N}} > 0\) if \([3]\) is binding with nonzero probability. Otherwise, \(\Sigma_{F_{W|N}} = 0\).

\(20\)Note that \(E[g_{W|N}(Z_i,Z_j,Z_k,Z_\ell)|N_i,Z_j,Z_k,Z_\ell] = 0\) and \(E[g_{D|W|N}(Z_i,Z_j,Z_k,Z_\ell)|N_i,Z_j,Z_k,Z_\ell] = 0\).
(ii) If Assumptions (T1), (T2') and (T3) hold, and if $\nabla_1 \Omega(F_{W|N}(w|n'), n, n')$ is bounded for all $n, n' \in S_N^2$ and all $w \in S_W$ (otherwise we can modify our test and restrict ourselves to a region of $S_W \times S_N$ where this is assumed to hold), then

$$\sqrt{L} \cdot U_{L(3)}^{\Omega_{W|N}} = \sqrt{L} \cdot \mu_{\Omega_{W|N}} + O_p(1).$$

Let

$$\delta_{\Omega_{W|N}}(Z_i) = E[T_{\Omega_{W|N}}(Z_i, Z_j, Z_k)|Z_i] + E[g_{\Omega_{W|N}}(Z_i, Z_j, Z_k)|Z_i].$$

Suppose (4) is satisfied with probability one. Then,

$$\sqrt{L} \cdot U_{L(3)}^{\Omega_{W|N}} = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta_{\Omega_{W|N}}(Z_i) + o_p(1),$$

and

$$\frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta_{\Omega_{W|N}}(Z_i) \xrightarrow{d} N(0, \Sigma_{\Omega_{W|N}}), \quad \text{with} \quad \Sigma_{\Omega_{W|N}} > 0 \quad \text{if and only if} \quad E\left[\mathbb{I}\{N_i > N_j\} \cdot \mathbb{I}\{\Phi_{W|N}(W_i, N_i, N_j) = 0\}\right] > 0.$$

That is, $\Sigma_{\Omega_{W|N}} > 0$ if (4) is binding with nonzero probability. Otherwise, $\Sigma_{\Omega_{W|N}} = 0$.

### Analytical Estimators for $E[\delta_{F_{W|N}}(Z_i)^2]$ and $E[\delta_{\Omega_{W|N}}(Z_i)^2]$

Let $g_{F_{W|N}}$ and $g_{\Omega_{W|N}}$ be nonparametric estimates of $g_{F_{W|N}}$ and $g_{\Omega_{W|N}}$. These can be constructed analogously to $\hat{T}_{F_{W|N}}$ and $\hat{T}_{\Omega_{W|N}}$ (see Equation 10). Let

$$\hat{\delta}_{F_{W|N}}(Z_i) = \frac{1}{(L-1)(L-2)} \sum_{j \neq i} \sum_{k \neq i, k \neq j} \hat{T}_{F_{W|N}}(Z_i, Z_j, Z_k) - \frac{1}{(L-1)(L-2)(L-3)} \sum_{j \neq i} \sum_{k \neq i, k \neq j} \sum_{\ell \neq i, \ell \neq j, \ell \neq k} \hat{g}_{F_{W|N}}(Z_i, Z_j, Z_k, Z_\ell),$$

and

$$\hat{\delta}_{\Omega_{W|N}}(Z_i) = \frac{1}{(L-1)(L-2)} \sum_{j \neq i} \sum_{k \neq i, k \neq j} \hat{T}_{\Omega_{W|N}}(Z_i, Z_j, Z_k) - \frac{1}{(L-1)(L-2)(L-3)} \sum_{j \neq i} \sum_{k \neq i, k \neq j} \sum_{\ell \neq i, \ell \neq j, \ell \neq k} \hat{g}_{\Omega_{W|N}}(Z_i, Z_j, Z_k, Z_\ell).$$

We can use

$$\hat{E}[\delta_{F_{W|N}}(Z_i)^2] \equiv \hat{\Sigma}_p = \frac{1}{L} \sum_{i=1}^{L} \hat{\delta}_{F_{W|N}}(Z_i)^2, \quad \text{and} \quad \hat{E}[\delta_{\Omega_{W|N}}(Z_i)^2] \equiv \hat{\Sigma}_n = \frac{1}{L} \sum_{i=1}^{L} \hat{\delta}_{\Omega_{W|N}}(Z_i)^2. \quad (24)$$

The same arguments leading to the proof of Theorem 8 can be used to establish consistency of these estimators.

### A.6.2 Preliminary results to prove Theorem 8

We will prove in detail part (i) of Theorem 8. Then we will outline how part (ii) is proved by using parallel arguments. Let $P_N = \min_{w \in S_N} \{P_N(n)\}$. We have $P_N > 0$ by compactness of $S_N$. Take any $i \neq j \neq k$. Using
results from empirical process theory (see Pakes and Pollard (1989) and Andrews (1994)), compactness of \( \mathcal{S}_N \) and \( \mathcal{S}_W \), and a linear approximation yield

\[
\hat{F}_{W|N}(w|n) = F_{W|N}(w|n) + \frac{1}{L-3} \sum_{\ell \neq i,j,k} \phi_{W|N}(Z_\ell, w, n) + \xi_{-i,j,k}(w, n),
\]

where \( \sup_{\ell \in \mathcal{S}_W} |\xi_{-i,j,k}(w, n)| = O_p(L^{-1}). \) (25)

From here we have

\[
\hat{\Delta}_{W|N}(w, n, n') = \Delta_{W|N}(w, n, n') + \frac{1}{L-3} \sum_{\ell \neq i,j,k} \phi_{W|N}(Z_\ell, w, n, n') + \xi_{-i,j,k}(w, n, n'),
\]

where \( \sup_{(n,n') \in \mathcal{S}_N^2} |\xi_{-i,j,k}(w, n, n')| = O_p(L^{-1}). \) (26)

Furthermore, for any \((i, j, k)\) and \(\ell\) in 1, ..., \(L\) we also have

\[
\sup_{w \in \mathcal{S}_W \atop (n,n') \in \mathcal{S}_N^2} |\xi_{-i,j,k}(w, n, n') - \xi_{-i,j,k,\ell}(w, n, n')| = O_p\left(\frac{1}{L}\right).
\]

Letting \(\gamma_{-i,j,k}(Z_\ell, w, n, n') \equiv \xi_{-i,j,k}(w, n, n') - \xi_{-i,j,k,\ell}(w, n, n')\), Assumption (T2) yields

\[
\sup_{w \in \mathcal{S}_W \atop (n,n') \in \mathcal{S}_N^2} \left|E\left[\phi_{W|N}(Z_\ell, w, n, n') \xi_{-i,j,k}(w, n, n')\right]\right| \leq \frac{1}{L} \zeta_L, \quad \text{where} \quad \zeta_L = O_p(1). \quad (27)
\]

From \(26\), there exists a sequence \(\zeta_L = O_p(L^{-1})\) such that

\[
\max_{(i,j,k) \in \{1, \ldots, L\}} \left\{ \sup_{w \in \mathcal{S}_W \atop (n,n') \in \mathcal{S}_N^2} |\xi_{-i,j,k}(w, n, n')| \right\} = \zeta_L, \quad (28)
\]

A quick inspection of Equation 22 shows that \(\zeta_L \leq 4 + \frac{8}{L} \sum_n\mathcal{D}_\ell\), where \(\mathcal{D}_\ell = O_p(1)\). Since \(\zeta_L\) is bounded, it must satisfy \(E[\zeta_L^4] < \infty\) for every \(L\). This is sufficient (but not necessary) to ensure that \(\lim_{a \to \infty} a^2 \cdot \Pr[\mathcal{D}_\ell \geq a] < \infty\) for each \(L\). This implies, in turn, that \((b_L L)^2 \cdot \Pr[\mathcal{D}_\ell \geq b_L] = O(1)\). Therefore,

\[
\Pr[\zeta_L \geq b_L] = \Theta\left(\frac{1}{(b_L L)^2}\right), \quad \text{and} \quad \Pr[\zeta_L \geq b_L] = o_p(L^{-1/2}), \quad (29)
\]

where the last result follows from Chebyshev’s inequality and Assumption T3.

**A convenient expression for** \(\hat{U}_{L(0)}^{\hat{F}_{W|N}}\)

Let

\[
\varepsilon_{-i,j,k}(w, n, n') \equiv b_L + |\xi_{-i,j,k}(w, n, n')|, \quad \text{and} \quad \tau_L \equiv b_L + \zeta_L. \quad (30)
\]
It will be convenient to decompose \( \hat{T}_{L}^{F_{W|N}} \) as follows,

\[
\hat{T}_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) = T_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) + \tilde{V}_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) + S_{1}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) + S_{2}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) + \bar{Q}_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) + \bar{M}_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}),
\]

where

\[
\tilde{V}_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) = - \left( \tilde{F}_{L}^{N}(W_{k}|N_{j}) - F_{W|N}(W_{k}|N_{j}) \right) \mathbb{1} \{ N_{i} > N_{j} \} \mathbb{1} \{ \Delta_{W|N}(W_{k}, N_{i}, N_{j}) \geq 0 \},
\]

\[
S_{1}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) = \left( \mathbb{1} \{ W_{i} \leq W_{k} \} - F_{W|N}(W_{k}|N_{j}) \right) \mathbb{1} \{ N_{i} > N_{j} \} \cdot \mathbb{1} \{ \tilde{\Delta}_{W|N}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \geq -b_{L} \}
\times \mathbb{1} \{ \Delta_{W|N}(W_{k}, N_{i}, N_{j}) < -\varepsilon_{L}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \},
\]

\[
S_{2}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) = \left( \mathbb{1} \{ W_{i} \leq W_{k} \} - F_{W|N}(W_{k}|N_{j}) \right) \mathbb{1} \{ N_{i} > N_{j} \} \cdot \mathbb{1} \{ \tilde{\Delta}_{W|N}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \geq -b_{L} \}
\times \mathbb{1} \{ -\varepsilon_{L}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \leq \Delta_{W|N}(W_{k}, N_{i}, N_{j}) < 0 \},
\]

\[
\bar{Q}_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) = - \left( \mathbb{1} \{ W_{i} \leq W_{k} \} - F_{W|N}(W_{k}|N_{j}) \right) \mathbb{1} \{ N_{i} > N_{j} \} \cdot \mathbb{1} \{ \tilde{\Delta}_{W|N}^{-i, j, k}(W_{k}, N_{i}, N_{j}) < -b_{L} \} \cdot \mathbb{1} \{ \Delta_{W|N}(W_{k}, N_{i}, N_{j}) \geq 0 \},
\]

\[
\bar{M}_{L}^{F_{W|N}}(Z_{i}, Z_{j}, Z_{k}) = - \left( \tilde{F}_{L}^{N}(W_{k}|N_{j}) - F_{W|N}(W_{k}|N_{j}) \right) \cdot \mathbb{1} \{ N_{i} > N_{j} \} \cdot \left( \mathbb{1} \{ \tilde{\Delta}_{W|N}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \geq -b_{L} \} - \mathbb{1} \{ \Delta_{W|N}(W_{k}, N_{i}, N_{j}) \geq 0 \} \right)
\times \mathbb{1} \{ \Delta_{W|N}(W_{k}, N_{i}, N_{j}) < -\varepsilon_{L}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \}.
\]

Accordingly, we can express

\[
U_{L}^{F_{W|N}} = U_{L}^{T_{L}^{F_{W|N}}} + U_{L}^{\tilde{V}_{L}^{F_{W|N}}} + U_{L}^{S_{1}^{F_{W|N}}} + U_{L}^{S_{2}^{F_{W|N}}} + U_{L}^{\bar{Q}_{L}^{F_{W|N}}} + U_{L}^{\bar{M}_{L}^{F_{W|N}}}. \quad (32)
\]

Sections A.6.3 and A.6.6 will be devoted to showing that \( U_{L}^{S_{1}^{F_{W|N}}} = o_{p}(L^{-1/2}) \), \( U_{L}^{S_{2}^{F_{W|N}}} = o_{p}(L^{-1/2}) \), \( U_{L}^{\bar{Q}_{L}^{F_{W|N}}} = o_{p}(L^{-1/2}) \) and \( U_{L}^{\bar{M}_{L}^{F_{W|N}}} = o_{p}(L^{-1/2}) \).

\[
A.6.3 \quad U_{L}^{S_{1}^{F_{W|N}}} = o_{p}(L^{-1/2})
\]

By (26), if \( \Delta_{W|N}(W_{k}, N_{i}, N_{j}) < -\varepsilon_{L}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \), we have \( \tilde{\Delta}_{W|N}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \geq -b_{L} \) only if

\[
\left\{ \frac{1}{(L-3)} \sum_{\ell \neq i, j, k} \varphi_{\Delta_{W|N}}(Z_{i}, W_{k}, N_{i}, N_{j}) \geq 2b_{L} \right\} \cdot \mathbb{1} \{ \Delta_{W|N}(W_{k}, N_{i}, N_{j}) < -\varepsilon_{L}^{-i, j, k}(W_{k}, N_{i}, N_{j}) \}.\]
Thus, for any $\delta > 0$ we have
\[
\Pr[\sqrt{L} | \beta_{L,(3)}^W | > \delta] \\
\leq \Pr\left[ \sum_{\ell \neq i,j,k} \varphi^{W,N}(Z_{\ell}, W_k, N_i, N_j) \geq 2b_k \right] \cdot \mathbf{1}\left\{ \Delta_{W|N}(W_k, N_i, N_j) < -\varepsilon_{L}^{-i,j,k}(W_k, N_i, N_j) \right\} \neq 0
\]
for some $(i,j,k)$ in $(1, \ldots, L)$
\[
\leq \sum_{i=1}^{L} \sum_{j=1}^{L} \sum_{k=1}^{L} \Pr\left[ \sum_{\ell \neq i,j,k} \varphi^{W,N}(Z_{\ell}, W_k, N_i, N_j) \geq 2b_k \right] \cdot \mathbf{1}\left\{ \Delta_{W|N}(W_k, N_i, N_j) < -\varepsilon_{L}^{-i,j,k}(W_k, N_i, N_j) \right\} \neq 0
\]
\[
\leq \sum_{i=1}^{L} \sum_{j=1}^{L} \sum_{k=1}^{L} \Pr\left[ \sum_{\ell \neq i,j,k} \varphi^{W,N}(Z_{\ell}, W_k, N_i, N_j) \geq 2b_k \right].
\]

The second bound follows from a Bonferroni inequality. Recall that $\varphi^{W,N}(Z_{\ell}, W_k, N_i, N_j) \leq \frac{8}{L^2} \equiv \overline{\varphi}$ w.p.1. Also note that $\varphi^{W,N}(Z_{\ell}, W_k, N_i, N_j) \leq \frac{8}{L^2} \equiv \overline{\varphi}$ w.p.1. Therefore, Hoeffding’s inequality (see Hoefding (1963), or Lemma 2.2.7 in van der Vaart and Wellner (1996)) yields
\[
\Pr\left[ \frac{1}{(L-3)} \sum_{\ell \neq i,j,k} \varphi^{W,N}(Z_{\ell}, W_k, N_i, N_j) \geq 2b_k \right] \leq \exp\left\{ -\frac{2(L-3)b_k^2}{\overline{\varphi}^2} \right\} \quad \text{w.p.1.}
\]

By dominance and Assumption (T3), we obtain
\[
\sum_{i=1}^{L} \sum_{j=1}^{L} \sum_{k=1}^{L} \Pr\left[ \frac{1}{(L-3)} \sum_{\ell \neq i,j,k} \varphi^{W,N}(Z_{\ell}, W_k, N_i, N_j) \geq 2b_k \right] \leq L^3 \cdot \exp\left\{ -\frac{2(L-3)b_k^2}{\overline{\varphi}^2} \right\} \\
= \exp\left\{ 3 \log L - \frac{2(L-3)b_k^2}{\overline{\varphi}^2} \right\} \rightarrow 0 \quad \text{as } L \rightarrow \infty,
\]

Therefore, $\lim_{L \to \infty} \Pr[\sqrt{L} | \beta_{L,(3)}^W | > \delta] = 0 \ \forall \ \delta > 0$. That is, $U_{L,(3)}^{\beta_{W,N}} = o_p(L^{-1/2})$.

### A.6.4 $U_{L,(3)}^{\beta_{W,N}} = o_p(L^{-1/2})$

Take any constant $\overline{\varphi} > 0$. From \[39\] we have $\Pr[\varepsilon_{L} > \overline{\varphi}] = \Pr[\varepsilon_{L} \geq \overline{\varphi}]$. The same arguments leading to \[30\] yield $\Pr[\tau_{L} > \overline{\varphi}] = O(L^{-1})$ and $\Pr[\tau_{L} > \overline{\varphi}] = o_p(L^{-1/2})$. Let $\tau$ and $\mathcal{H}$ be as described in Assumption (T2). We have
\[
\mathbb{E}[Z_{2}^{W,N}(Z_{i}, Z_{j}, Z_{k})] = \mathbb{E}\left[ \mathbb{E}[Z_{2}^{W,N}(Z_{i}, Z_{j}, Z_{k})|N_{i}, N_{j}, N_{k}] \right] \\
= \mathbb{E}\left[ \Delta_{W|N}(W_k, N_i, N_j) \mathbf{1}\{ N_i > N_j \} \mathbf{1}\{ \Delta_{W|N}(W_k, N_i, N_j) \geq -b_k \} \cdot \mathbf{1}\{ -\varepsilon_{L}^{-i,j,k}(W_k, N_i, N_j) \leq \Delta_{W|N}(W_k, N_i, N_j) < 0 \} \right] \\
\leq \Pr[\varepsilon_{L}^{-i,j,k}(W_k, N_i, N_j) > \tau] \\
+ \mathbb{E}\left[ \mathbf{1}\{ -\varepsilon_{L}^{-i,j,k}(W_k, N_i, N_j) \leq \Delta_{W|N}(W_k, N_i, N_j) < 0 \} \cdot \mathbf{1}\{ \varepsilon_{L}^{-i,j,k}(W_k, N_i, N_j) \leq \Delta_{W|N}(W_k, N_i, N_j) \leq \tau \} \right] \\
\leq \Pr[\varepsilon_{L} \geq \tau] + \mathcal{H} \cdot \mathbb{E}\left[ \varepsilon_{L}^{-i,j,k}(W_k, N_i, N_j)^2 \right] \\
\leq \Pr[\tau_{L} > \tau] + \mathcal{H} \cdot \mathbb{E}[\tau_{L}^2] = O(L^{-1}) + O(b_k^2) = o(L^{-1/2}),
\]
(33)
where $E[S^3_{FW}] = O(b_2^L)$ follows from the arguments leading to (29). From here, the last equality follows from Assumption (T3). Combined with Chebyshev’s inequality, this result yields

$$U_{L}^{SFW_{i,k}} = O_p \left( \sqrt{\text{Var} \left[ U_{L}^{SFW_{i,k}} \right]} \right) = O_p \left( \sqrt{E \left[ U_{L}^{SFW_{i,k}} \right]^2} + o(L^{-1}) \right). \quad (34)$$

Let $c$ denote the set of all $L_3$ distinct combinations $(i, j, k)$ of $(1, \ldots, L)$. We have

$$E \left[ U_{L}^{SFW_{i,k}} \right] = \left( \begin{array}{c} L \end{array} \right)^{-2} \sum_{c} \sum_{c} E \left[ H_{S_{2}^{FW_{i,k}}(Z_i, Z_j, Z_k) : h_{S_{2}^{FW_{i,k}}(Z_{i'}, Z_{j'}, Z_{k'})} \right].$$

Note first that, for any pair $(i, j, k)$ and $(i', j', k')$ in $(1, \ldots, L)$, we have

$$E \left[ H_{S_{2}^{FW_{i,k}}(Z_i, Z_j, Z_k) : h_{S_{2}^{FW_{i,k}}(Z_{i'}, Z_{j'}, Z_{k'})} \right] \leq \Pr[|z_{i,k}| \leq \Delta(W_k, N_i, N_j) < 0] \leq \Pr[z_{i,k} > \tau] + \Pi \cdot E[|z_{i,k}|] = O(L^{-1}) + O(b_L) = O(b_L).$$

where $E[|z_{i,k}|] = O(b_L)$ follows from the arguments leading to (29). Since the above bound is valid for any pair $(i, j, k)$ and $(i', j', k')$ in $(1, \ldots, L)$, it immediately implies that

$$E \left[ H_{S_{2}^{FW_{i,k}}(Z_i, Z_j, Z_k) : h_{S_{2}^{FW_{i,k}}(Z_{i'}, Z_{j'}, Z_{k'})} \right] \leq O(b_L). \quad (35)$$

For $d \in \{0, 1, 2, 3\}$, let $c_d$ denote the collection of all pairs $\{(i, j, k) ; (i', j', k')\}$ in $c$ that have exactly $d$ elements in common. It is convenient to express $E \left[ U_{L}^{SFW_{i,k}} \right]$ as

$$E \left[ U_{L}^{SFW_{i,k}} \right] = \left( \begin{array}{c} L \end{array} \right)^{-2} \sum_{d=0}^{\infty} \sum_{c_d} E \left[ H_{S_{2}^{FW_{i,k}}(Z_i, Z_j, Z_k) : h_{S_{2}^{FW_{i,k}}(Z_{i'}, Z_{j'}, Z_{k'})} \right].$$

It is not hard to verify that each $c_d$ has a total of $L_3 \binom{L-3}{d}$ elements. Given this, for $1 \leq d \leq 3$, the bound in (35) will suffice for our purposes. From there we have

$$\left( \begin{array}{c} L \end{array} \right)^{-2} \sum_{d=1}^{3} \sum_{c_d} E \left[ H_{S_{2}^{FW_{i,k}}(Z_i, Z_j, Z_k) : h_{S_{2}^{FW_{i,k}}(Z_{i'}, Z_{j'}, Z_{k'})} \right] \leq O(b_L) \cdot \left( \begin{array}{c} L \end{array} \right)^{-1} \cdot \sum_{d=1}^{3} \binom{L-3}{d} = O \left( \frac{b_L}{L} \right) = o(L^{-1}),$$

and therefore

$$E \left[ U_{L}^{SFW_{i,k}} \right] = \left( \begin{array}{c} L \end{array} \right)^{-2} \sum_{c_0} E \left[ H_{S_{2}^{FW_{i,k}}(Z_i, Z_j, Z_k) : h_{S_{2}^{FW_{i,k}}(Z_{i'}, Z_{j'}, Z_{k'})} \right] + o(L^{-1}). \quad (36)$$

Thus, it all boils down to verifying the properties of $E \left[ H_{S_{2}^{FW_{i,k}}(Z_i, Z_j, Z_k) : h_{S_{2}^{FW_{i,k}}(Z_{i'}, Z_{j'}, Z_{k'})} \right]$ when $(i, j, k)$ and $(i', j', k')$ have zero elements in common. First, let

$$\Delta_{W_{i,k}^{i',j',k'}}(w, n, n') = \Delta_{W_{i,k}^{i',j',k'}}(w, n, n') + \frac{1}{L-3} \sum_{\ell \neq i, j, k' \neq j', k'} \varphi_{W_{i,k}^{i',j',k'}}(w, n, n').$$

Our approach will rely on the rate at which the following probability goes to zero,

$$\Pr \left\{ \left\{ \Delta_{W_{i,k}^{i',j',k'}}(W_k, N_i, N_j) \geq -b_L \right\} \neq \left\{ \Delta_{W_{i,k}^{i',j',k'}}(W_k, N_i, N_j) \geq -b_L \right\} \right\} \quad \text{or} \quad \Pr \left\{ \left\{ \Delta_{W_{i,k}^{i',j',k'}}(W_k, N_i, N_j) \geq -b_L \right\} \neq \left\{ \Delta_{W_{i,k}^{i',j',k'}}(W_k, N_i, N_j) \geq -b_L \right\} \right\}.$$
To this end, note that the same arguments that led to (27) can be used to show that

\[
\sup_{w \in S_N} \left( \max \left\{ \hat{\Delta}_{W,N}^{-i,j,k',l'}(w,n,n') - \hat{\Delta}_{W,N}^{-i,j,k}(w,n,n') \right\} \right) \leq \frac{1}{L} \cdot \mathcal{S}_L, \text{ where } \mathcal{S}_L = o_p(1).
\]

This yields

\[
\left\lfloor \frac{1}{L} \mathcal{F}_L \left( \hat{\Delta}_{W,N}^{-i,j,k}(w,k',N_j) \right) \right\rfloor \leq \frac{1}{L} \mathcal{S}_L, \text{ where } \mathcal{S}_L = o_p(1). \tag{37}
\]

Define \( \nu_L^{-i,j,k',l'}(w,n,n') = \sqrt{\mathcal{L}} \left( \hat{\Delta}_{W,N}^{-i,j,k',l'}(w,n,n') - \mathcal{L}^{-i,j,k}(w,n,n') \right) \) and let \( \Gamma_L(w|n,n') = -\sqrt{\mathcal{L}}(b_L + \mathcal{L}^{-i,j,k}(w,n,n')) \). We have

\[
\left\lfloor \frac{1}{L} \mathcal{F}_L \left( \hat{\Delta}_{W,N}^{-i,j,k',l'}(w,n,n') \right) \right\rfloor \leq \frac{1}{L} \mathcal{S}_L, \text{ where } \mathcal{S}_L = o_p(1). \tag{38}
\]

From the uniform representation result in Equation (26), for any pair of scalars \( t < t' \) we have

\[
\text{Pr} \left\{ t \leq \nu_L^{-i,j,k',l'}(w,k,N_j) \leq t' \big| W_k, N_i, N_j, \mathcal{F}_L, \mathcal{S}_L \right\} = o_p \left( \left| t' - t \right| \right).
\]

The same property is satisfied by \( \text{Pr} \left\{ t \leq \nu_L^{-i,j,k',l'}(w,k',N_j) \leq t' \big| W_k', N_i', N_j' \right\} = o_p \left( \left| t' - t \right| \right) \). From here, (37) and (38), along with \( \text{Pr} \left\{ \mathcal{F}_L > \mathcal{M} \right\} = O(L^{-1}) \) and Assumption (T3) yield

\[
E \left[ \hat{\mathcal{S}}_{F|W}^2(Z_i, Z_j, Z_k) \cdot \hat{\mathcal{S}}_{F|W}^2(Z_i', Z_j', Z_k') \right] = O \left( \frac{b_L^4}{L} \right) + O \left( \frac{b_L^2}{\sqrt{L}} \right) + o(L^{-1}) = o(L^{-1}),
\]

for any \((i,j,k)\) and \((i',j',k')\) with zero elements in common. Consequently,

\[
\left( \frac{L}{3} \right)^2 \sum_{n''} E \left[ \hat{\mathcal{S}}_{F|W}^2(Z_i, Z_j, Z_k) \cdot \hat{\mathcal{S}}_{F|W}^2(Z_i', Z_j', Z_k') \right] = \left( \frac{L}{3} \right)^2 \left( \frac{L - 3}{3} \right) \cdot o(L^{-1}) = o(L^{-1}). \tag{39}
\]

From here, (34) and (36) yield \( \hat{U}_L^{F|W} \cdot \hat{U}_L^{F|W} = o_p(L^{-1/2}) \).
A.6.5 \( U_0^{\hat{W}} = o_p(L^{-1/2}) \)

First we note from (20) that if \( \Delta_{W|N}(W_k, N_i, N_j) \geq 0 \), then we will have \( \hat{\Delta}_{W|N}^{-i,j,k}(W_k, N_i, N_j) < -b_L \) only if

\[
\frac{1}{(L-3)} \sum_{\ell \neq i,j,k} \psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) < -b_L - \xi_\Delta^{-i,j,k}(W_k, N_i, N_j).
\]

It follows that

\[
|\hat{Q}^{W|N}_L(Z_t, Z_k) | < \mathbb{I}\left\{ -b_L - \xi_\Delta^{-i,j,k}(W_k, N_i, N_j) \right\}.
\]

We will decompose this upper bound as follows. Let

\[
\hat{Q}^{W|N}_L(Z_t, Z_k) = \mathbb{I}\left\{ \frac{1}{(L-3)} \sum_{\ell \neq i,j,k} \left( \psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) - E[\psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) \mid W_k, N_i, N_j, \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)] \right) \right\}
\]

where \( \hat{Q}^{W|N}_L(Z_t, Z_k) \) is described in Assumption (T2). (29) implies

\[
|\hat{Q}^{W|N}_L(Z_t, Z_k) | \leq \mathbb{I}\left\{ (1 + 2 \bar{T}) \cdot \bar{T} \geq b_L \right\}
\]

By construction of \( \bar{T} \), this becomes

\[
|\hat{Q}^{W|N}_L(Z_t, Z_k) | \leq \mathbb{I}\left\{ (1 + 2 \bar{T}) \cdot \bar{T} \geq b_L \right\}
\]

where \( \tau \) is described in Assumption (T2). \( (29) \) implies

\[
|\hat{Q}^{W|N}_L(Z_t, Z_k) | \leq \mathbb{I}\left\{ (1 + 2 \bar{T}) \cdot \bar{T} \geq b_L \right\} = o_p(L^{-1/2}).
\]

We will show that \( U_0^{\hat{W}} = o_p(L^{-1/2}) \) by establishing a much stronger (but easy to show) result. For any \( \delta > 0 \) we have \( \Pr\left[ \sqrt{L} |U_0^{\hat{W}}| \geq \delta \right] \leq \Pr\left[ \sum_{i=1}^{L} \sum_{j=1}^{L} \sum_{k=1}^{L} \Pr[\hat{Q}^{W|N}_L(Z_t, Z_j, Z_k) \neq 0] \right] \).

A Bonferroni inequality bounds this probability by \( \sum_{i=1}^{L} \sum_{j=1}^{L} \sum_{k=1}^{L} \Pr[\hat{Q}^{W|N}_L(Z_t, Z_j, Z_k) \neq 0] \).

A Hoeffding inequality in turn bounds this probability by \( 2L^3 E\left[ \exp\left( -2L(b_L + \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)) + E[\psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) \mid W_k, N_i, N_j, \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)] \right)^2 \right] \)

\[
= 2 E\left[ \exp\left( 3 \log L - \frac{2(\sqrt{T} \cdot b_L + \sqrt{T} \cdot \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)) + \sqrt{T} \cdot E[\psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) \mid W_k, N_i, N_j, \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)]}{\sqrt{2L}} \right)^2 \right],
\]

where \( \psi \equiv \sum_{i=1}^{L} \sum_{j=1}^{L} \sum_{k=1}^{L} \).

By (26), (27) and Assumption (T3), independence across observations implies that, conditional on \( \langle W_k, N_i, N_j \rangle \),

\[
\exp\left( 3 \log L - \frac{2(\sqrt{T} \cdot b_L + \sqrt{T} \cdot \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)) + \sqrt{T} \cdot E[\psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) \mid W_k, N_i, N_j, \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)]}{\sqrt{2L}} \right)^2 \quad \rightarrow 0,
\]

\[
\text{Note that } \left| \xi_\Delta^{-i,j,k}(W_k, N_i, N_j) \right| + E[\psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) \mid W_k, N_i, N_j, \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)] < b_L \text{ necessarily implies that } -b_L - \xi_\Delta^{-i,j,k}(W_k, N_i, N_j) - E[\psi_\Delta^{W|N}(Z_t, W_k, N_i, N_j) \mid W_k, N_i, N_j, \xi_\Delta^{-i,j,k}(W_k, N_i, N_j)] < 0, \text{ allowing us to invoke Hoeffding’s inequality.}
\]
and this result holds w.p.1 for any \((W_k, N_i, N_j)\). By Equations (26)–(29), we have
\[
\exp\left(3 \log L - \frac{2(\sqrt{V} \cdot b_L + \sqrt{U} \cdot \xi_{i,j,k}^N(W_k, N_i, N_j))}{(2\pi)^2}\right)
\leq K \exp\left\{\frac{4(1+2\tau)b_L D_L}{(2\pi)^2}\right\}
\]
where \(K \geq \exp\left\{3 \log L - \frac{2b_L^2}{(2\pi)^2}\right\}\) \(\forall L \geq 1\), and \(D_L\) is as described between Equations (27) and (29). By the properties of \(D_L\) described there, there exists a random variable \(\mathcal{Y}\) with finite expectation that dominates \(\exp\left\{\frac{4(1+2\tau)b_L D_L}{(2\pi)^2}\right\}\) \(\forall L \geq 1\). From here, the Dominated Convergence Theorem yields
\[
\mathbb{E}\left[\exp\left(3 \log L - \frac{2(\sqrt{V} \cdot b_L + \sqrt{U} \cdot \xi_{i,j,k}^N(W_k, N_i, N_j))}{(2\pi)^2}\right)\right] \to 0.
\]
Therefore \(\Sigma_{i=1}^L \sum_{j} \sum_{k} \mathbb{P}[\mathcal{Q}_{L}^{F_{W|N}}(Z_i, Z_j, Z_k) \neq 0] \to 0\), and in particular \(U_L^{\mathcal{Q}_{L}^{F_{W|N}}} = o_p(L^{-1/2})\). Finally by (40) this shows that \(U_L^{\mathcal{Q}_{L}^{F_{W|N}}} = o_p(L^{-1/2})\).

A.6.6 \(U_L^{\mathcal{Q}_{L}^{F_{W|N}}} = o_p(L^{-1/2})\)

Let
\[
\mathcal{P}_{L|N}^{F_{W|N}} = \max_{(i,j,k)} \left\{ \sup_{w \in \mathcal{S}_W} \left\| \sqrt{L} \mathcal{F}_{W|N}(w) - F_{W|N}(w) \right\| \right\}.
\]

By Equation (25) we have \(\mathcal{P}_{L|N}^{F_{W|N}} = O_p(1)\). Next note that
\[
|\mathcal{M}^{F_{W|N}}(Z_i, Z_j, Z_k)| \leq \frac{1}{\sqrt{L}} \cdot \mathcal{P}_{L|N}^{F_{W|N}} \cdot \left( \hat{B}_1^{F_{W|N}}(Z_i, Z_j, Z_k) + \hat{B}_2^{F_{W|N}}(Z_i, Z_j, Z_k) + \hat{B}_3^{F_{W|N}}(Z_i, Z_j, Z_k) \right),
\]
where
\[
\hat{B}_1^{F_{W|N}}(Z_i, Z_j, Z_k) = \mathbb{I}\left\{ \Delta_{W|N}(W_k, N_i, N_j) \geq -b_L \right\} \cdot \mathbb{I}\left\{ \Delta_{W|N}(W_k, N_i, N_j) \leq \Delta_{W|N}(W_k, N_i, N_j) \right\},
\]
\[
\hat{B}_2^{F_{W|N}}(Z_i, Z_j, Z_k) = \mathbb{I}\left\{ -\xi_{i,j,k}^N(W_k, N_i, N_j) \leq \Delta_{W|N}(W_k, N_i, N_j) \right\},
\]
\[
\hat{B}_3^{F_{W|N}}(Z_i, Z_j, Z_k) = \mathbb{I}\left\{ \Delta_{W|N}(W_k, N_i, N_j) \leq \Delta_{W|N}(W_k, N_i, N_j) \right\}.
\]

Accordingly we have
\[
|U_L^{\mathcal{Q}_{L}^{F_{W|N}}}| \leq \frac{1}{\sqrt{L}} \cdot \mathcal{P}_{L|N}^{F_{W|N}} \cdot (U_L^{\hat{B}_1^{F_{W|N}} + U_L^{\hat{B}_2^{F_{W|N}}} + U_L^{\hat{B}_3^{F_{W|N}}}}).
\]

The arguments and results used in Section A.6.4 are sufficient to show that \(U_L^{\hat{B}_1^{F_{W|N}}} = O_p(b_L)\). The exponential bounds derived in Sections A.6.3 and A.6.5 are sufficient to show that \(U_L^{\hat{B}_2^{F_{W|N}}} \) and \(U_L^{\hat{B}_3^{F_{W|N}}} \) disappear in probability much faster. We have \(U_L^{\hat{B}_1^{F_{W|N}}} + U_L^{\hat{B}_2^{F_{W|N}}} + U_L^{\hat{B}_3^{F_{W|N}}} = O_p(b_L)\) and consequently by (41), we obtain \(U_L^{\mathcal{Q}_{L}^{F_{W|N}}} = O_p(L^{-1/2} \cdot b_L) = o_p(L^{-1/2})\).
A.6.7 Final step to prove part (i) of Theorem 8

Equation (32) and the results in Sections A.6.3-A.6.6 yield $U_{L(3)}^{Fw|N} = U_{L(3)}^{Fw|N} + U_{L(3)}^{Fw|N} + o_p(L^{-1/2})$. Let $i \neq j \neq k \neq \ell$. The Hoeffding decomposition of $U_{L(3)}^{Fw|N}$ (see Lemma 5.1.5.A in Serfling (1980)) yields

$$U_{L(3)}^{Fw|N} = \mu_{Fw|N} + \frac{3}{L} \sum_{i=1}^{L} \left( E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k) | Z_i] - \mu_{Fw|N} \right) + O_p(L^{-1}),$$

and therefore $\sqrt{L} \cdot U_{L(3)}^{Fw|N} = \sqrt{L} \cdot \mu_{Fw|N} + O_p(1)$. By inspection, if (3) holds w.p.1 we have $\mu_{Fw|N} = 0$ and $E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k) | Z_i] = E[\mathcal{T}^{Fw|N}(Z_i, Z_j, Z_k) | Z_i]$, and in this case our Hoeffding decomposition implies $\sqrt{L} \cdot U_{L(3)}^{Fw|N} = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} E[\mathcal{T}^{Fw|N}(Z_i, Z_j, Z_k) | Z_i] + o_p(L^{-1/2}).$

Using Equation (25) we obtain $U_{L(3)}^{Fw|N} = -U_{L(4)}^{Fw|N} + o_p(L^{-1/2})$, with $U_{L(4)}^{Fw|N}$ as defined by (20) and (23). Notice that $E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k, Z_l) | Z_i, Z_j, Z_k, Z_l] = 0$ and therefore $E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k, Z_l) | Z_i, Z_j, Z_k, Z_l] = 0$. Furthermore, by inspection we have $E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k, Z_l) | Z_i] = E[g^{Fw|N}(Z_i, Z_j, Z_k, Z_l) | Z_i]$. The Hoeffding decomposition of $U_{L(4)}^{Fw|N}$ yields

$$U_{L(4)}^{Fw|N} = E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k, Z_l)] + \frac{4}{L} \sum_{i=1}^{L} \left( E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k, Z_l) | Z_i] - E[\mathcal{H}^{Fw|N}(Z_i, Z_j, Z_k, Z_l)] \right) + O_p(L^{-1})$$

$$= \frac{1}{L} \sum_{i=1}^{L} E[g^{Fw|N}(Z_i, Z_j, Z_k, Z_l) | Z_i] + O_p(L^{-1}).$$

Part (i) of Theorem 8 follows from these results.

Proving part (ii) of Theorem 8

The proof follows steps analogous to those in Sections A.6.2-A.6.7 very closely. As it was the case there, the main building block is a linear representation result analogous to (25) and (26). Let $\varphi^{Fw|N}$ and $\varphi^{Fw|N}$ be as defined in (21). If $\nabla_1 \Omega(F_{w|N}(w|n'), n, n')$ is bounded $n, n' \in S_N^2$ and all $w \in S_w$, we can now show that

$$\Omega(\tilde{F}^{-i,j,k}_{w|N}(w|n'), n, n') = \Omega(F_{w|N}(w|n'), n, n') + \frac{1}{L - 3} \sum_{\ell \neq i, j, k} \varphi^{Fw|N}(Z_{\ell}, w, n') \cdot \nabla_1 \Omega(F_{w|N}(w|n'), n, n') + \tilde{\zeta}_L^{-i,j,k}(w, n, n'),$$

where $\sup_{w \in S_w} \left| \tilde{\zeta}_L^{-i,j,k}(w, n, n') \right| = O_p(L^{-1}).$

and

$$\tilde{\Phi}^{-i,j,k}_{w|N}(w, n, n') = \Phi_{w|N}(w, n, n') + \frac{1}{L - 3} \sum_{\ell \neq i, j, k} \varphi^{Fw|N}(Z_{\ell}, w, n, n') + \tilde{\zeta}_L^{-i,j,k}(w, n, n'),$$

where $\sup_{w \in S_w} \left| \tilde{\zeta}_L^{-i,j,k}(w, n, n') \right| = O_p(L^{-1}).$

From here, using Assumption (T2') in place of (T2) is sufficient to prove part (ii) of the theorem by taking the same type of steps described in Sections A.6.2-A.6.7.
A.7 Asymptotic properties of $U_{L(4)}^{\mathcal{F}w|Y}$ and $U_{L(4)}^{\mathcal{F}Pw|Y}$

A.7.1 Assumptions

We begin by formalizing the basic distribution assumptions, all of which are compatible with the conditions in Section 3.2.

Assumption C1

We observe an iid sample $(W_i, N_i, X_i)_{i=1}^n \equiv (Z_i)_{i=1}^n$ produced by a data generating process that satisfies the assumptions in Section 3.2. The support $S_Z$ of $Z$ is compact. For each $n \in S_N$: (i) the support of $X|N=n$ is equal to $S_X$ (the unconditional support of $X$), and (ii) the conditional distribution of $X$ given $N=n$ is absolutely continuous with respect to Lebesgue measure. The distribution $F_{W|Y}(w|n, x)$ is continuous in $w$ for almost every $(n, x) \in S_N \times S_X$. Let $p_Y(x, n) = f_{X|N}(x|n) \cdot P_N(n)$, then $\inf_{(x, n) \in S_X \times S_N} p_Y(x, n) = p_Y > 0$.

Take any $(w, n, n', x)$ such that $(n, n') \in S_N^2$ and $x \in S_X$. We will now define

$$\varphi_{F_{W|Y}}(Z, w, n, x) = \frac{\mathbb{I} \{ W \leq u \} - F_{W|Y}(w|n, x)}{p_Y(x, n)} \mathbb{I} \{ N = n \} \frac{1}{h_n} \frac{1}{K} \left( \frac{X - x}{h_n} \right),$$

$$\varphi_{\Delta_{W|Y}}(Z, w, n, n', x) = \varphi_{F_{W|Y}}(Z, w, n, x) - \varphi_{F_{W|Y}}(Z, w, n', x),$$

$$\varphi_{\Phi_{W|Y}}(Z, w, n, n', x) = \varphi_{F_{W|Y}}(Z, w, n', x) \cdot \nabla_1 \Omega \left( F_{W|Y}(w|n', x), n, n' \right) - \varphi_{F_{W|Y}}(Z, w, n, x)$$

These objects are analogous to those described in Equation (21). Under our full set of assumptions (see below), an $M$th-order Taylor approximation yields $E[\varphi_{F_{W|Y}}(Y, w, n, x)] = O(\tilde{h}_n^M)$, $E[\varphi_{\Delta_{W|Y}}(Y, w, n, n', x)] = O(\tilde{K}_n^M)$ and $E[\varphi_{\Phi_{W|Y}}(Y, w, n, n', x)] = O(\tilde{K}_n^M)$. Now let

$$\xi_{-i,j,k,\ell}(w, n, n', x) = \Delta_{W|Y}(w, n, n', x) - \Delta_{W|Y}(w, n, n', x) - \frac{1}{L-4} \sum_{m \neq i,j,k,\ell} \varphi_{\Delta_{W|Y}}(Z_m, w, n, n', x),$$

$$\tilde{\xi}_{-i,j,k,\ell}(w, n, n', x) = \Phi_{W|Y}(w, n, n', x) - \Phi_{W|Y}(w, n, n', x) - \frac{1}{L-4} \sum_{m \neq i,j,k,\ell} \varphi_{\Phi_{W|Y}}(Z_m, w, n, n', x)$$

$\xi_{-i,j,k,\ell}(\cdot)$ and $\tilde{\xi}_{-i,j,k,\ell}(\cdot)$ have a similar interpretation as the object defined in (22). Its asymptotic properties correspond to those of the remainder term in a first order Taylor expansion of $\widehat{\Delta}_{W|Y}(\cdot)$ around $\Delta_{W|Y}(\cdot)$. The same interpretation holds for $\tilde{\xi}_{-i,j,k,\ell}(\cdot)$ and a first order approximation of $\widehat{\Phi}_{W|Y}(\cdot)$ around $\Phi_{W|Y}(\cdot)$.

Assumption C2

The type of regularity conditions described in Assumption (T2) hold true for $\varphi_{\Delta_{W|Y}}(\cdot)$ and $\xi_{-i,j,k,\ell}(\cdot)$. In addition, there exists a $\eta_L$ such that, for any $i \neq j \neq k \neq \ell$,

$$\sup_{w \in S_W} \left| \xi_{-i,j,k,\ell}(w, n, n', x) \right| \leq \eta_L, \quad \text{and} \quad \lim_{L \to \infty} E[\eta_L^4] < \infty.$$ 

Assumption C2'

The type of regularity conditions described in Assumption (T2') hold true for $\varphi_{\Phi_{W|Y}}(\cdot)$ and $\tilde{\xi}_{-i,j,k,\ell}(\cdot)$. In
addition, there exists a \( \eta_L \) such that, for any \( i \neq j \neq k \neq \ell \),

\[
\sup_{w \in S_W, (n,n') \in S_N^2, x \in S_X} \left| \tilde{c}_{L,i,j,k,\ell}(w,n,n',x) \right| \leq \eta_L, \quad \text{and} \quad \lim_{L \to \infty} E[f_{\eta_L}^4] < \infty.
\]

The existence-of-moments conditions in Assumptions (C2) and (C2') were not necessary in Assumptions (T2) and (T2') because, as we noted there, the objects defined in Equation (22) were bounded w.p.1. This is no longer the case now, due to our use of bias-reducing kernels in the estimation of \( \tilde{\Delta}_{W|Y}^{i,j,k,\ell}(\cdot) \) and \( \tilde{\Phi}_{W|Y}^{i,j,k,\ell}(\cdot) \). Hence, the need to impose these conditions. As we will discuss below, they will enable us to obtain a result analogous to Equation (29).

**Assumption C3**

The density \( f_X(x) \) is \( M \) times differentiable with bounded derivatives almost everywhere (i.e., except for a set of Lebesgue-measure zero) in \( S_X \). Furthermore, for any given \( n \in S_N \), the conditional density \( f_{X|N}(x|n) \) is also \( M \) times differentiable with respect to \( x \) with bounded derivatives almost everywhere \( S_X \). For any \( w \in S_W, (x,x') \in S_X^2 \) and \( (n,n') \in S_N^2 \), let

\[
\phi(w,n,n',x) = E_w \left[ \mathbb{1}\{u \leq W\} \mathbb{1}\{\Delta_{W|Y}(W,n,n',x) \geq 0\} \right]
\]

\[
\Psi(n,n',x) = E_w \left[ F_{W|Y}(W,n',x) \mathbb{1}\{\Delta_{W|Y}(W,n,n',x) \geq 0\} \right]
\]

\[
\Upsilon(w,n,x,x') = E_{N|X} \left[ \mathbb{1}\{N > n\} \mathbb{1}\{\Delta_{W|Y}(w,N,n,x) \geq 0\} \right] \quad \text{for } X = x'
\]

For almost every \( w \in S_W \) (i.e., except for a subset of \( S_W \) with Lebesgue measure zero) and for every \( (n,n') \in S_N^2 \): (i) \( F_{W|Y}(w,n,x) \), \( \phi(w,n,n',x) \) and \( \Psi(n,n',x) \) are \( M \)-times differentiable with respect to \( x \) with bounded derivatives almost everywhere in \( S_X \). (ii) \( \Upsilon(w,n,x,x') \) is \( M \) times differentiable with respect to both \( x \) and \( x' \) with bounded derivatives almost everywhere in \( S_X^2 \).

**Assumption C3’**

The density \( f_X(x) \) is \( M \) times differentiable with bounded derivatives almost everywhere in \( S_X \). Furthermore, for any given \( n \in S_N \), the conditional density \( f_{X|N}(x|n) \) is also \( M \) times differentiable with respect to \( x \) with bounded derivatives almost everywhere \( S_X \). For any \( w \in S_W, (x,x') \in S_X^2 \) and \( (n,n') \in S_N^2 \), let

\[
\tilde{\phi}(w,n,n',x) = E_w \left[ \mathbb{1}\{u \leq W\} \mathbb{1}\{\Phi_{W|Y}(W,n,n',x) \geq 0\} \right]
\]

\[
\tilde{\Psi}(n,n',x) = E_w \left[ \mathbb{1}\{\Phi_{W|Y}(W,n',x),n'\} \mathbb{1}\{\Phi_{W|Y}(W,n,n',x) \geq 0\} \right]
\]

\[
\tilde{\Upsilon}(w,n,x,x') = E_{N|X} \left[ \mathbb{1}\{N > n\} \nabla_1 \mathbb{1}\{\Phi_{W|Y}(w,n,x),N,n\} \mathbb{1}\{\Phi_{W|Y}(w,N,n,x) \geq 0\} \right] \quad \text{for } X = x'
\]

For almost every \( w \in S_W \) and for every \( (n,n') \in S_N \): (i) \( F_{W|Y}(w,n,x) \), \( \tilde{\phi}(w,n,n',x) \) and \( \tilde{\Psi}(n,n',x) \) are \( M \)-times differentiable with respect to \( x \) with bounded derivatives almost everywhere in \( S_X \). (ii) \( \tilde{\Upsilon}(w,n,x,x') \) is \( M \) times differentiable with respect to both \( x \) and \( x' \) with bounded derivatives almost everywhere in \( S_X^2 \).
Assumption C4

Let $M$ be the constant described in Assumptions (C3) and (C3'). Both $K(\cdot)$ and $\tilde{K}(\cdot)$ 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 $\tilde{K}$), and they are both bounded by some constant $\overline{K}$. The sequences $b_L$, $h_L$ and $\tilde{h}_L$ satisfy

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

We use bandwidth sequences of the form $b_L = c_L L^{-\alpha_1}$, $h_L = c_L L^{-\alpha_2}$ and $\tilde{h}_L = c_L L^{-\alpha_3}$, where $c_L > 0$ and $\alpha_j > 0$. The conditions in Assumption (C4) can be satisfied if and only if $M \geq 16$. If we assume $M = 16$, the conditions can be satisfied by setting $\alpha_1 = 0.344$, $\alpha_2 = 0.073$ and $\alpha_3 = 0.073$. Assuming higher values of $M$ can lead to different values of $\alpha_1$, $\alpha_2$ and $\alpha_3$ consistent with Assumption (C4).

**Theorem 9** Take $i \neq j \neq k \neq \ell \neq m$.

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

$$\sqrt{L} \cdot U_{L(4)}^{F_{W|Y}} = \sqrt{L} \cdot \gamma_{F_{W|Y}} + o_p(1).$$

Let

$$\delta_{F_{W|Y}}(Z_i) = E \left[ \Phi(W_i, N_i, N_j, X_i) - \Psi(N_i, N_j, X_i) \right] f_X(X_i)$$

$$- E \left[ \Phi(W_i, W_j, \Gamma(W_j|N_j, X_i)) \right] f_X(X_i) \frac{f_X(X_i)^2}{f_X|N_i|}.$$

Suppose the conditional version of $\Phi$ is satisfied with probability one. Then,

$$\sqrt{L} \cdot U_{L(4)}^{F_{W|Y}} = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta_{F_{W|Y}}(Z_i) + o_p(1),$$

and $\frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta_{F_{W|Y}}(Z_i) \rightarrow N(0, \Sigma_{F_{W|Y}})$, with $\Sigma_{F_{W|Y}} > 0$ if

$$E \left[ \mathbb{I} \{N_i > N_j\} \mathbb{I} \{\Delta_{W|Y}(W_k, N_i, N_j, X_i) = 0\} \right] > 0.$$

That is, $\Sigma_{F_{W|Y}} > 0$ if the conditional version of $\Phi$ is binding with nonzero probability.

(ii) If Assumptions (C1), (C2'), (C3') and (C4) hold, and if $\nabla_1 \Omega(F_{W|Y}(w|x), n, n')$ is bounded for all $n, n' \in S_{N}$ and almost every $w \in S_{W}$ and $x \in S_{X}$ (otherwise we can modify our test and restrict ourselves to a region of $S_{N} \times S_{X}$ where this is assumed to hold), then

$$\sqrt{L} \cdot U_{L(4)}^{\Omega_{W|Y}} = \sqrt{L} \cdot \gamma_{\Omega_{W|Y}} + O_p(1).$$

Let

$$\delta_{\Omega_{W|Y}}(Z_i) = E \left[ \Phi(W_i, N_i, N_j, X_i) - \Phi(W_i, N_i, N_j, X_i) \right] f_X(X_i)$$

$$+ E \left[ \Phi(W_i, W_j, \Gamma(W_j|N_j, X_i)) \right] f_X(X_i) \frac{f_X(X_i)^2}{f_X|N_i|}.$$
Suppose the conditional version of (4) is satisfied with probability one. Then,
\[ \sqrt{L} \cdot U_{L(i)}^{\xi^{w|y}} = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta^{w|y}(Z_i) + o_p(1), \]
and \[ \frac{1}{\sqrt{L}} \sum_{i=1}^{L} \delta^{w|y}(Z_i) \xrightarrow{d} \mathcal{N}(0, \Sigma_{w|y}), \quad \text{with} \quad \Sigma_{w|y} > 0 \]
if
\[ E \left[ \mathbb{I}\{N_i > N_j\} \mathbb{I}\{\Phi_{w|y}(W_k, N_i, N_j, X_{\ell}) = 0\} \right] > 0. \]
That is, \( \Sigma_{w|y} > 0 \) if the conditional version of (4) is binding with nonzero probability.

### A.7.2 Proof of Theorem 9

We will prove part (i). As it was the case in Theorem 8 the proof of part (ii) follows parallel steps and we will omit it for reasons of space. The proof begins by obtaining results analogous to (25) and (26) in Section A.6.1. Take any \( w \in S_W, n \in S_N \) and \( x \in S_X \). Under our assumptions, using a linear approximation and \( M^{th} \)-order expansion for \( E[\varphi^{w|y}_L(Z_m, w, n, x)] \), we can now show (see e.g., Lemma 3 in [Collomb and Hardle (1986)] and Theorem 1’ in Lewbel (1997) and the references cited there) that
\[ \hat{F}_{w|y}^{-i,j,k,l}(w|n, x) = F_{w|y}(w|n, x) + \frac{1}{L-4} \sum_{m \neq i,j,k,l}^{L} \varphi_{w|y}^{m}(Z_m, w, n, x) + \zeta_{i,j,k,l}^{-i,j,k,l}(w, n, x), \]
where \( \sup_{w \in S_W, n \in S_N, x \in S_X} \left| \xi_{i,j,k,l}^{-i,j,k,l}(w, n, x) \right| = O_p(L^{\delta-1} \cdot \tilde{h}_L^{-1}) \quad \forall \delta > 0, \tag{43} \]
and
\[ \hat{\Delta}_{w|y}^{-i,j,k,l}(w, n, n', x) = \Delta_{w|y}(w, n, n', x) + \frac{1}{L-4} \sum_{m \neq i,j,k,l}^{L} \varphi_{w|y}^{m}(Z_m, w, n, x) + \xi_{i,j,k,l}^{-i,j,k,l}(w, n, n', x), \]
where \( \sup_{(m, n) \in S_W, (n, x) \in S_N, x \in S_X} \left| \xi_{i,j,k,l}^{-i,j,k,l}(w, n, n', x) \right| = O_p(L^{\delta-1} \cdot \tilde{h}_L^{-1}) \quad \forall \delta > 0. \tag{44} \]
The linear representations in (43) and (44) will serve the same purpose as the results in (25) and (26). Note that under our assumptions we have
\[ E[\varphi_{w|y}^{w|y}(Z_m, w, n, n', x)] = O(\tilde{h}_L^M). \tag{45} \]
The next step in the proof is to show that, under our assumptions,
\[ U_{L(i)}^{\hat{F}_{w|y}} = U_{L(i)}^{\hat{F}_{w|y}} + U_{L(i)}^{\hat{F}_{w|y}} + o_p(L^{-1/2}), \tag{46} \]
where
\[ \hat{V}_{L(i)}^{F_{w|y}}(Z_i, Z_j, Z_k, Z_{\ell}) = -\left( \hat{F}_{w|y}^{-i,j,k,l}(W_k|N_j, X_{\ell}) - F_{w|y}(W_k|N_j, X_{\ell}) \right) \cdot \mathbb{I}\{N_i > N_j\} \mathbb{I}\{\hat{\Delta}_{w|y}(W_k, N_i, N_j, X_{\ell}) \geq 0\}. \]
To establish (46), we can decompose $U_{L(n)}^{\tilde{s}_{W}/y}$ in an analogous manner to Equation (32). Namely,

$$U_{L(n)}^{\tilde{s}_{W}/y} = U_{L(n)}^{S_{W}/y} + U_{L(n)}^{\tilde{V}_{L}/y} + U_{L(n)}^{\tilde{S}_{W}/y} + U_{L(n)}^{\tilde{Q}_{L}/y} + U_{L(n)}^{\tilde{M}_{W}/y}.$$  

The expressions for $U_{L(n)}^{S_{W}/y}$, $U_{L(n)}^{\tilde{V}_{L}/y}$, $U_{L(n)}^{\tilde{S}_{W}/y}$ and $U_{L(n)}^{\tilde{M}_{W}/y}$ are analogous to the ones described in Equation (31). Our assumptions enable us to follow steps parallel to those in Sections A.6.3-A.6.6 to show that each one of these four terms is of order $O_p(L^{-1/2})$. As it was the case there, we rely heavily on the linear representations of these four terms is of order $O_p(L^{-1/2})$. As it was the case there, we rely heavily on the linear representations

From (45), it follows that $E\left[\tilde{s}_{W}/y(Z_i, Z_j, N_k, N_l, X_m)\right] = O_h(L)$ for any $i \neq j \neq k \neq l \neq m$. Using this, plus now also on the relative convergence properties of the bandwidths employed. In particular, we note that under our bandwidth assumptions, there exists $\tilde{b} > 0$ such that

$$L^{\frac{1}{2}} \cdot \tilde{h}_l \cdot \tilde{h}_l \xrightarrow{P} \infty, \quad \frac{L^{\frac{1}{2}} \cdot \tilde{h}_l \cdot \tilde{h}_l}{\tilde{h}_l \cdot \tilde{h}_l} \xrightarrow{P} 0,$$

(47)

As stated in (44), for this particular $\tilde{b}$ we have

$$\sup_{w \in S_{W}} \left| \frac{1}{n} \tilde{h}_l \cdot \tilde{h}_l \right| = O_p(L^{-1} \cdot \tilde{h}_l^{-1}), \quad \text{and} \quad \sup_{w \in S_{W}} \left| \frac{1}{n} \tilde{h}_l \cdot \tilde{h}_l \right| = O_p(L^{-1} \cdot \tilde{h}_l^{-1}).$$

From (45), it follows that $E\left[\tilde{s}_{W}/y(Z_i, Z_j, N_k, N_l, X_m)\right] = O_h(L)$ for any $i \neq j \neq k \neq l \neq m$. Using this, plus the bandwith condition $L^{1/2} \tilde{h}_l \cdot \tilde{h}_l \xrightarrow{P} \infty$, a Hoeffding inequality argument parallel to that in Section A.6.3 can be used to show that $U_{L(n)}^{S_{W}/y} = o_p(L^{-1/2})$. We study the properties of $U_{L(n)}^{S_{W}/y}$ by taking the same approach as in Section A.6.4. We start by noting that, under Assumption (C2), and (44), we can express

$$\tilde{Z}_l = D_l/(L^{1-\tilde{b}} \cdot \tilde{h}_l),$$

where $b \cdot L^{1-\tilde{b}} \cdot \tilde{h}_l)^2 \Pr[D_l \geq b \cdot L^{1-\tilde{b}} \cdot \tilde{h}_l] = O(1)$, where $\tilde{b}$ is as described in (47).

From here we obtain $\Pr[\tilde{Z}_l \geq b \cdot L^{1-\tilde{b}} \cdot \tilde{h}_l] = O((b \cdot L^{1-\tilde{b}} \cdot \tilde{h}_l)^{-2}) = o(L^{-1})$. As in Equation (33), this yields

$$E\left[\tilde{Z}_l \cdot \tilde{Z}_l \right] = O\left(\frac{b^2}{\tilde{h}_l} + O\left(\frac{b^2}{\tilde{h}_l}\right)\right),$$

where the last equality follows from our bandwidth conditions. Using this result, we have

$$E\left[\tilde{Z}_l \cdot \tilde{Z}_l \right] = O\left(\frac{b^2}{\tilde{h}_l} + O\left(\frac{b^2}{\tilde{h}_l}\right)\right),$$

and $c$ is the collection of all $\binom{4}{2}$ distinct combinations $(i, j, k, \ell)$ of $(1, \ldots, L)$, and let $c_0$ denote the collection of all pairs $\{(i, j, k, \ell); (i', j', k', \ell')\}$ in $c$ that have zero elements in common. Steps like those we took between Equations (33) and (36) now yield

$$E\left[\tilde{Z}_l \cdot \tilde{Z}_l \right] = O\left(\frac{b^2}{\tilde{h}_l} + O\left(\frac{b^2}{\tilde{h}_l}\right)\right),$$

By taking steps analogous to those leading to (39), we now obtain

$$\binom{L}{2}^{-2} \sum_{c_0} E\left[\tilde{Z}_l \cdot \tilde{Z}_l \right] = O(b^4) + O(h^4) + O\left(\frac{b^2}{L^{\frac{1}{2}} \cdot \tilde{h}_l \cdot \tilde{h}_l}\right) = O\left(\frac{1}{L}\right).$$

51
And from here, we obtain $U_{L(W)}^{E} = o_p(L^{-1/2})$. We study the term $U_{L(W)}^{E}$ like we did in Section A.6.5. First, we note that our assumptions yield a result analogous to [27].

\[
\sup_{w \in S, (n,n') \in S_N} \left| E \left[ \varphi^A_{L}(Z_m,w,n,n',x) \right| \xi^{-j,k,l}(w,n,n',x) \right| = o_p(L^{-1/2}) \quad \forall \delta > 0.
\]

Using the previously stated result (see above) $P_r(\xi \geq b) = O((b L^{1-\delta} - \theta))$, a Hoeffding inequality argument similar to the one used in Section A.6.5 plus the bandwidth condition $L^{1/2} \theta - \theta \to \infty$ yield $U_{L(W)}^{E} = o_p(L^{-1/2})$. Finally, our linear representation result in (43) and the same type of arguments as in Section A.6.6 yield $U_{L(W)}^{E} = o_p(L^{1/2})$. Combined, these steps establish (46). Namely, $U_{L(W)}^{E} = U_{L(W)}^{E} + U_{L(W)}^{E} + o_p(L^{-1/2})$. The rest of the proof consists of characterizing the Hoeffding decompositions of $U_{L(W)}^{E}$ and $U_{L(W)}^{E}$. We begin with $U_{L(W)}^{E}$. For $i \neq j \neq k \neq \ell$, we have

\[
E \left[ \mathbb{I} \{ W_i \leq W_k \} \mathbb{I} \{ N_i > N_j \} \mathbb{I} \{ \Delta_{W|Y}(W_k, N_i, N_j, X_i) \geq 0 \} \frac{1}{h} \left( \frac{X_i - X_k}{h} \right) \right] = 0.
\]

The last equality follows from an $M^{th}$-order approximation and our assumptions. Similarly, we have

\[
E \left[ F_{W|Y}(W_k N_j, X_i) \mathbb{I} \{ N_i > N_j \} \mathbb{I} \{ \Delta_{W|Y}(W_k, N_i, N_j, X_i) \geq 0 \} \frac{1}{h} \left( \frac{X_i - X_k}{h} \right) \right] \Rightarrow E \left[ \mathbb{I} \{ N_i > N_j \} (\mathbb{I} \{ W_i \leq W_k \} \mathbb{I} \{ N_i > N_j \} \mathbb{I} \{ \Delta_{W|Y}(W_k, N_i, N_j, X_i) \geq 0 \}) \right] f_X(X_i) + O(h^{M}).
\]

and therefore

\[
E \left[ S_{L}^{E}(Z_i, Z_j, Z_k, Z_l) \right] = E \left[ \mathbb{I} \{ N_i > N_j \} (\mathbb{I} \{ W_i \leq W_k \} \mathbb{I} \{ N_i > N_j \} \mathbb{I} \{ \Delta_{W|Y}(W_k, N_i, N_j, X_i) \geq 0 \}) \right] f_X(X_i) + O(h^{M}).
\]

Analogous steps show that under our assumptions,

\[
\begin{align*}
E \left[ S_{L}^{E}(Z_i, Z_k, Z_l) \right] & = E \left[ \Delta_{W|Y}(W_k, N_i, N_k, X_i) \mathbb{I} \{ N_i > N_k \} f_{X|Y}(X_i | N_k) \mathbb{I} \{ \Delta_{W|Y}(W_k, N_i, N_k, X_i) \geq 0 \} \right] + O(h^{M}), \\
E \left[ S_{L}^{E}(Z_i, Z_j, Z_k, Z_l) \right] & = E \left[ \Delta_{W|Y}(W_i, N_j, N_k, X_i) \mathbb{I} \{ N_j > N_k \} f_{X|Y}(X_i | N_j) \mathbb{I} \{ \Delta_{W|Y}(W_i, N_j, N_k, X_i) \geq 0 \} \right] + O(h^{M}), \\
E \left[ S_{L}^{E}(Z_i, Z_k, Z_l, Z_j) \right] & = E \left[ \Delta_{W|Y}(W_i, N_j, N_k, X_i) \mathbb{I} \{ N_j > N_k \} f_{X|Y}(X_i | N_j) \mathbb{I} \{ \Delta_{W|Y}(W_i, N_j, N_k, X_i) \geq 0 \} \right] + O(h^{M}).
\end{align*}
\]
It follows from (48) and (49) that \( \frac{1}{L} \sum_{i=1}^{L} \left( E \left[ H_{(4)}^{S_L}(Z_i, Z_j, Z_k, Z_t) | Z_i \right] - E \left[ H_{(4)}^{S_L}(Z_i, Z_j, Z_k, Z_t) \right] \right) = O_p(L^{-1/2}) + O(h_L^M) = O_p(L^{-1/2}) \), where the last equality follows from the condition \( L^{1/2}h_L^M \rightarrow 0 \). Since \( E \left[ H_{(4)}^{S_L}(Z_i, Z_j, Z_k, Z_t) \right] = \gamma_{FW,Y} \), the Hoeffding decomposition of \( U_{L(4)}^{FW,Y} \) yields

\[
U_{L(4)}^{FW,Y} = \gamma_{FW,Y} + O \left( \frac{1}{h_L} \right) + O_p \left( \frac{1}{L} \right) = \gamma_{FW,Y} + O(h_L^M) + O_p(L^{-1/2}).
\]

By Theorem 3, if the conditional version of Equation (50) holds w.p.1 we have \( \gamma_{FW,Y} = 0 \). In this case, the Hoeffding decomposition of \( U_{L(4)}^{FW,Y} \) simplifies to

\[
U_{L(4)}^{FW,Y} = \frac{1}{L} \sum_{i=1}^{L} E \left[ S_L^{FW}(Z_i, Z_j, Z_k, Z_t) | Z_i \right] + O(h_L^M) = \frac{1}{L} \sum_{i=1}^{L} E \left[ S_L^{FW}(Z_i, Z_j, Z_k, Z_t) | Z_i \right] + o_p(L^{-1/2}).
\]

We move on to \( U_{L(5)}^{FW,Y} \). We can immediately extend our notation in Equations (17)-(20) to define \( f^{th} \) order symmetric U-statistics. For any function \( g \) with five arguments \( (x_1, x_2, x_3, x_4, x_5) \) we will let \( H_{(5)}^{g}(x_1, x_2, x_3, x_4, x_5) = \frac{1}{5!} \sum_{c_{5}} g(x_{r_1}, x_{r_2}, x_{r_3}, x_{r_4}, x_{r_5}) \), where \( \sum_{c_{5}} \) denotes the sum over the 5! permutations \( (r_1, r_2, r_3, r_4, r_5) \) of \( (1, 2, 3, 4, 5) \). We define

\[
U_{L(5)}^{g} = \left( \frac{L}{5} \right)^{-1} \sum_{i<j<k<\ell<\ell_m} H_{(5)}^{g}(Z_i, Z_j, Z_k, Z_{\ell}, Z_{\ell_m}).
\]

Let

\[
g_L^{FW,Y}(Z_i, Z_j, Z_k, Z_{\ell}, Z_{\ell_m}) = \varphi_{FW,Y}^{L}(Z_i, W_{\ell}, N_k, X_m) \cdot \| \{ N_j \neq N_k \} \cdot \| \{ \Delta_{W,Y}(W_{\ell}, N_j, N_k, X_m) \geq 0 \} \left( \frac{X_j - X_m}{h_L} \right)
\]

From Equation (43), we have

\[
U_{L(5)}^{FW,Y} = -U_{L(5)}^{g} + O_p \left( \frac{1}{L^{1/3}} \frac{1}{h_L} \right) = -U_{L(5)}^{g} + o_p(L^{-1/2}).
\]
And therefore

\[ E\left[ g_{W|Y}^{Z_i, Z_j, Z_k, Z_t, Z_m}\right]_{Z_j, Z_k, Z_t, Z_m} = O(\tilde{h}_M) \times \frac{1}{p_y(x_m, N_k)} \{ N_j > N_k \} \{ \Delta_{W|Y}(W_t, N_j, N_k, X_m) \geq 0 \} \frac{1}{h_L} K \left( \frac{X_j - X_m}{h_L} \right) = O(\tilde{h}_L) \]  

(53)

where the last equality follows from our assumptions, which imply

\[ \frac{1}{p_y(x_m, N_k)} \{ N_j > N_k \} \{ \Delta_{W|Y}(W_t, N_j, N_k, X_m) \geq 0 \} \frac{1}{h_L} K \left( \frac{X_j - X_m}{h_L} \right) \leq \frac{1}{h_L} \frac{\mathcal{K}}{p_y} \]

The result in (53) also implies \( E\left[ U_{L(5)}^{W|Y} \right] = O(\tilde{h}_M) \). Next note that under our assumptions,

\[ E \left[ \{ N_j > N_k \} \{ \Delta_{W|Y}(W_t, N_j, N_k, X_m) \geq 0 \} \frac{1}{h_L} K \left( \frac{X_j - X_m}{h_L} \right) \right]_{Z_k, Z_t, Z_m} = E \left[ \{ N_j > N_k \} \{ \Delta_{W|Y}(W_t, N_j, N_k, X_m) \geq 0 \} \frac{1}{h_L} K \left( \frac{X_j - X_m}{h_L} \right) \right]_{Z_k, Z_t, Z_m} = E \left[ (W_t, N_k, X_m, X_j) \frac{1}{h_L} K \left( \frac{X_j - X_m}{h_L} \right) \right]_{Z_k, Z_t, Z_m} = \int E \left[ (W_t, N_k, X_m, x) \frac{1}{h_L} K \left( \frac{x - X_m}{h_L} \right) \right] f_X(x) dx \]

\[ = \int E \left[ (W_t, N_k, X_m, x) \frac{1}{h_L} K \left( \frac{x - X_m}{h_L} \right) \right] f_X(x) dx + O(h_L^M) \]

Using this, we have

\[ E\left[ g_{W|Y}^{Z_i, Z_j, Z_k, Z_t, Z_m}\right]_{Z_i} = \int E \left[ \{ W_i \leq W_t \} - F_{W|Y}(W_t|N_k, X_m) \right]_{Z_i} = \int E \left[ \{ W_i \leq W_t \} - F_{W|Y}(W_t|N_k, X_m) \right]_{Z_i} + O(\tilde{h}_L^M) \]

Under our assumptions,

\[ E \left[ \{ W_i \leq W_t \} - F_{W|Y}(W_t|N_k, X_m) \right]_{Z_i} \geq \int E \left[ \{ W_i \leq W_t \} - F_{W|Y}(W_t|N_k, X_i) \right]_{Z_i} = \int E \left[ \{ W_i \leq W_t \} - F_{W|Y}(W_t|N_i, X_i) \right]_{Z_i} f_X(x) dx + O(\tilde{h}_L^M) \]

And we obtain

\[ E\left[ g_{W|Y}^{Z_i, Z_j, Z_k, Z_t, Z_m}\right]_{Z_i} = E \left[ \{ W_i \leq W_t \} - F_{W|Y}(W_t|N_k, X_i) \right]_{Z_i} + O(\tilde{h}_L^M) \]

(54)

\[ = E \left[ \{ W_i \leq W_t \} - F_{W|Y}(W_t|N_i, X_i) \right]_{Z_i} + O(\tilde{h}_L^M) \]

\[ = \int E \left[ \frac{P_y(X_i) f_X(X_i)^2}{p_y(x_i, N_i)} \right]_{Z_i} + O(\tilde{h}_L^M) \]
From [53] we have $E \left[ \mathcal{H}^{g_{L}}_{U_L} (Z_i, Z_j, Z_k, Z_t, Z_m) | Z_i \right] = \frac{4}{M} E \left[ g_{L}^{F_{W|Y}} (Z_i, Z_j, Z_k, Z_t, Z_m) | Z_i \right] + O \left( \frac{h_L}{M} \right)$. Using (54), the Hoeffding decomposition of $U_{L(5)}^{g_{W|Y}}$ becomes

\[
U_{L(5)}^{g_{W|Y}} = E \left[ U_{L(5)}^{g_{W|Y}} \right] + 5 \cdot \frac{1}{L} \sum_{i=1}^{L} \left( E \left[ \mathcal{H}^{g_{L}}_{U_L} (Z_i, Z_j, Z_k, Z_t, Z_m) | Z_i \right] - E \left[ U_{L(5)}^{g_{W|Y}} \right] \right) + O \left( \frac{1}{h_L h_L L} \right)
\]

\[
= \frac{1}{L} \sum_{i=1}^{L} E \left[ g_{L}^{F_{W|Y}} (Z_i, Z_j, Z_k, Z_t, Z_m) | Z_i \right] + O \left( \frac{h_L^M}{h_L} \right) + O \left( \frac{1}{h_L h_L L} \right)
\]

\[
= \frac{1}{L} \sum_{i=1}^{L} E \left[ \left( \mathbf{1} \{ W_i \leq W_i \} - F_{W|Y}(W_i | X_i) \right) \mathbf{Y}(W_i | N_i, X_i, X_i) | Z_i \right] \frac{f_X(X_i)^2}{J_X(N_i | X_i)} + O(\tilde{h}_L^M) + O \left( \frac{h_L^M}{h_L} \right) + O \left( \frac{1}{h_L h_L L} \right)
\]

\[
= \frac{1}{L} \sum_{i=1}^{L} E \left[ \left( \mathbf{1} \{ W_i \leq W_i \} - F_{W|Y}(W_i | X_i) \right) \mathbf{Y}(W_i | N_i, X_i, X_i) | Z_i \right] \frac{f_X(X_i)^2}{J_X(N_i | X_i)} + o_p(L^{-1/2}).
\]

(55)

Combining (46) with (52), (50–51) and (55) proves part (i) of Theorem 9. Part (ii) is shown following analogous steps, and we omit the details for reasons of space. \qed

References


55


