A Shape Constrained Estimator of Bidding Function of First-Price Sealed-Bid Auctions

Yu Yvette Zhang

Abstract

This paper is concerned with economic analysis of first-price sealed-bid auctions with risk averse bidders. The identification is based on exogenous variations in the number of bidders across auctions. We present a shape constrained estimator of the bidding function, which satisfies the theoretical properties of passing through origin, positivity and monotonicity. The underlying utility function and bidder value distribution are readily obtained from the estimated bidding function. Monte Carlo simulations demonstrate good performance of the proposed estimator.

Key words: First-price auctions, Bidding functions, Risk aversion, Shaped constrained estimator.

JEL classification: C5 ; C14; D44

*Department of Agricultural Economics, Texas A&M University, College Station, TX 77843, USA; yzhang@tamu.edu
1 Introduction

This study concerns economic analysis of auction data from first-price sealed-bid auctions with risk averse bidders. Guerre et al. (2000) studied the nonparametric identification of utility function of first-price sealed-bid auctions under the independent private value paradigm. They showed that the equilibrium bidding strategy depends on bidders’ utility, value distribution, and the number of bidders. Provided that number of bidders is independent of value distribution, there exists a quantile-by-quantile equality, through the inverse bidding function, across otherwise identical auctions with different number of bidders. This inverse bidding function identifies the latent utility function and value distribution.

Based on the identification result of Guerre et al. (2000), Kim (2015) suggested a method that nonparametrically calculates the utility function via repeated applications of a contraction mapping operator; Zincenko (2016) proposed a sieves estimator and established its large sample properties. The current study employs the same identification strategy and proposes a shape constrained estimator of the inverse bidding function. This novel estimator is constructed to be positive, monotone and pass through the origin, satisfying the theoretical properties of the (inverse) bidding function. The underlying utility function and bid value density are readily inferred from the estimated bidding function. Our Monte Carlo simulations demonstrate good performance of the proposed method.

2 Auction model and identification

In this section, we briefly describe the model of first-price sealed-bid auctions and its identification. Readers are referred to Guerre et al. (2000, 2009) and Campo et al. (2011) for rigorous treatments. We consider first-price sealed-bid auctions of a single indivisible good within the independent private values paradigm. Each bidder has a private value drawn independently from a common distribution \( F(\cdot) \) defined on a compact support \([0, \bar{v}]\) with a density \( f(\cdot) > 0\). For simplicity, we assume there is no reservation price in the auctions. Bidders’ preference is represented by a utility function \( U(\cdot) \) satisfying \( U'(\cdot) > 0, U''(\cdot) \leq 0\) and \( U(0) = 0\).

In a first-price auction with \( n \geq 2 \) bidders, bidder \( i \) with a private value \( v_i \), chooses to bid \( b_i \) to maximize her expected utility. In particular, her objective function is given by \( U(v_i - b_i) \times \Pr(b_i \geq b_j, \forall j \neq i) \). There exists a strictly increasing symmetric Bayesian Nash
equilibrium strategy $s(\cdot)$ that satisfies

$$s'(v_i) = (n - 1) \frac{f(v_i)}{F(v_i)} \lambda(v_i - b_i), \; v_i \in [0, \bar{v}],$$  \hspace{1cm} (1)$$

where $\lambda(\cdot) = U(\cdot)/U'(\cdot)$. It can be shown that $\lambda(0) = 0$ and $\lambda'(\cdot) > 0$ on a compact support $[0, \bar{v} - b]$, where $\bar{v} - b$ corresponds to the maximum value of a bidder’s gain; see e.g., Athey (2001) and Maskin and Riley (2003). Under a general utility function, (1) does not provide a closed form solution to the bidding strategy. Extra information is needed for identification. For instance, Lu and Perrigne (2008) assumed that the value distribution is inferred from a separate set of ascending auctions with the same value distribution; Campo et al. (2011) estimated the utility function under some parametric identifying assumptions.

Kim (2015) and Zincenko (2016) exploited identification based on variations in the number of bidders across auctions with the same value distribution. Let $b_n = s(v)$ be the bid corresponding to value $v$ in an auction with $n$ bidders. Denote by $G_n$ and $g_n$ the bid distribution and density functions of this auction. One can derive the inverse bidding function, from equation (1), as follows:

$$v = s^{-1}(b_n) = b_n + \lambda^{-1}\left(\frac{G_n(b_n)}{(n - 1)g_n(b_n)}\right).$$  \hspace{1cm} (2)$$

Next suppose there exist two otherwise identical auctions with different number of bidders $n_1$ and $n_2$, which are assumed to be independent of the value distribution. For $j = 1, 2$, let $v_{n_j}(\alpha)$ and $b_{n_j}(\alpha)$ be the $\alpha$th quantile of the value and bid distributions and define

$$R_{n_j}(\alpha) = \frac{\alpha}{(n_j - 1)g_{n_j}(b_{n_j}(\alpha))}, \; \alpha \in [0, 1].$$

It follows that $v_{n_j}(\alpha) = b_{n_j}(\alpha) + \lambda^{-1}(R_{n_j}(\alpha)), \; j = 1, 2$. Since the two auctions share a common value distribution, we have

$$b_{n_1}(\alpha) + \lambda^{-1}(R_{n_1}(\alpha)) = b_{n_2}(\alpha) + \lambda^{-1}(R_{n_2}(\alpha)), \; \forall \alpha \in [0, 1].$$  \hspace{1cm} (3)$$

Kim (2015) interpreted this quantile-by-quantile equality of the inverse bidding function as a transformation and established a contraction mapping that converges to the underling utility function. He proposed a method of calculating the utility function based on iterative
applications of a contractor operator. Zincenko (2016) considered the more general case of auctions with more than two different number of bidders and proposed an estimator that minimizes a certain global distance between the value quantile functions across auctions differing in number of bidders. He employed a sieve based estimator and established its large sample properties.

3 Estimation

Similarly to Kim (2015) and Zincenko (2016), we base identification on variations in number of bidders across auctions. To ease exposition, we present the case with two different bidder numbers; generalization to more than two bidder numbers is straightforward. Following Kim (2015), we make the simplifying assumption that bid distribution and density are known when introducing our estimator in this section. Interested readers are referred to, e.g., Guerre et al. (2000), Marmer and Shneyerov (2012) and Luo and Wan (forthcoming) for the estimation of bid density and related quantities.

Recall that $\lambda$ is a bounded increasing function with $\lambda(0) = 0$. It follows that $\lambda^{-1}$ shares the same set of properties. We seek a nonparametric estimator that satisfies these properties. Our strategy is to employ an integration transformation of a smooth non-negative function as follows

$$
\Psi(x) = \int_0^x \psi(y)dy,
$$

where $\psi : \mathcal{R} \to (0 \cup \mathcal{R}^+)$ and $\psi(x) < \infty$ for $x \in (0, \infty)$. It can be easily verified that $\Psi(0) = 0$, $0 \leq \Psi(x) < \infty$ and $\Psi'(x) = \psi(x) \geq 0$ for $x \in (0, \infty)$. Ramsay (1998) used this integration device to model monotone functions. This approach was adopted by Zhang et al. (2011) and Liu et al. (2015) to model monotone auction bidding processes. Henderson et al. (2012) considered a kernel-based monotone estimator of bidding function.

Ramsay (1998) set $\psi(x) = \exp(w(x))$, where $w$ is a smooth real valued function. In this study, we choose to use $\psi(x) = w^2(x)$ mainly because it offers the advantage that $\Psi$ admits a simple analytical form when $w$ is a polynomial or spline function. The polynomial case is particularly simple to handle; for instance, $w(x) = c_0 + c_1 x$ yields $\Psi(x) = c_0^2 x + c_0 c_1 x^2 + \frac{c_1^2}{3} x^3$. 

3
Generally under a $K$-degree polynomial $w(x; c) = \left( \sum_{s=0}^{K} c_s x^s \right)$, we have

$$
\Psi(x; c) = \int_{0}^{x} w^2(y; c) dy = \sum_{s=0}^{K} \sum_{t=0}^{K} \frac{c_s c_t}{s + t + 1} x^{s+t+1}, \quad x > 0.
$$

Alternatively, we can use a spline function. For instance, the commonly used truncated power series is given by

$$
w(x; c) = \sum_{s=0}^{K} c_{1,s} x^s + \sum_{s=1}^{J} c_{2,s} (x - z_s)^K,
$$

where $(x)_+ = \max(0, x)$ and $0 < z_1 < \cdots < z_J < \infty$ are a set of spline knots. Although slightly more tedious, the integration of a squared truncated power series also admits a simple analytical form, which remains a truncated power series; see Appendix for details.

We use $\Psi(x; c) = \int_{0}^{x} w^2(y; c) dy$ to approximate $\lambda^{-1}(x)$, where $w(\cdot; c)$ is a polynomial or spline function given above. We then construct an estimator based on the quantile-to-quantile equality (3). Let

$$
v_{n_j}(\alpha; c) = b_{n_j}(\alpha) + \Psi(R_{n_j}(\alpha); c), \quad j = 1, 2.
$$

We consider a minimum sum of squares estimator

$$
\min_c \sum_{j=1}^{2} \sum_{\alpha \in A} \left\{ v_{n_j}(\alpha; c) - b_{n_j}(\alpha) - \Psi(R_{n_j}(\alpha); c) \right\}^2,
$$

where $j = 3 - j, j = 1, 2$ and $A$ is a set of user selected quantile levels.

The optimization problem (5) can be solved by a standard nonlinear least squares routine or a general purpose optimization routine. Since $\Psi(\cdot; c) = \Psi(\cdot; -c)$, we impose the identification restriction $c_0 \geq 0$. Denote the solution to (5) by $\hat{c}$. The estimated inverse bidding function is given by $\hat{\lambda}^{-1}(\cdot) = \Psi(\cdot; \hat{c})$. We can then proceed to calculate $\hat{\lambda}(\cdot)$ and its associated utility function. Under the condition that $\lambda(\cdot) = U(\cdot)/U'(\cdot)$ and $U(1) = 1$, we have

$$
\hat{U}(x) = \exp \left( \int_{1}^{x} \frac{1}{\hat{\lambda}(y)} dy \right), \quad x \in [0, v - b].
$$
Lastly, the latent value associated with a bid $b_n$ in an auction with $n$ bidders is estimated as

$$
\hat{v}(b_n) = b_n + \lambda^{-1}(R_n(G_n(b_n))), \ n = n_1, n_2.
$$

The value density can then be estimated, parametrically or nonparametrically, based on all estimated values pooled across the two types of auctions.

4 Monte Carlo simulations

We follow the experiment design of Kim (2015) in our Monte Carlo simulations. Two value distributions are considered: (i) the exponential distribution with mean 3, truncated from above at 10; (ii) the standard log-normal distribution, truncated from below at its 5% quantile and from above at its 95% quantile, and rescaled to be in $[0, 10]$. In all experiments, bidders’ utility is taken to be $U(x) = x^{1-\theta}$, with $\theta = 0.3, 0.5$ or 0.7. Under this utility, the bidding function in an auction with $n$ bidders is conveniently given by

$$
s(v) = v - \int_0^v \left\{ \frac{F(y)}{F(v)} \right\}^{\frac{n-1}{\theta}} dy.
$$

We consider two bidder numbers $n_1 = 2$ and $n_2 = 4$. The number of auctions are set to $T = 200$ or $500$. Each experiment is repeated 1,000 times.

To implement the proposed estimator, we replace the population bid quantile and density functions in (2) by their nonparametric estimates. For $j = 1, 2$, we estimate the bid quantile function by the empirical quantile function $\hat{b}_{n_j}(\cdot)$ and the bid density function by a kernel estimate $\hat{g}_{n_j}(\cdot)$ with a triweight kernel and Silverman’s rule of thumb bandwidth. Both estimates are calculated based on $(n_j \times T)$ bids from all $T$ auctions. We then plug

$$
\hat{R}_{n_j}(\alpha) = \frac{\alpha}{(n_j - 1)\hat{g}_{n_j}(\hat{b}_{n_j}(\alpha))}, \ j = 1, 2,
$$

into objective function (5) and use a 99-point quantile grid $\mathcal{A} = \{i/100\}_{i=1}^{99}$ in the estimation. We experiment with both polynomial and spline specification of the proposed estimator. In particular, we consider $w_p(x; c) = c_0 + c_1x + c_2x^2$ and $w_s(x; c) = c_0 + c_1x + c_2(x-z)_+$, where the spline knot $z$ is set to be the median of $\{\hat{R}_{n_j}(\alpha), \alpha \in \mathcal{A}, j = 1, 2\}$. We conduct the simulations using the optimization routine `optim` of R language. The computation is very
fast and convergence is obtained in all experiments.

Equipped with the estimated bidding function, we proceed to calculate the utility function and bidder values. We then estimate the underlying value density based on \((n_1 + n_2) \times T\) pseudo values. We use the kernel density estimator with the triweight kernel and Silverman’s rule of thumb bandwidth. We summarize the estimation results by plotting the pointwise averages, 2.5% and 97.5% percentiles of the estimated utility functions on \([0, 2]\) and the same set of quantities for the estimated value densities on \([0, 10]\). These summary plots for the polynomial-based and spline-based estimators are virtually identical. To save space, we only report the polynomial estimation results; those of spline estimation are available from the author upon request.

In Figures 1 and 2, we report the estimated utility and value densities on auctions with a truncated exponential value distribution. In each plot, the estimation results are represented by dotted lines while the true function by a solid line. It is seen that the average utility estimates essentially coincides with the true utility in all experiments. The 95% pointwise frequency bands are rather tight and appear to get narrower with the risk averse parameter \(\theta\). As expected, the frequency bands shrink when the number of auctions increases from 200 to 500. Similarly, the estimated value density functions closely track the true density and their frequency bands improve with sample size. Results for auctions with a truncated lognormal value distribution are reported in Figures 3 and 4. Similar general pattern and good performance are observed.

Lastly to demonstrate the consequence of ignoring risk aversion, we conduct our analysis under the false assumption of risk neutrality. When bidders are risk neutral, the (inverse) bidding function is reduced to the identity function, and bidder values are readily obtained using equation (2) with \(\lambda^{-1}(x) = x\). We then estimate the value densities based on these erroneously estimated values, using the same kernel estimator described above. For the case of truncated exponential value distribution with 200 bidders, we compare in Figure 5 the average estimates allowing risk aversion (black-dotted) and those under risk neutrality (red-dotted). The mis-specified estimates under risk neutrality are clearly biased to the right and the discrepancy increases with the degree of risk aversion. Similar patterns are observed in all other experiments.

Lastly we note that instead of \(L_2\) type criterion in (5), we may use \(L_1\) or \(L_\infty\) criterion in the estimation. It is also possible to incorporate exogenous variables into the estimation to accommodate auction and bidder heteroskedasticity. We leave these topics to future studies.
Figure 1: Utility estimation with truncated exponential value distribution

(a) $T = 200, \theta = 0.3$  (b) $T = 200, \theta = 0.5$  (c) $T = 200, \theta = 0.7$

(d) $T = 500, \theta = 0.3$  (e) $T = 500, \theta = 0.5$  (f) $T = 500, \theta = 0.7$

Figure 2: Value density estimation with truncated exponential value distribution

(a) $T = 200, \theta = 0.3$  (b) $T = 200, \theta = 0.5$  (c) $T = 200, \theta = 0.7$

(d) $T = 500, \theta = 0.3$  (e) $T = 500, \theta = 0.5$  (f) $T = 500, \theta = 0.7$
Figure 3: Utility estimation with truncated lognormal value distribution

Figure 4: Value density estimation with truncated lognormal value distribution
Figure 5: Value densities estimation with truncated exponential distribution (black solid: true density; black dotted: average estimates allowing risk aversion; red dotted: average estimates under risk neutrality)

References


**Appendix**

We provide in this appendix an analytical formula for integrated squared spline functions. For simplicity, we present the case of linear splines. Generalization to higher order splines is straightforward. Let \( z_0 = 0 \) and \( z_{J+1} = \infty \). Consider \( w(t) = \beta_0 + \beta_1 t + \sum_{j=1}^{J} b_j (t - z_j)_+ \), where \( z_0 < z_1 < \cdots < z_J < z_{J+1} \). Let \( (\alpha_0 = \beta_0, a_0 = \beta_1) \), \( (\alpha_j = \beta_0 - \sum_{i=1}^{j} b_i z_i, a_j = \beta_1 + \sum_{i=1}^{j} b_i) \) for \( j = 1, \ldots, J \), and \( w(t) = \sum_{j=0}^{J} (\alpha_j + a_j t) I(z_j \leq t < z_{j+1}) \). Define \( h(\alpha, a; s_1, s_2) = \int_{s_1}^{s_2} (\alpha + at)^2 dt = (\alpha^2 t + a\alpha t^2 + \frac{a^2}{3} t^3)|_{s_1}^{s_2} \). Let \( z_0 = 0 \) and \( h_0 = h(\alpha_0, a_0; z_0, z_1) \), \( h_1 = h(\alpha_1, a_1; z_1, z_2) \), \ldots, \( h_J = h(\alpha_J, a_J; z_J, 1) \). Next define \( J(s) = \max\{j : z_j \leq s, j = 1, \ldots, J\} \). We then have

\[
\int_0^s w(t) dt = \sum_{j=0}^{J(s)-1} h_j + h(\alpha_{J(s)}, a_{J(s)}; z_{J(s)}, s) \equiv H_{J(s)-1} + h(\alpha_{J(s)}, a_{J(s)}; z_{J(s)}, s).
\]

Thus the integrated squared linear spline evaluated at a given point \( s \) is a third degree polynomial, whose coefficients depend the value of \( s \) and the spline knots \( \{z_1, \ldots, z_J\} \).