

USING GRID DISTRIBUTIONS TO TEST FOR AFFILIATION IN MODELS OF FIRST-PRICE AUCTIONS WITH PRIVATE VALUES*

Luciano I. de Castro
Department of Economics
University of Illinois

Harry J. Paarsch[†]
Department of Economics
University of Melbourne

April 2008

Abstract

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.

JEL Classification Numbers: C20, D44, L1.

Keywords: first-price, sealed-bid auctions; affiliation; MTP₂.

* 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, and Elie Tamer for useful advice and helpful suggestions during our early research. However, Timothy P. Hubbard deserves special recognition and thanks for his advice and help with AMPL as well as his comments and insights on earlier drafts of this paper.

[†] Corresponding author: Fifth Floor, Economics and Commerce Building; Melbourne, Victoria, 3010, AUSTRALIA. Telephone: 61 3 8344 4655; E-mail: hjp@paarsch.ecom.unimelb.edu.au

1. Motivation and Introduction

During the past half century, economists have made considerable progress in understanding the theoretical structure of strategic behaviour under market mechanisms, such as auctions, when the number of potential participants is relatively small; see Krishna (2002) 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 often 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 (1982): “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 started 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 non-decreasing in their arguments. de Castro (2007) has noted that affiliation implies that

- 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), often referred to as a decreasing *inverse hazard rate*,

which implies that

- b) $\Pr(V_2 \leq v_2|V_1 = v_1)$ is non-increasing in v_1 (and v_2 in the other case), also referred to as *positive regression dependence*,

which implies that

- c) $\Pr(V_2 \leq v_2|V_1 \leq v_1)$ is non-increasing in v_1 (and v_2 in the other case), also referred to as *left-tail decreasing* in v_1 (v_2),

which implies that

- d) $\text{cov}[g(V_1), h(V_2)]$ is positive,

which implies that

- e) $\text{cov}[g(V_1), h(V_2)]$ is positive,

which implies that

- f) $\text{cov}(V_1, V_2)$ 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, the three commonly-studied auction formats—English as well as first- and second-price—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 second-price (Vickrey) auctions which are weakly greater than those at first-price auctions (either sealed-bid or Dutch).

Investigating equilibrium behaviour at auctions, empirically, when latent valuations are affiliated, has challenged researchers for some time. Laffont and Vuong (1996) showed that identification is 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, Perrigne, and Vuong (2000) have demonstrated nonparametric identification within the conditional IPVP, a special case of the APVP, while Li, Perrigne, and Vuong (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 indirect estimation strategy need not constitute an equilibrium. Li, Paarsch, and Hubbard (2007) 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 bid-

ding 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 (TP₂). Subsequently, following the theoretical work of de Castro (2007,2008), 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. In section 4, we develop a test of affiliation, while in section 5 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. We summarize and conclude in section 6, the final section of the paper.

2. Affiliation and TP₂

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-person auction game with two values is depicted in figure 2.1. 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 TP₂—*viz.*,

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

Put another way, TP₂ 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 determinantal inequality. Note, too, that affiliation restricts distributions to a part of the simplex depicted in figure 2.2. In that figure, it is the region on the simplex that appears to be 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 2.2, 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.

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 player; five additional points then appear, as is depicted in figure 2.3. Affiliation requires that the probabilities at all collections of four points satisfy TP₂; *i.e.*, the following additional six inequalities must hold:

$$p_{12}p_{23} \geq p_{13}p_{22}$$

$$p_{22}p_{33} \geq p_{23}p_{32}$$

$$p_{21}p_{32} \geq p_{22}p_{31}$$

$$p_{11}p_{33} \geq p_{13}p_{31}$$

$$p_{12}p_{33} \geq p_{13}p_{32}$$

$$p_{21}p_{33} \geq p_{23}p_{31}.$$

Of course, symmetry would imply that p_{ij} equal p_{ji} for all i and j , so the joint mass function for two players 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

Figure 2.1
Affiliation with Two Players and Two Values for Signals

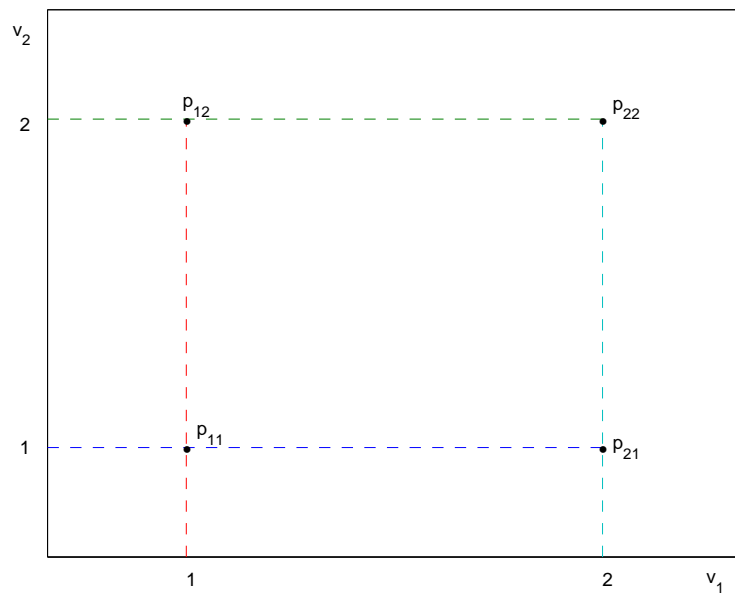


Figure 2.2
Probability Set: Affiliation and Alternative

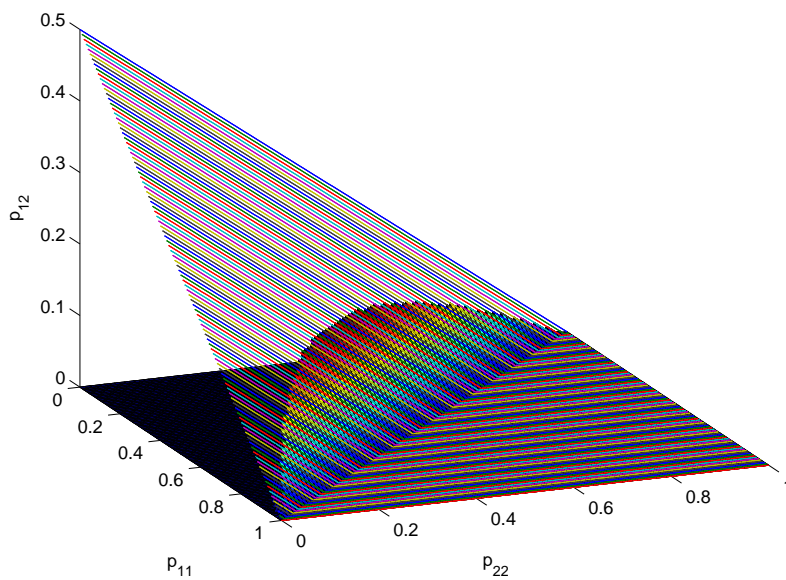
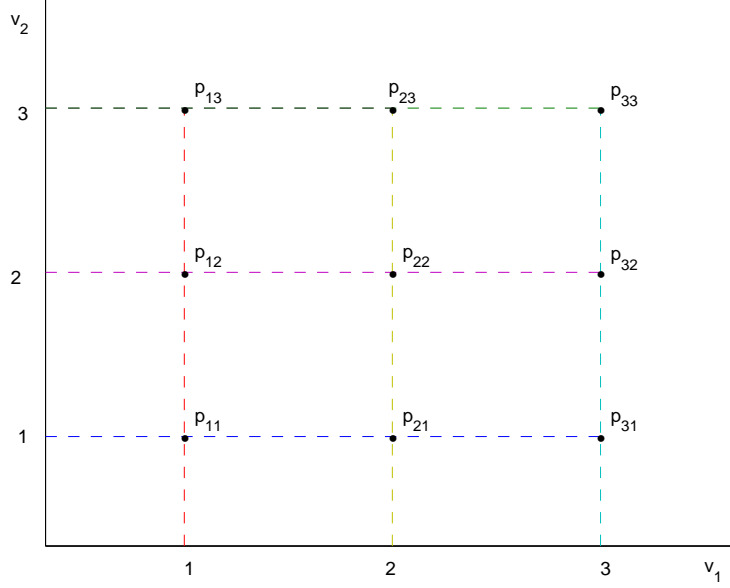


Figure 2.3
Affiliation with Two Players and Three Values for Signals



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 $(1, 2) \times (1, 2) : \mathbf{ab} \geq \mathbf{d^2}$ and $(2, 3) \times (1, 2) : \mathbf{df} \geq \mathbf{be}$. Finally, inequality $(2, 3) \times (1, 3) : dc \geq ef$ derives from the other two, previously established—*viz.*, $(2, 3) \times (1, 2) : \mathbf{df} \geq \mathbf{be}$ and $(2, 3) \times (2, 3) : bc \geq f^2$. All other inequalities can be obtained from the adjacent ones in this fashion, and we relegate to an appendix of the paper the necessary tedious calculations for other cases.

Adding values to the type spaces of players expands the number of determinantal restrictions required to satisfy TP_2 , thus restricting the space of distributions that can be entertained. Likewise, adding players to the game, particularly if the players are assumed symmetric, also restricts the space of distributions that can be entertained. For example, suppose a third player 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 players 1 and 2, indexed by the values of player 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 players 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 determinantal 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.*,

$$\begin{aligned} ab &\geq d^2, \quad bc \geq f^2, \quad df \geq be, \\ dh &\geq b^2, \quad hi \geq g^2, \quad bg \geq fh, \\ eg &\geq f^2, \quad gj \geq i^2, \quad fi \geq cg. \end{aligned}$$

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. The proof of this claim is contained in an appendix to the paper.

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 $\{1, 2, \dots, N\}$. Now, bidder n is assumed to draw V_n , his private valuation of the object for sale, from the closed interval $[\underline{v}, \bar{v}]$. We note that, without loss of generality, one can reparametrize the valuations from $[\underline{v}, \bar{v}]$ to $[0, 1]$. Below, we do this. We collect the valuations in the vector \mathbf{v} which equals (v_1, \dots, v_N) and denote this vector without the n^{th} element by \mathbf{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. Now, \mathbf{V} lives in $[0, 1]^N$. We assume that the values are distributed according to the probability density function $f_{\mathbf{V}} : [0, 1]^N \rightarrow \mathbb{R}_+$ which is symmetric; *i.e.*, for the permutation $\varphi : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$, we have $f_{\mathbf{V}}(v_1, \dots, v_N)$ equals $f_{\mathbf{V}}(v_{\varphi(1)}, \dots, v_{\varphi(N)})$. Letting $f_n(v_n)$ denote the marginal probability density function of V_n , we note that it equals $\int_0^1 \cdots \int_0^1 f_{\mathbf{V}}(\mathbf{v}_{-n}, v_n) d\mathbf{v}_{-n}$. (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 have not enough information to estimate the case with varying f_n s.) Our main interest is the case when $f_{\mathbf{V}}$ is *not* the product of its marginals—the case where the types are dependent. We denote the conditional density of \mathbf{V}_{-n} given v_n by

$$f_{\mathbf{V}_{-n}|V_n}(\mathbf{v}_{-n}|v_n) = \frac{f_{\mathbf{V}}(\mathbf{v}_{-n}, v_n)}{f_n(v_n)}.$$

Finally, we denote the largest order statistic of \mathbf{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 $\sigma : [0, 1] \rightarrow \mathbb{R}_+$ 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 $\sigma : [0, 1] \rightarrow \mathbb{R}_+$, is

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

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

$$\Pi[v_n, \sigma(v_n), \boldsymbol{\sigma}_{-n}] \geq \Pi(v_n, s, \boldsymbol{\sigma}_{-n}) \quad \forall s, v_n. \quad (3.1)$$

As mentioned above, in most theoretical models of auctions that admit dependence in valuation draws, researchers have assumed that $f_{\mathbf{V}}$ satisfies affiliation. Affiliation can be formally defined as follows: for all \mathbf{v} and \mathbf{v}' , the random variables \mathbf{V} are said to be affiliated if

$$f_{\mathbf{V}}(\mathbf{v} \vee \mathbf{v}') f_{\mathbf{V}}(\mathbf{v} \wedge \mathbf{v}') \geq f_{\mathbf{V}}(\mathbf{v}) f_{\mathbf{V}}(\mathbf{v}')$$

where

$$(\mathbf{v} \vee \mathbf{v}') = [\max(v_1, v'_1), \max(v_2, v'_2), \dots, \max(v_N, v'_N)]$$

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

$$(\mathbf{v} \wedge \mathbf{v}') = [\min(v_1, v'_1), \min(v_2, v'_2), \dots, \min(v_N, v'_N)]$$

denotes the component-wise minima, sometimes referred to as the *meet*. We do not restrict ourselves to $f_{\mathbf{V}}$ s that satisfy affiliation. We assume only that $f_{\mathbf{V}}$ belongs to a set of distributions \mathcal{P} which guarantees the existence and uniqueness of a MPSE. This set \mathcal{P} is partially characterized in de Castro (2008).

Let \mathcal{C} denote the set of continuous density functions $f_{\mathbf{V}} : [0, 1]^N \rightarrow \mathbb{R}_+$ 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

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

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

We now introduce a transformation $\mathbb{T}^k : \mathcal{D} \rightarrow \mathcal{D}$ which is the workhorse of our method. To define \mathbb{T}^k , let $\mathbb{I} : [0,1] \rightarrow \{1, 2, \dots, k\}$ 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 $\mathbb{S}(\mathbf{v})$ denote the “square” (hypercube) $\prod_{n=1}^N \left(\frac{\mathbb{I}(v_n)-1}{k}, \frac{\mathbb{I}(v_n)}{k}\right]$ where \mathbf{v} collects $(v_1, v_2, \dots, v_N) \in [0,1]^N$. From this, we define $\mathbb{T}^k : \mathcal{D} \rightarrow \mathcal{D}$ as the transformation that associates to each $f_{\mathbf{V}} \in \mathcal{D}$ the probability density function $\mathbb{T}^k(f_{\mathbf{V}})$ given by:

$$\mathbb{T}^k(f_{\mathbf{V}})(\mathbf{v}) = k^N \int_{\mathbb{S}(\mathbf{v})} f_{\mathbf{V}}(\mathbf{u}) \, d\mathbf{u}.$$

Observe that $\mathbb{T}^k(f_{\mathbf{V}})$ is constant over each square $\prod_{n=1}^N \left(\frac{m_n-1}{k}, \frac{m_n}{k}\right]$, for all combinations of $m_n \in \{1, \dots, k\}$. 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 $f_{\mathbf{V}} \in \mathcal{D}$, $\mathbb{T}^1(f_{\mathbf{V}})(\mathbf{v})$ equals one for all $\mathbf{v} \in [0,1]^N$; *i.e.*, $\mathbb{T}^1(f_{\mathbf{V}})$ is the uniform distribution on $[0,1]^N$.

We now need to introduce a compact notation to represent arrays of dimension $\overbrace{N \text{ times}}^{N \text{ times}}$ $k \times k \times \dots \times k$. 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 players, an array is obviously just a matrix, while in our application N is three. The $(i_1, i_2, \dots, i_N)^{\text{th}}$ element of an array is denoted $[\mathbf{P}](i_1, i_2, \dots, i_N)$, or $[\mathbf{P}](\mathbf{i})$ for short, where \mathbf{i} denotes the vector (i_1, i_2, \dots, 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

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

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

$$f_{\mathbf{V}}(v_1, v_2) = [\mathbf{P}](i, j) \text{ if } (v_1, v_2) \in \left(\frac{i-1}{k}, \frac{i}{k}\right] \times \left(\frac{j-1}{k}, \frac{j}{k}\right] \quad (3.2)$$

for $i, j \in \{1, 2, \dots, k\}$. The definition of $f_{\mathbf{V}}$ 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 0.4 wide, while the next one can be 0.2 wide, and the last can be 0.4 wide. 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 3.1. Let $0 = r_0 < r_1 < r_2 < \dots < r_{k-1} < r_k = 1$ be an arbitrary partitioning of the interval $[0, 1]$. Now, define $\mathbb{I} : [0, 1] \rightarrow \{1, 2, \dots, k\}$ by $\mathbb{I}(v) = j$ if and only if $v \in (r_{j-1}, r_j]$. Define $\mathbb{B}(\mathbf{v})$ as the rectangle (box) where \mathbf{v} collects $(v_1, \dots, v_N) \in [0, 1]^N$ lies. Thus, $\mathbb{B}(\mathbf{v}) \equiv \prod_{n=1}^N (r_{\mathbb{I}(v_n)-1}, r_{\mathbb{I}(v_n)}]$. Then, define

$$\mathbb{T}_{\mathbb{B}}^k(f_{\mathbf{V}}^0)(\mathbf{v}) = \frac{\int_{\mathbb{B}(\mathbf{v})} f_{\mathbf{V}}^0(\mathbf{u}) \, d\mathbf{u}}{\int_{\mathbb{B}(\mathbf{v})} d\mathbf{u}}.$$

The following theorem was proven by de Castro (2007):

Theorem 1. Let $f_{\mathbf{V}}^0$ be a symmetric and continuous probability density function. Then, $f_{\mathbf{V}}^0$ is affiliated if and only if for all k , $\mathbb{T}_{\mathbb{B}}^k(f_{\mathbf{V}}^0)$ is also affiliated.

In our notation,

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

or

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

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 3.1 when N is two.

One can also subdivide the space of valuations asymmetrically and irregularly as illustrated in figure 3.2. In this case, we impose a fine grid from which we then build

Figure 3.1
Symmetric Non-Equi-Spaced Grid

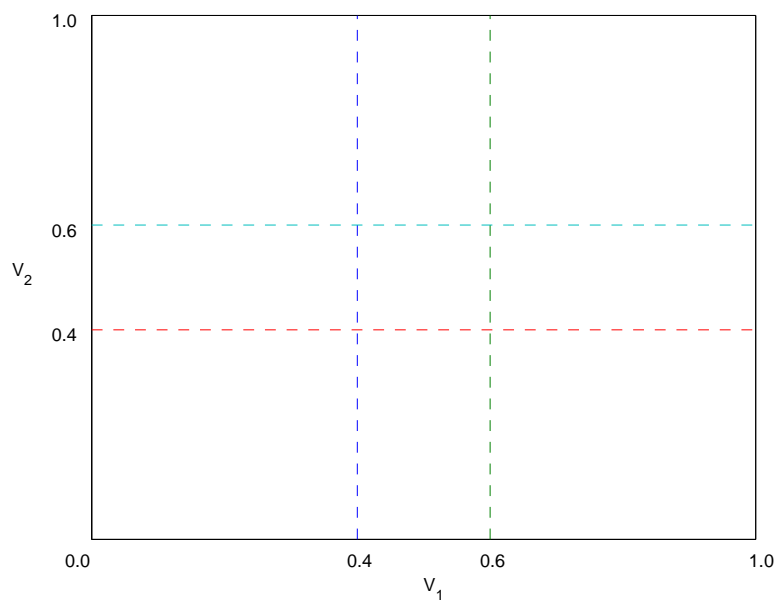
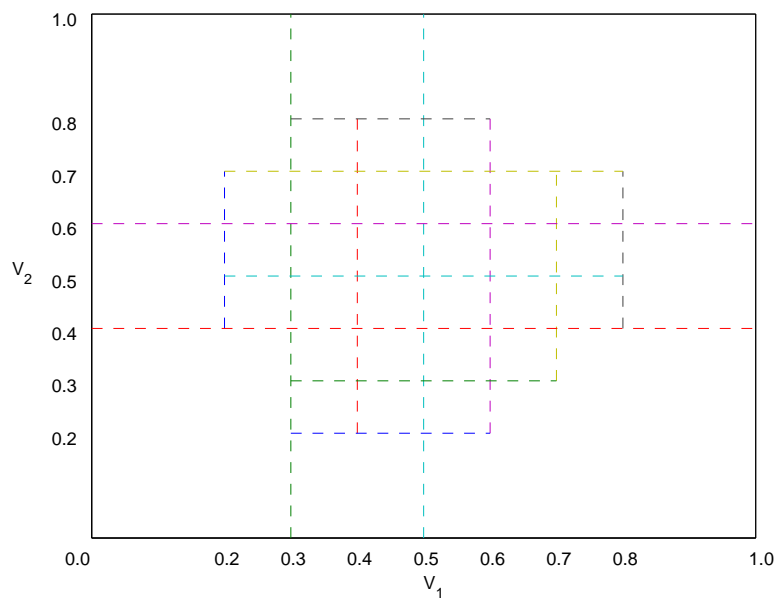


Figure 3.2
Asymmetric and Irregular Non-Equi-Spaced Grid



rectangles for the areas in which we are interested. To accomplish this, fix a grid of rectangles \mathbb{B} . We say that a generalized grid \mathbb{G} is generated by \mathbb{B} if the element $\mathbb{G}(\mathbf{v})$ where \mathbf{v} lies is formed by the union of rectangles $\mathbb{B}(\mathbf{v}')$ adjacent to $\mathbb{B}(\mathbf{v})$ or adjacent to rectangles adjacent to $\mathbb{B}(\mathbf{v})$, and so on. We define

$$\mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)(\mathbf{v}) = \frac{\int_{\mathbb{G}(\mathbf{v})} f_{\mathbf{V}}^0(\mathbf{u}) \, d\mathbf{u}}{\int_{\mathbb{G}(\mathbf{v})} d\mathbf{u}}.$$

Thus, with some modifications, we can still test affiliation using $\mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)$. The following corollary states what should be tested:

Corollary 1. Let $f_{\mathbf{V}}^0$ be a symmetric, continuous, and affiliated probability density function. Let \mathbb{G} be an area generated by \mathbb{B} . If $\mathbb{G}(\mathbf{u}) = \mathbb{B}(\mathbf{u})$, for $\mathbf{u} = \mathbf{v}$, \mathbf{v}' , $\mathbf{v} \wedge \mathbf{v}'$, and $\mathbf{v} \vee \mathbf{v}'$, then $\mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)$ satisfies the affiliated inequality

$$\mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)(\mathbf{v} \vee \mathbf{v}') \mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)(\mathbf{v} \wedge \mathbf{v}') \geq \mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)(\mathbf{v}) \mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)(\mathbf{v}').$$

Proof: This is an immediate implication of Theorem 1. In fact, since $f_{\mathbf{V}}^0$ is affiliated, $\mathbb{T}_{\mathbb{B}}^k(f_{\mathbf{V}}^0)$ is also affiliated, and $\mathbb{G}(\mathbf{u}) = \mathbb{B}(\mathbf{u})$ implies $\mathbb{T}_{\mathbb{G}}^k(f_{\mathbf{V}}^0)(\mathbf{u}) = \mathbb{T}_{\mathbb{B}}^k(f_{\mathbf{V}}^0)(\mathbf{u})$ for $\mathbf{u} = \mathbf{v}$, \mathbf{v}' , $\mathbf{v} \wedge \mathbf{v}'$, and $\mathbf{v} \vee \mathbf{v}'$.

At this point, it is useful to note that Corollary 1 does not state the converse implication, as Theorem 1 does. However, the stated direction is what we need in order to construct a test that can reject affiliation. Indeed, if the true distribution is affiliated, then all inequalities obtained of the form specified in Corollary 1 must be satisfied. If we are able to conclude that some of these inequalities do not hold, then this would imply that the true distribution is not affiliated. Also, note that the implications in the Corollary 1 are *necessary*, but not *sufficient* for affiliation. In other words, the Corollary provides weaker conditions than Theorem 1. Why then is this corollary helpful? Because it allows us to test in regions where estimation is precise—where we have more observed data.

The results presented here also hold when the interval $[0, 1]$ is partitioned with different sets of numbers $0 < r_{n,1} < r_{n,2} < \dots < r_{n,k-1} < 1$, for each direction $n = 1, \dots, N$. In this case, however, the grid distribution $\mathbb{T}^k(f_{\mathbf{V}}^0)$ will not be

symmetric when $r_{n,j} \neq r_{n',j}$ for $n \neq n'$. Also, as noted previously, symmetry of the regions is unimportant to a test of affiliation: only if symmetry is important in the application need the areas be chosen to generate symmetric grid density functions.

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. 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 $t = 1, \dots, T$ at which N bidders participated by submitting the NT bids $\{\{s_{nt}\}_{n=1}^N\}_{t=1}^T$. At this point, at least two different strategies can be pursued, the first a conventional approach following along the path of Guerre, Perrigne, and Vuong (2000) as well as Li *et al.* (2000,2002), and another which we have pursued. We describe our approach first and then, later in this section, we relate our approach to an implementation of Guerre *et al.* (2000).

Under our approach, we note that affiliation is preserved under a monotonic transformation, so examining a discretization of $g_S^0(\mathbf{s})$, the true probability density function of bids under the hypothesis of expected-profit maximizing equilibrium behaviour, is the same as examining $f_V^0(\mathbf{v})$. 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

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

where \underline{s} is the smallest observed bid and \bar{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.¹ Now, the random vector \mathbf{Y} , which

¹ We know that the support of g_S^0 is strictly positive at \bar{s}^0 , the true upper bound of support

represents the number of outcomes that fall in each of the cells and equals the vector $(Y_1, Y_2, \dots, Y_L)^\top$, follows a multinomial distribution having the joint probability mass function

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

where π_ℓ equals $\Pr(Y_\ell = y_\ell)$, with $y_\ell = 0, 1, \dots, T$, while $\boldsymbol{\pi}$ collects (π_1, \dots, π_L) and lives on the simplex—*viz.*, the set

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

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

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

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

subject to

- 1) the vector $\boldsymbol{\pi}$ lies in the simplex \mathcal{S}_L and
- 2) all of the determinantal inequalities required for TP₂ hold.

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

While the determinantal constraints required for TP₂ are convex sets of the parameters when the sub-matrices are symmetric, they are not for general sub-matrices. However, by taking logarithms of both sides of any general determinantal inequality

$$ab \geq cd,$$

of bids, and we assume that f_V^0 is strictly positive at \underline{v} , so g_S^0 is strictly positive at \underline{s}^0 , the true lower bound of support of bids. Consequently, the sample estimators of the lower and upper bounds of support of \mathbf{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.

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. Of course, the adding-up constraint for the simplex must be finessed—*e.g.*, by considering the following:

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

which gives rise to a convex set. Thus, the problem is almost a linear programme. One needs also to worry about the “logarithm variables” zipping off to minus infinity, which can occur if any of the weights in the objective function (the elements of \mathbf{y}) equal zero. To finesse this numerically, one needs to restrict the logarithm variables to be in the set, say $[-20, 0]$, but this is not without complications for the statistical specification.

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 provide one strategy to deal with this later.

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

$$\sum_{\mathbf{i}} [\mathbf{E}](\mathbf{i}) \{\log[\mathbf{P}](\mathbf{i})\}. \quad (4.1)$$

Now, the following inequalities be met:

$$-20 \leq \log\{[\mathbf{P}](\mathbf{i})\} \leq 0 \text{ and } \sum_{\mathbf{i}} \exp(\log\{[\mathbf{P}](\mathbf{i})\}) \leq 1, \quad (4.2)$$

while symmetry requires the following linear restrictions:

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

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

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

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).

Consider now adapting the research of Guerre *et al.* (2000), which was within the symmetric IPVP, to the APVP, as was done by Li *et al.* (2000,2002). Guerre *et al.* focused on the first-order condition for an equilibrium which, in our notation with affiliation, can be written as

$$v = s + \frac{F[\sigma^{-1}(s)|\sigma^{-1}(s)]\sigma'[\sigma^{-1}(s)]}{f[\sigma^{-1}(s)|\sigma^{-1}(s)]} \quad (4.5)$$

Following Guerre *et al.*, Li *et al.* noted that the term

$$\frac{F[\sigma^{-1}(s)|\sigma^{-1}(s)]\sigma'[\sigma^{-1}(s)]}{f[\sigma^{-1}(s)|\sigma^{-1}(s)]}$$

can be consistently estimated using observed bids. In particular, letting $G(w|s)$ denote the conditional cumulative distribution function of W , the largest of the $(N-1)$ opponents' bids given a bid S , and $g(w|s)$ its corresponding conditional probability density function, then we know that

$$G(w|s) = F[\sigma^{-1}(w)|\sigma^{-1}(s)]$$

and

$$g(w|s) = \frac{f[\sigma^{-1}(w)|\sigma^{-1}(s)]}{\sigma'[\sigma^{-1}(w)]},$$

so

$$\frac{G(w|s)}{g(w|s)} = \frac{F[\sigma^{-1}(w)|\sigma^{-1}(s)]\sigma'[\sigma^{-1}(w)]}{f[\sigma^{-1}(w)|\sigma^{-1}(s)]}.$$

Li *et al.* adapted the estimator of Guerre *et al.* to admit affiliation. Substituting their estimators $\hat{G}(w|s)$ and $\hat{g}(w|s)$ into equation (4.5), Li *et al.* then created the pseudo values according to

$$\hat{v}_{nt} = s_{nt} + \frac{\hat{G}(s_{nt}|s_{nt})}{\hat{g}(s_{nt}|s_{nt})} \quad n = 1, \dots, N; \quad t = 1, \dots, T. \quad (4.6)$$

Unfortunately, the kernel-smoothed estimators that Li *et al.* employed do not necessarily respect the inherent structure that affiliation would bestow on both $\hat{G}(\cdot|\cdot)$ and $\hat{g}(\cdot|\cdot)$, at least in small samples. Thus, the kernel-smoothed estimators that Li *et al.* used to estimate $f_{\mathbf{V}}(\cdot)$ from the estimated pseudo-values $\{\{\hat{v}_{nt}\}_{n=1}^N\}_{t=1}^T$ defined by equation (4.6) need not inherit affiliation.

One can now take the nonparameteric estimate of $f_{\mathbf{V}}^0$ and then discretize it on the interval $[0, 1]^N$, and then apply the steps described above. Of course, as mentioned before, the pre-estimation error in the pseudo-values $\{\{\hat{v}_{nt}\}_{n=1}^N\}_{t=1}^T$ could be a problem. By working directly with the bids, our approach avoids this pre-estimation error. Of course, the kernel-smoothing approach deals effectively with the fact that, in small samples, any given cell $[\mathbf{E}]$ may be empty, which cause numerical problems. In our application, we illustrate one way to deal with the problem that occurs when equi-spaced cells are empty.

4.1. Sampling Distribution of Test Statistic

To get some notion concerning the potential effect that sampling variation can have on our proposed test, consider first the example from figure 2.2. In this figure, there are three parameters (p_{11}, p_{22}, p_{12}) which we shall refer to as a , b and c , respectively. Now, $a + b + 2c$ equals one, so the c implied by the simplex is $[(1 - a - b)/2]$. Also, under TP_2 ,

$$ab \geq \frac{(1 - a - b)^2}{4}.$$

Focus on the equation

$$4ab = 1 - 2a - 2b + 2ab + a^2 + b^2.$$

In this simple example, the “centre” of the simplex for (a, b) , the probability subvector $(0.25, 0.25)$, is one possible distribution within the symmetric IPVP. Note that this point is at the maximum of the arc in figure 2.2. The point $(0.3, 0.3)$, on the other hand, lives in the affiliated set, while the point $(0.2, 0.2)$ lives outside the affiliated set.

Consider the maximum-likelihood estimator (\hat{a}, \hat{b}) in this multinomial model where \hat{c} is derived from the simplex. For large enough T , we can invoke a central

Table 4.1
Probability $\hat{\mathbf{d}} \in \mathcal{A}$ Given \mathbf{d}^0

T	(0.2, 0.2)	(0.25, 0.25)	(0. $\dot{3}$, 0. $\dot{3}$)
1	0.308	0.357	0.442
10	0.229	0.447	0.799
25	0.141	0.466	0.934
100	0.020	0.485	0.999
225	0.002	0.491	1.000

limit theorem, so

$$\sqrt{T} \begin{pmatrix} \hat{a} - a^0 \\ \hat{b} - b^0 \end{pmatrix} \xrightarrow{d} \mathcal{N}[\mathbf{0}_2, \Sigma(a^0, b^0)]$$

where

$$\Sigma(a^0, b^0) = \begin{pmatrix} a^0(1 - a^0) & -a^0b^0 \\ -a^0b^0 & b^0(1 - b^0) \end{pmatrix}.$$

Here, as above, the superscript “0” on a and b denote the true values. Collect (a, b) as \mathbf{d} . Because

$$\hat{\mathbf{d}} \xrightarrow{d} \mathcal{N}\left[\mathbf{d}^0, \frac{1}{T}\Sigma(\mathbf{d}^0)\right],$$

its joint probability density function is approximately

$$h(\hat{\mathbf{d}}|\mathbf{d}^0) = \frac{1}{2\pi \left|\frac{1}{T}\Sigma(\mathbf{d}^0)\right|} \exp\left[-\frac{T(\hat{\mathbf{d}} - \mathbf{d}^0)^\top \Sigma(\mathbf{d}^0)^{-1}(\hat{\mathbf{d}} - \mathbf{d}^0)}{2}\right].$$

What fraction of the $\hat{\mathbf{d}}$ s lives in the set of affiliated distributions \mathcal{A} ? To calculate this, integrate $h(\hat{\mathbf{d}}|\mathbf{d}^0)$ for all affiliated distributions; *i.e.*, find

$$\int_{\hat{\mathbf{d}} \in \mathcal{A}} h(\hat{\mathbf{d}}|\mathbf{d}^0) \, d\hat{\mathbf{d}}. \tag{4.7}$$

In table 4.1, we present estimates of expression (4.7) for various sample sizes when \mathbf{d}^0 is (0.2, 0.2), (0.25, 0.25), and (0. $\dot{3}$, 0. $\dot{3}$). Obviously, sampling error can result in mis-classification, particularly within the symmetric IPVP. On the other hand, a non-affiliated distribution, such as (0.2, 0.2), is rejected often, while an affiliated one, such as (0. $\dot{3}$, 0. $\dot{3}$), is rarely rejected, even when the sample size is relatively small.

Our test of symmetric affiliation is based on the difference between the maximum of the logarithm of the likelihood function $\mathcal{L}(\hat{\mathbf{P}})$ minus the maximum of the logarithm

of the likelihood function under symmetric affiliation $\mathcal{L}(\tilde{\mathbf{P}})$. 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 MTP₂ in binary models—suggests that the statistic

$$2[\mathcal{L}(\hat{\mathbf{P}}) - \mathcal{L}(\tilde{\mathbf{P}})] \quad (4.8)$$

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:

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

where $\mathbf{h} : \mathbb{R}^L \rightarrow \mathbb{R}^J$ 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 $\frac{\partial h_i(\text{vec}[\mathbf{P}])}{\partial \text{vec}[\mathbf{P}]_j}$. 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 $\mathcal{I}(\text{vec}[\mathbf{P}^0])$ Fisher's information matrix which is defined by

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

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

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

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

Theorem 2. Consider the local hypothesis testing problem

$$H_0 : \mathbf{h}(\text{vec}[\mathbf{P}]) \geq \mathbf{0}_J \text{vec}[\mathbf{P}] \in N_\delta(\text{vec}[\mathbf{P}^0])$$

$$H_1 : \text{not } H_0.$$

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

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

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 to verify that the assumptions of Theorem 4.2 in Wolak (1989b) are satisfied; we do this in an appendix.

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, but these bounds are typically quite far apart, particularly when the maximal number of constraints that can bind is quite large, as would be the case in any application to data from field auctions.

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, Romano, and Wolf (1999) as was done by Romano and Shaikh (2008).

4.2. Sensitivity of Test to Choice of k

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 along the lines of research in time-series analysis by Guay, Guerre, and Lazarova (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. Monte Carlo Experiment

Below, we describe a small-scale Monte Carlo experiment used to investigate the 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 four 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 with parameter α ;
- AA) (U_1, U_2, U_3) are asymmetrically-affiliated random variables according to two Frank copulas with parameters α and γ ;
- 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, SA denotes symmetric affiliation, AA denotes asymmetric affiliation, and AN denotes asymmetric non-affiliation.

While Nelsen (1999) has provided a detailed introduction to the theory of copulas, we repeat here some basic facts that important in understanding our experiments. In what follows, for expositional reasons, for the most part, we restrict our discussion to

bivariate copulas, but the results generalize to the case of N variables easily. Given two variables, U_1 and U_2 , a bivariate copula $\mathbb{C}(u_1, u_2)$ is a continuous function having the following properties:

1. $\text{Domain}(\mathbb{C}) = [0, 1]^2$;
2. $\mathbb{C}(u_1, 0) = 0 = \mathbb{C}(0, u_2)$;
3. $\mathbb{C}(u_1, 1) = u_1$ and $\mathbb{C}(1, u_2) = u_2$;
4. \mathbb{C} is a twice-increasing function, so

$$\mathbb{C}(u_1^1, u_2^1) - \mathbb{C}(u_1^0, u_2^1) - \mathbb{C}(u_1^1, u_2^0) + \mathbb{C}(u_1^0, u_2^0) \geq 0$$

for any $u_1^0, u_2^0, u_1^1, u_2^1 \in [0, 1]^2$, such that $u_1^0 \leq u_1^1$ and $u_2^0 \leq u_2^1$.

Because U_1 and U_2 are both defined on the unit interval, they can be viewed as uniform random variables with $\mathbb{C}(u_1, u_2)$ being their joint distribution function. Alternatively, U_1 and U_2 can be viewed as the cumulative distribution functions of two random variables V_1 and V_2 which are collected in the vector \mathbf{V} . In this case, their marginal distribution functions $F_1(v_1)$ and $F_2(v_2)$ are linked to their joint distribution $F_{\mathbf{V}}(v_1, v_2)$ by

$$F_{\mathbf{V}}(v_1, v_2) = \mathbb{C}[F_1(v_1), F_2(v_2)].$$

One attractive feature of copulas is that the marginal cumulative distribution functions do not depend on the choice of the dependence function for the two random variables in question. When one is interested in the association between random variables, copulas are a useful device because the dependence structure is easily separated from the marginal cumulative distribution functions.

From Sklar's Theorem, we know that a unique function \mathbb{C} , the copula, exists such that

$$F_{\mathbf{V}}(v_1, v_2) = \mathbb{C}[F_1(v_1), F_2(v_2)].$$

Under symmetry,

$$F_{\mathbf{V}}(v_1, v_2) = \mathbb{C}[F_V(v_1), F_V(v_2)]$$

because the two marginal cumulative distribution functions are identical. Also, one can nest copulas. For example, consider two appropriately-defined bivariate copulas

\mathbb{C}_1 and \mathbb{C}_2 , then the joint cumulative distribution function of (V_1, V_2, V_3) can be represented

$$F_{\mathbf{V}}(v_1, v_2, v_3) = \mathbb{C}_2(\mathbb{C}_1[F_V(v_1), F_V(v_2)], F_V(v_3)).$$

When V_1 and V_2 are independent, the copula $\mathbb{C}(\cdot)$ is trivial as

$$F_{\mathbf{V}}(v_1, v_2) = F_1(v_1) \times F_2(v_2).$$

Different families of copulas exist. For example, perhaps the best known family is the *Gaussian* family in which the dependence is completely determined by the linear correlation coefficient ρ . Thus, the cumulative distribution function of a standard normal bivariate distribution with linear correlation coefficient ρ is

$$\mathbb{C}_\rho(u_1, u_2) = \Phi_{12,\rho}[\Phi^{-1}(u_1), \Phi^{-1}(u_2)],$$

while the joint probability density function is

$$\mathbf{c}_\rho(u_1, u_2) = \frac{\phi_{12,\rho}[\Phi^{-1}(u_1), \Phi^{-1}(u_2)]}{\phi[\Phi^{-1}(u_1)]\phi[\Phi^{-1}(u_2)]}$$

where

$$\Phi(v) = \int_{-\infty}^v \phi(z) dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^v \exp(-z^2/2) dz$$

and

$$\begin{aligned} \phi_{12,\rho}(v_1, v_2) &= \frac{\partial^2 \Phi_{12,\rho}(v_1, v_2)}{\partial v_1 \partial v_2} \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(v_1^2 + v_2^2 - 2\rho v_1 v_2)\right]. \end{aligned}$$

Affiliation within this bivariate normal family of distributions requires that ρ be non-negative; within multivariate normal distributions affiliation requires that all of the off-diagonal parameters of the variance-covariance matrix be non-negative.

A commonly-used family of copulas is the *Archimedean* family, which is uniquely characterized by its generator function $\zeta(\cdot)$ where

$$\mathbb{C}_\zeta(u_1, u_2, \dots, u_N) = \zeta^{-1}[\zeta(u_1) + \zeta(u_2) + \dots + \zeta(u_N)]. \quad (4.9)$$

Here, $\zeta(\cdot)$ is a convex, decreasing function. Note, too, that $\zeta(1)$ must equal zero and $\zeta^{-1}(u)$ must be zero for any u exceeding $\zeta(0)$. These conditions are both necessary and sufficient for \mathbb{C}_ζ to be a distribution function.

A commonly-used member of the Archimedean family of copulas is the *Frank* copula, which has the following generator function:

$$\zeta(u) = -\log \left[\frac{\exp(-\alpha u) - 1}{\exp(-\alpha) - 1} \right], \quad (4.10)$$

and inverse-generator function

$$\zeta^{-1}(\tau) = -\frac{1}{\alpha} \log (1 + \exp(\tau)[\exp(-\alpha) - 1]). \quad (4.11)$$

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 non-monotonic inverse-generator function of the Frank copula; see example 4.22 in Nelsen (1999, p. 123). The Frank copula has the following N -variate form:

$$\mathbb{C}_\zeta(u_1, \dots, u_N) = -\frac{1}{\alpha} \log \left(1 + \frac{\prod_{i=1}^N [\exp(-\alpha u_i) - 1]}{[\exp(-\alpha) - 1]^{N-1}} \right) \quad \alpha > 0. \quad (4.12)$$

Müller and Scarsini (2005) have characterized various notions of positive dependence, such as MTP_2 , for Archimedean copulas. 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 (1987) 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 Cherubini, Luciano, and Vecchiato (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 $\mathbb{C}_2(w_2|u_1)$, and u_3 is drawn from $\mathbb{C}_3(w_3|u_1, u_2)$ where all the w_i s are independent $U(0, 1)$ draws. We implemented conditional sampling using the parameterization of the Frank copula given in equation (4.12) in conjunction with the generator function defined in equation (4.10) and the inverse-generator function defined in equation (4.11). 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

$$u_2 = -\frac{1}{\alpha} \left(1 + \frac{w_2[1 - \exp(-\alpha)]}{w_2[\exp(-\alpha u_1) - 1] - \exp(-\alpha u_1)} \right); \quad (4.13)$$

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

$$w_3 = D_2^{-1}[\exp(-\alpha u_3) - 1][\exp(-\alpha) - 1] \times \\ ([\exp(-\alpha) - 1] + [\exp(-\alpha u_1) - 1][\exp(-\alpha u_2) - 1])^2 \quad (4.14)$$

where

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

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 4.1, 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 scatterplot looks uniform. In

Figure 4.1
Simulated Data Under Independence, $\alpha = 0$

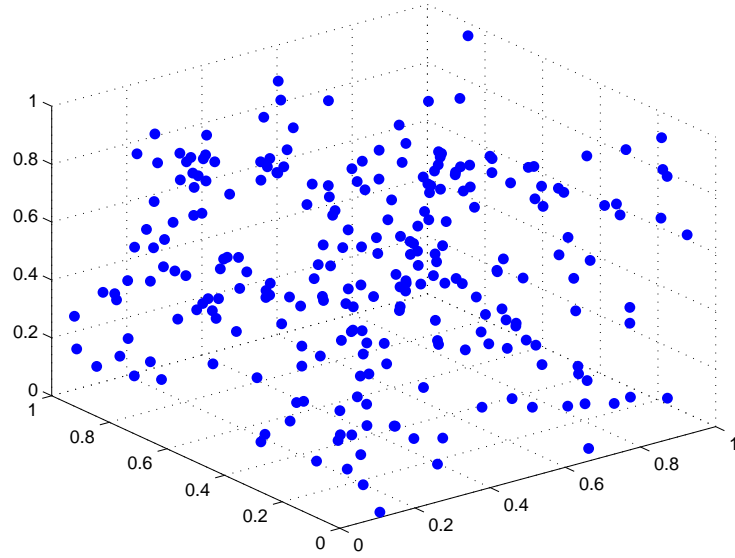


Figure 4.2
Simulated Data from Frank Copula, $\alpha = 2$

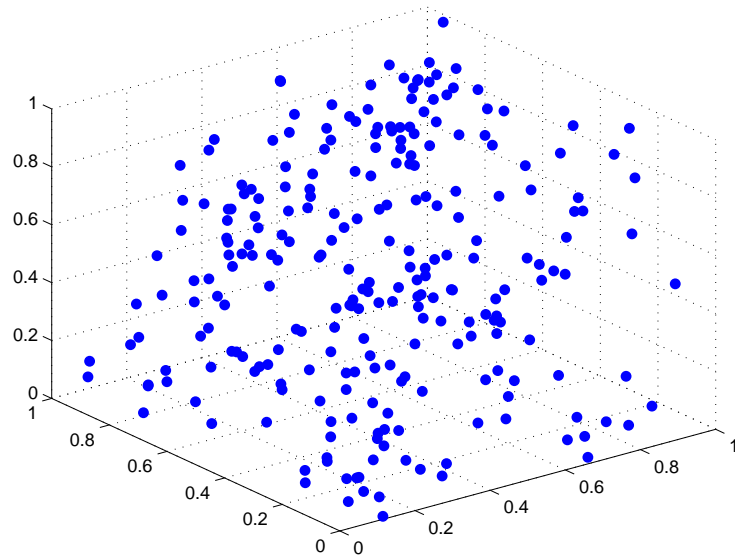


figure 4.2, we present a plot of all the points, when N is three and T is 250, for one replication generated when α is 2, which means only modest affiliation. Our test appears able to distinguish relatively well between weak and modest affiliation, and to detect non-affiliation extremely well.

We next describe the algorithm used to implement conditional sampling using the parameterization of two Frank copulas. Specifically, to generate asymmetrically-affiliated draws (u_1, u_2, u_3) , 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

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

4. form u^* which equals $\mathbb{C}(u_1, u_2; \alpha)$;
5. use w_3 and u^* to calculate

$$u_3 = -\frac{1}{\gamma} \left(1 + \frac{w_3[1 - \exp(-\gamma)]}{w_3[\exp(-\gamma u^*) - 1] - \exp(-\gamma u^*)} \right).$$

The above algorithm yields three asymmetrically-affiliated random draws; this procedure was also repeated either 100 or 250 times for each of 1,000 replications.

In figures 4.3 and 4.4 are presented the frequency distributions of the LR test statistics for SI, SA, AA, and NA when k is three and T is either 100 or 250. The test has relatively high power in the case NA, non-affiliation. As expected, there are fewer rejections with SA, symmetric affiliation, than under SI, weak affiliation (independence). What is more surprising is that the test rejects less frequently under AA, asymmetric affiliation, than under SA, suggesting that it has low power in this direction.

For k of three and T of 100 with symmetric independence, one can calculate the weights $\{\omega(j, J - j, \mathbf{\Pi}^0)\}_{j=0}^J$ in Theorem 2. In figure 4.2, we present the exact

Figure 4.3
SI, SA, AA, NA: $k = 3$, $T = 100$

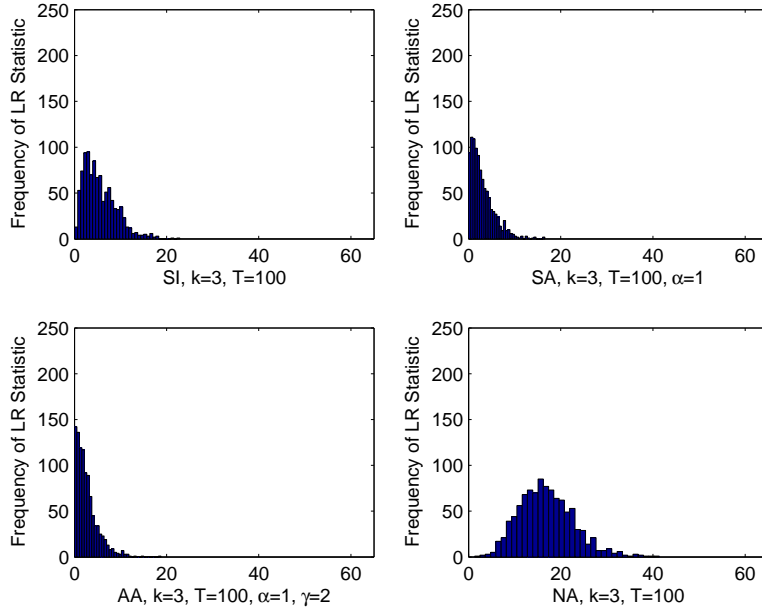


Figure 4.4
SI, SA, AA, NA: $k = 3$, $T = 250$

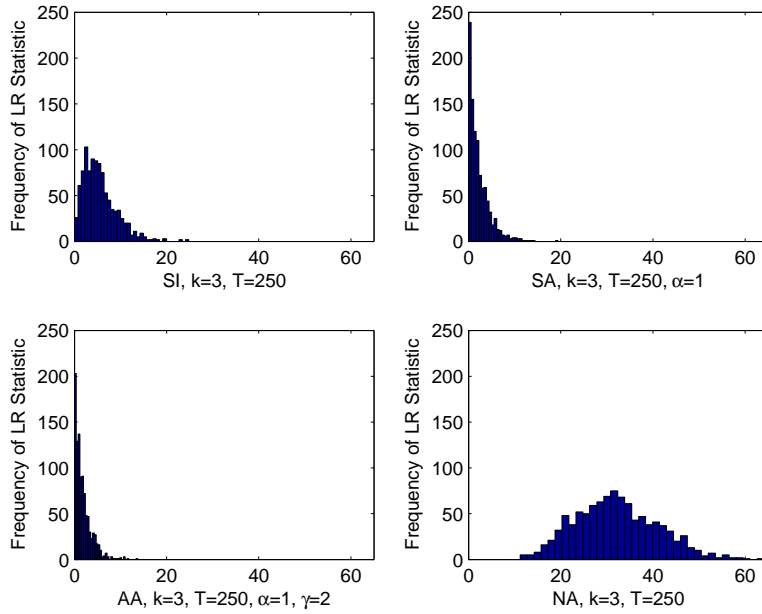
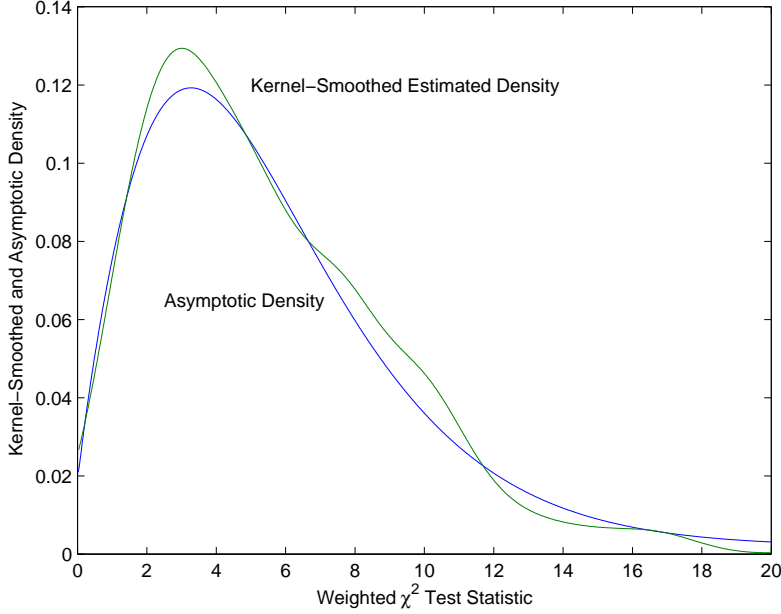


Figure 4.5
Asymptotic and Kernel-Smoothed Density of Test Statistic:
SI, $k = 3$, $T = 100$



probability density function of the asymptotic approximation as well as the kernel-smooth 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.

5. Empirical Application

To demonstrate the feasibility of our testing strategy, we have chosen to implement it using 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. The marginal cost per mile C_{nt} at auction t can be then written as:

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

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 $(\delta_y W_t + \delta_z X_t)$ is log-concave, which is discussed 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 have eschewed investigating issues relating to asymmetries across bidders because introducing F_n s that vary across bidders is computationally quite arduous and yields little pedagogically. Also, we do not have the 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 Li, Paarsch, and Hubbard (2007); a complete description of these data is provided in that paper. In table 5.1, we present the summary descriptive statistics concerning our sample of 834 observations—278 auctions that involved three

Table 5.1
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

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, presumably 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.² We assume that C_{nt} , the cost to bidder n at auction t , can be factored as follows:

$$C_{nt} = \lambda^0(p_t)\varepsilon_{nt} \tag{5.1}$$

² Of course, besides p , it is possible that other covariates, which are common knowledge to all the bidders, exist. Unfortunately, we do not have access to any additional information. Were such information available, then we would condition on it as well.

Figure 5.1
Scatterplot and Nonparametric, LS, and LAD Regressions
Logarithm of Bids versus Logarithm of Engineer's Estimate

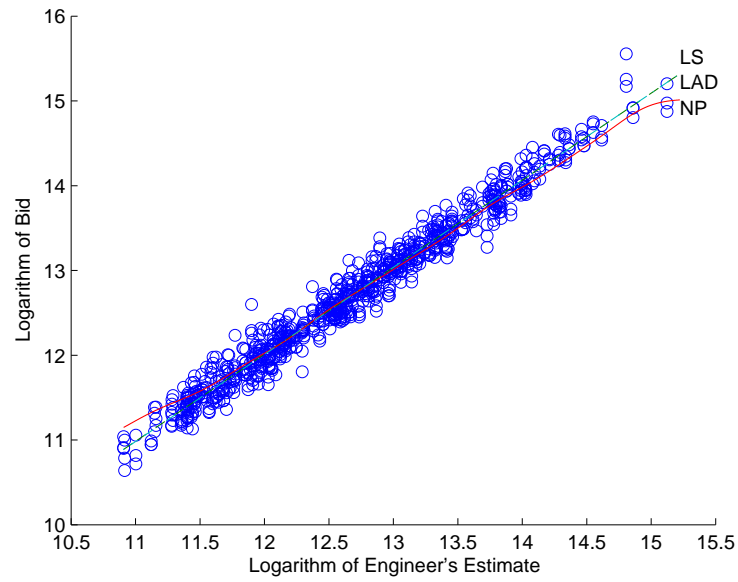
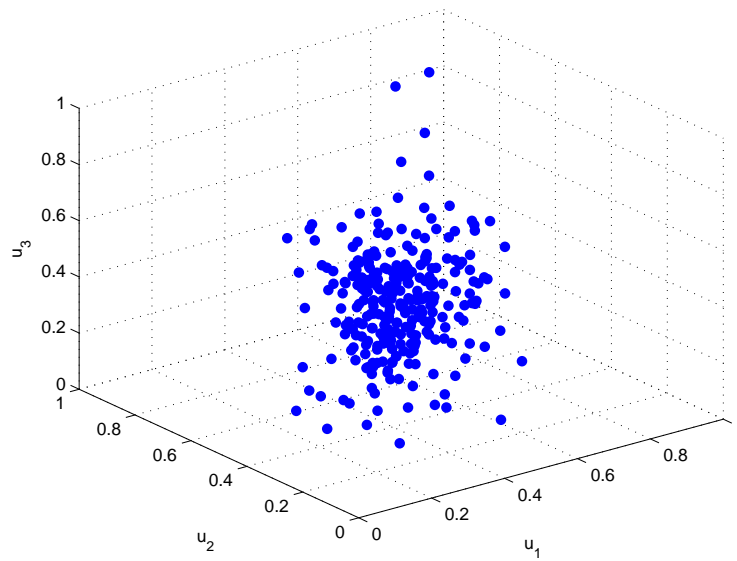


Figure 5.2
Scatterplot of Transformed Fitted LS Residuals



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

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

Another is

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

Under (5.1), 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 5.1. 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 5.2, 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 2.38.³ At size 0.05,

³ 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.

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.

Because a k of two is unusually small, we introduced a symmetric, but non-equispaced, grid distribution like the one depicted in figure 3.1; see figures 5.3, 5.5, and 5.7 for the definitions of the elements of the probability array. As before,

$$\begin{aligned} ab &\geq d^2, \quad bc \geq f^2, \quad df \geq be, \\ dh &\geq b^2, \quad hi \geq g^2, \quad bg \geq fh, \\ eg &\geq f^2, \quad gj \geq i^2, \quad fi \geq cg. \end{aligned}$$

Now, however, the adding-up inequality must be re-written, in this case as

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

We depict the conditional scatterplots for each slice of the probability array in figures 5.4, 5.6, and 5.8. In these scatterplots, the symbol \circ denotes $(\hat{u}_{1t}, \hat{u}_{2t} | \hat{u}_{3t} \in (r_{j-1}, r_j])$, while the symbol \triangle denotes $(\hat{u}_{1t}, \hat{u}_{3t} | \hat{u}_{2t} \in (r_{j-1}, r_j])$, and the symbol ∇ denotes $(\hat{u}_{2t}, \hat{u}_{3t} | \hat{u}_{1t} \in (r_{j-1}, r_j])$.

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 .⁴ 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 .

⁴ 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.

Figure 5.3
Symmetric, Non-Equi-Spaced Grid Distribution:
 U_i versus U_j , given $U_l \in (0.0, 0.4]$

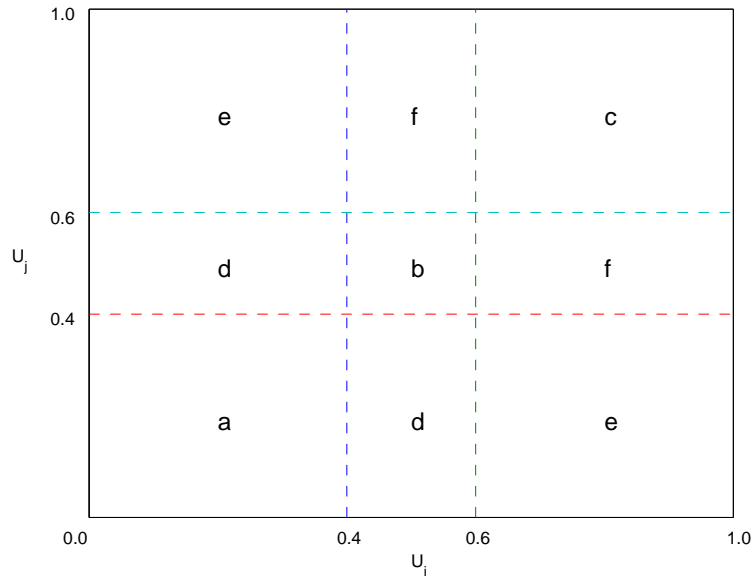
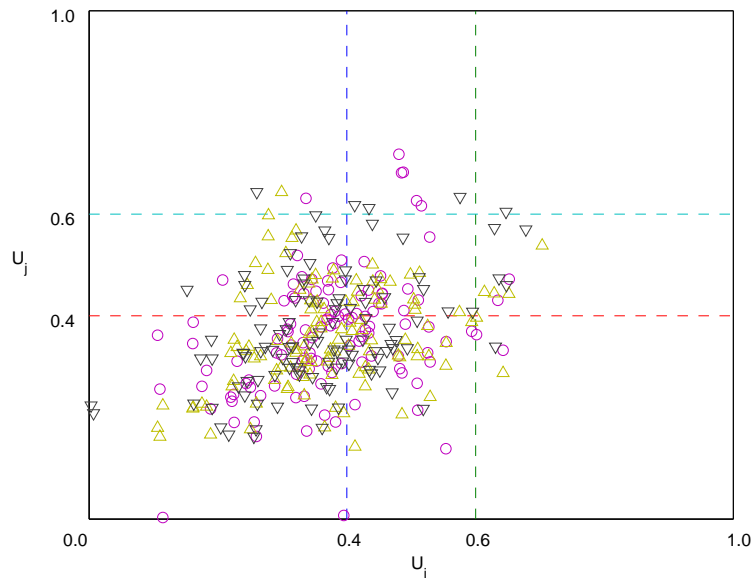


Figure 5.4
Scatterplot of Fitted LS Residuals:
 U_i versus U_j , given $U_l \in (0.0, 0.4]$



6. 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.

Figure 5.5
Symmetric, Non-Equi-Spaced Grid Distribution:
 U_i versus U_j , given $U_l \in (0.4, 0.6]$

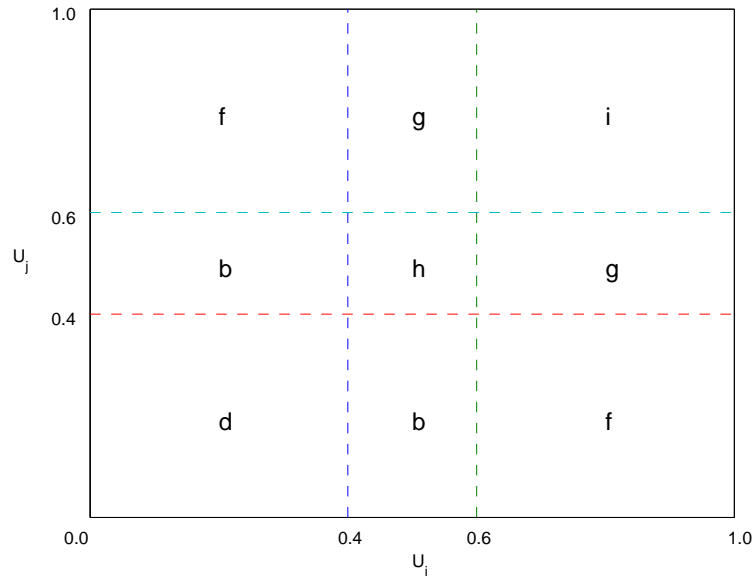
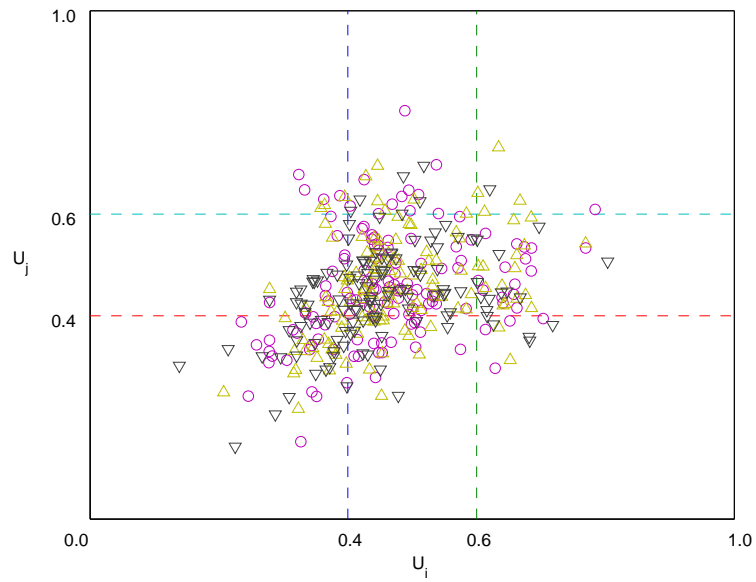


Figure 5.6
Scatterplot of Fitted LS Residuals:
 U_i versus U_j , given $U_l \in (0.4, 0.6]$



A. Appendix

In this appendix, we present the calculations necessary to determine the maximum number of binding constraints required to satisfy MTP_2 as well as the proof of Theorem 2.

A.1. Maximum Number of Binding Constraints

When N is two, for k of four, let us represent the matrix in the following tableau:

	1	2	3	4
1	a	e	h	j
2	e	b	f	i
3	h	f	c	g
4	j	i	g	d

Again, under the convention of section 2, the inequalities can be collected in the following tableau:

	(1,2)	(1,3)	(2,3)	(2,4)	(3,4)	(1,4)
(1,2)	$ab \geq e^2$	$af \geq eh$	$eg \geq bh$	$ei \geq bj$	$hi \geq fj$	$ai \geq ej$
(1,3)	$af \geq eh$	$ac \geq h^2$	$ec \geq fh$	$eg \geq fj$	$gh \geq cj$	$ag \geq hj$
(2,3)	$ef \geq bh$	$ec \geq fh$	$bc \geq f^2$	$bg \geq fi$	$fg \geq ci$	$eg \geq hi$
(2,4)	$ei \geq bj$	$eg \geq fj$	$bg \geq fi$	$bd \geq i^2$	$df \geq gi$	$ed \geq ij$
(3,4)	$hi \geq fj$	$hg \geq cj$	$fg \geq ci$	$df \geq gi$	$cd \geq g^2$	$hd \geq gj$
(1,4)	$ai \geq ej$	$ag \geq hj$	$eg \geq hi$	$ed \geq ij$	$hd \geq gj$	$ad \geq j^2$

In general, to obtain non-bold elements from bold ones, assume the inequalities in the bold cells and do the following: if a cell is between two cells with previously obtained inequalities (either diagonally or horizontally or vertically), then its inequality is obtained from these two cells.⁵ In this fashion, we complete the tableau, with the exception of row (1,4). Row (1,4) is obtained by combining rows (1,2) and (2,4).

Consider now the case of N equal three when k is two. A symmetric density can be represented as

1	1	2
1	a	c
2	c	b
2	1	2
1	c	b
2	b	d

The inequalities are $ab \geq c^2$ and $cd \geq b^2$ in each of the tableaux indexed by the bold numbers **1** and **2**, which denote the value of player 3, plus $ad \geq bc$ for the three-dimensional

⁵ By diagonally, we mean the case that the cell is to the right and above cells already obtained. For example, $(1,3) \times (1,3)$ derives from $(1,3) \times (1,2)$ and $(2,3) \times (1,3)$.

Figure 5.7
Symmetric, Non-Equi-Spaced Grid Distribution:
 U_i versus U_j , given $U_l \in (0.6, 1.0]$

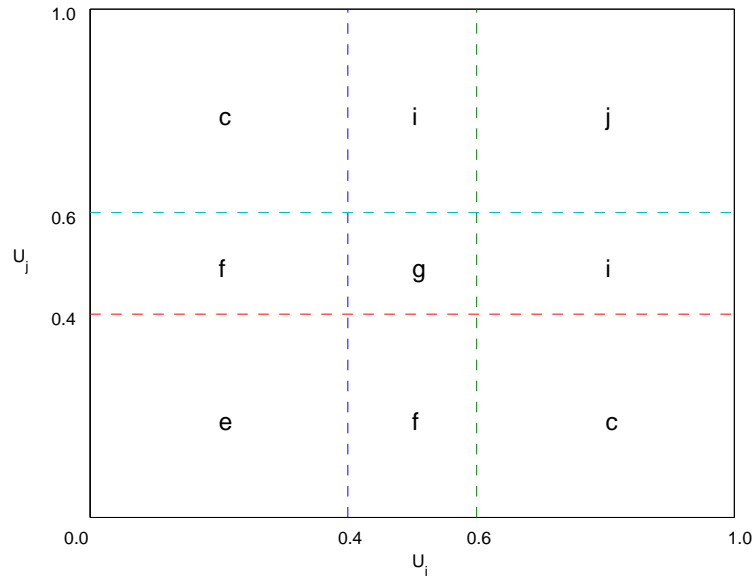
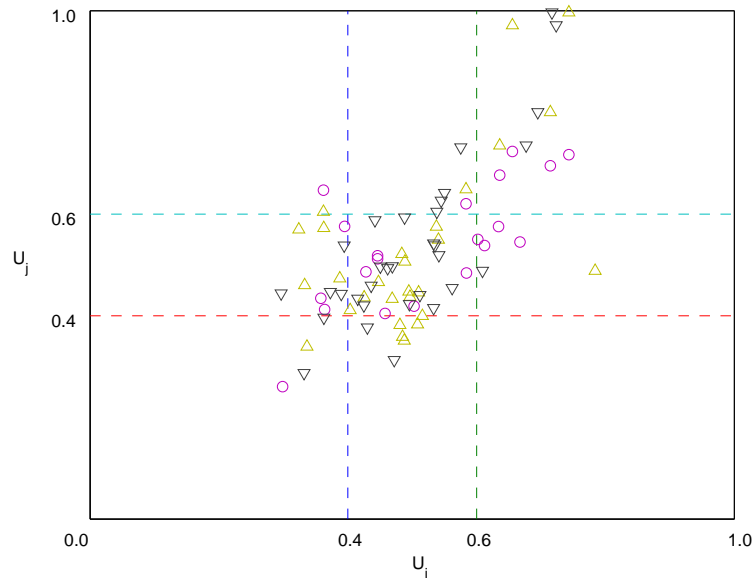


Figure 5.8
Scatterplot of Fitted LS Residuals:
 U_i versus U_j , given $U_l \in (0.6, 1.0]$



(3D) combinations. However, it is sufficient to impose just $ab \geq c^2$ and $cd \geq b^2$ because the others follow from these two.

For the cases below, it is also useful to consider the asymmetric case when N is three and k is two, so

1	1	2
1	a	c
2	d	b
2	1	2
1	e	g
2	h	f

Note that the cube has six faces, and the inequalities for the faces are:

$$ab \geq cd, \quad ef \geq gh, \quad ah \geq de, \quad ag \geq ce, \quad cf \geq bg, \quad \text{and} \quad df \geq bh.$$

We call these the two-dimensional (2D) inequalities because they correspond to faces of the cube. In addition to these, however, one need also consider the following 3D inequalities:

$$af \geq ch, \quad af \geq dg, \quad \text{and} \quad af \geq be.$$

Now consider the case of N equal three and k equal three. A total of $\binom{3+3-1}{3-1}$ or ten different variables exist as depicted in the following three tableaux:

1	1	2	3
1	a	d	e
2	d	b	f
3	e	f	c
2	1	2	3
1	d	b	f
2	b	h	g
3	f	g	i
3	1	2	3
1	e	f	c
2	f	g	i
3	c	i	j

Because of symmetry, the 2D inequalities (for just one level of the the third variable) can be represented in the following three tableaux:

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

2	(1,2)	(1,3)	(2,3)
(1,2)	$dh \geq b^2$	$dg \geq bf$	$bg \geq bg$
(1,3)	$dg \geq bf$	$di \geq f^2$	$bi \geq fg$
(2,3)	$bg \geq fh$	$bi \geq fg$	$hi \geq g^2$
3	(1,2)	(1,3)	(2,3)
(1,2)	$eg \geq f^2$	$ei \geq cf$	$fi \geq cg$
(1,3)	$ei \geq cf$	$ej \geq c^2$	$fj \geq ci$
(2,3)	$fi \geq cg$	$fj \geq ci$	$gj \geq i^2$

While not difficult, it is indeed tedious to verify that the above nine inequalities in bold are independent. The 3D inequalities, which correspond to combining two levels of the third variable, result in the following three tableaux:

1,2	(1,2)	(1,3)	(2,3)
(1,2)	$ah \geq bd$		
(1,3)	$ag \geq be, ag \geq fd$	$ai \geq ef, ai \geq dc$	
(2,3)	$dg \geq he, dg \geq bf$	$di \geq ge, di \geq f^2, di \geq bc$	$bi \geq fg, bi \geq hc$
2,3	(1,2)	(1,3)	(2,3)
(1,2)	$gd \geq bf, gd \geq eh$		
(1,3)	$di \geq f^2, di \geq bc, di \geq eg$	$dj \geq cf, dj \geq ei$	
(2,3)	$bi \geq ch, bi \geq fg$	$bj \geq if, bj \geq cg$	$hj \geq ig$
1,3	(1,2)	(1,3)	(2,3)
(1,2)	$ag \geq df, ag \geq be$		
(1,3)	$ai \geq ef, ai \geq cd$	$aj \geq ce$	
(2,3)	$di \geq f^2, di \geq eg, di \geq bc$	$dj \geq cf, dj \geq ie$	$bj \geq if, bj \geq cg$

It is also easy (but tedious) to verify that all 3D inequalities follow from the 2D inequalities. Note that, in principle, all cells in the above tableaux should have three inequalities, but symmetry reduces this to two or, in some cases, just one.

A.2. Proof of Theorem 2

We need to verify that our framework satisfies the regularity conditions of Theorem 4.2 in Wolak (1989b). First, $\mathbf{h} \in \mathbb{R}^L \rightarrow \mathbb{R}^J$ is linear for all but one condition, which is the adding-up condition—*viz.*, that the integral is one or the array defines a density. This is true because the affiliation inequalities are log-linear and \mathbf{h} takes the logarithms as inputs. Thus, \mathbf{h} is well-behaved and satisfies the regularity conditions stated in section 2 of Wolak (1989b). It remains to verify the assumptions on pp.31–33 of Wolak (1989b).

The following assumptions involve only continuity and are trivially satisfied within our framework:

Assumption 1. For all T , $T^{-1}\mathcal{L}(\mathbf{P})$ is a continuous function of \mathbf{P} .⁶

⁶ Note that Wolak (1989b) denotes the sample size T by n and the parameters \mathbf{P} by β .

Assumption 3. The partial derivatives $\frac{\partial \mathcal{L}(\text{vec}[\mathbf{P}])}{\partial \text{vec}[\mathbf{P}]}$ exist and are continuous with probability one.

Assumption 4. The second partial derivatives $\frac{\partial^2 \mathcal{L}(\text{vec}[\mathbf{P}])}{\partial \text{vec}[\mathbf{P}] \partial \text{vec}[\mathbf{P}]^\top}$ exist and are continuous with probability one.

Now, we consider the other assumptions.

Assumption 2. $T^{-1} \mathcal{L}(\mathbf{P})$ converges almost surely to a function $\mathcal{L}_\infty(\mathbf{P}, \mathbf{P}^0)$ which is $\mathcal{E}_{\mathbf{P}^0}[\mathcal{L}(\mathbf{P})]$ for all \mathbf{P} . The function $\mathcal{L}_T(\mathbf{P}, \mathbf{P}_T^0)$ is then $\mathcal{E}_{\mathbf{P}_T^0}[\log g_Y(\mathbf{P})]$ has a unique maximum at \mathbf{P}_T^0 . In addition, as $T \rightarrow \infty$, the function $\mathcal{L}_T(\mathbf{P}, \mathbf{P}_T^0)$ converges to $\mathcal{L}_\infty(\mathbf{P}, \mathbf{P}^0)$, which has a unique local maximum at \mathbf{P} equal \mathbf{P}^0 .

Convergence derives from the strong law of large numbers. Uniqueness of the maximum derives from the fact that $\mathcal{L}_T(\mathbf{P}, \mathbf{P}_T^0)$ is less than $\mathcal{L}_T(\mathbf{P}_T^0, \mathbf{P}_T^0)$ for $\mathbf{P} \neq \mathbf{P}_T^0$, and $\mathcal{L}_\infty(\mathbf{P}, \mathbf{P}^0)$ is less than $\mathcal{L}_\infty(\mathbf{P}^0, \mathbf{P}^0)$ for $\mathbf{P} \neq \mathbf{P}^0$, which holds because there exist realizations \mathbf{y} such that $\mathcal{L}(\mathbf{y}, \boldsymbol{\pi}) \neq \mathcal{L}(\mathbf{y}, \boldsymbol{\pi}^0)$ for $\boldsymbol{\pi} \neq \boldsymbol{\pi}^0$; see, for example, the comments following assumption 6A in Silvey (1959), p.391.

The following two assumptions are consequences of the definition of

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

In section 4, we denoted $\mathcal{I}(\text{vec}[\mathbf{P}]; \text{vec}[\mathbf{P}^0])$ by $\mathcal{I}(\text{vec}[\mathbf{P}^0])$ as a short-hand; here, we explicitly distinguish between \mathbf{P} and \mathbf{P}^0 .

Assumption 5. The matrix

$$-\frac{1}{T} \frac{\partial^2 \mathcal{L}(\text{vec}[\mathbf{P}])}{\partial \text{vec}[\mathbf{P}] \partial \text{vec}[\mathbf{P}]^\top}$$

converges almost surely and uniformly for all $\text{vec}[\mathbf{P}]$ to the matrix

$$\mathcal{I}(\text{vec}[\mathbf{P}]; \text{vec}[\mathbf{P}^0]).$$

Assumption 6. The matrix $\mathcal{I}(\text{vec}[\mathbf{P}]; \text{vec}[\mathbf{P}^0])$ is positive definite.

The asymptotic normality of the estimator (assumption 7 below) is verified by a standard application of an appropriate central limit theorem.

Assumption 7. The vector

$$\frac{1}{T} \frac{\partial \mathcal{L}(\text{vec}[\mathbf{P}^0])}{\partial \text{vec}[\mathbf{P}]}, \text{vec}[\mathbf{P}^0] \in N_{\delta_T}(\text{vec}[\mathbf{P}^0])$$

where δ_T equals $O(T^{-1/2})$ is asymptotically normal with the mean zero vector and covariance matrix $\mathcal{I}(\text{vec}[\mathbf{P}^0])$.

As Wolak (1989b) states, the above conditions can be directly verified by observing that assumptions 1–12 of Silvey (1959) are satisfied.

We shall not repeat here Wolak’s assumption 8—the Abadie constraint qualification condition—because it requires the introduction of much additional notation. Essentially, this assumption requires that the cone of tangents of the set $\mathbf{h}(\text{vec}[\mathbf{P}]) \geq \mathbf{0}_J$ be the same as the intersection of directions \mathbf{d} such that $\frac{\partial h_i(\text{vec}[\mathbf{P}])}{\partial \mathbf{P}} \mathbf{d} \geq 0$ (if i is such that $h_i(\text{vec}[\mathbf{P}]) \geq 0$) and $\frac{\partial h_i(\text{vec}[\mathbf{P}])}{\partial \text{vec}[\mathbf{P}]} \mathbf{d} = 0$ (if i is such that $h_i(\text{vec}[\mathbf{P}]) = 0$), where the derivatives are taken at $\hat{\mathbf{P}}$, the estimated \mathbf{P} . There is only one equality constraint, which is also the only one that is non-linear (it is the condition that the integral is one). The directions in the cone for this constraint is the same as the direction given by the above condition. Since the other conditions are linear, then the equivalence of the directions is immediate. Thus, we conclude verifying the conditions.

B. Bibliography

- Andrews, D., 2000. Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica* 68, 399–405.
- Andrews, D., Guggenberger, P., 2007. Validity of subsampling and “plug-in asymptotic” inference for parameters defined by moment inequalities. Typescript, Department of Economics, Yale University.
- Bartolucci, F., Forcina, A., 2000. A likelihood ratio test for MTP_2 within binary variables. *Annals of Statistics* 28, 1206–1218.
- Bishop, Y., Fienberg, S., Holland, P., 1975. *Discrete Multivariate Analysis*. MIT Press, Cambridge.
- Brendstrup, B., Paarsch, H., 2007. Semiparametric identification and estimation multi-object, English auctions. *Journal of Econometrics* 141, 84–108.
- Bugni, F., 2008. Bootstrap inference in partially identified models. Typescript, Department of Economics, Northwestern University.
- Canay, I., 2008. EL inference for partially identified models: large deviations optimality and bootstrap validity. Typescript, Department of Economics, University of Wisconsin–Madison.
- Cherubi, U., Luciano, E., Vecchiato, W., 2004. *Copula Methods in Finance*. John Wiley & Sons, New York.
- de Castro, L., 2007. Affiliation, equilibrium existence, and the revenue ranking of auctions. Typescript, Department of Economics, Carlos III University.
- de Castro, L., 2008. Grid distributions to study single object auctions. Typescript, Department of Economics, University of Illinois.
- Genest, C., 1987. Frank’s family of bivariate distributions. *Biometrika* 74, 549–555.
- Guay, A., Guerre, E., Lazarova, S., 2008. Optimal adaptive detection of small correlation functions. Typescript, Department of Economics, Queen Mary University London.
- Guerre, E., Perrigne, I., Vuong, Q., 2000. Optimal nonparametric estimation of first-price auctions. *Econometrica* 68, 525–574.
- Karlin, S., 1968. *Total Positivity, Volume 1*. Stanford University Press, Stanford.
- Kodde, D., Palm, F., 1986. Wald criteria for jointly testing equality and inequality restrictions. *Econometrica* 54, 1243–1248.
- Krishna, V., 2002. *Auction Theory*. Academic Press, San Diego.
- Laffont, J.-J., Vuong Q., 1996. Structural analysis of auction data. *American Economic Review* 86, 414–420.
- Lehmann, E., 1966. Some concepts of dependence. *Annals of Mathematical Statistics* 37, 1137–1153.

- Lehmann, E., 1983. *Theory of Point Estimation*. John Wiley & Sons, New York.
- Li, T., 2005. Econometrics of first-price auctions with entry and binding reservation prices. *Journal of Econometrics* 126, 173–200.
- Li, T., Paarsch, H., Hubbard, T., 2007. Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. Typescript, Department of Economics, University of Iowa.
- Li, T., Perrigne, I., Vuong, Q., 2000. Conditionally independent private information in OCS wildcat auctions. *Journal of Econometrics* 98, 129–161.
- Li, T., Perrigne, I., Vuong, Q., 2002. Structural estimation of the affiliated private value model. *RAND Journal of Economics* 33, 171–193.
- Milgrom, P., 1981a. Good news and bad news: representation theorems and applications. *Bell Journal of Economics* 12, 380–391.
- Milgrom, P., 1981b. Rational expectations, information acquisition, and competitive bidding. *Econometrica* 49, 921–943.
- Milgrom, P., Weber, R., 1982. A theory of auctions and competitive bidding. *Econometrica* 50, 1089–1122.
- Müller, A., Scarsini, M., 2005. Archimedean copulae and positive dependence. *Journal of Multivariate Analysis* 93, 434–445.
- Nelsen, R., 1999. *An Introduction to Copulas*. Springer, New York.
- Politis, D., Romano, J., Wolf, M., 1999. *Subsampling*. Springer, New York.
- Romano, J., Shaikh, A., 2008. Inference for identifiable parameters in partially identified econometric models. *Journal of Statistical Planning and Inference* 138, 2786–2807.
- Silvapulle, M., Sen, P.K., 2004. *Constrained Statistical Inference: Inequality, Order and Shape Restrictions*. Wiley–Interscience, New York.
- Silvey, S., 1959. The Lagrangian multiplier test. *Annals of Mathematical Statistics* 30, 389–407.
- Wolak, F., 1987. An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association* 82, 782–793.
- Wolak, F., 1989a. Testing inequality constraints in linear econometric models. *Journal of Econometrics* 41, 205–235.
- Wolak, F., 1989b. Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models. *Econometric Theory* 5, 1–35.
- Wolak, F., 1991. The local nature of hypothesis tests involving inequality constraints in nonlinear models. *Econometrica* 59, 981–995.