Flexible Bayesian Analysis of First Price Auctions Using Simulated Likelihood (Job Market Paper)*

Dong-Hyuk Kim†

November 24, 2009

Abstract

To analyze bid data from first price auctions, this paper develops an empirical framework that fully exploits all the shape restrictions arising from economic theory: bidding monotonicity and density affiliation. We directly model the valuation density so that bidding monotonicity is automatically satisfied, and restrict the parameter space to rule out all the non-affiliated densities. Our method uses a simulated likelihood to allow for a very flexible specification, but our posterior analysis is exact for the chosen likelihood. Our method controls the smoothness and tail behavior of the vaulation density and provides a decision theoretic framework for auction design. We reanalyze the sample from Outer Continental Shelf auctions that has been widely used. Our approach gives significantly different policy prescriptions on the choice of reserve price than previous methods, suggesting the importance of the theoretical shape restrictions.

Keywords: first price auctions, affiliated private values, auction design, optimal reserve price, Bayesian analysis, flexible specification, simulated likelihood, density estimation, shape restriction

JEL classification: C11, C13, C14, C15, C44, D44, D78, D81, L38

*I am greatly indebted to Keisuke Hirano, my advisor, for his guidance, patience, and encouragement. I am grateful to my professors, Jonah Gelbach, Gautam Gowrisankaran, John Wooders, and Mo Xiao for their invaluable advices and suggestions. I would also like to thank Tong Li for providing me with the data.

†Economics Department at the University of Arizona, Tucson, AZ. Email: donghyuk@email.arizona.edu.

1 Introduction

Auctions are a frequently used market institution for allocating a variety of economic resources and account for a significant portion of the U.S. economy. For example, since 1954 the U.S. government has leased approximately thirty thousand tracts for offshore oil and gas development in areas of the Gulf of Mexico. Each of these *Outer Continental Shelf* (OCS) auctions has a sale price which ranges from hundreds of thousands to several million dollars.[1] Reflecting the economic significance of auctions, many economists have extensively investigated this mechanism. In particular, auction theory has characterized equilibrium bidding behavior and optimal auction design for a given valuation distribution, while empirical research has focused on estimating the valuation distribution and using these estimates for auction design, e.g., the estimation of optimal reserve prices.

For first price sealed bid auctions with independent private values (IPV), the pioneers in the empirical auction literature specify the valuation distribution using strong parametric assumptions.[2] However, a flexible specification is often preferred because inference strongly relies on the shape of valuation distribution. For this reason Guerre, Perrigne, and Vuong (2000) indirectly recover the valuation density by using the estimated bid distribution. Li, Perrigne, and Vuong (2002) generalize this method to affiliated private value auctions (APV). These indirect approaches have been widely used because they are fully flexible and computationally simple.

However, for the estimation of bid distribution, the indirect methods do not fully exploit the shape restrictions arising from economic theory: bidding monotonicity and density affiliation. Thus, these methods may rely on an estimated bid density that is either nonaffiliated

[1]For more information, visit the website of the Offshore Energy and Minerals Management (OEMM) at `www.mms.gov/offshore/`. As can be found there, similar offshore oil and gas auctions have being held in other areas such as Alaska, Pacific, and Atlantic. The government also has auctioned off various onshore mineral rights, timber rights, magnetic spectrums, etc. In private sector, countlessly many commodities are auctioned on/off-line.

[2]See Donald and Paarsch (1993, 1996), Laffont, Ossard, and Vuong (1995), and Li and Vuong (1997).

or associated with a nonincreasing inverse bidding function. This problem is empirically important for the datasets of interest to researchers and policymakers. For instance, Li, Perrigne, and Vuong (2003) propose an optimal reserve price for the OCS auctions using a bid sample of 217 observations. We find that their policy recommendations are based on an estimated bid density that cannot be generated by an equilibrium. Note that recently Handerson, List, Millimet, Parmeter, and Price (2008) develop a kernel approach to impose bidding monotonicity, but do not discuss affiliated private values.

Motivated by this, we develop a Bayesian framework that satisfies both bidding monotonicity and density affiliation to provide more precise inference. Specifically, we directly parametrize the valuation density so that bidding monotonicity is satisfied for every parameter value and put zero prior weight on every density that violates density affiliation. For a reasonably flexible analysis we employ a series representation for the valuation density. To handle such a rich specification, we use a simulated likelihood. In particular, we employ a multinomial likelihood defined on a finely discretized sample space. Then, as Flury and Shephard (2008) discuss, we can obtain an exact posterior even with a finite number of simulation draws for each candidate parameter value.[3] Note that the simulation error would typically inflate the asymptotic variance of other simulation methods.

We revisit a sample of OCS auctions that Li, Perrigne, and Vuong (2003) have analyzed. Our methodology simulates the posterior distribution of the valuation density for the OCS wildcat auctions. The resulting bid density estimate fits the data very well. We select a reserve price to maximize the seller's future revenue using the Bayesian decision method introduced by Kim (2008). We find that a reserve price of $462 per acre is optimal given our likelihood and prior. According to our counterfactual analysis, this price increases the predictive revenue for each tract by $262,414 relative to the actual reserve price of $15. Note also that our recommended reserve price is much larger than the value of $273 obtained by

[3]See Andrieu, Doucet, and Holenstein (2007), Andrieu, Doucet, and Roberts (2007). Note that the information loss from the discretization can be minimal, if we use small bins.

Li, Perrigne, and Vuong (2003), suggesting the importance of the additional theoretical shape restrictions.

2 Auction Models and Empirical Method

This section defines the APV/IPV auctions, develops our empirical methodology, and discusses its difference from the indirect approaches.

2.1 First Price Auctions under the APV/IPV Paradigm

Consider $N \geq 2$ risk neutral bidders in an auction with a reserve price ρ. The bids are collected simultaneously and the bidder with the highest bid obtains the auctioned item at the price equal to his own bid. Let $(v_1, \ldots, v_N) \in \Re_+^N$ be a vector of valuations drawn from a joint distribution F assumed to be absolutely continuous with density f. Each bidder $i = 1, \ldots, N$, after observing his own valuation v_i, bids b_i to maximize his expected utility, $(v_i - b_i)\Pr(b_i > \max_{j \neq i} b_j)$. The number of bidders N, reserve price ρ, and the valuation distribution F are common knowledge among the bidders.

The auction is said to be under the APV paradigm if the valuations $v_1, \ldots, v_N$ are affiliated. Informally, affiliation implies that if some elements of $(v_1, \ldots, v_N)$ are large, others are more likely to be also large. (See Milgrom and Weber (1982) for a rigorous treatment.) Hence, each bidder gets some information on the distribution of other bidders' valuation from his own valuation and takes it into account to bid optimally. We assume that F is exchangeable so that every bidder is ex-ante identical. This auction mechanism implies a symmetric game of incomplete information among the bidders. For this game, Milgrom and Weber (1982) derive a symmetric Bayesian Nash equilibrium with a strictly increasing bidding function. Let $f_{y_i|v_i}(\cdot|\cdot)$ be the conditional density of $y_i := \max_{j \neq i} v_j$ given v_i. If $v \geq \rho$, the equilibrium bidding function is given by

$$\beta(v|\rho, F) := v - \int_{\rho}^{v} \exp\left\{-\int_{\alpha}^{v} \frac{f_{y_1|v_1}(u|u)}{\int_0^u f_{y_1|v_1}(t|u)dt} du\right\} d\alpha \tag{1}$$

Otherwise, any value strictly less than ρ is optimal. Note that the IPV paradigm is a special case of the APV paradigm in which $v_1, \ldots, v_N$ are independent. Under the IPV, Equation (1) simplifies to

$$\beta(v|\rho, F) := v - \int_{\rho}^{v} \left\{\frac{F(\alpha)}{F(v)}\right\}^{N-1} d\alpha \tag{2}$$

where $F(\cdot)$ is the marginal distribution of an individual valuation, with a slight abuse of notation.[4]

Suppose we observe a random sample of T auctions with the common valuation distribution F and a zero reserve price. Let $\beta(\cdot|F) := \beta(\cdot|0, F)$. We assume that for each auction, the N bidders follow the equilibrium bidding function $\beta(\cdot|F)$. Let $z := \{(b_{1,t}, \ldots, b_{N,t})\}_{t=1}^{T}$ denote the data set. From these data we want to learn F.

2.2 Bayes with Simulated Likelihood

For this purpose, we model the valuation density using a flexible specification $f(\cdot|\theta)$ and rule out all θ's with a nonaffiliated $f(\cdot|\theta)$. (For the IPV case, we model only the marginal of the valuation density.) We take a Bayesian approach for the following reasons. First, we can formally control smoothness and tail behavior reflecting our prior beliefs. Thus, our data analysis, while flexible, would not result in an unreasonably noisy or thick-tailed density. Second, for the choice of reserve price, we can use Bayes rule that may produce higher revenue

[4] Under the IPV paradigm with the exchangeability assumption, the joint valuation distribution is the product of N identical marginals. i.e., $F(v_1, \ldots, v_N) = \prod_{i=1}^{N} F(v_i)$. This abuse of notation should not lead to any confusion.

by formally considering the structure of payoffs and parameter uncertainty.[5] Third, we can handle the restricted parameter space (due to affiliation) simply by putting zero prior over the excluded area of the parameter space.

However, the likelihood is challenging to evaluate. Consider the likelihood

$$\prod_{t=1}^{T}\left\{\frac{f\left(\beta^{-1}(b_{1,t}|\theta),\ldots,\beta^{-1}(b_{N,t}|\theta)|\theta\right)}{\prod_{i=1}^{N}\beta'\left(\beta^{-1}(b_{i,t}|\theta)|\theta\right)}\right\}\cdot 1\left\{\max(z)\leq\bar{b}(\theta)\right\} \tag{3}$$

where $\bar{b}(\theta) := \lim_{v\to\infty}\beta(v|\theta)$ and $1\{A\} = 1$ if A is true, otherwise, zero. Since (3) lacks a closed form expression for our specification, we would need to numerically evaluate $\beta^{-1}(\cdot|\theta)$ at each data point for many parameter values. This is impractical because even for a single evaluation of $\beta^{-1}(\cdot|\theta)$, we need to evaluate the triple integrals in (1) repeatedly. For this reason, we simulate the equilibrium bids to construct a likelihood; generate random values from $f(\cdot|\theta)$, evaluate (1) at each drawn value, and then estimate the likelihood using these simulated bids.

When the likelihood is simulated, the Bayesian method is even more useful. Andrieu, Doucet, and Holenstein (2007) and Andrieu, Doucet, and Roberts (2007) show that using an unbiased simulated likelihood a Markov Chain Monte Carlo (MCMC) algorithm converges to the exact posterior distribution. The idea is simple. Let $p(\theta)$ and $p(z|\theta)$ denote a prior and a conditional density of data z given θ, and let $p_u(z|\theta)$ be an unbiased estimator for $p(z|\theta)$ constructed by a uniform random vector u. $E[p_u(z|\theta)] = \int p_u(z|\theta)p(u)du = p(z|\theta)$ and $\int p_u(z|\theta)p(u)du = \int p_u(z|\theta)du$. Thus, $\int p_u(z|\theta)du = p(z|\theta)$ and hence $p_u(z|\theta)$ can be regarded as a joint density of (u, z), say, $p(u, z|\theta)$. Suppose we apply an MCMC algorithm to $p(\theta)p_u(z|\theta)$ instead of $p(\theta)p(z|\theta)$ drawing a new u at each iteration. Then, this algorithm

[5] See Kim (2008) for more detail. We also use Bayes rule to choose a reserve price to maximize the seller's predictive revenue for the OCS wildcat auctions in section 4.

can be seen as an sampling scheme generating $(u_1, \theta_1), \ldots, (u_S, \theta_S)$ from

$$p(u, \theta|z) \propto p(\theta)p(u, z|\theta) = p(\theta)p_u(z|\theta)$$

Therefore, $\theta_1, \ldots, \theta_S$ are draws from the exact posterior $p(\theta|z)$. This method is called the Bayes with Simulated Likelihood (BSL) in this paper.[6]

To construct an unbiased simulated likelihood, we discretize the sample space into D bins.[7] Let y_d denote the number of sample points in the d-th bin. In addition, let $\pi_d(\theta)$ be the probability of the d-th bin under θ. Then, the discretized sample space with the implied histogram $y := (y_1, \ldots, y_D)$ leads to a multinomial likelihood given by

$$L(\theta|y) = \prod_{d=1}^{D} \{\pi_d(\theta)\}^{y_d} \tag{4}$$

Let $\widehat{\pi}_d(\theta)$ be the fraction of simulation draws belonging to the d-th bin. Then, we estimate (4) unbiasedly using

$$\widehat{L}(\theta|y) := \prod_{d=1}^{D} \{\widehat{\pi}_d(\theta)\}^{y_d} \tag{5}$$

This approach is very easy to compute and fairly flexible.[8] For a given discretization, we obtain the exact posterior by using a new simulation draw at each iteration. The information loss of using y rather than z would be minimal if the bins are small.

An alternative method would be the maximum likelihood estimator (MLE) to maximize

[6]Flury and Shephard (2008) introduce this method to econometrics. Note that Fernandez-Villaverde, Rudio-Ramirez, and Santos (2006) and Fernandez-Villaverde and Rudio-Ramirez (2007) have employed simulation to form a likelihood.

[7]Note that, for an APV paradigm, one datum is an observed auction, or its bid profile $(b_{1,t}, \ldots, b_{N,t}) \in z$. But, for an IPV model, a single bid, $b_{i,t}$, can be used as a sample point, since all the observed bids are independently and identically distributed.

[8]Note that it might seem to be natural to estimate the theoretical bid density using a kernel method or a method of series over the simulated auctions to construct likelihoods. However, these methods are biased and, moreover, the latter would not be computationally simple, if flexible.

(3) or its simulation estimate. However, as Hirano and Porter (2003) show, the MLE is not efficient because the bid data has parameter dependent support, $[0, \bar{b}(\theta)]$. Moreover, maximizing (3) under many constraints (due to affiliation) over high dimensional parameter space would be much more involved. Lastly, simulation typically inflates the asymptotic variance of the MLE and also the number of simulation draws needs to grow to infinity unlike BSL.

2.3 Valuation Density Specification

We discuss our choice for $f(\cdot|\theta)$ for the IPV and APV, separately. The simpler goes first.

IPV case

We use the specification of Verdinelli and Wasserman (1998). Let $\tilde{F}$ be a simple parametric distribution indexed by μ with density $\tilde{f}$, h be a flexible density on $[0,1]$ with parameter ψ, and $\theta := (\mu, \psi)$. The specification is given by

$$f(v|\theta) = \tilde{f}(v|\mu) h(\tilde{F}(v|\mu)|\psi) \tag{6}$$

Note that (6) can be seen as a derivative of a flexible CDF, $H(\tilde{F}(v|\mu)|\psi)$.

We construct $h(\cdot|\psi)$ as follows. Let $\{\phi_i\}$ denote a sequence of functions on $[0,1]$ such that we can approximate any function on $[0,1]$ using their linear combination $\sum_{i \in I} \psi_i \phi_i$ for some index I and some real numbers $\{\psi_i\}_{i \in I}$. Polynomials, splines, or Fourier functions can construct such $\{\phi_i\}$. Then,

$$h(x|\psi) \propto \exp\left(\sum_{i \in I} \psi_i \phi_i(x)\right) \cdot 1\{x \in [0,1]\} \tag{7}$$

approximates a density on $[0,1]$. This approach is a method of series. Note that the semi-nonparametric method of Gallant and Nychka (1987) is an example of series estimation in econometrics.

We employ this particular specification for the following reasons. First, it is a parsimonious way to specify a flexible model. To see this, take a logarithm on each side of (6). Then, $\log f(v|\theta) = \log \tilde{f}(v|\mu) + \psi_1\phi_1(\tilde{F}(v|\mu)) + \psi_2\phi_2(\tilde{F}(v|\mu)) + \cdots$. Thus, we approximate the valuation density first using $\log \tilde{f}$ and then explain the difference between the true density and $\log \tilde{f}$ using the additional terms. Obviously, for a given accuracy, if $\tilde{f}$ is already a good approximation to the density of valuations, we would not need many additional terms. Hence, we may deal with low dimensional parameter space. In this case, a computational advantage also follows. In practice, one may choose $\tilde{f}$ to reflect beliefs about the form of the true valuation density. For example, if bid data are exponential-like distributed, one could use an exponential with hazard rate μ, believing the valuations may be similarly distributed. Second, when it extends to the APV, exchangeability and density affiliation can be simply characterized as follows.[9]

APV case

We focus on two bidder case ($N = 2$) for clarity. Then, (6) extends to

$$f(v_1, v_2|\theta) = \tilde{f}(v_1|\mu)\tilde{f}(v_2|\mu)h\left(\tilde{F}(v_1|\mu), \tilde{F}(v_2|\mu)|\psi\right) \tag{8}$$

and (7) turns to

$$h(x_1, x_2|\psi) \propto \exp\left(\sum_{i\in I}\sum_{j\in I}\psi_{i,j}\phi_i(x_1)\phi_j(x_2)\right) \cdot 1\left\{(x_1, x_2) \in [0,1]^2\right\} \tag{9}$$

[9] Though the normal mixture studied by Ferguson (1973), Escobar (1994), and Escobar and West (1995) are popularly used, it is not appropriated for our purpose because it is hard to extend to the APV. Especially, there is no convenient way to impose density affiliation.

Recall that (8) must be affiliated and exchangeable.[10] The affiliation (or equivalently density log-supermodular) holds if and only if

$$\frac{\partial^2}{\partial x_1 \partial x_2} \sum_{i \in I} \sum_{j \in I} \psi_{i,j} \phi_i(x_1) \phi_j(x_2) \geq 0 \tag{10}$$

for every $(x_1, x_2) \in [0,1] \times [0,1]$, infinitely many constraints. We use normalized B splines to construct $\{\phi\}$, because, then, (10) reduces to a finite number of linear inequality constraints. (See Beresteanu (2007).) In addition, the exchangeability restriction is equivalent to $\psi_{i,j} = \psi_{j,i}$ for all $i, j \in I$. Let ψ be a vector of $\psi_{i,j}$ with $i \geq j$. Then, there is a matrix A such that (8) is affiliated and exchangeable if and only if

$$A\psi \geq 0 \tag{11}$$

Now, we simply put zero prior weight on the set of ψ's violating (11) to impose the affiliation.

Outline of Simulation Algorithm

For the IPV case, Guerre, Perringe, and Vuong (2000) show that random bids are also independent. Thus, we can handle a one dimensional sample size. We estimate the bin probabilities using the simulation algorithm given by

1. draw $\tilde{v}_1, \ldots, \tilde{v}_R$ from $f(\cdot|\theta)$.
2. compute $\tilde{b}_1, \ldots, \tilde{b}_R$ by evaluating (2) at each $\tilde{v}_r$, $r = 1, \ldots, R$.
3. Then, $\hat{\pi}_d(\theta) := R^{-1} \sum_{r=1}^{R} 1 \left\{ \tilde{b}_r \in d\text{-th bin} \right\}$

[10] One might want to use a copula to account for a correlation structure. In this case, the nonparametric (Bernstein) copula of Sancetta and Satchell (2004) should be employed, because we want the specification to be flexible. However, there is no convenient way to characterize the density affiliation. Even if there is, the specification would be needlessly complicated, because we need to use another flexible specification for the marginal density separately. Only when we use the Bernstein density of Petrone (1999a,1999b) for the marginal, there is a simplification, which, however, would actually lead back to (8) with a slightly different h than (9).

For the APV case, we use similar algorithm. But, we need to jointly draw $N(=2)$ dimensional valuations and to discretize N dimensional sample space. Appendix A provides the simulation algorithm for both the IPV and APV models.

2.4 Discussion

Guerre, Perrigne, and Vuong (2000) first develop the nonparametric indirect method for the IPV. They observe that the inverse of bidding function can be expressed as

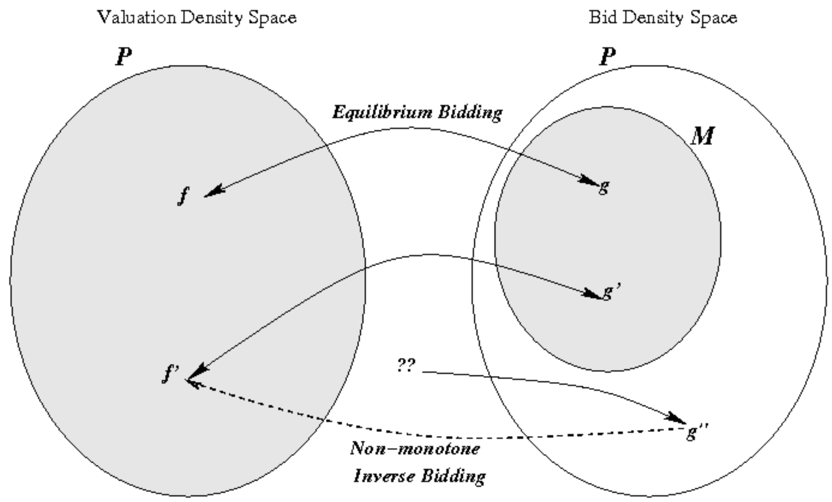
$$\beta^{-1}(b) = b + \frac{G(b)}{(N-1)g(b)} \tag{12}$$

where G and g are the marginal bid distribution and its density.[11] Thus, if we knew G and g, we could uncover $v_{i,t}$ for each $b_{i,t}$ in z. Observing this, Guerre, Perrigne, and Vuong first estimate the bid distribution, say $\widehat{G}$ and $\widehat{g}$, from the observed bids and then estimate the valuation density over $\widehat{v}_{i,t} := b_{i,t} + \frac{\widehat{G}(b_{i,t})}{(N-1)\widehat{g}(b_{i,t})}$. This method is very useful, because it is fully flexible and computationally simple.

However, since they estimate $\widehat{g}$ without considering bidding monotonicity, the estimated inverse bidding may not be increasing. To see the implication of this, consider Figure 1. P denotes the set of all densities on $\Re_+$, which can be seen as a valuation density space (left panel). Let $M \subset P$ be the set of bid densities constructing a strictly monotone inverse bidding function. Guerre, Perrigne, and Vuong show that the equilibrium creates a one to one mapping between M and P. That is, any $g \in M$ can be rationalized as an equilibrium by some valuation density $f \in P$. But, a bid density outside M, say g'', cannot be an equilibrium outcome. Note that (12) would connect g'' with a density in P, say f'. But, the equilibrium links f' and some other bid density $g' \in M$.

[11] Its derivation is simple. The expected utility of bidder i bidding b_i can be expressed by $(v_i - b_i)F(\beta^{-1}(b_i))^{N-1}$. The first order condition is $-F(\beta^{-1}(b_i))^{N-1} + (v_i - b_i)(N-1)F(\beta^{-1}(b_i))^{N-2}f(\beta^{-1}(b_i))/\beta'(\beta^{-1}(b_i)) = 0$. From this, we obtain $v_i = b_i + \frac{F(\beta^{-1}(b_i))}{(N-1)f(\beta^{-1}(b_i))/\beta'(\beta^{-1}(b_i))}$. Note that $G(b) = F(\beta^{-1}(b))$ and $g(b) = f(\beta^{-1}(b))/\beta'(\beta^{-1}(b))$.

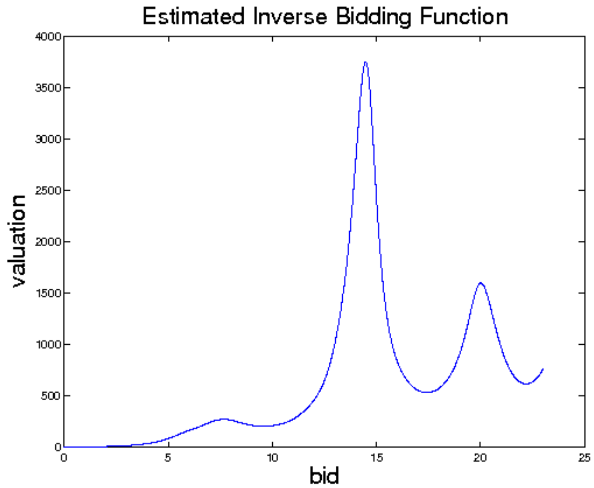
Figure 1: Mapping between Bid Density Space and Valuation Density Space (For the IPV)



Li, Perrigne, and Vuong (2002) extend this indirect method to the APV. But, they exploit neither bidding monotonicity nor density affiliation for the estimation of bid distribution. Li, Perrigne, and Vuong (2003) estimate the optimal reserve price exclusively based on the estimated bid distribution for the OCS wildcat auctions. We find that the estimated bidding function for this sample seriously violate the bidding monotonicity. See Figure 2. This suggests that the estimated bid density would be fairly different from the true one. Hence, inference based on this estimated bid density would not be accurate, and policy recommendations may not be reliable.

The idea of this paper is to rule out all the bid densities associated with a nonincreasing inverse bidding function, e.g., M^c. We find that it is very difficult to nonparametrically estimate the bid density while imposing the monotonicity of (12). Thus, we take a direct approach and use a simulated likelihood to allow for a flexible specification. Since our method satisfies

Figure 2: Nonincreasing Inverse Bidding Function for OCS wildcat auctions



all the shape conditions, it would provide more precise inference. However, we consider that our method complements the indirect approach because it is not fully nonparametric and imposes more computational costs. Note that our method may be particularly useful for small samples, because the contribution of the additional theoretical shape restrictions would be significant.

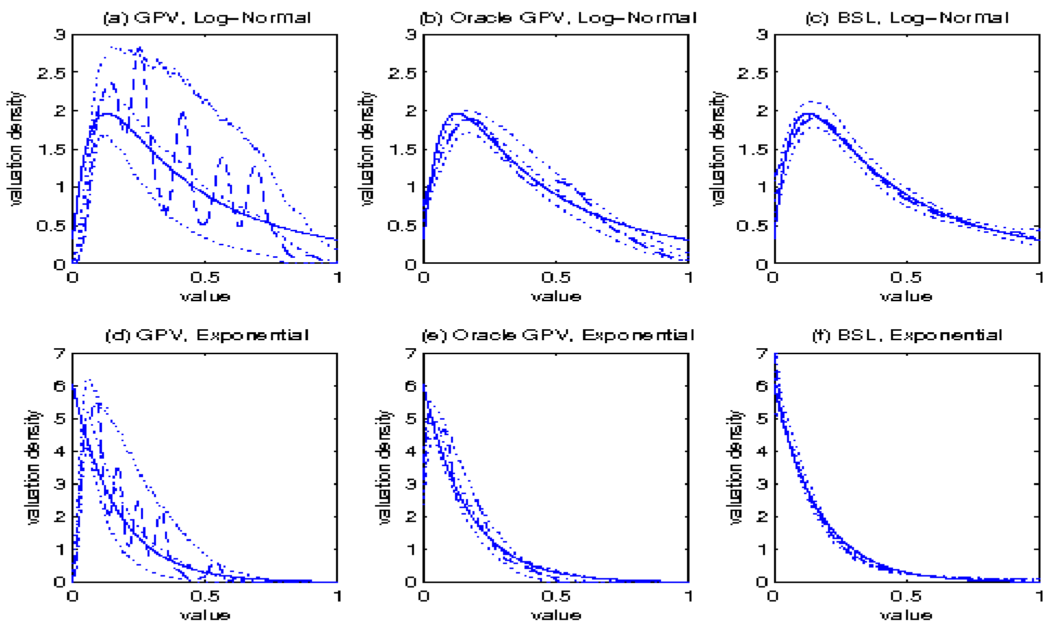
3 Monte Carlo

We compare our method (BSL) with Guerre, Perrigne, and Vuong (2000) (GPV). For BSL we employ Legendre polynomials for $\{\phi\}$ in (7) and use priors on parameters to control smoothness of the valuation density.[12] We report the predictive density estimate $E[f(\cdot|\theta)|y]$.[13] For GPV we use the rule of thumb driven by Guerre, Perrigne, and Vuong (2000) to choose bandwidths. In addition, since one can use different bandwidths, we also use the bandwidths

[12] See Appendix B.1 for Legendre polynomials.

[13] Alternatively, we could use $f(\cdot|\widehat{\theta}_B)$ with $\widehat{\theta}_B = E[\theta|y]$. We find this gives only slightly different estimates.

Figure 3: Monte Carlo Results for Log-Normal and Exponential

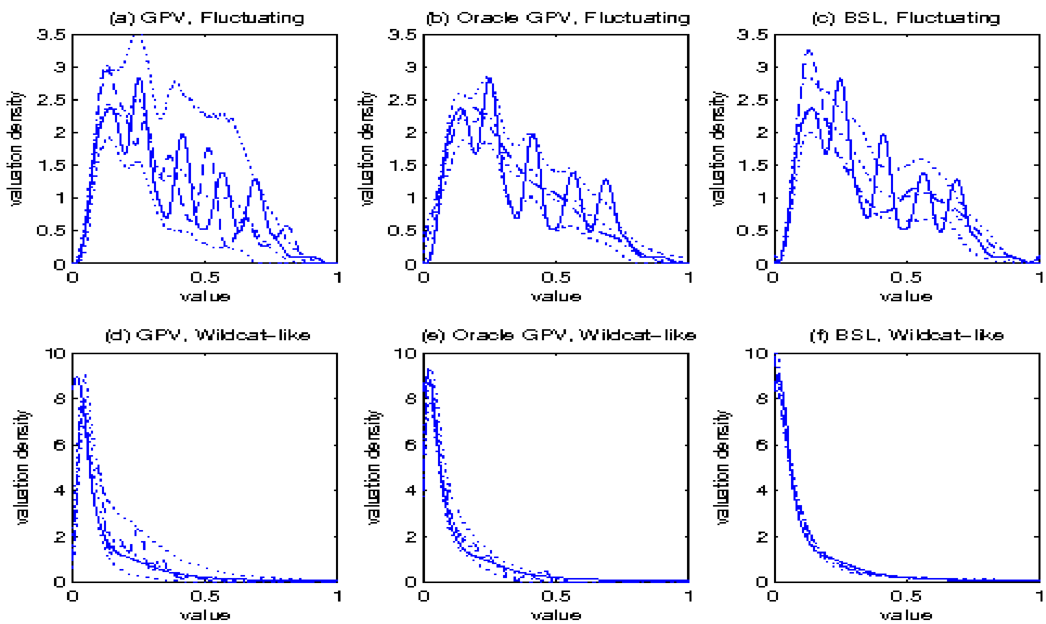


minimizing the mean integrated squared error (MISE).[14] We name this procedure 'Oracle GPV.' Note that Oracle GPV is infeasible in the real world because the true valuation density is unknown. For a fixed valuation density, we employ 1,000 Monte Carlo replications. For each replication, we generate a new sample of size $T \times N = 200 \times 2$ (similar to the OCS wildcat data) and run BSL, GPV, and Oracle GPV.

We try four different valuation densities. First, Guerre, Perrigne, and Vuong (2000), for their Monte Carlo study, employ a truncated log-normal with parameter 0 and 1 with support $[0.055, 2.5]$. We rescale it so that its support is $[0, 1]$. We call it 'Log-Normal.' Second, we use a truncated exponential distribution with mean $1/6$ and support $[0, 1]$, denoted by 'Exponential.' Figure 3 summarizes the results for Log-Normal and Exponential. Each

[14]The MISE is a precision measure of the density estimate. Let $\hat{f}_z$ be an estimate for the true density f_0. Then, $MISE(\hat{f}_z) = \int E_z \left(\hat{f}_z(x) - f_0(x)\right)^2 dx$ which is decomposed into $\int V_z \left(\hat{f}_z(x)\right) dx + \int \left(E_z \hat{f}_z(x) - f_0(x)\right)^2 dx$. Thus, the MISE is small, only when both variance and bias are small.

Figure 4: Monte Carlo Results for Nonsmooth and Wildcat-like



panel plots the true valuation density (plain), point-wise 5-th, 95-th percentiles, and point-wise mean (dotted) along with a typical density estimate (dashed). (from the last Monte Carlo replication) Apparently, BSL is more precise than GPV, providing much narrower 90% frequency bands and better behaving near the boundaries. Strikingly, Table 1 shows that BSL has even smaller MISE than Oracle GPV. That is, GPV cannot be more precise than BSL for these valuation densities.

However, this comparison may seem to be a little unfair because the true valuation densities are very smooth and BSL imposes this smoothness, whereas GPV is designed to work well for a wide range of density functions. Thus, we introduce another valuation density for which GPV might be more accurate. We use the typical GPV estimate under Log-Normal (dashed-line on Figure 3(a)) for the third data generating process (DGP). We call this 'Nonsmooth.' GPV might perform better, because (i) BSL "incorrectly" imposes more smoothness,

Table 1: MISE Comparisons

DGP's	GPV	Oracle GPV	BSL	$\frac{\text{BSL}}{\text{GPV}} \times 100\%$
Log-Normal	0.23805	0.02955	0.01034	4.344
Exponential	0.95726	0.15037	0.02430	2.539
Nonsmooth	0.29562	0.14500	0.16283	55.080
Wildcat-like	1.04260	0.20158	0.06406	6.144

(ii) BSL would not be sufficiently flexible, and (iii) Nonsmooth is actually from GPV. The Monte Carlo results in Figure 4 show that BSL is still superior to GPV, but very similar to Oracle GPV. In fact, MISE of Oracle GPV is slightly smaller than BSL (Table 1), indicating that for some bandwidths GPV can be more precise.

The fourth DGP is the marginal valuation density estimate of the OCS wildcat auctions (Figure 6(b)).[15] Overall, the Monte Carlo results are very similar to Exponential.[16] To make a connection to the next section, we compare the expected revenues for the policy implications on reserve prices under GPV and BSL. For GPV we choose a reserve price maximizing the seller's revenue for the estimated bid distribution following Li, Perrigne, and Voung (2003) and for BSL we maximize the posterior mean of the seller's revenue following Kim (2008). Then, we compare the averages of true revenue under each policy prescription. We find that BSL produces 18.30% higher revenue than GPV for the Wildcat-like distribution. For other experiments, BSL gains revenue increase by 5.52% for Log-Normal, 8.43% for Exponential, and 3.43% for Nonsmooth. Appendix D provide more detailed discussions for the Monte Carlo studies in this section.

[15]Though this estimate comes from the APV model, we conduct this Monte Carlo under the IPV, because, otherwise, it would take too long.

[16]Figure 4 second row displays the experiment results. See also Table 1 and Table 2.

4 Estimation and Auction Design for OCS Wildcats

We apply our methodology to the OCS wildcat auctions to simulate the posterior of the valuation distribution. We provide the valuation/bid density estimates and choose a reserve price maximizing the seller's future expected revenue using the decision theoretic method introduced by Kim (2008).

4.1 Data and Sample Space Discretization

We discuss the OCS wildcat data briefly.[17] The sample consists of 217 auctions, each having two bids in 1972 dollars. It contains the auctions between 1954 and 1969 among the sales held by the U.S. federal government to sell its mineral right on oil and gas on offshore lands off the Texas and Louisiana coasts.

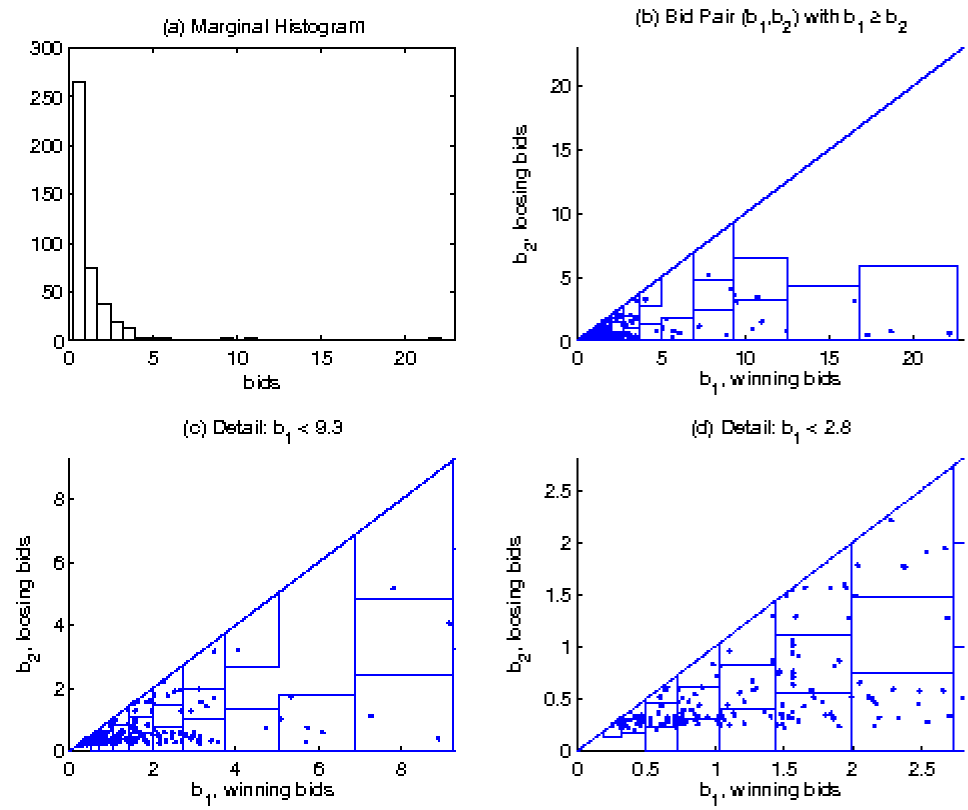
Figure 5(a) is the marginal histogram of the sample, which roughly resembles an exponential density. The sample mean and standard deviation are 1.458 and 2.557 ($\times 100$ dollars for each). Panel (b) shows how the bid pairs (b_1, b_2) with $b_1 > b_2$ are scattered. Most bids are condensed around the origin, while the sample has a long tail. The sample correlation is 0.412, but some negative correlation pattern is observed in the tail area. To examine closely, panel (c) and (d) enlarge the parts for $b_1 < 9.3$ and $b_1 < 2.8$, respectively.

We follow the assumptions that Li, Perrigne, and Vuong (2000, 2003) maintain for this sample: nonbinding reserve price, no dynamic consideration, the auction homogeneity, and the symmetric APV paradigm. Note that they explain the reason that these assumptions are reasonable for the dataset.[18] Though most of these assumptions have been taken by other researchers who use the same sample, there has been a disagreement on the APV

[17] For a thorough data description, see McAfee and Vincent (1992), Hendricks and Porter (1992), Hendricks, Pinkse, and Porter (2003), and Li, Perrigne, and Vuong (2000,2003). The dataset is publically available at `http://capcp.psu.edu/index.html`

[18] For example, they claim that the actual reserve price \$15 is too low to be an effective screening device and the game induced by the auction mechanism can be seen as symmetric, because all bidders have equal opportunity to access the same information on the auctioned tract.

Figure 5: OCS Wildcat Data and Bid Space Discretization



hypothesis. For example, Li, Perrigne, and Vuong (2000, 2003) and Campo, Perrigne, and Vuong (2003) take the APV paradigm, while Hendricks, Pinske, and Porter (2003) support the pure common value (PCV) paradigm. However, they all agree that the reality would be better explained by the affiliated value (AV) paradigm that embraces both APV and PCV as special cases.[19]

Last, to obtain an unbiased simulated likelihood, we discretize the sample space as indi-

[19]The AV paradigm is introduced by Wilson (1977).

cated in Figure 5. In panel (b) (and (c) and (d), too), for any two adjacent squared bins on the horizontal axis, the right one is 80% larger than the left one. For each squared bin, we put an equally sized bin right above it until the remaining trapezoid cannot contain another squared bin. We target to construct 30 nonempty bins.

4.2 Valuation/Bid Density Estimation

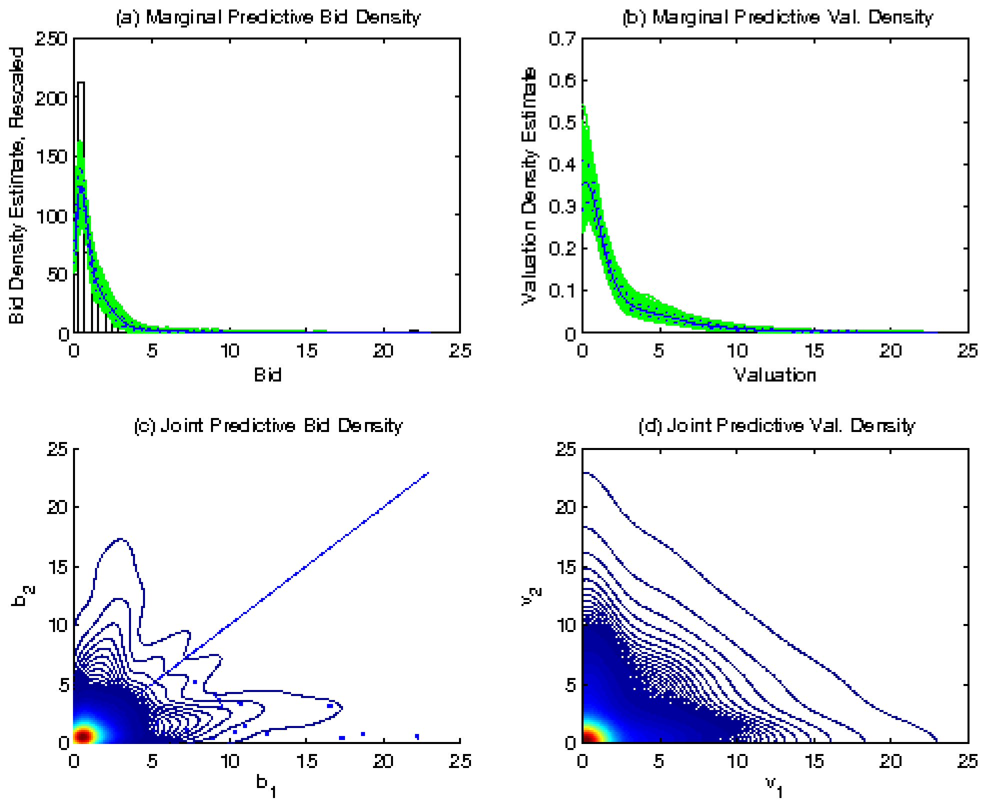
We model the valuation density using (8) with (9). In particular, to impose the affiliation restriction, we employ normalized B splines to construct $\{\phi_i\}$ with 91 components. Since the bids are exponential-like distributed, we take an exponential with hazard rate μ for $\tilde{f}(\cdot|\mu)$ in (8). We control the tail behavior of the valuatin density. (See 4.4.) The Adaptive Metropolis algorithm of Haario, Saksman, and Tamminen (2001) simulates the posterior.[20] Appendix E provides all the details for the specification and implementation.

To obtain a valuation density estimate that preserves density affiliation, we compute the posterior mean of log density, i.e., $E[\log f(v_1, v_2|\theta)|y]$. Then, we use exponential of this as our valuation density estimate. This posterior mean is consistently estimated by the AM output. In order to check the goodness-of-fit, it may be useful to construct a bid density estimate. We approximate the bid density using simulation; generate many auctions for each parameter value from the AM output $\theta_1, \ldots, \theta_S$ and run a kernel method over these simulated bids. $\widehat{g}(\cdot|\theta_s)$. Then, we use the average of these kernel estimates as our bid density estimate.

Figure 6 summarizes the estimation results. Panel (a) plots the marginal bid density estimate (solid) and a 90% credible band (dotted) along with the posterior distribution of bid density (shaded). This bid density estimate fits the data fairly well and its narrow credible band indicates high precision. But, it does not closely fit the tail because we control it. Nevertheless, panel (c), which shows 1,000 level curves of the joint density estimate, explains some negative correlation pattern over the tail area. Similarly, panel (b) plots the posterior

[20]See Appendix C for the Adaptive Metropolis algorithm.

Figure 6: Predictive Densities for OCS wildcat auctions



distribution of marginal valuation density. Since a bidder bids less than his valuation, the valuation densities spread out toward large values. The credible band is also narrow, which implies that the posterior is very condensed around its mean. Panel (d) is the 1,000 level curve contour of joint predictive valuation density.

4.3 Auction Design with Bayes Rule of Kim (2008)

This subsection uses the posterior distribution of the structural parameters to compute revenue-maximizing reserve prices. Riley and Samuelson (1981) show that when the seller's valuation is v_0, a reserve price $\rho_R(\theta) := \rho^*$ solving $\rho^* = v_0 + \frac{1-F(\rho^*|\theta)}{f(\rho^*|\theta)}$ maximizes the seller's expected revenue (payoff). Paarsch (1997) estimates an optimal reserve price by $\rho_R(\widehat{\theta})$ with a consistent estimate $\widehat{\theta}$. Similarly, Li, Perrigne, Vuong (2003) nonparametrically estimate the bid distribution and then derive an optimal price for this estimated bid distribution. This decision procedure is called a 'plug-in' rule because it makes a decision regarding an estimate as true.

However, the plug-in rule does not consider the payoff structure and parameter uncertainty. Consider a seller whose payoff increases slowly up to the optimal reserve price, but then drops sharply thereafter. For a given sample, plug-in rule must either overestimate or underestimate the true optimal price due to sampling error. Then, the seller should prefer underestimation to overestimation. Thus, we can make a higher payoff by formally considering the payoff structure as well as parameter uncertainty (that is related to sampling error).

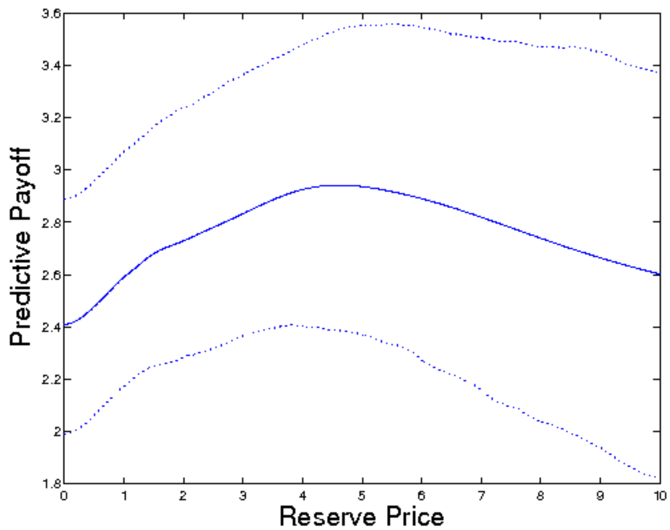
Kim (2008) introduces a Bayesian decision theoretic framework to auction design. Let $\Pi(\theta, \rho)$ denote the seller's payoff for reserve price ρ under θ and $\mathcal{A}$ be a set of all feasible reserve prices. The Bayes rule selects a Bayes action defined by

$$\rho_B(y) := \arg\max_{\rho \in \mathcal{A}} \int \Pi(\theta, \rho) p(\theta|y) d\theta \tag{13}$$

Observe that the posterior systematically quantifies parameter uncertainty and the Bayes rule considers the average payoff structure that is weighted by the posterior. Kim (2008) discusses the optimality of (13) as a decision rule. We revisit the auction design problem for the OCS wildcat auctions using the Bayes rule. Assume $v_0 = 0$ for simplicity. Then,

$$\Pi(\theta, \rho) = E\left[\beta(v_{(2)}|\rho, \theta) \cdot 1\left(v_{(2)} > \rho\right) |\theta\right] \tag{14}$$

Figure 7: Predictive Revenue for OCS wildcat auctions



where $v_{(1)} < v_{(2)}$. We consistently estimate (14) using Monte Carlo for each θ and ρ. Then, we approximate (13) using the MCMC output.

We find $\widehat{\rho}_B(y) = 4.62$. Thus, the posterior expected revenue-maximizing reserve price is \$462 per acre. We evaluate the predictive revenue at this price value; $\int \Pi(\theta, \widehat{\rho}_B(y))p(\theta|y)d\theta = 294.24$. Hence, the seller's predictive revenue for a typical tract of 5,000 acres would be \$1,471,197 ($= 294.24 \times 5,000$). We find that the predictive revenue of \$15, the actual price, is \$1,208,783. Therefore, our choice for the reserve price would increase the revenue by \$262,414. This revenue gain is significant, since hundreds of tracts are offered annually.

We also find that $\rho_R(\widehat{\theta}_B) = 4.47$ and its predictive revenue is \$1,470,872; the plug-in rule turns out to well approximate the Bayes rule. It must be the case that either the payoff is about symmetric or parameter uncertainty is negligible. In our case, the former is the answer. Figure 7 plots the predictive payoff (plain) along with its 90% credible band (dashed). the predictive revenue curve is roughly symmetric about its maximum, while parameter uncertainty is still large (wide credible band). Note that the predictive revenue for zero reserve price is very close to the sample mean of winning bids, 2.243, indicating the accuracy of our analysis. Note that our policy prescriptions are dramatically different from the estimate \$273 of Li, Perrigne, and Vuong (2003). This suggests that the contribution of the additional shape restrictions is significant.

Finally, though many researchers have pointed out that the actual price of \$15 is too low to be an effective screening device the U.S government still employs a very low reserve price. Our finding in this subsection supports the opinion of the previous researchers by suggesting an even higher reserve price.

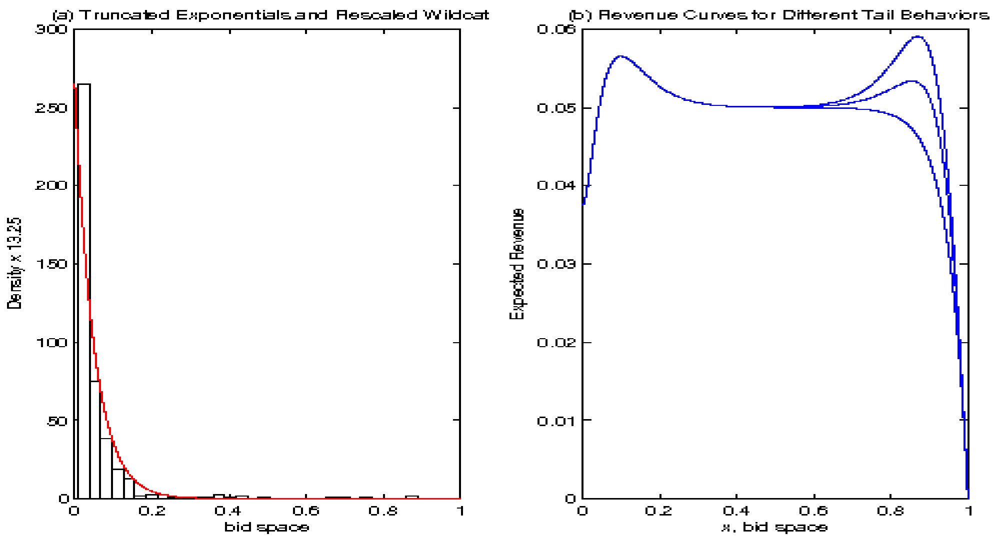
4.4 Note on discontinuity of $\rho_R(\cdot)$

We find that $\rho_R(\theta)$ can be discontinuous in the tail behavior of the underlying distribution. For clarity, consider an IPV auction with two bidders. Li, Perrigne, and Vuong (2003) show that when seller's valuation is zero, the expected revenue is given by

$$\tilde{\Pi}(x) := 2E\left[\frac{G(x)^2}{g(x)G(b_1)} \cdot 1\left\{b_1 > b_2\right\} 1\left\{b_1 > x\right\}\right] \tag{15}$$

for a point x on a bid space. Then, the optimal reserve price is given by $\rho^* = \beta^{-1}(x_0) := x_0 + G(x_0)/g(x_0)$ with x_0 maximizing (15). Now, suppose the bids are distributed as a density $g(b|c) \propto \left(0.05e^{-0.05b} + c\right) \cdot 1\left\{b \in [0,1]\right\}$. This is a variation of a truncated exponential for which we can choose the tail thickness $c > 0$. Figure 8 (a) plots three almost identical densities with different thickness $(c_1, c_2, c_3) = (0, 5 \cdot 10^{-8}, 10^{-7})$. Panel (b) displays the corresponding revenue curves, (15). They are very similar for $x < 0.6$ with a local maximum at $x_0 = 0.1$, but differ for $x > 0.6$: the second maximum grows as the tail gets thicker. in particular, the global maximum shifts from $x_0 = 0.1$ to $x_1 = 0.8703$. When we evaluate the associated reserve prices, $\rho^*(c_1) \approx \rho^*(c_2) \approx 0.4195$, but, strikingly, $\rho^*(c_3) \approx 1,534,500$! Thus, if the tail behavior is not properly controlled, an unacceptable policy suggestion can arise.

Figure 8: Discontinuity of Revenue Maximizing Reserve Price



Our discussion here is relevant to the wildcat auctions since these densities explain the data histogram fairly well. In fact, when we do not control the tail behavior, we have two local maxima; one with a reasonable value but the other with about 4 million dollars, which amounts to 20 billion dollars for a 5,000 acre tract. We consider this to be unreasonable. For this reason, we control the tail behavior using a prior.

5 Conclusion

Our methodology formally considers the additional theoretical shape restrictions arising from economic theory, such as bidding monotonicity and density affiliation. These are not fully exploited by the current nonparametric methods. We directly parametrize the valuation density allowing for a flexible specification so that bidding monotonicity is satisfied. We also restrict the parameter space so that the posterior selects only affiliated densities. We simulate the likelihood to handle such a rich specification. Since our method exploits more information, it can give more precise inference especially when the sample is small. We reanalyze the sample from the OCS auctions that Li, Perrigne, and Vuong (2003) investigate. While the density estimate fits the data very well, our choice of reserve price is drastically different from the previous results. This indicates that the additional theoretical shape restrictions play an significant role.

We develop our framework under some assumptions. Li, Perrigne, and Vuong (2003) argues that these assumptions are quite reasonable for the sample from the OCS wildcat auctions. However, in order to provide more convincing policy recommendations, it would be useful to consider more features of the OCS auctions, such as endogenous entry and multi-unit auctions. That is, there are many auctions with more than or less than just two bidders. Not all the potential bidders participate in every auction, while reserve prices are very low. If bidders make entry decisions first, the theoretical optimal reserve price will be dramatically different.[21] Relatedly, the government holds many auctions at a given date, creating another strategic situation among the bidders. Hence, it is necessary to consider these issues in order to provide more convincing policy analysis. We plan to address these issues extending our method developed in this paper. Some economists have argued that the pure common value better approximates the OCS wildcat auctions than the APV. From the policymaker's point of view, it would be important to obtain a policy advice considering the risk under all possible alternative models. Hence, we also plan to develop a decision method that incorporates alternative models into decision process.

[21]See Levin and Smith (1994) and Moreno and Wooders (2008).

Appendix

A Simulation Algorithm

A.1 Sampling Valuations and Bidding Function Evaluation for IPV

The auction simulation algorithm consists of two steps as follows. H denotes the CDF of h and $x_{(r)}$ is the r-th smallest out of $x_1, x_2, \ldots$. Under given θ,

1. Draw a random sample of size R, $\tilde{v}_1, \ldots, \tilde{v}_R$.

   Let $\{\tilde{u}_r\}_{r=1}^{R}$ be uniform draws. Then, $\tilde{v}_r := F^{-1}(\tilde{u}_r|\theta) = \tilde{F}^{-1}(H^{-1}(\tilde{u}_r|\psi)|\mu)$. To approximate $H^{-1}(\tilde{u}_r|\psi)$, we evaluate $H(\cdot|\psi)$ on each point in $\{0, 0.01, \ldots, 0.99, 1\}$ and apply a monotonicity preserving interpolation.[22] Most statistical softwares can evaluate $\tilde{F}^{-1}(\cdot|\mu)$ numerically.

2. Compute the equilibrium bids $\tilde{b}_1, \ldots, \tilde{b}_R$ using (2).

   Note that $F(\tilde{v}_r|\theta) = \tilde{u}_r$. Hence, a trapezoidal rule over $\left((0,0), \left\{\left(\tilde{v}_{(r)}, \tilde{u}_{(r)}^{N-1}\right)\right\}_{r=1}^{R}\right)$ solves $\int_0^{\tilde{v}_r} F(\alpha|\theta)^{N-1} d\alpha$. That is,

$$\int_0^{\tilde{v}_{(r)}} F(\alpha|\theta)^{N-1} d\alpha \approx \int_0^{\tilde{v}_{(r-1)}} F(\alpha|\theta)^{N-1} d\alpha + \frac{1}{2}\left(\tilde{u}_{(r-1)}^{N-1} + \tilde{u}_{(r)}^{N-1}\right)\left(\tilde{v}_{(r)} - \tilde{v}_{(r-1)}\right)$$

   starting from $\int_0^{\tilde{v}_{(1)}} F(\alpha|\theta)^{N-1} d\alpha \approx \frac{1}{2}\tilde{u}_{(1)}^{N-1}\tilde{v}_{(1)}$.

This trapezoidal rule is fairly accurate. For the part where F quickly increases, we may have many reference points because they are random sample from F. For the part where F is almost flat, we may not have many reference points, because the probability density is small. However, the area to be integrated is almost a rectangle.

[22] For example, a piecewise linear function interpolation or the piecewise cubic interpolation method of Fritsch and Carlson (1980).

A.2 Sampling for APV: $(\tilde{v}_{1,1}, \tilde{v}_{2,1}), \ldots, (\tilde{v}_{1,R}, \tilde{v}_{2,R}) \sim f(\cdot|\theta)$

We employ an accept/reject method.[23] In particular, we use a piecewise uniform density mimicking (9) for the source function. That is, we use the kernel of (9) given by $k_h(x_1, x_2|\psi) := \exp\left\{\sum_{i\in I}\sum_{j\in I}\psi_{i,j}\phi_i(x_1)\phi_j(x_2)\right\}$. Then, construct the kernel for the piecewise uniform as follows

$$k_{pu}(x_1, x_2|\psi) = k_h\left(\frac{i}{10}, \frac{j}{10}\middle|\psi\right) + k_h\left(\frac{i}{10}, \frac{j+1}{10}\middle|\psi\right) + k_h\left(\frac{i+1}{10}, \frac{j}{10}\middle|\psi\right) + k_h\left(\frac{i+1}{10}, \frac{j+1}{10}\middle|\psi\right)$$

for $(x_1, x_2) \in [\frac{i}{10}, \frac{i+1}{10}] \times [\frac{j}{10}, \frac{j+1}{10}]$ $i, j \in \{0, 1, 2, \ldots, 9\}$ Then, the accept/reject algorithm is as follows. For $r = 1, \ldots, R$.

1. Draw a proposal $(\tilde{u}_1^*, \tilde{u}_2^*)$ from the density proportional to k_{pu}
2. Let $(\tilde{u}_{1,r}, \tilde{u}_{2,r}) = (\tilde{u}_1^*, \tilde{u}_2^*)$ with probability $\frac{k_h(\tilde{u}_1^*, \tilde{u}_2^*|\psi)}{Qk_{pu}(\tilde{u}_1^*, \tilde{u}_2^*|\psi)}$ with $Q \geq \sup_{(u_1,u_2)\in[0,1]^2} \frac{k_h(u_1,u_2|\psi)}{k_{pu}(u_1,u_2|\psi)}$
3. If the proposal is not accepted, go back to step 1.

Once $(\tilde{u}_{1,1}, \tilde{u}_{2,1}), \ldots, (\tilde{u}_{1,R}, \tilde{u}_{2,R})$ are drawn, then $(\tilde{v}_{1,1}, \tilde{v}_{2,1}), \ldots, (\tilde{v}_{1,R}, \tilde{v}_{2,R})$ are obtained by $\tilde{v}_{i,r} = \tilde{F}^{-1}(\tilde{u}_{i,r}|\mu)$ for each $i = 1, 2$ and $r = 1, \ldots, R$.

A.3 Evaluation of Equilibrium Bidding Function for APV

We compute the equilibrium bids $\left\{(\tilde{b}_{1,r}, \tilde{b}_{2,r})\right\}_{r=1}^{R}$ using (1). It may be quite time consuming to evaluate (1), since it requires us to estimate many triple integrals. However, we find that the following recursive algorithm reduces the computing time, significantly, which has three steps.

[23] The inverse CDF not only runs slowly in a multivariate case but also gives no help to computing (1) as in the IPV.

1. We compute $a_1, \ldots, a_{2R}$ with

$$a_j := \frac{f(\tilde{v}_{(j)}, \tilde{v}_{(j)}|\theta)}{\int_0^{\tilde{v}_{(j)}} f(\tilde{v}_{(j)}, t|\theta)dt}$$

   for $j = 1, \ldots, 2R$ where $\tilde{v}_{(j)}$ is the j-th smallest out of the $R \times 2$ simulated random values. Note that the integral on the denominator is quickly estimated by the Gaussian quadrature. [24]

2. Constuct

$$A_{i,j} := \int_{\tilde{v}_{(i)}}^{\tilde{v}_{(j)}} \frac{f(u,u|\theta)}{\int_0^u f(u,t|\theta)dt} du \approx A_{i,j-1} + \frac{1}{2}\left(a_j + a_{j-1}\right)\left(\tilde{v}_{(j)} - \tilde{v}_{(j-1)}\right)$$

   for $i = 1, \ldots, 2R$ and $j = i+1, \ldots, 2R$ with $A_{i,i+1} \approx \frac{1}{2}\left(a_i + a_{i+1}\right)\left(\tilde{v}_{(i+1)} - \tilde{v}_{(i)}\right)$.

3. $L(\alpha|v) := \exp\left\{-\int_\alpha^v \frac{f_{y_1|v_1}(u|u)}{\int_0^u f_{y_1|v_1}(t|u)dt} du\right\}$ is approximated by

$$\int_0^{\tilde{v}_{(j)}} L\left(\alpha|\tilde{v}_{(j)};\theta\right) d\alpha \approx \frac{1}{2}\sum_{i=1}^{j-1}(\exp\left(-A_{i,j}\right) + \exp\left(-A_{i+1,j}\right))(\tilde{v}_{i+1} - \tilde{v}_i)$$

The last two steps amount to a trapezoidal rule approximation with reference points $\tilde{v}_{(1)}, \ldots, \tilde{v}_{(2R)}$.

A.4 Alternative Method to Estimate Likelihoods

This bidding function evaluation is still computationally expensive. We use an alternative to estimate $\hat{\pi}_d(\theta)$ on the implied valuation space discretization. First, we approximate $\beta^{-1}(\cdot|\theta)$ using a monotonicity preserving interpolation: evaluate (1) at every percentile of

[24] The upper limit of the integral can be a very large number and, in that case, the Gaussian quadrature approximation with finite points may be poor. However, the integral can be expressed as $\int_0^{\tilde{F}(\tilde{v}_{(j)}|\mu)} h(\tilde{F}(\tilde{v}_{(j)}|\mu), s|\psi)ds$ for which the upper limit is always less than one. Then, for this integral, we find that the Gaussian quadrature even with three points provides fairly accurate estimates. Moreover, $\tilde{F}$ should not be evaluated, because we already have $\tilde{u}_{i,r} = \tilde{F}(\tilde{v}_{i,r}|\mu)$ $i = 1, 2,$ $r = 1, \ldots, R$.

$\left\{\tilde{v}_{(1)},\ldots,\tilde{v}_{(2R)}\right\}$ and find a strictly increasing function connecting $\left\{\left(\tilde{b}_{p\text{-th }\%},\tilde{v}_{p\text{-th }\%}\right)\right\}_{p=1}^{100}$. Second, we construct the valuation space discretization by evaluating this function at each corner point of the bins in the sample space. Then, we estimate $\hat{\pi}_d(\theta)$ by counting the simulated valuations in the d-th bin in the valuation space. Since this procedure evaluates (1) only 100 times regardless of R, we may use a large R.

B Basis Functions and Prior Specification

B.1 Legendre Polynomials and Smoothness Control

We employ the Lengendre polynomials, i.e., $\phi_j(u) := \sqrt{2j+1}\tilde{\phi}_j(2u-1)$ with $\tilde{\phi}_j(x) = \frac{d^j}{dx^j}(x^2-1)^j/(2^j j!)$. Note that the order of ϕ_j gets higher as j increases. Hence, to control the smoothness, we use the prior such that $\psi_j \sim N(0,\tau/2^j)$ for $\tau > 0$. Then, ψ_j is more condensed around zero for higher j and the posterior picks a smooth density more likely. We use Unif[0,1] for $\tilde{F}$ just for simplicity.[25]

B.2 Normalized B splines

We construct the basis functions as follows:

$$\left\{\phi_i(x) := K\left(\frac{x-i/k}{1/k}\right)\right\}_{i\in I} \tag{16}$$

where $K(x) := \sum_{j=0}^{l} \frac{(-1)^j}{(l-1)!}\left(\frac{l!}{j!(l-j)!}\right)\left(x+\frac{l}{2}-j\right)_+^{l-1} 1\{|x|<l/2\}$ with an integer $l>1$ and $x_+^a = x^a 1\{x>0\}$. K is a kernel symmetric about zero and $l-1$ times differentiable with support $[-\frac{l}{2},\frac{l}{2}]$. In addition, I is the set of all integers in $[-m,k+m]$, with an integer $k>l$ and $m := (\text{floor}(l/2) - 1\{l \text{ is even}\})$. ($|I| = 2m+k+1$, cardinality of I.) Then, $\{\phi\}$ are centered on equidistant grid points $\left\{\frac{-m}{k},\frac{-m+1}{k},\ldots,\frac{k+m}{k}\right\}$. Note that (16) are $l-2$ times differentiable and the $l-1$-th derivative does not exist at the grid point. Some are located

[25] Hence, the properly selected $\tilde{F}$ may result in smaller MISE's than presented below.

outside $[0,1]$, since every point in $[0,1]$ must have equal number of nonzero bases for the affiliation condition below.[26]

B.3 Affiliation Restrictions with Normalized B splines

Affiliation is equivalent to supermodularity of the log-density. $\kappa_\psi(x_1,x_2) := \sum_{i\in I}\sum_{j\in I}\psi_{i,j}\phi_i(x_1)\phi_j(x_2)$ for simplicity. Then, (8) is log-supermodular if and only if $\frac{\partial^2}{\partial x_1 \partial x_2}\kappa_\psi(x_1,x_2) \geq 0$ for all $(x_1,x_2) \in [0,1]\times[0,1]$, which requires infinitely many restrictions. However, Beresteanu (2007), using (16), characterizes supermodularity of the log-density by

$$\kappa_\psi\left(\frac{i}{k},\frac{j}{k}\right) + \kappa_\psi\left(\frac{i+1}{k},\frac{j+1}{k}\right) \geq \kappa_\psi\left(\frac{i+1}{k},\frac{j}{k}\right) + \kappa_\psi\left(\frac{i}{k},\frac{j+1}{k}\right)$$

for all $i,j \in \{0,1,\ldots,k\}$, which is now k^2 linear inequalities. Furthermore, the exchangeability reduces this to $\frac{k(k+1)}{2}$, because only the ones associated with grid points below (or above) the 45 degree line are relevant due to $\kappa_\psi(x_1,x_2) = \kappa_\psi(x_2,x_1)$. Recall that ψ denotes the vector of $\psi_{i,j}$ with $i \geq j$. Hence, it is $\frac{|I|(|I|+1)}{2}$ dimensional. Therefore, A in (11) is $\frac{k(k+1)}{2} \times \frac{|I|(|I|+1)}{2}$.

B.4 Tail Behavior

The discontinuity of optimal reserve price with respect to the tail behavior leads us to construct a prior to control the tail behavior. First of all, we assume $p(\mu,\psi) = p(\mu)p(\psi)$ for simplicity. Let $N(a,b)$ denote a normal density with mean vector a and covariance b. For the AM algorithm to converge, the posterior must have a bounded support. Hence, we use a prior whose supports are given by $\mu \in (0,30]$ and $\max(|\psi|) < 30$. In addition, the prior

[26] For example, ϕ_{-m} is the most left located with a positive tail at zero. That is, $\phi_{-m-1}(0) = 0$, if $-m-1 \in I$. Similarly, $\phi_{k+m}(1) > 0$, while $\phi_{k+m+1}(1) = 0$.

should be zero for ψ violating (11). Then, we employ the prior in the form of

$$
\begin{aligned}
p(\mu) &\propto N(25,1) \cdot 1\{\mu \in (0,30]\} \\
p(\psi) &\propto N(0,\Sigma_\psi) \cdot 1\{A\psi \geq 0 \text{ and } \max(|\psi|) < 30\}
\end{aligned}
$$

Now, recall that $\phi_j(\cdot)$ given in (16) is centered around $\frac{j}{k}$. Hence, $\psi_{i,j}$ is related with the shape of (9) around $\left(\frac{i}{k},\frac{j}{k}\right)$. Let $d_{i,j} := 10 \cdot \max\left\{\left((\frac{i}{k})^2 + (\frac{i}{k})^2\right)^{\frac{3}{2}}, 0.02\right\}$, which is increasing in the distance of $\left(\frac{i}{k},\frac{j}{k}\right)$ from the origin. We employ a diagonal matrix Σ_ψ whose element associated with $\psi_{i,j}$ equals $d_{i,j}^{-1}$. Then, this prior puts a smaller variance on $\psi_{i,j}$ located further from the origin so that the posterior selects a valuation density whose tail resembles the exponential $\tilde{f}$.

C The Adaptive Metropolis Algorithm

For a given prior $p(\theta)$, we can simulate the posterior using a Metropolis-Hastings algorithm. However, since choosing a good proposal density is hard for a high dimensional θ, we use an adaptive MCMC algorithm. in particular, we employ the Adaptive Metropolis (AM) algorithm of Haario, Saksman, and Tamminen (2001). Suppose θ is d dimensional. Let I_d be the $d \times d$ identity matrix and $s_d := (2.38)^2/d$. At each s-th iteration, we draw a proposal $\tilde{\theta}$ from $N(\theta_{s-1}, \Omega_{s-1})$ where, for a small number $\varepsilon > 0$ and a prespecified initial periods s_0,

$$
\Omega_{s-1} := \begin{cases} \Omega_0 & \text{if } s \leq s_0 \\ s_d \widehat{\text{Cov}}(\theta_0, \theta_1, \ldots, \theta_{s-1}) + s_d \varepsilon I_d & \text{otherwise} \end{cases}
$$

with Ω_0 and $\widehat{\text{Cov}}$ denoting some initial covariance and the sample covariance, respectively. Then, $\theta_s := \tilde{\theta} \cdot 1\{u < \alpha\} + \theta_{s-1} \cdot 1\{u > \alpha\}$ with $u \sim \text{Unif}[0,1]$ and

$$
\alpha := \min\left\{\frac{p(\tilde{\theta})\hat{L}(\tilde{\theta}|y)}{p(\theta_{s-1})\hat{L}(\theta_{s-1}|y)}, 1\right\} \tag{17}
$$

Haario, Saksman, and Tamminen (2001) show that the AM algorithm converges to the correct posterior for any θ_0 with $p(\theta_0) > 0$ provided that the posterior is bounded from above and has a bounded support. They also note that we can update Ω_s for any increasing subset of $\{\theta_s\}$

Sampling ψ under $A\psi > 0$

Hajivassiliou and McFadden (1990) and Keane (1990) develop a method, which is called the GHK sampler, to draw a random vector from a truncated normal distribution. Then, Geweke (1995) draws a random vector under a set of linear inequality restrictions using the GHK sampler. As noted earlier, we draw a candidate under the affiliation restrictions (11) in order to obtain a proposal function that generates an acceptable candidate more frequently. For this purpose, we use Geweke (1995) as follows. Let $\bar{A}$ denote a $\frac{|I|(|I|+1)}{2}$ dimensional invertable matrix that contains A in (11) as its first $\frac{k(k+1)}{2}$ rows. Then, (11) can be written as

$$\bar{A}\psi \geq \bar{a} \tag{18}$$

for which first $\frac{k(k+1)}{2}$ elements of $\bar{a}$ are zeros and the others are all $-\infty$. Therefore, the additional restrictions brought by $\bar{A}$ are not relevant.

Now, let ψ be the current parameter of the MCMC sequence. Then, we draw a candidate $\tilde{\psi}$ for the next iteration under the linear inequality constraints:

$$\tilde{\psi} \sim N(\psi, \Omega)1(\bar{A}\psi \geq \bar{a})$$

Let $u := \tilde{\psi} - \psi$. Then,

$$\bar{A}u = \bar{A}\tilde{\psi} - \bar{A}\psi \sim N(0, \bar{A}\Omega\bar{A}')1(\bar{A}u \geq \bar{a}^* := \bar{a} - \bar{A}\psi)$$

Let C be the cholesky decomposition of $\bar{A}\Omega\bar{A}'$. Then, it can be shown that $\bar{A}u = C\varepsilon$ with

$\varepsilon \sim N(0, I)1(\varepsilon \geq \underline{\varepsilon})$. The lower bounds are given by

$$\underline{\varepsilon}_j = c_{j,j}^{-1} \left(a_j^* - \sum_{i=1}^{j-1} c_{j,i} \varepsilon_i \right)$$

for $j > 1$ and $\underline{\varepsilon}_1 = c_{1,1}^{-1} \bar{a}_1^*$ where $c_{i,j}$ is the (i, j) element of C. We draw ε's recursively from the truncated standard normal distribution. Then, $\tilde{\psi} = \psi + \bar{A}^{-1} C \varepsilon$.

D More Discussion on Monte Carlo Studies

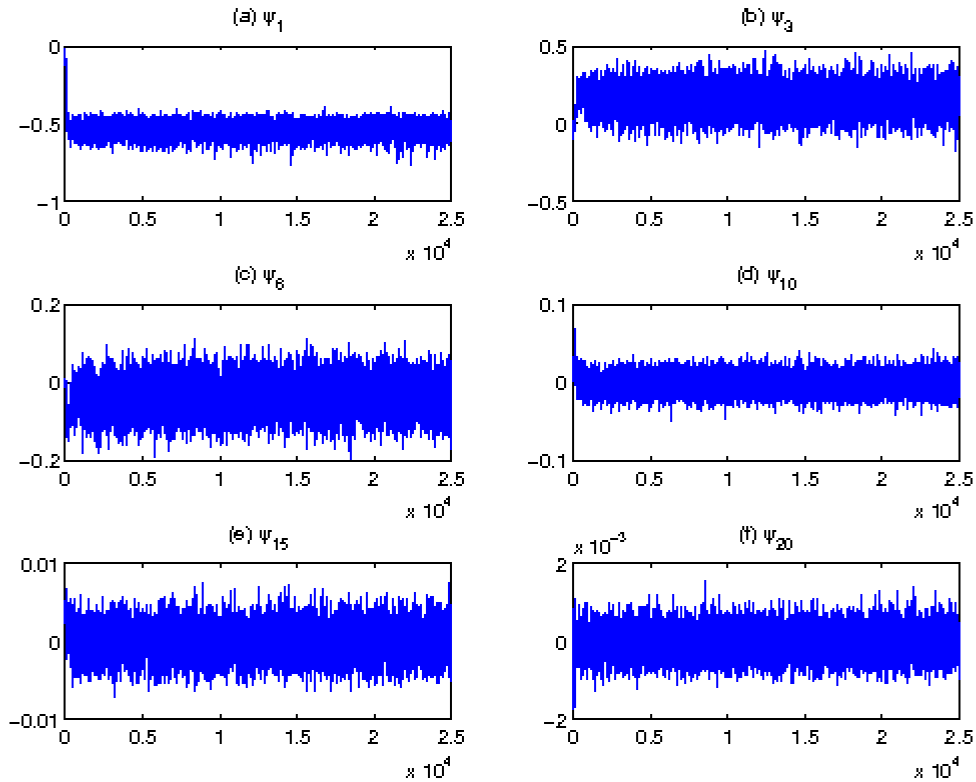
D.1 Implementation of BSL

For $\tilde{f}$ we use the uniform $[0, 1]$ and employ Legendre polynomials to construct $\{\phi\}$. Especially, we use 20 components. Then, the AM algorithm iterates 500,000 times and we choose every 20 parameters from the last 200,000 iterations to compute the predictive density $\hat{f}_{BSL}(v|y) := E_\theta[f(v|\theta)|y]$ and choose the reserve price maximizing the future revenue. Figure 9 to 12 plots some AM outputs for each Monte Carlos studies. (We record every 20 iterations.) For each iteration, $R = 10,000$ bids are simulated. The bid space is discretized into twenty bins with equal size.

For each data generating process, we use 1,000 Monte Carlo replications. If we run the BSL for each replication separately, it would be very time consuming. In this reason, we implement our method only for the first replication. Then, we estimate the predictive valuation density using an importance sampling method.

Specifically, the predictive valuation density estimate for j-th replication is estimated by $\hat{f}_{BSL}(v|y^j) := \frac{\sum_{s=1}^S w_j(\theta_s) f(v|\theta_s)}{\sum_{s=1}^S w_j(\theta_s)}$ where S is the number of random parameters drawn from the posterior of the first Monte Carlo replication and $w_j(\theta)$ is the ratio of target density (the j-th posterior) and the source density (the 1st posterior). That is, $w_j(\theta) = \frac{p(u,\theta|y^j)}{p(u,\theta|y^1)} = \frac{p(\theta)\hat{L}(\theta|y^j)}{p(\theta)\hat{L}(\theta|y^1)} = \prod_{d=1}^D \{\hat{p}_d(\theta)\}^{y_d^j - y_d^1}$ and y^j is the histogram implied by the discretization and the j-th sample for $j = 2, \ldots, 1,000$. Similarly, we conduct the Bayesian decision method to compute the

Figure 9: The AM outputs for Log-Normal

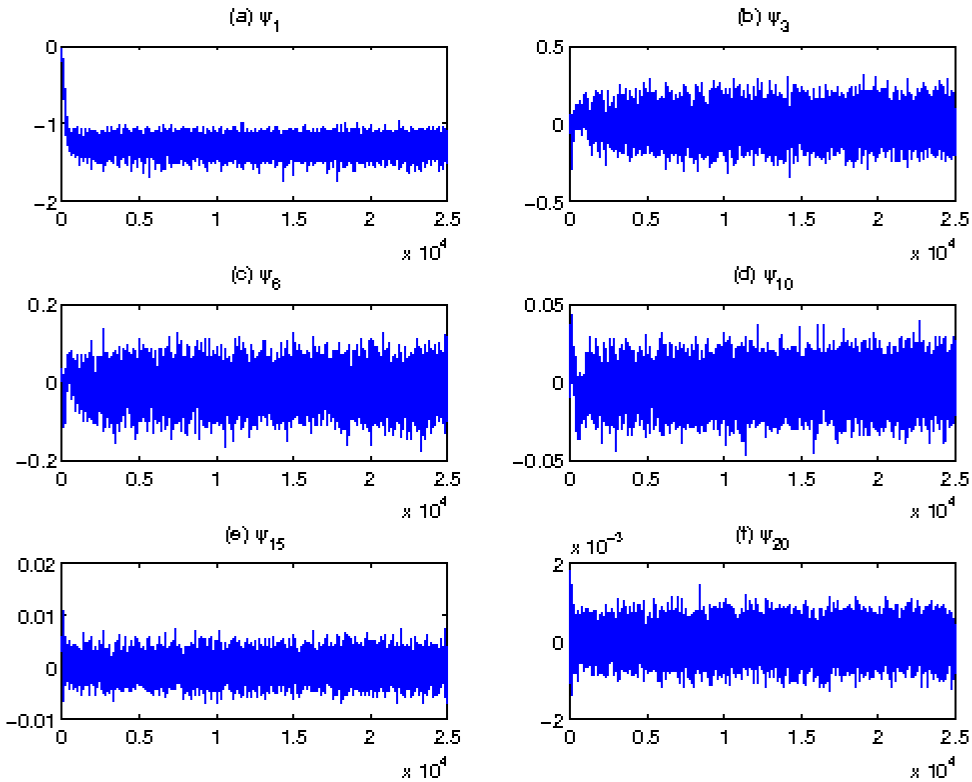


reserve price for last 999 replications.

D.2 Specification for GPV

We employ the estimation procedure that Guerre, Perrigne, and Vuong (2000) use for their Monte Carlo study. They take the triweight kernel $K(u) := \frac{35}{32}(1-u^2)^3 \cdot 1\{|u|<1\}$. Then, the bid density is given by $\widehat{g}(b|h_g) := \frac{1}{h_g TN} K\left(\frac{b-b_{i,t}}{h_g}\right)$ using the bandwidth $h_g = 1.06 \cdot \mathrm{stdv}(z) \cdot (TN)^{-1/5}$. The pseudo values are computed by $\hat{v}_{i,t} = \widehat{\beta^{-1}}(b_{i,t}) = b_{i,t} + \frac{\widehat{G}(b_{i,t})}{(N-1)\widehat{g}(b_{i,t})}$

Figure 10: The AM outputs for Exponential

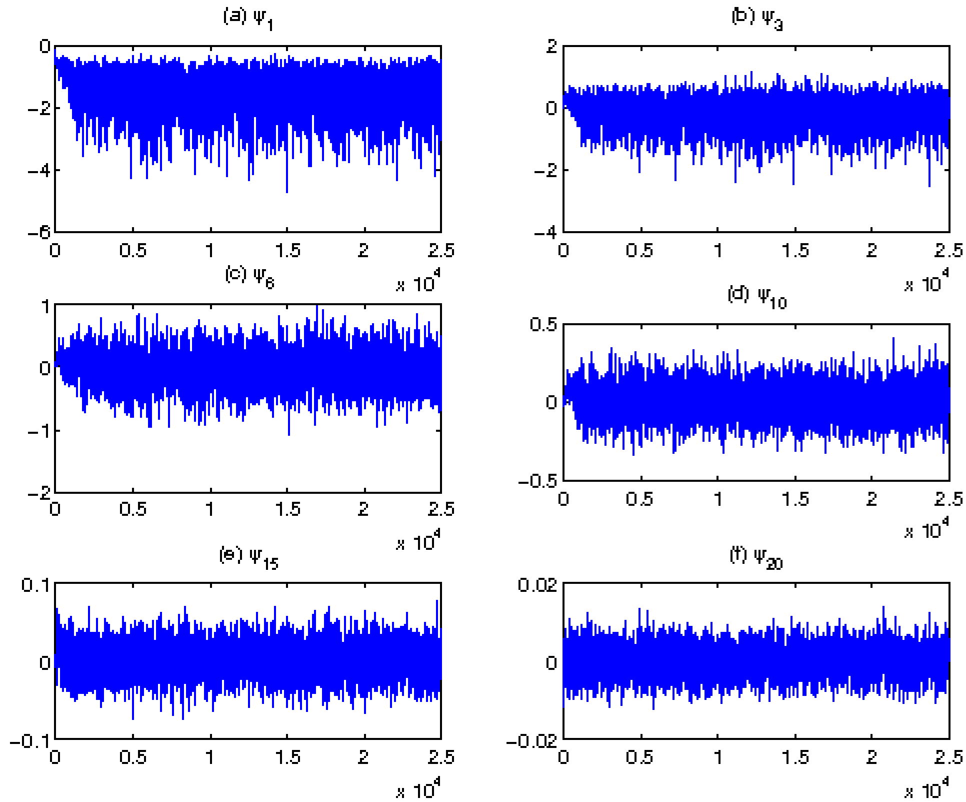


only for $b_{i,t} \in [\min(z) + h_g, \max z - h_g]$. Finally, they estimate the valuation density using $\widehat{f}(v|h_g, h_f) := \frac{1}{h_f|\{\hat{v}\}|} K\left(\frac{v-\hat{v}_{i,t}}{h_f}\right)$ with $h_f = 1.06 \cdot \text{stdv}(\{\hat{v}\}) \cdot (|\{\hat{v}\}|)^{-1/5}$.

D.3 Specification for Oracle GPV

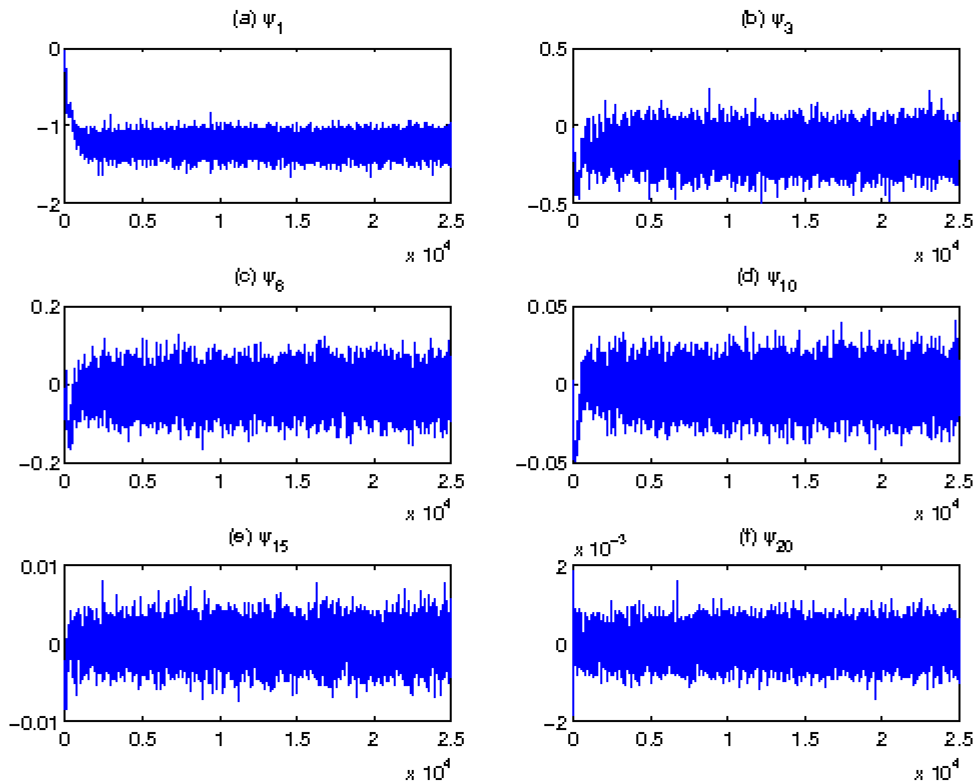
The bandwidths for Oracle GPV are given by $(h_g^*, h_f^*) := \arg\min_{(h_g,h_f)} \int_0^1 \left\{\widehat{f}(x|h_g, h_f) - f(x)\right\}^2 dx$. We run a grid search to solve this optimization problem. (h_g^*, h_f^*) does not imply a monotone inverse bidding function. Note that the smallest bandwidth for strictly monotone $\widehat{\beta^{-1}}$ is

Figure 11: The AM outputs for Nonsmooth



typically too large. Therefore, it leads to too flat bid density estimate which does not fit the data well. Instead, Oracle GPV chooses a very small h_g^* and a large h_f^*: this h_g^* implies a very noisy $\widehat{\beta^{-1}}$ which covers the true inverse bidding function so that the pseudo values from this $\widehat{\beta^{-1}}$ look roughly distributed as the true DGP. Then, the large h_f^* provides a smooth density estimate over these pseudo values.

Figure 12: The AM outputs for Wildcat-like



D.4 Information Loss of GPV

We summarize the loss of information from GPV in Table 2. The first column is the fraction of the sample points trimmed out the boundary problem of the kernel method. The second column is the fraction of data points whose order is distorted due to the nonincreasing inverse bidding function. The third column is simply sum of the first two columns.

Table 2: Information Loss of GPV (%)

DGP's	Triming Rate (boundaries)	Distortion Rate (Monotonicity Violation)	Total
Log-Normal	10.918	13.993	24.911
Exponential	18.420	15.503	33.923
Nonsmooth	7.567	14.772	22.339
Wildcat-like	18.883	14.404	33.288

E Appendix to Wildcat Auction Analysis

E.1 Specification for $f(\cdot|\theta)$

We employ $(l,k) = (4,10)$. Then, ψ has 91 components and (9) is twice differentiable.[27] This is differentiable enough for the affiliation condition to be written as (10). We judge this specification to be sufficiently flexible and computationally practical. We iterate the AM algorithm 100,000,000 times and throw away first 60,000,000 of them. Then, we take every 4,000 iterations to make the posterior inference. Hence, $S = 10,000$. Figure 12 plots the AM algorithm for some parameters. (We record every 1,000 iteration.)

E.2 Computation

Now, we discuss the computation of (13). We consistently estimate (14) using

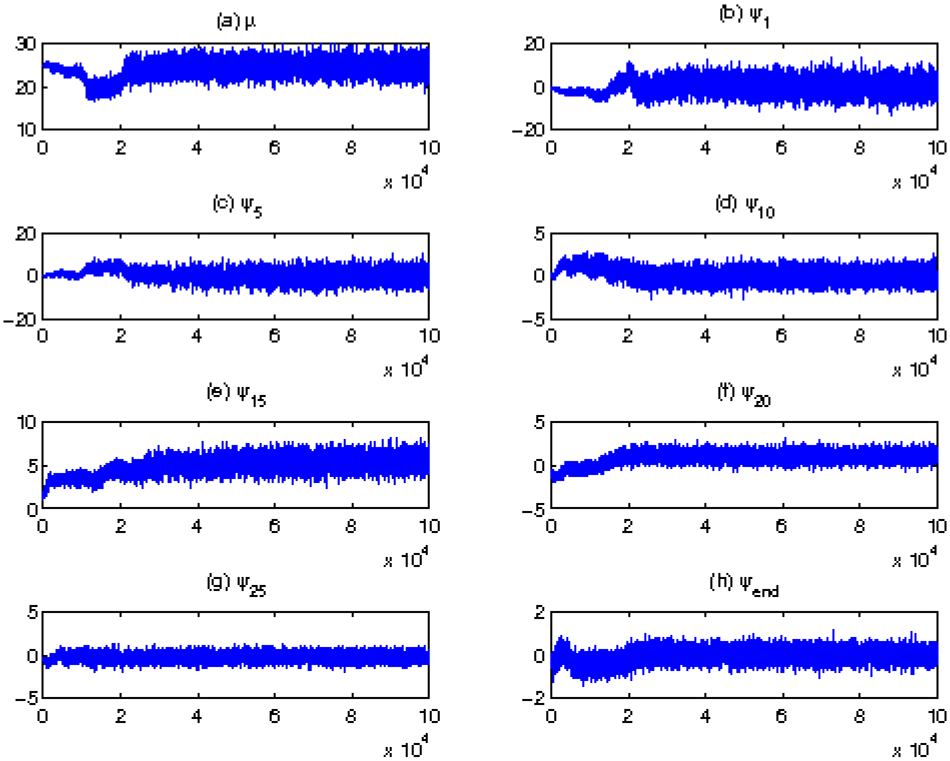
$$\widehat{\Pi}(\theta,\rho) := \frac{1}{\tilde{R}} \sum_{r=1}^{\tilde{R}} \left\{ \beta(\tilde{v}_{(2),r}|\rho,\theta) \cdot 1\left(\tilde{v}_{(2),r} > \rho\right) \right\} \tag{19}$$

where $\{(\tilde{v}_{1,r}, \tilde{v}_{2,r})\}_{r=1}^{\tilde{R}} \sim f(\cdot,\cdot|\theta)$. Hence, we approximate the Bayes action (13) with

$$\widehat{\rho}_B(y) := \arg\max_{\rho\in\mathcal{A}} \frac{1}{S} \sum_{s=1}^{S} \widehat{\Pi}(\theta_s,\rho) \tag{20}$$

[27] $|\psi| = \frac{|I|(|I|+1)}{2} = 91$ for which $|I| = 2m + k + 1 = 2(1) + 10 + 1 = 13$. The l-th derivative does not exist at each grid point $\frac{i}{k}$ for $i \in I$.

Figure 13: The AM outputs for Wildcat-like



where $\{\theta_s\}_{s=1}^{S}$ are the MCMC output.

Note that for the AM algorithm we have used new simulation draws at each iteration to obtain the unbiased likelihood estimates. However, we use a fixed uniform draw $\{(\tilde{u}_{1,r}, \tilde{u}_{2,r}\}_{r=1}^{\tilde{R}}$ to compute (19) for each ρ and θ, because, otherwise, the sample mean in (20) would not be smooth. Now, to evaluate (20), we employ the algorithm as follows. For a fixed θ

1. Compute $\{(\tilde{v}_{1,r}, \tilde{v}_{2,r})\}_{r=1}^{\tilde{R}}$ using an inverse CDF method:
   $\tilde{v}_{1,r} = F^{-1}(\tilde{u}_{1,r}|\theta)$ and $\tilde{v}_{2,r} = F^{-1}(\tilde{u}_{2,r}|\tilde{v}_{1,r}, \theta)$ with the marginal CDF, $F(v_1|\theta) :=$

$\int_0^{v_1}\int_0^{\infty} f(s,t|\theta)dtds$, and the conditional CDF, $F(v_2|v_1,\theta) := \int_0^{v_2} \frac{f(v_1,s|\theta)}{\int_0^{\infty} f(v_1,t|\theta)dt}ds$.

2. Compute $\left\{\beta(\tilde{v}_{(2),r}|\theta)\right\}_{r=1}^{\tilde{R}}$ using the recursive method we have employed for the AM algorithm.

   Note that $\tilde{v}_{(2),r} > \tilde{v}_{(1),r}$ for each $r = 1,\ldots,\tilde{R}$ and $\beta(v|\theta) := \beta(v|0,\theta)$

3. For each $\rho \in \mathcal{A}$, compute $\left\{\beta(\tilde{v}_{(2),r}|\rho,\theta)\right\}_{r=1}^{\tilde{R}}$ using $\beta(v|\rho,\theta) = \{\beta(v|\theta) + L(\rho|v,\theta)(\rho - \beta(v|\theta))\}\cdot 1(v > \rho)$ following Li, Perrigne, and Vuong (2003) and construct (19).

We do this for each θ_s, $s = 1,\ldots,S$. Then, solving (20) is easy.

Only the larger valuations $\left\{\tilde{v}_{(2),r}\right\}_{r=1}^{\tilde{R}}$ can be used for step 2 above. However, we compute all the bids $\left\{\left(\tilde{b}_{1,r}(\theta),\tilde{b}_{2,r}(\theta)\right)\right\}_{r=1}^{\tilde{R}}$ with $\tilde{b}_{i,r}(\theta) = \beta(\tilde{v}_{i,r}|\theta)$ for $i = 1,2$ and $r = 1,\ldots,\tilde{R}$. This is useful to estimate the predictive bid density. That is, we may estimate the marginal of bid density using

$$\widehat{g}(b|f(\cdot|\theta)) := \frac{1}{2\tilde{R}h_m}\sum_{r=1}^{\tilde{R}}\sum_{i=1}^{2} K\left(\frac{b - \tilde{b}_{i,r}(\theta)}{h_m}\right) \tag{21}$$

Then, the marginal predictive bid density is estimated by

$$\widehat{g}(b|y) := \frac{1}{S}\sum_{s=1}^{S}\widehat{g}(b|f(\cdot|\theta_s)) \tag{22}$$

In addition, to estimate the joint predictive bid density, we first estimate

$$\tilde{g}(b_1,b_2|y) := \frac{1}{S}\sum_{s=1}^{S}\left\{\frac{1}{\tilde{R}h_j}\sum_{r=1}^{\tilde{R}} K\left(\frac{b_1 - \tilde{b}_{1,r}(\theta_s)}{h_j}\right) K\left(\frac{b_2 - \tilde{b}_{2,r}(\theta_s)}{h_j}\right)\right\}$$

with a kernel K and some bandwidths h_m and h_j and symmetrize this using

$$\widehat{g}(b_1,b_2|y) = \frac{1}{2}\left\{\tilde{g}(b_1,b_2|y) + \tilde{g}(b_2,b_1|y)\right\} \tag{23}$$

We employ $\tilde{R} = 1,500$.

References

[1] Andrieu, C., A. Doucet, and R. Holenstein (2007), "Particle Markov chain Monte Carlo," *Unpublished paper: Department of Mathematics, University of Bristol.*

[2] Andrieu, C., A. Doucet, and G. Roberts (2007), "The expected auxiliary variable method for Monte Carlo simulation," *Unpublished paper*

[3] Beresteanu, A. (2007) "Nonparametric Estimation of Regression Functions under Restrictions on Partial Derivatives," *Working Paper*

[4] Campo, S., Perrigne, I., and Vuong, Q. (2003), "Asymmetry in First-Price Auctions with Affiliated Private Values," *Journal of Applied Econometrics,* **18**: 179–207

[5] Donald, S. and Paarsch, H. (1993), "Piecewise Pseudo-Maximum Likelihood Estimation in Empirical Models of Auctions," *International Economic Review,* **34** : 121–148.

[6] Donald, S. and Paarsch, H. (1996), "Identification, Estimation, and Testing in Parametric Empirical Models of Auctions within the Independent Private Values Paradigm," *Econometric Theory,* **12** : 517–567.

[7] Flury, T., and Shephard, N. (2008) "Bayesian Inference Based only on Simulated Likelihood: Particle Filter Analysis of Dynamic Economic Models," *Working Paper.*

[8] Fernandez-Villaverde, J. and J. F. Rudio-Ramirez (2007) "Estimating macroeconomics models: a likelihood approach," *Review of Economic Studies,* **74** : 1059–1087.

[9] Fernandez-Villaverde, J., J. F. Rudio-Ramirez, and M. S. Santos (2006) "Convergence properties of the likelihood of computed dynamic models," *Econometrica,* **74**: 93–119.

[10] Fritsch, F. N. and R. E. Carlson, (1980) "Monotone Piecewise Cubic Interpolation," *SIAM Journal of Numerical Analysis,* **17**: 238–246

[11] Gallant, A.R., and Nychka, D.W. (1987) "Seminonparametric maximum likelihood estimation," *Econometrica,* **55,** 363–390.

[12] Geweke, J. (1995), "Bayesian Inference for Linear Models Subject to Linear Inequality Constraints" *Working Paper.*

[13] Guerre, E., Perrigne, I., and Vuong, Q. (2000), "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica,* **68**: 525–574.

[14] Haario, H., Saksman, E., and Tamminen, J. (2001) "An Adaptive Metropolis Algorithm," *Bernoulli,* **7** : 223–242.

[15] Hajivassiliou, V., and McFadden, D. (1998), "The Method of Simulated Scores for the Estimation of LDV Models," *Econometrica,* **66**: 863 – 896.

[16] Henderson, D., List, J., Millimet, D., Parmeter, C., and Price, M. (2008), "Imposing Monotonicity Nonparametrically in First-Price Auctions," *Working Paper.*

[17] Hendricks, K., Pinkse, J., and Porter, R. (2003) "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions," *Review of Economic Studies,* **70**: 115–145.

[18] Hirano, K. and Porter, J. (2003) "Asymptotic Efficiency in Parametric Structural Models with Parameter-Dependent Support," *Econometrica,* **71**: 1307–1338.

[19] Keane, M. (1990), "Four Essays in Empirical Macro and Labor Economics" Ph.D. dissertation, Brown University.

[20] Kim, D. (2008) "Auction Design using Bayesian Methods," *Working Paper: University of Arizona,* **08-11**.

[21] Laffont,J., Ossard, H., and Vuong, Q. (1995) “Econometrics of First-Price Auctions,” *Econometrica,* **63** : 953–980

[22] Laffont,J. and Vuong, Q., (1996) “Structural Analysis of Auction Data,” *The American Economic Review,* **86**, : 414–420.

[23] Li,T., Perrigne, I., and Vuong, Q., (2002) “Structural Estimation of the Affiliated Private Value Auction Model,” *RAND Journal of Ecoomics,* **33**, : 171–193.

[24] Li,T., Perrigne, I., and Vuong, Q., (2003) “Semiparametric Estimation of the Optimal Reserve Price in First-Price Auctions,” *Journal of Business & Economic Statistics,* **21**(1), : 53–64.

[25] Li,T. and Vuong, Q., (1997) “Using All Bids in Parametric Estimation of First Price Auctions,” *Economics Letters,* **55** : 321–325.

[26] Milgrom, P. R., and Weber, R. J. (1982), “A Theory of Auctions and Competitive Bidding,” *Econometrica,* **50**, 1089–1122.

[27] Newey, N. K., and McFadden, D. (1994) “Large Sample Estimation and Hypothesis Testing,” *Handbook of Econometrics,* **4**, 2111 – 2245

[28] Paarsch, H. J. (1997), “Deriving an Estimate of the Optimal Reserve Price: An Application to British Columbian Timber Sales,” *Journal of Econometrics,* **78**, 333–357.

[29] Riley, J., and Samuelson, W. (1981), “Optimal Auctions,” *American Economic Review,* **71**, 381–392.

[30] Verdinelli, I. and Wasserman, L. (1998), “Bayesian Goodness-Of-Fit Testing Using Infinite-Dimensional Exponential Families,” *The Annals of Statistics,* **26**, 1215–1241.

[31] Wilson, R., (1977), “A Bidding Model of Perfect Competition,” *Review of Economic Studies,* **44**: 511–518.