Center for Economic Institutions

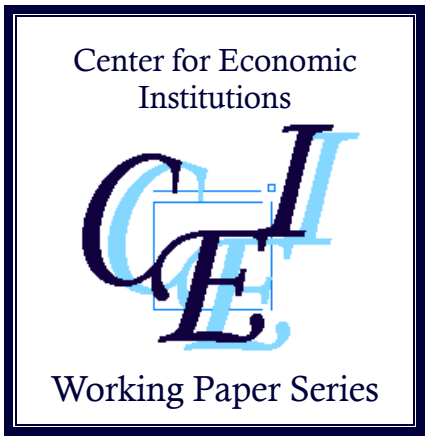
Working Paper Series

No. 2009-11

"Testing Affiliation in Private-values Models of First-price Auctions Using Grid Distributions"

Luciano I. de Castro and Harry J. Paarsch

December 2009



Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitachi, Tokyo, 186-8603 JAPAN
http://cei.ier.hit-u.ac.jp/English/index.html
Tel:+81-42-580-8405/Fax:+81-42-580-8333

28 November 2009

TESTING AFFILIATION IN PRIVATE-VALUES MODELS OF FIRST-PRICE AUCTIONS USING GRID DISTRIBUTIONS

BY LUCIANO I. DE CASTRO AND HARRY J. PAARSCH

University of Illinois at Urbana-Champaign
University of Melbourne

Within the private-values paradigm, we construct a tractable empirical model of equilibrium behaviour at first-price auctions when bidders' valuations are potentially dependent, but not necessarily affiliated. We develop a test of affiliation and apply our framework to data from low-price, sealed-bid auctions held by the Department of Transportation in the State of Michigan to procure road-resurfacing services: we do not reject the hypothesis of affiliation in cost signals.

1. Motivation and Introduction. During the past half century, economists have made considerable progress in understanding the theoretical structure of equilibrium strategic behaviour under market mechanisms, such as auctions, when the number of potential participants is relatively small; see Krishna [2010] for a comprehensive presentation and evaluation of progress.

One analytic device, commonly used to describe bidder motivation at single-object auctions, is a continuous random variable which represents individual-specific heterogeneity in valuations. The conceptual experiment involves each potential bidder's receiving a draw from a distribution of valuations. Conditional on his draw, a bidder is then assumed to act purposefully, maximizing either the expected profit or the expected utility of profit from winning the auction. Another frequently-made assumption is that the valuation draws of bidders are independent and that the bidders are *ex ante* symmetric—their draws being from the same distribution of valuations. This framework is often referred to as the *symmetric independent private-values paradigm* (symmetric IPVP). Under these assumptions, a researcher can then focus on a representative agent's decision rule when describing equilibrium behaviour.

At many real-world auctions, the latent valuations of potential bidders are probably dependent in some way. In auction theory, it has been assumed that dependence satisfies *affiliation*, a term coined by Milgrom and Weber [1982]. Affiliation is a condition concerning the joint distribution of signals. Often, affiliation is described using the intuition presented by Milgrom and Weber: "roughly, this [affiliation] means that a high value of one bidder's estimate makes high values of the others' estimates more likely." Thus described, affiliation seems like a relatively innocuous condition. In the case of continuous random variables, following the path blazed by Karlin [1968], some refer to affiliation as *multivariate total positivity of order two*, or MTP_2 for short. Essentially, under affiliation, with continuous random variables, the off-diagonal elements of the Hessian of the logarithm of the joint probability density function of signals are all non-negative; i.e., the joint probability density function is log-supermodular. Under joint normality of signals, affiliation requires that *all* the pair-wise covariances be weakly positive.

How is affiliation related to other forms of dependence? Consider two continuous random variables V_1 and V_2, having joint probability density function $f_{V_1,V_2}(v_1,v_2)$ as well as conditional probability density functions $f_{V_2|V_1}(v_2|v_1)$ and $f_{V_1|V_2}(v_1|v_2)$ and conditional cumulative distribution functions $F_{V_2|V_1}(v_2|v_1)$ and $F_{V_1|V_2}(v_1|v_2)$. Introduce $g(\cdot)$ and $h(\cdot)$, functions that are nondecreasing in their arguments. de Castro [2007] has noted that affiliation implies a) $[F_{V_2|V_1}(v_2|v_1)/f_{V_2|V_1}(v_2|v_1)]$ is decreasing in v_1 (and v_2 in the other case),

AMS 2000 subject classifications: Primary 62P20, 91B26; secondary 91B44.
Keywords and phrases: first-price, sealed-bid auctions; affiliation; MTP_2.

often referred to as a decreasing *inverse hazard rate*, which implies b) $\Pr(V_2 \leq v_2 | V_1 = v_1)$ is nonincreasing in v_1 (and v_2 in the other case), also referred to as *positive regression dependence*, which implies c) $\Pr(V_2 \leq v_2 | V_1 \leq v_1)$ is nonincreasing in v_1 (and v_2 in the other case), also referred to as *left-tail decreasing* in v_1 (v_2), which implies d) $\mathrm{cov}[g(V_1,V_2),h(V_1,V_2)]$ is positive, which implies that e) $\mathrm{cov}[g(V_1),h(V_2)]$ is positive, which implies f) $\mathrm{cov}(V_1,V_2)$ is positive. The important point to note is that affiliation is a much stronger form of dependence than positive covariance. In addition, de Castro [2007] has demonstrated that, within the set of all signal distributions, the set satisfying affiliation is small, both in the topological sense and in the measure-theoretic sense.

Affiliation delivers several predictions and results: first, under affiliation, the existence and uniqueness of a monotone pure-strategy equilibrium (MPSE) is guaranteed. Also, four commonly-studied auction formats—the oral, ascending-price (often referred to by economists as the *English*) and the second-price, sealed-bid (often referred to by economists as the *Vickrey*) as well as two first-price ones—can be ranked in terms of the revenues they can be expected to generate. Specifically, the expected revenues at English auctions are weakly greater than those at Vickrey auctions which are greater than those at first-price auctions—either the sealed-bid or the oral, descending-price (often referred to by economists as the *Dutch*) formats. Note, however, that when bidders are asymmetric, their valuation draws being from different marginal distributions, these rankings no longer apply. In fact, in general, very little can be said about the revenue-generating properties of the various auction formats and pricing rules under asymmetries.

Investigating equilibrium behaviour at auctions, empirically, when latent valuations are affiliated, has challenged researchers for some time. Laffont and Vuong [1996] showed that identification has been impossible to establish in many models when affiliation is present. In fact, Laffont and Vuong demonstrated that any model within the affiliated-values paradigm (AVP) is observationally equivalent to a model within the affiliated private-values paradigm (APVP). For this reason, virtually all empirical workers who have considered some form of dependence have worked within the APVP.

Only a few researchers have dealt explicitly with models within the APVP. In particular, Li et al. [2000] have demonstrated nonparametric identification within the conditional IPVP, a special case of the APVP, while Li et al. [2002] have demonstrated nonparametric identification within the APVP. One of the problems that Li et al. faced when implementing their approach is that nonparametric kernel-smoothed estimators are often slow to converge. In addition, Li et al. do not impose affiliation in their estimation strategy, so the first-order condition used in their two-step estimation strategy need not constitute an equilibrium. Hubbard et al. [2009] have sought to address some of these technical problems using semiparametric methods which sacrifice the full generality of the nonparametric approach *in lieu* of additional structure.

To date, except for Brendstrup and Paarsch [2007], no one has attempted to examine, empirically, models in which the private values are potentially dependent, but not necessarily affiliated. Incidentally, using data from sequential English auctions of two different objects, Brendstrup and Paarsch found weak evidence against affiliation in the valuation draws of two objects for the same bidder.

de Castro [2007] has noted that, within the private-values paradigm, affiliation is unnecessary to guarantee the existence and uniqueness of a MPSE. In fact, he has demonstrated existence and uniqueness of a MPSE under a weaker form of dependence, one where the inverse hazard rate is decreasing in the conditioned argument.

Because affiliation is unnecessary to guarantee existence and uniqueness of bidding strategies in models of first-price auctions with private values, expected revenue predictions based on empirical models in which affiliation is imposed are potentially biased. Knowing whether valuations are affiliated is central to ranking auction formats in terms of the expected revenues generated. In the absence of affiliation, the expected-revenue rankings delivered by the *linkage principle* of Milgrom and Weber [1982] need not hold: the expected-revenue rankings across auctions formats remain an empirical question. Thus, investigating the

empirical validity of affiliation appears both an important and a useful exercise.

In next section of this paper, we present a brief description of affiliation and its soldier—total positivity of order two (TP_2). Subsequently, following the theoretical work of de Castro [2007, 2008], who introduced the notion of the grid distribution, in section 3 we construct a tractable empirical model of equilibrium behaviour at first-price auctions when the private valuations of bidders are potentially dependent, but not necessarily affiliated.[1] In section 4, we develop a test of affiliation, which is based on grid distributions, rather than kernel-smoothing methods, thus avoiding the drawback encountered by Li et al. [2000, 2002], while in section 5, we conduct a small-scale Monte Carlo to investigate the numerical as well as small-sample properties of our proposed test. In section 6, we apply our methods in an empirical investigation of low-price, sealed-bid, procurement-contract auctions held by the Department of Transportation in the State of Michigan, and do not reject the null hypothesis of affiliation.

This information is potentially useful to a policy maker. The apparent high degree of estimated affiliation also explains why low levels of observed competition are often sufficient to maintain relatively low profit margins: strong affiliation is akin to fierce competition. Under strong affiliation, a potential winner knows that his nearest competitor probably has a valuation (cost) close to his, and this disciplines his bidding behaviour: he becomes more aggressive than under independence. We summarize and conclude in section 7, the final section of the paper.

2. Affiliation and TP_2. As mentioned above, affiliation is often described using an example with two random variables that can take on either a low or an high value. The two random variables are affiliated if high (low) values of each are more likely to occur than high and low or low and high values. A commonly-used graph of the four possible outcomes in a two-bidder auction game with two values is depicted in figure 1. The $(1,1)$ and $(2,2)$ points are more likely than the $(2,1)$ or $(1,2)$ points. Letting p_{ij} denote the probability of (i,j), affiliation in this example then reduces to TP_2—viz.,

$$p_{11}p_{22} \geq p_{12}p_{21}.$$

Put another way, TP_2 means that the determinant of the matrix

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

must be weakly positive. Independence in valuations obviously satisfies the lower bound on this determinental inequality. Note, too, that affiliation restricts distributions to a part of the simplex depicted in figure 2. In that figure, it is the region of the simplex that appears below a semi-circle rising from the line where $p_{11}+p_{22}$ equals one. In order to draw this figure, we needed to impose symmetry, so p_{12} and p_{21} are equal; thus, the intercept for p_{12} is one half. Conditions that are weaker than affiliation, but that also guarantee existence and uniqueness of equilibrium, are depicted in figure 2, too. In fact, in this simple example, the entire simplex satisfies these weaker conditions. In richer examples, however, it is a subset of the simplex, but one that contains the set of affiliated distributions. Thus, the assumption of affiliation could be important in determining the revenues a seller can expect from a particular auction format.

Slavković and Fienberg [2009] have discussed geometric representations of 2×2 distributions, like some of those considered here. Their representations are based on tetrahedrons, while ours reduce to triangles because of symmetry.

[1] The grid distributions discussed and used in this paper can also be modelled as contingency tables, which have been used extensively in applications in the social sciences; see Douglas et al. [1990] for the connections between contingency tables and positive dependence properties, including affiliation (TP_2), which is the focus of this paper.

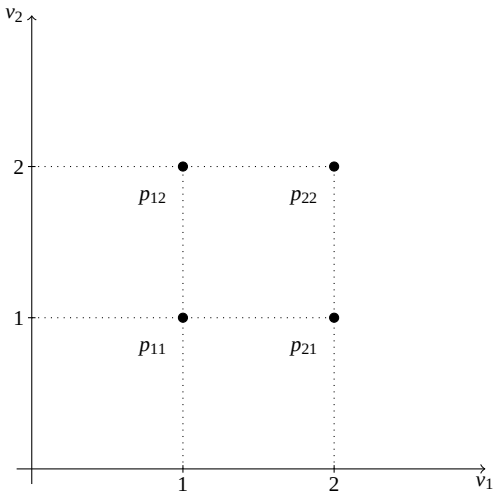


FIG 1. *Affiliation with Two Bidders and Two Values for Signals*

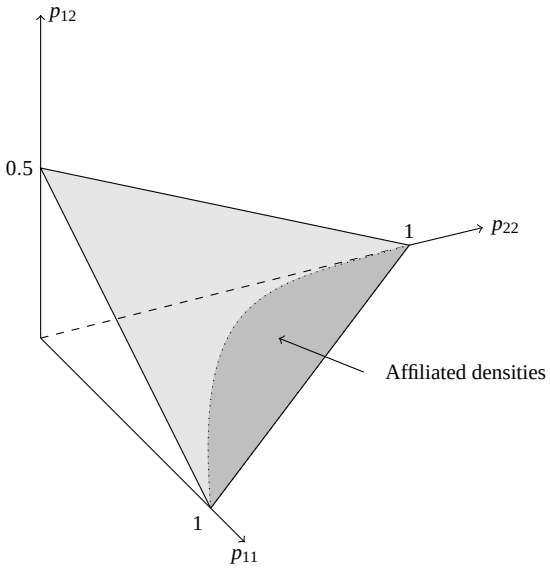


FIG 2. *Probability Set: Affiliation and Alternative*

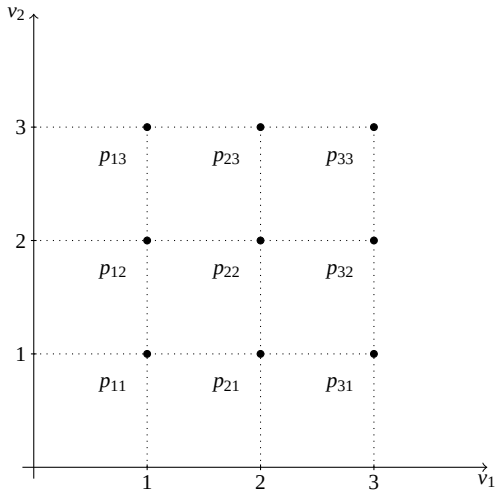


FIG 3. *Affiliation with Two Bidders and Three Values for Signals*

Another important point to note is that affiliation is a global restriction. To see the importance of this fact, introduce the valuation 3 for each bidder; five additional points then appear, as is depicted in figure 3. Affiliation requires that the probabilities at all collections of four points satisfy TP_2; i.e., the following additional six inequalities must hold:

$$p_{12}p_{23} \geq p_{13}p_{22},\ p_{22}p_{33} \geq p_{23}p_{32},\ p_{21}p_{32} \geq p_{22}p_{31},$$
$$p_{11}p_{33} \geq p_{13}p_{31},\ p_{12}p_{33} \geq p_{13}p_{32},\ \text{and } p_{21}p_{33} \geq p_{23}p_{31}.$$

Of course, symmetry would imply that p_{ij} equal p_{ji} for all i and j, so the joint mass function for two bidders and three valuations under symmetric affiliation can be written as the following matrix:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}$$

where the determinants of all (2×2) submatrices must be positive. Note, too, that all the points must also live on the simplex, so

$$0 \leq a, b, c, d, e, f < 1 \text{ and } a + b + c + 2d + 2e + 2f = 1.$$

How many inequalities are relevant? Let us represent the above matrix in the following tableau:

	1	2	3
1	a	d	e
2	d	b	f
3	e	f	c

where the row and column numbers will be used later to define TP_2 inequalities. There are $\binom{3}{2} \times \binom{3}{2}$ or nine possible combinations of four cells—i.e., nine inequalities. However, by symmetry, three are simply duplicates of others. The following tableau represents all of the inequalities:

	(1,2)	(1,3)	(2,3)
(1,2)	$ab \geq d^2$	$af \geq de$	$df \geq be$
(1,3)	$af \geq de$	$ac \geq e^2$	$dc \geq ef$
(2,3)	$df \geq be$	$dc \geq ef$	$bc \geq f^2$

where $(i, j) \times (\ell, m)$ means form a matrix with elements from rows i and j and columns ℓ and m of the first tableau. Observe that when the three inequalities highlighted in bold are satisfied, all others will be also satisfied. In fact, the inequality $(1,3) \times (1,2) : af \geq de$ derives from $(1,2) \times (1,2) : ab \geq d^2$ and $(2,3) \times (1,2) : df \geq be$. Finally, inequality $(2,3) \times (1,3) : dc \geq ef$ derives from the other two, previously established—viz., $(2,3) \times (1,2) : df \geq be$ and $(2,3) \times (2,3) : bc \geq f^2$. All other inequalities can be obtained from the adjacent ones in this fashion.

Adding values to the type spaces of bidders expands the number of determinental restrictions required to satisfy TP_2, thus restricting the space of distributions that can be entertained. Likewise, adding bidders to the game, particularly if the bidders are assumed symmetric, also restricts the space of distributions that can be entertained. For example, suppose a third bidder is added, one who is symmetric to the previous two. The probability mass function for triplets of values (v_1, v_2, v_3), where $v_n = 1, 2, 3$ and $n = 1, 2, 3$, can be represented as an array whose slices can then be represented by the following three matrices for bidders 1 and 2, indexed by the values of bidder 3:

$$\mathbf{P}_1 = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}, \mathbf{P}_2 = \begin{pmatrix} d & b & f \\ b & h & g \\ f & g & i \end{pmatrix}, \text{ and } \mathbf{P}_3 = \begin{pmatrix} e & f & c \\ f & g & i \\ c & i & j \end{pmatrix}.$$

In general, if the number of bidders is N and the number of values is k, then, without symmetry or affiliation, probability arrays have $(k^N - 1)$ unique elements. Also, de Castro [2008] has shown that symmetry reduces this to $\binom{k+N-1}{k-1}$ elements, while affiliation restricts where these $\binom{k+N-1}{k-1}$ probabilities can live on the simplex via the determinental inequalities required by TP_2. For it well-known that a function is MTP_2 (affiliated), if and only if, it is TP_2 in all relevant collections of four points. As an aside, in this three-by-three example, only nine constraints are relevant—viz.,

$$ab \geq d^2,\ bc \geq f^2,\ df \geq be,\ dh \geq b^2,\ hi \geq g^2,\ bg \geq fh,\ eg \geq f^2,\ gj \geq i^2, \text{ and } fi \geq cg.$$

If these hold, then the remainder are satisfied, too. Knowing the maximum number of binding constraints is relevant later in the paper when we discuss our test statistic.

Consider now the random N-vector V which equals $(V_1, \ldots, V_N)$, having joint density (mass) function f_V with realization v equal to $(v_1, \ldots, v_N)$. Affiliation can be formally defined as follows: for all v and v', the random variables V are said to be affiliated if

$$f_V(v \vee v') f_V(v \wedge v') \geq f_V(v) f_V(v')$$

where

$$(v \vee v') = [\max(v_1, v_1'), \max(v_2, v_2'), \ldots, \max(v_N, v_N')]$$

denotes the component-wise maxima of v and v', sometimes referred to as the *join*, while

$$(v \wedge v') = [\min(v_1, v_1'), \min(v_2, v_2'), \ldots, \min(v_N, v_N')]$$

denotes the component-wise minima, sometimes referred to as the *meet*.

3. Theoretical Model. We develop our theoretical model within the private-values paradigm, assuming away any interdependencies. We consider a set $\mathcal{N}$ of bidders $\{1, 2, \ldots, N\}$. Now, bidder n is assumed to draw V_n, his private valuation of the object for sale, from the closed interval $[\underline{v}, \overline{v}]$. We note that, without loss of generality, one can reparametrize the valuations from $[\underline{v}, \overline{v}]$ to $[0,1]$. Below, we do this. We collect the valuations in the vector v which equals $(v_1, \ldots, v_N)$ and denote this vector without the n^{th} element by v_{-n}. Here, we have used the now-standard convention that upper-case letters denote random variables, while lower-case ones denote their corresponding realizations. Note, too, that V lives in $[0,1]^N$.

We assume that the values are distributed according to the probability density function $f_V : [0,1]^N \to \mathbb{R}_+$ which is symmetric; i.e., for the permutation $\varphi : \{1, \ldots, N\} \to \{1, \ldots, N\}$, we have $f_V(v_1, \ldots, v_N)$ equals $f_V(v_{\varphi(1)}, \ldots, v_{\varphi(N)})$. Letting $f_n(v_n)$ denote the marginal probability density function of V_n, we note that it equals $\int_0^1 \cdots \int_0^1 f_V(v_{-n}, v_n)\, \mathrm{d}v_{-n}$. (Below, we constrain ourselves to the case where $f_n(\cdot)$ is the same for all n, but this is unnecessary and done only because, when we come to apply the method, we do have not enough information to estimate the case with varying f_ns.) Our main interest is the case when f_V is *not* the product of its marginals—the case where the types are dependent. We denote the conditional density of V_{-n} given v_n by

$$f_{V_{-n}|V_n}(v_{-n}|v_n) = \frac{f_V(v_{-n}, v_n)}{f_n(v_n)}.$$

Finally, we denote the largest order statistic of V_{-n} given v_n by Z_n and its probability density and cumulative distribution functions by $f(z_n|v_n)$ and $F(z_n|v_n)$, respectively.

We assume that bidders are risk neutral and abstract from a reserve price. Given his value v_n, bidder n tenders a bid $s_n \in \mathbb{R}_+$. If his tender is the highest, then bidder n wins the object and pays what he bid. A pure strategy is a function $\sigma : [0,1] \to \mathbb{R}_+$ which specifies the bid $\sigma(v_n)$ for each value v_n. The interim pay-off of bidder n, who bid s_n when his opponents follow $\sigma : [0,1] \to \mathbb{R}_+$, is

$$\Pi(v_n, s_n, \sigma) = (v_n - s_n) \int_{\underline{v}}^{\sigma^{-1}(s_n)} f(z_n|v_n)\, \mathrm{d}z_n = (v_n - s_n) F[\sigma^{-1}(s_n)|v_n].$$

We focus on symmetric, increasing pure-strategy equilibria (PSE) which are defined by $\sigma : [0,1] \to \mathbb{R}_+$ such that

$$\Pi[v_n, \sigma(v_n), \sigma_{-n}] \geq \Pi(v_n, s, \sigma_{-n}) \;\forall\; s,\; v_n.$$

As mentioned above, in most theoretical models of auctions that admit dependence in valuation draws, researchers have assumed that f_V satisfies affiliation. We do not restrict ourselves to f_Vs that satisfy affiliation. We assume only that f_V belongs to a set of distributions $\mathcal{P}$ which guarantees the existence and uniqueness of a MPSE. This set $\mathcal{P}$ was fully characterized by de Castro [2008] in the particular case of grid distributions, which are considered in our Assumption 4.1, below.

Let $\mathcal{C}$ denote the set of continuous density functions $f_V : [0,1]^N \to \mathbb{R}_+$ and let $\mathcal{A}$ denote the set of affiliated probability functions. For convenience and consistency with the notation used in later sections, we include in $\mathcal{A}$ the set of all affiliated probability functions, not just the continuous ones. Endow $\mathcal{C}$ with the topology of the uniform convergence—i.e., the topology defined by the norm of the supremum

$$\|f_V\| = \sup_{v \in [0,1]^N} |f_V(v)|.$$

Let $\mathcal{D}$ be the set of probability functions $f_V : [0,1]^N \to \mathbb{R}_+$ and assume there is a measure μ over it.

We now introduce a transformation $\mathbb{T}^k : \mathcal{D} \to \mathcal{D}$ which is the workhorse of our method. To define $\mathbb{T}^k$, let $\mathbb{I} : [0,1] \to \{1,2,\ldots,k\}$ denote the function that associates to $v \in [0,1]$ the ceiling $\lceil kv \rceil$—viz., the smallest integer at least as large as kv. Thus, for each $v \in [0,1]$, we have $v \in \left(\frac{\mathbb{I}(v)-1}{k}, \frac{\mathbb{I}(v)}{k}\right]$. Similarly, let $\mathbb{S}(v)$ denote the "square" (hypercube) $\prod_{n=1}^N \left(\frac{\mathbb{I}(v_n)-1}{k}, \frac{\mathbb{I}(v_n)}{k}\right]$ where v collects $(v_1, v_2, \ldots, v_N) \in [0,1]^N$. From this, we define $\mathbb{T}^k : \mathcal{D} \to \mathcal{D}$ as the transformation that associates to each $f_V \in \mathcal{D}$ the probability density function $\mathbb{T}^k(f_V)$ given by:

$$\mathbb{T}^k(f_V)(v) = k^N \int_{\mathbb{S}(v)} f_V(u)\, \mathrm{d}u.$$

Observe that $\mathbb{T}^k(f_V)$ is constant over each square $\prod_{n=1}^N \left(\frac{m_n-1}{k}, \frac{m_n}{k}\right]$, for all combinations of $m_n \in \{1,\ldots,k\}$. The term k^N above derives from the fact that each square $\prod_{n=1}^N \left(\frac{m_n-1}{k}, \frac{m_n}{k}\right]$ has volume $(1/k^N)$. Note that for all probability density functions $f_V \in \mathcal{D}$, $\mathbb{T}^1(f_V)(v)$ equals one for all $v \in [0,1]^N$; i.e., $\mathbb{T}^1(f_V)$ is the uniform distribution on $[0,1]^N$.

We now need to introduce a compact notation to represent arrays of dimension $\overbrace{k \times k \times \cdots \times k}^{N \text{ times}}$. We denote by $\mathcal{M}^{k^N}$ the set of arrays and by $[\mathbf{P}]$ an array in that set. When there are but two bidders, an array is obviously just a matrix. In the application of this model to field data, which we describe in section 6, N is three. The $(i_1, i_2, \ldots, i_N)^{\text{th}}$ element of an array is denoted $[\mathbf{P}](i_1, i_2, \ldots, i_N)$, or $[\mathbf{P}](i)$ for short, where i denotes the vector $(i_1, i_2, \ldots, i_N)$. Now, $\mathbb{I}(v) = i$ if $v \in \left(\frac{i-1}{k}, \frac{i}{k}\right]$. Thus, for $k \in \mathbb{N}$, we define the finite-dimensional subspace $\mathcal{D}^k \subset \mathcal{D}$ as

$$\mathcal{D}^k = \left\{ f_V \in \mathcal{D} : \exists\, [\mathbf{P}] \in \mathcal{M}^{k^N},\ f_V(v) = [\mathbf{P}][\mathbb{I}(v_1), \ldots, \mathbb{I}(v_N)] \right\}.$$

Observe that $\mathcal{D}^k$ is a finite-dimensional set. In fact, when N is two, a probability density function $f_V \in \mathcal{D}^k$ can be described by a $(k \times k)$ matrix $\mathbf{P}$ as follows:

$$f_V(v_1, v_2) = [\mathbf{P}](i, j) \text{ if } (v_1, v_2) \in \left(\frac{i-1}{k}, \frac{i}{k}\right] \times \left(\frac{j-1}{k}, \frac{j}{k}\right] \tag{3.1}$$

for $i, j \in \{1, 2, \ldots, k\}$. The definition of f_V at the zero measure set of points $\{(v_1, v_2) = (\frac{i}{k}, \frac{j}{k}) : i = 0 \text{ or } j = 0\}$ is arbitrary.

Note, too, that the width of the cells can be allowed to vary. For example, one might be 0.3 wide, while the next one can be 0.2 wide, the third 0.1 wide, the next 0.25, and the last 0.15. In fact, the transformation can be defined in terms of rectangles, instead of squares as above. To illustrate this, consider again the symmetric case and introduce figure 4. Let $0 = r_0 < r_1 < r_2 < \ldots < r_{k-1} < r_k = 1$ be an arbitrary partitioning of the interval $[0,1]$. Now, define $\mathbb{I} : [0,1] \to \{1,2,\ldots,k\}$ by $\mathbb{I}(v) = j$ if and only if $v \in (r_{j-1}, r_j]$. Define $\mathbb{B}(v)$ as the rectangle (box) where v collects $(v_1, \ldots, v_N) \in [0,1]^N$ lies. Thus, $\mathbb{B}(v) \equiv \prod_{n=1}^N (r_{\mathbb{I}(v_n)-1}, r_{\mathbb{I}(v_n)}]$. Now, define

$$\mathbb{T}^k_{\mathbb{B}}(f_V^0)(v) = \frac{\int_{\mathbb{B}(v)} f_V^0(u)\, \mathrm{d}u}{\int_{\mathbb{B}(v)} \mathrm{d}u}.$$

The following theorem was proven by de Castro [2007]:

THEOREM 3.1. *Let f_V^0 be a symmetric and continuous probability density function. f_V^0 is affiliated if and only if for all k, $\mathbb{T}^k_{\mathbb{B}}(f_V^0)$ is also affiliated.*

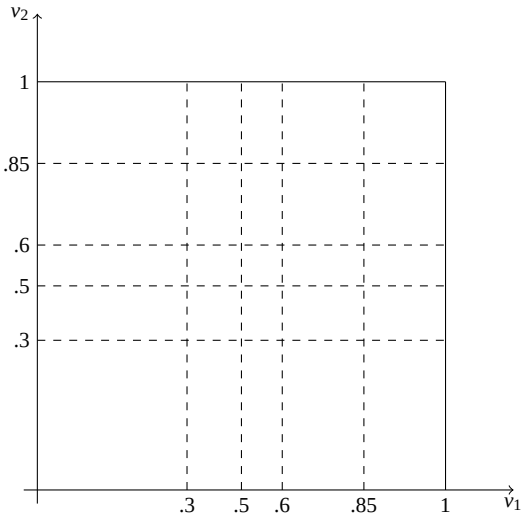


FIG 4. *Symmetric Non-Equi-Spaced Grid*

In our notation,

$$f_V^0 \in \mathcal{A} \Leftrightarrow \mathbb{T}_{\mathbb{B}}^k(f_V^0) \in \mathcal{A},\ \forall\, k \in \mathbb{N}$$

or

$$\mathcal{A} = \cap_{k\in\mathbb{N}} \mathbb{T}_{\mathbb{B}}^{-k}\left(\mathcal{A} \cap \mathcal{D}^k\right).$$

Why is this important? Well, in many applications, the set of hypercubes defined by a large k will have many empty cells, which causes problems in both estimation and inference. Thus, one may want to subdivide the space of valuations irregularly, but symmetrically, as illustrated in figure 4 when N is two.

4. Test of Affiliation. The key result from de Castro [2007] that allows us to develop our test of symmetric affiliation is the following: if the true probability density function f_V^0 exhibits affiliation, then $\mathbb{T}_{\mathbb{B}}^k(f_V^0)$, a discretized version of it, will too. (See Theorem 3.1, above.) To the extent that the grid distribution $\mathbb{T}_{\mathbb{B}}^k(f_V^0)$ can be consistently estimated from sample data, one can then test whether the estimated grid distribution exhibits affiliation. Of course, sampling error will exist, but presumably one can evaluate its relative importance using first-order asymptotic methods.

Consider a sequence of T auctions indexed $t = 1,\ldots,T$ at which N bidders participated by submitting the NT bids $\{\{s_{nt}\}_{n=1}^N\}_{t=1}^T$. We note that affiliation is preserved under a monotonic transformation, so examining a discretization of $g_S^0(s)$, the true probability density function of bids under the hypothesis of expected-profit maximizing equilibrium behaviour, is the same as examining $f_V^0(v)$. Of course, neither f_V^0 nor g_S^0 is known. One can, however, construct an estimate of $\mathbb{T}_{\mathbb{B}}^k(g_S^0)$ on the interval $[0,1]^N$ by first transforming the observed bids according to

$$u_{nt} = \frac{s_{nt} - \underline{s}}{\overline{s} - \underline{s}} \quad n = 1,\ldots,N \text{ and } t = 1,\ldots,T$$

where $\underline{s}$ is the smallest observed bid and $\overline{s}$ is the largest observed bid, and then by breaking-up this hypercube

into $L(=k^N)$ cells and counting the number of times that a particular N-tuple falls in that cell.[2] Now, the random vector Y, which represents the number of outcomes that fall in each of the cells and equals the vector $(Y_1, Y_2, \ldots, Y_L)^\top$, follows a multinomial distribution having the joint probability mass function

$$g_Y(y|\pi) = \frac{T!}{y_1! \cdots y_L!} \prod_{\ell=1}^{L} \pi_\ell^{y_\ell}$$

where π_ℓ equals $\Pr(Y_\ell = y_\ell)$, with $y_\ell = 0, 1, \ldots, T$, while π collects $(\pi_1, \ldots, \pi_L)$ and lives on the simplex—viz., the set

$$\mathcal{S}_L = \{\pi | \pi \geq \mathbf{0}_L,\ \iota_L^\top \pi = 1\}$$

with ι_L being an $(L \times 1)$ vector of ones. Note, too, that $\iota^\top y$ equals T, the number of observations.

For $\ell = 1, \ldots, L$, the unconstrained maximum-likelihood estimates of the π_ℓs are the (y_ℓ/T)s. To test for affiliation, maximize the following logarithm of the likelihood function (minus a constant):

$$\mathcal{L}(\pi) = y^\top \log(\pi)$$

subject to

1) the vector π lies in the simplex $\mathcal{S}_L$;
2) all of the determinental inequalities required for TP_2 hold.

Then compare this value of $\mathcal{L}$ with the unconstrained one.

While the determinental constraints required for TP_2 are convex sets of the parameters when the submatrices are symmetric, they are not for general submatrices. However, by taking logarithms of both sides of any general determinental inequality

$$ab \geq cd,$$

one can convert this into a linear inequality, which does give rise to convex constraint sets, albeit in variables that are logarithms of the original variables. To wit,

$$\log a + \log b - \log c - \log d \geq 0$$

defines a convex set (in the transformed variables $\log a, \ldots, \log d$). Of course, the adding-up constraint for the simplex must be finessed—e.g., by considering the following:

$$\exp(\log a) + \exp(\log b) + \exp(\log c) + \exp(\log d) + \ldots \leq 1,$$

which gives rise to a convex set. Thus, the problem is almost a linear programme.

For known N and fixed k, the specific steps involved in implementing the test in this problem are the following. First, form the grid distribution of the joint density as the unknown array $[\mathbf{P}]$. Letting $[\mathbf{E}]$ denote the array of counts for the grid distribution, the logarithm of the likelihood function for this multinomial process is

$$\sum_i [\mathbf{E}](i) \log\{[\mathbf{P}](i)\}. \tag{4.1}$$

[2]We know that the support of g_S^0 is strictly positive at $\sigma^0(\overline{v})$, the true upper bound of support of bids, and we assume that f_V^0 is strictly positive at $\underline{v}$, so g_S^0 is strictly positive at $\sigma^0(\underline{v})$, the true lower bound of support of bids. Consequently, the sample estimators of the lower and upper bounds of support of S converge at rate T, which is faster than the rate of convergence of sample averages—rate $\sqrt{T}$. Hence, when using sample averages in our estimation, we can ignore this first-stage, pre-estimation error—at least under first-order asymptotic analysis.

Now, the following inequalities must be met:

$$\log\{[\mathbf{P}](i)\} \leq 0 \text{ and } \sum_i \exp\left(\log\{[\mathbf{P}](i)\}\right) \leq 1, \tag{4.2}$$

while symmetry requires the following linear restrictions:

$$[\mathbf{P}](i) = [\mathbf{P}][\varphi(i)] \tag{4.3}$$

where $\varphi(\cdot)$ is any permutation, and affiliation requires the following determinental inequalities:

$$\log\left\{\frac{[\mathbf{P}](i \vee i')[\mathbf{P}](i \wedge i')}{[\mathbf{P}](i)[\mathbf{P}](i')}\right\} \geq 0 \tag{4.4}$$

hold. A test of affiliation, within a symmetric environment, involves comparing the maximum of equation (4.1), subject to the constraints in (4.2) and (4.3), with the maximum of equation (4.1), subject to the constraints in (4.2), (4.3), and (4.4).

Our test of symmetric affiliation is based on the difference between the maximum of the logarithm of the likelihood function $\mathcal{L}([\hat{\mathbf{P}}])$ and the maximum of the logarithm of the likelihood function under symmetric affiliation $\mathcal{L}([\tilde{\mathbf{P}}])$. Obviously, the sampling theory associated with the difference in these two values of the objective function $\mathcal{L}$ is not straightforward because not all of the inequality constraints required by MTP_2 may hold and, from sample to sample, the ones that do hold can change, but we shall suggest several strategies to deal with this, below.

Experience gleaned from other models with a related structure—e.g., Wolak [1987, 1989a,b, 1991] as well as Bartolucci and Forcina [2000], who investigated MTP_2 in binary models—suggests that the likelihood ratio (LR) statistic

$$2[\mathcal{L}([\hat{\mathbf{P}}]) - \mathcal{L}([\tilde{\mathbf{P}}])] \tag{4.5}$$

is not distributed according to a standard χ^2 random variable.

Introducing $\text{vec}[\mathbf{P}]$ as a short-hand notation, for the L-vector created from the array $[\mathbf{P}]$, our constrained-optimization problem can be summarized in a notation similar to that of Wolak (1989b) as:

$$\max_{\text{vec}[\mathbf{P}]} y^\top \log(\text{vec}[\mathbf{P}]) \quad \text{subject to} \quad \mathbf{h}(\text{vec}[\mathbf{P}]) \geq \mathbf{0}_J$$

where $\mathbf{h} : \mathbb{R}^L \to \mathbb{R}^J$ is the function representing all J relevant constraints where $J \leq L$ and L is the total number of variables under the alternative hypothesis. (Here, for notational parsimony, we have ignored the adding-up condition, which is implicit.)

Consider $N_\delta(\text{vec}[\mathbf{P}^0])$, a neighbourhood of the true value $\text{vec}[\mathbf{P}^0]$. Denote by $\mathbf{H}(\text{vec}[\mathbf{P}^0])$ the matrix of partial derivatives whose (i,j)-element is $\frac{\partial h_i(\text{vec}[\mathbf{P}])}{\partial \text{vec}[\mathbf{P}]_j}$. Now, let us define the set $\mathcal{B} = \{\text{vec}[\mathbf{P}] : \mathbf{H}(\text{vec}[\mathbf{P}^0])\text{vec}[\mathbf{P}] \geq \mathbf{0},\ \text{vec}[\mathbf{P}] \in \mathbb{R}^L\}$. Denote by $I(\text{vec}[\mathbf{P}^0])$ Fisher's information matrix which is defined by

$$\lim_{T\to\infty} T^{-1}\mathcal{E}_{[\mathbf{P}^0]}\left[-\frac{\partial^2 \mathcal{L}(\text{vec}[\mathbf{P}])}{\partial \text{vec}[\mathbf{P}]\partial \text{vec}[\mathbf{P}]^\top}\right]$$

evaluated at $\text{vec}[\mathbf{P}^0]$. Finally, denote by

$$\Pi^0 = \mathbf{H}(\text{vec}[\mathbf{P}^0])\, I(\text{vec}[\mathbf{P}^0])^{-1}\mathbf{H}(\text{vec}[\mathbf{P}^0])^\top$$

the variance-covariance matrix of $\mathbf{h}(\text{vec}[\hat{\mathbf{P}}])$ and by $\omega(j, J-j, \Pi^0)$, the probability that j constraints bind, that $(J-j)$ constraints are strictly satisfied; i.e., they are nonbinding. We have the following:

THEOREM 4.1. *Consider the local hypothesis testing problem*

$$\begin{aligned} \mathrm{H}_0 &: \mathbf{h}(\mathrm{vec}[\mathbf{P}]) \geq \mathbf{0}_J \ \mathrm{vec}[\mathbf{P}] \in N_\delta(\mathrm{vec}[\mathbf{P}^0]) \\ \mathrm{H}_1 &: \textit{not } \mathrm{H}_0. \end{aligned}$$

The asymptotic distribution of the likelihood-ratio statistic satisfies the following property:

$$\sup_{b\in\mathcal{B}} \Pr_{[\mathbf{P}^0],I(\mathrm{vec}[\mathbf{P}^0])^{-1}}(D \geq c) = \Pr_{[\mathbf{P}^0]}(D \geq c) = \sum_{j=0}^{J} \Pr(W_j \geq c)\omega(j, J-j, \Pi^0).$$

where D is the asymptotic value of the test statistic, while W_j is an independent χ^2 random variable having j degrees of freedom.

PROOF. It is sufficient and straightforward to verify that the assumptions of Theorem 4.2 in Wolak [1989b] are satisfied. □

Because this statistic depends on the unknown population grid distribution $[\mathbf{P}^0]$, the statistic is not pivotal. Kodde and Palm [1986] have provided lower and upper bounds for this test statistic for tests of various sizes and different numbers of maximal constraints.

According to Wolak [1989b], the best way to evaluate the weights is using Monte Carlo simulation. Wolak also offered lower and upper bounds for the probabilities above (see his equations 19 and 20, p.215); these bounds are based on Kodde and Palm [1986]. An alternative strategy would be to adapt the bootstrap methods of Bugni [2008] to get the appropriate p-values of the test statistic. Yet a third strategy would be to adapt the subsampling methods described in Politis et al. [1999] as was done by Romano and Shaikh [2008].

4.1. *Some Comparisons with Other Nonparametric Methods.* It should be noted, too, that our proposed estimation strategy involves nothing more than estimating an histogram using a special class of grids. Scott [1992, p. xi] has argued that the classical histogram "remains the most widely applied and most intuitive nonparametric estimator." In other words, the procedure proposed here is not based on any unfamiliar concepts. Of course, there are more statistically efficient methods, but they also have limitations, as Scott [1992] has discussed. Also, although the rate of convergence of histogram estimation is slow, it is still reasonable; see Scott [1992, Theorem 3.5, p. 82].

Note, too, the similarities between grid-distribution and kernel-smoothed estimators. Kernel-smoothed density estimators are well-behaved and have good rates of convergence when the probability density functions to be estimated are continuously differentiable $\mathbb{C}^1$.[3] The set $\mathbb{C}^1$ is dense in the set of all probability density functions. Similarly, grid distribution estimators are well behaved for probability density functions in $\mathcal{D}^\infty = \cup_{k=1}^\infty \mathcal{D}^k$, which is also a dense set in the set of all probability density functions.[4] While $\mathbb{C}^1$ probability density functions form a familiar and well-known class probability density functions, the probability density functions in $\mathcal{D}^k$ are also familiar because they are just (a special class of) simple functions, which are fundamental, such as in the definition of the Lebesgue integral. When estimating grid distributions, one has to choose k or, equivalently, the size of the bin $(1/k)$, which is nothing more than the bandwidth of the grid-distribution estimator. Similarly, kernel-smoothing requires a choice of bandwidth parameter, too. In sum, nonparametric estimation using either grid distributions or smoothed kernels is very similar.

[3] Methods exist that require fewer smoothness conditions—e.g., the function need just be continuous $\mathbb{C}^0$; others require additional smoothness, $\mathbb{C}^2$ or higher. This does not change our claims.

[4] Recall that $\mathcal{D}^k$ is the set of grid distributions where the interval is subdivided into k intervals; i.e., $\mathcal{D}^k \equiv \mathbb{T}^k(\mathcal{D})$.

4.2. *Consistency and Power of the Proposed Test.* Of course, one concern is that k appears fixed in our analysis, but T is increasing, so our test is potentially inconsistent. We imagine a sequence of $\{k_T\}$ with values increasing as T increases, but not as fast as T. Below, we discuss in detail what we have in mind. Another worry is that the test statistic will be ill-behaved if k_T tends to infinity. Thus, an upper bound $\bar{k}$ must exist. This discussion leads us to introduce the following assumption concerning f_V^0 which allows us to side-step these technical problems:

ASSUMPTION 4.1. *The true data-generating process f_V^0 is a grid distribution; i.e., there exists $\bar{k} \in \mathbb{N}$ such that $f_V^0 \in \mathcal{D}^{\bar{k}}$.*

As the discussion above made clear, this assumption is similar to the assumptions of smoothness concerning f_V^0 which kernel-smoothing methods require. In addition to this analogy, we offer two additional justifications for Assumption 4.1.

First, the set of grid distributions is dense in the set of all distributions: even if the data-generating process (DGP) f_V^0 were not a grid distribution, there is a grid distribution that is arbitrary close to it. To wit, no finite amount of data could reject Assumption 4.1. In this sense, Assumption 4.1 is almost "no assumption."

Second, the DGP in question is a distribution of values, which are discrete (up to, say, dollars or cents or Yen or Won or whatever units one wants). When one assumes a smooth probability density function, one is making an approximation, for computational convenience: such an approximation does not seem, to us at least, any more appealing than the one we make. On the contrary, it seems more natural to us to assume simple probability density functions rather than any smoothness conditions. In general, smoothness is just a tool used to lighten the burden in the technical analysis of a particular problem. In our case, by assuming that the distribution is simple (i.e., a grid distribution), we can stay closer to reality.

Under Assumption 4.1, our test is consistent. For Assumption 4.1 implies that a k exists such that $f_V^0 \in \mathcal{D}^k$. Therefore, the number of inequalities required for affiliation remains fixed. We are then in the standard framework considered by Wolak, which has a fixed set of inequalities. Thus, consistency follows directly from Wolak's research. A technically sophisticated reader may feel that our consistency result is trivial, once Assumption 4.1 is made. The point of this paper (and this subsection, in particular) is not to provide a technical proof of consistency, but rather to remove any doubts concerning the consistency of our test under a reasonable assumption.

The power of the proposed test clearly depends on the choice of k. Were k chosen to be one (i.e., a uniform distribution on the N-dimensional hypercube), then affiliation would never be rejected. On the other hand, given a finite sample of T observations, a large k will result in many cells having no elements. While the choice of k is obviously important and certainly warrants additional theoretical investigation, perhaps along the lines of research in time-series analysis by Guay et al. [2008] concerning optimal adaptive detection of correlation functions, it is beyond the scope of this paper. In fact, in most applications to auctions, where samples are often quite small, k will be dictated by practical considerations—viz., the relative size of T.

4.3. *Bounding the Number of Inequalities.* For our test statistic to be well-behaved, it is important to know that an upper bounds exists on the number of inequalities. For arbitrary N and k, assuming a symmetric distribution, we can construct a bound on how many inequalities there are. Because we focus on symmetric distributions,

$$\left[f_V^0\right](i_1, i_2, \ldots, i_N) = \left[f_V^0\right](i_1', i_2', \ldots, i_N')$$

where $(i_1', i_2', \ldots, i_N')$ is a permutation of $(i_1, i_2, \ldots, i_N)$. Thus, we need only consider sorted indices, indices $(i_1, i_2, \ldots, i_N)$ for which $i_1 \geq i_2 \geq \cdots \geq i_N$. Consider $(i_1, i_2, \ldots, i_N)$, a sorted index having ℓ different numbers; let $r_1, \ldots, r_\ell$ denote the number of repetitions of the different numbers in $(i_1, i_2, \ldots, i_N)$. Obviously, $r_1 + \ldots +$

$r_\ell = N$. Using this notation, the number of permutations of $(i_1, i_2, \ldots, i_N)$ is then $\frac{N!}{r_1!\ldots r_\ell!}$. For instance, the index $(4,3,3,2,2,2)$ has $\frac{6!}{1!2!3!}$ or 120 different permutations.

Given the above, we can now focus our attention to sorted indices only. Consider the lexicographic order of them. In this way, we can attribute an unambiguous natural number to each sorted index of length N. For example, consider $N = 3$ in which case (1,1,1) corresponds to 1; (2,1,1), to 2; (2,2,1), to 3; (3,1,1) to 5; (3,2,2), to 7; and (4,1,1) to 11. It is important to develop an algorithm to convert a sorted index into a corresponding number, which we describe now.

First, let us define $\mathrm{Num}(j,N)$ as the number of all indices that are weakly below (in the lexicographic order) to the index $(j, j, \ldots, j)$; i.e., the index that has j in all positions and has length N. It is easy to see that $\mathrm{Num}(1,N) = 1$, because there is just one index weakly below $(1,1,1,\ldots,1) : (1,1,1,\ldots,1)$, itself. Also, $\mathrm{Num}(2,2) = 3$, because $(1,1)$, $(2,1)$, and $(2,2)$ are the sorted indices weakly below $(2,2)$. Similarly, $\mathrm{Num}(2,3) = 4$, because $(1,1,1)$, $(2,1,1)$, $(2,2,1)$, $(2,2,2)$ are the sorted indices weakly below $(2,2,2)$. From this argument, it is not difficult to see that $\mathrm{Num}(2,N) = N+1$. Observe, too, that $\mathrm{Num}(j,1) = j$, because there are only the indexes (1), (2), …, (j) weakly below (j). de Castro [2008] has proven the following:

LEMMA 4.1. $\mathrm{Num}(j,N) = \binom{N+j-1}{j-1}$.

Thus, if we fix the number of bidders N and the number of intervals k, then there are $M \equiv \mathrm{Num}(k,N) = \binom{N+k-1}{k-1}$ different indices. Affiliation will be satisfied if the corresponding inequality is satisfied for any pair of indices (i,i'). Since there are $\binom{M}{2}$ or $\frac{M(M-1)}{2}$ pair of such indices, it is sufficient to test $(M^2-M)/2$ inequalities. Note, however, that this is an upper bound because some inequalities are implied by others. The above discussion also provides some guidance concerning how to choose the inequalities; however, in an effort to conserve space, we leave the discussion of what the minimal set of sufficient inequalities is to another paper.

4.4. *Two Related Papers.* Like us, Li and Zhang [2008] have examined some important economic implications of affiliation. Instead of considering bids, however, Li and Zhang examined the entry behaviour of potential bidders whose signals may be affiliated. Theirs is a parametric analysis and they implemented their test using simulation methods, examining timber sales organized by the Department of Forestry in the State of Oregon. Li and Zhang found only a small degree of affiliation, perhaps because the zero/one entry decision is not as informative as bid data.

Jun et al. [2009] have developed a consistent nonparametric test designed for continuous data. By avoiding discretization, Jun et al. presumably have more information than we do. On the other hand, having rejected affiliation with their test, it is unclear what to do within their framework because an alternative hypothesis is unspecified. In contrast, our approach augments the theoretical work of de Castro [2008] where the alternative hypothesis is clearly outlined.

4.5. *Policy Uses for Grid Distributions.* de Castro [2008] has developed a complete theoretical treatment of grid distributions, even in the absence of affiliation. His idea is as follows: first, assume that $f_V^0 \in \mathcal{D}^k$ for some k; i.e, the DGP is a grid distribution—Assumption 4.1 holds. Standard estimation methods (histograms) can be used to calculate $[\hat{\mathbf{P}}] \in \mathcal{D}^k$ that best approximate f_V^0.

Under de Castro's method, one can then test whether $[\hat{\mathbf{P}}]$ has a symmetric MPSE. The method developed by de Castro is exact: to wit, modulo sampling error, $[\hat{\mathbf{P}}]$ has a symmetric MPSE if and only if the method detects it. Errors can occur only in simple numerical operations such as sums, divisions and square roots. It turns out that determining the existence of a symmetric MPSE is nontrivial when affiliation is absent.

If $[\hat{\mathbf{P}}]$ has a symmetric MPSE, then it can be used to calculate expected revenues under the first- and second-price auctions, denoted $R^1_{[\hat{\mathbf{P}}]}$ and $R^2_{[\hat{\mathbf{P}}]}$, respectively. In this way, one can determine which auction format yields an higher expected revenue for $[\hat{\mathbf{P}}]$ and, also, the magnitude of the revenue difference $(R^2_{[\hat{\mathbf{P}}]} - R^1_{[\hat{\mathbf{P}}]})$, to decide whether it is significant.

The procedure can then be repeated using $[\tilde{\mathbf{P}}]$, which is obtained under the constraint that the distribution satisfies affiliation. We know that, under affiliation, a symmetric MPSE exists and that $(R^2_{[\tilde{\mathbf{P}}]} - R^1_{[\tilde{\mathbf{P}}]}) \geqslant 0$, but the method also allows one to decide whether the magnitudes of the differences $(R^1_{[\hat{\mathbf{P}}]} - R^1_{[\tilde{\mathbf{P}}]})$ and $(R^2_{[\hat{\mathbf{P}}]} - R^2_{[\tilde{\mathbf{P}}]})$ are economically important.

Finally, the method allows one to examine sampling variability by repeating the above procedures using resampled draws from $[\hat{\mathbf{P}}]$ or $[\tilde{\mathbf{P}}]$.

Thus, the grid distributions proposed in this paper have many advantages because a theory exists that can be used for policy analysis. Such theories have not yet been developed for the methods proposed by Li and Zhang [2008] or Jun et al. [2009]; if affiliation is rejected under their methods, then what to do?

5. Monte Carlo Experiment. Below, we describe a small-scale Monte Carlo experiment used to investigate the numerical as well as small-sample properties of our testing strategy. Our simulation study involved samples of size T equal 100 and 250 with N of three bidders; each sample was then replicated $1,000$ times. In all of the experiments, the building blocks were triplets of independently- and identically-distributed uniform random variables on the interval $[0,1]$. We considered the following three types of experiments:

SI) (U_1, U_2, U_3) are independent uniform random variables;

SA) (U_1, U_2, U_3) are symmetric and affiliated random variables according to the Frank copula which has the following generator function:

$$\zeta(u) = -\log\left[\frac{\exp(-\alpha u) - 1}{\exp(-\alpha) - 1}\right], \tag{5.1}$$

and inverse-generator function

$$\zeta^{-1}(\tau) = -\frac{1}{\alpha}\log\left(1 + \exp(\tau)[\exp(-\alpha) - 1]\right). \tag{5.2}$$

where the parameter α controls affiliation;

AN) (U_1, U_2, U_3) are negatively correlated random variables having the following correlation matrix:

$$\Sigma = \begin{pmatrix} 1.0 & -0.1 & -0.2 \\ -0.1 & 1.0 & -0.3 \\ -0.2 & -0.3 & 1.0 \end{pmatrix} = \mathbf{F}\mathbf{F}^\top$$

where

$$\mathbf{F} = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 \\ -0.1000 & 0.9950 & 0.0000 \\ -0.2000 & -0.3216 & 0.9255 \end{pmatrix}$$

Above, SI denotes symmetric independence, which is the knife-edge case of affiliation when all J of the TP_2 determinants are exactly zero; SA denotes symmetric affiliation, which is the case when all of the TP_2 determinants are positive; and AN denotes asymmetric nonaffiliation, which is the case when none of the TP_2 determinental inequalities are satisfied.

What interpretation can be given to the dependence parameter α? In the bivariate case, the larger is a positive value of α, the greater the concordance, positive dependence. On the other hand, a very negative value of α indicates negative dependence. Independence obtains when α approaches zero. Note, however, that when N exceeds two, α is restricted to be positive because a negative α would mean a nonmonotonic inverse-generator function of the Frank copula; see example 4.22, page 123 in Nelsen [1999]. The Frank copula has the following N-variate form:

$$\mathbf{C}_\zeta(u_1,\ldots,u_N) = -\frac{1}{\alpha}\log\left(1+\frac{\prod_{i=1}^{N}[\exp(-\alpha u_i)-1]}{[\exp(-\alpha)-1]^{N-1}}\right) \quad \alpha>0. \tag{5.3}$$

Mueller and Scarsini [2005] have characterized various notions of positive dependence, such as MTP_2, for the Archimedean family of copulas, of which the Frank copula is a member. They have also presented a general condition that the generator of an arbitrary Archimedean copula must satisfy in order to guarantee that MTP_2 holds (cf. Theorem 2.11 in their paper). Genest [1987] has shown that the relevant condition for the Frank copula coincides with the condition that guarantees a monotonic inverse-generator function when N exceeds two; viz., α must be positive. Genest's condition requires that the Frank copula satisfy TP_2 as he was only concerned with the bivariate Frank copula. As mentioned above, however, it is well-known that a function is MTP_2 if and only if it is TP_2 in all pairs.

To simulate data from a Frank copula with affiliation, we followed the approach described by Luciano et al. [2004]. Theirs involves *conditional sampling* where, initially, w_1, a $U(0,1)$ random draw is taken, and then u_1 is set equal to it. The next (dependent) draw is taken from $\mathbf{C}_2(w_2|u_1)$, and u_3 is drawn from $\mathbf{C}_3(w_3|u_1,u_2)$ where all the w_is are independent $U(0,1)$ draws. We implemented conditional sampling using the parameterization of the Frank copula given in equation (5.3) in conjunction with the generator function defined in equation (5.1) and the inverse-generator function defined in equation (5.2). Specifically, to generate symmetrically-affiliated draws (u_1,u_2,u_3) from the trivariate Frank copula, we did the following:

1. simulate the independent random variables (w_1,w_2,w_3) from $U(0,1)$;
2. set u_1 equal to w_1;
3. use w_2 and u_1 to calculate

$$u_2 = -\frac{1}{\alpha}\left(1+\frac{w_2[1-\exp(-\alpha)]}{w_2[\exp(-\alpha u_1)-1]-\exp(-\alpha u_1)}\right); \tag{5.4}$$

4. use w_3 as well as u_1 and u_2 to define the following polynomial equation of order two in the variable $[\exp(-\alpha u_3)-1]$:

$$\begin{aligned} w_3 \quad = \quad & D_2^{-1}[\exp(-\alpha u_3)-1][\exp(-\alpha)-1]\times \\ & \left([\exp(-\alpha)-1]+[\exp(-\alpha u_1)-1][\exp(-\alpha u_2)-1]\right)^2 \end{aligned}$$

where

$$D_2 = \left([\exp(-\alpha)-1]^2+[\exp(-\alpha u_1)-1][\exp(-\alpha u_2)-1][\exp(-\alpha u_3)-1]\right)^2$$

which is then solved for u_3.

The above algorithm yields three symmetrically-affiliated random draws from the trivariate Frank copula for one simulation draw; this procedure was repeated either 100 or 250 times for each of $1,000$ replications.

In figure 5, we present a plot of all the points, when N is three and T is 250, for one replication generated under independence, weak affiliation. (Remember: an α of zero is the independent case.) Note that the

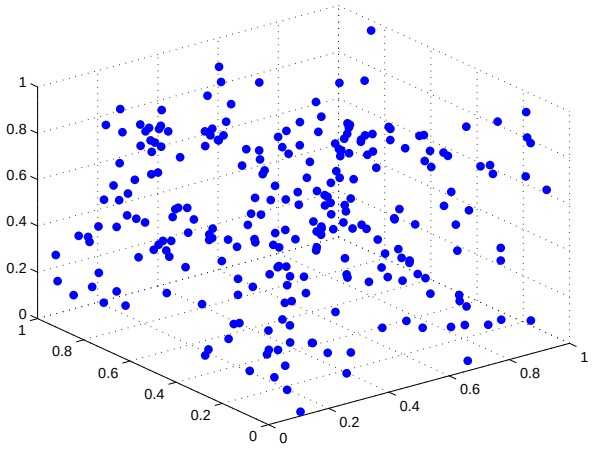


FIG 5. *Simulated Data Under Independence,* $\alpha = 0$

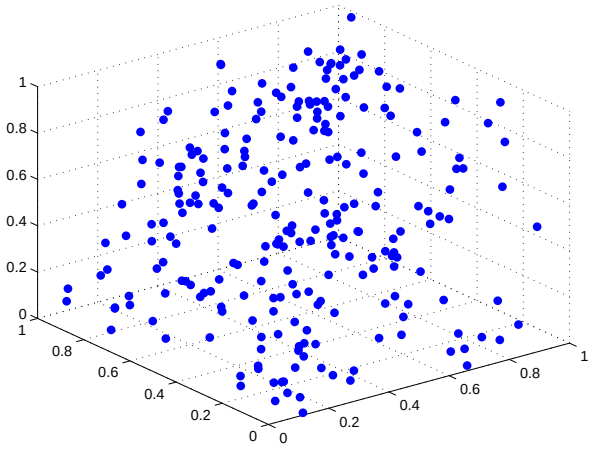


FIG 6. *Simulated Data Under Independence,* $\alpha = 2$

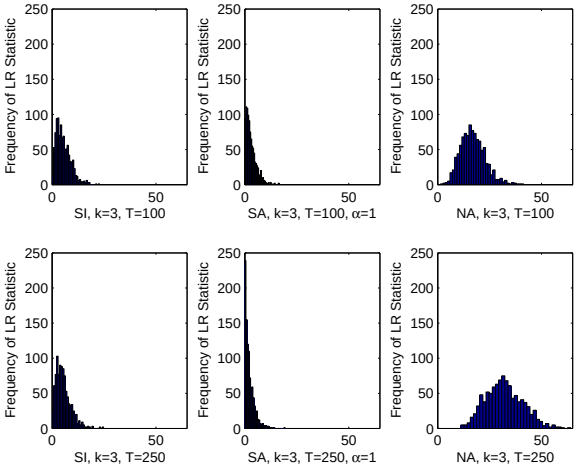


FIG 7. *SI, SA, NA:* $k = 3$, $T = 100$ *and* $T = 250$

scatterplot looks as one would expect—uniform. In figure 6, we present a plot of all the points, when N is three and T is 250, for one replication generated when α is 2, which means only modest affiliation. Note, in the scatterplot, a bit of white space exists in the extremes—e.g., near $(0,1,1)$.

We solved the constrained nonlinear optimization problems specified by equations (4.1), (4.2), (4.3), and (4.4) by implementing our methods using the programming language AMPL; for additional details concerning AMPL, see Fourer et al. [2002]. Using AMPL has a number of advantages: first, its user interface admits choice among a variety of nonlinear optimization solvers, including SNOPT and MINOS, without having to modify code significantly; second, AMPL can also perform automatic differentiation on nonlinear programming problems; and, third, the language is free. In fact, users can run the code for free using the NEOS server online. The code for a representative estimation problem ran in around one second on an unremarkable desktop computer.

In figure 7 are presented the frequency distributions of the LR test statistics for SI, SA, and NA when k is three and T is either 100 or 250. Our test appears able to distinguish relatively well between weak and modest affiliation, and to detect nonaffiliation extremely well. The test has relatively high power in the case NA, nonaffilation. Note, too, that as T increases, the distribution of the test statistic under SI remains constant, the one under SA shifts to the left, and the one under NA shifts to the right.

For k of three and T of 100 with symmetric independence, one can calculate the weights $\{\omega(j, J - j, \Pi^0)\}_{j=0}^{J}$ in Theorem 4.1. In figure 8, we present the exact probability density function of the asymptotic approximation as well as the kernel-smoothed estimate using the Monte Carlo data. The approximation appears quite close to the actual process, suggesting that the first-order asymptotics are working quite well.

6. Empirical Application. Above, in section 3, in the tradition of the theoretical literature concerning auctions, we developed our model of bidding in terms of valuations for an object to be sold at auction under the first-price, sealed-bid format. Sealed-bid tenders are often used in procurement—i.e., low-price, sealed-bid auctions at which a buyer (often a government agency) seeks to find the lowest-cost producer of some good or service. In this section, we report results from an empirical investigation of procurement tenders

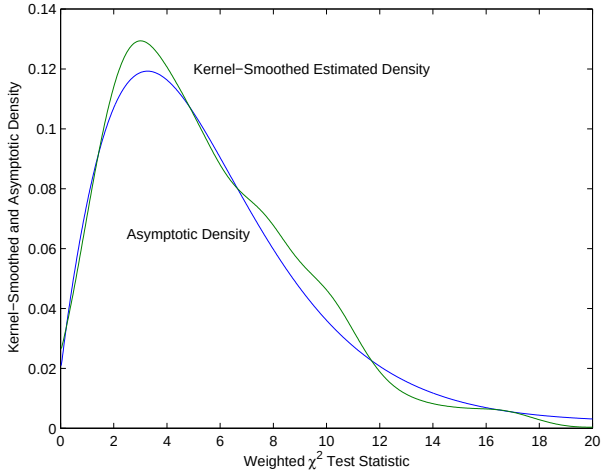


FIG 8. *Asymptotic and Kernel-Smoothed Density of Test Statistic: SI,* $k = 3$, $T = 100$

for road resurfacing by a government agency. Although it is well-known that results from auctions can be translated to procurement, and vice versa, sometimes this translation is tedious. We direct the interested reader to the work of de Castro and de Frutos [forthcoming] who have developed a procedure to translate results from auctions to procurement.

We have applied our empirical framework to data from low-price, sealed-bid, procurement auctions held by the Department of Transportation (DOT) in the State of Michigan. At these auctions, qualified firms are invited to bid on jobs that involve resurfacing roads in Michigan. We have chosen this type of auction because the task at hand is quite well-understood. In addition, there are reasons to believe that firm-specific characteristics make the private-cost paradigm a reasonable assumption; e.g., the reservation wages of owners/managers, who typically are paid the residual, can vary considerably across firms. On the other hand, other features suggest that the cost signals of individual bidders could be dependent, perhaps even affiliated; e.g., these firms hire other factor services in the same market and, thus, face the same costs for inputs such as energy as well as paving inputs. For example, suppose paving at auction t has the following Leontief production function for bidder n:

$$q_{nt} = \min\left(\frac{h_{nt}}{\delta_h}, \frac{y_{nt}}{\delta_y}, \frac{z_{nt}}{\delta_z}\right)$$

where h denotes the managerial labour, while y and z denote other factor inputs which are priced competitively at W_t and X_t, respectively, at auction t. Assume that R_n, bidder n's marginal value of time, is an independent, private-cost draw from a common distribution. In addition, assume that the other factor prices W_t and X_t are draws from another joint distribution, and that they are independent of R_n. The marginal cost per mile C_{nt} at auction t can be then written as:

$$C_{nt} = \delta_h R_n + \delta_y W_t + \delta_z X_t,$$

which is a special case of an affiliated private-cost (APC) model, known as a *conditional private-cost* model. The costs in this model are affiliated only when the distribution of R_n is log-concave, which is discussed

TABLE 1
Sample Descriptive Statistics—Dollars/Mile: $N = 3$; $T = 278$

Variable	Mean	St.Dev.	Median	Minimum	Maximum
Engineer's Estimate	475,544.54	491,006.52	307,331.26	54,574.41	3,694,272.59
Winning Bid	466,468.63	507,025.81	286,102.57	41,760.32	3,882,524.81
All Tendered Bids	507,332.42	564,842.58	317,814.77	41,760.32	5,693,872.81

extensively in de Castro [2008]. Li et al. [2000] have studied this model, extensively. In short, the affiliated private-cost paradigm (APCP) seems a reasonable null hypothesis.

We did not investigate issues relating to asymmetries across bidders because we do not know bidder identities, data necessary to implement such a specification. Because no reserve price exists at these auctions, we treat the number of participants as if it were the number of potential bidders and focus on auctions at which three bidders participated. Thus, we are ignoring the potential importance of participation costs which others, including Li [2005], have investigated elsewhere.

The data for this part of the paper were provided by the Michigan DOT and were organized and used by Hubbard et al. [2009]; a complete description of these data is provided in that paper. In table 1, we present the summary descriptive statistics concerning our sample of 834 observations—278 auctions that involved three bidders each. We chose auctions with just three bidders not only to illustrate the general nature of the method (if we can do three, then we can do N), but also to reduce the data requirements. When we subdivide the unit hypercube into k^N cells, the average number of bids in a cell is proportional to (k^N/T). When N is very large, the sample size must be on the order of k^N in order to expect at least one observation in each cell. This example also illustrates the potential limitations of our approach; viz., even in relatively large samples, some of the cells will not be populated, so k will need to be kept small. However, one can circumvent this problem by varying the width of the subdivisions as we do below. Of course, one must then adjust the conditions which define the determinental inequalities. We describe this below, too.

Our bid variable is the price per mile. Notice that both the winning bids as well as all tendered bids vary considerably, from a low of \$41,760.32 per mile to a high of \$5,693,872.81 per mile. What explains this variation? Well, presumbably heterogeneity in the tasks that need to be performed. One way to control for this heterogeneity would be to retrieve each and every contract and then to construct covariates from those contracts. Unfortunately, the State of Michigan cannot provide us with this information, at least not any time soon.

How can we deal with this heterogeneity? Well, in our case, we have an engineer's estimate p of the price per mile to complete the project.[5] We assume that C_{nt}, the cost to bidder n at auction t, can be factored as follows:

$$C_{nt} = \lambda^0(p_t)\varepsilon_{nt} \tag{6.1}$$

where λ^0 is a known function. One example of this is

$$C_{nt} = p_t\varepsilon_{nt}.$$

Another is

$$C_{nt} = \delta_0 p_t^{\delta_1}\varepsilon_{nt}.$$

[5]Of course, besides p, it is possible that other covariates, which are common knowledge to all the bidders, exist. If these other common-knowledge covariates exist, then we could wrongly conclude that the signals have a strong form of correlation when, in fact, the correctly-specified model of signals (conditioned on the common-knowledge information) would have only small correlation. Unfortunately, we do not have access to any additional information. Were such information available, then we would condition on it as well.

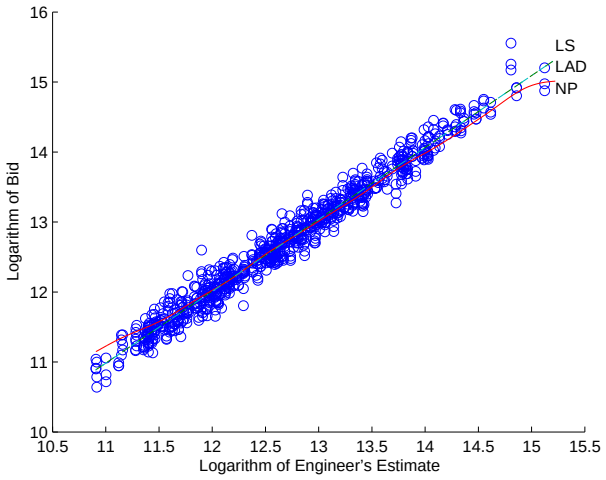


FIG 9. *Data as well as NP, LS, and LAD Regressions: Logarithm of Bids versus Logarithm of Engineer's Estimate*

Under equation (6.1), the equilibrium bid B_{nt} at auction t for bidder n takes the following form:

$$B_{nt} = \lambda^0(p_t)\beta(\varepsilon_{nt}),$$

so

$$\frac{B_{nt}}{\lambda^0(p_t)} = \beta(\varepsilon_{nt}).$$

Of course, we do not know λ^0, but we can estimate λ^0 either parametrically, under an appropriate assumption, or nonparametrically, using the following empirical specification:

$$\log B_{nt} = \psi(p_t) + U_{nt}$$

where $\psi(p_t)$ denotes $-\log[\lambda^0(p_t)]$ and U_{nt} denotes $\log[\beta(\varepsilon_{nt})]$.

Empirical results from this exercise are presented in figure 9. In this figure are presented results for the nonparametric regression (NP), the least-squares regression (LS), the least-absolute-deviations (LAD) regression. To get some notion of the relative fit, note that the R^2 for the LS regression is around 0.97. The LS estimates of the constant and slope coefficients are -0.3114 and 1.0268, respectively, while LAD estimates of the constant and slope coefficients are -0.3221 and 1.0276, respectively.

Subsequently, we took the normalized fitted residuals, which (for the LS case) are depicted in figure 10, and applied the methods described in section 4 above for a k of two. Our test results are as follows: the maximum of the logarithm of the likelihood function (minus a constant) without symmetry was -442.50, while the maximum of the logarithm of the likelihood function under symmetry was -444.88, and under symmetric affiliation it was also -444.88—a total difference of 2.38.[6] At size 0.05, twice the above difference is above the lower bound provided by Kodde and Palm [1986], but below the upper bound, so the test is inconclusive.

[6]The results for the LAD residuals were identical: the probability array obtained by discretizing the LAD residuals was exactly the same as in the LS case because none of the fitted residuals was classified differently. This is not, perhaps, surprising given the similar fits of the two empirical specifications.

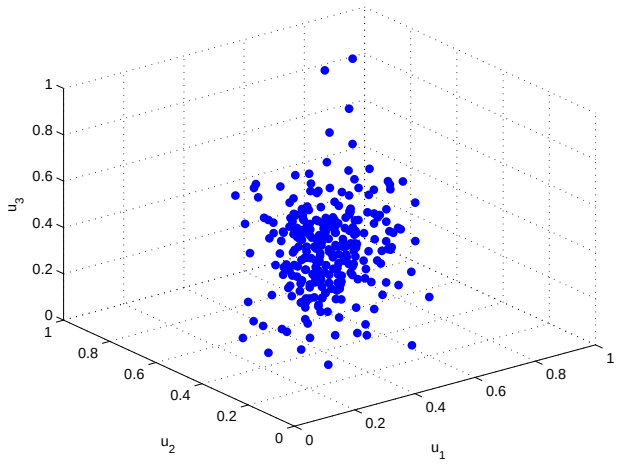


FIG 10. *Scatterplot of Transformed Fitted LS Residuals*

Because a k of two is unusually small, we introduced a symmetric, but nonequispaced, grid distribution—like the one depicted in figure 4, but with intervals $[0.,0.4)$, $[0.4,0.6)$, and $[0.6,1.0]$. The TP_2 inequalities can be derived in the usual way, but the adding-up inequality must be rewritten, in this case as

$$a+2b+8c+8d+16e+8f+4d+2h+4i+4b+16f+8g+8e+2g+8j+8f+16c+8i \leq 1.$$

Again, we applied our methods. Our test results are as follows: the maximum of the logarithm of the likelihood function (minus a constant) under symmetry was -715.72, while the maximum under symmetric affiliation was -716.49—a difference of 0.77.[7] At size 0.05, twice the above difference is below the lower bound provided by Kodde and Palm, so we do not reject the hypothesis of symmetric affiliation. To put these results into some context, the centre of the simplex had a logarithm of the likelihood function of -916.24; using the marginal distribution of low, medium, and high costs $(0.4233, 0.4808, 0.0959)$ and imposing independence yielded a logarithm of the likelihood function of -784.67.

7. Summary and Conclusions. We have constructed a tractable empirical model of equilibrium behaviour at first-price auctions when bidders' private valuations are dependent, but not necessarily affiliated. Subsequently, we developed a test of affiliation and then investigated its small-sample properties. We applied our framework to data from low-price, sealed-bid auctions used by the Michigan DOT to procure road-resurfacing: we do not reject the hypothesis of affiliation in cost signals.

This information is potentially useful to a policy maker. The apparent high degree of estimated affiliation also explains why low levels of observed competition are often sufficient to maintain relatively low profit margins: strong affiliation is akin to fierce competition. Under strong affiliation, a potential winner knows that his nearest competitor probably has a valuation (cost) close to his, and this disciplines his bidding behaviour: he become more aggressive than under independence.

[7] The results for the LAD residuals were virtually identical: the probability array obtained by discretizing the LAD residuals was almost the same as in the LS case.

Our research has other policy implications, too. As mentioned above, it is well-known that, under affiliation, the English auction format, on average, generates more revenue for the seller than the first-price, sealed-bid format. In procurement, under affiliation, an English or a Vickrey auction would get the job done more cheaply than the low-price, sealed-bid format, Were the English or Vickrey formats being used and affiliation not rejected, then the procurement agency would be justified in its choice of mechanism. What remains a bit of a puzzle is why the low-price, sealed-bid format is used in the presence of such strong affiliation. Perhaps, other features, such as the ability of the low-price, sealed-bid auction format to thwart collusion are important, too. Alternatively, perhaps other moments of the bid distribution, such as the variance, are important to the procurement agency.

On the other hand, had affiliation been rejected, then the procedures described in section 4 could be used to determine which auction format would get the job done most cheaply, on average. Again, it is possible that the English or Vickrey formats would still be preferred. In any case, the methods described in section 4 permit a better understanding of the bidding differences, which can aid in choosing the best auction format.

Acknowledgements. Previous versions of this paper circulated under de Castro and Paarsch [2007] and de Castro and Paarsch [2008]. For their comments on that research, we thank Kurt Anstreicher, Samuel Burer, Victor Chernozhukov, Srihari Govindan, Emmanuel Guerre, Han Hong, Joel L. Horowitz, Ali Hortaçsu, Kenneth L. Judd, Roger Koenker, David Prentice, Joseph P. Romano, Che-Lin Su, Elie Tamer, and Michael Wolf. Timothy P. Hubbard deserves special recognition and thanks for his advice and help with AMPL as well as his comments and insights on the earlier research. We are also grateful to Stephen E. Fienberg, Vijay Krishna, Charles F. Manski, E. Glen Weyl, and Frank A. Wolak as well as two anonymous referees who provided helpful comments and useful suggestions on the penultimate draft of this paper.

References.

Francesco Bartolucci and Antonio Forcina. A likelihood ratio test for MTP_2 within binary variables. *Annals of Statistics*, 28:1206–1218, 2000.

Bjarne O. Brendstrup and Harry J. Paarsch. Semiparametric estimation of multi-object, sequential, English auctions. *Journal of Econometrics*, 141:84–108, 2007.

Federico Bugni. Bootstrap inference in partially identified models, Northwestern University, Department of Economics, typescript, 2008.

Luciano I. de Castro. Affiliation, equilibrium existence, and the revenue ranking of auctions, Carlos III University, Department of Economics, typescript, 2007.

Luciano I. de Castro. Grid distributions to study single object auctions, University of Illinois Urbana-Champaign, Department of Economics, typescript, 2008.

Luciano I. de Castro and Maria-Angeles de Frutos. How to translate results from auctions to procurements. *Economic Letters*, forthcoming.

Luciano I. de Castro and Harry J. Paarsch. Testing for affiliation and monotone pure-strategy equilibrium in models of first-price auctions with private values, University of Iowa, Department of Economics, typescript, 2007.

Luciano I. de Castro and Harry J. Paarsch. Using grid distributions to test for affiliation in models of first-price auctions with private values, University of Iowa, Department of Economics, typescript, 2008.

Ruth Douglas, Stephen E. Fienberg, Mei Ling T. Lee, Allan R. Sampson, and Lyn R. Whitaker. Positive dependence concepts for ordinal contingency tables. In Henry W. Block, Allan R. Sampson, and Thomas H. Savits, editors, *Topics in Statistical Dependence*. Hayward, California: Institute of Mathematical Statistics, 1990.

Robert Fourer, David M. Gay, and Brian W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*, Second Edition. Pacific Grove, CA: Brooks/Cole Publishing Company, 2002.

Christian Genest. Frank's family of bivariate distributions. *Biometrika*, 74:549–555, 1987.

Alain Guay, Emmanuel Guerre, and Stepana Lazarova. Optimal adaptive detection of small correlation functions, Queen Mary University of London, Department of Economics, typescript, 2008.

Timothy P. Hubbard, Tong Li, and Harry J. Paarsch. Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation, University of Melbourne, Department of Economics, typescript, 2009.

Sung Jae Jun, Joris Pinkse, and Yuanyuan Wan. A consistent nonparametric test of affiliation in auction models, Pennsylvania State University, Department of Economics, typescript, 2009.

Samuel Karlin. *Total Positivity*, Volume 1. Stanford, CA: Stanford University Press, 1968.
David A. Kodde and Franz C. Palm. Wald criteria for jointly testing equality and inequality restrictions. *Econometrica*, 54:1243–1248, 1986.
Vijay Krishna. *Auction Theory*, Second Edition. San Diego: Academic Press, 2010.
Jean-Jacques Laffont and Quang H. Vuong. Structural analysis of auction data. *American Economic Review*, 86:414–420, 1996.
Tong Li. Econometrics of first-price auctions with binding reservation prices. *Journal of Econometrics*, 126:173–200, 2005.
Tong Li and Bingyu Zhang. Testing for affiliation in first-price auctions using entry behavior, Vanderbilt University, Department of Economics, typescript, 2008.
Tong Li, Isabelle M. Perrigne, and Quang H. Vuong. Conditionally independent private information in OCS wildcat auctions. *Journal of Econometrics*, 98:129–161, 2000.
Tong Li, Isabelle M. Perrigne, and Quang H. Vuong. Structural estimation of the affiliated private value model. *RAND Journal of Economics*, 33:171–193, 2002.
Umberto Luciano, Elisa Vecchiato, and Walter Cherubini. *Copula Methods in Finance*. New York: John Wiley & Sons, 2004.
Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50:1089–1122, 1982.
Alfred Mueller and Marco Scarsini. Archimedean copulae and positive dependence. *Journal of Multivariate Analysis*, 93:434–445, 2005.
Roger B. Nelsen. *An Introduction to Copulas*. New York: Springer, 1999.
Dimitris N Politis, Joseph P. Romano, and Michael Wolf. *Subsampling*. Berlin: Springer–Verlag, 1999.
Joseph P. Romano and Azeem M. Shaikh. Inference for identifiable parameters in partially identified econometric models. *Journal of Statistical Planning and Inference*, 138:2786–2807, 2008.
David W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. New York: Wiley–Interscience, 1992.
Aleksandra B. Slavković and Stephen E. Fienberg. Algebraic geometry of 2×2 contingency tables. In Paolo Gibilisco, Eva Riccomagno, Maria-Piera Rogantin, and Henry P. Wynn, editors, *Algebraic and Geometric Methods in Statistics*. New York: Cambridge University Press, 2009.
Frank A. Wolak. An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association*, 82:782–793, 1987.
Frank A. Wolak. Testing inequality constraints in linear econometric models. *Journal of Econometrics*, 41:205–235, 1989a.
Frank A. Wolak. Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models. *Econometric Theory*, 5:1–35, 1989b.
Frank A. Wolak. The local nature of hypothesis tests involving inequality constraints in nonlinear models. *Econometrica*, 59, 1991.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, USA
DEPARTMENT OF ECONOMICS, UNIVERSITY OF MELBOURNE, AUSTRALIA
E-MAIL: deCastro.Luciano@gmail.com
HPaarsch@UniMelb.edu.au