Nonparametric Tests for Conditional Affiliation in Auctions and Other Models *

Andrés Aradillas-López
Pennsylvania State University

November 21, 2016

Abstract

The property of affiliation of agents' private signals introduced in Milgrom and Weber (1982) is key for a number of essential results in structural auction models, ranging from equilibrium existence in first-price auctions to the revenue dominance of second-price auctions. Affiliation has testable implications for observable decision variables (e.g, bids, participation decisions, etc.) which can be used as the basis for econometric tests. To this end, having at hand an econometric procedure that can test whether affiliation holds conditional on observable characteristics is crucial since rational agents construct their beliefs conditional on all the information available to them. Therefore the relevant question is whether affiliation holds conditional on common knowledge characteristics. Presently there do not exist tests for conditional affiliation which do not rely on either parametric assumptions or an artificial discretization/categorization of observed heterogeneity. In this paper we propose nonparametric tests for affiliation that allow us to condition on observables in a very general way. Our procedures rely on one-sided statistics which are estimated nonparametrically. Under regularity and smoothness conditions our tests have an asymptotically pivotal property, they are easy to compute and perform well in Monte Carlo experiments. We apply our methodology to Timber auction data in the southwern region of the United States (Region 8 of the USFS) and we show that, while an unconditional test fails to reject affiliation of bids, once we control for observable auction characteristics affiliation is rejected.
Keywords: *Affiliation, auctions, nonparametric test, conditional moments.*

1 Introduction

Affiliation is a stronger notion than positive correlation (although it includes independence as a special case). Introduced by Milgrom and Weber (1982) in the context of auction models, it captures

*518 Kern Bldg. Department of Economics Penn State University. University Park, PA 16802. email: aaradill@psu.edu.
The author is grateful to the Human Capital Foundation (www.hcfoundation.ru), and especially Andrey P. Vavilov, for support of the Department of Economics, the Center for the Study of Auctions, Procurements, and Competition Policy (CAPCP, http://capcp.psu.edu/), and the Center for Research in International Financial and Energy Security (CRIFES, http://crifes.psu.edu/) at Penn State University.

the intuitive notion that, as a particular bidder's value rises, higher values for other bidders' estimates become relatively more likely. Since its introduction, affiliation has been a crucial assumption in many important results in auction theory. For instance, it implies the existence of a symmetric pure-strategy equilibrium in first-price auctions and is also leads to a well-known revenue ranking result whereby second-price auctions revenue-dominate first-price auctions and are in turn dominated by English auctions. Subsequent work (e.g, de Castro (2010)) has shown that equilibrium existence in first-price auctions and the revenue dominance of second-price auctions do not hold under other definitions of positive dependence. Affiliation has played a key role in the structural analysis of auction data. A partial list of papers that either maintain affiliation or use it as a basis of specification tests between competing auction models include Paarsch (1992), Li, Perrigne, and Vuong (2000), Hendricks, Pinkse, and Porter (2003), Haile, Hong, and Shum (2003), Li, Perrigne, and Vuong (2002), Levin and Smith (1994), Pinkse and Tan (2005), Li, Paarsch, and Hubbard (2007).

The results obtained from affiliation in auction models arise through agents' optimal (i.e, expected-payoff maximizing) strategies. Rational agents (bidders) compute their expected payoffs conditional on all their relevant information. This includes their own private signal as well as everything that is common knowledge among agents. Therefore, what is relevant is whether agents' signals (value estimates or informational variables) are affiliated *conditional on common knowledge variables*. For this reason it is important to have an econometric procedure that can test for affiliation conditional on observable covariates. This can also allow us to control for possible asymmetries across agents, which is important since many of the results that follow from affiliation presuppose symmetry[1]. Existing tests for affiliation include Roosen and Hennessy (2004), de Castro and Paarsch (2010), Jun, Pinkse, and Wan (2010) and Li and Zhang (2010). While these tests vary in their scope and nature, they are all limited in their ability to condition on observable characteristics. Roosen and Hennessy (2004), de Castro and Paarsch (2010) and Jun, Pinkse, and Wan (2010) would require a discretization of observable covariates into categories or cells, which could lead to inconsistent results when the true support of these variables is continuous, as is often the case in existing data sets. On the other hand the test in Li and Zhang (2010) is parametric in nature –with the concomitant risk of misspecification– and is explicitly limited to entry models of auctions. In addition, only Jun, Pinkse, and Wan (2010) is designed to test for affiliation between continuous variables (e.g, bids). Our goal here is to develop testing procedures with the following features:

[1] As de Castro and Paarsch (2010, p. 2074) point out, "very little can be said about the revenue-generating properties of the various auction formats and pricing rules under asymmetries".

(i) Tests that are nonparametric in nature and consistent.

(ii) Tests that allow for rich observable covariates, including those with continuous support, without having to discretize or categorize them.

(iii) Can be used to test affiliation between discrete and/or continuous variables.

Our tests will be based on one-sided statistics similar to those used in Jun, Pinkse, and Wan (2010). However, while they relied on unconditional probabilities, our procedure requires conditioning, which poses nontrivial technical challenges. We describe conditions under which our test statistics possess pivotal asymptotic properties which make them easy to implement. These conditions involve smoothness and a special type of regularity of the relevant conditional functionals, as well as precise restrictions on the features of tuning parameters (bandwidths) involved. The paper proceeds as follows. Section 2 describes our econometric tests and their asymptotic properties. Section 3 explores the finite sample features of our tests through several Monte Carlo experiments. In Section 4 we apply our procedures to test for affiliation in USFS data. Section 5 concludes. Proofs are contained in the appendix.

2 A test for conditional affiliation

We begin by presenting the characterization of affiliation given in Jun, Pinkse, and Wan (2010), which is equivalent to the definition in Milgrom and Weber (1982, Lemma 1). Let $U \in \mathbb{R}^L$ be a vector of random variables with joint distribution F_U. Following convention, let $a \vee b$ and $a \wedge b$ denote, respectively, the element-wise maximum and minimum of a and b. Now take $\delta \in \mathbb{R}_+^L$. For a given $u \in \mathbb{R}^L$ let $G(u, \delta) = P\left(u - \delta \leq U \leq u + \delta\right)$ and denote

$$\tau(u_1, u_2, \delta) = G(u_1, \delta) \cdot G(u_2, \delta) - G(u_1 \vee u_2, \delta) \cdot G(u_1 \wedge u_2, \delta).$$

The elements of U are affiliated if and only if $\tau(u_1, u_2, \delta) \leq 0 \; \forall \, u_1, u_2$ and δ. Based on this characterization, Jun, Pinkse, and Wan (2010) describe a nonparametric test for affiliation.

Our goal here is to test whether affiliation holds conditional on a vector of observable characteristics X. Having the ability to do this is important since many existing data sets (particularly in applications involving auctions) include a rich collection of observable covariates X which may capture common knowledge characteristics of the environment as well as asymmetries across agents. Our specific goal is to introduce tests that satisfy three key features:

(i) Are nonparametric in nature and consistent.

(ii) Allow for the vector of conditioning covariates X to include a large collection continuous and/or discrete random variables with a rich support.

(iii) Can be applied to either discrete or continuous variables in U.

To date there is no test in the literature that has all these features. Some tests (e.g, de Castro and Paarsch (2010) and Roosen and Hennessy (2004)) are nonparametric but can only be used when U consists exclusively of discrete variables. The test in Jun, Pinkse, and Wan (2010) can handle continuous elements in U but it would require a discretization of the support of X into categories or cells, which can lead to inconsistent results if X includes continuous covariates. The affiliation test in Li and Zhang (2010) can condition on continuous covariates but their procedure is parametric, with the concomitant risk of misspecification, and it focuses specifically on participation decisions in auctions. The tests proposed here are, to the best of our knowledge, the first ones to have all three features described above.

We begin by extending the definition of affiliation given above to a conditional setting. Let $U \in \mathbb{R}^L$ be a vector of random variables with joint distribution given by F_U and let X be a vector of conditioning variables with distribution F_X. We will group $Z \equiv (U, X)$ and we will maintain that we observe an iid sample $(Z_i)_{i=1}^n$. We will let Supp$(U) \equiv \mathcal{U}$, Supp$(X) \equiv \mathcal{X}$ and Supp$(Z) \equiv \mathcal{Z}$ denote the supports of U, X and Z respectively. Let $\delta \equiv (\delta_1, \ldots, \delta_L) \in \mathbb{R}_+^L$. For a given $u \in \mathbb{R}^L$, $x \in \mathcal{X}$ and $\delta \in \mathbb{R}_+^L$ we will denote

$$G(u, \delta|x) = P\left(u - \delta \leq U \leq u + \delta \middle| X = x\right).$$

We will refer to the cubes $[u - \delta,\, u + \delta]$ in $\mathbb{R}^L$ as our "contact sets". Our procedure will test whether the conditions of affiliation hold over this class of sets. We will let $f_X(\cdot)$ denote the density of X and

$$\mu(u, \delta|x) = G(u, \delta|x) \cdot f_X(x)$$

Our test will be based on this type of density-weighted probabilities.

2.1 A one-sided functional as the basis for our test

Our test will be based on an extension of the type of one-sided statistic used in Jun, Pinkse, and Wan (2010) to a case where we condition on observables. Let $u_1 \wedge u_2 = \min\{u_1, u_2\}$ and $u_1 \vee u_2 =$

$\max\{u_1, u_2\}$ (both element-wise). For any pair u_1, u_2 in $\mathbb{R}^L$ we will denote

$$\begin{aligned}\tau(u_1, u_2, \delta|x) &= \mu(u_1, \delta|x) \cdot \mu(u_2, \delta|x) - \mu(u_1 \vee u_2, \delta|x) \cdot \mu(u_1 \wedge u_2, \delta|x) \\ &= \Big[G(u_1, \delta|x) \cdot G(u_2, \delta|x) - G(u_1 \vee u_2, \delta|x) \cdot G(u_1 \wedge u_2, \delta|x)\Big] \cdot f_X(x)^2.\end{aligned}$$

Note that, by construction, for any u_1, u_2, δ and a.e $x \in \mathcal{X}$,

$$G(u_1, \delta|x) \cdot G(u_2, \delta|x) - G(u_1 \vee u_2, \delta|x) \cdot G(u_1 \wedge u_2, \delta|x) \leq 0 \iff \tau(u_1, u_2, \delta|x) \leq 0.$$

By using density-weighted functionals we can also avoid trimming as well as assumptions about the density of X being bounded away from zero. If the elements of U are affiliated conditional on X then we must have $\tau(u_1, u_2, \delta|x) \leq 0$ for all u_1, u_2, δ and a.e $x \in \mathcal{X}$. From now on we will refer to this restriction as the (conditional) *affiliation inequality*. No trimming allows us to test whether the conditions of affiliation hold over the entire support $\mathcal{X}$, leading to a consistent test given our class of contact sets.

Let Q be a pre-specified probability measure with support $Supp(Q) \equiv \mathcal{S}_Q \subset \mathbb{R}^L_+$. We will use Q to generate our contact sets. Let

$$R^Q(u_1, u_2|x) = \int \max\left\{\tau(u_1, u_2, \delta|x), 0\right\} dQ(\delta)$$

If the variables in U are affiliated conditional on X then we must have $R^Q(u_1, u_2|x) = 0$ for a.e x and $\forall\, u_1, u_2$. We will design a procedure to test whether $R^Q(u_1, u_2|x) = 0$ for $F_U \times F_U \times F_X-$a.e $(u_1, u_2, x) \in \mathcal{U} \times \mathcal{U} \times \mathcal{X}$. For simplicity we will denote

$$H = F_U \times F_U \times F_X \times Q, \quad \text{and} \quad \mathcal{S} = \mathcal{U} \times \mathcal{U} \times \mathcal{X} \times \mathcal{S}_Q.$$

Our procedure will test whether $R^Q(u_1, u_2|x) = 0$ $H-$a.s over $\mathcal{S}$. We refer to $\mathcal{S}$ as our *testing range*. Note that it contains the supports of U and X.

Let $\omega(u_1, u_2, x)$ denote a pre-specified, bounded and positive weighting function and suppose for simplicity that it satisfies the symmetry condition $\omega(u_1, u_2, x) = \omega(u_2, u_1, x)$. Define

$$\mathcal{T}^Q = \int\int\int R^Q(u_1, u_2|x) \cdot \omega(u_1, u_2, x) dF_U(u_1) dF_U(u_2) dF_X(x) \tag{1}$$

By construction:

(i) $\mathcal{T}^Q \geq 0$.

(ii) $\mathcal{T}^Q = 0$ if and only if $\tau(u_1, u_2, \delta|x) \leq 0$ $H-$a.s over $\mathcal{S}$.

2.1.1 Strict conditional affiliation

We will say that the variables in U are *strictly* conditionally affiliated if the affiliation inequality holds *strictly* almost surely. That is, if

$$\tau(u_1, u_2, \delta|x) < 0 \quad H-\text{a.s over } \mathcal{S}.$$

The distinction between affiliation and strict affiliation will play an important role in our results

2.2 An estimator for $\mathcal{T}^Q$

We assume that we observe an iid sample[2] $(U_i, X_i)_{i=1}^n \equiv (Z_i)_{i=1}^n$. We assume that X can be split as $X = \left(X^c, X^d\right)$, where X^c have absolutely continuous distribution with respect to Lebesgue measure and X^d have a discrete distribution. We will denote the dimension of X^c by q. We will employ kernel-based nonparametric estimators. $K : \mathbb{R}^q \to \mathbb{R}$ will denote our kernel function. For a given $x \equiv \left(x^c, x^d\right)$ and $h > 0$ we will define[3]

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

Let $h_n \to 0$ be a nonnegative bandwidth sequence. For a given $x \equiv \left(x^c, x^d\right)$, and given u, u_1, u_2 and δ, our estimators are of the form

$$\begin{aligned}
\widehat{\mu}(u, \delta|x) &= \frac{1}{n}\sum_{i=1}^n \mathbb{1}\left\{u - \delta \leq U_i \leq u + \delta\right\} \cdot \mathcal{H}\left(X_i - x; h_n\right), \\
\widehat{\tau}(u_1, u_2, \delta|x) &= \widehat{\mu}(u_1, \delta|x) \cdot \widehat{\mu}(u_2, \delta|x) - \widehat{\mu}(u_1 \vee u_2, \delta|x) \cdot \widehat{\mu}(u_1 \wedge u_2, \delta|x), \\
\widehat{R}^Q(u_1, u_2|x) &= \int \widehat{\tau}(u_1, u_2, \delta|x) \cdot \mathbb{1}\left\{\widehat{\tau}(u_1, u_2, \delta|x) \geq -b_n\right\} dQ(\delta),
\end{aligned}$$

[2]in Appendix A.3 we generalize this to a setting where $\{(U_i, X_i)\ , 1 \leq i \leq n\}$ is a row-wise iid triangular array.

[3]We can also smooth over the discrete covariates X^d instead of using indicator functions. We focus on estimators of this type for simplicity.

where $b_n \to 0$ is a nonnegative bandwidth sequence whose convergence rates are described below. We estimate $\mathcal{T}^Q$ as

$$\widehat{\mathcal{T}}^Q = \frac{1}{(n)_3} \sum_{i \neq j \neq k} \widehat{R}^Q(U_j, U_k | X_i) \cdot \omega(U_j, U_k, X_i), \tag{2}$$

where $(n)_k \equiv n \cdot (n-1) \cdots (n-k+1)$. The use of b_n will allow us to deal with the kink at zero of $\max\{z, 0\}$ and lead to asymptotically pivotal properties. In this sense, it plays the same role as the sequence β_n in Jun, Pinkse, and Wan (2010) except that in our case it needs to satisfy a more precise set of restrictions involving its rate of convergence due to the nonparametric conditional nature of our problem. The intuition behind our use of the bandwidth sequence b_n is the following. First we note that any econometric test must allow for the possibility that $\tau(U_1, U_2, \delta|X)$ has a point mass at zero. For example, if the elements in U are independent conditional on X then $\tau(U_1, U_2, \delta|X) = 0$ w.p.1. Now, a point mass at zero complicates the asymptotic properties of

$$\left| \mathbb{1}\left\{ \widehat{\tau}(U_1, U_2, \delta|X) \geq 0 \right\} - \mathbb{1}\left\{ \tau(U_1, U_2, \delta|X) \geq 0 \right\} \right|$$

On the other hand if we use the sequence b_n we have

$$\begin{aligned} &\left| \mathbb{1}\left\{ \widehat{\tau}(U_1, U_2, \delta|X) \geq -b_n \right\} - \mathbb{1}\left\{ \tau(U_1, U_2, \delta|X) \geq 0 \right\} \right| \\ \leq &\mathbb{1}\left\{ \left| \widehat{\tau}(U_1, U_2, \delta|X) - \tau(U_1, U_2, \delta|X) \right| \geq b_n \right\} + \mathbb{1}\left\{ -2b_n \leq \tau(U_1, U_2, \delta|X) < 0 \right\}. \end{aligned}$$

The asymptotic features of these two terms are more tractable. We will introduce smoothness and manageability (empirical process) assumptions that will ensure an exponential probability bound for $\mathbb{1}\left\{ \left| \widehat{\tau}(U_1, U_2, \delta|X) - \tau(U_1, U_2, \delta|X) \right| \geq b_n \right\}$, and we will introduce a regularity condition that will help us deal with $\mathbb{1}\left\{ -2b_n \leq \tau(U_1, U_2, \delta|X) < 0 \right\}$. We describe our assumptions next.

Assumption 1. ***(Smoothness)***
$f_X(x) \leq \overline{f} < \infty$ for all $x \in \mathcal{X}$. As before, express any $x \in \mathcal{X}$ generically as $x \equiv \left(x^c, x^d\right)$ with x^c corresponding to the continuously distributed elements in X. For every u, δ and a.e $x \in \mathcal{X}$, both $f_X(x)$ and $G(u, \delta|x)$ are M times differentiable w.r.t x^c with bounded derivatives.

Assumption 2. ***(A regularity condition)*** *There exists constants $\overline{C} > 0$ and $\overline{b} > 0$ such that for $F_U \times F_X \times F_X$−a.e $(u, x_1, x_2) \in \mathcal{U} \times \mathcal{X} \times \mathcal{X}$ and Q−a.e $\delta \in \mathcal{S}_Q$,*

$$Pr\left(-b \leq \tau(U, u, \delta|x_1) < 0 \middle| X = x_2 \right) \leq \overline{C} \cdot b \quad \forall\, 0 < b \leq \overline{b}$$

Note that Assumption 2 allows for $\tau(U_1, U_2, \delta|X)$ to have a point mass at zero. It only assumes, in essence, that $\tau(U_1, U_2, \delta|X)$ has a *finite density* in a neighborhood of the form $[-\overline{b}, 0)$.

Assumption 3. ***(Manageability)*** *For some $\overline{b} > 0$, the classes of functions*

$$\mathscr{F}^1_{\delta,b} = \left\{ f(u_1, u_2, x) = \mathbb{1}\left\{-b \le \tau(u_1, u_2, \delta|x) < 0\right\} \text{ for some } b \in (0, \overline{b}] \text{ and } \delta \in \mathcal{S}_Q \right\},$$
$$\mathscr{F}^2_{\delta,b} = \left\{ f(u_1, u_2, x) = \mathbb{1}\left\{\tau(u_1, u_2, \delta|x) \ge -b\right\} \text{ for some } b \in (0, \overline{b}] \text{ and } \delta \in \mathcal{S}_Q \right\},$$

are Euclidean (see Pakes and Pollard (1989, Definition 2.7)) for envelope 1. *And the class of functions*

$$\mathscr{G}_{u,\delta} = \{g(x) = G(u, \delta|x) \text{ for some } u \in \mathcal{U},\ \delta \in \mathcal{S}_Q\}$$

is Euclidean for envelope 1.

Assumption 4. ***(Kernels and bandwidths)*** *Let M be as described in Assumption 1. We use a bias-reducing kernel K of order M with bounded support. The kernel is a function of bounded variation, symmetric around zero and satisfies $\sup_{v \in \mathbb{R}^q} \left|K(v)\right| \le \overline{K} < \infty$. The bandwidth sequences b_n and h_n are such that, for a small enough $\epsilon_1 > 0$,*

$$n^{1/2-\epsilon_1} \cdot h_n^q \cdot b_n \longrightarrow \infty, \quad n^{1/2+\epsilon_1} \cdot b_n^2 \longrightarrow 0, \quad n^{1/2+\epsilon_1} \cdot h_n^M \longrightarrow 0.$$

Consider bandwidths of the type $h_n \propto n^{-\alpha_h}$ and $b_n \propto n^{-\alpha_b}$. Let $\overline{\epsilon} > 0$ be an arbitrarily small, but strictly positive constant and let $\alpha_h = \frac{1}{2M} + \overline{\epsilon}$ and $\alpha_b = \frac{1}{4} + \overline{\epsilon}$. The conditions in Assumption 4 will be satisfied if

$$M \ge \left\lceil \frac{2 \cdot q}{1 - 4 \cdot \overline{\epsilon}(2+q)} \right\rceil.$$

For example, suppose $q = 4$. Then we need $M \ge 9$. Recall that M is the number of derivatives assumed to exist in Assumption 1 and it also corresponds to the order of the kernel employed.

We can outline the role played by our assumptions (all details are in the appendix). The type of smoothness conditions in Assumption 1 are common in nonparametric problems. Combined with Assumption 4 they ensure that

$$\sup_{\substack{(u,x) \in \mathcal{Z} \\ \delta \in \mathcal{S}_Q}} \left| E\left[\widehat{\mu}(u, \delta|x)\right] - \mu(u, \delta|x) \right| = O\left(h_n^M\right) = O\left(n^{-1/2-\epsilon}\right),$$

for an $\epsilon > 0$. From here, the properties (bounded variation) of our kernel function and known results from empirical process theory yield a useful probability bound,

$$Pr\left(\underset{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}{sup}\left|\widehat{\tau}(u_1,u_2,\delta|x)-\tau(u_1,u_2,\delta|x)\right|\geq b_n\right)$$
$$\leq D_1\exp\left\{-\left(\sqrt{n}h_n^q\left(D_2\cdot b_n-D_3\cdot h_n^M\right)\right)^2\right\}.$$

Assumption 3 combined with the regularity condition in Assumption 2 help ensure that

$$\underset{\delta\in\mathcal{S}_Q}{sup}\frac{1}{(n)_3}\sum_{i\neq j\neq k}\left|\tau(U_j,U_k,\delta|X_i)\right|\cdot\mathbb{1}\left\{-2b_n\leq\tau(U_j,U_k,\delta|X_i)<0\right\}\cdot\omega(U_j,U_k,X_i)$$
$$=O_p\left(b_n^2\right)=O_p\left(n^{-1/2-\epsilon}\right)$$

for some $\epsilon > 0$. These results combined are then used to show that

$$\widehat{\mathcal{T}}^Q=\frac{1}{(n)_3}\sum_{i\neq j\neq k}\int\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\tau(U_j,U_k,\delta|X_i)\geq 0\right\}dQ(\delta)\cdot\omega(U_j,U_k,X_i)+\varphi_n,$$
$$\text{where}\quad\left|\varphi_n\right|=O_p\left(n^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0$$

Adding zero, the above expression can be written as

$$\widehat{\mathcal{T}}^Q-\mathcal{T}^Q=\frac{1}{(n)_3}\sum_{i\neq j\neq k}\left(R^Q(U_j,U_k|X_i)\cdot\omega(U_j,U_k,X_i)-\mathcal{T}^Q\right)$$
$$+\frac{1}{(n)_3}\sum_{i\neq j\neq k}\int\Big(\widehat{\tau}(U_j,U_k,\delta|X_i)-\tau(U_j,U_k,\delta|X_i)\Big)\cdot\mathbb{1}\left\{\tau(U_j,U_k,\delta|X_i)\geq 0\right\}dQ(\delta)\cdot\omega(U_j,U_k,X_i)+\varphi_n.$$

The second term can be represented as a U-statistic whose asymptotic properties are characterized using the manageability conditions in Assumption 3 and known results (maximal inequalities) from U-process theory. We describe our main asymptotic result next.

2.2.1 Asymptotic properties of $\widehat{\mathcal{T}}^Q$

Our first main result will be to show that, under our assumptions $\widehat{\mathcal{T}}^Q$ has a linear representation of the form

$$\widehat{\mathcal{T}}^Q=\mathcal{T}^Q+\frac{1}{n}\sum_{i=1}^n\psi_n^Q(Z_i)+O_p\left(n^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0,$$

where $E\big[\psi_n^Q(Z_i)\big] = 0$. Let us describe the structure of the "influence function" ψ_n^Q. Take any triple of observations $(i \neq j \neq k)$ in $1, \ldots, n$ and denote

$$\widetilde{S}^a(U_j, U_k, X_i, \delta) = \max\left\{\tau(U_j, U_k, \delta | X_i), 0\right\} \cdot \omega(U_j, U_k, X_i),$$

Let $E\big[\widetilde{S}^a(U_j, U_k, X_i, \delta)\big] \equiv \mathcal{T}(\delta)$ and note that $\mathcal{T}^Q = \int \mathcal{T}(\delta) dQ(\delta)$. Next let

$$\mathbb{1}\left\{u - \delta \leq U_\ell \leq u + \delta\right\} \equiv \mathbb{I}(U_\ell, u, \delta).$$

Fix (u_1, u_2, x) and δ and define

$$\begin{aligned}
&\psi_\tau(u_1, u_2, x, Z_\ell, \delta; h_n) = \\
&\mu(u_2, \delta | x) \cdot \Big(\mathbb{I}(U_\ell, u_1, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_1, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big) \\
&+ \mu(u_1, \delta | x) \cdot \Big(\mathbb{I}(U_\ell, u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big) \\
&- \mu(u_1 \vee u_2, \delta | x) \cdot \Big(\mathbb{I}(U_\ell, u_1 \wedge u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_1 \wedge u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big) \\
&- \mu(u_1 \wedge u_2, \delta | x) \cdot \Big(\mathbb{I}(U_\ell, u_1 \vee u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_1 \vee u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big).
\end{aligned}$$

Note that $E_Z\left[\psi_\tau(u_1, u_2, x, Z, \delta; h_n)\right] = 0$ by construction. Now let

$$\widetilde{S}^b(U_j, U_k, X_i, Z_\ell, \delta; h_n) = \psi_\tau(U_j, U_k, X_i, Z_\ell, \delta; h_n) \cdot \mathbb{1}\left\{\tau(U_j, U_k, \delta | X_i) \geq 0\right\} \cdot \omega(U_j, U_k, X_i).$$

And define

$$\begin{aligned}
\psi^a\left(Z_i, \delta\right) &= \Big(E\big[\widetilde{S}^a(U_j, U_k, X_i, \delta)\big|X_i\big] - \mathcal{T}(\delta)\Big) + 2 \cdot \Big(E\big[\widetilde{S}^a(U_i, U_j, X_k, \delta)\big|U_i\big] - \mathcal{T}(\delta)\Big), \\
\psi^b\left(Z_i, \delta; h_n\right) &= E\left[\widetilde{S}^b(U_j, U_k, X_\ell, Z_i, \delta; h_n)\big|Z_i\right], \\
\psi\left(Z_i, \delta; h_n\right) &= \psi^a\left(Z_i, \delta\right) + \psi^b\left(Z_i, \delta; h_n\right).
\end{aligned}$$

Note that $E\left[\psi\left(Z_i, \delta; h_n\right)\right] = 0$ for any δ. We have

$$\psi_n^Q(Z_i) \equiv \int \psi(Z_i, \delta; h_n) dQ(\delta).$$

We state our main result next.

Theorem 1. Under Assumptions 1-4,

$$\widehat{\mathcal{T}}^{Q}=\mathcal{T}^{Q}+\frac{1}{n}\sum_{i=1}^{n}\psi_{n}^{Q}(Z_i)+O_p\left(n^{-1/2-\epsilon}\right)\quad \text{for some } \epsilon>0.$$

Proof: In Appendix A.1. □

Note that if affiliation is satisfied we have $\psi^{a}\left(Z_i,\delta\right)=0$ and therefore $\psi_n^{Q}(Z_i)\equiv\int\psi^{b}(Z_i,\delta;h_n)dQ(\delta)$.

2.3 Constructing a test

We will test the null hypothesis of conditional affiliation using the results in Theorem 1. Note that $\psi_n^{Q}(Z_i)$ satisfies two key properties,

1.– $E\left[\psi_n^{Q}(Z_i)\right]=0$.

2.– $\psi_n^{Q}(Z_i)=0$ a.s if *strict* conditional affiliation holds.

Denote $\sigma^2_{Q,n}=Var\left(\psi_n^{Q}(Z_i)\right)$. The degeneracy property described above implies that $\sigma^2_{Q,n}=0$ if strict affiliation holds. Let $\kappa_n\to 0$ be a nonnegative sequence satisfying $\kappa_n\cdot n^{\epsilon}\to\infty$ for any $\epsilon>0$. Now let

$$t_n=\frac{\sqrt{n}\cdot\widehat{\mathcal{T}}^{Q}}{\max\left\{\kappa_n,\sigma_{Q,n}\right\}}.$$

By Theorem 1, we have

(i) $t_n\xrightarrow{p}\infty$ if the elements of U violate conditional affiliation.

(ii) $t_n=O_p\left(\frac{1}{n^{\epsilon}\cdot\kappa_n}\right)=o_p(1)$ if the elements of U are strictly conditionally affiliated.

(iii) $t_n\xrightarrow{d}\mathcal{N}(0,1)$ if the elements of U are conditionally affiliated, but not strictly.

Let α denote the target significance level for our test and let $z_{1-\alpha}$ denote the $(1-\alpha)^{th}$ quantile of the $\mathcal{N}(0,1)$ distribution. Consider the rejection rule

“Reject H_0 if $t_n\geq z_{1-\alpha}$”.

Based on (i)-(iii) this rejection rule satisfies the following features over our testing range $\mathcal{S}$,

$$\lim_{n\to\infty}Pr\left(H_0 \text{ is rejected by } t_n \text{ when it is false}\right)=1, \tag{3A}$$

and

$$\lim_{n\to\infty} Pr\left(H_0 \text{ is falsely rejected by } t_n\right) = \begin{cases} 0 & \text{if strict conditional affiliation holds,} \\ \alpha & \text{if conditional affiliation holds, but not strictly.} \end{cases}$$

Therefore,

$$\lim_{n\to\infty} Pr\left(H_0 \text{ is falsely rejected by } t_n\right) \leq \alpha \tag{3B}$$

2.3.1 Our test statistic

$\sigma^2_{Q,n}$ is not known and therefore t_n is infeasible. However $\sigma^2_{Q,n}$ can be estimated using a sample analog of[4] $\psi^Q_n(Z_i)$. We describe how to estimate $\sigma^2_{Q,n}$ in Appendix A.2. We replace t_n with

$$\widehat{t}_n = \frac{\sqrt{n}\cdot\widehat{\mathcal{T}}^Q}{\max\left\{\kappa_n, \widehat{\sigma}_{Q,n}\right\}}. \tag{4}$$

Our proposed test rejects H_0 (conditional affiliation) if $\widehat{t}_n \geq z_{1-\alpha}$. Under the assumptions leading to Theorem 1 we have $\left|\widehat{\sigma}^2_{Q,n} - \sigma^2_{Q,n}\right| \xrightarrow{p} 0$. Therefore a test based on (4) inherits the properties described in (3A)-(3B). Note that our test is *consistent* in the following sense: If the affiliation inequalities are violated for some contact set generated by H over some range of values in $\mathcal{X}$ with positive probability mass, then our test will reject the null hypothesis of affiliation with probability one asymptotically.

2.3.2 Uniform asymptotic properties of test (4)

In Appendix A.3 we study the asymptotic properties (size and power) of test (4) over *sequences* of distributions $F_n \in \mathcal{F}$. We assume that every distribution in $\mathcal{F}$ satisfies the smoothness and regularity assumptions described previously and we add conditions that ensure that the appropriate CLTs and LLNs for triangular arrays hold. Let us index the various functionals in our test by F and the objects that depend on the bandwidth by h. We assume that the underlying space of distributions $\mathcal{F}$ is such

[4]Note that under the null hypothesis of affiliation, $\psi^a\left(Z_i,\delta\right) = 0$ and therefore under the null hypothesis we can use $\psi^Q_n(Z_i) \equiv \int \psi^b(Z_i,\delta;h_n)dQ(\delta)$.

that, for some $\Delta > 0$ and $b < \infty$,

$$\underset{\substack{F \in \mathcal{F} \setminus \overline{\mathcal{F}}^Q \\ h > 0}}{sup} E_F \left[\frac{\left|\psi_F^Q(Z,h)\right|^{2+\Delta}}{\sigma_{Q,F}^{2+\Delta}(h)} \right] \leq b,$$

where $\overline{\mathcal{F}}^Q$ is the collection of distributions in $\mathcal{F}$ that satisfy strict conditional affiliation. This restriction ensures that a Lindeberg condition holds over the set of distributions that do not satisfy strict conditional affiliation. Given the linear representation in Theorem 1, this will help ensure that the asymptotic size properties of our test hold uniformly over $\mathcal{F}$.

The linear representation in Theorem 1 facilitates the analysis of power of our test. Take a sequencee $\{F_n\}$ where affiliation is violated. The asymptotic power of our test along any such sequence will depend critically on the limit of the sequence

$$s_{2,n}(F_n) \equiv \frac{\sqrt{n} \cdot \mathcal{T}_{F_n}^Q}{\max\{\kappa_n, \sigma_{Q,F_n}(h_n)\}}.$$

Take any sequence $\{F_n\}$ such that $s_{2,n}(F_n)$ converges (possibly to $+\infty$). Then we have the following:

(i) If $s_{2,n}(F_n) \to \infty$ then our test will have asymptotic power of 1. This includes all sequences such that $\mathcal{T}_{F_n}^Q = O\big(n^{-\alpha}\big)$ for $\alpha < 1/2$.

(ii) If $s_{2,n}(F_n) \to 0$ then the asymptotic power of our test will be at most α. This includes all sequences such that $\mathcal{T}_{F_n}^Q = O\big(n^{-\alpha}\big)$ for $\alpha > 1/2$.

(iii) Consider a sequence such that $\mathcal{T}_{F_n}^Q = t/\sqrt{n}$ for some $t > 0$. The asymptotic power of our test for any such sequence will be at least α. It will be 1 if $\sigma_{Q,F_n}(h_n) \to 0$, which will occur if and only if $\Delta_{F_n}^Q \to 0$. If $\Delta_{F_n}^Q \nrightarrow 0$, then we will have $\sigma_{Q,F_n}(h_n) \to \sigma^* > 0$. In this case, $s_{1,n}(F_n) \to 1$ and the asymptotic power of our test will be $1 - \Phi\left(z_{1-\alpha} - \frac{t}{\sigma^*}\right) \geq \alpha$.

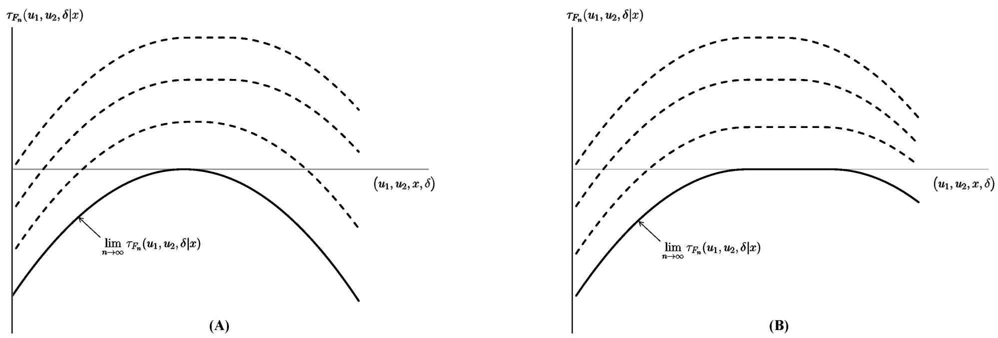
Now denote

$$\Delta_F^Q = \int\int\int\int \mathbb{1}\left\{\tau_F(u_1, u_2, \delta|x) = 0\right\} dQ(\delta) dF_U(u_1) dF_U(u_2) dF_X(x).$$

This is the probability that our affiliation inequalities are binding over our testing range. Now suppose $\mathcal{T}_{F_n}^Q = t/\sqrt{n}$. The asymptotic power of our test for any such sequence will be at least α. It will be 1 if $\Delta_{F_n}^Q \longrightarrow 0$ and it will be smaller than 1 otherwise. In general, the asymptotic power of our test for local sequences of the form $\mathcal{T}^Q_{F_n} = t/\sqrt{n}$ will suffer if $\Delta^Q_{F_n}$ remains bounded away from zero. If such limit is zero then the asymptotic power will be 1.

Figure 1 depicts this result in an informal way. There we have reduced (u_1, u_2, x, δ) into a one-dimensional aggregate representation on the horizontal axis. Segments of positive length along the horizontal axis are meant to represent ranges of values of (u_1, u_2, x, δ) with positive measure. In the figure we depict $\tau_{F_n}(u_1, u_2, \delta|x)$. Panel (A) is meant to describe the properties of a sequence F_n satisfying $\mathcal{T}^Q_{F_n} = t/\sqrt{n}$ for which our test would have asymptotic power of 1. This would be any sequence for which $\Delta^Q_{F_n} \to 0$. Panel (B) describes a sequence for which this would not be true, which would be any sequence for which $\Delta^Q_{F_n} \nrightarrow 0$.

Figure 1: Graphical depiction of power properties for distribution sequences $\{F_n\}$ satisfying $\mathcal{T}^Q_{F_n} = t/\sqrt{n}$. Our test will have asymptotic power 1 for sequences that behave like (A) and asymptotic power smaller than 1 for sequences that behave like (B).



2.4 A computationally simpler test

Implementing the test in (4) requires computing a U-statistic of order four (to estimate $\sigma^2_{Q,n}$). This can be computationally costly for moderately large sample sizes. We can construct ann alternative test that requires U-statistics of lower order. First recall that the statistic $\widehat{t}_n$ is designed to test whether the affiliation inequalities hold $H-$a.s over $\mathcal{S}$, where $H = F_U \times F_U \times F_X \times Q$, and $\mathcal{S} = \mathcal{U} \times \mathcal{U} \times \mathcal{X} \times \mathcal{S}_Q$. Suppose instead that we replace H and $\mathcal{S}$ with

$$H' = F_U \times F_Z \times Q \quad \text{and} \quad \mathcal{S}' = \mathcal{U} \times \mathcal{Z} \times \mathcal{S}_Q.$$

As before let $R^Q(u_1, u_2|x) = \int \max\{\tau(u_1, u_2, \delta|x), 0\}\, dQ(\delta)$, but instead of focusing on $\mathcal{T}^Q$, consider the alternative functional

$$\mathcal{V}^Q = \int\int\int R^Q(u_1, u_2|x)dF_U(u_1)dF_Z(u_2, x) \tag{5}$$

Similar to $\mathcal{T}^Q$, we have

(i) $\mathcal{V}^Q \geq 0$.

(ii) $\mathcal{V}^Q = 0$ if and only if $\tau(u_1, u_2, \delta|x) \leq 0\ \ H'$−a.s over $\mathcal{S}'$.

The set $\mathcal{S}'$ is a subset of the testing range $\mathcal{S}$ used in the construction of $\widehat{t}_n$, but it contains the joint support of (U, X). The main practical difference is that $\mathcal{V}^Q$ can be estimated by a U-statistic of order 2:

$$\widehat{\mathcal{V}}^Q = \frac{1}{(n)_2}\sum_{i\neq j} \widehat{R}^Q(U_j, U_i|X_i)\cdot \omega(U_j, U_i, X_i), \tag{6}$$

which reduces the number of computations by a factor of n.

2.4.1 Asymptotic properties of $\widehat{\mathcal{V}}^Q$

The estimator $\widehat{\mathcal{V}}^Q$ satisfies a linear representation result analogous to that of $\widehat{\mathcal{T}}^Q$. Let $\widetilde{S}^a$ and $\widetilde{S}^b$ be as described previously and let

$$\begin{aligned}
\varphi^a(Z_i, \delta) &= 2\cdot\Big(E\big[\widetilde{S}^a(U_i, U_j, X_i, \delta)\big|Z_i\big] - \mathcal{V}(\delta)\Big),\\
\varphi^b(Z_i, \delta; h_n) &= E\left[\widetilde{S}^b(U_j, U_k, X_j, Z_i, \delta; h_n)\big|Z_i\right],\\
\varphi(Z_i, \delta; h_n) &= \varphi^a(Z_i, \delta) + \varphi^b(Z_i, \delta; h_n).
\end{aligned}$$

Again we have $E[\varphi(Z_i, \delta; h_n)] = 0$ for any δ. Let

$$\varphi_n^Q(Z_i) \equiv \int \varphi(Z_i, \delta; h_n)dQ(\delta).$$

Theorem 2. Under Assumptions 1-4,

$$\widehat{\mathcal{V}}^Q = \mathcal{V}^Q + \frac{1}{n}\sum_{i=1}^n \varphi_n^Q(Z_i) + O_p\left(n^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.$$

Proof: The steps of the proof are very similar to those of Theorem 1. Thus we omit it for simplicity. □

φ_n has properties analogous to ψ_n,

1.– $E\big[\varphi_n^Q(Z_i)\big] = 0.$

2.– $\varphi_n^Q(Z_i) = 0$ a.s if strict conditional affiliation holds.

Let $Var\big(\varphi_n^Q(Z_i)\big) \equiv \Omega^2_{Q,n}$. Appendix A.5 shows how to construct an estimator for $\Omega^2_{Q,n}$. Our test statistic has then the same structure as the one described above,

$$\widetilde{t}_n = \frac{\sqrt{n}\cdot\widehat{\mathcal{V}}^Q}{\max\left\{\kappa_n,\,\widehat{\Omega}_{Q,n}\right\}}, \tag{7}$$

where, as before, $\kappa_n \to 0$ is a nonnegative sequence satisfying $\kappa_n \cdot n^{\epsilon} \to \infty$ for all $\epsilon > 0$. We reject affiliation if $\widetilde{t}_n \geq z_{1-\alpha}$. This test will the asymptotic properties in (3A) and (3B) for the testing range $\mathcal{S}'$.

2.5 Tests when *U* is discrete

The tests described above are valid for continuous or discrete U. However, if U is discrete then generating the relevant contact sets can be greatly simplified since they consist simply of the union of points. This greatly simplifies our computations. Let

$$\overline{\mathcal{U}} = \{(u_1,u_2)\in\mathcal{U}\times\mathcal{U} :\ (u_1\vee u_2,\, u_1\wedge u_2) \neq (u_1,u_2) \text{ and } (u_1\vee u_2,\, u_1\wedge u_2)\neq(u_2,u_1)\}$$

$\overline{\mathcal{U}}$ represents the relevant range of pairwise values of U over which we need to test affiliation. We now have $G(u|x) = Pr\,(U = u|x)$. As before let $\mu(u|x) = G(u|x)\cdot f_X(x)$ and $\tau(u_1,u_2|x) = \mu(u_1|x)\cdot\mu(u_2|x) - \mu(u_1\vee u_2|x)\cdot\mu(u_1\wedge|x)$. Let

$$\mathcal{T}(u_1,u_2) = E_X\big[\max\{\tau(u_1,u_2|x),0\}\cdot\omega(u_1,u_2,x)\big].$$

Our functional of interest is now

$$\mathcal{T} = \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} \mathcal{T}(u_1,u_2).$$

Our estimators are of the form

$$\begin{aligned}
\widehat{\mu}(u|x) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left\{U_i = u\right\} \cdot \mathcal{H}\left(X_i - x; h_n\right), \\
\widehat{\tau}(u_1,u_2|x) &= \widehat{\mu}(u_1|x)\cdot\widehat{\mu}(u_2|x) - \widehat{\mu}(u_1 \vee u_2|x)\cdot\widehat{\mu}(u_1 \wedge u_2|x), \\
\widehat{\mathcal{T}}(u_1,u_2) &= \frac{1}{n}\sum_{i=1}^{n} \widehat{\tau}(u_1,u_2|X_i)\cdot \mathbb{1}\left\{\widehat{\tau}(u_1,u_2|X_i) \geq -b_n\right\}\cdot\omega(u_1,u_2,X_i), \\
\widehat{\mathcal{T}} &= \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} \widehat{\mathcal{T}}(u_1,u_2).
\end{aligned} \tag{8}$$

In this case we need a simpler version of the regularity condition in Assumption 2.

Assumption 2'.

There exist constants $\overline{C} > 0$ and $\bar{b} > 0$ such that for every $(u_1,u_2) \in \overline{\mathcal{U}}$,

$$Pr\left(-b \leq \tau(u_1,u_2,\delta|X) < 0\right) \leq \overline{C}\cdot b \quad \forall\, 0 < b \leq \bar{b}.$$

Let

$$\begin{aligned}
\lambda^a(u_1,u_2,X_i) &= \max\left\{\tau(u_1,u_2|X_i),0\right\}\cdot\omega(u_1,u_2,X_i) - \mathcal{T}(u_1,u_2), \\
\lambda^a(X_i) &= \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} \lambda^a(u_1,u_2,X_i),
\end{aligned}$$

and define

$$\begin{aligned}
&\zeta_\tau(u_1,u_2,x,Z_i;h_n) = \\
&\mu(u_2|x)\cdot\Big(\mathbb{1}\left\{U_i = u_1\right\}\cdot\mathcal{H}(X_i - x;h_n) - E\left[\mathbb{1}\left\{U_i = u_1\right\}\cdot\mathcal{H}(X_i - x;h_n)\right]\Big) \\
&+\mu(u_1|x)\cdot\Big(\mathbb{1}\left\{U_i = u_2\right\}\cdot\mathcal{H}(X_i - x;h_n) - E\left[\mathbb{1}\left\{U_i = u_2\right\}\cdot\mathcal{H}(X_i - x;h_n)\right]\Big) \\
&-\mu(u_1 \vee u_2|x)\cdot\Big(\mathbb{1}\left\{U_i = u_1 \wedge u_2\right\}\cdot\mathcal{H}(X_i - x;h_n) - E\left[\mathbb{1}\left\{U_i = u_1 \wedge u_2\right\}\cdot\mathcal{H}(X_i - x;h_n)\right]\Big) \\
&-\mu(u_1 \wedge u_2|x)\cdot\Big(\mathbb{1}\left\{U_i = u_1 \vee u_2\right\}\cdot\mathcal{H}(X_i - x;h_n) - E\left[\mathbb{1}\left\{U_i = u_1 \vee u_2\right\}\cdot\mathcal{H}(X_i - x;h_n)\right]\Big), \\
&\widetilde{q}_\tau(u_1,u_2,X_i,Z_j;h_n) = \zeta_\tau(u_1,u_2,X_i,Z_j;h_n)\cdot\mathbb{1}\left\{\tau(u_1,u_2|X_i) \geq 0\right\}\cdot\omega(u_1,u_2,X_i),
\end{aligned}$$

and

$$\begin{aligned}
\lambda_n^b(u_1,u_2,Z_i) &= E\left[\widetilde{q}_\tau(u_1,u_2,X_j,Z_i;h_n)\big|Z_i\right], \\
\lambda_n^b(Z_i) &= \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} \lambda_n^b(u_1,u_2,Z_i).
\end{aligned}$$

Finally let

$$\lambda_n(Z_i) = \lambda^a(X_i) + \lambda_n^b(Z_i)).$$

Note that $E[\lambda_n(Z_i)] = 0$ and $\lambda_n(Z_i)$ satisfies the same type of degeneracy properties as our previous influence functions under strict conditional affiliation. We have the following result.

Theorem 3. Under Assumptions 1, 2', 3 and 4,

$$\widehat{\mathcal{T}} = \mathcal{T} + \frac{1}{n}\sum_{i=1}^{n}\lambda_n(Z_i) + \xi_n, \quad \text{where} \quad \left|\xi_n\right| = O_p\left(n^{-1/2-\epsilon}\right) \;\; \text{for some } \epsilon > 0.$$

Proof: In Appendix A.4. □

A consistent estimator for $\sigma_n^2 \equiv Var\left(\lambda_n(Z_i)\right)$ can be easily constructed using sample analogs. We show how to construct it in Appendix A.5. Our test statistic is,

$$\bar{t}_n = \frac{\sqrt{n} \cdot \widehat{\mathcal{T}}}{\max\left\{\kappa_n, \widehat{\sigma}_n\right\}}, \tag{9}$$

where, once again $\kappa_n \to 0$ is a nonnegative sequence satisfying $\kappa_n \cdot n^{\epsilon} \to \infty$ for all $\epsilon > 0$. We reject affiliation if $\bar{t}_n \geq z_{1-\alpha}$. This test will have the asymptotic properties in (3A)-(3B).

2.5.1 A test for affiliation of binary choices using aggregate decisions

Suppose U represents a collection of L binary choice decisions. For example, participation decisions in an auction with L potential bidders. Suppose however that we do not observe the individual choices but rather their aggregate, $A = \sum_{\ell=1}^{L} U_\ell$. For example, we only observe the number of bidders but we do not distinguish their identities. Suppose that we maintain that all decisions (bidders) are *symmetric*. If we let $Pr(A = k|X, L) = G_{k,L}(X)$, symmetry and affiliation imply the inequality

$$\binom{L}{k}^{-2} \cdot G_{k,L}^2(X) \leq \binom{L}{k-1}^{-1} \cdot \binom{L}{k+1}^{-1} \cdot G_{k-1,L}(X) \cdot G_{k+1,L}(X),$$

for all $1 \leq k \leq L-1$ and all $L \geq 2$. As before, we will use a density-weighted version of the above inequality. Let $\mu_{k,L}(X) = G_{k,L}(X) \cdot f_{X,L}(X,L)$ and

$$\begin{aligned}\tau_{k,L}(X) &= \binom{L}{k}^{-2} \cdot \mu_{k,L}^2(X) - \binom{L}{k-1}^{-1} \cdot \binom{L}{k+1}^{-1} \cdot \mu_{k-1,L}(X) \cdot \mu_{k+1,L}(X) \\ &= \left(\binom{L}{k}^{-2} \cdot G_{k,L}^2(X) - \binom{L}{k-1}^{-1} \cdot \binom{L}{k+1}^{-1} \cdot G_{k-1,L}(X) \cdot G_{k+1,L}(X) \right) \cdot f_{X,L}^2(X,L)\end{aligned}$$

and let

$$\begin{aligned}\mathcal{Q}_{k,L} &= E_X \left[\max\left\{ \tau_{k,L}(X), 0 \right\} \cdot \omega(X,k,L) \right], \\ \mathcal{Q} &= \sum_{L=\underline{L}}^{\overline{L}} \sum_{k=1}^{L-1} \mathcal{Q}_{k,L}\end{aligned}$$

where $\omega(X,k,L)$ is a pre-specified, positive weighting function and $2 \leq \underline{L} \leq \overline{L}$ represent the range of values of L over which we wish to test affiliation. Affiliation implies $\mathcal{Q} = 0$, with $\mathcal{Q} > 0$ otherwise. Our estimators are of the same type as before. Namely,

$$\begin{aligned}\widehat{\mu}_{k,L}(x) &= \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}\left\{A_j = k\right\} \cdot \mathbb{1}\left\{L_j = L\right\} \cdot \mathcal{H}(X_j - x; h_n), \\ \widehat{\tau}_{k,L}(x) &= \binom{L}{k}^{-2} \cdot \widehat{\mu}_{k,L}^2(x) - \binom{L}{k-1}^{-1} \cdot \binom{L}{k+1}^{-1} \cdot \widehat{\mu}_{k-1,L}(x) \cdot \widehat{\mu}_{k+1,L}(x), \\ \widehat{\mathcal{Q}}_{k,L} &= \frac{1}{n} \sum_{i=1}^{n} \widehat{\tau}_{k,L}(X_i) \cdot \mathbb{1}\left\{\widehat{\tau}_{k,L}(X_i) \geq -b_n\right\} \cdot \omega(X_i,k,L), \\ \widehat{\mathcal{Q}} &= \sum_{L=\underline{L}}^{\overline{L}} \sum_{k=1}^{L-1} \widehat{\mathcal{Q}}_{k,L}\end{aligned}$$

Let $Z_i \equiv (A_i, L_i, X_i)$ and

$$\begin{aligned}&\phi_{k,L}(Z_i, x; h_n) = \\ &2\binom{L}{k}^{-2} \mu_{k,L}(x) \Big(\mathbb{1}\left\{A_i = k\right\} \mathbb{1}\left\{L_i = L\right\} \mathcal{H}(X_i - x; h_n) - E\big[\mathbb{1}\left\{A_i = k\right\} \mathbb{1}\left\{L_i = L\right\} \mathcal{H}(X_i - x; h_n) \big] \Big) \\ &- \binom{L}{k-1}^{-1} \binom{L}{k+1}^{-1} \times \\ &\Bigg\{ \mu_{k+1,L}(x) \cdot \Big(\mathbb{1}\left\{A_i = k-1\right\} \mathbb{1}\left\{L_i = L\right\} \mathcal{H}(X_i - x; h_n) - E\big[\mathbb{1}\left\{A_i = k-1\right\} \mathbb{1}\left\{L_i = L\right\} \mathcal{H}(X_i - x; h_n) \big] \Big) \\ &+ \mu_{k-1,L}(x) \cdot \Big(\mathbb{1}\left\{A_i = k+1\right\} \mathbb{1}\left\{L_i = L\right\} \mathcal{H}(X_i - x; h_n) - E\big[\mathbb{1}\left\{A_i = k+1\right\} \mathbb{1}\left\{L_i = L\right\} \mathcal{H}(X_i - x; h_n) \big] \Big) \Bigg\}\end{aligned}$$

Note that $E\left[\phi_{k,L}(Z_i, x; h_n)\right] = 0$. Let

$$\begin{aligned}
\gamma^a_{k,L}(X_i) &= \max\left\{\tau_{k,L}(X_i), 0\right\} \cdot \omega(X_i, k, L) - \mathcal{Q}_{k,L}, \\
\gamma^b_{k,L,n}(Z_i) &= E\big[\phi_{k,L}(Z_i, X_j; h_n) \cdot \mathbb{1}\left\{\tau_{k,L}(X_j) \geq 0\right\} \cdot \omega(X_j, k, L)\big| Z_i\big], \\
\gamma_n(Z_i) &= \sum_{L=\underline{L}}^{\overline{L}} \sum_{k=1}^{L-1} \Big(\gamma^a_{k,L}(X_i) + \gamma^b_{k,L,n}(Z_i)\Big)
\end{aligned}$$

Note that $E\left[\gamma_n(X_i)\right] = 0$. If $G_{k,L}(X)$ and $f_{X,L}(X, L)$ satisfy the same type of smoothness properties in Assumption 1 with respect to X^c, and if we maintain a regularity condition analogous to Assumption 2: for some $\overline{b} > 0$ and $\overline{C}$,

$$Pr\left(-b \leq \tau_{k,L}(X) < 0\right) \leq \overline{C} \cdot b \quad \forall\, 0 < b \leq \overline{b}$$

for all $1 \leq k \leq L-1$ and all $\underline{L} \leq \overline{L}$, then the bandwidth and kernel restrictions in Assumption 4 yield

$$\widehat{\mathcal{Q}} = \mathcal{Q} + \frac{1}{n}\sum_{i=1}^{n} \gamma_n(Z_i) + O_p\left(n^{-1/2-\epsilon}\right)$$

Letting $\Sigma_n^2 = Var\left(\gamma_n(Z_i)\right)$, our test-statistic would be of the form

$$\widehat{s}_n = \frac{\sqrt{n} \cdot \widehat{\mathcal{Q}}}{\max\{\kappa_n, \widehat{\Sigma}_n\}}. \tag{10}$$

And once again, for a target level α we reject affiliation if $\widehat{s}_n > z_{1-\alpha}$. We show how to construct $\widehat{\Sigma}_n$ in Appendix A.5. This test will have the same type of asymptotic properties described in equations (3A)-(3B).

2.6 Choice of tuning parameters

Our test involves three sets of bandwidths: h_n, b_n and κ_n. The first one, h_n is used to construct our nonparametric estimators of the conditional probabilities involved. Therefore we can choose among existing bandwidth selection procedures for kernel estimators summarized, for example, in Ichimura and Todd (2007, Section 6). The bandwidth sequences b_n and κ_n are specific to our problem. Even though a general theory to select these tuning parameters appears to be infeasible, we advocate choosing them based on scale considerations in a least favorable configuration.

Note that κ_n is asymptotically relevant only if affiliation is strict because it is only then that the variance of our various test-statistics is zero. On the other hand, if the elements of U are independent conditional on X then the variance of our test-statistics is always nonzero. For this reason we use the variance under independence as our benchmark in our choice of κ_n since it captures the scale of the variance of our estimators in a least favorable configuration. With independence we have $\mathbb{1}\{\tau(u_1, u_2, \delta|x) \geq 0\} = 1$ for all u_1, u_2, δ and all a.e $x \in \mathcal{X}$. Let ψ_τ, ζ_τ and $\phi_{k,L}$ be as described previously. Under independence we have

- For test $\widehat{t}_n$ in (4),

$$\psi_n^Q(Z_i) = \int E\left[\psi_\tau(U_j, U_k, X_\ell, Z_i, \delta; h_n) \cdot \omega(U_j, U_k, X_\ell)\big|Z_i\right] dQ(\delta) \equiv \overline{\psi}_n^Q(Z_i),$$
$$\sigma_{Q,n}^2 = Var\left(\overline{\psi}_n^Q(Z_i)\right) \equiv \overline{\sigma}_{Q,n}^2$$

- For test $\widetilde{t}_n$ in (7),

$$\varphi_n^Q(Z_i) = \int E\left[\psi_\tau(U_j, U_k, X_j, Z_i, \delta; h_n) \cdot \omega(U_j, U_k, X_j)\big|Z_i\right] dQ(\delta) \equiv \overline{\varphi}_n^Q(Z_i),$$
$$\Omega_{Q,n}^2 = Var\left(\overline{\varphi}_n^Q(Z_i)\right) \equiv \overline{\Omega}_{Q,n}^2$$

- For test $\overline{t}_n$ in (9),

$$\lambda_n(Z_i) = \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} E\left[\zeta_\tau(u_1, u_2, X_j, Z_i; h_n) \cdot \omega(u_1, u_2, X_j)\big|Z_i\right] \equiv \overline{\lambda}_n(Z_i),$$
$$\sigma_n^2 = Var\left(\overline{\lambda}_n(Z_i)\right) \equiv \overline{\sigma}_n^2$$

- For test $\widehat{s}_n$ in (10),

$$\gamma_n(Z_i) = \sum_{L=\underline{L}}^{\overline{L}} \sum_{k=1}^{L-1} E\big[\phi_{k,L}(Z_i, X_j; h_n) \cdot \omega(X_j, k, L)\big|Z_i\big] \equiv \overline{\gamma}_n(Z_i),$$
$$\Sigma_n^2 = Var\left(\overline{\gamma}_n(Z_i)\right) \equiv \overline{\Sigma}_n^2$$

Let c_b and c_κ be constants chosen by the econometrician and $r_n \to 0$ be such that $r_n \cdot n^\epsilon \to \infty$ for any $\epsilon > 0$. Let α_b be as described in the paragraph following Assumption 4. The bandwidth sequences we use are of the form

$$b_n = c_b \cdot \widehat{\widehat{\sigma}}_{Q,n} \cdot n^{-\alpha_b}, \quad \kappa_n = c_\kappa \cdot \widehat{\widehat{\sigma}}_{Q,n} \cdot r_n \quad \text{(for test } \widehat{t}_n)$$
$$b_n = c_b \cdot \widehat{\widehat{\Omega}}_{Q,n} \cdot n^{-\alpha_b}, \quad \kappa_n = c_\kappa \cdot \widehat{\widehat{\Omega}}_{Q,n} \cdot r_n \quad \text{(for test } \widetilde{t}_n)$$
$$b_n = c_b \cdot \widehat{\widehat{\sigma}}_n \cdot n^{-\alpha_b}, \quad \kappa_n = c_\kappa \cdot \widehat{\widehat{\sigma}}_n \cdot r_n \quad \text{(for test } \overline{t}_n)$$
$$b_n = c_b \cdot \widehat{\widehat{\overline{\Sigma}}}_n \cdot n^{-\alpha_b}, \quad \kappa_n = c_\kappa \cdot \widehat{\widehat{\overline{\Sigma}}}_n \cdot r_n \quad \text{(for test } \widehat{s}_n).$$

In our Monte Carlo experiments we fix $c_\kappa = 1$ and let $r_n = \frac{1}{\log(\log(n))}$ and we use $c_b \in \{0.1, 0.01, 0.001\}$, with $\alpha_b = \frac{1}{4} + 10^{-6}$. We show that our tests perform well in all the cases considered.

3 Monte Carlo experiments

Our tests involve the estimation of various nonparametric functionals and the choice of several tuning parameters. Therefore a study of their finite-sample properties is essential. To illuminate this we investigate the performance of our tests under different designs. Each design includes a rich collection of observable and unobservable covariates.

3.1 Experiments involving affiliation between continuous variables

Here we study a case where the econometrician observes a continuous decision variable (e.g, a bid in an auction). For simplicity we assume $L = 2$. For the i^{th} experiment (auction) the structure of our design is as follows,

$$U_{\ell,i} = W^a_{\ell,i} - W^b_{\ell,i} + D_i + \varepsilon_{\ell,i} + \eta_i.$$

We assume that the econometrician observes $X_i \equiv \left(D_i, W^a_{1,i}, W^a_{2,i}, W^b_{1,i}, W^b_{2,i}\right)$ and $(U_{1,i}, U_{2,i})$. We describe the DGP of X_i next.

DGP for $X_i \equiv \left(D_i, W^a_{1,i}, W^a_{2,i}, W^b_{1,i}, W^b_{2,i}\right)$

We generated $W^b_{\ell,i}$ simply as $W^b_{\ell,i} \sim U[0,1]$. The DGP of $W^a_{\ell,i}$ is the following: For each i we generated two random variables, φ^I_i and φ^{II}_i, iid $\mathcal{N}(0,1)$ and then we constructed the signals

$$\Xi^I_i = -\frac{1}{2} \cdot \varphi^I_i + \varphi^{II}_i, \quad \Xi^{II}_i = \varphi^I_i - \frac{1}{2} \cdot \varphi^{II}_i.$$

Each agent ℓ observes either signal Ξ_i^I or Ξ_i^{II} and $W_{\ell,i}^a$ corresponds to the signal observed by ℓ. We assume that agents always observe different signals from each other, each with probability 1/2. That is,

$$\left(W_{1,i}^a, W_{2,i}^a\right) \in \left\{\left(\Xi_i^I, \Xi_i^{II}\right),\ \left(\Xi_i^{II}, \Xi_i^I\right)\right\},$$

with probability $1/2$ for each of the two possible combinations. The mechanism $\mathcal{M}_w$ that determines the draw of the signals is independent of everything else in the design. Note that $W_{1,i}^a$ and $W_{2,i}^a$ are negatively correlated by design.

D_i is discrete and its value is associated with $W_{\ell,i}^a$ and $W_{\ell,i}^b$ in the following way,

$$D_i = \begin{cases} 2 & \text{if } \left(W_{1,i}^a - W_{1,i}^b\right) + \left(W_{2,i}^a - W_{2,i}^b\right) \leq -2 \\ 3 & \text{if } -2 < \left(W_{1,i}^a - W_{1,i}^b\right) + \left(W_{2,i}^a - W_{2,i}^b\right) \leq 0 \\ 4 & \text{if } \left(W_{1,i}^a - W_{1,i}^b\right) + \left(W_{2,i}^a - W_{2,i}^b\right) > 0 \end{cases}$$

DGP for η_i

Both $\varepsilon_{\ell,i}$ and η_i are assumed to be unobserved by the econometrician. η_i was simply generated as $\eta_i \sim U[0,1]$. We used two different designs to generate $\varepsilon_{\ell,i}$, which we describe next.

DGP 1 for $\varepsilon_{\ell,i}$

For each i we generated two random variables, ν_i^I and ν_i^{II}, both iid $\mathcal{N}(0,1)$ and the corresponding pair of signals

$$\theta_i^I = D_i \cdot \nu_i^I - \nu_i^{II}, \quad \theta_i^{II} = -\nu_i^I + D_i \cdot \nu_i^{II}.$$

Each ℓ observes one of these signals and $\varepsilon_{\ell,i}$ corresponds to the signal observed by ℓ. In DGP1 we assume that both $\ell = 1,2$ observe the *same signal*, which can be either θ_i^I or θ_i^{II} with probability 1/2 each. Therefore in DGP1 we have

$$(\varepsilon_{1,i}, \varepsilon_{2,i}) \in \left\{\left(\theta_i^I, \theta_i^I\right),\ \left(\theta_i^{II}, \theta_i^{II}\right)\right\},$$

with probability $1/2$ for each of the two possible combinations. The mechanism $\mathcal{M}_\varepsilon$ that determines the draw is independent of everything else.

DGP 2 for $\varepsilon_{\ell,i}$

θ_i^I and θ_i^{II} are generated in the same manner as in DGP1, but now we assume that agents always observe a different signal from each other. Thus in DGP2 we have,

$$(\varepsilon_{1,i}, \varepsilon_{2,i}) \in \left\{ \left(\theta_i^I, \theta_i^{II}\right),\ \left(\theta_i^{II}, \theta_i^I\right) \right\},$$

with probability $1/2$ for each of the two possible combinations. The mechanism $\mathcal{M}_\varepsilon$ that determines which signal is drawn is independent of everything else. Note that $\varepsilon_{1,i}$ and $\varepsilon_{2,i}$ are negatively correlated by design.

Conditional affiliation in our designs

$U_{1,i}$ and $U_{2,i}$ are not affiliated unconditionally. Affiliation conditional on X_i is satisfied in DGP 1 and violated in DGP 2. In fact, DGP 1 satisfies common values conditional on X_i.

3.2 Implementation of test (7)

We employed the test-statistic described in (7):

$$\widetilde{t}_n = \frac{\sqrt{n} \cdot \widehat{\mathcal{V}}^Q}{\max\left\{\kappa_n, \widehat{\Omega}_{Q,n}\right\}},$$

with $\omega(u_1, u_2, x) = 1$ and $\widehat{\Omega}_{Q,n}$ constructed as shown in Appendix A.5. Our conditioning covariates are $X_i \equiv \left(D_i, W_{1,i}^a, W_{2,i}^a, W_{1,i}^b, W_{2,i}^b\right)$ which include $q = 4$ continuous covariates, and a discrete one.

3.2.1 Kernels and bandwidths used

We adhered to the conditions of Assumption 4 as follows.

Kernels

Our vector of observable covariates X_i includes $q = 4$ continuous variables. Accordingly we used a bias-reducing kernel of order $M = 12$, which was constructed as

$$k(\psi_j) = \left(\sum_{j=0}^{10} a_j \cdot \psi^j\right) \cdot \mathbb{1}\left\{|\psi_j| \leq s\right\}, \tag{11}$$

with $K(\psi_1, \ldots, \psi_q) = \prod_{j=1}^{q} k(\psi_j)$. The coefficients $(a_j)_{j=0}^{10}$ are constructed to satisfy the bias-reducing conditions of a 12^{th}-order kernel as well as bias-adjustments at the boundary. The support of the kernel was the interval $[-s, s] = [-30, 30]$.

Bandwidths

We begin with the bandwidth h_n used to construct our conditional probabilities. We employed a bandwidth of the form $h_n = c \cdot \widehat{\sigma}(X) \cdot n^{-\alpha_h}$ (note that each X has its own bandwidth), where $\alpha_h = \frac{1}{2M} + \overline{\epsilon}$ and $\overline{\epsilon} = 10^{-6}$. This convergence rate will satisfy Assumption 4. As a guidance to select the constant 'c' we used the "rule of thumb" formula (Silverman (1986)), using the Normal distribution as the reference distribution. We chose

$$\begin{aligned} &h_n = c \cdot \widehat{\sigma}(X) \cdot n^{-\alpha_h}, \quad \text{where} \\ &c = 2 \cdot \left(\frac{\pi^{1/2}\,(M!)^3 \cdot R_k}{(2M) \cdot (2M)! \cdot \left(k_{_M}^2\right)}\right)^{\frac{1}{2M+1}}, \quad R_k = \int_{-s}^{s} k^2(u)du, \quad \text{and} \quad k_{_M} = \int_{-s}^{s} u^M k(u)du \end{aligned} \tag{12}$$

This yielded $c \approx 0.15$. We chose the bandwidths b_n and κ_n as in Section 2.6. We used

$$b_n = c_b \cdot \widehat{\widehat{\Omega}}_{Q,n} \cdot n^{-\left(\frac{1}{4}+\overline{\epsilon}\right)}, \quad \text{and} \quad \kappa_n = \frac{\widehat{\widehat{\Omega}}_{Q,n}}{\log\left(\log(n)\right)}, \tag{13}$$

where

$$\begin{aligned} \widehat{\widehat{\varphi}}_n^Q(Z_i) &= \frac{1}{(n-1)_2} \sum_{\substack{j \neq k \\ (j,k) \neq i}} \int \widehat{\psi}_\tau(U_j, U_k, X_j, Z_i, \delta; h_n) dQ(\delta), \\ \widehat{\widehat{\Omega}}_{Q,n}^2 &= \frac{1}{n} \sum_{i=1}^{n} \widehat{\widehat{\varphi}}_n^Q(Z_i)^2, \end{aligned}$$

as described in Section 2.6

3.2.2 Choice of Q

To generate the "contact sets" we chose Q as the uniform measure in the rectangle $\left[\underline{\delta}_1, \overline{\delta}_1\right] \times \left[\underline{\delta}_2, \overline{\delta}_2\right] \times \cdots \times \left[\underline{\delta}_L, \overline{\delta}_L\right]$, where

$$\underline{\delta}_\ell = \frac{1}{10} \cdot \max_{i=1,\ldots,n} \{U_{\ell,i}\} \quad \text{and} \quad \overline{\delta}_\ell = \frac{1}{2} \cdot \max_{i=1,\ldots,n} \{U_{\ell,i}\}. \tag{14}$$

To compute the test-statistic we simulated draws for δ from the distribution Q described above. In order to produce thousands of Monte Carlo simulations, we generated n cubes (where n is the sample size of each design). While this produces a relatively small collection of contact sets (cubes), it helps us evaluate whether our procedure has good power properties even when the class of contact sets used is not very rich.

3.2.3 Monte Carlo results for our continuous affiliation designs

Given our choice of tuning parameters, our test is constructed as in (7) for three different values of the constant c_b used in the construction of b_n: $c_b = \{0.1, 0.01, 0.001\}$. Our nominal target sizes were $\alpha = 1\%$ and $\alpha = 5\%$. We generated 500 simulations each for the following sample sizes: $n \in \{500, 750, 1000\}$. The results are presented in Table 1. They can be summarized as follows.

- The results for DGP 1 were very much in line with the asymptotic predictions under strict affiliation, with rejection frequences near or below the nominal size values of 1% and 5%.
- Our test showed good power properties for the design in DGP 2 for all sample sizes. This is encouraging given the relatively simple class of contact sets (cubes) generated by it.
- As one would expect, the rejection frequencies are sensitive to the choice of c_b. Still, the qualitative features of power and size were robust to the range of values used for this constant. For the designs analyzed here, using smaller values of c_b increased power for DGP 2 but it did not do so at the expense of increases in terms of size (false rejections) for DGP 1 above the nominal size levels.

3.3 Experiments involving affiliation between discrete variables

Let U be generated according to the designs DGP1 and DGP2 described above. Accordingly, suppose the observable conditioning covariates X are exactly as described above. However, suppose that instead of observing U we only observe the aggregate variable

$$A_i = \mathbb{1}\{U_{1,i} \geq 2\} + \mathbb{1}\{U_{2,i} \geq 2\}.$$

We can think of the above indicator functions as participation decisions and A_i as the aggregate participation. This corresponds to the setting described in Section 2.5.1 and therefore we apply test $\widehat{s}_n$ described in (10). We use the same kernels and bandwidths described above.

3.3.1 Monte Carlo results for our discrete affiliation designs

We use the test-statistic $\widehat{s}_n$ with the same kernels and bandwidths described in Section 2.5.1. In this case we have $\underline{L} = \overline{L} = 2$. We also use $\omega(X, k, L) = 1$ as our weights. Note that this test does not involve Q. We use 2000 simulations for this computationally much simpler test. Table 2 summarizes the results from our experiments. As we can see there, using the same tuning parameters as in our continuous affiliation designs leads to slightly more conservative results in this case. In particularly, it leads to comparatively less power for small sample sizes. However, power quickly picks up for moderately larger sample sizes. The sensitivity of our results is also relatively less sensitive to the choice of c_b compared to our continuous affiliation tests. Overall, the results in both the continuous and discrete designs are in line with our asymptotic predictions.

Table 1: Summary of Monte Carlo experiments for our continuous designs.

	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
	DGP 1					
	$c_b = 0.1$					
$n = 500$	0.0000	0.0000	−1.2761	1.4673	1.0855	1.2564
$n = 750$	0.0000	0.0000	−1.9507	1.5015	0.9464	1.2656
$n = 1000$	0.0000	0.0000	−2.5148	1.2877	0.7236	1.0351
	$c_b = 0.01$					
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.0040	0.0000	0.0667	1.7911	1.3887	1.5956
$n = 750$	0.0358	0.0119	0.3263	2.6997	1.5602	2.2354
$n = 1000$	0.0159	0.0039	0.3624	2.4686	1.4994	1.7280
	$c_b = 0.001$					
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.0159	0.0000	0.0694	1.8284	1.4423	1.6551
$n = 750$	0.0558	0.0199	0.3487	2.8282	1.6801	2.3562
$n = 1000$	0.0478	0.0039	0.3951	2.5280	1.6044	1.8796
	DGP 2					
	$c_b = 0.1$					
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.9204	0.6972	0.2079	4.5064	3.5679	4.0032
$n = 750$	0.9681	0.8929	0.7985	6.6961	4.7469	5.4928
$n = 1000$	1.0000	0.9920	1.6607	10.2750	7.7911	9.1137
	$c_b = 0.01$					
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.9442	0.8207	0.2247	4.7172	3.8846	4.3442
$n = 750$	0.9801	0.9482	0.8212	7.1461	5.1828	5.9585
$n = 1000$	1.0000	0.9920	1.9201	10.8540	8.4623	9.7130
	$c_b = 0.001$					
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.9482	0.8366	0.2381	4.7357	3.9078	4.3928
$n = 750$	0.9801	0.9641	0.8413	7.5844	5.4113	6.1193
$n = 1000$	1.0000	0.9920	1.9663	10.9733	8.5493	9.9151

Table 2: Summary of Monte Carlo experiments for our discrete designs.

	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
DGP 1						
$c_b = 0.1$						
$n = 500$	0.0000	0.0000	0.0298	1.5728	1.1400	1.3117
$n = 750$	0.0002	0.0000	0.0872	1.9222	1.1840	1.6901
$n = 1000$	0.0003	0.0000	0.0094	2.0622	1.2147	1.5995
$c_b = 0.01$						
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.0000	0.0000	0.0320	1.6217	1.1821	1.3520
$n = 750$	0.0021	0.0000	0.0898	2.0120	1.2107	1.7842
$n = 1000$	0.0026	0.0000	0.0256	2.1471	1.3713	1.7043
$c_b = 0.001$						
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.0010	0.0000	0.0481	1.7388	1.2217	1.3806
$n = 750$	0.0022	0.0000	0.1162	2.0975	1.2964	1.8950
$n = 1000$	0.0040	0.0002	0.0798	2.3409	1.4640	1.9069
DGP 2						
$c_b = 0.1$						
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.6062	0.3240	0.2001	3.1422	2.7785	2.9536
$n = 750$	0.7135	0.4612	0.2560	3.8318	3.1706	3.3540
$n = 1000$	0.8412	0.7102	0.2897	4.7910	4.0990	4.4952
$c_b = 0.01$						
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.6440	0.3640	0.2318	3.3033	2.8098	3.0111
$n = 750$	0.7468	0.4972	0.2669	3.9110	3.3404	3.6019
$n = 1000$	0.8740	0.7356	0.2910	4.8124	4.1717	4.5702
$c_b = 0.001$						
	% rejections, 5% nominal size	% rejections, 1% nominal size	smallest value observed	largest value observed	95th percentile	99th percentile
$n = 500$	0.6870	0.4012	0.2488	3.7916	2.9920	3.1590
$n = 750$	0.7849	0.5460	0.2707	4.0466	3.5762	3.8156
$n = 1000$	0.9015	0.7660	0.3108	5.0136	4.3302	4.6263

4 Testing for conditional affiliation in Timber auctions

We test for affiliation of bids in sealed-bid auctions of timber tracts by the USFS in the southwern region of the United States (Region 8 of the USFS). The object being auctioned is the right to harvest timber in a specific tract of land. We have available two randomly drawn bids from sealed-bid auctions between 1982 and 1990. The continuous conditioning covariates we observe include: appraisal value (calculated by the USFS as a function of all the information in the government cruise report of the tract in question), species concentration (the Herfindahl index compiled based on the relative volumes of the different species present) and a measure of timber density. We also observe the number of bidders in each auction. Thus we have a total of $q = 3$ continuous and one discrete variables in X. It is safe to assume that all the covariates included here are observable to the bidders at the time of the auction and are thus common knowledge. Our sample consisted of $n = 1048$ observations. We applied the test described in (7). Kernels and bandwidths were chosen as described in Section 3.2.1 (with $q = 3$ we used a bias-reducing kernel of order $M = 8$, with boundary bias correction). As in our Monte Carlo experiments, we computed the test for three values of c_b: 0.1, 0.01 and 0.001 (used to construct the bandwidth b_n). The distribution Q used to produce our contact sets (rectangles) was the same described in Section 3.2.2. In our tests we generated $30,000$ random draws of δ from Q to construct the contact sets. Our results are shown in Table 3.

Table 3: Conditional affiliation test in Timber auctions.

	$c_b = 0.1$		$c_b = 0.01$		$c_b = 0.001$	
n	test statistic	p-value	test statistic	p-value	test statistic	p-value
1048	3.2101	0.0006	3.2876	0.0005	3.3669	0.0004

• c_b is used in the construction of $b_n = c_b \cdot \widehat{\widehat{\Omega}}_{Q,\mathcal{Z}} \cdot n^{-\left(\frac{1}{4}+10^{-6}\right)}$.

Affiliation of bids is rejected with p-values close to zero. Since the covariates used to condition are common knowledge to bidders at the time of the auction and they include relevant information about the good being auctioned[5], rejection of affiliation is puzzling, appearing to call into question whether the equilibrium behavior in Milgrom and Weber (1982) is an adequate representation of bidding behavior in this particular case. Another possibility is that bidders' signals (values) are indeed affiliated, but the bids observed correspond to the selection of a non-monotonic bidding

[5] In timber auctions, Lu and Perrigne (2008) treated appraisal value itself as a sufficient statistic for the auction's characteristics, motivated by the fact that appraisal value summarizes all the information in the government cruise report. They also showed that that appraisal value alone explained most of the variation in transaction price relative to the other observed covariates.

equilibrium. McAdams (2007) shows that in this case we should observe ties in bids with positive probability. While this feature is not apparent in the data, we did not formally test it.

4.1 Conditional vs unconditional affiliation

To illustrate the importance of controlling for common knowledge characteristics in an auction, we performed an unconditional test for affiliation where we did not control for any of the observable covariates described above (i.e, we set $X = 1$). As the results in Table 4 show, by failing to control for common knowledge characteristics in the auction we do not reject affiliation of bids.

Table 4: Unconditional affiliation test in Timber auctions.

	$c_b = 0.1$		$c_b = 0.01$		$c_b = 0.001$	
n	test statistic	p-value	test statistic	p-value	test statistic	p-value
1048	0.0038	0.4984	0.0041	0.4983	0.0047	0.4981

• c_b is used in the construction of $b_n = c_b \cdot \widehat{\widehat{\Omega}}_{Q,\mathcal{Z}} \cdot n^{-\left(\frac{1}{4}+10^{-6}\right)}$.

As we stated previously, we believe that in the specific context of an auction an appropriate test for affiliation should condition on all available covariates that capture common knowledge at the time of the auction since this is how agents condition their beliefs and it is therefore the relevant distribution. We also tried to argue that the particular set of observable covariates that we included in our application is an adequate control for the relevant information available to bidders (since it includes appraisal value, which summarizes all the information in the government cruise report).

5 Concluding remarks

We presented nonparametric, consistent tests for affiliation that enable the researcher to condition on observables with rich support in a very general way. Doing so is important because it helps us control for common knowledge characteristics that rational agents condition their beliefs on. Our tests are based on one-sided L_1 statistics which possess asymptotically pivotal properties under smoothness and regularity conditions. Our procedures displayed good properties in Monte Carlo analysis. We applied our procedures to test affiliation in Timber auction data and we showed that, while an unconditional test fails to reject affiliation of bids, once we control for observable auction characteristics our test rejects affiliation. This finding could still be reconciled with affiliated bidder signals (values) if the equilibrium selected is non-monotonic. Selection of such an equilibrium has observable implications, but we leave a formal test for it to future research.

A Appendix– Proofs

A.1 Proof of Theorem 1

The proof proceeds in three steps.

Step 1

Our first step is to show that under our assumptions, there exist $D_1 > 0$, $D_2 > 0$ and $D_3 > 0$ such that

$$Pr\left(\sup_{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\tau}(u_1,u_2,\delta|x)-\tau(u_1,u_2,\delta|x)\right|\geq b_n\right)\leq D_1\exp\left\{-\left(\sqrt{n}h_n^q\left(D_2\cdot b_n-D_3\cdot h_n^M\right)\right)^2\right\}.$$

For a given u, x and δ we have defined

$$\mu(u,\delta|x)=G(u,\delta|x)\cdot f_X(x),$$
$$\widehat{\mu}(u,\delta|x)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}\left\{u-\delta\leq U_i\leq u+\delta\right\}\cdot\mathcal{H}\left(X_i-x;h_n\right).$$

Using an M^{th} order approximation, our smoothness restrictions in Assumption 1 imply the existence of a finite constant $\overline{M}$ such that

$$\sup_{\substack{(u,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|E\left[\widehat{\mu}(u,\delta|x)\right]-\mu(u,\delta|x)\right|\leq\overline{M}\cdot h_n^M. \qquad \text{(A-1)}$$

Invoking Lemma 22 in Nolan and Pollard (1987) and Lemmas 2.4 and 2.14 in Pakes and Pollard (1989), having a kernel of bounded variation implies that the class of functions

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

is *Euclidean*[6] with respect to the constant envelope $\overline{K}$. Lemma 2.4 in Pakes and Pollard (1989) also implies that the class of functions

$$\mathscr{G}_{\delta,v} = \left\{g : g(u) = \mathbb{1}\left\{v - \delta \leq u \leq v + \delta\right\} \text{ for some } (v,\delta) \in \mathbb{R}^{2L}\right\}$$

is Euclidean with respect to the envelope 1. From here, Lemma 2.14 in Pakes and Pollard (1989) implies that the class of functions

$$\mathscr{F}_{v,y,\delta,h} = \left\{f : f(u,x) = \mathbb{1}\left\{v - \delta \leq u \leq v + \delta\right\} \cdot \mathcal{H}(x-y;h) \text{ for some } (v,\delta) \in \mathbb{R}^{2L}, y \in d_x, h > 0\right\}$$

is Euclidean with respect to the envelope $\overline{K}$. Since this envelope is constant, the maximal inequality results in Pollard (1990, Chapter 7) combined with the bias conditions in A-1 imply that there exist positive constants A_1, A_2 and A_3 such that for any $b > 0$,

$$Pr\left(\sup_{\substack{(u,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\mu}(u,\delta|x) - \mu(u,\delta|x)\right| \geq b\right) \leq A_1 \cdot \exp\left\{-\left(\sqrt{n}\cdot h_n^q\left(A_2\cdot b - A_3\cdot h_n^M\right)\right)^2\right\}. \tag{A-2}$$

And by Sherman (1994, Corollary 4),

$$\sup_{\substack{(u,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\mu}(u,\delta|x) - \mu(u,\delta|x)\right| = O_p\left(\frac{\log(n)}{\sqrt{n\cdot h_n^q}}\right) + O\left(h_n^M\right) = O_p\left(\frac{\log(n)}{\sqrt{n\cdot h_n^q}}\right). \tag{A-3}$$

where the last equality follows from the bandwidth convergence conditions in Assumption 4. Note that

$$\begin{aligned}\sup_{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\tau}(u_1,u_2,\delta|x) - \tau(u_1,u_2,\delta|x)\right| \leq & 4\cdot\overline{f}\sup_{\substack{(u,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\mu}(u,\delta|x) - \mu(u,\delta|x)\right| \\ & + 2\sup_{\substack{(u,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\mu}(u,\delta|x) - \mu(u,\delta|x)\right|^2.\end{aligned}$$

[6] See Pakes and Pollard (1989, Definition 2.7).

Therefore for any $b > 0$,

$$Pr\left(\sup_{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\tau}(u_1,u_2,\delta|x)-\tau(u_1,u_2,\delta|x)\right|\geq b\right)$$
$$\leq Pr\left(4\cdot\overline{f}\sup_{\substack{(u,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\mu}(u,\delta|x)-\mu(u,\delta|x)\right|\geq\frac{b}{2}\right)+Pr\left(2\sup_{\substack{(u,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\mu}(u,\delta|x)-\mu(u,\delta|x)\right|^2\geq\frac{b}{2}\right)$$
$$\leq 2\cdot A_1\cdot\exp\left\{-\left(\sqrt{n}\cdot h_n^q\left(A_2\cdot\min\left\{\frac{b}{8\cdot\overline{f}},\frac{b^{1/2}}{2}\right\}-A_3\cdot h_n^M\right)\right)^2\right\}$$

Since $b_n\to 0$, for large enough n we will have $\min\left\{\frac{b_n}{8\cdot\overline{f}},\frac{b_n^{1/2}}{2}\right\}=\frac{b_n}{8\cdot\overline{f}}$ and consequently for large enough n,

$$\begin{aligned}&Pr\left(\sup_{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\tau}(u_1,u_2,\delta|x)-\tau(u_1,u_2,\delta|x)\right|\geq b_n\right)\\&\leq 2\cdot A_1\cdot\exp\left\{-\left(\sqrt{n}\cdot h_n^q\cdot\left(A_2\cdot\frac{b_n}{8\cdot\overline{f}}-A_3\cdot h_n^M\right)\right)^2\right\}\\&\equiv D_1\cdot\exp\left\{-\left(\sqrt{n}\cdot h_n^q\left(D_2\cdot b_n-D_3\cdot h_n^M\right)\right)^2\right\},\end{aligned}\tag{A-4}$$

where $D_1\equiv 2\cdot A_1$, $D_2\equiv\frac{A_2}{8\cdot\overline{f}}$ and $D_3\equiv A_3$. Furthermore, by (A-3) and our results above we also obtain

$$\sup_{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}\left|\widehat{\tau}(u_1,u_2,\delta|x)-\tau(u_1,u_2,\delta|x)\right|=O_p\left(\frac{\log(n)}{\sqrt{n\cdot h_n^q}}\right)+O\left(h_n^M\right)=O_p\left(\frac{\log(n)}{\sqrt{n\cdot h_n^q}}\right).\tag{A-5}$$

Step 2

Here we use the results from Step 1 to show that

$$\widehat{\mathcal{T}}^Q=\frac{1}{(n)_3}\sum_{i\neq j\neq k}\int\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\tau(U_j,U_k,\delta|X_i)\geq 0\right\}dQ(\delta)\cdot\omega(U_j,U_k,X_i)+\varphi_n,$$
$$\text{where}\quad\left|\varphi_n\right|=O_p\left(n^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0$$

Take any $\delta \in \mathcal{S}_Q$. Then,

$$\begin{aligned}&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\widehat{\tau}(U_j,U_k,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_k,X_i)=\\&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\tau(U_j,U_k,\delta|X_i)\geq 0\right\}\cdot\omega(U_j,U_k,X_i)+\varphi_n(\delta),\end{aligned} \tag{A-6}$$

where

$$\begin{aligned}\big|\varphi_n(\delta)\big| \leq &\underbrace{\left|\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{-2b_n\leq\tau(U_j,U_k,\delta|X_i)<0\right\}\cdot\omega(U_j,U_k,X_i)\right|}_{\equiv|\varphi_{1,n}(\delta)|}\\&+\underbrace{\left|\frac{2}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\big|\widehat{\tau}(U_j,U_k,\delta|X_i)-\tau(U_j,U_k,\delta|X_i)\big|\geq b_n\right\}\cdot\omega(U_j,U_k,X_i)\right|}_{\equiv|\varphi_{2,n}(\delta)|}.\end{aligned} \tag{A-7}$$

We begin by analyzing $|\varphi_{2,n}(\delta)|$. Using (A-3) we have $\underset{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}{sup}\big|\widehat{\tau}(u_1,u_2,\delta|x)\big| = O_p(1)$. Therefore,

$$\begin{aligned}&\underset{\delta\in\mathcal{S}_Q}{sup}\ \big|\varphi_{2,n}(\delta)\big|\\&\leq O_p(1)\cdot\underset{\delta\in\mathcal{S}_Q}{sup}\ \left|\frac{1}{(n)_3}\sum_{i\neq j\neq k}\mathbb{1}\left\{\big|\widehat{\tau}(U_j,U_k,\delta|X_i)-\tau(U_j,U_k,\delta|X_i)\big|\geq b_n\right\}\cdot\omega(U_j,U_k,X_i)\right|.\end{aligned}$$

Take any $\alpha > 0$ and any $\varepsilon > 0$. Then,

$$
\begin{aligned}
&Pr\left(n^{\alpha}\cdot \underset{\delta\in\mathcal{S}_Q}{sup}\left|\frac{1}{(n)_3}\sum_{i\neq j\neq k}\mathbb{1}\left\{\left|\widehat{\tau}(U_j,U_k,\delta|X_i)-\tau(U_j,U_k,\delta|X_i)\right|\geq b_n\right\}\cdot\omega(U_j,U_k,X_i)\right|>\varepsilon\right)\\
&\leq Pr\left(\mathbb{1}\left\{\underset{\delta\in\mathcal{S}_Q}{sup}\left|\widehat{\tau}(U_j,U_k,\delta|X_i)-\tau(U_j,U_k,\delta|X_i)\right|\geq b_n\right\}\cdot\omega(U_j,U_k,X_i)\neq 0 \text{ for some } i,j,k\right)\\
&\leq \sum_{i\neq j\neq k} Pr\left(\mathbb{1}\left\{\underset{\delta\in\mathcal{S}_Q}{sup}\left|\widehat{\tau}(U_j,U_k,\delta|X_i)-\tau(U_j,U_k,\delta|X_i)\right|\geq b_n\right\}\cdot\omega(U_j,U_k,X_i)\neq 0\right)\\
&\leq n^3\cdot Pr\left(\underset{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}{sup}\left|\widehat{\tau}(u_1,u_2,\delta|x)-\tau(u_1,u_2,\delta|x)\right|\geq b_n\right)\\
&\leq n^3\cdot D_1\cdot\exp\left\{-\left(\sqrt{n}\cdot h_n^q\left(D_2\cdot b_n-D_3\cdot h_n^M\right)\right)^2\right\}\\
&= D_1\cdot\exp\left\{-\left(\sqrt{n}\cdot h_n^q\left(D_2\cdot b_n-D_3\cdot h_n^M\right)\right)^2+3\cdot\log(n)\right\}\longrightarrow 0,
\end{aligned}
$$

where the last result follows from our bandwidth convergence restrictions. Therefore, $\underset{\delta\in\mathcal{S}_Q}{sup}\left|\varphi_{2,n}(\delta)\right| = o_p\left(n^{-\alpha}\right)$. In particular the much weaker (but useful for our purposes) result holds,

$$
\underset{\delta\in\mathcal{S}_Q}{sup}\left|\varphi_{2,n}(\delta)\right| = O_p\left(n^{-1/2-\epsilon}\right) \text{ for some } \epsilon > 0. \tag{A-8}
$$

To study the properties of $\left|\varphi_{1,n}(\delta)\right|$, the following object is relevant. Define,

$$
\begin{aligned}
\widetilde{q}(X_i,U_j,U_k,\delta,b) &= \mathbb{1}\left\{-2b\leq\tau(U_j,U_k,\delta|X_i)<0\right\}\cdot\omega(U_j,U_k,X_i),\\
q(Z_i,Z_j,Z_k,\delta,b) &= \sum_c \frac{\widetilde{q}(X_{m_1},U_{m_2},U_{m_3},\delta,b)}{3!},\\
V_n(\delta,b) &= \binom{n}{3}^{-1}\sum_{i<j<k} q(Z_i,Z_j,Z_k,\delta,b),
\end{aligned}
$$

where $\sum_c$ denotes the sum over all 3! combinations $\{m_1, m_2, m_3\}$ of $\{i, j, k\}$. Now let

$$\begin{aligned}
V(\delta, b) &= E\left[q(Z_i, Z_j, Z_k, \delta, b)\right], \\
\overline{q}_1(Z_i, \delta, b) &= E\left[q(Z_i, Z_j, Z_k, \delta, b)\big|Z_i\right] - V(\delta, b), \\
\overline{q}_2(Z_i, Z_j, \delta, b) &= E\left[q(Z_i, Z_j, Z_k, \delta, b)\big|Z_i, Z_j\right] - V(\delta, b) - \overline{q}_1(Z_i, \delta, b) - \overline{q}_1(Z_j, \delta, b), \\
\overline{q}_3(Z_i, Z_j, Z_k, \delta, b) &= q(Z_i, Z_j, Z_k, \delta, b) - V(\delta, b) - \overline{q}_1(Z_i, \delta, b) - \overline{q}_1(Z_j, \delta, b) - \overline{q}_1(Z_k, \delta, b) \\
&\quad - \overline{q}_2(Z_i, Z_j, \delta, b) - \overline{q}_2(Z_i, Z_k, \delta, b) - \overline{q}_2(Z_j, Z_k, \delta, b)
\end{aligned}$$

And denote

$$V_{1,n}(\delta, b) = \frac{1}{n}\sum_{i=1}^{n} \overline{q}_1(Z_i, \delta, b), \quad V_{2,n}(\delta, b) = \binom{n}{2}^{-1} \sum_{i<j} \overline{q}_2(Z_i, Z_j, \delta, b),$$

$$V_{3,n}(\delta, b) = \binom{n}{3}^{-1} \sum_{i<j<k} \overline{q}_2(Z_i, Z_j, Z_k, \delta, b).$$

The Hoeffding decomposition of $V_n(\delta, b)$ is given by,

$$V_n(\delta, b) = V(\delta, b) + 3 \cdot V_{1,n}(\delta, b) + 3 \cdot V_{2,n}(\delta, b) + V_{3,n}(\delta, b).$$

For each $b > 0$ we have,

$$\underset{\delta\in\mathcal{S}_Q}{sup} \left|V_n(\delta, b)\right| \leq \underset{\delta\in\mathcal{S}_Q}{sup} \left|V(\delta, b)\right| + 3 \cdot \underset{\delta\in\mathcal{S}_Q}{sup} \left|V_{1,n}(\delta, b)\right| + \underset{\delta\in\mathcal{S}_Q}{sup} \Big(3 \cdot \left|V_{2,n}(\delta, b)\right| + \left|V_{3,n}(\delta, b)\right|\Big).$$

Let $\overline{b}$ be as described in Assumption 2. The conditions described there imply that

$$\underset{\delta\in\mathcal{S}_Q}{sup} \left|V(\delta, b)\right| = O(b), \quad \underset{\delta\in\mathcal{S}_Q}{sup}\ Var\left[\overline{q}_1(Z_i, \delta, b)\right] = O(b) \quad \forall\, 0 < b \leq \frac{\overline{b}}{2} \tag{A-9}$$

Next note that by the manageability conditions in Assumption 3, the class of functions

$$\mathscr{F}_{\delta,b} = \left\{ f(z_1, z_2, z_3) = q(z_1, z_2, z_3, \delta, b) \text{ for some } \delta \in \mathcal{S}_Q \text{ and some } 0 < b \leq \frac{\overline{b}}{2} \right\}$$

is Euclidean for the constant envelope 1. The properties of $Var\left[\overline{q}_1(Z_i,\delta,b)\right]$ shown in (A-9) and Sherman (1994, Corollary 4) yield

$$\underset{\delta\in\mathcal{S}_Q}{sup}\left|V_{1,n}(\delta,b)\right| = O_p\left(\sqrt{\frac{b}{n}}\right) \quad \forall\, 0<b\leq\frac{\overline{b}}{2}.$$

Sherman (1994, Corollary 4) also implies,

$$\underset{\substack{\delta\in\mathcal{S}_Q\\ 0<b\leq\frac{\overline{b}}{2}}}{sup}\left(3\cdot\left|V_{2,n}(\delta,b)\right| + \left|V_{3,n}(\delta,b)\right|\right) = O_p\left(\frac{1}{n}\right).$$

Therefore,

$$\underset{\delta\in\mathcal{S}_Q}{sup}\left|V_n(\delta,b)\right| \leq O(b) + O_p\left(\sqrt{\frac{b}{n}}\right) + O_p\left(\frac{1}{n}\right) \quad \forall\, 0<b\leq\frac{\overline{b}}{2}. \tag{A-10}$$

Since $b_n \longrightarrow 0$, for large enough n we will have $b_n \leq \frac{\overline{b}}{2}$. Using (A-10), it follows that

$$\underset{\delta\in\mathcal{S}_Q}{sup}\left|V_n(\delta,b_n)\right| \leq O(b_n) + O_p\left(\sqrt{\frac{b_n}{n}}\right) + O_p\left(\frac{1}{n}\right) = O_p\left(b_n\right), \tag{A-11}$$

where the last equality follows from our bandwidth convergence restrictions which imply that $\sqrt{\frac{b_n}{n}} = o\left(b_n\right)$ and $\frac{1}{n} = o\left(b_n\right)$. Now let us analyze $\varphi_{1,n}(\delta)$ in (A-7). We have

$$\begin{aligned}
&\left|\varphi_{1,n}(\delta)\right| \\
&\leq \underbrace{\frac{1}{(n)_3}\sum_{i\neq j\neq k}\left|\tau(U_j,U_k,\delta|X_i)\right|\cdot\mathbb{1}\left\{-2b_n\leq\tau(U_j,U_k,\delta|X_i)<0\right\}\cdot\omega(U_j,U_k,X_i)}_{\equiv|\varphi^a_{1,n}(\delta)|} \\
&+ \underbrace{\frac{1}{(n)_3}\sum_{i\neq j\neq k}\left|\widehat{\tau}(U_j,U_k,\delta|X_i)-\tau(U_j,U_k,\delta|X_i)\right|\cdot\mathbb{1}\left\{-2b_n\leq\tau(U_j,U_k,\delta|X_i)<0\right\}\cdot\omega(U_j,U_k,X_i)}_{\equiv|\varphi^b_{1,n}(\delta)|}
\end{aligned}$$

Note that

$$\left|\varphi^a_{1,n}(\delta)\right| \leq 2b_n\cdot\frac{1}{(n)_3}\sum_{i\neq j\neq k}\mathbb{1}\left\{-2b_n\leq\tau(U_j,U_k,\delta|X_i)<0\right\}\cdot\omega(U_j,U_k,X_i) = 2b_n\cdot V_n(\delta,b_n),$$

where $V_n(\delta, b_n)$ is the U-statistic described previously. By the result in (A-11),

$$\underset{\delta\in\mathcal{S}_Q}{sup}\left|\varphi_{1,n}^{a}(\delta)\right| \leq 2b_n \cdot O_p\left(b_n\right) = O_p\left(b_n^2\right).$$

Next, by (A-5) we have

$$\begin{aligned}
&\left|\varphi_{1,n}^{a}(\delta)\right| \leq \\
&\underset{\substack{(u_1,u_2,x)\in\mathcal{Z}\\ \delta\in\mathcal{S}_Q}}{sup}\left|\widehat{\tau}(u_1,u_2,\delta|x)-\tau(u_1,u_2,\delta|x)\right| \cdot \frac{1}{(n)_3}\sum_{i\neq j\neq k}\mathbb{1}\left\{-2b_n \leq \tau(U_j,U_k,\delta|X_i)<0\right\}\cdot\omega(U_j,U_k,X_i)\\
&= O_p\left(\frac{\log(n)}{\sqrt{n\cdot h_n^q}}\right)\cdot V_n(\delta,b_n).
\end{aligned}$$

Once again, by (A-11) we have

$$\underset{\delta\in\mathcal{S}_Q}{sup}\left|\varphi_{1,n}^{b}(\delta)\right| \leq O_p\left(\frac{\log(n)}{\sqrt{n\cdot h_n^q}}\right)\cdot O_p\left(b_n\right) = O_p\left(\frac{\log(n)\cdot b_n}{\sqrt{n\cdot h_n^q}}\right).$$

Combining these results,

$$\underset{\delta\in\mathcal{S}_Q}{sup}\left|\varphi_{1,n}(\delta)\right| \leq O_p\left(b_n^2\right)+O_p\left(\frac{\log(n)\cdot b_n}{\sqrt{n\cdot h_n^q}}\right) = O_p\left(b_n^2\right) = O_p\left(n^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon>0, \tag{A-12}$$

where the second-to-last equality follows from our bandwidth convergence conditions in Assumption 4, which imply that $\frac{\log(n)}{\sqrt{n\cdot h_n^q\cdot b_n}} \longrightarrow 0$ and the last equality follows from the bandwidth convergence restrictions of b_n. Combining (A-7), (A-8) and (A-12), we obtain $\underset{\delta\in\mathcal{S}_Q}{sup}\left|\varphi_{1,n}(\delta)\right| = O_p\left(n^{-1/2-\epsilon}\right)$ for some $\epsilon > 0$ and therefore from (A-6) this implies that

$$\begin{aligned}
&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\widehat{\tau}(U_j,U_k,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_k,X_i) = \\
&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\tau(U_j,U_k,\delta|X_i)\geq 0\right\}\cdot\omega(U_j,U_k,X_i)+\varphi_n(\delta),
\end{aligned} \tag{A-13}$$

where $\underset{\delta\in\mathcal{S}_Q}{sup}\left|\varphi_n(\delta)\right| = O_p\left(n^{-1/2-\epsilon}\right)$ for some $\epsilon > 0$.

Therefore, $\int |\varphi_n(\delta)|\, dQ(\delta) = O_p\left(n^{-1/2-\epsilon}\right)$ for some $\epsilon > 0$. From here it follows that

$$\widehat{\mathcal{T}}^Q = \frac{1}{(n)_3} \sum_{i \neq j \neq k} \int \widehat{\tau}(U_j, U_k, \delta | X_i) \cdot \mathbb{1}\left\{\tau(U_j, U_k, \delta | X_i) \geq 0\right\} dQ(\delta) \cdot \omega(U_j, U_k, X_i) + \varphi_n,$$
$$\text{where} \quad \left|\varphi_n\right| = O_p\left(n^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0$$

Step 3

As before we denote

$$\mathbb{1}\left\{u - \delta \leq U_\ell \leq u + \delta\right\} \equiv \mathbb{I}(U_\ell, u, \delta).$$

Fix (u_1, u_2, x) and δ and define

$$\begin{aligned}
&\psi_\tau(u_1, u_2, x, Z_\ell, \delta; h_n) = \\
&\mu(u_2, \delta|x) \cdot \Big(\mathbb{I}(U_\ell, u_1, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_1, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big) \\
&+ \mu(u_1, \delta|x) \cdot \Big(\mathbb{I}(U_\ell, u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big) \\
&- \mu(u_1 \vee u_2, \delta|x) \cdot \Big(\mathbb{I}(U_\ell, u_1 \wedge u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_1 \wedge u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big) \\
&- \mu(u_1 \wedge u_2, \delta|x) \cdot \Big(\mathbb{I}(U_\ell, u_1 \vee u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n) - E\left[\mathbb{I}(U_\ell, u_1 \vee u_2, \delta) \cdot \mathcal{H}(X_\ell - x; h_n)\right]\Big)
\end{aligned}$$

Using (A-2) we have

$$\begin{aligned}
&\widehat{\tau}(u_1, u_2, \delta|x) - \tau(u_1, u_2, \delta|x) = \frac{1}{n}\sum_{i=1}^{n} \psi_\tau(u_1, u_2, x, Z_i, \delta; h_n) + \vartheta_\tau(u_1, u_2, x, \delta), \quad \text{where} \\
&\underset{\substack{(u_1, u_2, x) \in \mathcal{Z} \\ \delta \in \mathcal{S}_Q}}{sup} \left|\vartheta_\tau(u_1, u_2, x, \delta)\right| = O_p\left(\frac{\log(n)}{\sqrt{n \cdot h_n^q}}\right) + O\left(h_n^M\right) = O_p\left(n^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.
\end{aligned} \tag{A-14}$$

Where the last equality follows from our bandwidth convergence conditions. Let

$$\widetilde{S}^b(U_j, U_k, X_i, Z_\ell, \delta; h_n) = \psi_\tau(U_j, U_k, X_i, Z_\ell, \delta; h_n) \cdot \mathbb{1}\left\{\tau(U_j, U_k, \delta | X_i) \geq 0\right\} \cdot \omega(U_j, U_k, X_i).$$

Combining (A-13) and (A-14) we have

$$\begin{aligned}&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\widehat{\tau}(U_j,U_k,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_k,X_i)=\\&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\max\left\{\tau(U_j,U_k,\delta|X_i),0\right\}\cdot\omega(U_j,U_k,X_i)\\&+\frac{1}{(n)_3}\sum_{i\neq j\neq k}\frac{1}{n}\sum_{\ell=1}^{n}\widetilde{S}^b(U_j,U_k,X_i,Z_\ell,\delta;h_n)+\vartheta_n(\delta),\\&\text{where }\sup_{\delta\in\mathcal{S}_Q}\left|\vartheta_n(\delta)\right|=O_p\left(n^{-1/2-\epsilon}\right)\text{ for some }\epsilon>0.\end{aligned}\tag{A-15}$$

Note that $\sup_{\delta\in\mathcal{S}_Q}\left|\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widetilde{S}^b(U_j,U_k,X_i,Z_m,\delta;h_n)\right|=O_p\left(\frac{1}{h_n^q}\right)$ for $m=\{i,j,k\}$. Therefore,

$$\frac{1}{(n)_3}\sum_{i\neq j\neq k}\frac{1}{n}\sum_{\ell=1}^{n}\widetilde{S}^b(U_j,U_k,X_i,Z_\ell,\delta;h_n)=\frac{1}{n\cdot(n)_3}\sum_{i\neq j\neq k\neq\ell}\widetilde{S}^b(U_j,U_k,X_i,Z_\ell,\delta;h_n)+\overline{\varsigma}_n(\delta),$$

where $\sup_{\delta\in\mathcal{S}_Q}\left|\overline{\varsigma}_n(\delta)\right|=O_p\left(\frac{1}{n\cdot h_n^q}\right)=O_p\left(n^{-1/2-\epsilon}\right)$ for some $\epsilon>0$.

Let $S^b(Z_i,Z_j,Z_k,Z_\ell,\delta;h_n)=\frac{1}{4!}\sum_c\widetilde{S}^b(U_{m_1},U_{m_2},X_{m_3},Z_{m_4},\delta;h_n)$, where $\sum_c$ denotes the sum over all 4! different combinations $\{m_1,m_2,m_3,m_4\}$ of $\{i,j,k,\ell\}$. And define

$$V_n^b(\delta)=\binom{n}{4}^{-1}\sum_{i<j<k<\ell}S^b(Z_i,Z_j,Z_k,Z_\ell,\delta;h_n).$$

Then (A-15) becomes,

$$\begin{aligned}&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\widehat{\tau}(U_j,U_k,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_k,X_i)=\\&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\max\left\{\tau(U_j,U_k,\delta|X_i),0\right\}\cdot\omega(U_j,U_k,X_i)+\left(\frac{n-3}{n}\right)\cdot V_n^b(\delta)+\overline{\vartheta}_n(\delta),\\&\text{where }\sup_{\delta\in\mathcal{S}_Q}\left|\overline{\vartheta}_n(\delta)\right|=O_p\left(n^{-1/2-\epsilon}\right)\text{ for some }\epsilon>0.\end{aligned}\tag{A-16}$$

Note that $E\left[\widetilde{S}^b(U_j,U_k,X_i,Z_\ell,\delta;h_n)\middle|Z_i,Z_j,Z_k\right]=0$ by construction and by independence across observations in the sample. Assumption 3 states that the class of functions

$$\mathscr{M}_\delta=\left\{m:m(u_1,u_2,z)=\mathbb{1}\left\{\tau(u_1,u_2,\delta|x)\geq 0\right\}\ \text{ for some }\delta\in\mathcal{S}_Q\right\}$$

is Euclidean with respect to envelope 1. Combined with the manageability conditions in Assumption 3, this implies that the following class of functions is Euclidean,

$$\mathscr{F}^b_{\delta,h} = \Big\{ f : f(z_1,z_2,z_3,z_4) = S^b(z_1,z_2,z_3,z_4,\delta;h) \text{ for some } \delta \in \mathcal{S}_Q,\, h>0 \Big\}$$

for envelope $4\overline{K}$. From here, Sherman (1994, Corollary 4) and the Hoeffding decomposition of $V^b_n(\delta)$ yield,

$$\begin{aligned} V^b_n(\delta) &= \frac{4}{n}\sum_{i=1}^{n}\frac{1}{4!}\cdot(3!)\,E\left[\widetilde{S}^b(U_j,U_k,X_\ell,Z_i,\delta;h_n)\big|Z_i\right] + \upsilon^b_n(\delta) \\ &= \frac{1}{n}\sum_{i=1}^{n}E\left[\widetilde{S}^b(U_j,U_k,X_\ell,Z_i,\delta;h_n)\big|Z_i\right] + \upsilon^b_n(\delta), \end{aligned}$$

where $\underset{\delta\in\mathcal{S}_Q}{sup}\ \left|\upsilon^b_n(\delta)\right| = O_p\left(\frac{1}{n\cdot h^q_n}\right) = O_p\left(n^{-1/2-\epsilon}\right)$ for some $\epsilon>0$. Now let

$$\widetilde{S}^a(U_j,U_k,X_i,\delta) = \max\left\{\tau(U_j,U_k,\delta|X_i),0\right\}\cdot\omega(U_j,U_k,X_i),$$

and $S^a(Z_i,Z_j,Z_k,\delta) = \frac{1}{3!}\sum_c \widetilde{S}^a(U_{m_1},U_{m_2},X_{m_3},\delta)$, where $\sum_c$ denotes the sum over the 3! different combinations $\{m_1,m_2,m_3\}$ of $\{i,j,k\}$. Let $V^a_n(\delta) = \binom{n}{3}^{-1}\sum_{i<j<k}S^a(Z_i,Z_j,Z_k,\delta)$. Then

$$\frac{1}{(n)_3}\sum_{i\neq j\neq k}\max\left\{\tau(U_j,U_k,\delta|X_i),0\right\}\cdot\omega(U_j,U_k,X_i) = V^a_n(\delta) + \varrho_n(\delta),$$

where $\underset{\delta\in\mathcal{S}_Q}{sup}\ |\varrho_n(\delta)| = O_p\left(\frac{1}{n}\right)$. As we did above, denote

$$E\left[\max\left\{\tau\left(U_j,U_k,\delta|X_i\right),0\right\}\cdot\omega(U_j,U_k,X_i)\right] = \mathcal{T}(\delta).$$

Our manageability assumptions and a Hoeffding decomposition argument now yield

$$\begin{aligned}V_n^a(\delta) &= \mathcal{T}(\delta) + \frac{3}{n}\sum_{i=1}^{n}\Bigg\{\frac{2}{3!}\cdot\Big(E\left[\widetilde{S}^a(U_j,U_k,X_i,\delta)\big|X_i\right] - \mathcal{T}(\delta)\Big) \\ &\qquad + \frac{4}{3!}\cdot\Big(E\left[\widetilde{S}^a(U_i,U_j,X_k,\delta)\big|U_i\right] - \mathcal{T}(\delta)\Big)\Bigg\} + \nu_n^a(\delta) \\ &= \mathcal{T}(\delta) + \frac{1}{n}\sum_{i=1}^{n}\Bigg\{\Big(E\left[\widetilde{S}^a(U_j,U_k,X_i,\delta)\big|X_i\right] - \mathcal{T}(\delta)\Big) \\ &\qquad + 2\cdot\Big(E\left[\widetilde{S}^a(U_i,U_j,X_k,\delta)\big|U_i\right] - \mathcal{T}(\delta)\Big)\Bigg\} + \nu_n^a(\delta)\end{aligned}$$

where $\underset{\delta\in\mathcal{S}_Q}{sup}\ |\upsilon_n^a(\delta)| = O_p\left(\frac{1}{n}\right)$. Let

$$\begin{aligned}\psi^a\left(Z_i,\delta\right) &= \Big(E\left[\widetilde{S}^a(U_j,U_k,X_i,\delta)\big|X_i\right] - \mathcal{T}(\delta)\Big) + 2\cdot\Big(E\left[\widetilde{S}^a(U_i,U_j,X_k,\delta)\big|U_i\right] - \mathcal{T}(\delta)\Big), \\ \psi^b\left(Z_i,\delta;h_n\right) &= E\left[\widetilde{S}^b(U_j,U_k,X_\ell,Z_i,\delta;h_n)\big|Z_i\right], \\ \psi\left(Z_i,\delta;h_n\right) &= \psi^a\left(Z_i,\delta\right) + \psi^b\left(Z_i,\delta;h_n\right).\end{aligned}$$

Note that $E\left[\psi\left(Z_i,\delta;h_n\right)\right] = 0$ for any δ. Combining our previous results we obtain

$$\begin{aligned}&\frac{1}{(n)_3}\sum_{i\neq j\neq k}\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\widehat{\tau}(U_j,U_k,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_k,X_i) \\ &= \mathcal{T}(\delta) + \frac{1}{n}\sum_{i=1}^{n}\psi(Z_i,\delta;h_n) + \nu_n(\delta), \quad where \quad \underset{\delta\in\mathcal{S}_Q}{sup}\left|\nu_n(\delta)\right| = O_p\left(n^{-1/2-\epsilon}\right) \text{ for some } \epsilon > 0\end{aligned}$$

We defined $\mathcal{T}^Q = \int\mathcal{T}(\delta)dQ(\delta)$. Let $\psi^Q(Z_i;h_n) = \int\psi(Z_i,\delta;h_n)dQ(\delta)$. Note that $E\left[\psi^Q(Z_i;h_n)\right] = 0$ since $E\left[\psi(Z_i,\delta;h_n)\right] = 0$ for all δ. This yields our final result,

$$\widehat{\mathcal{T}}^Q = \mathcal{T}^Q + \frac{1}{n}\sum_{i=1}^{n}\psi^Q(Z_i;h_n) + \nu_n, \quad where \quad \left|\nu_n\right| = O_p\left(n^{-1/2-\epsilon}\right) \text{ for some } \epsilon > 0.$$

This concludes the proof of Theorem 1. □

A.2 An estimator for $Var\left(\psi^Q(Z_i;h_n)\right) = \sigma^2_{Q,n}$ in the test statistic $\widehat{t}_n$

Here we describe an estimator for the variance used in the construction of our test-statistic $\widehat{t}_n$ in (4). Let

$$\begin{aligned}\widehat{\psi}_\tau(u_1,u_2,x,Z_\ell,\delta;h_n) = {}& \widehat{\mu}(u_2,\delta|x)\cdot\Big(\mathbb{I}(U_\ell,u_1,\delta)\cdot\mathcal{H}(X_\ell-x;h_n)-\widehat{\mu}(u_1,\delta|x)\Big)\\ &+\widehat{\mu}(u_1,\delta|x)\cdot\Big(\mathbb{I}(U_\ell,u_2,\delta)\cdot\mathcal{H}(X_\ell-x;h_n)-\widehat{\mu}(u_2,\delta|x)\Big)\\ &-\widehat{\mu}(u_1\vee u_2,\delta|x)\cdot\Big(\mathbb{I}(U_\ell,u_1\wedge u_2,\delta)\cdot\mathcal{H}(X_\ell-x;h_n)-\widehat{\mu}(u_1\wedge u_2,\delta|x)\Big)\\ &-\widehat{\mu}(u_1\wedge u_2,\delta|x)\cdot\Big(\mathbb{I}(U_\ell,u_1\vee u_2,\delta)\cdot\mathcal{H}(X_\ell-x;h_n)-\widehat{\mu}(u_1\vee u_2,\delta|x)\Big).\end{aligned}$$

And

$$\begin{aligned}\widehat{\psi}^a(Z_i,\delta) = {}& \frac{1}{(n-1)_2}\sum_{\substack{j\neq k\\(j,k)\neq i}}\Big(\widehat{\tau}(U_j,U_k,\delta|X_i)\cdot\mathbb{1}\left\{\widehat{\tau}(U_j,U_k,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_k,X_i)-\widehat{\mathcal{T}}(\delta)\Big)\\ &+\frac{2}{(n-1)_2}\sum_{\substack{j\neq k\\(j,k)\neq i}}\Big(\widehat{\tau}(U_i,U_j,\delta|X_k)\cdot\mathbb{1}\left\{\widehat{\tau}(U_i,U_j,\delta|X_k)\geq -b_n\right\}-\widehat{\mathcal{T}}(\delta)\Big),\\ \widehat{\psi}^b(Z_i,\delta;h_n) = {}& \frac{1}{(n-1)_3}\sum_{\substack{j\neq k\neq \ell\\(j,k,\ell)\neq i}}\widehat{\psi}_\tau(U_j,U_k,X_i,Z_\ell,\delta;h_n)\cdot\mathbb{1}\left\{\widehat{\tau}(U_j,U_k,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_k,X_i),\\ \widehat{\psi}(Z_i,\delta;h_n) = {}& \widehat{\psi}^a(Z_i,\delta)+\widehat{\psi}^b(Z_i,\delta;h_n),\\ \widehat{\psi}(Z_i;h_n) = {}& \int\widehat{\psi}(Z_i,\delta;h_n)dQ(\delta),\\ \widehat{\sigma}^2_{Q,n} = {}& \frac{1}{n}\sum_{i=1}^n\widehat{\psi}^2(Z_i;h_n).\end{aligned}$$

Under the conditions of Theorem 1 we have $\left|\widehat{\sigma}^2_{Q,n}-\sigma^2_Q\right|\xrightarrow{p}0$. Note that $\psi(Z_i,\delta;h_n)=\psi^b(Z_i,\delta;h_n)$ under the null hypothesis of affiliation and therefore under this null hypothesis we can use $\widehat{\psi}(Z_i,\delta;h_n)=\widehat{\psi}^b(Z_i,\delta;h_n)$.

A.3 Asymptotic properties of test (4)

Let us generalize our inferential setting. Suppose $\{(U_i,X_i)\ ,1\leq i\leq n\}$ is triangular array, with row-wise iid distribution $F_n\in\mathcal{F}$. For a generic element $F\in\mathcal{F}$ we will index all the relevant functionals by F. Accordingly denote

$$G_F(u,\delta|x),\quad \mu_F(u,\delta|x),\quad \tau_F(u_1,u_2,\delta|x),\quad R_F^Q(u_1,u_2|x),\quad \mathcal{T}_F^Q.$$

Fix (u_1,u_2,x), δ and $h>0$ and define

$$\begin{aligned}
&\psi_{\tau,F}(u_1,u_2,x,Z,\delta;h)=\\
&\mu_F(u_2,\delta|x)\cdot\Big(\mathbb{I}(U,u_1,\delta)\cdot\mathcal{H}(X-x;h)-E_F\left[\mathbb{I}(U,u_1,\delta)\cdot\mathcal{H}(X-x;h)\right]\Big)\\
&+\mu_F(u_1,\delta|x)\cdot\Big(\mathbb{I}(U,u_2,\delta)\cdot\mathcal{H}(X-x;h)-E_F\left[\mathbb{I}(U,u_2,\delta)\cdot\mathcal{H}(X-x;h)\right]\Big)\\
&-\mu_F(u_1\vee u_2,\delta|x)\cdot\Big(\mathbb{I}(U,u_1\wedge u_2,\delta)\cdot\mathcal{H}(X-x;h)-E_F\left[\mathbb{I}(U,u_1\wedge u_2,\delta)\cdot\mathcal{H}(X-x;h)\right]\Big)\\
&-\mu_F(u_1\wedge u_2,\delta|x)\cdot\Big(\mathbb{I}(U,u_1\vee u_2,\delta)\cdot\mathcal{H}(X-x;h)-E_F\left[\mathbb{I}(U,u_1\vee u_2,\delta)\cdot\mathcal{H}(X-x;h)\right]\Big).
\end{aligned}$$

In what follows let $(Z_1,Z_2,Z_3,Z_4)\sim F_Z\times F_Z\times F_Z\times F_Z$ and define

$$\begin{aligned}
\widetilde{S}_F^a(U_1,U_2,X,\delta)&=\max\left\{\tau_F(U_1,U_2,\delta|X),0\right\}\cdot\omega(U_j,U_k,X_i),\\
\widetilde{S}_F^b(U_1,U_2,X_3,Z_4,\delta;h)&=\psi_{\tau,F}(U_1,U_2,X_3,Z_4,\delta;h)\cdot\mathbb{1}\left\{\tau(U_1,U_2,\delta|X_3)\geq 0\right\}\cdot\omega(U_1,U_2,X_3).
\end{aligned}$$

And

$$\begin{aligned}
\psi_F^a\left(Z_1,\delta\right)&=\Big(E_F\big[\widetilde{S}_F^a(U_2,U_3,X_1,\delta)\big|X_1\big]-\mathcal{T}_F(\delta)\Big)+2\cdot\Big(E_F\big[\widetilde{S}_F^a(U_1,U_2,X_3,\delta)\big|U_1\big]-\mathcal{T}_F(\delta)\Big),\\
\psi_F^b\left(Z_1,\delta;h\right)&=E_F\left[\widetilde{S}_F^b(U_2,U_3,X_4,Z_1,\delta;h)\big|Z_1\right],\\
\psi_F\left(Z_1,\delta;h\right)&=\psi_F^a\left(Z_1,\delta\right)+\psi_F^b\left(Z_1,\delta;h\right),\\
\psi_F^Q\left(Z_1,h\right)&=\int\psi_F\left(Z_1,\delta;h\right)dQ(\delta),\\
\sigma_{Q,F}^2(h)&=Var_F\left(\psi_F^Q\left(Z_1,h\right)\right).
\end{aligned}$$

Suppose we strengthen Assumptions 1-3 in the following way.

Assumption A5. *Each $F\in\mathcal{F}$ has the same support, with $Supp(U)\equiv\mathcal{U}$, $Supp(X)\equiv\mathcal{X}$ and $Supp(Z)\equiv\mathcal{Z}$ where as before $Z\equiv(U,X)$. In addition,*

(i) The conditions in Assumptions 1-3 are satisfied for each $F\in\mathcal{F}$.

(ii) Let $H_F = F_U \times F_U \times F_X \times Q$ and $\mathcal{S} = \mathcal{U} \times \mathcal{U} \times \mathcal{X} \times \mathcal{S}_Q$, and define

$$\overline{\mathcal{F}}^Q = \Big\{F \in \mathcal{F}\text{: } \tau_F(u_1, u_2, \delta|x) < 0 \;\; for \;\; H_F - a.e \;\; (u_1, u_2, x, \delta) \in \mathcal{S}\Big\}.$$

That is, $\overline{\mathcal{F}}^Q$ is the collection of distributions in $\mathcal{F}$ that are consistent with strict conditional affiliation. Then for some $\Delta > 0$ and $b < \infty$,

$$\underset{\substack{F\in\mathcal{F}\setminus\overline{\mathcal{F}}^Q \\ h>0}}{sup} E_F\left[\frac{\left|\psi_F^Q(Z,h)\right|^{2+\Delta}}{\sigma_{Q,F}^{2+\Delta}(h)}\right] \leq b.$$

Part (i) of Assumption A5 is meant to ensure that the linear representation in Theorem 1 holds uniformly over $\mathcal{F}$. Its first implication is that

$$\underset{F\in\overline{\mathcal{F}}^Q}{sup}\left|\frac{\sqrt{n}\cdot\widehat{\mathcal{T}}^Q(F)}{\max\left\{\kappa_n, \sigma_{Q,F}(h_n)\right\}}\right| = o_p(1).$$

Part (ii) is sufficient for the Lindeberg condition to hold uniformly over the space of distributions that do not satisfy strict conditional affiliation,

$$\lim_{\lambda\to\infty} \underset{\substack{F\in\mathcal{F}\setminus\overline{\mathcal{F}}^Q \\ h>0}}{sup} E_F\left[\frac{\left|\psi_F^Q(Z,h)\right|^2}{\sigma_{Q,F}^2(h)}\cdot\mathbb{1}\left\{\frac{\left|\psi_F^Q(Z,h)\right|}{\sigma_{Q,F}(h)} > \lambda\right\}\right] = 0$$

(see Romano (2004)). Thus, parts (i) and (ii) of Assumption A5 and a CLT for triangular arrays imply that for any sequence $\{F_n\} \in \mathcal{F} \setminus \overline{\mathcal{F}}^Q$,

$$\frac{\sqrt{n}\cdot\widehat{\mathcal{T}}_{F_n}^Q}{\sigma_{Q,F}^2(h_n)} \xrightarrow{d} \mathcal{N}(0,1)$$

Next, part (i) of Assumption A5 and the nature of the functions involved in the construction of $\widehat{\sigma}_{Q,\mathcal{Z}}^2$ (see Appendix A.2) imply that the conditions for the Law of Large Numbers for triangular arrays described in Romano (2004) are satisfied and

$$\left|\widehat{\sigma}_{Q,F_n}^2 - \sigma_{Q,F_n}^2(h_n)\right| \xrightarrow{p} 0$$

for any $\{F_n\} \in \mathcal{F}$.

A.3.1 Asymptotic size

Let $\mathcal{F}_*$ denote the collection of distributions in $\mathcal{F}$ that satisfy conditional affiliation. From our previous results it follows that if Assumption A5 is satisfied, we have

$$\lim_{n\to\infty} P_{F_n}\left(\widehat{t}_n \geq z_{1-\alpha}\right) = 0 \quad \forall\ \{F_n\} \in \mathcal{F}_* \cap \overline{\mathcal{F}}^Q,$$
$$\lim_{n\to\infty} P_{F_n}\left(\widehat{t}_n \geq z_{1-\alpha}\right) \leq \alpha \quad \forall\ \{F_n\} \in \mathcal{F}_* \cap \mathcal{F} \setminus \overline{\mathcal{F}}^Q,$$

and from here we have

$$\begin{aligned} &\lim_{n\to\infty} \underset{F\in\mathcal{F}_*}{sup}\ P_F\left(\widehat{t}_n \geq z_{1-\alpha}\right) \leq \alpha, \quad \text{with} \\ &\lim_{n\to\infty} \underset{F\in\mathcal{F}_*}{sup}\ P_F\left(\widehat{t}_n \geq z_{1-\alpha}\right) = \alpha \quad \text{if } \mathcal{F}_* \cap \mathcal{F} \setminus \overline{\mathcal{F}}^Q \neq \emptyset \end{aligned} \tag{A-17}$$

$\mathcal{F}_* \cap \mathcal{F} \setminus \overline{\mathcal{F}}^Q \neq \emptyset$ implies the existence of distributions that satisfy conditional affiliation, but not strict affiliation over our testing range. (A-17) describes the asymptotic size properties of our test under Assumption A5.

A.3.2 Power properties

The linear representation in Theorem 1 facilitates the study of the power features of our test. We defined $\overline{\mathcal{F}}^Q$ as the subset of distributions in $\mathcal{F}$ that satisfy strict conditional affiliation. Therefore all distributions where the affiliation inequalities are violated must live in $\mathcal{F} \setminus \overline{\mathcal{F}}^Q$ (although some distributions that satisfy affiliation -but not strict affiliation- also live there) and so we can focus on sequences that belong there. Take any sequence $\{F_n\} \in \mathcal{F} \setminus \overline{\mathcal{F}}^Q$. By Assumption A5(ii), a CLT for triangular arrays yields (see Romano (2004, Lemma 1))

$$\lim_{n\to\infty} P_{F_n}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\psi^Q_{F_n}(Z_i,h_n)}{\sigma_{Q,F_n}(h_n)} > c\right) = 1 - \Phi(c)$$

for any c (with $\Phi(\cdot)$ denoting the Standard Normal distribution). If Assumption A5 holds, the asymptotic power of our procedure for any sequence $\{F_n\} \in \mathcal{F} \setminus \overline{\mathcal{F}}^Q$ will be given by

$$\lim_{n\to\infty} P_{F_n}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\psi^Q_{F_n}(Z_i,h_n)}{\sigma_{Q,F_n}(h_n)} \geq \frac{\max\{\kappa_n, \sigma_{Q,F_n}(h_n)\}}{\sigma_{Q,F_n}(h_n)} \cdot \left[z_{1-\alpha} - \frac{\sqrt{n}\cdot\mathcal{T}^Q_{F_n}}{\max\{\kappa_n,\sigma_{Q,F_n}(h_n)\}}\right]\right)$$

The asymptotic power of our test will depend on the limit of the sequences

$$s_{1,n}(F_n) \equiv \frac{\max\{\kappa_n, \sigma_{Q,F_n}(h_n)\}}{\sigma_{Q,F_n}(h_n)} \quad \text{and} \quad s_{2,n}(F_n) \equiv \frac{\sqrt{n}\cdot \mathcal{T}^Q_{F_n}}{\max\{\kappa_n, \sigma_{Q,F_n}(h_n)\}}.$$

To describe the relevant cases the following functional is key,

$$\Delta^Q_F = \int\int\int\int \mathbb{1}\{\tau_F(u_1,u_2,\delta|x)=0\}\, dQ(\delta)dF_U(u_1)dF_U(u_2)dF_X(x).$$

This is the probability that the affiliation inequalities are binding as equalities over our testing range. Inspecting each one of the terms that comprise the influence function $\psi^Q_F(Z_i,h)$ we can see that under Assumption A5 we have:

(i) If $\mathcal{T}^Q_{F_n} \to 0$ and $\Delta^Q_{F_n} \to 0$, then $\sigma^2_{Q,F_n}(h_n) \to 0$.

(ii) If $\mathcal{T}^Q_{F_n} \to 0$ but $\Delta^Q_{F_n} \nrightarrow 0$, then $\sigma^2_{Q,F_n}(h_n) \nrightarrow 0$.

(iii) If $\mathcal{T}^Q_{F_n} \nrightarrow 0$, then $\sigma^2_{Q,F_n}(h_n) \nrightarrow 0$.

Suppose $s_{1,n}(F_n) \to s_1$ and $s_{2,n}(F_n) \to s_2$. Note that we must have $s_1 \geq 1$ by construction. If Assumption A5 holds, the conditions in Romano (2004, Theorem 5) are satisfied and we can use this to show that

$$\lim_{n\to\infty} P_{F_n}\left(\widehat{t}_n > z_{1-\alpha}\right) = 1 - \Phi\left(s_1 \cdot (z_{1-\alpha} - s_2)\right).$$

Consider any sequences $\{F_n\}$ such that $s_{1,n}(F_n) \to s_1$ and $s_{2,n}(F_n) \to s_2$ (i.e, such that both $s_{1,n}(F_n)$ and $s_{w,n}(F_n)$ converge to a limit, possibly $+\infty$). We have:

(i) If $s_{2,n}(F_n) \to \infty$ then our test will have asymptotic power of 1. This includes all sequences such that $\mathcal{T}^Q_{F_n} = O\big(n^{-\alpha}\big)$ for $\alpha < 1/2$.

(ii) If $s_{2,n}(F_n) \to 0$ then the asymptotic power of our test will be at most α. This includes all sequences such that $\mathcal{T}^Q_{F_n} = O\big(n^{-\alpha}\big)$ for $\alpha > 1/2$.

(iii) Consider a sequence such that $\mathcal{T}^Q_{F_n} = t/\sqrt{n}$. The asymptotic power of our test for any such sequence will be at least α. It will be 1 if $\sigma_{Q,F_n}(h_n) \to 0$, which will occur if and only if $\Delta^Q_{F_n} \to 0$. If $\Delta^Q_{F_n} \nrightarrow 0$, then we will have $\sigma_{Q,F_n}(h_n) \to \sigma^* > 0$. In this case, $s_{1,n}(F_n) \to 1$ and the asymptotic power of our test will be $1 - \Phi\left(z_{1-\alpha} - \frac{t}{\sigma^*}\right) \geq \alpha$.

A.4 Proof of Theorem 3

Let

$$
\begin{aligned}
&\zeta_\tau(u_1,u_2,x,Z_i;h_n) = \\
&\mu(u_2|x)\cdot\Big(\mathbb{1}\{U_i=u_1\}\cdot\mathcal{H}(X_i-x;h_n) - E\left[\mathbb{1}\{U_i=u_1\}\cdot\mathcal{H}(X_i-x;h_n)\right]\Big)\\
&+\mu(u_1|x)\cdot\Big(\mathbb{1}\{U_i=u_2\}\cdot\mathcal{H}(X_i-x;h_n) - E\left[\mathbb{1}\{U_i=u_2\}\cdot\mathcal{H}(X_i-x;h_n)\right]\Big)\\
&-\mu(u_1\vee u_2|x)\cdot\Big(\mathbb{1}\{U_i=u_1\wedge u_2\}\cdot\mathcal{H}(X_i-x;h_n) - E\left[\mathbb{1}\{U_i=u_1\wedge u_2\}\cdot\mathcal{H}(X_i-x;h_n)\right]\Big)\\
&-\mu(u_1\wedge u_2|x)\cdot\Big(\mathbb{1}\{U_i=u_1\vee u_2\}\cdot\mathcal{H}(X_i-x;h_n) - E\left[\mathbb{1}\{U_i=u_1\vee u_2\}\cdot\mathcal{H}(X_i-x;h_n)\right]\Big).
\end{aligned}
$$

And let

$$
\begin{aligned}
\widetilde{q}_\tau(u_1,u_2,X_i,Z_j;h_n) &= \zeta_\tau(u_1,u_2,X_i,Z_j;h_n)\cdot\mathbb{1}\{\tau(u_1,u_2|X_i)\geq 0\}\cdot\omega(u_1,u_2,X_i),\\
q_\tau(u_1,u_2,Z_i,Z_j;h_n) &= \frac{\widetilde{q}_\tau(u_1,u_2,X_i,Z_j;h_n)+\widetilde{q}_\tau(u_1,u_2,X_j,Z_i;h_n)}{2},\\
q_\tau(Z_i,Z_j;h_n) &= \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} q_\tau(u_1,u_2,Z_i,Z_j;h_n),\\
V_{\tau,n} &= \binom{n}{2}^{-1}\sum_{i<j} q_\tau(Z_i,Z_j;h_n).
\end{aligned}
$$

Note that $E\left[\widetilde{q}_\tau(u_1,u_2,X_i,Z_j;h_n)\middle|X_i\right]=0$. Under the assumptions of Theorem 3, the same type of arguments in Steps 1 and 2 of the proof of Theorem 1 can be used to show that

$$
\widehat{\mathcal{T}} = \frac{1}{n}\sum_{i=1}^n\left[\sum_{(u_1,u_2)\in\overline{\mathcal{U}}}\max\{\tau(u_1,u_2|X_i),0\}\cdot\omega(u_1,u_2,X_i)\right] + V_{\tau,n}+\varepsilon_n,
$$

where $|\varepsilon_n| = O_p\left(n^{-1/2-\epsilon}\right)$ for some $\epsilon>0$. Letting

$$
\begin{aligned}
\mathcal{T}(u_1,u_2) &= E\left[\max\{\tau(u_1,u_2|X_i),0\}\cdot\omega(u_1,u_2,X_i)\right],\\
\lambda^a(u_1,u_2,X_i) &= \max\{\tau(u_1,u_2|X_i),0\}\cdot\omega(u_1,u_2,X_i)-\mathcal{T}(u_1,u_2),\\
\lambda^a(X_i) &= \sum_{(u_1,u_2)\in\overline{\mathcal{U}}}\lambda^a(u_1,u_2,X_i),
\end{aligned}
$$

and recalling that $\mathcal{T} \equiv \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} \mathcal{T}(u_1,u_2)$, the expression above can be re-written as

$$\widehat{\mathcal{T}} = \mathcal{T} + \frac{1}{n}\sum_{i=1}^{n} \lambda^a(X_i) + V_{\tau,n} + \varepsilon_n,$$

Let

$$\lambda_n^b(u_1,u_2,Z_i) = E\left[\widetilde{q}_\tau(u_1,u_2,X_j,Z_i;h_n)\big|Z_i\right],$$
$$\lambda_n^b(Z_i) = \sum_{(u_1,u_2)\in\overline{\mathcal{U}}} \lambda_n^b(u_1,u_2,Z_i)$$

Note that $E\left[\lambda_n^b(Z_i)\right]$. The Hoeffding decomposition of $V_{\tau,n}$ is given by

$$V_{\tau,n} = \frac{1}{n}\sum_{i=1}^{n} \lambda_n^b(Z_i) + O_p\left(\frac{1}{n\cdot h_n^q}\right).$$

Letting $\lambda_n(Z_i) = \lambda^a(X_i) + \lambda_n^b(Z_i)$ we obtain

$$\widehat{\mathcal{T}} = \mathcal{T} + \frac{1}{n}\sum_{i=1}^{n} \lambda_n(Z_i) + V_{\tau,n} + \xi_n,$$

where $\left|\xi_n\right| = O_p\left(n^{-1/2-\epsilon}\right)$ for some $\epsilon > 0$. This is the result in Theorem 3. $\square$

A.5 Estimators for the variances in $\widetilde{t}_n$, $\overline{t}_n$ and $\widehat{s}_n$

All our asymptotic results are constructive, leading to straightforward ways to construct consistent estimators of the variances of our test statistics using sample analogs of the influence functions. We describe them next.

A.5.1 An estimator for $Var\left(\varphi_n^Q(Z_i)\right) = \Omega_{Q,n}^2$ in test statistic $\widetilde{t}_n$

Let $\widehat{\psi}_\tau$ be as defined in Appendix A.2 and let

$$\widehat{\varphi}^a(Z_i,\delta) = \frac{2}{n-1}\sum_{j:j\neq i}\Big(\widehat{\tau}(U_j,U_i,\delta|X_i)\cdot \mathbb{1}\left\{\widehat{\tau}(U_j,U_i,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_i,X_i)-\widehat{\mathcal{V}}(\delta)\Big)$$

$$\widehat{\varphi}^b(Z_i,\delta;h_n) = \frac{1}{(n-1)_2}\sum_{\substack{j\neq k\\(j,k)\neq i}}\widehat{\psi}_\tau(U_j,U_i,X_i,Z_k,\delta;h_n)\cdot \mathbb{1}\left\{\widehat{\tau}(U_j,U_i,\delta|X_i)\geq -b_n\right\}\cdot\omega(U_j,U_i,X_i),$$

$$\widehat{\varphi}(Z_i,\delta;h_n) = \widehat{\varphi}^a(Z_i,\delta)+\widehat{\varphi}^b(Z_i,\delta;h_n),$$

$$\widehat{\varphi}(Z_i;h_n) = \int\widehat{\varphi}(Z_i,\delta;h_n)dQ(\delta).$$

We can estimate $\Omega_{Q,n}^2$ with

$$\widehat{\Omega}_{Q,n}^2 = \frac{1}{n}\sum_{i=1}^n\widehat{\varphi}^2(Z_i;h_n).$$

A.5.2 An estimator for $Var\left(\lambda_n(Z_i)\right) = \sigma_n^2$ in test statistic $\bar{t}_n$

Denote

$$\widehat{\lambda}^a(u_1,u_2,X_i) = \widehat{\tau}(u_1,u_2|X_i)\cdot \mathbb{1}\left\{\widehat{\tau}(u_1,u_2|X_i)\geq -b_n\right\}\cdot\omega(u_1,u_2,X_i)-\widehat{\mathcal{T}}(u_1,u_2),$$

$$\widehat{\lambda}^a(X_i) = \sum_{(u_1,u_2)\in\overline{\mathcal{U}}}\widehat{\lambda}^a(u_1,u_2,X_i)$$

Let

$$\begin{aligned}
\widehat{\zeta}_\tau(u_1,u_2,x,Z_i;h_n) = {}& \widehat{\mu}(u_2|x)\cdot\Big(\mathbb{1}\left\{U_i=u_1\right\}\cdot\mathcal{H}(X_i-x;h_n)-\widehat{\mu}(u_1|x)\Big)\\
&+\widehat{\mu}(u_1|x)\cdot\Big(\mathbb{1}\left\{U_i=u_2\right\}\cdot\mathcal{H}(X_i-x;h_n)-\widehat{\mu}(u_2|x)\Big)\\
&-\widehat{\mu}(u_1\vee u_2|x)\cdot\Big(\mathbb{1}\left\{U_i=u_1\wedge u_2\right\}\cdot\mathcal{H}(X_i-x;h_n)-\widehat{\mu}(u_1\wedge u_2|x)\Big)\\
&-\widehat{\mu}(u_1\wedge u_2|x)\cdot\Big(\mathbb{1}\left\{U_i=u_1\vee u_2\right\}\cdot\mathcal{H}(X_i-x;h_n)-\widehat{\mu}(u_1\vee u_2|x)\Big),
\end{aligned}$$

$$\widehat{\lambda}_n^n(u_1,u_2,Z_i) = \frac{1}{n-1}\sum_{j:j\neq i}\widehat{\zeta}_\tau(u_1,u_2,X_j,Z_i;h_n)\cdot \mathbb{1}\left\{\widehat{\tau}(u_1,u_2|X_j)\geq -b_n\right\}\cdot\omega(u_1,u_2,X_j),$$

$$\widehat{\lambda}_n^b(Z_i) = \sum_{(u_1,u_2)\in\overline{\mathcal{U}}}\widehat{\lambda}_n^n(u_1,u_2,Z_i),$$

$$\widehat{\lambda}_n(Z_i) = \widehat{\lambda}^a(Z_i)+\widehat{\lambda}_n^b(Z_i).$$

We can estimate σ_n^2 with

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\widehat{\lambda}_n^2(Z_i)$$

A.5.3 An estimator for $Var\left(\gamma_n(Z_i)\right) = \Sigma_n^2$ in test statistic $\widehat{s}_n$

Let

$$\widehat{\gamma}_{k,L}^a(X_i) = \widehat{\tau}_{k,L}(X_i)\cdot \mathbb{1}\left\{\widehat{\tau}_{k,L}(X_i) \geq -b_n\right\}\cdot \omega(X_i,k,L) - \widehat{\mathcal{Q}}_{k,L},$$

and

$$\begin{aligned}
&\widehat{\phi}_{k,L}(Z_i, x; h_n) = \\
&2\binom{L}{k}^{-2}\widehat{\mu}_{k,L}(x)\Big(\mathbb{1}\left\{A_i = k\right\}\mathbb{1}\left\{L_i = L\right\}\mathcal{H}(X_i - x; h_n) - \widehat{\mu}_{k,L}(x)\Big)\\
&-\binom{L}{k-1}^{-1}\binom{L}{k+1}^{-1}\times\Bigg\{\mu_{k+1,L}(x)\cdot\Big(\mathbb{1}\left\{A_i = k-1\right\}\mathbb{1}\left\{L_i = L\right\}\mathcal{H}(X_i - x; h_n) - \widehat{\mu}_{k-1,L}(x)\Big)\\
&+\mu_{k-1,L}(x)\cdot\Big(\mathbb{1}\left\{A_i = k+1\right\}\mathbb{1}\left\{L_i = L\right\}\mathcal{H}(X_i - x; h_n) - \widehat{\mu}_{k+1,L}(x)\Big)\Bigg\},\\
&\widehat{\gamma}_{k,L,n}^b(Z_i) = \frac{1}{n-1}\sum_{j:j\neq i}\widehat{\phi}_{k,L}(Z_i, X_j; h_n)\cdot\mathbb{1}\left\{\widehat{\tau}_{k,L}(X_j)\geq -b_n\right\}\cdot\omega(X_j,k,L),\\
&\widehat{\gamma}_n(Z_i) = \sum_{L=\underline{L}}^{\overline{L}}\sum_{k=1}^{L-1}\left(\widehat{\gamma}_{k,L}^a(X_i) + \widehat{\gamma}_{k,L,n}^b(Z_i)\right)
\end{aligned}$$

We can estimate Σ_n^2 as

$$\widehat{\Sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\widehat{\gamma}_n(Z_i)^2.$$

References

de Castro, L. (2010). Affiliation, equilibrium existence and revenue ranking of auctions. Northwestern University.

de Castro, L. and H. Paarsch (2010). Testing affiliation in private-values models of first-price auctions using grid distributions. *The Annals of Applied Statistics 4*, 2073–2098.

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

Hendricks, K., J. Pinkse, and R. Porter (2003). Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. *Review of Economic Studies 70*, 115 – 145.

Ichimura, H. and P. Todd (2007). Implementing nonparametric and semiparametric estimators. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, vol. 6B*, Chapter 74, pp. 5369–5468. Elsevier.

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

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

Li, T., H. Paarsch, and T. Hubbard (2007). Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. Vanderbilt University.

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

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

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

Lu, J. and I. Perrigne (2008). Estimating Risk Aversion From Ascending and Sealed-Bid Auctions: The Case of Timber Auction Data. *Journal of Applied Econometrics 23*, 871–896.

McAdams, D. (2007). Monotonicity in asymmetric first-price auctions with affiliation. *International Journal of Game Theory 35*, 427–453.

Milgrom, P. and R. Weber (1982). A theory of auctions and competitive bidding. *Econometrica 50*, 1089–1122.

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

Paarsch, H. (1992). Deciding between the common and private value paradigms in empirical models of auctions. *Journal of Econometrics 51*(1), 191 – 215.

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

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

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

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

Roosen, J. and D. Hennessy (2004). Testing for the monotone likelihood ratio assumption. *Journal of Business and Economic Statistics 22*, 358–366.

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

Silverman, B. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall/CRC.