A Survey of Recent work on Identification, Estimation and Testing of Structural Auction Models*

Samita Sareen
Research and Risk Management
Financial Markets Department
Bank of Canada

December 24, 2000

Abstract

This survey contributes to the existing surveys on empirical work in auction models in three ways. First, it incorporates recent progress in establishing nonparametric identification of several structural auctions models not identified from bid data. Bidder identity is used to establish identification of asymmetric models. Variation in the ex post value of the auctioned object or the number of potential bidders is used to distinguish the private-values model from the common-value model. Second, the power of Bayesian tools in estimation and testing of structural auctions models is discussed. The focus is on different methods of evaluating the likelihood function and specification of noninformative priors in models where the support of the data depends on the parameters as in structural auction models. Finally, approaches to modelling unobserved heterogeneity are described.

*I have benefitted from several clarifications by Susan Athey, Patrick Bajari, Phil Haile and Matthew Shum.

1 Introduction

Auctions are a fast, transparent, fair and economically efficient mechanism for allocating goods. This is demonstrated by the long and impressive list of commodities being bought and sold through auctions. Auctions are used to sell agricultural products like eggplant and flowers. Procurement auctions are being conducted by government agencies to sell the right to fell timber, drill offshore areas for oil, procure crude-oil for refining, sell milk quotas, etc. In the area of finance auctions have been used by central banks to sell Treasury bills and bonds. The assets of bankrupt firms are being sold through auctions. Transactions between buyers and sellers in a stock-exchange and foreign-exchange markets takes the form of double auctions. Licenses for radio spectrum were awarded on the basis of "beauty contests" previously; spectrum auctions have now become routine in several countries. Recently, five third-generation mobile telecommunication licenses have been auctioned off in the United Kingdom for an unprecedented and unexpected sum of £ 22.47 billion indicating that a carefully designed auction mechanism will provide incentives for the bidders to reveal their true value. Finally, internet sites like eBay, Yahoo and Amazon are being used to bring together buyers and sellers across the world to sell goods ranging from airline tickets to laboratory ventilation hoods (Lucking-Reiley, 2000). The experience with spectrum and internet auctions has demonstrated that auction mechanism can be designed in a manner to make bidder's reveal their "true" value for the auctioned object not just in laboratories but in real-life as well.

Insights into the design of auctions has been obtained by modelling auctions as games of incomplete information. A survey of the theory developed for single-unit auctions can be found in McAfee and McMillan (1987) and Milgrom and Weber (1982) and for multi-unit auctions in Weber (1983). A structural or game-theoretic auction model emerges once a seller pre-commits to certain auction form and rules and assumptions are made about the nature of "incomplete" information possessed by the bidders. The latter includes assumptions about the risk attitude of the bidders, the relationship between bidder's valuations, whether the "true" value of the object is known to them and whether bidders are symmetric up to information differences. The rules and form of the auctions are determined by the seller in a manner which provides the bidders with incentives to reveal their valuation for the auctioned object.

Empirical work in auctions using field data has taken two directions. The first direction, referred to as the structural approach, attempts to recover the structural elements of the game-theoretic model. For example, if bidders are assumed to be risk-neutral, the structural element of the game-theoretic model would be the underlying probability law of valuations of the bidders. Any game-theoretic or structural model makes ceratin nonparametric predictions. For example, the revenue equivalence theorem is a prediction of a single-unit, independent-private-values auction model with symmetric and risk-neutral players. The second approach, referred to as the reduced-form approach, tests the predictions of the underlying game-theoretic model.

Identification of a private-values, second-price auction is trivial if there is a nonbinding reserve price and each bidders' bid in an auction is available. In a private-values auction a bidder knows the value of the auctioned object to herself.

In a second-price auction the bidder pays the value of the object to the second-highest bidder. Hence in a private-values, second-price auction, a bidder submits a bid equal to her value of the auctioned object. If the bids of all bidders in an auction were observed, then in essence the private values of all bidders are observed. Identification then amounts to identifying the distribution of bidders' valuations from data on these valuations.

Outside this scenario identification, estimation and testing of structural models is a difficult exercise. First, outside the private-values, second-price auction, bidders do not bid their valuation of the object. In most cases, an explicit solution is not available for the Bayesian-Nash equilibrium strategy. Second, since all bids are not observed, identification and estimation now involves recovering the distribution of bidder valuations when you observe a sample of order statistics from this distribution. The properties of these estimators have to be investigated since they do not have standard $\sqrt{T}$asymptotics, where T is the number of auctions in the sample.

Laffont (1997), Hendricks and Paarsch (1995) and Perrigne and Vuong (1999) provide a survey of issues in empirical work in auctions using field data. Perrigne and Vuong focus on identification and estimation of single-unit, first-price, sealed-bid structural auctions when data on all bids in an auction is available. Bidders are assumed to be risk-neutral and know the "true" value of the auctioned object. In addition to surveying the empirics of the above mentioned structural auction model, Laffont and Hendricks and Paarsch discuss several approaches to testing predictions of various single-unit structural auction models through the reduced-form approach. The focus of the former is the work by Hendricks, Porter and their coauthors on drainage sales in the OCS auctions.

The current survey adds to these surveys in *at least* three ways.

First, Guerre, Perrigne and Vuong (2000) have established that the common value model and private value model cannot be distinguished if only data on all bids in an auction are observed.[1] Distinguishing between the two models of valuation is important for several reasons. First, the two models of valuations imply different bidding and participation behavior; this may be of interest by itself. Second, depending on the model of valuation, the optimal auction is different.

The need to distinguish the two models of valuation raises the question as to what additional data would be required to distinguish the two models. Several papers have addressed this issue recently: Hendricks, Pinkse and Porter (1999), Athey and Haile (2000) and Hong, Shum and Haile (2000). Section 3 integrates these papers with the currently existing work of work by Guerre, Perrigne and Vuong (2000) on identification of several classes of private-values models from bid data. Exploiting results on competing risks models as in Berman (1963) and the multi-sector Roy model as in Heckman and Honoré (1990), Athey and Haile establish identification of several asymmetric models if additional data in the form of the identity of the winner is observed. Hendricks, Pinkse and Porter (1999) and Hong, Shum and Haile (2000) use the ex post value of the auctioned object and the variation in the number of potential bidders, respectively to distinguish the common-value model from the private-values model.

[1] Parametric assumptions on the joint distribution of private and public signals about the the value of the auctioned object will identify the two models.

Second, existing surveys make a cursory mention of the power of Bayesian tools in estimation and testing of structural auction models. This gap in the literature is remedied in this survey in Section 4.

Third, several field data sets exhibit heterogeneity in the auctioned object violating one of the assumption made in the theory of auctions that homogeneous objects are being auctioned. The issue is how to model object heterogeneity and still be able to use the results in auction theory that a single homogenous object is being auctioned. Available approaches to model object heterogeneity are discussed in Section 5.

Section 6 concludes with some directions for future research.

2 Definitions and Notation

The framework used in this paper follows Milgrom and Weber (1982). Unless otherwise mentioned, a single and indivisible object is auctioned by a single seller to n potential bidders. The subscript i will indicate these potential bidders with $i = 1, ..., n$. X_i is bidder i's private signal about the auctioned object with $\mathbf{X} = (X_1, ..., X_n)$ indicating the private signals of the n potential bidders. $\mathbf{S} = (S_1, ..., S_m)$ are additional m random variables which affect the value of the auctioned object for *all* bidders. These are the *common* components affecting the utility function of all bidders. $U_i(\mathbf{X}, \mathbf{S})$ is the utility function of bidder i; $u_i \geq 0$; it assumed to be continuous and non decreasing in its variables. The probability density function and the distribution function of the random variable $\mathbf{X}$ will be indicated by $f_{\mathbf{X}}(\mathbf{x})$ and $F_{\mathbf{X}}(\mathbf{x})$, respectively. The joint distribution of $(\mathbf{X}, \mathbf{S})$ is defined on the support $[\underline{x}, \overline{x}]^n \times [\underline{s}, \overline{s}]^m$. If $\mathbf{X}$ is an n dimensional vector with elements X_i, then the $n-1$ dimensional vector *excluding* element X_i will be indicated by $\mathbf{X}_{-i}$. The support of a variable Z, if it depends on θ, will be indicated by $\zeta_Z(\theta)$.

The i^{th} order statistic from a sample of size n will be denoted by $X^{i:n}$, with $X^{1:n}$ and $X^{n:n}$ being the lowest and the highest order statistic, respectively. If $X^{i:n}$ is the i^{th} order statistic from the distribution F_x, it's distribution function will be indicated by $F_x^{i:n}$ and density function by $f_x^{i:n}$.

Once the seller has pre-committed to a set of rules, the genesis of a structural model is in the following optimization exercise performed by each of the $n+1$ players; each player i chooses his equilibrium bid b_i by maximizing

$$E_{\mathbf{X}_{-i}\mathbf{S}|x_i}\left[(U_i(\mathbf{X}, \mathbf{S}) - b_i)\Pr(\text{Bidder i wins})\right]. \tag{1}$$

The concept of equilibrium employed is the Bayesian-Nash equilibrium. For example, if the seller pre-commits to the rules of a first-price auction, the first-order conditions for the maximization problem are the following set of differential equations,

$$v(x_i, x_i; n) = b_i + \frac{\Pr\left(\max_{j\neq n} B_j \leq b_i | B_i = b_i\right)}{\frac{\partial}{\partial\tau}\Pr\left(\max_{j\neq n} B_j \leq \tau | B_i = b_i\right)|_{\tau=b_i}} \equiv \xi(b_i, F_{\mathbf{B}}(\mathbf{b}); n), \tag{2}$$

where $v(x_i, y_i; n) = E\left(U_i | X_i = x_i, X_{-i}^{n-1:n-1} = y_i\right)$. $X_{-i}^{n-1:n-1}$ is the maximum over the variables $X_1, ..., X_n$ *excluding* X_i. This set of n differential equations will simplify to a single equation if bidders are symmetric. The first-order conditions involve variables that are both observed and unobserved to the econometrician.[2] The solution of this set of differential equations gives the equilibrium strategy

$$b_i = e_i\left(v(x_i, y_i; n), F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s}), n\right). \tag{3}$$

The equilibrium strategy of a bidder differs depending on the assumptions made about the auction format or rules and the model of valuation.[3] The structure of the utility function and the relationship between the variables $(\mathbf{X}, \mathbf{S})$ comprises the model of valuation. For example, the simplest scenario is a second-price or an English auction with independent private values; a bidder bids her "true" valuation for the auctioned object in these auctions. Here the auction format is a second-price or English auction. The utility function is $U_i(\mathbf{X}, \mathbf{S}) = X_i$ and the $X_i s$ are assumed to be independent.

The **models of valuation** which will be encountered in this paper at various points are now defined. The definitions are from Milgrom and Weber (1982), Laffont and Vuong (1996), Li, Perrigne and Vuong (2000) and Athey and Haile (2000).

Model 1: Affiliated-Values (AV)
This model is defined by the pair $[U_i(\mathbf{X}, \mathbf{S}), F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s})]$, with variables $(\mathbf{X}, \mathbf{S})$ being affiliated. A formal definition of affiliation is given in Milgrom and Weber (1982, p. 8). Roughly when the variables $(\mathbf{X}, \mathbf{S})$ are affiliated, large values for some of the variables will make other variables large rather than small. Independent variables are always affiliated; only strict affiliation rules out independence.

The private-values model and common value model are the two polar cases of the **AV.**

Model 2: Private-Values (PV)
Assuming risk neutrality, the private-values model emerges from the affiliated-values model if $U_i(\mathbf{X}, \mathbf{S}) = X_i$. Hence from equation (2)

$$x_i = e^{-1}(b_i) = \xi(b_i, F_{\mathbf{B}}(\mathbf{b}); n),$$

is also the inverse of the equilibrium bidding rule with respect to x_i. Different variants of this model emerge depending on the assumption made about the relationship between $(\mathbf{X}, \mathbf{S})$.

The **affiliated private values** $(\mathbf{AV})$ model emerges if $(\mathbf{X}, \mathbf{S})$ are affiliated. The structural element of this model is $F_{\mathbf{XS}}$.

It is possible that the private values have a component common to all bidders and an idiosyncratic component. Li, Perrigne and Vuong (2000) justify that this is the case of the OCS wildcat auctions. The common component would be the unknown value of the tract. The idiosyncratic component could be due to a differences in operational cost as well as the interpretation of geological surveys between bidders. Conditional on this common component, the idiosyncratic cost components could be independent. This was also the case in the procurement auctions

[2] Either some or bids or bids of all n potential bidders may be observed. If participation is endogenous, then the number of potential bidders, n, is not observed. The private signals of the bidders are unobserved as well.

[3] If bidders are symmetric $e_i(\bullet) = e(\bullet)$ and $\xi_i(\bullet) = \xi(\bullet)$ for all n bidders.

for crude-oil studied by Sareen (1999); the cost of procuring crude-oil by bidders was independent conditional on past crude-oil prices. This gives rise to the **conditionally independent private values** (**CIPV** henceforth) model. In a **CIPV** model,
(1) $U_i(\mathbf{X},S) = X_i$ with $m = 1$. S is the auction-specific or common component that affects the utility of all bidders;
(2) X_i are independent conditional on S.

The structural elements of the **CIPV** are $\left(F_S(s), F_{X|S}(x_i)\right)$. In the **CIPV** model an explicit relationship could be specified for the common and idiosyncratic components of the private signals X_i. Thus
(1) $U_i(\mathbf{X},S) = X_i = S + A_i$ where A_i is the component idiosyncratic to bidder i.
(2) A_i are independent conditional on S.

This is the **CIPV** model with additive components (**CIPV-A**); it's structural elements are $\left(F_S(s), F_{A|S}(a_i)\right)$.

If the assumption of conditional independence of A_i in the **CIPV-A** model is replaced with mutual independence of $(A_1, ..., A_n, S)$, the **CIPV** with independent components (**CIPV-I**) emerges. The independent private value (**IPV**) model is a special case of the **CIPV-I** with $S = 0$, so that $U_i = X_i = A_i$.

Model 3: Common Value (CV)
The pure **CV** model emerges from the **AV** model if $U_i(\mathbf{X}, \mathbf{S}) = S$. S is the "true" value of the object which is unknown but the *same* for all bidders; bidders receive private signals X_i about the unknown value of this object. If the OCS auctions are interpreted as common-value auctions as in Hendricks and Porter (1988) and Hendricks, Porter and Pinkse (1999), then an example of s is the ex post value of the tract. The structural elements of the pure **CV** model are $\left(F_S(s), F_{X|S}(x_i)\right)$. Notice its similarity to the **CIPV** model with $U_i(\mathbf{X}, \mathbf{S}) = X_i$, and no further structure on X_i.

A limitation of the pure **CV** model is that it does not allow for variation in utility across the n bidders. This would be the case if a bidder's utility for the auctioned object depended on the private signals of other bidders; thus $U_i = U_i(X_{-i})$. Alternatively, bidders tastes could vary due to factors idiosyncratic to a bidder, $U_i = U_i(A_i, S)$. Athey and Haile (2000) consider an additively separable variant of the latter. Specifically, $U_i = S + A_i$ with U_i, X_j *strictly* affiliated but not *perfectly* correlated. Excluding perfect correlation between U_i, X_j rules out *exclusively* private-values, where $U_i = X_i = A_i$, and guarantees that the winners curse exists. This will be referred to as the **CV** model in this survey.

In the **pure CV** model $U_i(\mathbf{X}, \mathbf{S}) = S$. S is the "true" value of the object which is unknown but the *same* for all bidders; bidders receive private signals X_i about the unknown value of this object. If the OCS auctions are interpreted as common-value auctions as in Hendricks and Porter (1988) and Hendricks, Porter and Pinkse (1999), then an example of s is the ex post value of the tract. The structural elements of the pure **CV** model are $\left(F_S(s), F_{X|S}(x_i)\right)$. Notice its similarity to the **CIPV** model with $U_i(\mathbf{X}, \mathbf{S}) = X_i$, and no further structure on X_i.

The **Linear Mineral Rights (LMR)** model puts more structure on the pure **CV** model. Using the terminology of Li, Perrigne and Vuong (2000), the **LMR** model assumes (1) $U_i(\mathbf{X},\mathbf{S}) = S$; (2) $X_i = SA_i$, with A_i independent conditional on S; and (3) $V(x_i, x_i)$ is loglinear in $\log x_i$. If in addition $(S, \mathbf{A})$ are mutually independent the **LMR with independent components** (**LMR-I**) emerges. The **LMR** model is analogous to the **CIPV-A** and the **LMR-I** to the **CIPV-I**.

All the above mentioned models of valuation can be either symmetric or asymmetric. A model of valuation is **symmetric** if the distribution $F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s})$ is exchangeable in its first n arguments; otherwise the model is **asymmetric**. Thus a symmetric model says that bidders are *ex ante* symmetric. That is, the index i which indicates a bidder is exchangeable: it does not matter whether a bidder gets the label $i = 1$ or $i = n$. Note that symmetry of $F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s})$ in it first n arguments $\mathbf{X}$ implies that $F_{\mathbf{X}}(\mathbf{x})$ is exchangeable as well.

The **auction formats** discussed in this survey are the first-price auction, Dutch auction, second-price and the ascending auction of Milgrom and Weber (1982). In both the first-price and second-price auction the winner is the bidder who submits the highest bid; the transaction price is the highest bid in a first-price auction but the second-highest bid in the second-price auction. In a Dutch auction an auctioneer calls out an initial high price and continuously lowers it till a bidder indicates that she will buy the object for that price and stops the auction. In the ascending auction or the "button" auction the price is raised continuously by the auctioneer. Bidders indicate to the auctioneer and the other bidders when they are exiting the auction;[4] when a single bidder is left, the auction comes to an end. Thus the price level and the number of active bidders is observed continuously.

Once an auction format is combined with a model of valuation, an equilibrium bidding rule emerges as described in equation (3). The bidding rules for the **AV** model with some auction formats are given below.

Example 1: AV model and Second-Price Auction

$$e\left(x_i\right) = v\left(x_i, x_i; n\right) = E\left(U_i | X_i = x_i, X_{-i}^{n-1:n-1} = x_i\right) \forall i. \tag{4}$$

$X_{-i}^{n-1:n-1} = \max_{j \neq i, j=1,\ldots,n} X_j$, is the maximum of the private signals over $n-1$ bidders excluding bidder i; and

$$v(x, y; n) = E\left(U_i | X_i = x_i, X_{-i}^{n-1:n-1} = y_i\right). \tag{5}$$

The subscript i does not appear in the bidding rule $e\left(x_i\right)$ since bidders are symmetric. $e\left(x_i\right)$ is increasing in x_i.[5] If $(\mathbf{X}, S)$ are *strictly* affiliated, as in the **CV** model, $e\left(x_i\right)$ is *strictly* increasing in x_i. For a **PV** model, $U_i(\mathbf{x}, \mathbf{s}) = X_i$. Hence $v\left(x_i, x_i\right) = x_i$. The bidding rule simplifies to

$$e\left(x_i\right) = x_i. \tag{6}$$

Example 2: AV model and First-Price, Sealed-Bid/Dutch Auction

[4] For example, bidders could press a button or lower a card to indicate exit.

[5] A reasoning for this follows from Milgrom and Weber. By definition U_i is a nondecreasing function of its arguments $(\mathbf{X}, \mathbf{S})$. Since $(\mathbf{X}, \mathbf{S})$ are affiliated, $(X_i, X_{-i}^{n-1:n-1})$ are affiliated as well from Theorem 4 of Milgrom and Weber. Then from Theorem 5, $E\left(U_i(\mathbf{X}, \mathbf{S}) | X_i, X_{-i}^{n-1:n-1}\right)$ is strictly increasing in all its arguments.

$$b_i = e\left(x_i\right) = v\left(x_i, x_i\right) - \int_{\underline{x}}^{x_i} L\left(\alpha|x_i\right) dv\left(\alpha, \alpha\right), \tag{7}$$

where $L\left(\alpha|x_i\right) = \exp\left(-\int_\alpha^{x_i} \frac{f_{x_{-i}}^{n-1:n-1}(t|t)}{F_{x_{-i}}^{n-1:n-1}(t|t)} dt\right)$. Again from affiliation of $(\mathbf{X}, \mathbf{S})$ and that U_i is non decreasing in its arguments, it follows that $e\left(x_i\right)$ is strictly increasing in x_i.[6]

Again for the **PV** model the bidding rule simplifies to

$$e\left(x_i\right) = x_i - \int_{\underline{x}}^{x_i} L\left(\alpha|x_i\right) d\alpha. \tag{8}$$

3 Identification, Estimation and Testing: An Overview

Identification and estimation of a structural model involves recovering the elements $[\mathbf{U}(\mathbf{X}, \mathbf{S}), F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s})]$ from data on the observables. Identification, estimation and testing of structural auction models is complicated for several reasons. First, an explicit solution for the Bayesian-Nash equilibrium strategy $e_i(\bullet)$ is available for few models of valuation. In most instances all an empirical researcher has are a set of differential equations which are a first-order conditions for the optimization exercise given by (1). Second, even if an explicit solution for $e_i(\bullet)$ is available, complication is introduced by the fact that the equilibrium bid is not a function of $(\mathbf{x}, \mathbf{s})$ *exclusively*. $(\mathbf{x}, \mathbf{s})$ affects the bidding rule through the distribution function $F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s}|\bullet)$ as well. Section 3.1-3.3 provide a brief description of the approaches to identification and estimation; these approaches have been discussed at length in the surveys by Hendricks and Paarsch (1995) and Perrigne and Vuong (1999). Testing is discussed in Section 3.4.

3.1 Identification and Estimation

There have been two approaches to identification and estimation of structural auction models when data on all bids is available and the number of potential bidders, n, is known. The **direct approach** starts by making some distributional assumption, $f_{\mathbf{XS}}(\mathbf{x}, \mathbf{s}|\theta)$, about the signals $(\mathbf{X}, \mathbf{S})$ of the auctioned object and the utility function U. Assuming risk-neutrality, the structural element is now θ. From here there are three alternatives. First, the likelihood function of the observed data can be used. Alternatively, the posterior distribution of the parameters that characterize the distribution of the signals $(\mathbf{X}, \mathbf{S})$ could be used; this is the product of the

[6] Since $(\mathbf{X}, \mathbf{S})$ are affiliated, $(X_i, X_{-i}^{n-1:n-1})$ are affiliated as well. Then it follows that $\frac{f_{x_{-i}}^{n-1:n-1}(t|t)}{F_{x_{-i}}^{n-1:n-1}(t|t)}$ is increasing in t. Hence $L(\alpha|x_i)$ is decreasing in x_i; which implies that $b(x_i)$ is increasing in x_i since the first term in the equilibrium bid function is increasing in x_i from Example 1a.

likelihood function and the prior distribution of the parameters. Both these alternatives are essentially *likelihood based* methods. These methods are relatively easy as long as an explicit one-to-one transformation relating the unobserved signals with the observed bids is available. The former approach has been used in Paarsch (1992, 1997) and Donald and Paarsch (1996) for identification and estimation assuming an **IPV** model of valuation. The latter approach has been used in several papers by Bajari (1999), Bajari and Hortacsu (2000) and Sareen (1999) and Van den Berg and Van der Klaauw (2000).

When an explicit solution is not available for the optimization problem in equation (1), it may be preferable to work with simulated features of the distribution specified for the signals $(\mathbf{X}, \mathbf{S})$; centred moments and quantiles are some examples. Simulated features are used since the analytical moments or quantiles of the likelihood function are likely to be intractable. These are called *simulation-based* methods.

For certain structural models the simulation of the moments could be simplified through the predictions made by the underlying game-theoretic models; Laffont, Ossard and Vuong (1995) and Donald, Paarsch and Robert (1997) provide two excellent examples.

Example 3: Laffont, Ossard and Vuong (1995)
In this paper, the revenue equivalence theorem is used to simulate the expected winning bid for a Dutch Auction of eggplants with independent-private-values and risk neutral bidders. The revenue equivalence theorem is a prediction of the independent-private-values model with risk neutral bidders. It says that the expected revenue of the seller is identical in Dutch, first-price, second-price and English auctions. The winning bid in a second-price and English auction is the second-highest value. If N and $F_X(x)$ were known, this second-highest value could be obtained by making n draws from $F_X(x)$ and retaining the second-highest draw. This process could be repeated to generate a sample of second-highest private values and the average of these would be the expected winning bid. Since N and $F_X(x)$ are not known, an importance function is used instead of $F_X(x)$; the above process is repeated for each value of N in a fixed interval.

Example 4: Donald, Paarsch and Robert (1997)
The idea in the previous example is extended to a sequential, ascending-price auction for Siberian timber-export permits assuming independent-private-values. The key feature of this data set is that bidders demand multiple-units in an auction. This complicates single-unit analysis of auctions since bidders may decide to bid strategically across units in an auction. Indeed the authors demonstrate that winning prices form a sub-martingale. An efficient equilibrium is isolated from a class of equilibria of these auctions. The direct incentive compatible mechanism which will generate this equilibrium is found. This mechanism is used to simulate the expected winning bid in their simulated nonlinear least squares objective function. The incentive compatible mechanism comprises each bidder revealing her true valuation. The T lots in an auction are allocated to bidders with the T highest valuation with the price of the $t-th$ lot being the $(t+1)th$ valuation. Having specified a functional form for the distribution of private values, they recreate T highest values for each of the N bidders from some importance function. The T highest valuations

of these NT draws then identifies the winners in an auction and hence the price each winner pays for the lot. This process is repeated several times to get a sample of winning prices for each of the T lots. The average of the sample is an estimate of the expected price in the simulated nonlinear least squares objective function.

The key issue in using any simulation based estimator is the choice of the importance function. It is important that the tails of the importance function be fatter than the tails of the target density to ensure that tails of the target density get sampled. For example, Laffont, Ossard and Vuong (1995) ensure this by fixing the variance of the lognormal distribution from the data and using this as the variance of the importance function.

In structural models where the bidding rule is a **monotonic** transformation of the unobserved signals, Hong and Shum (2000) suggest the use of simulated quantiles instead of centred moments. Using quantiles has two advantages. First, quantile estimators are more robust to outliers in the data than estimators based on centred moments. Second, quantile estimators dramatically reduce the computational burden associated with simulating the moments of the equilibrium bid distribution when the bidding rule is a monotonic transformation of the unobserved signals. This follows from the quantiles of a distribution being invariant to monotonic transformations so that the quantiles of the unobserved signals and the equilibrium bidding rule are identical. This is not the case for centred moments in most instances.

Instead of recovering θ, the **indirect approach** recovers the distribution of the unobserved signals, $f_{\mathbf{XS}}(x,s)$. It is indirect since it does not work directly off the likelihood function of the observables. Hence it avoids the problem of inverting the equilibrium bidding rule to evaluate the likelihood function. The key element in establishing identification is equation (2). An example illustrates this point.

Example 5: Symmetric **PV** model
From Guerre, Perrigne and Vuong (2000) and Li, Perringne and Vuong (2000) the first-order condition in equation (2) is

$$x_i = \xi_i \left(b_i, f_{B_i|b_i}(b_i|b_i), F_{B_i|b_i}(b_i|b_i), n\right) = b_i + \frac{F_{B_i|b_i}(b_i|b_i)}{f_{B_i|b_i}(b_i|b_i)}. \tag{9}$$

$\xi_i(\bullet)$ is the inverse of the bidding rule given in equation (3) with respect to the private signal x_i in a **PV** model. Identification of the APV model is based on this equation since $F_{\mathbf{X}}(\mathbf{x})$, the distribution of private signals is completely determined by n, $f_{B_i|b_i}(b_i|b_i)$ and $F_{B_i|b_i}(b_i|b_i)$. This idea underlies their estimation procedure as well. First, $\frac{f_{B_i|b_i}(\bullet)}{F_{B_i|b_i}(\bullet)}$ is nonparametrically estimated through kernel density estimation for each n. Then a sample of nT pseudo private-values are recovered using the above relation. Finally this sample of pseudo private-values is used to obtain a kernel estimate of the distribution $F_{\mathbf{X}}(\mathbf{x})$. There are several technical and practical issue with this class of estimators. First, the kernel density estimator of bids, $\widehat{f_{\mathbf{B}}}(\mathbf{b})$, is biased at the boundary of the support. This implies that obtaining pseudo private values from relation (9) for observed bids that are close to the boundary will be problematic. The pseudo private-values are as a result defined by the relation in equation (9) only for that part of the support of $f_{\mathbf{B}}(\mathbf{b})$ where $\widehat{f_{\mathbf{B}}}(\mathbf{b})$ is its unbiased estimator. The second problem concerns the rate of convergence of the estimator $\widehat{f_{\mathbf{X}}}(\mathbf{x})$. Since the density that is being estimated nonparametrically is not the density of observables, standard results on the convergence of nonparametric estimators of density do not apply. Guerre, Perrigne and Vuong (2000) prove that the best rate of convergence that these estimators can achieve is $(T/\log T)^{r/(2r+3)}$, where r is the number of bounded continuous derivatives of $f_{\mathbf{X}}(\mathbf{x})$.

The **indirect approach** has concentrated on identification and estimation from bid data. Laffont and Vuong (1996) show that several models of valuations described in Section 2 are not identified from bid data. Traditionally identification of non-identified models has been obtained either through additional data or dogmatic assumptions about the model or both. Eventually, the data that is available will be a guide as to which course an empirical researcher has to take to achieve identification. For a particular auction format, identification and estimation of more general models of valuation will require more detailed data sets. These issues are discussed in Sections 3.2 and 3.3 for symmetric models; I turn to asymmetric models in Section 3.4. Nonparametric identification implies parametric identification but the reverse is not true; hence the focus is on nonparametric identification.

3.2 Identification from Bids

When bids of all participants are observed, as in this Subsection, identification of a **PV** model, second-price auction is trivial. This follows from the bidders submitting bids identical to their valuation of the auctioned object. formats of interest are the first-price and the second-price auctions. For a model of valuation other than the

PV model, the identification strategy from data on bids for the second-price auction will be similar to that of a first-price auction. The structure of the bidding rules in Examples 1 and 2 makes this obvious. The second-price auction can be viewed as a special case of the first-price auction. The second term in the bidding rule of a first-price auction is the padding a bidder does over the expected value of the auctioned object. This is zero in a second-price auction since a bidder knows that he will have to pay only the second-highest bid. For example, for the **PV** model, while the bids are strictly increasing transformations of the private signals under both auction formats, this increasing transformation is the identity function for a second-price auction. A model of valuation identified for a first-price auction will be identified for a second-price auction. In Identifying auction formats other than

the first-price and second-price auction from data on all bids is not feasible since for both the the Dutch and the "button" auctions all bids can never be observed. In a Dutch auction only the transaction price is observed. The drop-out points of all except the highest valuation bidder are observed in the "button" auction.

Hence the discussion in this section is confined to first-price auctions.

Within the class of symmetric, risk-neutral, **private-values** models with a non-binding reserve price, identification based on the **indirect approach** has been established for models as general as the **APV** model (Perrigne and Vuong, 1999). Since risk neutrality is assumed the only unknown structural element is the latent distribution, $F_{\mathbf{X}}(\mathbf{x})$.

The **CIPV** model is a special case of the **APV** model. The interesting identification question in the **CIPV** model is whether the structural elements $\left[F_S(s), F_{\mathbf{X}|S}(\mathbf{x})\right]$ can be determined uniquely from data on bids. From Li, Perrigne and Vuong (2000), the answer is no; an observationally equivalent **CV** model can always be found. For example replace the conditioning variable S by an increasing transformation λS in the **CIPV** model. The utility functions is unchanged since $U_i(\mathbf{X}, \mathbf{S}) = X_i$. The distribution of bids generated by this model would be identical to the **CV** model where $U_i(\mathbf{X}, \mathbf{S}) = S$ and the X_i are scaled by $(\lambda)^{1/n}$. Thus $F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s})$ is identified but not its individual components. Data could be available on S; this is the case in the OCS auctions where the *ex post* value of the tract is available. Even though $F_S(s)$ can be recovered, this would still not guarantee identification since the relationship between the unknown private signals of the bidders, X_i, and the common component which is affecting all private values S is not observed by the econometrician. Without further structure on the private signals the **CIPV** model is not identified *irrespective* of the auction form and the data available.

Both Li, Perrigne and Vuong (2000) and Athey and Haile (2000) assume $X_i = S + A_i$. Note that $U_i = X_i$ from the **PV** assumption. With this additional structure the **CIPV-A** emerges when $(A_1, ..., A_n)$ are independent conditional on S. This again is not enough to guarantee identification of the structural elements $\left(F_S(s), F_{A|S}(a)\right)$ from data on all bids in an auctions since bids give information about the private valuations but not the individual components of these valuations; hence the observational equivalence result of the **CIPV** model is valid here as well. At this point identification of the **CIPV-A** model can be accomplished by either putting further structure on it or through additional data. These are discussed in turn.

In addition to (1) $U_i = X_i = S + A_i$, assume (2) $(A_1, ..., A_n, S)$ are mutually independent with A_is identically distributed with mean equal to one; and (3) the characteristic functions of $\log S$ and $\log A_i$ are nonvanishing everywhere. This is the **CIPV-I** model defined above. Li, Perrigne and Vuong (2000) and Athey and Haile (2000) draw upon the results in the measurement error literature as in Kotlaraski (1966) and Li and Vuong (1998) to prove that (F_s, F_A) are identified.[7] The importance of the assumptions about the decomposability of the X_i and the mutual independence is that $\log X_i$ can be written as

$$\log X_i = \log c + \log \epsilon_i, \tag{10}$$

where $\log c \equiv [\log s + E(\log A_i)]$ and $\log \epsilon_i \equiv [\log A_i - E(\log \eta_i)]$. Hence the $\log X_i s$ are indicators for the unknown $\log c$ which are observed with a measurement error $\log \epsilon_i$. $\log X_i$ or X_i can be recovered from data on bids since the bids are strictly increasing functions of X_i. Identification is then straightforward from results concerning error-in-variable models with multiple indicators, as in Li and Vuong (1998, Lemma 2.1).

S is the common component through which the private signals of the bidders are affiliated. There are several examples where data on S could be available. In the crude-oil auctions studied by Sareen (1999) the past prices of crude-oil would

[7] Li, Perrigne and Vuong (2000) consider a multiplicative decomposition of X_i with $X_i = SA_i$. This will be a linear decomposition if the logarithm of the variables is taken.

be a candidate for S. In the timber auctions studied by Sareen (2000d) and Athey and Levin (1999), the species of timber, tract location, etc would be examples of S. In the OCS auctions S could be the ex post value of the tract. If data was available on S then the **CIPV-A** model with $U_i = X_i = S + A_i$ would be identified; refining **CIPV-A** to **CIPV-I** is not needed. For each s, $A_i = X_i - s$; hence identification of $F_{A|s}(a)$ follows from the identification of the **IPV** model. $F_S(s)$ is identified from data on S.[8]

The same argument goes through if instead of S some auction-specific covariates are observed conditional on which the A_i are independent. If $S = 0$, but auction specific covariates Z are observed, then again $F_{A|z}(a)$ is identified for each z. It could also be the case that $U_i = g(Z) + A_i$; then $g(Z)$ and $F_{A|z}(a)$ for each z are identified. In fact when either S or auction-specific characteristics are observed, all bids are not needed in either the first-price or the second-price auctions to identify the **CIPV-A** model. Any bid, for example the transaction price, will do in a second-price auction (Athey and Haile, 2000, Proposition 6).

The picture is not as promising when one comes to **Common Value** auctions when the data comprises exclusively of bids in a first-price auction. The **CV** model is not identified from data on bids; the bids can be rationalized by an **APV** model as well. This follows from Proposition 1 of Laffont and Vuong (1996).

The **symmetric pure CV** model is not identified from all n bids; this is not surprising in view of the observational equivalence of the **pure CV** and the **CIPV** model discussed above. The identification strategy is similar to the one followed for the **CIPV** model: more structure on the model of valuation or additional data.

The **LMR** and the **LMR-I** models puts more structure on the **pure CV** model. The **LMR** model is analogous to the **CIPV-A** and the **LMR-I** to the **CIPV-I;** hence identification strategies for the two sets of models is similar. The **LMR** model is not identified from bids. Assuming mutual independence of $(S, \mathbf{A})$, the **LMR-I** emerges; Li, Perrigne and Vuong (2000) prove it's identification for the first-price auction and Athey and Haile (2000, Proposition 4) for the second-price auction. Again if data on the realized value of the object, S, was available, the **LMR** model is identified from just the transaction price (Athey and Haile, 2000, Proposition 15).

3.2.1 Binding reserve price

Once the assumption of a nonbinding reserve price is relaxed there is an additional unknown structural element besides the distribution of the latent private signals. Due to the binding reserve price p_o the number of players who submit bids, p_j, is less than the number of potential bidders which is now unknown. Bidders with a valuation for the auctioned object greater than p_o will submit bids in the auction; thus a truncated sample of bids will be observed.

The question now is whether the joint distribution of $(\mathbf{N}, \mathbf{V})$ is identified from the truncated sample of p_j bids in an auction. Guerre, Perrigne and Vuong (2000) in the context of an **IPV** model examine a variant of this issue. Using the **indirect**

[8]Since more than one realization of S is observed, $F_{\mathbf{A}}$ is identified too. If the auctions are assumed to be independent and *identical, then the variation in* S across auctions can be used to identify the conditional distribution $F_{\mathbf{A}|S}$.

approach they establish the identification of the *number of potential bidders* and the nonparametric identification of the distribution of valuations. A necessary and sufficient condition, in addition to those obtained for the case of the nonbinding reserve price emerges; the distribution of the number of potential bidders is binomial with parameters $(n, [1 - F_v(p_o)])$. In essence though there is an extra parameter here, there is also an additional observable present: the number of potential bidders. This extra observable pins down the parameter n.

3.2.2 Risk averse models of valuation

Another noteworthy extension here is the relaxation of the assumption of risk neutrality. The structural model is now characterized by $[U, F_{\mathbf{X}}(\mathbf{x})]$. Campo, Perrigne and Vuong (2000) show that for nonbinding reserve prices, an **IPV** model with constant relative risk aversion cannot be distinguished from one with constant absolute risk aversion from data on observed bids. Since the risk-neutral model is a special case of the risk averse models, this implies that a risk neutral model is observationally equivalent to a risk averse model. The redeeming aspect of this exercise is that the converse is not true. This follows from the risk neutral model imposing a strict monotonicity restriction on the inverse bidding rule $\xi(\bullet)$, a restriction that the risk averse model does **not** impose.

Identification of $[U, F_{\mathbf{X}}(\mathbf{x})]$ is achieved through two assumptions. First, the utility function is parametrized with some k dimensional parameter vector η; this on it's own does not deliver identification. Then heterogeneity across the auctioned objects is exploited by identifying η through k distinct data points. Second, the upper bound of $F_{\mathbf{X}}(\mathbf{x})$ is assumed to be known but does not depend on the observed covariates. This seems logical since object heterogeneity can be exploited for identification only one - either to identify η, the parameters of the utility function or some parts of the distribution $F_{\mathbf{X}}(\mathbf{x})$, but not both.

3.3 Identification with less than all bids

In Section 3.2 data was available on the bids of all n agents. As a result identification of a model of valuation for a first-price and second-price auction was similar. Outside his scenario identification of a model of valuation for a first-price auction does not automatically lead to the identification of the same model for a second-price auction. The reason is that unlike the first-price auction the transaction price is not the highest bid but the second-highest bid. I start with identification when only the transaction price is recorded in an auction. For models that cannot be identified, the additional data or parametrization needed to identify the models are described. The discussion in Sections 3.3.1 and 3.3.2 is conditional on n, the number of potential bidders; identification of n raises additional issues which are discussed in Section 3.3.3. Thus n is common knowledge, fixed across auctions and observed by the empirical researcher.

3.3.1 First-Price, Sealed-Bid Auction

The most general **PV** model that has been identified from winning bids across auctions is the **symmetric IPV** model. Guerre, Perrigne and Vuong (1995) discuss identification and estimation of F_X from data on winning bids. The discussion is analogous to identification from bids in the last subsection with b replaced with w, $f_B(b)$ with $f_W(w)$ and $F_B(b)$ with $F_W(w)$.[9]

It is independence of private values and the symmetry of bidders that makes identification of the distribution of private values feasible from only the winning bid. Beyond this scenario, all bids will be needed if the independence assumption is relaxed but symmetry is retained (Athey and Haile, 2000, Corollary 3). If the latter is relaxed as well, identity of the bidders will needed for any further progress; asymmetry is the subject of Section 3.4. Also note that observing s or auction-specific covariates, in addition to the winning bid, does not help; the nature of the additional data should be such that it gives information about either the affiliation structure and/or the asymmetry of the bidders. For example, Athey and Haile give several instances when the top two bids are recorded.[10] Assuming symmetry, the top two bids should give information about the affiliation structure compared to observing just the winning bid. Under the additional assumption of the **CIPV-A** model, Athey and Haile (2000, Corollary 2) show that observing the top two bids and S identifies only the joint distribution of the idiosyncratic components, $F_{\mathbf{A}}(\mathbf{a})$.

3.3.2 Second-Price Auctions

Unlike the first-price auction, the winning bid and the transaction price are different in a second-price and an ascending auction. The transaction price is the second-highest bid. Hence to establish identification, the properties of second-highest order statistic have to be invoked.

The symmetric **APV** model is not identified from transaction prices; this is not surprising in view of the comments in Section 3.3.1. A formal proof is provided in Proposition 8 of Athey and Haile (2000); linking it to the discussion for first-price auctions they prove that all n bids are sufficient to identify the symmetric **APV** model.

Independence of the private signals buys identification. The **symmetric IPV** model is a special case of the **APV** with strict affiliation between the X_i replaced by independence. Athey and Haile (2000) prove that the **symmetric IPV** model is identified from the transaction price. This follows from the distribution function of any i^{th} order statistic from a distribution function F_X being an increasing function of F_X,

$$F_X^{2:n}(z) = \frac{n!}{(n-2)!} \int_0^{F_X(z)} t(1-t)^{n-2} dt. \tag{11}$$

The **CIPV-A** puts parametric structure on the **APV** model in this sense of decomposing the private values into a common and an idiosyncratic component, $X_i = S + A_i$. If s was observed, the identification of the **CIPV-A** would follow from the identification of the **IPV** model with transaction prices exclusively since the A_i are independent conditional on s. Athey and Haile (2000, Propositions 6,7) prove

[9] Since the estimator of $F_w()$ converges at a slow rate of $(T/\log T)^{r/(2r+3)}$, the data requirements and hence estimation could be burdensome. As a result, parametric or direct methods may be necessary.

[10] The difference between the top two bids is the "money left on the table".

that the **CIPV-A** model is identified from the transaction prices and observing the the *ex post* realization of S or auction-specific covariates, Z, conditional on which the A_i are independent. Alternatively if $X_i = g(Z) + A_i$, then both $g(Z)$ and $F_{A|z}$ are identified up to a locational normalization from the transaction prices and the auction-specific covariate Z.

The prognosis for the **CV** model can be obtained from the **PV** model since the **CIPV-A** is similar to the **LMR** model and the **CIPV-I** to the **LMR-I** model. Athey and Haile (2000, Proposition 15) prove that if the *ex post* realization of the "true" value of the object was observed, then the **LMR** model (and consequently the **LMR-I** model) is identified from the transaction price.

3.3.3 Identifying the number of potential bidders

In the previous Sections the number of potential bidders is assumed to be given. With n given, the **IPV** model is identified from transaction prices for both the first-price and the second-price auctions. In many auctions the empirical researcher does not observe the number of potential bidders even though it is common knowledge for the bidders. Inference about N may be of interest since it affects the average revenue of the seller in **PV** auctions; larger the number of bidders, more competitive is the bidding. In **CV** auctions the number of potential bidders determines the magnitude of the winner's curse.

Issues in identification of n will differ depending on whether it is **endogenous** or **exogenous**. If the number of potential bidders is systematically correlated with the underlying heterogeneity of the auctioned object, then n is endogenous. For example, this was the case in Bajari and Hortacsu (2000) when they model the entry decision of bidders in eBay auctions. If these object characteristics are observed, identification of n involves conditioning on these characteristics. Unobserved object heterogeneity will cause problems for nonparametric identification.[11] This is similar in spirit to the identification of the **CIPV-A** in Athey and Haile (2000) that is discussed in Section 3.3.2; nonparametric identification of the **CIPV-A** model is possible in case the *ex post* realization of S or auctions-specific covariates are observed.

Exogeneity of n implies that the distribution of unobserved values in an auction is the same for all potential bidders conditional on the symmetry of the bidders. This accommodates both the case when the number of actual bidders differs from the number of potential bidders and the case of a random n. Donald and Paarsch (1996), Donald, Paarsch and Robert (1999) and Laffont, Ossard and Vuong (1995) take the former approach. In these papers n is fixed but the actual number of bidders varies across auctions due to either the existence of a reserve price or the costs of preparing and submitting bids. Hendricks, Pinkse and Porter (1999) take the latter approach; they make a serious effort to elicit the number of potential bidders on each tract in the OCS wildcat auctions.

Assuming **exogenity** of the number of potential bidders and in view of the **IPV** model being the most general model of valuation that is identified from the trans-

[11] Li (2000) endogenizes entry but makes parametric assumptions about the distribution of private values. Bajari and Hortacsu (2000) make parametric assumptions about both the distribution of unobserved values and the entry process.

action price, I now examine whether the highest bid identifies both the distribution of private-values, $F_X(x_i)$, and the number of potential bidders N. The proposition below formalizes these ideas when the highest bid is observed.

Proposition 1: For a single-unit first-price, Dutch or Second-price symmetric **APV** auction, the distribution of the winning bid can be expressed as

$$F_w(u) = Q\left([F_{\mathbf{X}_n}(u)]^{1/n}\right),$$

where $Q(s) = \sum_{n=0}^{\infty} s^n \Pr(N = n)$ and $F_{\mathbf{X}_n}(u)$ is the joint distribution of n private-values.

Proof: Let $W = Max\{B_1, ..., B_N\}$. Then

$$\begin{aligned}
F_w(u) &= \Pr(W \leq u), \\
&= \sum_{n=0}^{\infty} \Pr(W \leq u | N = n) \Pr(N = n), \\
&= \sum_{n=0}^{\infty} \Pr(B_1 \leq u, ..., B_{n-1} \leq u, W \leq u) \Pr(N = n), \\
&= \sum_{n=0}^{\infty} \Pr(X_1 \leq u, ..., X_{n-1} \leq u, W \leq u) \Pr(N = n), \\
&= \sum_{n=0}^{\infty} F_{\mathbf{X}_n}(u) \Pr(N = n).
\end{aligned}$$

$F_{\mathbf{X}_n}(u)$ is the joint distribution of $\mathbf{X}$ conditional on $N = n$. The second last step follows from the bids being strictly increasing monotonic transformations of the private signals of the bidders. With $s = [F_{\mathbf{v}_n}(u)]^{1/n}$ in the function $Q(s)$,

$$F_w(u) = Q\left([F_{\mathbf{v}_n}(u)]^{1/n}\right).$$

From Proposition 1 it is obvious that the distribution of private signals is not identified as in Athey and Haile (2000, Corollary 3) given n.[12]

Corollary 1: For a single-unit first-price, Dutch or Second-price symmetric **IPV** auction, either $F_X(x_i)$ or N but not both are identified from the highest bid.

Proof: Suppose n is fixed. For a symmetric **IPV** auction

$$F_{\mathbf{X}_n}(u) = [F_X(u)]^n.$$

Then from Proposition 1 $F_X(u)$ is uniquely determined,

$$F_X(u) = Q^{-1}(F_w(u)).$$

[12] Choose

$$F_{\mathbf{X}_n}(u) = \left[Q^{-1}(F_w(u))\right]^n;$$

then from Proposition 1,

$$\begin{aligned}
F_w(u) &= Q\left([F_{\mathbf{X}_n}(u)]^{1/n}\right) = Q\left(\left[\left[Q^{-1}(F_w(u))\right]^n\right]^{1/n}\right), \\
&= F_w(u).
\end{aligned}$$

Assume $F_X(u)$ is fixed. A counter example shows that n is determined uniquely from

$$F_w(u) = Q\left([F_X(u)]\right) = \sum_{n=0}^{\infty} \left(F_X(u)\right)^n \Pr\left(N = n\right).$$

Suppose $\Pr\left(N = n\right)$ puts point mass on n and n is not uniquely determined. Then there exits n^* with $n \neq n^*$ and

$$F_w(u; n) = F_w\left(u; n^*\right).$$

But that cannot be true since

$$F_w(u; n) = \left(F_X(u)\right)^n \neq \left(F_X(u)\right)^{n^*} = F_w\left(u; n^*\right).$$

In most cases if the structural model is not identified nonparametrically, parametric assumptions may identify the model; this is not the case for the result in Corollary 1. Even if parametric assumptions are made about $F_X\left(x_i|\theta\right)$, Corollary 1 holds.[13] Corollary 1 forms the basis for identification and estimation when data on just the winning bids is available. Laffont, Ossard and Vuong (1995) when studying a Dutch auction for eggplants use a simulated nonlinear least squares function to obtain $\widehat{\theta}$; an estimate of N is obtained conditional on $\widehat{\theta}$. Similarly Guerre, Perrigne and Vuong (1995) study nonparametric identification and estimation of F_v, the distribution of private values, from winning bids, for a first-price, sealed-bid auctions with independent private values, conditional on n. n is obtained by solving

$$\overline{v} = \xi\left(\overline{b}, F_W(w), n\right),$$

where $\overline{v}$ and $\overline{b}$ are the *known* upper bounds of the distribution $F_X(x)$ and $F_B(b)$.

3.4 Asymmetric Models of Valuation

Since symmetric models of valuation are special cases of asymmetric models, if a symmetric model is not identified, the corresponding asymmetric model will not be identified as well. Thus the asymmetric pure **CV** model or the asymmetric **CIPV-A** model are not identified from n bids. The asymmetric **APV** model will not be identified from transaction prices.

What about symmetric models of valuation that are identified when all n bids are observed? As long as a complete set of n bids is available, the asymmetric counterparts of the identified symmetric models should be identified. The **indirect method** is used to identify the asymmetric **APV** model by Campo, Perrigne and Vuong (1998) for a first-price auction and by Athey and Haile (2000) for a second-price auction.[14]

The interesting question with regard to asymmetric **PV** models is their identification from transaction prices. Asymmetric models cannot be identified from

[13] Guerre, Perrigne and Vuong (1995, pp. 27-28) give a numerical example of two specifications of the symmetric **IPV** model which generate the same distribution of winning bids.

[14] Proposition 8 of Athey and Haile shows that the **APV** model cannot be identified from an incomplete set of bids.

transaction prices only. An example clarifies this point. Suppose all auctions observed are won by a single bidder j. Then $F_{X_j}(x_j)$, bidder js value distribution will be identified; since bidders are not exchangeable now $F_{X_j}(x_j)$ will not identify $F_{\mathbf{X}}(\mathbf{x})$.

The example suggests that identification from transaction prices may be possible if the identity of the winner was observed as well. Borrowing the work on competing risk models (Berman, 1963 and Prakasa Rao, 1982) and the multi-sector Roy model (Heckman and Honoré, 1990), Athey and Haile (2000) use the additional data on the identity of the winner to establish identification of several **PV** models with a caveat.[15] They assume that the support of the distribution of valuation of each of the n potential bidders is identical. If this is not the case, then the identity of the bidders could be used to establish identification from transaction prices if each bidder won some auctions. It is very rare that a researcher gets to observe all potential bidders winning several auctions in the data set. The identification result of Athey and Haile (2000) are important since in many data sets it is possible to establish symmetry with respect to groups of bidders. This was the case in Sareen (2000) where one observed three nonfringe firms and several fringe firms some of whom participated only once; in essence there are two non-exchangeable groups of bidders in this auction.

3.5 Testing of Structural Auction Models

A Structural model comprises of many elements. For example, a structural model makes assumptions about bidder rationality, that bidders bid according to Bayesian-Nash equilibrium strategies, about the correlation between bidder's valuations, common value or private-values components, exchangeability etc. Testing of structural models has followed three broad directions.

The first direction focuses on whether a structural auction model imposes any testable restrictions on the observables.[16] For example, for an **IPV** first-price auction, the structural elements $[U_i(\mathbf{X}, \mathbf{S}), F_{\mathbf{XS}}(\mathbf{x}, \mathbf{s})] = [X_i, F_X(x_i)]$. Laffont and Vuong (1996) show that this structure imposes two testable restrictions on the distribution of bids: the bids in an auction are *iid* and that $\xi(b)$ is strictly increasing in b.[17] The former restriction ascertains the independence of bidders valuations irrespective of whether they are playing Bayesian-Nash strategies; the latter tests if they are playing Bayesian-Nash strategies irrespective of the model of valuation. Through these restrictions, individual tests can be designed for different aspects of a structural model.

[15]For a first-price and second-price auction, Athey and Haile (2000, Propositions 16, 2 respectively) establish the identification of the **IPV** model from the highest bid and the identity of the winner. For ascending and second-price auctions the **IPV** model is identified from the lowest bid and the identity of the loser as well. Additional data on the identity of the winner can also be used for identifying **APV** model for a second-price or ascending auction if further structure, described in Proposition 11, is imposed on the private signals; the similarity of the auction model to the mutli-sector Roy model is exploited for this. A similar structure on the private signals for first-price auctions does *not* identify the **APV** model (Corollary 2).

[16]According to Laffont and Vuong, a distribution of bids $F_{\mathbf{B}}$ is rationalized by a structural model $[U(\mathbf{X}, \mathbf{S}), F_{\mathbf{XS}}]$ if $F_{\mathbf{B}}$ is the equilibrium bid distribution of the corresponding game.

[17]These two restrictions comprise necessary and sufficient for the bids to be generated by a symmetric **IPV** first-price auction.

If all bidders were bidding in accordance with a proposed structural model, estimates of this model from different subsets of the data should be identical. If that is the case the data are consistent with the model; if not the data are inconsistent with the model. This idea is used by Athey and Haile (2000) for testing several structural models. For example, assuming risk-neutrality, they test a symmetric **IPV** model for a second-price auction by ascertaining if $F_X(x)$ recovered from the transaction prices is identical to the $F_X(x)$ recovered from data on the highest bid. In case it is not, the null hypothesis that the data is consistent with the **IPV** model is rejected. From the decision to reject a structural model it is unclear as to which component of the model led to its rejection; rejection could be due to violation of independence or symmetry or playing Bayesian-Nash strategies. This problem arises because the alternative model being tested is not explicit.

A third direction specifies explicit structural models under the null and the alternative hypothesis. The focus of this research has been on distinguishing the **CV** model from the **PV** model and a symmetric model from its from asymmetric counterpart. These are discussed in turn.

Laffont and Vuong (1996) have proved that for first-price auctions, conditional on the number of potential bidders, the **CV** model cannot be distinguished from the **PV** model on the basis of bids of the n potential bidders. Distinguishing the two models of valuation is important for several reasons one of which is mechanism design. Hendricks, Pinkse and Porter (1999), Haile, Hong and Shum (2000) and Athey and Haile (2000) examine what additional data would distinguish the two models in a first-price auction. As the discussion below will show, there is no unique way of distinguishing between the two models; eventually which test to use would be decided by the data that is available.

Haile, Hong and Shum (2000) and Athey and Haile (2000) work with bidding data like Laffont and Vuong (1996). However they do not condition on the number of potential bidders. Rather the comparative statics implication of varying the number of potential bidders on the equilibrium bidding rule under the two models is exploited for testing. The winner's curse is an adverse selection phenomenon arising in **CV** but not **PV** auctions. In a **CV** auction winning is bad news since the winner is the bidder who has been most optimistic about the unknown value of the auctioned object. A rational bidder will therefore account for the winner's curse by lowering his expected value of the auctioned object and hence his bid. The more competition a bidder expects, the more severe is the curse and hence larger is the adjustment in expected value of the auctioned object to mitigate the curse. The winner's curse does not arise in the **PV** setting since the value of the auctioned object to a bidder does not depend on the information that his opponents have. On this basis Haile, Hong and Shum (2000) suggest the following test:

$$\begin{aligned} PV &: \quad F_{v,2} = F_{v,3} = ... = F_{v,N}; \\ CV &: \quad F_{v,2} < F_{v,3} < ... < F_{v,N}, \end{aligned} \tag{12}$$

in the sense of first-order stochastic dominance. $F_{v,n}$ is the distribution of $v(x, x; n)$ induced by the distribution specified for the private signals $\mathbf{X}$; $v(x, x; n)$ is given in equation (4). Instead of testing the stochastic dominance hypothesis, Haile, Hong and Shum (2000) compare various features like quantiles of the distributions as the number of potential bidders varies. The *key* data requirements for their test is to observe auctions in which the number of potential bidders vary from 1 to N; in an auction with n bidders, all n bids should be observed.

Athey and Haile (2000, Corollary 3) exploit recurrence relations between order statistics and reduce the data requirements of Haile, Hong and Shum (2000); all they need is the top two bids in an auction with $n \geq 3$ bidders and the top bid in an auction with $n-1$ bidders.[18]

A key feature of the OCS data set is the observation of the *ex post* value of the tract. Therefore Hendricks, Pinkse and Porter (1999) can compute the average rent over tracts and average bid markdowns over tracts and bidders under the **CV** and the **PV** models.[19] Since entry is determined by a zero expected profit condition, the average rents under the **PV** and the **CV** models are compared with the total entry cost on a tract.[20] They find that the average rents under the **CV** model are comparable with the entry costs. They find no difference in the average bid markdowns under the **CV** hypothesis between tracts with a small and large number of potential bidders. In both cases they are comparable to the average rents. Thus bidders seems to have anticipated the winner's curse.

An alternative test suggested by Hendricks, Pinkse and Porter (1999) applies to first-price and second-price auctions with nonbinding reserve prices. This involves observing the behavior of $\xi(x, F_B)$ near the reserve price. In a **PV** model since $v(x,x) = x$, $\xi(x, F_B)$ will satisfy the boundary condition, $\lim_{b \downarrow r} \xi(x, F_B) = r$ from equation (2). Thus

$$\lim_{b \downarrow r} \frac{\Pr\left(\max_{j \neq n} B_j \leq b | B_i = b\right)}{\frac{\partial}{\partial \tau} \Pr\left(\max_{j \neq n} B_j \leq b | B_i = b\right)|_{\tau = b}} = 0.$$

For the **CV** model since $v(x,x) \neq x$, $\lim_{b \downarrow r} \xi(x, F_B) = v(x,x) > r$. Hence the limit above approaches a nonzero constant.

Finally, can symmetric models be distinguished from asymmetric models. On the basis of bidding data alone progress is limited. Laffont and Vuong (1996) give an example to show that asymmetric pure **CV** models may not be distinguished from their symmetric counterparts for first-price auctions. This result is not surprising

[18] If in addition, the identity of the bidders corresponding to the observed bids is available the asymmetric version of this test is possible.

[19] If the sample size is T, the average rents under the PV model are

$$R = T^{-1} \sum_{t=1}^{T} \left[\widehat{x}_t^{n:n} - w_t\right],$$

where $\widehat{x}_t^{n:n} = \widehat{\xi}(w_t)$ in the **PV** model and $\widehat{x}_t^{n:n} = s_t$ in the **CV** model where s_t is the observed ex post value of tract t. The bid markdowns are given by

$$M = T^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n_t} \left[\widehat{x}_{it}^{n:n} - b_{it}\right],$$

where $\widehat{x}_{it}^{n:n} = \widehat{\xi}(b_{it})$ in the **PV** model and $\widehat{x}_t^{n:n} = E(S|w = b_{it})$ in the **CV** model.

[20] Entry costs are the price of conducting the seismic survey pr acre hiring engineers to study the survey data and prepare a bid.

since their asymmetric model comprises one informed bidder and $n-1$ uninformed bidders. The uninformed players will play a mixed strategy whereby their maximum bid mimics the bid distribution of the informed bidder. Thus if $n = 2$, the observational equivalence noted by Laffont and Vuong becomes obvious.

In Section 3.4 I have explained how Athey and Haile have used bidder identity to identify asymmetric models. It should be possible to exploit bidder identity to test for symmetry *vs* asymmetry. For example, let I_i indicates the identity of the bidder who submits bid b_i. Then the conditional distribution $F_X(x_i|I_i)$ should be invariant with respect to I_i for a symmetric model but not an asymmetric model.

4 Bayesian Inference of Structural Auction Models

Bayesian methods are **direct methods** in that they require specification of a likelihood function. Parametric assumption about the unobservable $(\mathbf{X}, \mathbf{S})$ and the utility function U and a one-to-one mapping from the unobservable to the observables ensures the specification of the likelihood. The structural element of the auction model is now θ, where θ characterizes the distribution of $(\mathbf{X}, \mathbf{S})$. In *addition* to the likelihood function, Bayesian tools require specification of prior distribution for the parameters θ. Estimation proceeds through the posterior distribution of the parameters which is a product of the likelihood function and the prior distribution. The posterior odds ratio or the posterior predictive distributions may be used for testing. These concepts will be explained below.

There are several reasons why Bayesian tools may be attractive for inference in structural auction models.

Despite progress on nonparametric identification of structural auction models, the results pertain to single-unit auctions. A large number of important auctions are however multi-unit auctions. For example, the auctioning of T-bills, spectrum auctions, timber export licenses, auctions to sell the right to fell timber in Southern Ontario, etc. Internet auctions are multi-unit auctions too since more than one seller could be conducting an auction for the same object at different sites at almost the same time. Further in most cases identification is achieved conditional on the number of potential bidders. In several auctions for which field data is available, participation is endogenous. In many cases it is linked to the characteristics of the auctioned object some of which may not be observed or even obvious to an econometrician. The section on identification has illustrated that even with single-unit auctions and fixed number of potential bidders, it is difficult to establish identification outside the private-values model without further assumptions about the model of valuation. The key point is that each auction has characteristics that are idiosyncratic to it; establishing general results on identification may not be of much help while doing applied work. Hence making parametric assumptions may be a necessity to proceed with applied work. A structural model is identified from a Bayesian perspective if the prior and the posterior distribution of the structural elements are different. Equivalently a likelihood that is not flat in any direction of the parameter space will ensure identification; prior beliefs about the structural elements will *always* be revised through the likelihood.

Coming to estimation, the **indirect methods** are a comprehensive method for estimating structural auction models nonparametrically. However proving the properties of these estimators is a difficult exercise; Perrigne and Vuong (1999) point out the second step of these two-step estimators involves proving properties of the density estimator of unobserved variables. This leads them to comment that it may be more attractive to become parametric in the second-stage by specifying a parametric distribution for the pseudo-sample of private-values generated in the first stage. An undesirable consequence could be that the parametric distribution for the pseudo private-values may not correspond to the distribution of bids obtained in the first step since the equilibrium bidding rule has not been used to obtain this distribution.

A well recognized problem for maximum likelihood estimation is the dependence of the support of the data on the parameters. Standard asymptotic theory breaks down when the support of the data depends on the parameters. Donald and Paarsch (1996) and Hong (1998) have obtained the asymptotic theory of the maximum likelihood estimator for a first-price, sealed-bid auction with independent private values. The key point in these papers is that the sample minimum or maximum is a superconsistent estimator of the support of the data; that is, the sample minimum or maximum converges to the support of the data evaluated at the "true" parameter values at the same rate as the sample size. Since the support of the data is a function of some or all the parameters, the properties of the maximum likelihood estimator are, as a result, based on the properties of the sample maximum or minimum.

Compared to estimators based on **indirect methods** and maximum likelihood estimators, the properties of the simulation-based estimators are relatively easy to establish. However simulation-based estimators are method-of-moment estimators; drawbacks of method-of-moment estimators apply to the simulation-based methods as well.

Bayesian estimation proceeds by examining the posterior distribution of parameters. As a result even though they are likelihood based, Bayesian methods are not affected by the dependence of the support of the data on the parameters. A prior sensitivity analysis can be used to ensure the robustness of the estimates of the elements of the structural model.

Testing of structural auction models involves non-nested hypothesis testing with the support of the data being different under the null and the alternative structural auction model. The key challenge in testing is to find pivotal quantities in this scenario and to prove their properties. I have described testing in Section 3.5. The properties of the test statistics described in Section 3.5 are currently unknown. Further, the tests are extremely sensitive to the data that is available.. Bayesian testing can proceed either through the posterior odds ratio or comparing the posterior predictive density function of the structural models being considered. Since the emphasis is on comparing two structural models *conditional* on the observed data, the problems that plague testing described in Section 3.5 are not encountered.

Bayesian inference involves specifying the likelihood and the prior; these are discussed in the two subsections that follow. Section 4.3 deals with testing.

4.1 Likelihood specification

The likelihood for structural auction models is a function of the observed data and the unobserved variables $(\theta, \mathbf{X}, \mathbf{S})$. There are two ways in which the likelihood could be specified. One could work directly with the likelihood of observed data; this has been done in several papers by Bajari (1999) and Bajari and Hortacsu (2000). There are three problems with this approach.

First, to obtain the posterior "probability" of a particular value of θ, the equilibrium bidding rule has to be solved to evaluate the likelihood function. Ordinary differential equation solvers could be used for this. An alternative is to approximate the equilibrium bidding rule in a way to bypass inverting it to obtain the unobserved $(\mathbf{X}, \mathbf{S})$. Bajari (1998) suggests the use of quantal response equilibrium (QRE) approach first suggested by McKelvey and Palfrey (1995) to solve normal form games. A Bayesian-Nash equilibrium is a set of probability measures $B_1^*(b_1|x_1), ..., B_n^*(b_n|x_n)$ that maximizes expected utility for all i and all x_i. Note that x_i indicates a player's type now.[21] $B_i(b_i|x_i)$ is a probability measure over agent i's strategy set $B_i = \left(b_i^1, ..., b_i^\tau, ..., b_i^{J_i}\right)$. If $B_i(b_i|x_i)$ puts a point mass on an element b_i^τ of B_i and zero on the other elements of B_i, then b_i^τ will be interpreted as a probability measure. If the set of possible types of a player and the set of strategies is finite, expected utility is discretized; I will indicate expected utility of bidder i by $\overline{u}(b_i, \mathbf{B}_{-i}; x_i, \theta)$, where $\mathbf{B}_{-i}$ is an $n-1$ dimensional vector each element of which is the set of strategies for the n players excluding i. The basic idea underlying the QRE approach is that bidder i observe this expected utility with an error ϵ_i,

$$\widehat{u}_i(x_i) = \overline{u}(b_i, \mathbf{B}_{-i}; x_i, \theta) + \epsilon_i(x_i). \tag{13}$$

Assuming that $\epsilon_i(x_i)$ follows an extreme value distribution, player i's QRE strategy is

$$B_i\left(b_i|x_i, \lambda\right) = \frac{\exp\left(\lambda \overline{u}(b_i, \mathbf{B}_{-i}; x_i, \theta)\right)}{\sum_{\tau=1}^{J_i} \exp\left(\lambda \overline{u}(b_i^\tau, \mathbf{B}_{-i}; x_i, \theta)\right)}. \tag{14}$$

The likelihood function based on the QRE is

$$f_{b_1,\ldots b_n}\left(b_1^\tau, ..., b_n^\tau|\theta, \lambda\right) = \sum_{x_i \in X} \left\{ \prod_{i=1}^n B_i\left(b_i|x_i, \lambda\right) f_{\mathbf{x}}\left(\mathbf{x}|\theta\right) \right\}, \tag{15}$$

where X is the Cartesian product of all possible types for all n players in the game. λ is an additional parameter in the QRE approach; as λ becomes large, the QRE approaches the Bayesian-Nash equilibrium.

This approach, like all direct approaches, is implemented by specifying a distribution for $\mathbf{X}|\theta$. For each value of θ, the x_i's and b_i are obtained from assumptions about the underlying game-theoretic model; this allows the likelihood to be evaluated. Note that the simulated bids are the QRE bids and not the Bayesian-Nash equilibrium bids. They key is to put a prior on the parameter λ so that the simulated bids mimic the observed Bayesian-Nash bids with a few draws of θ. For

[21] For example in a procurement auction the set of possible types would be the possible set of costs for an agent. The private signal x_i for an agent i that has been referred to in Section 2 is just one element of the set of types.

example, Bajari (1998) puts a dogmatic prior of $\lambda = 25$ to obtain QRE bids which are good approximations to the observed bids. The idea is to sample the parameter space where the likelihood puts most of its mass.

Second, methods based on the likelihood of the bids have an additional problem in that the support of the bids is truncated with the truncation point depending on the parameter vector θ. For example, in a first-price, symmetric **IPV** auction, a bidders bid will be no lower than the expected value of the second-highest bid. Suppose $l(\theta)$ is the lower bound of the bid distribution and a bid b_j less than $l(\theta)$ is observed; then either the first moment of the distribution of private-values will be underestimated or the second moment will be overestimated to accommodate the outlier b_j. Unlike the likelihood function based on the Bayesian-Nash equilibrium bidding rule, the QRE is robust to the presence of outliers in the bidding data since it is based on the full likelihood.

Third, it is rarely the case that the bids of all n potential bidders is observed. Preparing and submitting bids is not a costless activity; there could be other reasons like the characteristics of the auctioned object which could prevent a bidder from submitting a bid. Hence additional problems of working with censored data arise whether one works with the likelihood based on the QRE or the Bayesian-Nash bidding rule.

Indirect methods shed an important insight into an alternative way of specifying the likelihood. First a pseudo-sample of private-values is generated; then this pseudo-sample is used to obtain an estimate of the private signals of the bidders. The key is to work with the unobserved private signals.

Making use of this idea, Sareen (1998) suggests working with the likelihood of the latent data $(\mathbf{X}, \mathbf{S})$ of all n potential bidders; this will be referred to as the *full* likelihood of the latent data. Since the support of the latent data does not depend on parameters θ the kind of problem mentioned with outliers for the Bayesian-Nash likelihood will not arise. Further, the *full* likelihood of the latent data includes all n potential bidders; hence censoring of the data will not be an issue. Working with the latent structure is also simplified because the likelihood of the latent data does not involve the Jacobian of the transformation from the distribution of the signals to that of the bids.

Assuming that the number of participants and potential bidders are identical, estimation is based on the following posterior distribution,

$$f_{\theta|data}(\theta|\mathbf{X}, \mathbf{S}, data) \propto f_\theta(\theta)\, f_{\mathbf{XS}|\theta}(\mathbf{x}, \mathbf{s})\, 1_{b=e(\bullet)}. \tag{16}$$

$f_\theta(\theta)$ is the prior density function for the parameters θ. $f_{\mathbf{XS}|\theta}(\mathbf{x}, \mathbf{s})$ is the specified density function of the signals $(\mathbf{X}, \mathbf{S})$. $1_{(\bullet)}$ is an indicator function. It is equal to one if $(\theta, \mathbf{X}, \mathbf{S}$,data) solve the differential equations that are the first-order conditions for the optimization exercise given by equation (1); otherwise it is zero. The indicator function is the likelihood function of the bids conditional on realization of $(\mathbf{X}, \mathbf{S})$ that solve the Bayesian-Nash equilibrium bidding rule given by (2). The basic idea of the method is to sample $(\mathbf{X}, \mathbf{S})$ for each draw of θ and retain only those draws of $(\mathbf{X}, \mathbf{S}, \theta)$ which along with the observed bids solve the system of differential equations given by (2).

If participation is endogenous the above scheme can be modified. Indicate by n_* the number of bidders who submit bids. N is the number of potential bidders; it is now a random variable which has to be estimated. First, consider estimation conditional on n; even if n is not observed, a close approximation is available.

$$f_{\theta|data}(\theta|\mathbf{X}, data) \propto f_\theta(\theta)\, f_{\mathbf{X}_{n_*}\mathbf{S}|\theta}(\mathbf{x}_{n_*}, \mathbf{s})\, 1_{b_*=e_*(\bullet)} f_{\mathbf{X}_{n-n_*}|\mathbf{X}_{n_*},\mathbf{S},\theta}(\mathbf{x}_{n-n_*})\,. \quad (17)$$

$\mathbf{X}_{n_*}$ and $\mathbf{X}_{n-n_*}$ are n_* and $n-n_*$ dimensional vectors of private signals, respectively. For each auction only n_* bids, $\mathbf{b}_{n_*}$ are observed. Given θ, $(\mathbf{x}_{n_*}, \mathbf{s})$ is drawn from the distribution specified for the signals $(\mathbf{X}_{n_*}, \mathbf{S})$;[22] note that unlike equation (16) n_* instead of n values of $\mathbf{X}$ are drawn. If a draw of $(\mathbf{x}_{n_*}, \mathbf{s}, \theta)$ solves the equilibrium bidding rule in equation (2) it is retained; otherwise it is discarded. This is what the indicator function is doing. Next $n-n_*$ signals are drawn from the distribution $\mathbf{X}_{n-n_*}|\mathbf{x}_{n_*}, \mathbf{s}, \theta$. Each of these draws will be less than the minimum of $\mathbf{x}_{n_*}$; conditioning on $\mathbf{x}_{n_*}$ ensures this. For example in the **IPV** model n_* draws will be made from the distribution $X|\theta$; they will be accepted if they solve the equilibrium bidding rule. Suppose the minimum of these accepted draws is $x^{\min}_{n_*}$. Next $n-n_*$ draws are made from $X|\theta$ such that each draw is less than $x^{\min}_{n_*}$.

In case a model of participation is specified, then sampling can be done iteratively in two blocks. Conditional on n, $(\mathbf{X}, \mathbf{S}, \theta)$ can be sampled in the manner above; then condition on $(\mathbf{x}, \mathbf{s}, \theta)$, the specified model of participation can be used to sample N. The problematic scenario of Corollary 1 will emerge here too if N has to be estimated from the high bid data without a model of participation.[23]

The key point is to work with the latent data. The latent data need not be obtained in the manner discussed above. It could be obtained nonparametrically as in the indirect methods. Alternatively in models like the **IPV** the equilibrium bidding rule is solved in a few iterations; here ordinary differential equations solvers could be used to obtain the latent data conditional on θ. Van den Berg and Van der Klaauw (2000) also use the *full* likelihood of the latent data to study Dutch auction for flowers assuming the **IPV** model. They observe not just the winning bid but all bids made within 0.2 seconds of the winning bid. Their full likelihood of the latent data is a modified version of that given by equation (17),

$$f_{\theta|data}(\theta|\mathbf{X}, data) \propto f_\theta(\theta)\, f_{\mathbf{X}|\theta}(\xi(\bullet)|\theta)\, 1_{b^{1:n} \leq E(V_{n-1}|\theta)}. \quad (18)$$

$1_{b^{1:n} \leq E(V_{n-1}|\theta)}$ is an indicator function that equals one if the smallest observed bid is less than the expected value of the second-highest private value. To simplify the exposition, let $n = n_*$ and that n is given.[24] The basic idea is to set up a Gibbs sampler with conditional distributions, $\mathbf{X}|\theta$ and $\theta|\mathbf{X}$. But $\mathbf{X}$ is unobserved;

[22] The distribution of $\mathbf{X}_{n_*}, \mathbf{S}|\theta$ is obtained from the distribution specified for $(\mathbf{X}, \mathbf{S})$ by integrating out $\mathbf{X}_{n-n_*}$.

[23] The inability of the winning bid to identify *both* N and θ shows up in the posterior distribution for N. For a simulated Dutch auction, Sareen (1998) observes a spike in the posterior density function of N at the lower boundary for the parameter space for N. Suppose N is defined on the *known* interval $[l_n, u_n]$. To estimate N, for each value of $N \in [l_n, u_n]$, the posterior for $\theta(n)$ can be obtained; features of the posterior will decide the estimate of N. This approach is similar in spirit to the SNLLS method of Laffont, Ossard and Vuong (1995).

[24] Van den Berg and Van der Klaauw (2000) assume participation is endogenous. Since the private signals are independent, $\mathbf{X}_{n-n_*}|\mathbf{X}_{n_*}$ is sampled by taking $n-n_*$ draws less than $\min\{\mathbf{x}_{n_*}\}$ from $f_{X|\theta}(\bullet)$. The sampling scheme is in two blocks as described above.

so for some $\widetilde{\theta}$ they obtain $x(\mathbf{b},\widetilde{\theta}) = \xi(\mathbf{b},\widetilde{\theta})$, the private signals of the bidders by numerically solving the equilibrium bidding rule; hence $\mathbf{X}|\theta$ is a spike at $x(\mathbf{b},\widetilde{\theta})$. Next substituting $x(\mathbf{b},\widetilde{\theta})$ in the likelihood for private signals, a draw from the posterior $\theta|x(\mathbf{b},\widetilde{\theta})$ given by equation (18) is obtained. Ideally what is required are $(\theta, \mathbf{X}, \mathbf{S}, \text{data})$ combinations that solve the equilibrium bidding rule; but then this sampling scheme would not work since it will remain at $\widetilde{\theta}$. Hence Van den Berg and Van der Klaauw (2000) accept a draw from $\theta|x(\mathbf{b},\widetilde{\theta})$ as long as it satisfies the constraint imposed by the likelihood of bids: the smallest observed bid is less than the expected value of the second highest private value. This will ensure that the posterior samples that part of the parameter space where the likelihood function puts most of its mass.

4.2 Prior Specification

Prior specification for structural auction models has to be done with care for several reasons.

Structural auction models are nonregular models in that the support of the data depends on θ. A prior on θ implies a prior on the support of the data $\zeta_B(\theta)$ as well. This implicit prior on the support of the data should be consistent with the observed data; that is, it should put nonzero mass on the part of the parameter space within which the observed data falls. An example will help to clarify this issue. Suppose the support of the data $\zeta_B(\theta)$ is the interval, $\zeta_B(\theta) = [\underline{b}(\theta), \infty)$. A proper prior, f_θ, is specified for θ such that the implied prior on the function $\underline{b}(\theta)$ is uniform on the interval $[\underline{s}, \overline{s}]$; thus, $\underline{s} \leq \underline{b}(\theta) \leq \overline{s}$. If the observed bids b_i lie in the interval $[\underline{s}, \overline{s}]$, the prior on θ is consistent with the observed data. Suppose some $b_i < \underline{s}$; then the prior for θ is not consistent with the observed data. f_θ puts a mass of zero on the part of the parameter space where the likelihood function of the observed bid b_i is non-zero.

They key point that emerges from this discussion is that the part of the parameter space where the likelihood and the prior put most of their mass should be similar. This is feasible in many auctions since the parameters have a natural interpretation as the limiting form of some observables.[25] Bajari (1999), Bajari and Hortacsu (2000) and Sareen's (1999) work provide many examples of this.

Priors specified in the manner above require a serious prior elicitation effort on part of the empirical researcher. For example, interpreting parameters in terms of observable to impose reasonable priors may not be obvious. It may also be the case that an empirical researcher may want to be noninformative to conduct a prior sensitivity analysis since the prior specification above is highly informative. I next turn to specificating of noninformative priors.

Jeffreys' prior is used as a standard noninformative prior in many instances. If $\mathbf{J_Z}(\theta)$ is the Fisher information matrix of the likelihood of the data $\mathbf{Z}$, the Jeffreys' prior for θ is

$$f^z_{jeff}(\theta) \propto \sqrt{\det \mathbf{J}_Z(\theta)}, \tag{19}$$

[25]The notion of parameters being limiting form of observables is from de Finetti's Representation theorem. Bernardo and Smith (1994, 172-181) provide a lucid explanation of this theorem.

where "det" indicated determinant of the matrix.

It is attractive to represent "vague" prior beliefs through the Jeffreys' prior for several reasons. It is easy to understand and implement. Unlike uniform priors it is *invariant*; Inference about θ is identical whether it is based on the Jeffreys' prior for θ or a one-to-one transformations of θ. Jeffreys recommended this prior to represent "noninformativeness" for single parameter problems from *iid* data. Beyond the single parameter case it is less ceratin whether Jeffreys' prior is a candidate for the position of "noninformative" priors. This has led to several modifications and reinterpretations of Jeffreys' prior.[26]

A modification proposed by Bernardo (1979) to represent "noninformativeness" is the reference prior. The reference prior emerges from maximizing an asymptotic expansion of Lindley's measure of information. Lindley's (1965) measure of information is defined as the expected Kulback-Liebler divergence between the posterior and the prior; the larger the measure, more informative the data and hence less informative the prior. Reference priors are appealing to both nonsubjective Bayesians and frequentists since the posterior probabilities agree with sampling probabilities to a certain order (Ghosh, Ghosal and Samanta, 1994). They are appealing to subjective Bayesians as well since they serve as a reference point in a prior sensitivity analysis. When there are no nuisance parameters and certain regularity conditions are satisfied, Bernardo's reference prior is the Jeffreys' prior.

In several papers, Ghosal and Samanta (1995) and Ghosal (1997), extending Bernardo's work, have obtained the reference prior when the support of the data depends on the parameters are in structural auction models; for an overview of the reference prior idea and it's extension to the nonregular case see Sareen (2000c, pp. 51-57). Suppose $\theta = [\eta,\varphi]$ with η being a scalar. The support of the likelihood of bids $\zeta_B(\theta) = \zeta_B(\eta)$ is strictly monotonic in η; η is referred to as the "nonregular" parameter. Since the support of the likelihood of the bids conditional on η does not depend on a parameter, φ is called the "regular" parameter. The reference prior for $\theta = [\eta,\varphi]$ is

$$\pi^B_{ref}(\eta,\varphi) \propto |c(\eta,\varphi)| \sqrt{\det \mathbf{J}_B^{\varphi\varphi}(\eta,\varphi)}; \tag{20}$$

$c(\eta,\varphi)$ is the score function of the likelihood of bids

$$c(\eta,\varphi) = E_{B|\eta,\varphi}\left[\frac{\partial}{\partial\eta}\log f_B(b_i|\eta,\varphi)\right],$$

and $\mathbf{J}_B^{\varphi\varphi}(\eta,\varphi)$ is the lower right hand block of $\mathbf{J}_B(\eta,\varphi)$, the Fisher information from bids. Intuitively, both the "regular" and the "nonregular" parameters contribute the standard deviation of the asymptotic distribution of the relevant estimator. Since the asymptotic distribution of the estimator of the "regular" and the "nonregular" parameter are different, the "regular" parameter φ contributes $\sqrt{\mathbf{J}_{\mathbf{Z}}^{\varphi\varphi}(\eta,\varphi)}$ and the "nonregular" parameter η contributes $c(\eta,\varphi)$ to the reference prior. [27]

[26] See Kleinberger (1994), Zellner (1971, pp. 216-220) and Phillips (1991) for examples.

[27] The MLE of the "regular" parameter φ after "concentrating" out the "nonregular" parameter from the likelihood function converges to a Normal distribution; $\sqrt{T}\left(\widehat{\varphi}_{ml} - \varphi_o\right) \sim N\left(0, \mathbf{J}_{\mathbf{B}}^{\varphi\varphi}\left(\widehat{\eta},\varphi\right)\right)$, where $\widehat{\eta} = w_*$. The MLE of the "nonregular" parameter η is the minimum

The reference prior is invariant like the Jeffreys' prior. In addition, it provides a way for handling "nuisance" parameters which Jeffreys' prior does not. The distinction between "nuisance" parameters and "parameters of interest" could be important in structural auction models since the support $\zeta_B(\theta)$ is like a "nuisance" parameter. Further, since the reference prior is based on the sampling density of the data, the nonregularity in the likelihood is taken into account while constructing the prior; inconsistency between the prior and the data of the kind discussed above will not arise with the reference prior.

When there are no "nuisance" parameters and certain regularity conditions are satisfied it is well known that Jeffrey's prior and the reference prior coincide. When the support of the data depends on the parameters the reference prior is *not* the Jeffreys' prior in general. Sareen (2000a) proves that the two **coincide** even when the support of the data depends on the parameters if the winning bid is sufficient for a scalar parameter θ. The necessary conditions under which a sample maximum/minimum is sufficient for a scalar parameter have been established by Huzurbazar (1976). These conditions restrict the functional form of the support and the density function from which the order statistics are drawn. Specifically if $B_j^i \geq \zeta_B(\theta)$, then the necessary and sufficient conditions under which the winning bid is sufficient for θ are

(1) $\zeta_B(\theta)$ is a *strictly* monotonic, continuous and differentiable function of θ; and
(2) The form of the density function of bids is

$$f_b(b_j^i \mid \theta) = m(b_j^i)q(\theta), \tag{21}$$

$m(b_j^i)$, $q(\theta)$ are strictly positive functions of b_j^i and θ, respectively.

Outside of this scenario the use of Jeffrey's prior to represent "noninformativeness" may lead to pathologies of the kind discussed above.

4.3 Testing

The posterior odds ratio is a means of comparing two models. It gives the odds of one model compared with another conditional on the observed data. Formally, let M_1 and M_2 be two models. Then the posterior odds ratio between model 1 and 2, P_{12} is

$$P_{12} = \frac{f_{M_1|data}(\bullet)}{f_{M_2|data}(\bullet)} = \frac{f_{M_1}(\bullet)f_{data|M_1}(\bullet)}{f_{M_2}(\bullet)f_{data|M_2}(\bullet)}. \tag{22}$$

$f_{M_\tau}(\bullet)$, $f_{data|M_\tau}(\bullet)$, $\tau = 1, 2$, are the prior probability and the marginal likelihood of model τ, respectively. The ratio $\frac{f_{M_1}(\bullet)}{f_{M_2}(\bullet)}$ are a researcher'sss beliefs, prior to observing the data, about which model is a more probable explanation of the auctions she is studying. These prior beliefs are revised on observing the data through the marginal likelihood

$$f_{data|M_\tau}(\bullet) = \int_{\Theta} f_{\theta|M_\tau}(\bullet) l\left(\theta; data, M_\tau\right) d\theta. \tag{23}$$

winning bid, W_*. It converges to an Exponential distribution; $T\left(W_* - \eta_o\right) \sim EXP\left(c(\eta, \varphi)\right)$, where $c(\eta, \varphi)$ is the sampling expectation of the score function of the bids.

$l(\theta; data, M_\tau)$is the likelihood function under model M_τ; $f_{\theta|M_\tau}(\bullet)$ is the prior for θ under model M_τ.

As long as $f_{\theta|M_\tau}(\bullet)$, the prior for θ under model M_τ, is proper the marginal likelihood of the data can be evaluated using the draws from the posterior distribution of the parameters (Chib and Jeliazkov, 2000). Improper priors for θ cause problems in calculating the marginal likelihood; since improper priors are defined up to a constant of proportionality, the scaling of the marginal likelihood is arbitrary. Using "vague proper priors" does not solve this problem. For example, Bajari (1999) and Bajari and Hortacsu (2000) uses uniform priors on $[l_i, u_i], i = 1, ..., k$ for each of the k parameters. The Bayes factor is proportional to a constant so that the resultant answer will again be arbitrary like the "vague improper prior"; Berger and Pericchi (1997, p. 2) point out this out with several examples.

Several "default" Bayes factors have been suggested to get around this problem; Kass and Raftrey (1995, pp. 773-795) provide a survey of these options. The basic idea underlying these "default" options is to use a subset of the data, called a training sample, to "convert" the improper priors under each model into a proper posterior; this is referred to as an intrinsic prior. As its name suggests, the intrinsic prior is used as a prior to define a Bayes factor for the remaining data. Since the resultant Bayes factor will depend on the size of the training sample, an "average" over all possible training samples is taken.

They key point is that a "default" option cannot be used indiscriminately; what will work best depends on the applied problem at hand. For example, for a scalar nonregular parameter and truncated exponential likelihoods, Berger and Pericchi (1997, pp. 12-13) discuss "default" Bayes factors for a one-sided hypothesis testing. The problem they encounter is that the intrinsic prior is unbounded under the null hypothesis.

Informal means to compare models could be used to circumvent the problem of working with the posterior odds ratio and improper priors. For example, Sareen (1999) compares the posterior predictive distribution under the pure **CV** and the **IPV** models. The posterior predictive distribution represents beliefs about an out-of-sample bid after observing the data; this is in contrast to the prior predictive distribution which represents beliefs about this out-of-sample bid before observing the data. The extent of "divergence" between the prior and posterior predictive distribution for a model indicates the learning from data about that model. A model for which this "divergence" is large should be favored as there is more learning from the data about this model.

5 Unobserved Heterogeneity

Empirical work requires data on several auctions. It is rarely the case that these auctions are identical; either the auctioned object or the environment under which an auction is held or both could differ across auctions. Some of this auction/object heterogeneity may be observed by the econometrician. Many characteristics which make the auctioned object different will not be observed; this is termed as unobserved heterogeneity. Since the value of an object determines whether a bidder will participate or not, these unobserved characteristics would affect not just the bidding behavior but participation by bidders as well.

A standard approach to modelling unobserved heterogeneity is to give it a parametric form. This makes direct estimation more complicated since this unobserved heterogeneity enters the Bayesian-Nash equilibrium bid for each auction in addition to the unobserved private values already present in the equilibrium bidding rule. Indirect estimation methods have been unable to address this issue; this is one of the reasons that Perrigne and Vuong (1999) recommend a parametric second stage in their survey.

An alternative approach has been suggested and implemented for the **IPV** model by Chakraborty and Deltas (2000). It falls within the class of direct methods in that parametric assumptions are made about the distribution of $\mathbf{X}|\theta_j$; it is assumed to belong to the location-scale class either unconditionally or conditional on a shape parameter. The subscript j in θ_j indicates that the estimation of the distribution of valuations is auction-specific; *only* bids within an auction are used to obtain $\widehat{\theta}_j$ for each auction j. θ_j incorporates both observed and unobserved heterogeneity despite the fact that no auction-specific covariates are used to obtain $\widehat{\theta}_j$. Indicating observed heterogeneity in auction j by $\mathbf{Z}_j$ and unobserved heterogeneity by ν_j the following relationship

$$\theta_j = \mathbf{r}\left(\mathbf{Z}_j, \nu_j|\delta\right), \tag{24}$$

is used to recover estimates of the coefficients of observed heterogeneity δ in the second step. $\mathbf{r}(\bullet)$ is a k dimensional function of the observed and unobserved heterogeneity in each auction j. The estimates of θ_j obtained in the first step are used instead of θ_j. The estimates of δ obtained in the second step are robust to unobserved heterogeneity in the sense that they do not depend on the distribution of ν. In addition, breaking the estimation into two parts simplifies the estimation of δ. The first stage of the estimation which obtains θ is structural as it is based on the Bayesian-Nash equilibrium bidding rule. Reduced form estimation is done in the second stage to obtain $\widehat{\delta}$.

The two-stage procedure of Chakraborty and Deltas (2000) is similar to the following three-stage Bayesian hierarchical model,

$$\begin{aligned} &\mathbf{B}_j|\theta_j \\ &\theta_j|\mathbf{Z}_j, \delta \\ &\delta, \end{aligned} \tag{25}$$

where $\mathbf{b}_j = \left(b_1, ..., b_{n_j}\right)$ are the bids observed for auction j. The first stage is obtained from the parametric assumptions made about the distribution of the signals $\mathbf{X}|\theta_j$ and the Bayesian-Nash equilibrium bidding rule. The second stage is obtained from the relation specified in equation (24). The last stage is the specification of the prior for δ. Sampling the posterior for (θ_j, δ) can be done in blocks. The following blocks fall out naturally from the ideas of Chakraborty and Deltas (2000),

$$\begin{aligned} &\theta_j|\mathbf{B}_j && \text{for each auction } j, \\ &\delta|\mathbf{Z}, \theta && \text{across auctions,} \end{aligned} \tag{26}$$

where $\mathbf{Z} = (\mathbf{Z}_1, ..., \mathbf{Z}_n)$ and $\theta = (\theta_1, ..., \theta_n)$. In case participation is endogenous an iterative sampling scheme could be set up. First, the blocks in equation (25) are sampled conditional on n. Then conditional on draws of (θ, δ), the parameters of the process used to model entry could be sampled.

6 Conclusion

This survey describes the menu of techniques for identification, estimation and testing available to an empirical researcher. The choice between structural *vs* reduced form, direct methods *vs* indirect methods will depend on the goal of the empirical exercise, the available data and the theory developed for the underlying game-theoretic model. Each of these aspects is commented on in turn.

If the eventual goal of the empirical exercise is mechanism design, the structural approach would be preferred over the reduced form approach. In many auctions, a seller may be interested in goals other than revenue maximization. For example in the timber auctions in Southern Ontario, the county of Simcoe wants to encourage a viable local timber industry. Since a when-issued market and a secondary market exists for T-bills and other government debt, one of the aims of a central bank in conducting T-bill auctions would be to promote liquidity in the market for government debt. In these instances reduced-form estimation may be preferred especially if the theory for the underlying game-theoretic model is not well developed.

In certain instances various features of the distribution of private-values could be of interest; for example, Li, Perrigne and Vuong (2000) were interested in estimation the magnitude of the "money left on the table" for the OCS wildcat auctions. For estimation indirect methods could be preferred since they do not impose any distributional assumption on the private signals of the bidders. Since the rate of convergence of estimators based on indirect methods is slow, large data sets are required to implement these estimators. For example, if data on winning bids is observed, one may be forced to turn to direct methods to establish the identification of the underlying game-theoretic model in the first instance.. In general when testing a structural model against an explicit alternative, Bayesian methods may be preferred since they allow testing *conditional* on the observed data.

Estimating and testing structural models is a commendable goal; it presupposes that the underlying game-theoretic model exists and that it's properties have been established. With few exceptions, our understanding of game-theoretic models is confined to single-unit, symmetric private-values auctions. The challenge in doing empirical work is how to adapt the existing theory to the idiosyncratic features of a specific auction; here the role of conditioning and the details observed about each auction will be the guiding tools. Chakraborty and Deltas (2000) incorporate unobserved heterogeneity without altering the equilibrium bidding rule of a symmetric **IPV** first-price auction. Bajari and Hortacsu (2000) explaining last minute bidding in eBay auctions through a two-stage game; conditional on a stochastic entry process in the first-stage, the second stage is similar to a symmetric sealed-bid, second-price auction. Athey and Haile (2000) and Hendricks, Pinkse and Porter (1999) identify and estimate standard auction models by observing additional details of that auction Thus *ex post* value of the auctioned object is used to identify the **CV** model; bidder identity is used to identify asymmetric models. Similarly, if the component of private-values that makes bidders asymmetric was observed, then it may be possible to use the theory of symmetric games by conditioning on this asymmetric component.

References

[1] S Athey, PA Haile. Identification of standard auction models. Woking paper, University of Wisconsin-Madison, 2000.

[2] S Athey, J. Levin. Information and competition in U.S. forest service timber auctions. Journal of Political Economy, forthcoming, 2000.

[3] P Bajari. Econometrics of sealed bid auctions. American Statistical Association, Proceedings of the Business and Economic Statistics Section, 1998.

[4] P Bajari. Econometrics of first price auctions with asymmetric bidders. Working paper, Stanford University, 1998.

[5] P Bajari, A. Hortacsu. Winner's curse, reserve price and endogenous entry: empirical insights from eBay auctions. Working paper, Stanford University, 2000.

[6] JO Berger, L Pericchi. On criticisms and comparisons of default Bayes factors for model selection and hypothesis testing. In: W Racugno, ed. Proceedings of the workshop on model selection, Bologna: Pitagora Editrice, 1997.

[7] SM Berman. Note on extreme values, competing risks and semi-markov processes. Annals of mathematical statistics 34:1104-1106, 1963.

[8] JM Bernardo. Reference Posterior Distributions for Bayesian Inference. J Roy Statist Soc *B*, 41:113-147, 1979.

[9] JM Bernardo, AFM Smith. Bayesian Theory. New York: Wiley, 1994.

[10] S Campo, I Perrigne, Q Vuong. Asymmetry and joint bidding in OCS wildcat auctions. Working paper, University of Southern California, 1998.

[11] S Campo, I Perrigne, Q Vuong. Semiparametric estimation of first-price auctions with risk-averse bidders. Working paper, University of Southern California, 2000.

[12] I Chakraborty, G Deltas. A two-satge approach to structural econometric analysis of first-price auctions. Working paper, University of Illinois, 2000

[13] S Chib, I Jeliazkov. Marginal likelihood from the Metropolis-Hastings output. Working paper, Washington University, 2000.

[14] S Donald, HJ Paarsch. Identification, estimation and testing in parametric models of auctions within the independent private values paradigm. Econometric Theory 12:517-567, 1996.

[15] S Donald, HJ Paarsch, J Robert. Identification, estimation and testing in empirical models of sequential, ascending-price auctions with multi-unit demand: an application to Siberian timber-export permits. Working paper, University of Iowa, 1999.

[16] S Ghosal. Reference Priors in Multiparameter Nonregular case. Test 6:159-186, 1997.

[17] S Ghosal, T Samanta. Asymptotic Behavior of Bayes Estimates and Posterior Distributions in Multiparameter Nonregular Case. Math. Methods Statist. 4:361-388, 1995.

[18] JK Ghosh, S Ghosal, T Samanta. Stability and Convergence of Posterior in Nonregular Problems. In: S.S Gupta and J.O. Berger, eds. Statistical Decision Theory and Related Topics V. New York: Springer-Verlag, 1994, pp 183-199.

[19] E Guerre, I Perrigne, Q Vuong. Nonparametric estimation of first-price auctions. Working paper, Economie et Sociologie Rurales, Toulouse no. 95-14D, 1995.

[20] E Guerre, I Perrigne, Q Vuong. Optimal nonparametric estimation of first-price auctions. Econometrica 68:525-574, 2000.

[21] JJ Heckman, BE Honoré. The empirical content of the Roy model. Econometrica: 58:1121-1149, 1990.

[22] K Hendricks HJ Paarsch. A survey of recent empirical work concerning auctions. Canadian Journal of Economics 28:403-426, 1995.

[23] K Hendricks, J Pinkse, RH Porter. Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. Working paper, Northwestern University, 1999.

[24] K Hendricks, RH Porter. An empirical study of an auction with asymmetric information. American Economic Review 78:865-883, 1988.

[25] K Hendricks, RH Porter, B Boudreau. Information, Returns, and Bidding Behavior in OCS Auctions: 1954-1969. The Journal of Industrial Economics 34:517-542, 1987.

[26] PA Haile, H Hong and M Shum. Nonparametric tests for common values in first-price auctions. Working paper, University of Wisconsin-Madison, 2000.

[27] H Hong, M Shum. Increasing competition and the winner's curse: evidence from procurement. Working paper, University of Toronto, 2000.

[28] H Hong. Nonregular maximum likelihood estimation in auction, job search and production frontier models. Working paper, Princeton University, 1997.

[29] RE Kass and AE Raftrey. Bayes Factors. Journal of the American Statistical Association 90:773-795, 1995.

[30] F Kleibergen. Identifiability and nonstationary in classical and Bayesian econometrics. PhD dissertation, Erasmus University, Rotterdam, 1994.

[31] I Kotlarski. On some characterization of probability distributions in Hilbert spaces. Annali di Matematica Pura ed applicata 74:129-134, 1966.

[32] JJ Laffont. Game theory and empirical economics: the case of auction data. European Economic Review 41:1-35, 1997.

[33] JJ Laffont, H Ossard, Q Vuong. Econometrics of first-price auctions. Econometrica 63:953-980, 1995.

[34] JJ Laffont, Q Vuong. Structural analysis of auction data. American Economics Review, Papers and Proceedings 86:414-420, 1996.

[35] T Li, I Perrigne, Q Vuong. Conditionally independent private information in OCS wildcat auctions. Mimeo, University of Southern California, 1999.

[36] T Li, Q Vuong. Nonparametric estimation of the measurement error model using multiple indicators. Journal of Multivariate Analysis 65:139-165, 1998.

[37] DV Lindley. On a Measure of Information Provided by an Experiment. Annals of Mathematical Statistics 27:986-1005, 1956.

[38] D Lucking-Reiley. Auctions on the internet: what is being auctioned and how?. Journal of Industrial Economics, forthcoming, 2000.

[39] RP Mcafee, J McMillan. Auctions and bidding. Journal of Economic Literature 25:699-738, 1987.

[40] R McKelvey, T Palfrey. Quantal response equilibria for extensive form games. Working paper, CalTech, 1997.

[41] PR Milgrom, J Weber. A theory of auctions and competitive bidding. Econometrica 50:1089-1122, 1982.

[42] HJ Paarsch. Deciding between the common and private values paradigm in empirical models of auctions. Journal of Econometrics 51:191-215, 1992.

[43] HJ Paarsch. Deciding an estimate of the optimal reserve price: an application to British Columbian timber sales. Journal of Econometrics 78:333-357, 1997.

[44] I Perrigne, Q Vuong. Structural economics of first-price auctions: a survey 47:203-233, 1999.

[45] PCB Phillips. To criticize the critics: an objective Bayesian analysis of stochastic trends. Journal of Applied Econometrics 6:333-364, 1991.

[46] BLP Rao. Identifiability in stochastic models: characterization of probability distributions. Academic Press: San Diego, 1992.

[47] S Sareen. Relating Jeffreys' prior and reference prior in nonregular problems: an application to a structural, low-price, sealed-bid auction. Unpublished manuscript, 2000a.

[48] S Sareen. Evaluating data in structural parametric auctions, job-search and Roy models. Unpublished manuscript, 2000b.

[49] S Sareen. Estimation and testing of structural parametric sealed-bid auctions. PhD dissertation, University of Toronto, Toronto, 2000c.

[50] S Sareen. The economics of timber auctions in Southern Ontario. Unpublished manuscript, 2000d.

[51] S Sareen. Posterior odds comparison of a symmetric low-price, sealed-bid auction within the common-value and the independent-private-values paradigms. Journal of Applied Econometrics 14:651-676, 1999.

[52] S Sareen. Likelihood based estimation of symmetric, parametric structural auction models with independent private values. Unpublished manuscript, 1998.

[53] G Van den Berg, B Van der Klaauw. Structural empirical analysis of Dutch flower auctions. Working paper, Free University Amsterdam, 2000.

[54] R Weber. Multiple-object auctions. In: R Engelbrecht-Wiggans, M Shubik, R Stark, ed. Auctions, Bidding and Contracting: Uses and Theory. New York: New York University Press, 1983, pp 165-191.

[55] A Zellner. An introduction to Bayesian inference in econometrics. New York: Wiley, 1971.