1993

Identification, Estimation, and Testing in Empirical Models of Auctions Within the Independent Private Values Paradigm

Stephen G. Donald

Harry J. Paarsch

ISSN: 0318-725X
ISBN: 0-7714-1591-5

RESEARCH REPORT 9319

Identification, Estimation, and Testing in Empirical Models of Auctions Within the Independent Private Values Paradigm

by

Stephen G. Donald
and
Harry J. Paarsch

December 1993

Department of Economics

Social Science Centre

University of Western Ontario

London, Ontario, CANADA

N6A 5C2

IDENTIFICATION, ESTIMATION, AND TESTING IN EMPIRICAL MODELS OF AUCTIONS WITHIN THE INDEPENDENT PRIVATE VALUES PARADIGM*

Stephen G. Donald
Department of Economics
Boston University
Boston, Massachusetts 02215
U.S.A.
Telephone: (617) 353-4824
Internet Address: ocker@acs.bu.edu

and

Harry J. Paarsch
Department of Economics
University of Western Ontario
London, Ontario N6A 5C2
Canada
Telephone: (519) 661-3494
Internet Address: paarsch@vaxr.sscl.uwo.ca

December 1993

Abstract

Recent advances in the application of game theory to the study of auctions have spawned a growing empirical literature involving both experimental and field data. In this paper, we focus upon four different mechanisms (the Dutch, English, first-price sealed-bid, and Vickrey auctions) within one of the most commonly used theoretical models (the independent private values paradigm) to investigate issues of identification, estimation, and testing in structural econometric models of auctions.

---

*This paper is a substantial revision of research contained in University of Western Ontario Research Reports 92–11 and 92–16. Paarsch is grateful to the SSHRC of Canada for continued financial support. Both authors would like to thank Joel Horowitz and two anonymous referees for helpful comments and useful suggestions.

1. Motivation and Introduction

Over the past thirty-five years, economists have made considerable progress in understanding the factors influencing the prices realized from goods sold at auction. For example, they have found that the seller's expected revenue depends upon the type of auction employed, the rules that govern bidding, the number of potential bidders, the information available to the potential bidders, and the attitudes of the bidders toward risk. Perhaps the most remarkable and surprising result to emerge from this research was one first derived by Vickrey (1961). Vickrey showed that if potential bidders are risk neutral with respect to winning the auction, if each potential bidder knows only his own valuation for the object, and if each valuation is an independent draw from a common distribution of valuations, then four quite different institutions yield the same average revenue to the seller. The four institutions are the oral ascending-price (English) auction, the oral decending-price (Dutch) auction, the first-price sealed-bid auction, and the second-price sealed-bid (Vickrey) auction. This result, known as the revenue equivalence proposition (REP), is of considerable practical interest to both buyers and sellers at auctions.

Vickrey's environment is often referred to as the independent private values paradigm (IPVP). The stark theoretical predictions concerning equilibrium outcomes at auctions within it have invited considerable empirical investigation using both experimental and field data. Coppinger, Smith, and Titus (1980) were the first to examine Vickrey's (1961) claims concerning Dutch, English, first-price sealed-bid, and Vickrey auctions using experimental data, although researchers before them had examined some of Vickrey's other propositions using experimental methods. In a series of subsequent papers (cited in Smith [1982]), Smith and other co-authors examined the implications of other factors, particularly risk aversion, on bidding behaviour. Because timber is one of the few commodities to be sold simultaneously in the same market using different auction mechanisms, researchers such as Mead (1967), Johnson (1979), and Hansen (1986) have employed regression methods to examine the REP using field data concerning timber sales.[1]

Although either experimental or field evidence concerning the REP appears mixed, some support for "rational" behaviour appears to exist. A problem with much of the empirical evidence discussed above is that most researchers have tested only one implication of the IPVP: a restriction upon the first moment of revenues. Revenue equivalence might fail to be rejected for alternatives

[1] Both English and first-price sealed-bid auctions have been used to sell timber.

that are close because of low power. A structural econometric framework for interpreting experimental and field data is required to address questions concerning the applicability of the IPVP when the REP is not rejected and to investigate directions in which behaviour differs from what is predicted theoretically.

Another goal of recent research, such as the work of Riley and Samuelson (1981), has been to construct optimal selling mechanisms. The literature concerning mechanism design has been criticized as lacking practical value because the optimal mechanism depends upon random variables whose distributions are typically unknown to the designer. At auctions within the IPVP (the most commonly used framework within which to investigate mechanism design), potential bidders' equilibrium bidding strategies are increasing functions of their valuations. Thus, it is possible to estimate the underlying probability law of valuations using bid data from a cross-section of auctions. Based upon such an estimate, one can then derive an estimate of the optimal auction. To carry out such an exercise, however, also requires a structural econometric framework.

In this paper, we investigate the structural econometrics of auctions within the IPVP, focusing upon the empirical analysis of data from the four types of auctions Vickrey examined. The paper is in eight more parts. In section 2, we introduce the IPVP and define the equilibrium strategies at the four auctions within that paradigm, while in section 3 we discuss the data that are typically available from auctions, and the minimum data required to specify the data generating processes implicit in theoretical models of auctions. In section 4, we develop the data generating processes implicit in the equilibrium strategies, while in section 5 we investigate parameter identification in the empirical specifications. Parameter estimation by the method of maximum likelihood is investigated in section 6,[2] while we examine hypothesis testing within this empirical framework in section 7. The results of a small Monte Carlo experiment are reported in section 8, and we summarize and conclude the paper in section 9.

2. Equilibrium Bidding Strategies within the Independent Private Values Paradigm

Consider an indivisible object that is to be sold at auction by one seller to N potential bidders. Suppose that the seller announces a minimum (reserve) price v_0. Assume that the i^{th} potential

[2] Laffont, Ossard, and Vuong (1991) have examined the method of simulated non-linear least squares in empirical specifications of Dutch and first-price sealed-bid auctions. The performance of their method was compared to the method of maximum likelihood discussed below by Paarsch (forthcoming). Gallant and Tauchen (1992) have examined the generalized method of moments in models that appear feasible for data from English auctions.

bidder has a valuation v_i for the object, which is known to him, but not to the $M = (N - 1)$ other potential bidders. Heterogeneity across potential bidders in valuations is ascribed to independent draws for a random variable V having a probability density function $f(v)$, with cumulative distribution function $F(v)$, and having support upon the interval $[\underline{v}, \overline{v}]$. Assume that v_0 exceeds $\underline{v}$, and that the number of potential bidders N and the distribution function $F(v)$ are common knowledge. Consider the case when potential bidders are potentially risk averse with respect to winning the object, having von Neumann-Morgenstern utility functions that fall within the hyperbolic absolute risk aversion (HARA) family.[3] Thus, for an uncertain prospect y utility $U(y)$ is

$$U(y) = \eta y^{1/\eta}$$

where $\eta \geq 1$, with $\eta = 1$ being the risk-neutral case.

2.1. English and Vickrey Auctions

English and Vickrey auctions are strategically equivalent. To see this, consider first an English auction where the seller sets the reserve price at v_0, and then lets it rise more or less continuously as long as at least two bidders are willing to pay the announced price. Each bidder indicates willingness to pay by some action that is observable not only to the seller but also to the other bidders. Within this setting, the dominant strategy for each potential bidder $\beta(v)$ is to participate as long as the announced sale price does not exceed that bidder's valuation. Thus,

$$\beta(v) = v \quad v \geq v_0. \tag{2.1}$$

Notice that risk aversion plays no role in determining the dominant bidding strategy at English auctions.

At a Vickrey auction, each participant must submit a sealed bid, with the highest bidder winning the auction but paying only what his next-nearest opponent has bid.[4] The dominant strategy at this auction is also to follow (2.1). To see why, consider what would happen if participant i were to bid less than his valuation v_i. He would then risk losing a non-negative rent, but gain

[3] This type of risk aversion was considered by Smith and his co-authors. Other parametric forms of risk aversion, such as constant absolute risk aversion, could also be implemented, but have not been considered below because they do not nest the risk neutral case.

[4] We shall assume that if only one potential bidder submits a bid, then that bidder gets the object at the reserve price.

nothing. Conversely, to bid more than v_i would be to reduce the amount of rent garnered if he won. Thus, bidding v_i when it exceeds v_0 is optimal from the perspective of bidder i.

Notice that the winner is the potential bidder with the highest valuation, provided it exceeds v_0. Letting $V_{(i:N)}$ denote the i^{th} highest order statistic for a sample of size N from the distribution of V and denoting the winning bid by W, the above bidding behaviour implies that W equals $V_{(2:N)}$ whenever W exceeds v_0.

When the reserve price v_0 exceeds $\underline{v}$, the number of actual bidders (participants) at an auction P is endogenous, and typically less than the number of potential bidders N. Only those potential bidders with valuations exceeding the reserve price v_0 participate. To calculate the number of participants at an auction, introduce the indicator variable

$$I_i = \begin{cases} 1 & \text{if } V_i \geq v_0, \\ 0 & \text{otherwise.} \end{cases}$$

The number of participants at an auction is then

$$P = \sum_{i=1}^{N} I_i.$$

Note that P is distributed binomially with parameters N and $\Pr[I_i = 1] = [1 - F(v_0)]$.

2.2. Dutch and First-Price Sealed-Bid Auctions

Dutch and first-price sealed-bid auctions are also strategically equivalent. To see this, consider the decision problem faced by a participant at a Dutch auction. At Dutch auctions, the price starts high and then falls continuously until some one stops it. Depending upon the participant's valuation for the object, he must decide at what point to stop the auction by signalling willingness to pay the existing price. This situation is identical to that faced by a participant at a first-price sealed-bid auction who must decide how high to bid for the object.

To analyze this case, we shall focus upon symmetric Bayesian-Nash equilibria. To construct the equilibrium, suppose that the M opponents of player i are using a common bidding rule $\sigma(v)$ which is increasing and differentiable in v. Since valuations are modelled as independent draws from a common distribution, the probability of player i winning with strategy s_i equals the probability that every other opponent bids lower because each has a lower valuation

$$F\big(\sigma^{-1}(s_i)\big)^M.$$

Here $\sigma^{-1}(s_i)$ denotes the inverse of the bid function. Given that player i's valuation v_i is determined before the bidding, that player's choice of strategy s_i has only two effects upon his expected utility

$$U(v_i - s_i)F\big(\sigma^{-1}(s_i)\big)^M = \eta(v_i - s_i)^{1/\eta}F\big(\sigma^{-1}(s_i)\big)^M.$$

The higher is s_i, the higher is player i's probability of winning the auction $F\big(\sigma^{-1}(s_i)\big)^M$, but the lower is the pay-off following a win $\eta(v_i - s_i)^{1/\eta}$. Maximizing behaviour implies that the optimal bidding strategy solves the first-order condition

$$-(v_i - s_i)^{(\eta-1)/\eta}F\big(\sigma^{-1}(s_i)\big)^M + M\eta(v_i - s_i)^{1/\eta}f\big(\sigma^{-1}(s_i)\big)F\big(\sigma^{-1}(s_i)\big)^{M-1}\frac{d\sigma^{-1}(s_i)}{ds_i} = 0. \quad (2.2)$$

Symmetry among bidders implies

$$s_i = \sigma(v_i). \quad (2.3)$$

Substituting (2.3) into (2.2), recalling that $d\sigma^{-1}(s_i)/ds_i = 1/\sigma'(v_i)$, and requiring (2.2) to hold for all feasible v_is, yields the following differential equation for σ:

$$\sigma'(v)F(v)^M + M\eta\sigma(v)f(v)F(v)^{M-1} = M\eta v f(v)F(v)^{M-1}. \quad (2.4)$$

Integrating (2.4), imposing the boundary condition $\sigma(v_0) = v_0$, yields

$$\sigma(v) = v - \frac{\int_{v_0}^{v} F(\xi)^{M\eta}\, d\xi}{F(v)^{M\eta}}. \quad (2.5)$$

Notice that the condition determining participation at Dutch and first-price sealed-bid auctions $(V \geq v_0)$ is identical to that at English and Vickrey auctions.

The winner at Dutch and first-price sealed-bid auctions will be the player with the highest valuation $V_{(1:N)}$. Because the winning bid function is monotonic in $V_{(1:N)}$, its distribution is related to that of the largest order statistic for a sample of size N from the distribution of V.

3. Data Availability and Requirements

The type of data available will typically determine whether a particular structural econometric analysis of an auction is possible. With experiments, data problems can be avoided by efficient design. This is not the case with field data. The way field data are sampled will typically determine whether structural econometric work can be completed. Thus, in this section we consider what kind of field data are typically available, and what are the minimum data requirements.

In the above characterization of English auctions, where the exit of each participant is observed, one can measure participants' valuations from their final bids. In addition, both the reserve price v_0 and the number of participants P can be observed. Other information must be typically used in order to find the number of potential bidders N. When exit is unobserved at English auctions because bids are only observed when a participant cries out, the last recorded bid for each participant only provides a bound upon the valuation of that participant. In particular, a participant's valuation is greater than or equal to his last observed bid.

At Dutch auctions, one may observe the number of participants P, but only one bid as only the winning bid is ever revealed. Whether the reserve price is observed will vary. As in the case of English auctions, the number of potential bidders N is typically unobserved.

The use of Vickrey auctions is rare, but when used they are like first-price sealed-bid auctions, in that both institutions provide the most complete information of the four mechanisms. Thus, in addition to the reserve price v_0, one typically observes all of the bids. Although one can measure the number of participants P, one cannot observe the number of potential bidders N. Thus, like English auctions, one must typically use other information to find the number of potential bidders N.

In order to carry out the analyses consider below, a researcher must know the number of potential bidders N, the reserve price v_0, and the winning bid w. As discussed above, obtaining the winning bid and the reserve price are typically straightforward. Thus, we shall often focus just upon the special case of the winning bid. In all of our work, we shall assume that a measure of the number of potential bidders N is available.[5] We shall also assume that the researcher has access to a sample of T independent auctions of a relatively homogeneous good or service as in Paarsch (1992) who considered low-price sealed-bid auctions of tree planting procurement auctions; Laffont, Ossard, and Vuong (1991) who considered the sale of eggplants using Dutch auctions; or Paarsch (1993) who considered the sale of timber at English auctions.

4. Data Generating Processes

One strategy for interpreting field data (see Paarsch 1989, 1991, 1992,1993) involves exploiting the fact that (2.1) and (2.5) are monotonic functions of V: players with higher valuations will bid

[5] Note that the number of participants at an auction P is not a useful proxy for the amount of competition N as P is endogenous. In the absence of a reserve price, researchers often argue that P is a good measure of N; see, for example, Paarsch (1992).

more. Because the bidding rules are functions of the random variable V, the bids are also random variables and their densities are related to $f(v)$.

4.1. English and Vickrey Auctions

At English and Vickrey auctions, finding the density of a bid B is straightforward because β is a trivial function of V; *viz.*, $B = \beta(V) = V$, if $V \geq v_0$. Thus by employing independence, one obtains the joint density of non-participation and bidding when $P > 1$ is similar to a Tobit model

$$\prod_{i=1}^{N} f(b_i)^{I_i} F(v_0)^{(1-I_i)}. \tag{4.1}$$

If exit from an English auction is imperfectly observed, then the joint density of non-participation and bidding for $P > 1$ is

$$\prod_{i=1}^{N} [1 - F(b_i)]^{I_i} F(v_0)^{(1-I_i)}. \tag{4.2}$$

When $P = 1$, the dominant bidding strategy at English auctions reveals no other information than $V_i > v_0$ for one potential bidder. The joint density in this case is then

$$\prod_{i=1}^{N} [1 - F(v_0)]^{I_i} F(v_0)^{(1-I_i)}.$$

The winning bid is also a function of the Vs. Thus, its density is related to $f(v)$. The density of the second-highest valuation for the object $Y = V_{(2:N)}$, when it exceeds v_0, is

$$NMF(y)^{M-1}[1 - F(y)]f(y),$$

so the density of the winning bid $W = \beta(Y) = Y$ when it exceeds v_0, denoted $h_B(w; M)$, is

$$h_B(w; M) = NMF(w)^{M-1}[1 - F(w)]f(w).$$

A single potential bidder participating at an English auctions bids the reserve price v_0, while at a Vickrey auction the solo bidder pays the reserve price v_0. Thus, there is a discrete mass point in W's distribution at v_0 having probability

$$h_B(v_0; M) = NF(v_0)^M[1 - F(v_0)].$$

When v_0 exceeds $\underline{v}$, there is also a chance of the object's going unsold, the "winning" bid being zero. This occurs when $P = 0$, the probability of which is

$$F(v_0)^N.$$

Introducing

$$D_0 = \begin{cases} 1 & \text{if } P = 0, \\ 0 & \text{otherwise;} \end{cases}$$

$$D_1 = \begin{cases} 1 & \text{if } P = 1, \\ 0 & \text{otherwise;} \end{cases}$$

and

$$D_{2+} = \begin{cases} 1 & \text{if } P \geq 2, \\ 0 & \text{otherwise;} \end{cases}$$

the density of W is

$$h_B(w; M) = \left[F(v_0)^N\right]^{D_0} \left[NF(v_0)^M[1 - F(v_0)]\right]^{D_1} \left[NMF(w)^{M-1}[1 - F(w)]f(w)\right]^{D_{2+}}. \tag{4.3}$$

4.2. Dutch and First-Price Sealed-Bid Auctions

The density of $S = \sigma(V)$ is more complicated to calculate because σ is a non-linear function of V. The joint density of non-participation and bidding is

$$\prod_{i=1}^{N} \left[\frac{f(\sigma^{-1}(s_i))}{\sigma'(\sigma^{-1}(s_i))}\right]^{I_i} F(v_0)^{(1-I_i)} = \prod_{i=1}^{N} \left[\frac{F(\sigma^{-1}(s_i))^{M\eta+1}}{M\eta \int_{v_0}^{\sigma^{-1}(s_i)} F(\xi)^{M\eta}\, d\xi}\right]^{I_i} F(v_0)^{(1-I_i)} \tag{4.4}$$

where $F(v_0)$ again denotes the probability of non-participation and where

$$\sigma'(v) = \frac{M\eta f(v) \int_{v_0}^{v} F(\xi)^{M\eta}\, d\xi}{F(v)^{M\eta+1}}$$

is the Jacobian for the transformation of v to $\sigma(v)$.

The winning bid is also a function of the Vs. Thus, its density is related to $f(v)$. The density of the highest valuation for the object $U = V_{(1:N)}$ is

$$NF(u)^M f(u),$$

so the density of the winning bid $W = \sigma(U)$, denoted $h_S(w; \eta, M)$, is

$$h_S(w; \eta, M) = \left[\frac{NF(\sigma^{-1}(w))^{M\eta+N}}{M\eta \int_{v_0}^{\sigma^{-1}(w)} F(\xi)^{M\eta}\, d\xi}\right]^{(1-D_0)} \left[F(v_0)^N\right]^{D_0}. \tag{4.5}$$

4.3. Revenue Equivalence Proposition

Assuming potential bidders are risk neutral with respect to winning the auction, the REP is a restriction upon the first moments of the winning bids at Dutch, English, first-price sealed-bid, and Vickrey auctions. In particular,

$$\int w h_B(w; M)\, dw = \int w h_S(w; 1, M)\, dw. \tag{4.6}$$

While many empirical applications of regression models based upon a moment condition like (4.6) are easy to implement, in the case of auction models (because of the discrete mass points in the density, *etc.*) it is often conceptually and analytically easier to work off the density of the individual bids or the density of just the winning bids.[6] In what follows, we shall adopt this approach.

5. Identification

In experimental work, researchers choose $F(v)$ and then examine whether observed behaviour is consistent with what (4.1), (4.2), (4.3), (4.4), and (4.5) would predict. When field data are used, however, the structure of $F(v)$ is typically unknown. Also, auctions can differ in ways that are observable to both potential bidders and researchers. In such cases, researchers often assume that the random variable V comes from some distribution that can be uniquely characterized by an $((r-1) \times 1)$ parameter vector θ as well as a vector of known, exogenous covariates denoted Z.[7] Thus,

$$F(v) = F(v; \theta, Z). \tag{5.1}$$

The parameter vector θ and the covariate vector Z will embed themselves in the densities of all of the bids as well as the winning bids for any of the auctions. Can one identify $F(v; \theta, Z)$ from observed field data?

5.1. English and Vickrey Auctions

At English and Vickrey auctions, an affirmative answer to this question is easy to see. For example, in the case of $P > 1$ when exit is observed perfectly, the joint density of non-participation and bidding is

$$\prod_{i=1}^{N} f(b_i; \theta, Z)^{I_i} F(v_0; \theta, Z)^{(1-I_i)},$$

[6] Computationally, the method of simulated non-linear least squares proposed by Laffont, Ossard, and Vuong (1991) may be preferred in the case of risk-neutral bidders at either Dutch or first-price sealed-bid auctions. No one, to our knowledge, has ever applied the generalized method of moments estimator proposed by Gallant and Tauchen (1992).

[7] Note that $\underline{v}$ and $\overline{v}$ may be known, or they may be contained in the unknown parameter vector θ.

while the density of just the winning bid is

$$h_B(w;\theta,M,Z) = \left[F(v_0;\theta,Z)^N\right]^{D_0}$$
$$\left[NF(v_0;\theta,Z)^M[1-F(v_0;\theta,Z)]\right]^{D_1}$$
$$\left[NMF(w;\theta,Z)^{M-1}[1-F(w;\theta,Z)]f(w;\theta,Z)\right]^{D_{2+}}.$$

For English and Vickrey auctions, the behavioural hypotheses of the model are that potential bidders bid independently and losers tell the truth. For example, at English auctions each potential bidder drops out of the auction when the price reaches the bidder's valuation, thereby revealing that valuation. The optimality of this strategy does not depend upon the strategies chosen by opponents. Because the equilibrium is a dominant strategy, there is an absence of strategic play, which makes it difficult to test the theory, assuming F is unknown. Thus, any assumption concerning F is an assumption concerning the distribution of bids (or the winning bid) since the bid function is essentially the identity function. Identification of θ can be assured by the choice of $F(v;\theta,Z)$.

5.2. Dutch and First-Price Sealed-Bid Auctions

In the case of Dutch and first-price sealed-bid auctions, the density of non-participation and bidding is

$$\prod_{i=1}^{N}\left[\frac{F(\sigma^{-1}(s_i;\theta,\eta,M,Z);\theta,Z)^{M\eta+1}}{M\eta\int_{v_0}^{\sigma^{-1}(s_i;\theta,\eta,M,Z)}F(\xi;\theta,Z)\,d\xi}\right]^{I_i} F(v_0;\theta,Z)^{(1-I_i)},$$

while the density of the winning bid is

$$h_S(w;\theta,\eta,M,Z) = \left[\frac{NF(\sigma^{-1}(w;\theta,\eta,M,Z),\theta,Z)^{M\eta+N}}{M\eta\int_{v_0}^{\sigma^{-1}(w;\theta,\eta,M,Z)}F(\xi;\theta,Z)^{M\eta}\,d\xi}\right]^{(1-D_0)}\left[F(v_0;\theta,Z)^N\right]^{D_0}.$$

In the following analysis of identification for Dutch and first-price sealed-bid auctions, we shall focus upon $h_S(w;\theta,\eta,M,Z)$ since at Dutch auctions only the winning bid is observed. Our results apply to the joint distribution of non-participation and bidding too.

The issue of identification for empirical specifications of Dutch and first-price sealed-bid auctions revolves around the following question: Is there a class of distributions $\mathcal{G}$ in which members G other than F solve

$$h_S(w;\theta,\eta,M,Z) = \frac{NG(S^{-1}(w))^{M\eta+N}}{M\eta\int_{v_0}^{S^{-1}(w)}G(\xi)^{M\eta}\,d\xi}$$

where

$$w = \mathcal{S}(u) = u - \frac{\int_{v_0}^{u} G(\xi)^{M\eta}\, d\xi}{G(u)^{M\eta}}?$$

To begin our analysis of identification, we assume that $F(v;\theta^0, Z)$ is the true distribution of valuations and that η^0 is the true value of the risk aversion parameter η. (For notational parsimony we shall often suppress the Z argument.) Note that the derived distribution of all the bids as well as the winning bid has support upon

$$\left[v_0, \int_{v_0}^{\overline{v}} F(\xi;\theta^0, Z)^{M\eta^0}\, d\xi\right] \equiv [v_0, \bar{s}(\theta^0, \eta^0, M, Z)].$$

In showing that the parameters vector (θ^0, η^0) is identifiable, we shall proceed in two steps. First, we shall show that (F, η^0) is the only distribution-risk aversion parameter pair that give rise to the true distribution for the winning bid. The identifiability of θ^0 will then follow from standard results concerning the uniqueness of the parameters defining F.

Being first concerned with the identifiability of the pair (F, η^0), we consider the question of whether (F, η^0) is the only pair that gives rise to the true distribution of the winning bid. Stated another way, is there another (G, η), such that either G is different from F (in a sense to be defined below), $\eta \neq \eta^0$, or both differ, that gives rise to the same probability law for the winning bid? The main regularity assumption on the class of distributions (of which F is a member) is contained in Assumption 1.

Assumption 1.

The class of distributions $\mathcal{G}$ contains distributions $F(v)$ defined upon $[\underline{v}, \overline{v}]$ that are monotonically increasing, continuously differentiable with continuous density $f(v)$ such that $f(v) > 0$ on $(\underline{v}, \overline{v})$.

Note that this assumption imposes no restrictions upon the value of the probability density function at its lower or upper bounds. A large number of families of probability density functions satisfies Assumption 1. Identification of F will be relative to the class $\mathcal{G}$, with F assumed to be a member of $\mathcal{G}$. We shall assume that η belongs to some set of real numbers $\mathcal{R} = [1, \Delta]$, for some large real number Δ. The sense in which we say that $G \in \mathcal{G}$ differs from F, written $G \neq F$, is given by the following definition:

Definition 1.

We say that $G \neq F$ if there exists an open interval $A \subset (\underline{v}, \overline{v})$ for which either $F(v) > G(v)$ or $G(v) > F(v)$ for all $v \in A$.

The sense in which (F, η^0) is identified is given in the final definition. Note that this definition requires that there exist some event (concerning the winning bid) for which the probability under (F, η^0) differs from the probability under (G, η) for any $(G, \eta) \neq (F, \eta^0)$. This is the standard notion of identification as in Wald (1949).

Definition 2.

(F, η^0) is identifiably unique relative to $\mathcal{G} \times \mathcal{R}$, if for any $G \in \mathcal{G}$ and $\eta \in \mathcal{R}$ such that $G \neq F$, $\eta \neq \eta^0$, or both,

$$\Pr[W \in A \mid F, \eta^0] \neq \Pr[w \in A \mid G, \eta]$$

for any Lebesgue measurable set A, where $\Pr[\cdot|F, \eta^0]$ denotes the probability calculated under (F, η^0).

Note that it will be sufficient to show that the density of W under (F, η^0) differs from that under (G, η) over some open set in $[v_0, \bar{s}]$.

Theorem 1.

Given Assumption 1, (F, η^0) is identifiably unique in $\mathcal{G} \times \mathcal{R}$.

The proof of this theorem, like the proofs of all of our results, is contained in an appendix to the paper.

A consequence of this result is that if we have proposed a family of distributions $F(v; \theta) = F_\theta$, then if the parameter θ^0 is identifiably unique, in the sense of generating a unique probability model for v, then it is also identifiably unique in the sense of generating a unique probability model for all of the bids or just the winning bid.

Corollary 1.

If for any $\theta \in \Theta$, such that $\theta \neq \theta^0$, we have $F_\theta \neq F$, then θ^0 is identifiably unique in the model for all of the bids or just the winning bid.

To obtain these results, we have not imposed any regularity conditions beyond those contained in Assumption 1. Additional assumptions concerning F may be required for an equilibrium to exist in the auction model, but these do not need to be imposed to identify the distribution of valuations given the distribution of all the bids or just the winning bids.

6. Estimation

Having established identification for each of the mechanisms in the case of field data and trivially for all four mechanisms in the case of experimental data, we shall now concentrate upon estimating the unknown parameters of these empirical specifications by the method of maximum likelihood. As in the case of identification, estimation is straightforward for data concerning English and Vickrey auctions, but not in the case for Dutch and first-price sealed-bid auctions.

6.1. English and Vickrey Auctions

Because the equilibrium bidding strategies at English and Vickrey auctions are trivial functions of the valuations for participants, deriving the likelihood function for a sample of data is straightforward. The econometrics of these auctions requires no more theoretical work than exists already in the literature. In addition, such empirical specifications can usually be estimated using existing software; *e.g.*, the maximum likelihood command available in TSP.

6.2. Dutch and First-Price Sealed-Bid Auctions

The straightforward econometric structure of English and Vickrey auctions does not carry over to Dutch and first-price sealed-bid auctions. The main technical problem is that the the support of both the individual bid and the winning bid distributions depends upon all of the parameters of interest. Thus, the standard distribution theory does not apply. Moreover, the standard way that consistency is demonstrated in econometrics is unsatisfactory for this purpose in this context. In what follows, we again focus upon the winning bids, but the methods we propose would apply were all of the bids used.

A simple example will illustrate the nonstandard nature of the problem a researcher faces when attempting to estimate structural econometric models of Dutch or first-price sealed-bid auctions. Consider a random sample of size T for a random variable W that is distributed uniformly upon the interval $[0, \alpha]$, where α is an unknown parameter which the investigator seeks to estimate. The probability density of function of W is

$$h(w; \alpha) = \alpha^{-1}\mathbb{1}[0 \leq w \leq \alpha]$$

where $1[\cdot]$ is the indicator function of the event argument. The standard approach to finding the maximum likelihood estimator of α would involve maximizing the following logarithm of the likelihood function with respect to α:

$$\frac{1}{T}\log L(\alpha; w_1, w_2, \ldots, w_T) = -\log\alpha + \frac{1}{T}\sum_{t=1}^{T}\log\left(1[0 \le w \le \alpha]\right).$$

The standard approach to demonstrating the parameter consistency of the maximum likelihood estimator (see, for example, White [1982] or Amemiya [1985]) would involve showing that this function converges uniformly over some parameter set to a function that is maximized at the true value.[8] The problem with this approach is that showing uniform convergence is difficult, unless the parameter set is restricted to be $[\alpha^0, \alpha^0 + \delta]$ for some value of δ which is greater than zero, where α^0 is the true value of α. This is because the usual dominance condition, such as Assumption A3 of White (1982), can only be satisfied on this set. This would then imply that the maximum is α^0, a somewhat unsatisfactory result that obviously depends upon knowledge of the true parameter value.

Our approach to demonstrating consistency avoids this difficulty. We note that an equivalent representation to maximizing the logarithm of the likelihood function is to solve the following constrained optimization problem:

$$\max_{<\alpha>} -T\log\alpha \quad \text{subject to} \quad \begin{cases} w_1 \le \alpha \\ w_2 \le \alpha \\ \vdots \\ w_T \le \alpha. \end{cases}$$

This approach is also more in line with the way the estimator will be calculated in practice. A further advantage to this approach is that treating the problem in this way will provide a link to the distribution theory. For, as we demonstrate below, the binding constraints are an important part of the solution. For example, in the uniform example considered above, it is easy to see that the solution for the maximum likelihood estimator involves a binding constraint. In fact, the maximum likelihood estimator is

$$\hat{\alpha} = \max[w_1, w_2, \ldots, w_T].$$

Abstracting from ties in the data, $T-1$ of the constraints do not bind. Also, the conventional methods used to determine the asymptotic distribution of $\hat{\alpha}$ do not apply. In particular, only the

[8] Note that some variant of Wald's (1949) proof of parameter consistency could potentially be used, although as Amemiya (1985, p. 118) and Dhrymes (1970, p. 121) have noted, some of Wald's conditions may be difficult to verify in practice.

largest w_t is important in determining the distribution of $\hat{\alpha}$. We shall discuss the problems which arise in performing the asymptotic analysis in detail below. Suffice to say here that we shall define the maximum likelihood estimator of the unknown parameter vector $\alpha = (\theta, \eta)$ in a similar fashion.

Optimization Problem and Parameter Consistency

To define the maximum likelihood estimator, we first denote the density of the t^{th} observed winning bid w_t conditional upon $x_t = (m_t, Z_t)$ by

$$h_S(w_t; \alpha, x_t)\ 1[v_0 \leq w_t \leq \bar{s}(\alpha, x_t)]$$

where α is an r dimensional parameter vector of interest. Letting

$$L_T(\alpha) = \frac{1}{T}\sum_{t=1}^{T} \log h_S(w_t; \alpha, x_t),$$

we denote the set of feasible values of α (that are consistent with the data in the sense described above) by

$$A_T^* = \{\alpha \in \bar{A} \mid v_0 \leq w_t \leq \bar{s}(\alpha, x_t)\ \forall\ t = 1, \ldots, T\}$$

where $\bar{A}$ is some compact set which contains the true value of the parameter α^0. Note that by definition $\alpha_0 \in A_T^*$ for all T. The maximum likelihood estimator of α, $\hat{\alpha}$, is defined as the solution to

$$\max_{<\alpha>} L_T(\alpha) \quad \text{subject to} \quad \alpha \in A_T^*.$$

In proving the parameter consistency of the maximum likelihood estimator, the main complication that arises is that the set A_T^* shrinks as the sample size increases. To prove consistency in this case, we shall show how A_T^* behaves as T grows. To begin, we make some assumptions regarding the distribution of the x_t.

Assumption 2.

The x_ts are independently and identically distributed and contain discrete variables x_{1t} with finite support X_1, and continuously distributed variables x_{2t} with compact support X_2. Moreover, for any non-empty subset A_1 of X_1 and any non-empty open subset A_2 of X_2, $\Pr[A_1 \times A_2] > 0$.

We shall let $X = X_1 \times X_2$ denote the complete set of x variables. Next, we shall make an assumption that restricts the behaviour of the true density function near the upper bound of the support. This will be useful in determining the limiting behaviour of the set A_T^*.

Assumption 3.

For any $\epsilon > 0$,

$$\inf_{x \in X} \Pr[W > \bar{s}(\alpha^0, x) - \epsilon] = \delta(\epsilon) > 0.$$

A final assumption that is used to analyze the behaviour of A_T^* regards the nature of the $\bar{s}(\alpha, x)$ function.

Assumption 4.

For any $\alpha \in \bar{A}$, $\bar{s}(\alpha, x)$ is continuous in x on X. Moreover,

$$v_0 < \inf_{x \in X} \bar{s}(\alpha^0, x) < \sup_{x \in X} \bar{s}(\alpha^0, x) < \infty.$$

In providing conditions for consistency, we shall show first that the set A_T^* converges to the set

$$A^* = \{\alpha \in \bar{A} \mid \bar{s}(\alpha^0, x) \leq \bar{s}(\alpha, x) \; \forall \; x \in X\}.$$

This is the set of α that obeys the constraints for all possible values of x and w. Notice that by construction $A^* \subset A_T^*$. After showing that A_T^* converges to A^*, we shall give a set of conditions in terms of A^* that guarantee consistency. The sense in which A_T^* converges to A^* is given in the following definition.

Definition 3.

$A_T^* \xrightarrow{\text{a.s.}} A^*$ if $A^* \subset A_T^*$ and if for any $\alpha \notin A^*$ there is a finite value $\bar{T}$ such that $\alpha \notin A_T^*$ for all $T \geq \bar{T}$ with probability 1.

Notice that in this definition $\bar{T}$ may depend upon the particular α chosen. Also note that the probability measure with respect to which this condition relates is the joint probability measure of the pair (w, x). Theorem 2 contains the consistency result.

Theorem 2.

Under Assumptions 2 to 4, $A_T^* \xrightarrow{\text{a.s.}} A^*$.

The following result contains the general consistency result for the maximum likelihood estimator. Notice that the result proved in Theorem 2 is one of the assumptions used to prove consistency.

Theorem 3.

Given the following:

a) $\bar{A}$ is compact;

b) $A_T^* \xrightarrow{\text{a.s.}} A^*$;

c) $L_T(\alpha) \xrightarrow{\text{a.s.}} L(\alpha)$ uniformly over $\bar{A}$;

d) if $L(\alpha) \geq L(\alpha^0)$ for any $\alpha \in A^*$, then $\alpha = \alpha^0$;

$\hat{\alpha} \xrightarrow{\text{a.s.}} \alpha^0$.

The conditions have been stated in such a way so as not to impose a strict set of primitive conditions on the underlying density function. This seems most desirable since there are now a large number of ways that one can verify the uniform convergence condition in c); see, for example, Newey (1991) and Andrews (1992). Conditions c) and d) have been stated in a way that requires verification with respect to the fixed sets A^* and $\bar{A}$ rather than with respect to the changing random set A_T^*. This should make it easier to verify them in practice.

In order to calculate the maximum likelihood estimator on a computer using non-linear programming techniques as well as to determine its asymptotic distribution, we need the following two assumptions:

Assumption 5.

The $\log h_S(\alpha; w, x)$ function is twice continuously differentiable in α.

Assumption 6.

The $\bar{s}(\alpha, x_t)$ functions are quasi-convex and twice continuously differentiable in α.

The maximum likelihood estimator $\hat{\alpha}$ can be computed by solving the following optimization problem:

$$\max_{<\alpha>} \sum_{t=1}^{T} \log h_S(\alpha; w_t, x_t) \quad \text{subject to} \quad \begin{cases} w_1 \leq \bar{s}(\alpha, x_1) \\ w_2 \leq \bar{s}(\alpha, x_2) \\ \quad\vdots \\ w_T \leq \bar{s}(\alpha, x_T). \end{cases}$$

In practice, solving for the maximum likelihood estimator will involve maximizing the following Lagrangean:

$$\mathcal{L}(\alpha, \lambda) = \sum_{t=1}^{T} \Big(\log h_S(\alpha; w_t, x_t) + \lambda_t \big(\bar{s}(\alpha, x_t) - w_t \big) \Big)$$

with respect to the vector α, where $\lambda = (\lambda_1, \ldots, \lambda_T)$ is the vector of T Lagrange multipliers. The maximum likelihood estimator $\hat{\alpha}$ satisfies the following conditions:[9]

$$\sum_{t=1}^{T} \left(\nabla_\alpha \log h_S(\hat{\alpha}; w_t, x_t) + \lambda_t \nabla_\alpha \bar{s}(\hat{\alpha}, x_t)\right) = \mathbf{0}$$

$$\lambda_1 \left(\bar{s}(\hat{\alpha}, x_1) - w_1\right) = 0$$

$$\lambda_2 \left(\bar{s}(\hat{\alpha}, x_2) - w_2\right) = 0$$

$$\vdots$$

$$\lambda_T \left(\bar{s}(\hat{\alpha}, x_T) - w_T\right) = 0,$$

where ∇_α denotes the gradient vector of the function to follow with respect to the vector α. At most, r of the T constraints will ever bind at one time; *i.e.*, $T - r$ of the Lagrange multipliers will be zero at the optimum. For the binding constraints, the Lagrange multipliers will be non-negative.

Asymptotic Distribution of the Estimator

A natural way to calculate the variance-covariance matrix of $\hat{\alpha}$ would be to consider the behaviour of the Hessian matrix of the Lagrangean

$$\nabla_{\alpha\alpha} \mathcal{L}(\hat{\alpha}) = \sum_{t=1}^{T} \left(\nabla_{\alpha\alpha} \log h_S(\hat{\alpha}; w_t, x_t) + \hat{\lambda}_t \nabla_{\alpha\alpha} \bar{s}(\hat{\alpha}, x_t)\right).$$

This is useful when the solution to the optimization problem occurs along a smooth and differentiable part of the constraint set, but typically the solution obtains at the intersection of the constraints. In this case, the Hessian is ill-defined. Moreover, the properties of the perturbed optimum are determined solely by the constraints. To see this, consider the simple problem introduced in the first part of this section. There, the properties of the maximum likelihood estimator $\hat{\alpha}$ were solely determined by the behaviour of the largest w_t in a sample of size T. In this case, the properties will often be determined by the solution to some set of the largest r order statistics of w_t given x_t.

As may be expected from the previous discussion, the distribution theory for the estimator can be quite complicated. Because of technical difficulties that arise with T constraints when T is

[9] Under the stated conditions, these Kuhn-Tucker conditions are necessary, but not sufficient for a global maximum. If the logarithm of the likelihood function is pseudo-concave, then it is well-known (see Mangasarian [1969]) that the Kuhn-Tucker conditions are both necessary and sufficient.

going to infinity, we have only analyzed the case of discrete covariates in which case a finite number of constraints exist asymptotically. Hence,

Assumption 7.

x is a discrete random vector with probability mass function $\pi(x)$, with k being the number of points that have $\pi(x) > 0$.

Denote each possible point in the set by $x(i)$, and let $\pi_i = \pi(x(i))$ for $i = 1, ..., k$.

Despite the assumption of discrete covariates, the results are of considerable interest. Indeed, as the following discussion will show, the limiting distributions of the estimators will depend upon the relationship between k and r, and will only fall into the usual normal limiting family in a special case.

The advantage of having discrete covariates is that the sample optimization problem, which we shall call (L), may be written as

$$\max_{<\alpha>} \sum_{i=1}^{k} \hat{\pi}_i L_T(\alpha, x(i)) \quad \text{subject to} \quad \underline{\hat{w}}(x(i)) \leq \bar{s}(\alpha, x(i)) \quad i = 1, \ldots, k$$

where $\underline{\hat{w}}(x(i)) = \max\{w_t : x_t = x(i)\}$ is the largest order statistic of w_t over all observations that have $x_t = x(i)$, $\hat{\pi}_i = T_i/T$ is the proportion of the sample with $x_t = x(i)$, and

$$L_T(\alpha, x(i)) = \frac{1}{T_i} \sum_{t=1}^{T} \log h_S(\alpha; w_t, x(i)) I[x_t = x(i)]$$

is the average of the contributions to the logarithm of the likelihood function contribution of observations with $x_t = x(i)$. The fact that the constraints involves order statistics and that some of the constraints bind at the maximum likelihood estimate will lead to the unusual limiting distributions that appear below.

Before proceeding, we first present a general result, contained in Galambos (1978) and discussed in Reiss (1989), which gives the limiting distribution of order statistics and their relationship to the Weibull distribution.

Lemma 1.

Suppose that the $\{w_t\}_{t=1}^{T}$ are drawn randomly from a population with probability density function f and cumulative distribution function F, on $[0, \tau]$ such that for all $z > 0$,

$$\lim_{t \to \infty} \frac{1 - F(\tau - \frac{1}{tz})}{1 - F(\tau - \frac{1}{t})} = \frac{1}{z^\gamma}$$

and $\tau < \infty$, then

$$\frac{1}{d_T^*}(\tau - \max_t\{w_t\}) \rightarrow \mathcal{W}(1,\gamma)$$

where $\mathcal{W}(1,\gamma)$ denotes a random variable that is distributed Weibull with parameters 1 and γ, and $d_T^* = \tau - F^{-1}(1 - \frac{1}{T})$.

This Lemma provides conditions under which the limiting distribution of the largest order statistic is

$$[1 - \exp(-z^\gamma)].$$

Note that it is only defined for positive values of z. This gives the well-known fact that extreme order statistics are biased estimators of the upper bound of the distribution, although they generally converge very quickly as shown in the previous result. Note also that if we can find alternative constants d_T such that $d_T^*/d_T \rightarrow 1$, then the result will still hold. The γ parameter will depend upon the behaviour of the density function near the upper bound of the support. There may be other types of limiting distributions of largest order statistics (depending upon the nature of the parent population), but this result is sufficient to characterize the limiting distributions of order statistics in the auction case. As the following result shows, the densities in the problem we consider will satisfy the condition in Lemma 2 with $\gamma = 1$ because the density is strictly positive and finite at the upper bound of the support.

Lemma 2.

Given Assumption 4,

$$\lim_{w \rightarrow \bar{s}(\alpha^0,x)} h_S(w;\alpha^0,x) = \frac{N}{M\eta^0 \int_{v_0}^{\bar{v}} F(\xi;\theta^0,Z)^{M\eta^0}\, d\xi} > 0.$$

This result shows that the distribution of the winning bid has a strictly positive density at its upper bound. This fact will be important to the proof of Corollary 2 below. There we show that the largest order statistic for each possible value of x consistently estimates the upper bound for each possible value of x. Moreover, these order statistics are consistent at rate T. This result is useful since later we show that the maximum likelihood estimator in these models depends upon order statistics. In some cases, the maximum likelihood estimator is obtained by solving for the parameters purely as functions of the order statistics. In such cases, the maximum likelihood estimator will also be consistent a rate T. The particular limiting distribution that results will, however, depend upon the number of possible values of x as well as the number of parameters.

Combining Lemma 1 and Lemma 2, we can show that the limiting distribution of the largest winning bid for each possible covariate value will be exponential with intensity parameter equal to one. In addition, a convenient form for the normalizing constant can always be found.

Corollary 2.

Under Assumptions 1 to 7,

$$\frac{1}{d_T}\left(\bar{s}(\alpha^0, x(i)) - \underline{\hat{w}}(x(i))\right) \xrightarrow{\text{d}} \mathcal{W}(1,1)$$

where $\mathcal{W}(1,1)$ is an exponential random variable with parameter 1, denoted $\mathcal{E}(1)$, and

$$d_T = \frac{M\eta^0(\overline{v} - \bar{s}(\alpha^0, x(i)))}{NT_i} = O_p(T^{-1}).$$

To make these notions concrete, consider the following example of a procurement auction that has no covariates and where the lowest bidder wins the right to perform the task and is paid the amount bid. If costs are distributed Pareto, then so too is the winning bid. Suppose that the parameters are α_1 and α_2, and

$$h_S(w;\alpha_1,\alpha_2) = \frac{\alpha_2\alpha_1^{\alpha_2}}{w^{\alpha_2+1}} \quad 0 < \alpha_1 < w \text{ and } 0 < \alpha_2,$$

then for a sample of size T

$$\frac{\alpha_2 T}{\alpha_1}(\min w_t - \alpha_1) \xrightarrow{\text{d}} \mathcal{E}(1)$$

is distributed exponentially in the limit with parameter 1 because the Pareto density function is strictly positive at the upper bound of the distribution.[10]

Letting the subscript 0 on the function denote population values, introduce the following notation

$$L_0(\alpha, x(i)) = \mathrm{E}_0[\log h_S(w;\alpha^0, x(i))]$$

where E_0 denotes that the expectation is taken at the true parameter values α^0. Also, define the following population optimization problem (P_i):

$$\max_{<\alpha>} \; L_0(\alpha, x(i)) \quad \text{subject to} \quad \bar{s}(\alpha, x(i)) \le \bar{s}(\alpha^0, x(i))$$

[10] The reader will remark that the normalization employed in this example is not the one implied by Lemma 1. We use it because it is equivalent in the limit and has a convenient form, as do the constants d_T in Corollary 2 for the auction case.

as well as the aggregate problem (P)

$$\max_{<\alpha>} \sum_{i=1}^{k} \pi_i L_0(\alpha, x(i)) \quad \text{subject to} \quad \bar{s}(\alpha, x(i)) \leq \bar{s}(\alpha^0, x(i)) \quad i = 1, \ldots, k.$$

We shall assume throughout that we can interchange integration and differentiation. We also introduce the following assumption regarding the problem P_i.

Assumption 8.

The solution to P_i is (α^0, λ_i^0) with $\lambda_i^0 > 0$, so that the constraint binds.

Typically, we find that the expectation of the gradient vector $\nabla_\alpha L_0$ has at least one strictly positive element, with remainder being 0 when evaluated at the true parameters α^0. Continuing with the Pareto example considered above, one can demonstrate easily that Assumption 8 is satisfied with $\lambda^0 = \alpha_2^0/\alpha_1^0 > 0$.

A condition that simplifies the calculation of the limiting distribution is contained in Assumption 8. It essentially permits use of the implicit function theorem to concentrate the likelihood function using the binding constraints, so that standard expansions can be used to obtain the asymptotic distribution. Consequently, when $k \geq r$, the maximum likelihood estimator is determined solely by the constraints, so its distribution will depend only upon the distribution of the order statistics.

Assumption 9.

The matrix whose columns contain the k-vectors

$$\nabla_\alpha \bar{s}(\alpha, x(i))$$

is of full rank $\min\{k, r\}$ uniformly in a neighbourhood of α^0.

In the remainder of this section we analyze two different cases.

Case 1: $\mathbf{k \leq r}$

Partition the vector α into (α_1, α_2) of dimensions k and $(r-k)$ respectively, in such a way that the $k \times k$ matrix whose i^{th} column is

$$\nabla_{\alpha_1} \bar{s}(\alpha, x(i))$$

is non-singular over a neighbourhood of α^0. Of course, when $k = r$ the α_2 component is non-existent. The next result shows that $\hat{\alpha}$ the solution to the sample maximization problem defined above, with probability one, will satisfy the Kuhn-Tucker conditions, with all k constraints binding as T tends to infinity.

Theorem 4.

Given Assumptions 1 to 9, and assuming that

$$\nabla_\alpha L_T(\alpha, x(i)) \xrightarrow{\text{a.s.}} \nabla_\alpha L_0(\alpha, x(i))$$

uniformly over a neighbourhood of α^0 for each i, then for large enough T all k constraints bind at the solution $\hat{\alpha}$ with probability 1.

The fact that all k constraints bind makes it possible to invert out a subset of parameters α_1 as a function of the remaining parameters by Assumption 9 and the implicit function theorem. The resulting solution will be twice continuously differentiable in both α_2 and the remaining arguments in a neighbourhood of α_2^0. Here, the implicit function will be denoted as

$$\alpha_1 = \psi(\alpha_2, \underline{\hat{w}}, x),$$

the solution to the set of equations

$$\underline{\hat{w}}(x(i)) = \bar{s}(\alpha, x(i)) \quad i = 1, \ldots, k$$

where we introduce the shorthand $\underline{\hat{w}}$ to denote the k vector of $\underline{\hat{w}}(x(i))$s and x to denote the vector of $x(i)$. Also, note that

$$\alpha_1^0 = \psi(\alpha_2^0, \bar{s}^0, x)$$

where $\bar{s}^0$ denotes the k vector of values of the upper bounds. When $k = r$, one can solve for α_1 just using the constraints, so that it can be written as a function of only the $\underline{\hat{w}}(x(i))$s, and its distribution will depend on the distributions of the $\underline{\hat{w}}(x(i))$s. In this case, this condition will give rise to limiting distributions related to those in Corollary 2.

We next introduce the following notation, which will be useful in characterizing the results when $k < r$. Define

$$\mathrm{V}[\nabla_\alpha \log h_S(\alpha^0; w, x(i)] = \begin{pmatrix} \Omega_1^i & \Omega_{12}^i \\ \Omega_{21}^i & \Omega_2^i \end{pmatrix}$$

for each i where the partition is conformable with that of α. In the case where $k < r$, standard mean value expansions will be used to find the limiting distribution. The terms involved will be of the form

$$d_t = \nabla'_{\alpha_2} \psi(\alpha_2) \nabla_{\alpha_1} \log h_S(\alpha; w_t, x_t) + \nabla_{\alpha_2} \log h_S(\alpha; w_t, x_t).$$

Note that Assumption 8 implies that $\mathrm{E}[d_t] = 0$ for each t. Define for $x_t = x(i)$,

$$\mathrm{V}[d_i] = \nabla'_{\alpha_2}\psi(\alpha_2)\Omega_1^i\nabla_{\alpha_2}\psi(\alpha_2) + \Omega_2^i + \nabla'_{\alpha_2}\psi(\alpha_2)\Omega_{12}^i + \Omega_{21}^i\nabla_{\alpha_2}\psi(\alpha_2)$$

Theorem 5.

Under Assumptions 1 to 9, and assuming that $k < r$

$$\sqrt{T}(\hat{\alpha}_1 - \alpha_1^0) \xrightarrow{\mathrm{d}} \mathcal{N}(0, V_1)$$

and

$$\sqrt{T}(\hat{\alpha}_2 - \alpha_2^0) \xrightarrow{\mathrm{d}} \mathcal{N}(0, V_2)$$

where

$$V_1 = \left[\sum_{i=1}^{k} \pi_i \mathrm{V}[d_i]\right]^{-1}$$

and

$$V_2 = \nabla'_{\alpha_2}\psi(\alpha_2)V_1\nabla_{\alpha_2}\psi(\alpha_2).$$

When $k = r$, things are very different. Since, by Theorem 4, the parameters are determined by the constraints, no averages are involved and the distribution is related to that of $\underline{\hat{w}}$, which are extreme order statistics. The limiting distributions in this case are related to the $\mathcal{E}(1)$ family, and the estimators will converge at rate T. This unusual result is contained in Theorem 6. To state this theorem succinctly, we first develop some notation. Note that in this case for large enough T, $\hat{\alpha}$ is the solution to

$$\underline{\hat{w}}(x(i)) = \bar{s}(\hat{\alpha}, x(i)),$$

so that, as noted above, we can write

$$\hat{\alpha} = \psi(\underline{\hat{w}}, x)$$

where $\psi(\cdot)$ is a smooth function of $\underline{\hat{w}}$ near the limiting values $\bar{s}^0$. To characterize the limiting distribution we expand the function about $\bar{s}^0$.

Theorem 6.

Under Assumptions 1 to 9 and assuming that $k = r$,

$$-\hat{D}_T^{-1}\hat{J}_T(\hat{\alpha} - \alpha^0) \xrightarrow{\mathrm{d}} (\mathcal{E}_1(1), \ldots, \mathcal{E}_k(1))$$

a vector of independent $\mathcal{E}(1)$ random variables where

$$\hat{D}_T = \text{diag}\{d_T(i)\}$$

of dimension k, where $d_T(i)$ is given in Corollary 2, and

$$\hat{J}_T = \nabla_\alpha \bar{s}(\hat{\alpha})$$

with $\nabla_\alpha \bar{s}(\alpha)$ being the matrix formed by the vectors $\nabla_\alpha \bar{s}(\alpha, x(i))$ for $i = 1, \ldots, k$.

Note that the standardization in Theorem 6 will be proportional to T, so that the estimators converge at the rate T, and the limiting distribution is that of a vector of independent exponential $\mathcal{E}(1)$ random variables. This result does not imply that the estimators themselves have one-sided distributions, only that there is a linear transformation of the estimators that has a one-sided distribution. One may be concerned about the fact that the limiting distribution is not normal, since this will make inference difficult. There is no need for this concern since the exponential distribution has a particularly convenient closed-form cumulative distribution function, which will make it even easier to form confidence intervals than would be the case with a normal limiting distribution.

These results can easily be adjusted to the case where only a subset of the parameters influence the upper bound of the distribution. Another case that can easily be examined is where $k < r$ and one can solve for a subset of α_1 as functions of only $\hat{w}$ and x. The result of Theorem 6 would imply that these parameter estimators have $\mathcal{E}(1)$ limiting distributions whereas the remaining parameter estimators will have distributions that fall in the normal limiting family of distributions.

Corollary 3.

Suppose the conditions of Theorem 5 hold, and that a subset of $\hat{\alpha}_1$ can be written as functions of only $\underline{\hat{w}}$ and x, then this subset will have limiting $\mathcal{E}(1)$ distributions, and the remaining parameters will have limiting normal distributions. Moreover, the first subset will converge at rate T and the remainder converge at $\sqrt{T}$.

Note that this corollary can be used to show that the simple Pareto example considered previously in the examples behaves like this.

Case 2: k > r

When $k > r$, with finite k, there will generally be more than one way of determining the parameters from the constraints. Moreover, in the population problem (P), the objective function may be tangent to one of the constraints. These facts make it possible for the solution to the sample problem to be such that r constraints bind, or less than r constraints bind, and this will be random from sample to sample. This introduces potential difficulties in the asymptotic analysis. To proceed, we shall make the following assumption, which will guarantee that for large T the solution to (P) has at least r constraints binding.

Assumption 10.

In problem (P), the matrix

$$(\nabla_\alpha L_0(\alpha), \{\nabla_\alpha \bar{s}(\alpha)\}_{r-1})$$

has full rank over a neighbourhood of α^0 where

$$\{\nabla_\alpha \bar{s}(\alpha)\}_{r-1}$$

is a collection of derivatives of any $r-1$ distinct upper bounds.

The type of situation that this assumption rules out is illustrated in Figure 1, which applies to the Pareto auction example when $k = 3$ and $r = 2$. The following Lemma then shows that given this assumption the solution to the sample problem for large T will occur where r constraints are binding. The advantage of this is that the asymptotics for the case with r binding constraints and the case with less than r binding constraints are quite different; with the first convergence is at rate T and with the latter convergence is at rate $\sqrt{T}$.

Lemma 4.

In the sample problem (L), assuming that

$$\nabla_\alpha L_T(\alpha, x(i)) \xrightarrow{\text{a.s.}} \nabla_\alpha L_0(\alpha, x(i))$$

uniformly over a neighbourhood of α^0, then the optimal solution occurs at a point such that for large T at least r constraints bind with probability 1.

Suppose there are k constraints and r parameters, then there will be $\mathcal{L} = \binom{k}{r}$ possible combinations of constraints at which the solution may occur. Denote the set of possibilities by

Ξ. (Note that Assumption 9 guarantees that each possible combination will possess a solution.) We shall let ℓ index each solution and let $\hat{\alpha}(\ell)$ be the solution to the ℓ^{th} set of constraints. Also, let $E(\ell)$ be the event that the solution to (L) is at $\hat{\alpha}(\ell)$. In this Lemma, we show that for large enough T the solution to problem (L), denoted $\hat{\alpha}$, is such that

$$\hat{\alpha} = \sum_{\ell \in \Xi} \hat{\alpha}(\ell) I[E(\ell)]$$

almost surely, where $I[E]$ denotes the indicator function for the event E. Note that with probability 1 only one of the $I[E(\ell)]$ will be 1.

Using arguments similar to those used previously, one can rule out some of the combinations as being likely to occur with probability 0. For example, consider Figure 2 representing the population problem with $k = 3$ and $r = 2$, and satisfying Assumption 10. In this case, we can rule out an optimum at the solution of C_2 and C_3 as being likely. The reason is that if we maximized the objective function subject to these two constraints then the optimum is actually at B rather than at A. Alternatively, at A it is impossible to find a positive Lagrange multiplier λ_2 to satisfy the first Kuhn-Tucker condition using these two constraints alone. Also note that at the actual optimum B, C_1 is not satisfied. Noting this, we are able to narrow down the set of possible solutions. In the example in Figure 2, there will in fact be two possible solutions. Maximizing the objective subject to either (C_1, C_2) or (C_1, C_3) we are able to satisfy the Kuhn-Tucker conditions; *i.e.*, find positive λ_is and satisfy all remaining constraints. The following Lemma makes this more precise in general. First, define the solution to the linear equation,

$$\nabla_\alpha L_0(\alpha) - \{\nabla_\alpha \bar{s}(\alpha)\}_{r\ell} \lambda = 0$$

to be λ_0^ℓ, where the notation $\{\cdot\}_{r\ell}$ is the ℓ^{th} possible combination of r elements of the argument.

Lemma 5.

In the sample problem (L), assuming that

$$\nabla_\alpha L_T(\alpha, x(i)) \xrightarrow{\text{a.s.}} \nabla_\alpha L_0(\alpha, x(i))$$

uniformly over a neighbourhood of α^0, then for large enough T $\Pr[E(\ell)] = 0$ if any element of λ_0^ℓ is negative.

Note that Assumption 10 rules out the possibility of any of the λ_0^ℓ being 0. Also, note that in the example in Figure 2, the combination that has (C_2, C_3) will have one of the Lagrange multipliers being negative. Denote the remaining set of possible combinations, not ruled out by Lemma 5, by Ξ_R. Also, define Ξ_ℓ, the ℓ^{th} element of Ξ_R, to be the collection of the indices of constraints used to obtain this solution. For example, in the case considered in Figure 2, there are two possible solutions, so Ξ_R has two elements, and we could define $\Xi_1 = \{1,2\}$ and $\Xi_2 = \{1,3\}$.

One potential problem that is raised by this representation is that the events $E(\ell)$ are random. Although Theorem 6 gives a nice characterization of the distribution for any *given* solution (unconditionally), the normalizing matrix $\hat{D}_T^{-1}\hat{J}_T$ will, in general, be different for each possible solution. This problem can be avoided, however, by using the following normalizing random variable,

$$\hat{B}_T = -\sum_{\ell \in \Xi_R} I[E(\ell)]\hat{D}_{T\ell}^{-1}\hat{J}_{T\ell}$$

where $\hat{D}_{T\ell}^{-1}\hat{J}_{T\ell}$ is the required normalization for the ℓ^{th} possible solution as in Theorem 6. Since almost surely only one of the $I[E(\ell)]$ will be 1, and $I[E(i)]I[E(j)] = 0$ for $i \neq j$, then we have that

$$\hat{B}_T(\hat{\alpha} - \alpha^0) = -\sum_{\ell \in \Xi_R} I[E(\ell)]\hat{D}_{T\ell}^{-1}\hat{J}_{T\ell}(\hat{\alpha}(\ell) - \alpha^0).$$

In order to characterize the limiting distribution of this quantity, we must determine for each value of ℓ, the limiting distribution of

$$\hat{D}_{T\ell}^{-1}\hat{J}_{T\ell}(\hat{\alpha}(\ell) - \alpha^0) = \hat{D}_{T\ell}^{-1}(\underline{\hat{w}}(x(\ell)) - \bar{s}_\ell)$$

conditional on the event $E(\ell)$. This requires more precise information on what actually determines $E(\ell)$ and its relationship to the above random variables. The following Lemma shows precisely how this is done.

Lemma 6.

In the sample problem (L), assuming that

$$\nabla_\alpha L_T(\alpha, x(i)) \xrightarrow{\text{a.s.}} \nabla_\alpha L_0(\alpha, x(i))$$

uniformly over a neighbourhood of α^0, then for large enough T $I[E(\ell)] = 1$, and hence $\hat{\alpha} = \hat{\alpha}(\ell)$, if and only if, for every constraint i

$$\underline{\hat{w}}(x(i)) \leq \bar{s}(\hat{\alpha}(\ell), x(i)).$$

In other words, after restricting attention to solutions that are not ruled out by Lemma 5, a particular solution will be the optimum when all of the constraints are satisfied at that particular solution. In the case where the logarithm of the likelihood function is pseudo-concave over a neighbourhood of α^0, this result would be obvious due to the fact noted in footnote 9. The constraints used to determine the ℓ^{th} solution will all be satisfied, so it remains to check any constraint not included in the ℓ^{th} solution. This fact makes characterizing the conditional distribution possible. This is because the joint distribution of all k^{th} order statistics is simply the product of each marginal distribution due to the independence assumption, and the conditional distribution is just the conditional distribution of r components of this conditional upon the fact that certain linear combinations of these r components exceed each of the remaining $(k - r)$ components. This is proved in Theorem 7, which contains the limiting distribution result for the quantity $\hat{B}_T(\hat{\alpha} - \alpha^0)$.

Theorem 7.

Under the assumptions made above

$$\hat{B}_T(\hat{\alpha} - \alpha^0) \xrightarrow{\text{d}} z = (z_1, \ldots, z_r)'$$

which has a joint density function given by

$$\prod_{i=1}^{r} e(z_i) \sum_{\ell \in \Xi_R} \prod_{i \notin \Xi_\ell} [1 - \Pr(k'_{\ell i} z)]$$

where the constants $k_{\ell i}$ are given by

$$k'_{\ell i} = \lim_{T \to \infty} d_T^{-1}(i) \frac{\partial \bar{s}_i}{\partial \alpha'} \hat{J}_{T\ell}^{-1} \hat{D}_{T\ell}.$$

Here, $e(\cdot)$ is the probability density function of an $\mathcal{E}(1)$ random variable, while

$$\Pr[k'_{\ell i} z] = I[k'_{\ell i} z > 0] E(k'_{\ell i} z)$$

and $E(\cdot)$ is the cumulative distribution function for an $\mathcal{E}(1)$ random variable.

The limiting distribution in Theorem 7 is non-standard, and to our knowledge has not appeared in any other problems. As it stands, the distribution depends upon various unknowns, but these unknowns are all estimable. One may determine consistently the set Ξ_R by using the result of Lemma 6 to see if Lagrange multipliers at the solutions in the sample problem are all positive. Given that this is possible, one may then determine the constants $k_{\ell i}$ using sample estimates.

A simple corollary that follows from this is that the estimator is consistent at the rate T rather than the usual $\sqrt{T}$.

Corollary 4.

Under the conditions of Theorem 7, for any $\delta > 0$

$$T^{1-\delta}(\hat{\alpha} - \alpha^0) = o_p(1)$$

7. Testing

Having specified, identified, and estimated empirical models of auctions, the natural next step is to test whether these models are consistent with theoretical predictions. In experimental work, this might involve deciding whether the estimated parameters are consistent with those set by the researcher. Alternatively, when field data are used one might be interested in deciding whether an existing implementation of an auction is close to the optimal selling mechanism. As in the previous two sections, testing using data from English and Vickrey auctions is straightforward, but when data from either Dutch or first-price sealed-bid auctions are used the analysis is more complicated.

7.1. English and Vickrey Auctions

Hypothesis testing using the likelihood ratio, the Lagrange multiplier, or the Wald tests is well understood in the case of English and Vickrey auctions.[11] For example, with experimental data, one could examine the behaviour of the score function at the truth via a Lagrange multiplier test. With field data, the Wald test could be used to decide whether an observed implementation of an auction is close to the optimal auction selling mechanism.

7.2. Dutch and First-Price Sealed-Bid Auctions

Except in special cases, the non-standard distribution of the maximum likelihood estimator in the case of Dutch and first-price sealed-bid auctions prevents one from using any of the conventional hypothesis testing procedures. How can the asymptotic results presented above be used to make inferences about parameters? In answering this question, we consider three cases: a) $k < r$; b) $k = r$; and c) $k > r$. Because Case a) fits into the standard framework with asymptotic normality,

[11] Engle (1984) provides a good summary of these procedures.

inference may proceed in the usual way. Cases b) and c), on the other hand, require additional analysis.

Case b): k=r

In this case, the maximum likelihood estimator is asymptotically an r vector of independently and identically distributed exponential random variables having intensity parameter 1, hereafter denoted $\mathcal{E}_r(1)$. Note that nuisance parameters are not an issue here. For individual elements, however, the limiting distribution does depend upon nuisance parameters. In particular, letting A_i denote the i^{th} row of A, the limit of $\hat{A}/T$, where $\hat{A}^{-1} = -\hat{D}_T^{-1}\hat{J}_T$, then by Theorem 6

$$T(\hat{\alpha}_i - \alpha_i^0) \xrightarrow{\text{d}} A_i\mathcal{E}_r(1).$$

Because A_i depends upon nuisance parameters, the limiting distribution is nuisance parameter dependent. How can one conduct inference in this case?

Consider first hypothesis tests concerning the complete vector of parameters and of the form

$$\text{H}_0 : \alpha = \alpha^*$$

for some specified value α^*. One sees immediately that

$$(\hat{\alpha} - \alpha^*)'\hat{A}^{-1\prime}\hat{A}^{-1}(\hat{\alpha} - \alpha^*) \xrightarrow{\text{d}} \mathcal{E}_r(1)'\mathcal{E}_r(1)$$

under the null, and diverges when the null is false. Although the limiting distribution of $\mathcal{E}_r(1)'\mathcal{E}_r(1)$ is non-standard (being the sum of squared independent exponential random variables), it does not depend upon nuisance parameters. The distribution of $\mathcal{E}_r(1)'\mathcal{E}_r(1)$ can be simulated easily, so it is relatively straightforward to calculate critical values or p-values in order to conduct the test. Note also that this result can be used to form a confidence region for the parameters α^0.

The presence of nuisance parameters in the single-element of α case prevents the direct application of the principles considered above. A useful alternative is to use a Lagrange multiplier type procedure. To illustrate this, suppose (without loss of generality) that one is interested in testing

$$\text{H}_0 : \alpha_1 = \alpha_1^*$$

for some specified value of α_1^*. In this case, the maximum likelihood estimator $\hat{\alpha}$ solves a set of r equations of the form

$$\bar{s}(\hat{\alpha}_1, \hat{\alpha}_2) = \underline{\hat{w}}$$

where α_2 corresponds to the remaining $(r-1)$ parameters. Consider the functions $\bar{s}(\cdot,\cdot)$ evaluated at the hypothesized value of α_1^*. One then obtains

$$\bar{s}(\alpha_1^*, \hat{\alpha}_2) \xrightarrow{\text{p}} \bar{s}(\alpha_1^*, \alpha_2^0)$$

and

$$\underline{\hat{w}} \xrightarrow{\text{p}} \bar{s}(\alpha_1^0, \alpha_2^0),$$

so a simple test of the null can be performed by examining the (normalized or weighted) differences

$$\left(\bar{s}(\alpha_1^*, \hat{\alpha}_2) - \underline{\hat{w}}\right)$$

which converge to zero under the null, and to some non-negative vector under alternative hypotheses. To construct a test statistic, define the following matrix:

$$J = [\nabla_\alpha \bar{s}(\alpha)]^{-1}$$

with $\hat{J}_T$ being the estimate of J based upon $\hat{\alpha}$. Also, denote the last $(r-1)$ rows of $\hat{J}_T$ by $\hat{J}_{T2}$ and the last $(r-1)$ rows of $\hat{J}_T^{-1}$ by $(\hat{J}_T)_2^{-1}$. Using results from the proof of Theorem 6, one can show that under the null

$$\hat{D}_T^{-1}\left(\bar{s}(\alpha_1^*, \hat{\alpha}_2) - \underline{\hat{w}}\right) \xrightarrow{\text{d}} Z\mathcal{E}_r(1)$$

where

$$Z = \operatorname*{plim}_{T\to\infty} \hat{Z} = \operatorname*{plim}_{T\to\infty} \hat{D}_T^{-1}\left(I_{r-1} + (\hat{J}_T^{-1})_2 \hat{J}_{T2}\right)\hat{D}_T$$

which we assume to exist and to be non-singular. It then follows that

$$\left(\bar{s}(\alpha_1^*, \hat{\alpha}_2) - \underline{\hat{w}}\right)' \hat{D}_T^{-1} (\hat{Z}'\hat{Z})^{-1} \hat{D}_T^{-1} \left(\bar{s}(\alpha_1^*, \hat{\alpha}_2) - \underline{\hat{w}}\right) \xrightarrow{\text{d}} \mathcal{E}_r(1)'\mathcal{E}_r(1)$$

which again does not depend upon nuisance parameters. It is straightforward to extend this result to tests concerning more than one parameter, and to show that the extension of this to tests concerning all r parameters gives a test that is asymptotically equivalent to the one based upon the result in Theorem 6.

Case c): k > r

In this case, the limiting distribution does depend upon nuisance parameters. Thus, although in principle one could proceed as in Case b), the limiting distribution is an inner product of random variables having a distribution given in Theorem 7. Whether it is possible to estimate p-values consistently based upon estimates of the nuisance parameters and simulation given these nuisance parameter estimates is a difficult question that we shall not address here. Although the regularity conditions needed and the actual implementation of this in practice are open to question, it does seem plausible that such an approach could work. In any case, Theorem 7 is interesting in its own right and is a step towards a general solution for a wide class of problems.

What other alternatives exist? One possibility is to note that the solution to any set of r constraints will be consistent at rate T. If one chooses a set randomly, then Theorem 6 would apply and inference could be conducted as in Case b). Moreover, it may be possible to combine test statistics in a deterministic way, such as averaging over all possible combinations of r constraints.

Thus, inference is possible to conduct within this framework. Alternative estimation strategies, that would result in $\sqrt{T}$ consistent and asymptotically normal estimates, are available but tests based upon such approaches are likely to be less powerful because of the slower rate of convergence. The fast convergence for the maximum likelihood estimator (or estimates based upon solving sets of support constraints) makes this approach most attractive in parametric models.

8. Some Monte Carlo Evidence

In this section, we use Monte Carlo methods to compare the small sample properties of the maximum likelihood estimator with those of the piecewise pseudo-maximum likelihood proposed by Donald and Paarsch (1993) and non-linear least squares estimators in the case of Dutch and first-price sealed-bid auctions.[12] For direct comparability, we have adopted the experimental design used in Donald and Paarsch (1993). Donald and Paarsch focus upon a procurement auction within the independent private values paradigm, where the lowest of the sealed bids is the winning bid. Potential bidders are assumed to be risk neutral. Thus, η equals one. In all of our simulation experiments we have assumed that $f(c)$, latent distribution of costs C, follows the Pareto law, so

$$f(c) = \frac{\alpha_2 \alpha_1^{\alpha_2}}{c^{\alpha_2+1}} \qquad 0 < \alpha_1 < c, \ 0 < \alpha_2.$$

In this case, the optimal bid function is

$$\sigma(c) = c + \frac{\int_c^\infty [1 - F(\xi)]^M \, d\xi}{[1 - F(c)]^M} = \frac{\alpha_2 M}{\alpha_2 M - 1} \cdot c.$$

[12] When the amount of simulation goes to infinity, the method of simulated non-linear least squares proposed by Laffont, Ossard, and Vuong (1991) becomes the method of non-linear least squares. Thus, the non-linear least squares results can be interpreted as the best-case scenarios for the estimator proposed by Laffont, Ossard, and Vuong (1991).

The density of $w = \beta(z)$ where $z = \min[c_1, \ldots, c_N]$ is then

$$h(w; \alpha_1, \alpha_2, M) = \frac{\alpha_2(M+1)\left(\frac{\alpha_1\alpha_2 M}{\alpha_2 M - 1}\right)^{\alpha_2(M+1)}}{w^{\alpha_2(M+1)+1}} \qquad \frac{\alpha_1\alpha_2 M}{\alpha_2 M - 1} < w,$$

while the j^{th} raw moment of w is

$$\mathrm{E}[w^j] = \left(\frac{\alpha_1\alpha_2 M}{\alpha_2 M - 1}\right)^j \frac{\alpha_2(M+1)}{\alpha_2(M+1) - j} \qquad j < \alpha_2(M+1),\ j = 1, 2, \ldots,$$

implying the following empirical specification for the first raw moment:

$$w = \left(\frac{\alpha_1\alpha_2 M}{\alpha_2 M - 1}\right) \frac{\alpha_2(M+1)}{\alpha_2(M+1) - 1} + u_1,$$

where u_1 has a mean of zero and a variance that depends upon M.

We fixed the values of (α_1^0, α_2^0) at $(1, 2)$. This implies that the expected value of c is two, while the variance of c does not exist. This latter implication has no effect upon our work, since it is the second raw moment of $z = \min[c_1, \ldots, c_N]$ that is important. Allowing c to have a very diffuse distribution also mimics some of the empirical evidence encountered in field data (see Paarsch [1992]). In any case, the second raw moment of z depends upon $N = (M+1)$ and exists in all of our experiments.

We considered three different sample sizes T: 50, 100, and 200. These values reflect the amount of data typically available. In particular, samples of size 50 would be common, while those of 200 would be considered large. In each of these samples, the number of bidders N could take on four different values: 3, 6, 9, and 12. This implies that the number of opponents M could take on the values 2, 5, 8, and 11. These values reflect the amount of competition that is often encountered in field data. Thus, in this model $k = 4 > 2 = r$.

We investigated three different patterns for the design matrix of the Ms, the probability distribution of the Ms, the $\{\pi(M)\}_{M=1}^{M}$s. In the first, each M was equally likely (Design A), while in the second, large Ms were more likely than small ones (Design B), and in the third, small Ms were more likely than large ones (Design C). In Table 0, we present the T_Ms and their corresponding $\pi(M)$s for the three different designs.

For the piecewise pseudo-maximum likelihood estimator, we partitioned the parameter vector $\alpha = (\alpha_1, \alpha_2)$ in two different ways, concentrating out first α_1 and then α_2. Below, we refer to these partitions as Partition 1 and Partition 2, respectively.

For Partition 1, the piecewise pseudo-maximum likelihood estimator of α_2 is

$$\hat{\alpha}_2^{\text{ppl}} = \frac{T}{\sum_{t=1}^{T}(M_t+1)\log\left(\frac{w_t}{\underline{\hat{w}}(M_t)}\right)},$$

but an estimator of α_1 can be defined in at least four different ways. First, consider any of

$$\hat{\alpha}_1^M = \underline{\hat{w}}(M)\left(\frac{\hat{\alpha}_2^{\text{ppl}}M - 1}{\hat{\alpha}_2^{\text{ppl}}M}\right) \quad M = 1,\ldots,\mathcal{M}.$$

Alternative estimators are

$$\hat{\alpha}_1^{\min} = \min[\hat{\alpha}_1^1, \hat{\alpha}_1^2, \ldots, \hat{\alpha}_1^{\mathcal{M}}],$$

$$\hat{\alpha}_1^{\text{a}} = \sum_{M=1}^{\mathcal{M}} \frac{T_M \times (M+1)}{\mathcal{N}_T}\hat{\alpha}_1^M,$$

and

$$\hat{\alpha}_1^{\text{b}} = \sum_{M=1}^{\mathcal{M}} \frac{T_M}{T}\hat{\alpha}_1^M,$$

where $\mathcal{N}_T = \sum_{t=1}^{T} N_t$. For Partition 2, the piecewise pseudo-maximum likelihood estimator of α_1 is defined implicitly (see Donald and Paarsch [1993]), and an estimator of α_2 can also be defined in at least four different ways. First, consider any of

$$\hat{\alpha}_2^M = \frac{\underline{\hat{w}}(M)}{\left(\underline{\hat{w}}(M) - \hat{\alpha}_1^{\text{ppl}}\right)M} \quad M = 1,\ldots,\mathcal{M}.$$

Alternative estimators are

$$\hat{\alpha}_2^{\min} = \min[\hat{\alpha}_2^1, \hat{\alpha}_2^2, \ldots, \hat{\alpha}_2^{\mathcal{M}}],$$

$$\hat{\alpha}_2^{\text{a}} = \sum_{M=1}^{\mathcal{M}} \frac{T_M \times (M+1)}{\mathcal{N}_T}\hat{\alpha}_2^M,$$

and

$$\hat{\alpha}_2^{\text{b}} = \sum_{M=1}^{\mathcal{M}} \frac{T_M}{T}\hat{\alpha}_2^M.$$

The random numbers for the experiments were generated using the multiplicative congruential method with modulus $(2^{31} - 1)$, multiplier 397204094, and initial seed 2420375. This method generates uniform pseudo-random numbers on the interval $(0, 1)$. (For more details, see Hall et al. 1988, pp. 232-235.) Using the property that the distribution function is distributed uniformly on the interval $(0, 1)$, we applied the inverse distribution function to obtain the pseudo-random ws.

We maximized the logarithm of the likelihood function subject to the T constraints using a slight modification of Schittkowski's (1981a,b) implementation of the (recursive) quadratic approximation method of Wilson (1963), Han (1976, 1977), and Powell (1978), see Vaesson (1984, pp. 57-66).

For Partition 2, we maximized the logarithm of the concentrated likelihood functions using the Newton-Raphson algorithm. We minimized the sum of squared residuals using the Gauss-Newton method. The true parameter values were used as starting values.

The results of the nine experiments are presented in Tables 1 to 9. The abbreviations St.Dev., L.Q., and U.Q. denote respectively the standard deviation, lower quartile, and upper quartile of the estimator's distribution. Also, in these tables the superscript upon an estimator denotes its type. For example, the 5 on $\hat{\alpha}_1^5$ implies that this is an estimator based upon Partition 1, for the case when $M = 5$. The "min" superscript denotes the minimum of all estimators with numeric superscripts and the same subscript. The superscripts "a" and "b" denote the type of averaging of the $\hat{\alpha}_i^M$s, where the "a" denotes the weights $(T_M \times (M + 1)/\mathcal{N}_T)$ and where the "b" denotes the weights (T_M/T). An estimator with the superscript "ppl" is the piecewise pseudo-maximum likelihood estimator (e.g., $\hat{\alpha}_1^{\text{ppl}}$ is the piecewise pseudo-maximum likelihood estimator based upon Partition 2), while one with the superscript "ml" is the maximum likelihood estimator, and one with the superscript "nls" is the non-linear least squares estimator.

As one can see, the rates of bias (when measured using either the mean or the median) for the maximum likelihood estimator are typically less that those of the other estimation methods. What is most stark about the performance of the maximum likelihood estimator is its quick convergence, which is suggested by the rate T convergence since $4 = k > r = 2$. Notice in Table 1 that for a sample size of 50 the standard deviation of $\hat{\alpha}_1^{\text{nls}}$ is 0.0236, while that of $\hat{\alpha}_1^{\text{ml}}$ is 0.0042. In Table 3, where the sample size is 200, the standard deviation of $\hat{\alpha}_1^{\text{nls}}$ is 0.0113, while that of $\hat{\alpha}_1^{\text{ml}}$ is 0.0005. These results are common across the the nine tables.

9. Summary and Conclusions

Under the assumption of self-interested, non-coöperative behaviour by potential bidders at auctions, game theory imposes a host of restrictions upon the data generating processes of either experimental or field data. Previous empirical research has examined only a few reduced-form predictions concerning these data generating processes. We have presented an integrated study of structural econometric methods for four different auction mechanisms within the IPVP, discussing issues of identification, estimation, and testing. This research provides a useful framework for deeper studies of strategic behaviour in markets.

A. Appendix

In this appendix, we present the proofs of the corollary, lemmata, and theorems contained in the paper.

Proof of Theorem 1.

Note that

$$\Pr[W < k \mid F, \eta^0] = F(\sigma^{-1}(k))^N$$

and

$$\Pr[W < k \mid G, \eta] = G(\mathcal{S}^{-1}(k))^N$$

where σ^{-1} depends upon the pair (F, η^0) and $\mathcal{S}^{-1}$ depends upon the pair (G, η). The result proceeds by showing that if $(F, \eta^0) \neq (G, \eta)$, then there exists a k for which these differ.

Introducing $\bar{s}^0 = \bar{s}(\theta^0, \eta^0, M, Z)$, first note that if $(F, \eta^0) \neq (G, \eta)$, and if

$$\bar{s}^0 = \bar{v} - \int_{v_0}^{\bar{v}} F(\xi)^{M\eta^0}\, d\xi \neq \bar{v} - \int_{v_0}^{\bar{v}} G(\xi)^{M\eta}\, d\xi,$$

then it is obvious that one can find a k such that $\Pr[W < k \mid F, \eta^0] \neq \Pr[W < k \mid G, \eta]$. Suppose that

$$\bar{s}^0 > \bar{v} - \int_{v_0}^{\bar{v}} G(\xi)^{M\eta}\, d\xi,$$

then let $k = \bar{s}^0$ and by Assumption 1,

$$\Pr[W < k \mid G, \eta] > \Pr[W < k \mid F, \eta^0] = 0,$$

while if

$$\bar{s}^0 < \bar{v} - \int_{v_0}^{\bar{v}} G(\xi)^{M\eta}\, d\xi$$

then, letting $k = \int_{v_0}^{\bar{v}} G(\xi)]^{M\eta}\, d\xi$, we have by Assumption 1,

$$0 = \Pr[W < k \mid G, \eta] < \Pr[W < k \mid F, \eta^0].$$

If, on the other hand, $(F, \eta^0) \neq (G, \eta)$ and

$$\bar{s}^0 = \bar{v} - \int_{v_0}^{\bar{v}} F(\xi)^{M\eta^0}\, d\xi = \bar{v} - \int_{v_0}^{\bar{v}} G(\xi)^{M\eta}\, d\xi, \qquad (A.1)$$

then there are two cases to consider. First, if $\eta \neq \eta^0$, then the densities at the upper bounds ($\bar{s}^0$) under (F, η^0) and (G, η) are such that

$$h_S(\bar{s}^0|F,\eta^0) = \frac{N}{M\eta^0(\bar{v}-\bar{s}^0)} \neq \frac{N}{M\eta(\bar{v}-\bar{s}^0)} = h_S(\bar{s}^0|G,\eta). \tag{A.2}$$

Since the density functions in both cases are continuous, there must exist an $\varepsilon > 0$ such that over the interval $[\bar{s}^0 - \varepsilon, \bar{s}^0)$ one density is larger than the other (depending upon whether $\eta^0 > \eta$ or $\eta^0 < \eta$). This implies that

$$\Pr[W < \bar{s}^0 - \varepsilon|F,\eta^0] \neq \Pr[W < \bar{s}^0 - \varepsilon|G,\eta]. \tag{A.3}$$

The second possibility assuming that (A.1) holds is for $\eta^0 = \eta$ and $F \neq G$. Under these conditions there exist disjoint intervals A_1 and A_2 such that for $v \in A_1$, $F < G$ and

$$F(v)^{M\eta^0} > G(v)^{M\eta} \tag{A.4}$$

and for $v \in A_2$, $F > G$

$$F(v)^{M\eta^0} < G(v)^{M\eta}. \tag{A.5}$$

Without loss of generality, suppose that A_1 occurs before A_2 as v increases. Let

$$v_l = \inf\{v : F(v) < G(v)\}.$$

Note that there exists a δ such that on $(v_l, v_l + \delta)$ $F(v) < G(v)$. This implies that $F(v)^{M\eta^0} > G(v)^{M\eta}$. Also, define

$$v_u = \inf\{v : v > v_l + \delta, F(v) = G(v)\}$$

where v_u exists because of (A.1). By construction $F(v_u) = G(v_u)$. By construction, for all $v \in (v_l, v_u)$, $F(v) < G(v)$. Thus,

$$F(v)^{M\eta^0} > G(v)^{M\eta}.$$

Let

$$k = k_u = \sigma(v_u)$$

so that $\sigma^{-1}(k_u) = v_u$. Note, however, that

$$\mathcal{S}^{-1}(k_u) > v_u.$$

To see why this is, note that by construction

$$\int_{v_u}^{\bar{v}} F(\xi)^{M\eta^0}\, d\xi < \int_{v_u}^{\bar{v}} G(\xi)^{M\eta}\, d\xi,$$

so that

$$k_u = \sigma(v_u) > \mathcal{S}(v_u)$$

since we also have that $F(v_u) = G(v_u)$. Hence, we have that

$$\begin{aligned}\Pr[W < k_u | F, \eta^0] &= F(v_u)^N \\ &= G(v_u)^N < G\big(\mathcal{S}^{-1}(k_u)\big)^N = \Pr[W < k_u | G, \eta]\end{aligned}$$

and the result follows.

Proof of Theorem 2.

Pick an $\alpha \in \bar{A}$ such that $\alpha \notin A^*$. Since A^* is a closed set, it must be that for some $\bar{x} \in X$,

$$\bar{s}(\alpha, \bar{x}) = \bar{s}(\alpha^0, \bar{x}) - \gamma \tag{A.6}$$

for some $\gamma > 0$. Pick some $\epsilon < \frac{1}{3}\gamma$. By Assumption 4, there exists $\delta_1 > 0$ and $\delta_2 > 0$ such that for all $x^1, x^2 \in X$ with $|x^1 - x^2| < \delta_1$ we have $|\bar{s}(\alpha, x^1) - \bar{s}(\alpha, x^2)| < \epsilon$ and when $|x^1 - x^2| < \delta_2$ we have $|\bar{s}(\alpha^0, x^1) - \bar{s}(\alpha^0, x^2)| < \epsilon$. Let $\delta = \min\{\delta_1, \delta_2\}$. Let $B(\bar{x}, \delta)$ be the open ball (in Euclidean space) around the point $\bar{x}$ of radius δ and let $N(\bar{x}) = B(\bar{x}, \delta) \cap X$ be a neighbourhood of the point $\bar{x}$ in X. By Assumption 2, $\Pr[x \in N(\bar{x})] > 0$. Let,

$$\bar{t} = \arg\max\{w_t - \bar{s}(\alpha^0, x_t) | x_t \in N(\bar{x})\}$$

and denote $\bar{\rho} = w_{\bar{t}} - \bar{s}(\alpha^0, x_{\bar{t}})$. Notice that $\bar{\rho}$ is negative. By Lemma A.1, $\bar{\rho} \xrightarrow{\text{a.s.}} 0$, so that there exists some finite $\bar{T}$ such that for $T \geq \bar{T}$ with probability one, $\bar{\rho} > -\epsilon$. We can now show that (also with probability 1) $\alpha \notin A^*_{\bar{T}}$ and since $A^*_T \subset A^*_{\bar{T}}$ for $T \geq \bar{T}$ the result will follow. To show this, suppose to the contrary, that $\alpha \in A^*_{\bar{T}}$. Then for $\bar{t}$ corresponding to $\bar{T}$ we have

$$\bar{s}(\alpha, x_{\bar{t}}) \geq w_{\bar{t}}.$$

But by construction (and using (A.6)), with probability 1,

$$\begin{aligned}\epsilon &\geq \bar{s}(\alpha^0, x_{\bar{t}}) - w_{\bar{t}} \\ &\geq \bar{s}(\alpha^0, x_{\bar{t}}) - \bar{s}(\alpha, x_{\bar{t}}) \\ &= \gamma + \big(\bar{s}(\alpha^0, x_{\bar{t}}) - \bar{s}(\alpha^0, \bar{x})\big) + \big(\bar{s}(\alpha, \bar{x}) - \bar{s}(\alpha, x_{\bar{t}})\big) \\ &> \gamma - 2\epsilon\end{aligned}$$

which implies that $3\epsilon > \gamma$, a contradiction. Therefore (with probability 1) $\alpha \notin A_T^*$ for $T \geq \bar{T}$ and the result follows.

Lemma A.1.

Given Assumptions 2 to 4, $\bar{\rho} \xrightarrow{\text{a.s.}} 0$.

Proof of Lemma A.1:

Note that by Assumptions 2-4,

$$\Pr[\bar{\rho} < -\epsilon | x_t \in N(\bar{x})] \leq \big(1 - \delta(\epsilon)\big)^{\hat{T}}$$

where $\hat{T}$ is the number of $x_t \in N(\bar{x})$ so that by Assumption 2

$$\frac{\hat{T}}{T} \xrightarrow{\text{a.s.}} \Pr[N(\bar{x})] > 0,$$

so that

$$\Pr[\bar{\rho} < -\epsilon | x_t \in N(\bar{x})] \xrightarrow{\text{a.s.}} 0.$$

Since $\Pr[\bar{\rho} > -\epsilon | x_t \in N(\bar{x})]$ is dominated by the integrable (with respect to the density function of $x_t \in N(\bar{x})$) function 1, then we have by the dominated convergence theorem that,

$$\Pr[\bar{\rho} < -\epsilon] = \mathrm{E}[\Pr[\bar{\rho} < -\epsilon | x_t \in N(\bar{x})]] \to 0,$$

so that $\bar{\rho} \xrightarrow{\text{P}} 0$ since ϵ is arbitrary. But since $\bar{\rho}$ is a monotonically increasing function we have by Proposition 1.2.1 of Rao (1986) that $\bar{\rho} \xrightarrow{\text{a.s.}} 0$.

Proof of Theorem 3.

This proof is similar to that of Gallant and Nychka (1987). Since $\hat{\alpha} \in \bar{A}$ for all T, a subsequence must converge to some point in $\bar{A}$; *i.e.*, $\hat{\alpha}_j \to \alpha^* \in \bar{A}$. But it must be that $\alpha^* \in A^*$ by condition b). All we need to show then is that $\alpha^* = \alpha^0$, which follows from c) and d) using the proof of Gallant and Nychka (1987).

Proof of Corollary 2.

Using Lemma 2 we can show that for all $z > 0$

$$\lim_{t \to \infty} \frac{1 - F\big(\bar{s}(\alpha^0, x(i)) - \frac{1}{zt}\big)}{1 - F\big(\bar{s}(\alpha^0, x(i)) - \frac{1}{t}\big)} = \frac{1}{z}$$

using L'Hôpital's rule, and $\bar{s}(x(i), \alpha^0) < \infty$ by Assumption 4, so the result follows by Lemma 1. It is also easy to show by L'Hôpital's rule that d_T is such that

$$\frac{d_T^*}{d_T} \to 1$$

almost surely where d_T^* is given in Lemma 1.

Proof of Lemma 1.

Note that

$$h_S(w) = \frac{NF(\sigma^{-1}(w))^M f(\sigma^{-1}(w))}{\sigma'(\sigma^{-1}(w))}$$

where we have written this density in terms of in terms of $U = \sigma^{-1}(W) = V_{(1:N)}$. The result follows by noting that

$$\sigma'(v) = M\eta f(v) \int_{v_0}^{v} F(\xi)^{M\eta}\, d\xi F(v)^{M\eta+1},$$

and that $w \to \bar{s}$ as $v \to \bar{v}$.

Proof of Theorem 4.

Note that for the corresponding population problem the Kuhn-Tucker conditions are satisfied at

$$(\alpha^0, \pi(x(1))\lambda_1^0, \pi(x(2))\lambda_2^0, ..., \pi(x(k))\lambda_k^0)$$

and by Assumptions 7 and 8 each of the Lagrange multipliers are positive, so each of the k constraints binds. Also, note that by Assumption 4, the matrix formed by the k vectors $\nabla_{\alpha_1}\bar{s}(\alpha, x(i))$ has full rank in a neighbourhood of α^0, which implies that the Kuhn-Tucker conditions are sufficient over this neighbourhood for a solution. Since $\hat{\alpha}$ is consistent, and since we have uniform convergence of the gradient vector, then the solution to the sample problem must also be such that all k constraints bind with Lagrange multipliers converging almost surely to the population values.

Proof of Theorem 5.

The proof is very similar to that in Donald and Paarsch (1993), Propositions 2 and 3. The result follows from standard mean value expansions, noting that since k constraints bind

$$\hat{\alpha}_1 = \psi(\hat{\alpha}_2, \underline{\hat{w}}, x).$$

One can then expand the first-order condition

$$\sqrt{T}\nabla_{\alpha_2} L_T(\psi(\hat{\alpha}_2, \underline{\hat{w}}, x, \hat{\alpha}_2)) = 0$$

about α_2^0, and ignore the pre-estimation error in $\hat{\underline{w}}$ since by Corollary 2 this is $O_p(T^{-1})$. The result for α_1 follows from the δ method applied to the function $\psi(\alpha_2, \hat{\underline{w}}, x)$, which is twice continuously differentiable in a neighbourhood of the true values.

Proof of Lemma 4.

Due to the uniform convergence of the gradient $\nabla_\alpha L$ over a neighbourhood of α^0 and the fact that $\hat{\alpha} \to \alpha^0$ almost surely the result follows by Assumption 9 using a similar argument to that used in the proof of Theorem 4.

Proof of Lemma 5.

Given the uniform convergence of the gradient $\nabla_\alpha L$ over a neighbourhood of α^0 and consistency of $\hat{\alpha}$ it must be the case that the solution to

$$\nabla_\alpha L(\hat{\alpha}) - \{\nabla_\alpha \bar{s}(\hat{\alpha})\}_{r\ell} \hat{\lambda}^\ell = 0$$

that exists for large T by Assumption 4, is such that $\hat{\lambda}^\ell \to \lambda_0^\ell$ almost surely. For those combinations that have an element of λ_0^ℓ that is negative, this implies that for large enough T the corresponding element of $\hat{\lambda}^\ell$ must be negative and hence the Kuhn-Tucker conditions cannot be satisfied at such a solution for large enough T. Hence, for large enough T for such combinations $\Pr[E(\ell)] = 0$.

Proof of Lemma 6.

For large enough T, all of the $\hat{\alpha}(\ell)$ are within a neighbourhood of α^0 using the result in Theorem 5, which applies to any of the $\hat{\alpha}(\ell)$. The uniform convergence and continuity of the derivatives over this neighbourhood and the fact that the constraint set is convex by Assumption 3 imply that if more than one $\hat{\alpha}(\ell)$ satisfy all of the Kuhn-Tucker conditions, and these two values differ, then there will be a violation of Assumption 9. The event that the two $\hat{\alpha}(\ell)$ solutions are identical occurs with probability 0, so the result holds.

Proof of Theorem 6.

By Theorem 3, the solution is such that all k constraints bind, so that $\hat{\alpha}$ can be determined by

$$\hat{\alpha} = \psi(\hat{\underline{w}}, x)$$

for some twice continuously differentiable function ψ. Note that $\alpha^0 = \psi(\bar{s}^0, x)$, so an expansion of $\psi(\hat{\underline{w}}, x)$ about $\bar{s}^0$ yields

$$\hat{\alpha} - \alpha^0 = J_T^{*-1}(\hat{\underline{w}} - \bar{s}^0)$$

where $J_T^* = \nabla_\alpha \bar{s}(\alpha^*)$ for some α^* lying between $\hat{\alpha}$ and α^0. Since $\hat{\alpha} \to \alpha^0$, then $\alpha^* \to \alpha^0$ and J_T^* is almost surely invertible by Assumption 3. Therefore,

$$-\hat{D}_T^{-1} J_T^*(\hat{\alpha} - \alpha^0) = \hat{D}_T^{-1}(\bar{s}^0 - \underline{\hat{w}})$$

which is distributed jointly asymptotically as a vector of independent $\mathcal{E}(1)$ random variables by Corollary 1. The result then follows by showing that $\hat{J}_T - J_T^* = o_p(1)$, which follows since both $\hat{J}_T$ and J_T^* converge to $\nabla \bar{s}(\alpha^0, x)$ because $\hat{\alpha} \xrightarrow{P} \alpha^0$ and $\alpha^* \xrightarrow{P} \alpha^0$.

Proof of Theorem 7.

It is easy to see that

$$\hat{B}_T(\hat{\alpha} - \alpha^0) - B_T^*(\hat{\alpha} - \alpha^0) \xrightarrow{P} 0$$

where

$$B_T^* = -\sum_{\ell \in \Xi_R} I[E(\ell)] D_{T\ell}^{-1} J_{T\ell}^*$$

with $D_{T\ell}^{-1}$ being the true normalization given in Corollary 1, and $J_{T\ell}^*$ being given in the mean value expansions in Theorem 5. Thus, it suffices to find the asymptotic distribution of $B_T^*(\hat{\alpha} - \alpha^0)$. Using the mean value expansions as in Theorem 5,

$$B_T^*(\hat{\alpha} - \alpha^0) = \sum_{\ell \in \Xi_R} I[E(\ell)] D_{T\ell}^{-1}(\bar{s}^0(\ell) - \underline{\hat{w}}(\ell))$$

where $\underline{\hat{w}}(\ell)$ and $\bar{s}^0(\ell)$ are the vectors formed using the ℓ^{th} combination of smallest order statistics and true lower bounds. Using similar expansions the event $E(\ell)$ occurs when

$$k_{\ell i}^{*\prime} D_{T\ell}^{-1}(\bar{s}^0(\ell) - \underline{\hat{w}}(\ell)) \leq -d_T^{-1}(i)(\bar{s}_i^0 - \underline{\hat{w}}_i)$$

for each $i \notin \Xi_\ell$ where

$$k_{\ell i}^{*\prime} = d_T^{-1}(i) \nabla_\alpha \bar{s}(\alpha^*, x(i)) J_{T\ell}^{*^{-1}} D_{T\ell}$$

for some α^* between $\hat{\alpha}$ and α^0 . Note that $k_{\ell i}^* \xrightarrow{P} k_{\ell i}$ since $\hat{\alpha} \xrightarrow{P} \alpha^0$. The vector consisting of

$$d_T^{-1}(i)(\bar{s}_i^0 \underline{\hat{w}}_i)$$

is distributed as a k vector of independent $\mathcal{E}(1)$ random variables. Thus, the asymptotic distribution reduces to finding the asymptotic distribution of

$$Z = \sum_{\ell \in \Xi_R} \bar{z}(\ell) I[E(\ell)]$$

where $\bar{z}(\ell)$ is the ℓ^{th} combination of the $\mathcal{E}(1)$ random variables and the event $E(\ell)$ occurs when

$$k'_{\ell i}\bar{z}(\ell) \leq \bar{z}_i$$

for all $i \notin \Xi_\ell$. The probability that $Z \leq z$ then can be computed as

$$\sum_{\ell \in \Xi_R} \Pr[E(\ell)] \Pr[\bar{z}(\ell) \leq z | E(\ell)]$$

$$= \sum_{\ell \in \Xi_R} \int_0^{z_1} \cdots \int_0^{z_r} (\prod_{i \notin \Xi_\ell} \int_{k'_{\ell i}\bar{z}(\ell)}^{\infty}) \prod_{j=1}^{k} f(\bar{z}_j) \prod_{j=1}^{k} d\bar{z}_j$$

$$= \sum_{\ell \in \Xi_R} \int_0^{z_1} \cdots \int_0^{z_r} \prod_{i \in \Xi_\ell} f(\bar{z}_i) \prod_{j \notin \Xi_\ell} [1 - \Pr(k'_{\ell i}\bar{z}(\ell))] \prod_{i \in \Xi_\ell} d\bar{z}_j$$

where the notation

$$\prod_{i=1}^{k} d\bar{z}_i = d\bar{z}_1 d\bar{z}_2 \ldots d\bar{z}_k$$

is used. This result implies that the density is as given in the Theorem. Note that the function Pr arises since $k'_{\ell i}\bar{z}(\ell)$ may be negative and the density for the $\mathcal{E}(1)$ variable is only defined over positive values.

B. Bibliography

Amemiya, T., 1985, *Advanced econometrics*, (Harvard University Press, Cambridge, MA)

Andrews, D., 1992, Generic uniform convergence, *Econometric Theory* 8, 241–257.

Coppinger, V., V. Smith, and J. Titus, 1980, Incentives and behavior in English, Dutch, and sealed-bid auctions, *Economic Inquiry* 18, 1–22

Donald, S. and H. Paarsch, 1992a, Maximum likelihood estimation in empirical models of auctions, Research report no. 9211 (Department of Economics, University of Western Ontario, London, Canada)

Donald, S. and H. Paarsch, 1992b, Identification in empirical models of auctions, Research report no. 9216 (Department of Economics, University of Western Ontario, London, Canada)

Donald, S. and H. Paarsch, 1993, Piecewise pseudo-maximum likelihood estimation in empirical models of auctions, *International Economic Review* 34, 121–148

Elyakime, B., J.-J. Laffont, P. Loisel, and Q. Vuong, forthcoming, First-price sealed-bid auctions with secret reservation prices, *Annales d'Economie et de Statistique*

Engle, R., 1984, Wald, likelihood ratio, and Lagrange multiplier tests in econometrics, in: Z. Griliches and M. Intriligator eds., *Handbook of Econometrics, Volume II*, (North Holland, New York)

Galambos J., 1978 *The asymptotic theory of extreme order statistics*, (John Wiley, New York)

Gallant, R. and D. Nychka, 1987, Semi-nonparametric maximum likelihood estimation, *Econometrica* 55, 363–390

Gallant, R. and G. Tauchen, 1992, Which moments to match? Typescript (Department of Economics, Duke University, Durham, North Carolina)

Hall, B., R. Schnake and C. Cummins, 1988, *TSP reference manual, Version 4.1*, (TSP International, Palo Alto)

Han, S., 1976, Superlinearly convergent variable metric algorithms for general nonlinear programming problems, *Mathematical Programming* 11, 263–282.

Han, S., 1977, A globally convergent method for nonlinear programming, *Journal of Optimization Theory and Applications* 22, 297–309.

Hansen, R., 1986, Sealed-bid versus open auctions: the evidence, *Economic Inquiry* 24, 125–142

Johnson, R., 1979, Oral auction versus sealed bids: an empirical investigation, *Natural Resources Journal* 19, 315–335

Laffont, J.-J., H. Ossard, and Q. Vuong, 1991, Econometrics of first-price auctions, Document de travail no. 7 (Institut d'Economie Industrielle, Toulouse, France)

LeCam, L., 1970, On the assumptions used to prove asymptotic normality of maximum likelihood estimates, *Annals of Mathematical Statistics* 41, 802–828

Mangasarian, O., 1969, *Nonlinear programming*, (McGraw-Hill, New York)

Mead, W., 1967, Natural resource disposal policy — oral auction versus sealed bids, *Natural Resources Journal* 7, 194–224

Newey, W., 1991, Uniform convergence in probability and stochastic equicontinuity, *Econometrica* 59, 1161–1167.

Paarsch, H., 1989, Empirical models of auctions within the independent private values paradigm and an application to British Columbian timber sales, Discussion paper no. 89–14 (Department of Economics, University of British Columbia, Vancouver, Canada)

Paarsch, H., 1991, Empirical models of auctions an application to British Columbian timber sales, Discussion paper no. 91–19 (Department of Economics, University of British Columbia, Vancouver, Canada)

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

Paarsch, H., 1993, Deriving an estimate of the optimal auction: an application to British Columbian timber sales, Research report no. 9315 (Department of Economics, University of Western Ontario, London, Ontario)

Paarsch, H., forthcoming, A comparison of estimators for empirical models of auctions, *Annales d'Economie et de Statistique*

Powell, M., 1978, A fast algorithm for nonlinearly constrained optimization calculations, in: G. Watson, ed., *Numerical analysis. Proceedings of the biennial conference held at Dundee, June 1977. Lecture notes in mathematics*, Volume 630, (Springer, Berlin)

Rao, B., 1987, *Asymptotic theory of statistical inference*, (John Wiley: New York)

Reiss, R.-D., 1989, *Approximate distributions of order statistics: with applications to nonparametric statistics*, (Springer-Verlag, New York)

Riley, J. and W. Samuelson, 1981, Optimal auctions, *American Economic Review* 71, 381–392

Schittkowski, K., 1981a, The nonlinear programming method of Wilson, Han, and Powell with an augmented Lagrangian type line search function. Part 1: Convergence analysis, *Numerische Mathematik* 38, 83–114

Schittkowski, K., 1981b, The nonlinear programming method of Wilson, Han, and Powell with an augmented Lagrangian type line search function. Part 2: An efficient implementation with linear least squares subproblems, *Numerische Mathematik* 38, 115–127

Smith, V., 1982, Microeconomic systems as an experimental science, *American Economic Review* 72, 923–955

Vaessen, W., 1984, *UBC NLP: Nonlinear function optimization*, (University of British Columbia Computing Centre, Vancouver, Canada)

Vickrey, W., 1961, Counterspeculation, auctions, and competitive sealed tenders, *Journal of Finance* 16, 8–37

Wald, A., 1949, Note on the consistency of the maximum likelihood estimate, *Annals of Mathematical Statistics*, 20, 595–601

White, H., (1982), Maximum likelihood estimation of misspecified models, *Econometrica* 50, 1–25.

Wilson, R., (1963), *A simplicial algorithm for concave programming*, Ph.D. dissertation, (Harvard University, Cambridge, MA)

Woodroofe, M., 1972, Maximum likelihood estimation of a translation parameter of a truncated distribution, *Annals of Mathematical Statistics* 43, 113–122

Table 0
Design Matrices of the T_Ms

Sample Size		50				100				200			
	M	2	5	8	11	2	5	8	11	2	5	8	11
Design A	T_M	13	12	12	13	25	25	25	25	50	50	50	50
	$\pi(M)$	0.26	0.24	0.24	0.26	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Design B	T_M	5	10	15	20	10	20	30	40	20	40	60	80
	$\pi(M)$	0.10	0.20	0.30	0.40	0.10	0.20	0.30	0.40	0.10	0.20	0.30	0.40
Design C	T_M	20	15	10	5	40	30	20	10	80	60	40	20
	$\pi(M)$	0.40	0.30	0.20	0.10	0.40	0.30	0.20	0.10	0.40	0.30	0.20	0.10

Table 1

Experiment 1: Design A, Sample Size = 50

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0394	0.0469	1.0098	1.0390	1.0733
$\hat{\alpha}_1^5$	1.0156	0.0165	1.0054	1.0160	1.0267
$\hat{\alpha}_1^8$	1.0098	0.0100	1.0034	1.0098	1.0164
$\hat{\alpha}_1^{11}$	1.0070	0.0072	1.0024	1.0069	1.0118
$\hat{\alpha}_1^{min}$	1.0012	0.0191	1.0009	1.0060	1.0109
$\hat{\alpha}_1^{a}$	1.0181	0.0195	1.0059	1.0187	1.0318
$\hat{\alpha}_1^{b}$	1.0128	0.0131	1.0045	1.0132	1.0219
$\hat{\alpha}_1^{ppl}$	1.0083	0.0098	1.0017	1.0086	1.0155
$\hat{\alpha}_1^{nls}$	1.0008	0.0236	0.9854	1.0007	1.0172
$\hat{\alpha}_1^{ml}$	1.0009	0.0042	0.9991	1.0004	1.0018
$\hat{\alpha}_2^2$	1.9764	0.0897	1.9215	1.9849	2.0395
$\hat{\alpha}_2^5$	2.0474	0.2114	1.8980	2.0397	2.1917
$\hat{\alpha}_2^8$	2.1752	0.3646	1.9165	2.1338	2.3973
$\hat{\alpha}_2^{11}$	2.3552	0.5608	1.9497	2.2648	2.6537
$\hat{\alpha}_2^{min}$	1.8941	0.1669	1.8033	1.9305	2.0160
$\hat{\alpha}_2^{a}$	2.1397	0.2796	1.9332	2.1113	2.3122
$\hat{\alpha}_2^{b}$	2.2049	0.3636	1.9348	2.1590	2.4215
$\hat{\alpha}_2^{ppl}$	2.2141	0.3256	1.9875	2.1806	2.4124
$\hat{\alpha}_2^{nls}$	2.0477	0.2738	1.8591	2.0362	2.2169
$\hat{\alpha}_2^{ml}$	1.9815	0.0925	1.9461	1.9817	2.0063

Table 2

Experiment 2: Design A, Sample Size = 100

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0189	0.0343	0.9969	1.0196	1.0413
$\hat{\alpha}_1^5$	1.0076	0.0113	1.0002	1.0074	1.0150
$\hat{\alpha}_1^8$	1.0047	0.0068	1.0003	1.0045	1.0093
$\hat{\alpha}_1^{11}$	1.0034	0.0050	1.0002	1.0033	1.0069
$\hat{\alpha}_1^{min}$	0.9975	0.0165	0.9960	1.0030	1.0064
$\hat{\alpha}_1^{a}$	1.0086	0.0140	0.9998	1.0088	1.0181
$\hat{\alpha}_1^{b}$	1.0062	0.0095	1.0001	1.0063	1.0124
$\hat{\alpha}_1^{ppl}$	1.0042	0.0072	0.9997	1.0045	1.0091
$\hat{\alpha}_1^{nls}$	1.0015	0.0172	0.9902	1.0024	1.0130
$\hat{\alpha}_1^{ml}$	1.0003	0.0013	0.9997	1.0001	1.0008
$\hat{\alpha}_2^2$	1.9866	0.0560	1.9535	1.9908	2.0232
$\hat{\alpha}_2^5$	2.0219	0.1427	1.9237	2.0209	2.1063
$\hat{\alpha}_2^8$	2.0861	0.2431	1.9245	2.0640	2.2284
$\hat{\alpha}_2^{11}$	2.1634	0.3587	1.9163	2.1151	2.3584
$\hat{\alpha}_2^{min}$	1.9247	0.1311	1.8692	1.9602	2.0190
$\hat{\alpha}_2^{a}$	2.0645	0.1899	1.9352	2.0449	2.1790
$\hat{\alpha}_2^{b}$	2.0943	0.2421	1.9313	2.0680	2.2321
$\hat{\alpha}_2^{ppl}$	2.0969	0.2163	1.9470	2.0772	2.2299
$\hat{\alpha}_2^{nls}$	2.0342	0.2017	1.9011	2.0312	2.1689
$\hat{\alpha}_2^{ml}$	1.9908	0.0294	1.9774	1.9926	2.0008

Table 3

Experiment 3: Design A, Sample Size = 200

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0109	0.0226	0.9956	1.0110	1.0274
$\hat{\alpha}_1^5$	1.0042	0.0076	0.9988	1.0040	1.0096
$\hat{\alpha}_1^8$	1.0026	0.0045	0.9994	1.0029	1.0056
$\hat{\alpha}_1^{11}$	1.0019	0.0033	0.9998	1.0020	1.0041
$\hat{\alpha}_1^{min}$	0.9976	0.0111	0.9956	1.0016	1.0040
$\hat{\alpha}_1^{a}$	1.0049	0.0094	0.9987	1.0051	1.0117
$\hat{\alpha}_1^{b}$	1.0035	0.0064	0.9992	1.0036	1.0080
$\hat{\alpha}_1^{ppl}$	1.0024	0.0047	0.9991	1.0027	1.0056
$\hat{\alpha}_1^{nls}$	1.0001	0.0113	0.9927	1.0004	1.0077
$\hat{\alpha}_1^{ml}$	1.0001	0.0005	1.0000	1.0001	1.0003
$\hat{\alpha}_2^2$	1.9955	0.0349	1.9710	1.9983	2.0206
$\hat{\alpha}_2^5$	2.0177	0.0892	1.9555	2.0179	2.0794
$\hat{\alpha}_2^8$	2.0521	0.1517	1.9429	2.0466	2.1503
$\hat{\alpha}_2^{11}$	2.0910	0.2165	1.9356	2.0780	2.2286
$\hat{\alpha}_2^{min}$	1.9538	0.0908	1.9134	1.9829	2.0187
$\hat{\alpha}_2^{a}$	2.0391	0.1193	1.9531	2.0335	2.1158
$\hat{\alpha}_2^{b}$	2.0551	0.1505	1.9451	2.0467	2.1512
$\hat{\alpha}_2^{ppl}$	2.0563	0.1422	1.9545	2.0434	2.1513
$\hat{\alpha}_2^{nls}$	2.0137	0.1381	1.9239	2.0162	2.1062
$\hat{\alpha}_2^{ml}$	1.9946	0.0107	1.9927	1.9949	1.9998

Table 4

Experiment 4: Design B, Sample Size = 50

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0616	0.0601	1.0215	1.0596	1.0982
$\hat{\alpha}_1^5$	1.0172	0.0173	1.0059	1.0168	1.0285
$\hat{\alpha}_1^8$	1.0088	0.0098	1.0026	1.0090	1.0152
$\hat{\alpha}_1^{11}$	1.0058	0.0067	1.0016	1.0059	1.0104
$\hat{\alpha}_1^{min}$	1.0019	0.0167	1.0008	1.0055	1.0101
$\hat{\alpha}_1^{a}$	1.0146	0.0138	1.0060	1.0146	1.0245
$\hat{\alpha}_1^{b}$	1.0101	0.0100	1.0038	1.0104	1.0173
$\hat{\alpha}_1^{ppl}$	1.0060	0.0085	1.0007	1.0063	1.0118
$\hat{\alpha}_1^{nls}$	1.0021	0.0263	0.9857	1.0047	1.0203
$\hat{\alpha}_1^{ml}$	0.9996	0.0044	0.9978	0.9999	1.0014
$\hat{\alpha}_2^2$	1.8644	0.1497	1.7817	1.8951	1.9749
$\hat{\alpha}_2^5$	1.9779	0.1995	1.8441	1.9775	2.1178
$\hat{\alpha}_2^8$	2.1130	0.2934	1.9100	2.0869	2.2881
$\hat{\alpha}_2^{11}$	2.2613	0.4595	1.9453	2.1970	2.5141
$\hat{\alpha}_2^{min}$	1.8038	0.1676	1.6971	1.8148	1.9309
$\hat{\alpha}_2^{a}$	2.1205	0.2983	1.9154	2.0911	2.2969
$\hat{\alpha}_2^{b}$	2.1658	0.3456	1.9283	2.1256	2.3668
$\hat{\alpha}_2^{ppl}$	2.2148	0.3264	1.9993	2.1841	2.4113
$\hat{\alpha}_2^{nls}$	2.1013	0.3812	1.8265	2.0774	2.3681
$\hat{\alpha}_2^{ml}$	1.9575	0.1354	1.9028	1.9682	2.0038

Table 5

Experiment 5: Design B, Sample Size = 100

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0284	0.0368	1.0048	1.0274	1.0533
$\hat{\alpha}_1^5$	1.0082	0.0118	1.0005	1.0082	1.0161
$\hat{\alpha}_1^8$	1.0043	0.0067	1.0000	1.0042	1.0086
$\hat{\alpha}_1^{11}$	1.0028	0.0048	0.9998	1.0028	1.0061
$\hat{\alpha}_1^{\min}$	0.9985	0.0142	0.9983	1.0026	1.0060
$\hat{\alpha}_1^{a}$	1.0069	0.0095	1.0009	1.0069	1.0135
$\hat{\alpha}_1^{b}$	1.0048	0.0071	1.0003	1.0049	1.0096
$\hat{\alpha}_1^{ppl}$	1.0030	0.0061	0.9991	1.0031	1.0072
$\hat{\alpha}_1^{nls}$	1.0006	0.0200	0.9885	1.0020	1.0156
$\hat{\alpha}_1^{ml}$	0.9997	0.0020	0.9989	0.9998	1.0006
$\hat{\alpha}_2^2$	1.9296	0.0870	1.8836	1.9466	1.9905
$\hat{\alpha}_2^5$	1.9850	0.1237	1.9015	1.9878	2.0651
$\hat{\alpha}_2^8$	2.0521	0.2007	1.9191	2.0354	2.1708
$\hat{\alpha}_2^{11}$	2.1210	0.2915	1.9291	2.0881	2.2858
$\hat{\alpha}_2^{\min}$	1.8834	0.1158	1.8187	1.9002	1.9702
$\hat{\alpha}_2^{a}$	2.0540	0.1984	1.9188	2.0371	2.1704
$\hat{\alpha}_2^{b}$	2.0758	0.2277	1.9234	2.0563	2.2063
$\hat{\alpha}_2^{ppl}$	2.0953	0.2149	1.9479	2.0756	2.2277
$\hat{\alpha}_2^{nls}$	2.0470	0.2836	1.8565	2.0343	2.2359
$\hat{\alpha}_2^{ml}$	1.9769	0.0596	1.9473	1.9834	2.0027

Table 6

Experiment 6: Design B, Sample Size = 200

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0161	0.0246	1.0005	1.0160	1.0315
$\hat{\alpha}_1^5$	1.0046	0.0077	0.9993	1.0047	1.0102
$\hat{\alpha}_1^8$	1.0025	0.0046	0.9994	1.0025	1.0056
$\hat{\alpha}_1^{11}$	1.0016	0.0032	0.9995	1.0017	1.0038
$\hat{\alpha}_1^{min}$	0.9982	0.0102	0.9977	1.0015	1.0038
$\hat{\alpha}_1^a$	1.0039	0.0065	0.9996	1.0038	1.0084
$\hat{\alpha}_1^b$	1.0027	0.0048	0.9996	1.0027	1.0061
$\hat{\alpha}_1^{ppl}$	1.0018	0.0040	0.9991	1.0019	1.0046
$\hat{\alpha}_1^{nls}$	1.0005	0.0135	0.9921	1.0021	1.0102
$\hat{\alpha}_1^{ml}$	0.9998	0.0008	0.9997	1.0000	1.0001
$\hat{\alpha}_2^2$	1.9634	0.0530	1.9377	1.9718	2.0003
$\hat{\alpha}_2^5$	1.9982	0.0810	1.9386	1.9964	2.0558
$\hat{\alpha}_2^8$	2.0327	0.1262	1.9423	2.0313	2.1188
$\hat{\alpha}_2^{11}$	2.0700	0.1840	1.9386	2.0595	2.1859
$\hat{\alpha}_2^{min}$	1.9322	0.0793	1.8855	1.9487	1.9938
$\hat{\alpha}_2^a$	2.0338	0.1274	1.9430	2.0302	2.1176
$\hat{\alpha}_2^b$	2.0457	0.1454	1.9401	2.0412	2.1395
$\hat{\alpha}_2^{ppl}$	2.0566	0.1430	1.9561	2.0437	2.1509
$\hat{\alpha}_2^{nls}$	2.0301	0.1955	1.9002	2.0326	2.1579
$\hat{\alpha}_2^{ml}$	1.9871	0.0247	1.9785	1.9932	1.9970

Table 7

Experiment 7: Design C, Sample Size = 50

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0343	0.0456	1.0060	1.0357	1.0648
$\hat{\alpha}_1^5$	1.0140	0.0159	1.0041	1.0141	1.0248
$\hat{\alpha}_1^8$	1.0106	0.0104	1.0036	1.0105	1.0174
$\hat{\alpha}_1^{11}$	1.0122	0.0107	1.0052	1.0109	1.0172
$\hat{\alpha}_1^{\min}$	1.0017	0.0207	1.0010	1.0075	1.0125
$\hat{\alpha}_1^{a}$	1.0213	0.0252	1.0057	1.0222	1.0381
$\hat{\alpha}_1^{b}$	1.0167	0.0177	1.0057	1.0174	1.0286
$\hat{\alpha}_1^{ppl}$	1.0127	0.0127	1.0046	1.0134	1.0220
$\hat{\alpha}_1^{nls}$	1.0018	0.0264	0.9846	0.9999	1.0183
$\hat{\alpha}_1^{ml}$	0.9997	0.0050	0.9983	1.0002	1.0016
$\hat{\alpha}_2^2$	2.0300	0.0914	1.9698	2.0312	2.0938
$\hat{\alpha}_2^5$	2.1743	0.2733	1.9719	2.1631	2.3652
$\hat{\alpha}_2^8$	2.3539	0.5242	1.9655	2.2761	2.6606
$\hat{\alpha}_2^{11}$	2.4680	0.9338	1.8133	2.2506	2.8571
$\hat{\alpha}_2^{\min}$	1.9032	0.2472	1.7656	1.9929	2.0781
$\hat{\alpha}_2^{a}$	2.1819	0.2846	1.9674	2.1511	2.3521
$\hat{\alpha}_2^{b}$	2.2580	0.4019	1.9551	2.2088	2.4817
$\hat{\alpha}_2^{ppl}$	2.2126	0.3285	1.9901	2.1799	2.3982
$\hat{\alpha}_2^{nls}$	2.0457	0.2430	1.8817	2.0291	2.1965
$\hat{\alpha}_2^{ml}$	1.9618	0.0773	1.9419	1.9761	2.0003

Table 8

Experiment 8: Design C, Sample Size = 100

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0161	0.0333	0.9949	1.0162	1.0384
$\hat{\alpha}_1^5$	1.0067	0.0111	0.9996	1.0064	1.0141
$\hat{\alpha}_1^8$	1.0052	0.0072	1.0005	1.0050	1.0102
$\hat{\alpha}_1^{11}$	1.0058	0.0063	1.0017	1.0055	1.0094
$\hat{\alpha}_1^{min}$	0.9976	0.0175	0.9948	1.0034	1.0076
$\hat{\alpha}_1^{a}$	1.0101	0.0183	0.9981	1.0101	1.0223
$\hat{\alpha}_1^{b}$	1.0079	0.0129	0.9996	1.0078	1.0167
$\hat{\alpha}_1^{ppl}$	1.0068	0.0094	1.0004	1.0071	1.0131
$\hat{\alpha}_1^{nls}$	1.0012	0.0187	0.9885	1.0011	1.0137
$\hat{\alpha}_1^{ml}$	0.9999	0.0014	0.9991	0.9999	1.0006
$\hat{\alpha}_2^2$	2.0170	0.0608	1.9788	2.0163	2.0566
$\hat{\alpha}_2^5$	2.0919	0.1896	1.9603	2.0810	2.2167
$\hat{\alpha}_2^8$	2.1733	0.3346	1.9343	2.1193	2.3882
$\hat{\alpha}_2^{11}$	2.2257	0.5366	1.8691	2.1311	2.4596
$\hat{\alpha}_2^{min}$	1.9296	0.1816	1.8512	1.9957	2.0524
$\hat{\alpha}_2^{a}$	2.0916	0.1920	1.9586	2.0697	2.2059
$\hat{\alpha}_2^{b}$	2.1281	0.2640	1.9401	2.0925	2.2746
$\hat{\alpha}_2^{ppl}$	2.0948	0.2146	1.9474	2.0716	2.2321
$\hat{\alpha}_2^{nls}$	2.0225	0.1747	1.8981	2.0142	2.1431
$\hat{\alpha}_2^{ml}$	1.9842	0.0204	1.9736	1.9906	1.9979

Table 9

Experiment 9: Design C, Sample Size = 200

Estimator	Mean	St.Dev.	L.Q.	Median	U.Q.
$\hat{\alpha}_1^2$	1.0096	0.0225	0.9940	1.0091	1.0253
$\hat{\alpha}_1^5$	1.0039	0.0076	0.9985	1.0040	1.0093
$\hat{\alpha}_1^8$	1.0029	0.0046	0.9998	1.0030	1.0060
$\hat{\alpha}_1^{11}$	1.0031	0.0038	1.0004	1.0032	1.0056
$\hat{\alpha}_1^{min}$	0.9976	0.0116	0.9940	1.0021	1.0046
$\hat{\alpha}_1^a$	1.0059	0.0124	0.9971	1.0059	1.0147
$\hat{\alpha}_1^b$	1.0046	0.0087	0.9986	1.0047	1.0106
$\hat{\alpha}_1^{ppl}$	1.0038	0.0062	0.9997	1.0040	1.0080
$\hat{\alpha}_1^{nls}$	1.0000	0.0122	0.9915	1.0007	1.0081
$\hat{\alpha}_1^{ml}$	1.0000	0.0005	0.9997	1.0000	1.0002
$\hat{\alpha}_2^2$	2.0117	0.0398	1.9831	2.0129	2.0380
$\hat{\alpha}_2^5$	2.0513	0.1190	1.9638	2.0454	2.1299
$\hat{\alpha}_2^8$	2.0961	0.2090	1.9441	2.0855	2.2285
$\hat{\alpha}_2^{11}$	2.1184	0.3012	1.9034	2.0803	2.3001
$\hat{\alpha}_2^{min}$	1.9526	0.1237	1.8900	2.0070	2.0374
$\hat{\alpha}_2^a$	2.0511	0.1201	1.9610	2.0440	2.1273
$\hat{\alpha}_2^b$	2.0703	0.1624	1.9499	2.0585	2.1733
$\hat{\alpha}_2^{ppl}$	2.0560	0.1425	1.9549	2.0450	2.1501
$\hat{\alpha}_2^{nls}$	2.0101	0.1160	1.9312	2.0133	2.0919
$\hat{\alpha}_2^{ml}$	1.9926	0.0080	1.9893	1.9949	1.9970

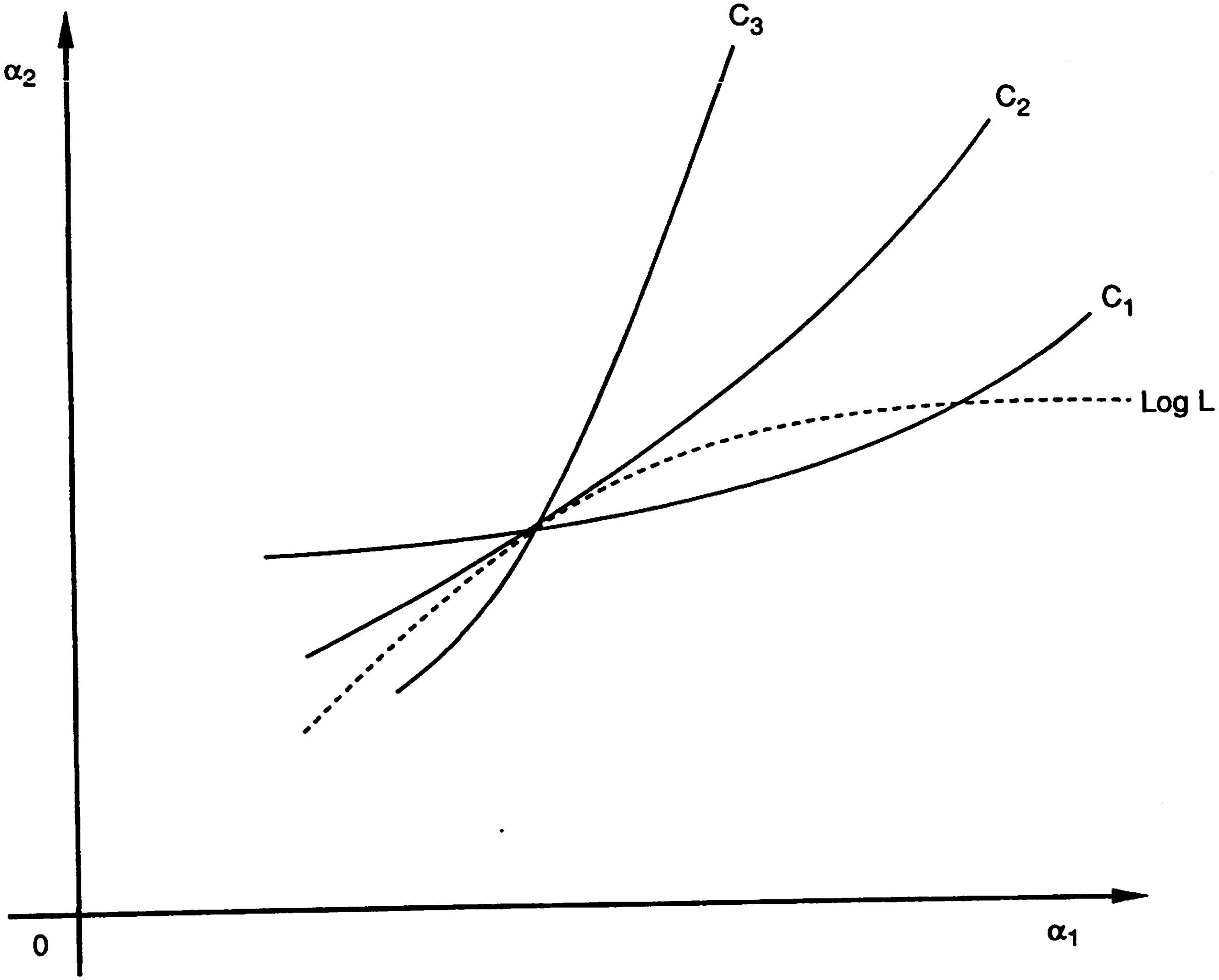

Figure 1

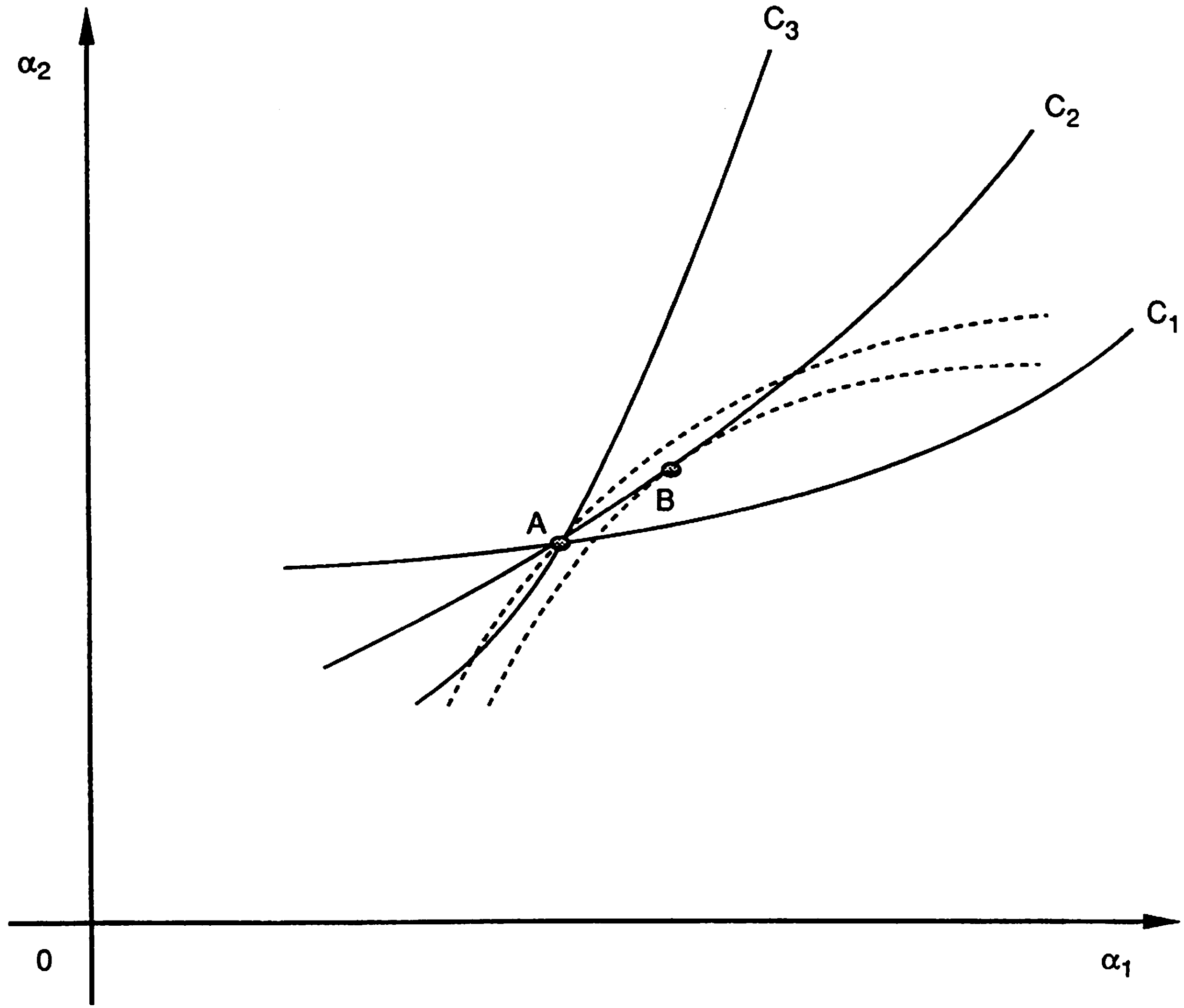

Figure 2