Journal of Econometrics 167 (2012) 345–357

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Quantile-based nonparametric inference for first-price auctions

Vadim Marmer[a,*], Artyom Shneyerov[b,c,d]

[a] Department of Economics, University of British Columbia, 997 - 1873 East Mall, Vancouver, BC, Canada V6T 1Z1
[b] CIREQ, University of Montreal, C.P. 6128, Succursale Centre-ville, Montreal, QC, Canada H3C 3J7
[c] CIRANO, 2020, University St., 25th Floor, Montreal, QC, Canada, H3A 2A5
[d] Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, QC, Canada H3G 1M8

ARTICLE INFO

Article history:
Available online 25 September 2011

JEL classification:
C14
D44

Keywords:
First-price auctions
Independent private values
Nonparametric estimation
Kernel estimation
Quantiles
Optimal reserve price

ABSTRACT

We propose a quantile-based nonparametric approach to inference on the probability density function (PDF) of the private values in first-price sealed-bid auctions with independent private values. Our method of inference is based on a fully nonparametric kernel-based estimator of the quantiles and PDF of observable bids. Our estimator attains the optimal rate of Guerre et al. (2000), and is also asymptotically normal with an appropriate choice of the bandwidth.

1. Introduction

Following the seminal article of Guerre et al. (2000), GPV hereafter, there has been an enormous interest in nonparametric approaches to auctions.[1] By removing the need to impose tight functional form assumptions, the nonparametric approach provides a more flexible framework for estimation and inference. Moreover, the sample sizes available for auction data can be sufficiently large to make the nonparametric approach empirically feasible.[2] This paper contributes to this literature by providing a fully nonparametric framework for making inferences on the density of bidders' valuations $f(v)$. The need to estimate the density of valuations arises in a number of economic applications, as for example the problem of estimating a revenue-maximizing reserve price.[3]

As a starting point, we briefly discuss the estimator proposed in GPV. For the purpose of introduction, we adopt a simplified framework. Consider a random, i.i.d. sample b_{il} of bids in first-price auctions each of which has n *risk-neutral* bidders; l indexes auctions and $i = 1, \dots, n$ indexes bids in a given auction. GPV assume independent private values (IPVs). In equilibrium, the bids are related to the valuations via the equilibrium bidding strategy $B : b_{il} = B(v_{il})$. GPV show that the inverse bidding strategy is identified directly from the observed distribution of bids:

$$v = \xi(b) \equiv b + \frac{1}{n-1}\frac{G(b)}{g(b)}, \tag{1}$$

where $G(b)$ is the cumulative distribution function (CDF) of bids in an auction with n bidders, and $g(b)$ is the corresponding density. GPV propose to use nonparametric estimators $\hat{G}$ and $\hat{g}$. When $b = b_{il}$, the left-hand side of (1) will then give what GPV call the pseudo-values $\hat{v}_{il} = \hat{\xi}(b_{il})$. The CDF $F(v)$ is estimated as the empirical CDF, and the PDF $f(v)$ is estimated by the method of kernels, both using $\hat{v}_{il}$ as observations. GPV show that, with an appropriate choice of the bandwidth, their estimator converges to the true value at the optimal rate (in the minimax sense; Khasminskii, 1979). However, the asymptotic distribution of this estimator is as yet unknown, possibly because both steps of the GPV method are nonparametric with estimated values $\hat{v}_{il}$ entering the second stage.

The estimator $\hat{f}(v)$ proposed in this paper avoids the use of pseudo-values. It builds instead on the insight of Haile et al.

* Corresponding author. Department of Economics, University of British Columbia, 997 - 1873 East Mall, Vancouver, BC, Canada V6T 1Z1.
E-mail addresses: vadim.marmer@ubc.ca (V. Marmer), achneero@alcor.concordia.ca (A. Shneyerov).

1 See a recent survey by Athey and Haile (2007).

2 For example, List et al. (2004) study bidder collusion in timber auctions using thousands of auctions conducted in the Province of British Columbia, Canada. Samples of similar size are also available for highway procurement auctions in the United States (e.g., Krasnokutskaya, 2011).

3 Several previous articles have studied that problem, see Paarsch (1997), Haile and Tamer (2003), and Li et al. (2003). In the supplement to this paper, we discuss how the approach developed here can be used for construction of confidence sets for the optimal reserve price. The supplement is available as Marmer and Shneyerov (2010) from the UBC working papers series and the authors' web-sites.

0304-4076/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2011.09.020

(2003).[4] They show that the quantiles of the distribution of valuations can be expressed in terms of the quantiles, PDF, and CDF of bids. We show below that this relation can be used for estimation of $f(v)$. Consider the τ-th quantile of valuations $Q(\tau)$ and the τ-th quantile of bids $q(\tau)$. The latter can be easily estimated from the sample by a variety of methods available in the literature. As for the quantile of valuations, since the inverse bidding strategy $\xi(b)$ is monotone, Eq. (1) implies that $Q(\tau)$ is related to $q(\tau)$ as follows:

$$Q(\tau) = q(\tau) + \frac{\tau}{(n-1)\, g\,(q(\tau))}, \tag{2}$$

providing a way to estimate $Q(\tau)$ by a plug-in method. The CDF $F(v)$ can then be recovered by inverting the quantile function, $F(v) = Q^{-1}(v)$.

Our estimator $\hat{f}(v)$ is based on a simple idea that by differentiating the quantile function we can recover the density: $Q'(\tau) = 1/f\,(Q(\tau))$, and therefore $f(v) = 1/Q'\,(F(v))$. Taking the derivative in (2) and using the fact that $q'(\tau) = 1/g\,(q\,(\tau))$, we obtain, after some algebra, our basic formula:

$$f(v) = \left(\frac{n}{n-1}\frac{1}{g\,(q\,(F\,(v)))} - \frac{1}{n-1}\frac{F(v)g'\,(q\,(F(v)))}{g^3\,(q\,(F(v)))}\right)^{-1}. \tag{3}$$

Note that all the quantities on the right-hand side, i.e. $g(b)$, $g'(b), q(\tau), F(v) = Q^{-1}(v)$ can be estimated nonparametrically, for example, using kernel-based methods. Once this is done, we can plug them in (3) to obtain our nonparametric estimator.

The expression in (3) can be also derived using the following relationship between the CDF of values and the CDF of bids:

$$F(v) = G\,(B(v))\,.$$

Applying the change of variable argument to the above identity, one obtains

$$\begin{aligned} f(v) &= g\,(B(v))\,B'\,(v) \\ &= g\,(B(v))\,/\xi'\,(B(v)) \\ &= \left(\frac{n}{n-1}\frac{1}{g\,(B(v))} - \frac{1}{n-1}\frac{F(v)g'\,(B(v))}{g^3\,(B(v))}\right)^{-1}. \end{aligned}$$

Note however, that from the estimation perspective, the quantile-based formula appears to be more convenient, since the bidding strategy function B involves integration of F (see Eq. (4)). Furthermore, replacing $B(v)$ with appropriate quantiles has no effect on the asymptotic distribution of the estimator.

Our framework results in the estimator of $f(v)$ that is both consistent and asymptotically normal, with an asymptotic variance that can be easily estimated. Moreover, we show that, with an appropriate choice of the bandwidth sequence, the proposed estimator attains the minimax rate of GPV.

In a Monte Carlo experiment, we compare finite sample biases and mean squared errors of our quantile-based estimator with that of the GPV's estimator. Our conclusion is that neither estimator strictly dominates the other. The GPV estimator is more efficient when the PDF of valuations has a positive derivative at the point of estimation and the number of bidders tends to be large. On the other hand, the quantile-based estimator is more efficient when the PDF of valuations has a negative derivative and the number of bidders is small. The Monte Carlo results suggest that the proposed estimator will be more useful when there are sufficiently many independent auctions with a small number of bidders.[5]

The rest of the paper is organized as follows. Section 2 introduces the basic setup. Similarly to GPV, we allow the number of bidders to vary from auction to auction, and also allow auction-specific covariates. Section 3 presents our main results. Section 4 discusses the bootstrap-based approach to inference on the PDF of valuations. In Section 5, we extend our framework to the case of auctions with a binding reserve price. We report Monte Carlo results in Section 6. Section 7 concludes. The proofs of the main results are given in the Appendix. The supplement to this paper contains the proof of the bootstrap result in Section 4, some additional Monte Carlo results, as well as an illustration of how the approach developed here can be applied for conducting inference on the optimal reserve price.

2. Definitions

The econometrician observes a random sample $\{(b_{il}, x_l, n_l) : l = 1, \ldots, L; i = 1, \ldots, n_l\}$, where b_{il} is the equilibrium bid of *risk-neutral* bidder i submitted in auction l with n_l bidders, and x_l is the vector of auction-specific covariates for auction l. The corresponding unobservable valuations of the object are given by $\{v_{il} : l = 1, \ldots, L; i = 1, \ldots, n_l\}$. We make the following assumption similar to Assumptions A1 and A2 of GPV (see also footnote 14 in their paper).

Assumption 1. (a) $\{(n_l, x_l) : l = 1, \ldots, L\}$ are i.i.d.

(b) The marginal PDF of x_l, φ, is strictly positive and continuous on its compact support $\mathcal{X} \subset R^d$, and admits up to $R \geq 2$ continuous derivatives on its interior.

(c) The distribution of n_l conditional on x_l is denoted by $\pi\,(n|x)$ and has support $\mathcal{N} = \left\{\underline{n}, \ldots, \bar{n}\right\}$ for all $x \in \mathcal{X}$, $\underline{n} \geq 2$.

(d) $\{v_{il} : l = 1, \ldots, L; i = 1, \ldots, n_l\}$ are i.i.d. and independent of the number of bidders conditional on x_l with the PDF $f\,(v|x)$ and CDF $F\,(v|x)$.

(e) $f\,(\cdot|x)$ is strictly positive and bounded away from zero and admits up to $R-1$ continuous derivatives on its support, a compact interval $\left[\underline{v}(x), \overline{v}(x)\right] \subset \mathbb{R}_+$ for all $x \in \mathcal{X}$; $f\,(v|\cdot)$ admits up to R continuous partial derivatives on Interior $(\mathcal{X})$ for all $v \in \left[\underline{v}(x), \overline{v}(x)\right]$.

(f) For all $n \in \mathcal{N}$, $\pi\,(n|\cdot)$ is strictly positive and admits up to R continuous derivatives on the interior of $\mathcal{X}$.

Under Assumption 1(c), the equilibrium bids are determined by

$$b_{il} = v_{il} - \frac{1}{(F\,(v_{il}|x_l))^{n-1}} \int_{\underline{v}}^{v_{il}} (F\,(u|x_l))^{n-1}\, du, \tag{4}$$

(see, for example, GPV). Let $g\,(b|n, x)$ and $G\,(b|n, x)$ be the PDF and CDF of b_{il}, conditional on both $x_l = x$ and the number of bidders $n_l = n$. Since b_{il} is a function of v_{il}, x_l, and $F\,(\cdot|x_l)$, the bids $\{b_{il}\}$ are also i.i.d. conditional on (n_l, x_l). Furthermore, by Proposition (i) and (iv) of GPV, for all $n = \underline{n}, \ldots, \overline{n}$ and $x \in \mathcal{X}$, $g\,(\cdot|n, x)$ has the compact support $\left[\underline{b}\,(n, x)\,, \overline{b}\,(n, x)\right]$ for some $\underline{b}\,(n, x) < \overline{b}\,(n, x)$, and $g\,(\cdot|n, \cdot)$ admits up to R continuous bounded partial derivatives.

The τ-th quantile of $F\,(v|x)$ is defined as

$$Q\,(\tau|x) = F^{-1}\,(\tau|x) \equiv \inf_v \{v : F\,(v|x) \geq \tau\}\,.$$

The τ-th quantile of G,

$$q\,(\tau|n, x) = G^{-1}\,(\tau|n, x)\,,$$

is defined similarly. The quantiles of the distributions $F\,(v|x)$ and $G\,(b|n, x)$ are related through the following conditional version of Eq. (2):

$$Q\,(\tau|x) = q\,(\tau|n, x) + \frac{\tau}{(n-1)\, g\,(q\,(\tau|n, x)\,|n, x)}. \tag{5}$$

Note that the expression on the left-hand side does not depend on n, since by Assumption 1(d) and as it is usually assumed in the

---

4 The focus of Haile et al. (2003) is a test of common values. Their model is therefore different from the IPV model, and requires an estimator that is different from the one in GPV. See also Li et al. (2002).

5 We thank a referee for pointing this out.