Quantile regression methods for first-price auctions

Nathalie Gimenes
Department of Economics
PUC-Rio
Brazil
ngimenes@econ.puc-rio.br

Emmanuel Guerre
School of Economics and Finance
Queen Mary University of London
United Kingdom
eoguerre@gmail.com

February 2021

Abstract

The paper proposes a quantile-regression inference framework for first-price auctions with symmetric risk-neutral bidders under the independent private-value paradigm. It is first shown that a private-value quantile regression generates a quantile regression for the bids. The private-value quantile regression can be easily estimated from the bid quantile regression and its derivative with respect to the quantile level. This also allows to test for various specification or exogeneity null hypothesis using the observed bids in a simple way. A new local polynomial technique is proposed to estimate the latter over the whole quantile level interval. Plug-in estimation of functionals is also considered, as needed for the expected revenue or the case of CRRA risk-averse bidders, which is amenable to our framework. A quantile-regression analysis to USFS timber is found more appropriate than the homogenized-bid methodology and illustrates the contribution of each explanatory variables to the private-value distribution. Linear interactive sieve extensions are proposed and studied in the Appendices.

JEL: C14, L70

Keywords: First-price auction; independent private values; dimension reduction; quantile regression; local polynomial estimation; specification testing; boundary correction; sieve estimation.

A previous version of this paper has been circulated under the title "Quantile regression methods for first-price auction:a signal approach". The authors acknowledge useful discussions and comments from Xiaohong Chen, Valentina Corradi, Yanqin Fan, Phil Haile, Xavier d'Haultfoeuille, Vadim Marmer, Isabelle Perrigne, Martin Pesendorfer and Quang Vuong, and the audience of many conferences and seminars. Nathalie Gimenes also thanks Ying Fan and Ginger Jin for encouragements. Many thanks to Elie Tamer and three anonymous referees, who all have been extremely stimulating and helpful to enrich the paper. All remaining errors are our responsibility. Both authors would like to thank the School of Economics and Finance, Queen Mary University of London, for generous funding. Nathalie Gimenes would like to express her gratitude to PUC Rio for generous support.

1 Introduction

Since Paarsch (1992), many parametric methods have been proposed to estimate first-price auction models under the independent private-value paradigm. See Laffont, Ossard and Vuong (1995), Athey and Levin (2001), Hirano and Porter (2003), Li and Zheng (2012), Paarsch and Hong (2012) and the references therein to name just a few. Validating specification choice is difficult and seldom attempted.

On the other hand, the nonparametric approach is very flexible and less subject to misspecification of functional form, so that it is commonly considered in applications and theoretical studies. See Guerre, Perrigne and Vuong (2000, hereafter GPV), Lu and Perrigne (2008), Krasnokutskaya (2011), Marmer and Shneyerov (2012), Hubbard, Paarsch and Li (2012), Campo, Guerre, Perrigne and Vuong (2013), Marmer, Shneyerov and Xu (2013a,b), Hickman and Hubbard (2015), Enache and Florens (2017), Liu and Luo (2017), Liu and Vuong (2018), Luo and Wan (2018), Zincenko (2018) and Ma, Marmer and Shneyerov (2019) among others. But the nonparametric approach comes with the burden of the curse of dimensionality, which considerably limits its scope of applications.

Haile, Hong and Shum (2003, HSS hereafter) and Rezende (2008) have proposed to circumvent the curse of dimensionality using a regression specification that purges the bids from the covariate effects. The resulting homogenized bids are then used as in GPV to backup the density of their private-value counterparts. This approach can tackle linear dependence, but is not appropriate to capture more complex interactions. The present paper proposes to use instead a more flexible quantile-regression specification.

The use of quantile in first-price auctions is not new. Milgrom (2001, Theorem 4.7) reformulates the identification relation of GPV using quantile function. See Guerre, Perrigne, Vuong (2009) and Campo et al. (2013) for the use of quantile in risk-aversion identification and, for related estimation methods, Menzel and Morganti (2013), Enache and Florens (2017), Liu and Vuong (2018), Luo and Wan (2018). Marmer and Shneyerov (2012) have proposed a quantile-based estimator of the private-value probability density function (pdf), which is an alternative to the two-step GPV method. See also Marmer, Shneyerov and Xu (2013b) who consider a nonparametric single-index quantile model. Guerre and Sabbah (2012) have noted that the private-value quantile function can be estimated using a one-step procedure from the estimation of the bid quantile function and its first derivative. Gimenes (2017) has developed a flexible but parsimonious quantile-regression estimation strategy for ascending auction. The present paper is however the first to develop a quantile inference framework in a first-price auction setting allowing for many covariates.

Using Koenker and Bassett (1978) quantile regression framework is appealing for several reasons. First, the quantile-regression specification is flexible enough to capture economically relevant effects as in Gimenes (2017), which could be ignored using parametric ones or less interpretable nonparametric models. These parsimonious specifications can be estimated with reasonable nonparametric rates, allowing implementation in small samples with rich covariate environment. Compared to GPV, this estimation method is one-step and only requests one bandwidth parameter, which theoretical choice follows from standard bias variance expansion. As detailed in Appendix A, quantile-regression specification can be enriched to include more nonparametric features using sieve extensions ranging from the additive specification of Horowitz and Lee (2005) to fully nonparametric one as in Belloni, Chernozhukov, Chetverikov and Fernández-Val (2019). Second, the quantile approach comes with a stability property of linear specifications, which ensures that a private-value quantile regression generates an bid quantile-regression. This is key for our estimation procedure and also for testing, as it transfers many null hypotheses of interest for the latent private-value distribution to the bid quantile-regression slopes. Tests derived from Koenker and Xiao (2002), Escanciano and Velasco (2010), Rothe and Wied (2013), Escanciano and Goh (2014) or Liu and Luo (2017) can be used to test correct specification of the quantile-regression or homogenized-bid models, or exogeneity of the auction format and of entry. Third, the quantile representation used in the paper can play the role of a reduced form generated by a more complex model, such as the random-coefficient model considered in Berry, Levinsohn and Pakes (1995), Hoderlein, Klemelä and Mammen (2010), or Backus and Lewis (2019) to name just a few. In particular, random coefficients drawn from an elliptical distribution generates a quantile specification which extends homogenized bid and can be easily estimated.

Fourth, the proposed augmented quantile-regression estimation methodology is based upon local polynomial for quantile levels, and is therefore not affected by asymptotic boundary bias. This permits better estimation of the upper tail distribution than most nonparametric methods, which is important as the winner's private value is high for a large number of bidders. Fifth, it can also be used to recover important parameters such as the probability or cumulative density functions (pdf and cdf hereafter), mitigating the curse of dimensionality that affects most nonparametric methods. Plug in estimation of the seller expected revenue, optimal reserve price and of agent constant relative risk-aversion parameter are also considered.

The rest of the paper is organized as follows. The next section 2 introduces our stability result for linear quantile specification. Section 2.3 considers the homogenized-bid and random-coefficient specifications. Section 2.4 reviews some testing strategies based upon the bid quantile regression. Section 3 explains how to use our quantile specification for estimating agent's risk-aversion, seller's expected revenue, and the cdf and pdf of the private values. A difficulty of the quantile approach for first-price auction is the need to estimate the bid quantile derivative with respect to quantile levels, see Guerre and Sabbah (2012) and the reference therein for related approaches. Section 4 introduces our new augmented quantile regression estimators, which use a quantile-level local-polynomial approach to jointly estimate the bid quantile regression and its higher-order derivatives. Sections 5 and 5.3 group our main theoretical results, including Integral Mean Squared Error (IMSE), optimal bandwidth choice, optimal uniform convergence rate and Central Limit Theorem for the proposed private-value quantile-regression estimators.

Our theoretical results are illustrated with a simulation experiment and an application to USFS first-price auctions in Sections 6 and 7. Some simulation experiments illustrate how the new estimation procedure improves on the GPV two-step density estimator and homogenized bids. A preliminary quantile-regression analysis of the bid quantile function suggests that the homogenized-bid technique should not be applied here because the quantile-regression slopes are not constant. The private-value quantile-regression slope functions reveal the covariates impact, and how strongly bidders in the top of the distribution can differ from the bottom. Section 8 concludes the paper.

The supplementary material groups the sieve extension and proofs. Appendix A details an interactive localized sieve quantile extension and related theoretical results which are the counterparts of the ones obtained for the quantile-regression specification. Appendix B briefly sketches the main proof arguments and states some preliminary lemmas used for the proofs of the two key bias and linearization results in Appendix C and Appendix D, from which our main results follow. The two remaining Appendices group the proofs of our main and intermediary results.

2 First-price auction and quantile specification

A single and indivisible object with some characteristic $X \in \mathbb{R}^D$ is auctioned to $I \geq 2$ buyers. The number of bidders I and X are known to the bidders and the econometrician. Bids B_i are sealed so that a bidder does not know the other bids when forming his own bid. The object is sold to the highest bidder who pays his bid to the seller, and all the bids B_i are then observed by the econometrician. Under the symmetric IPV paradigm, each bidder is assumed to have a private value V_i, $i = 1, \ldots, I$ for the auctioned object. A buyer knows his private value but not the other ones, the common distribution of the independent V_i being common knowledge. The private-value conditional cdf $F(\cdot|X, I)$ is continuous and supported by a compact interval, implying that the conditional private-value quantile function

$$V(\alpha|X, I) = F^{-1}(\alpha|X, I), \quad \alpha \text{ in } [0, 1],$$

is finite for $\alpha = 0$ and $\alpha = 1$.

The private-value quantile function $V(\alpha|x, I)$ plays an important economic role. The bidder's rent at quantile level α is $V(\alpha|x, I) - B(\alpha|x, I)$ where $B(\cdot|x, I)$ is the bid conditional quantile function, and assuming bids depend in a monotonous way on private values as considered below. The private-value quantile conditional function is important to compute counterfactuals, such as the bid quantile function in an alternative auction mechanism. In particular, it can be used to compute the seller expected revenue achieved with any reserve price, see (3.5) below. It allows, as a consequence, to compute an optimal reserve price, or more generally to propose suitable auction designs.

2.1 Private value quantile identification

It is well-known that the bidder i private-value rank

$$A_i = F(V_i|X, I)$$

has a uniform distribution over $[0,1]$ and is independent of X and I. It also follows from the IPV paradigm that the private-value ranks $A_i = 1, \ldots, I$ are independent. The dependence between the private value V_i and the auction covariates X and I is therefore fully captured by the non separable quantile representation

$$V_i = V(A_i|X, I), \quad A_i \overset{\text{iid}}{\sim} \mathcal{U}_{[0,1]} \perp (X, I), \tag{2.1}$$

which, when the private values are generated by an economic structural model, can be also viewed as a nonparametric reduced form. Following Milgrom and Weber (1982) or Milgrom (2001), $V(\cdot|X, I)$ can be also interpreted as a *valuation function,* the private-value rank A_i being the associated signal. In what follows, $G(\cdot|X, I)$ and $g(\cdot|X, I)$ stand for respectively the bid conditional cdf and pdf.

Maskin and Riley (1984) have shown that Bayesian Nash Equilibrium bids $B_i = \sigma(V_i; X, I)$ of symmetric risk-averse or risk-neutral bidders are strictly increasing and continuous in V_i. It follows that $B_i = B(A_i|X, i)$, where $B(\cdot; X, i) = \sigma(F(\cdot|X, I); X, I)$ can be viewed as a bidding strategy depending upon the rank A_i. If $F(\cdot|X, I)$ is also strictly increasing, so is $B(\cdot|X, I)$ and since A_i is uniform it holds

$$G(b|X, I) = \mathbb{P}[B(A_i|X, I) \leq b|X, I] = \mathbb{P}\left[A_i \leq B^{-1}(b|X, I)|X, I\right] = B^{-1}(b|X, I)$$

showing that the bidding strategy $B(\cdot|X, I)$ is also the bid quantile function.

A standard best response argument will show how to identify the private-value quantile function $V(\cdot|X, I)$ from $B(\cdot|X, I)$. Suppose bidder i signal A_i is equal to α, but that her bid is a suboptimal $B(a|X, I)$, all other bidders bidding $B(A_j|X, I)$. Then the probability that bidder i wins the auction is

$$\mathbb{P}\left[B(a|X,I) > \max_{1\leq j\neq i\leq I} B(A_j|X,I)\middle| A_i = \alpha, X, I\right] = \mathbb{P}\left[a > \max_{1\leq j\neq i\leq I} A_j \middle| A_i = \alpha, X, I\right]$$
$$= a^{I-1} \tag{2.2}$$

because the A_j's are independent $\mathcal{U}_{[0,1]}$ independent of X and I. It follows that the expected revenue of such a bid is, for a risk-neutral bidder, $(V(\alpha|X,I) - B(a|X,I))\, a^{I-1}$. If $B(\cdot|X,I)$ is a best-response bidding strategy, the optimal bid of a bidder with signal α is $B(\alpha|X,I)$, that is

$$\alpha = \arg\max_a \left\{(V(\alpha|X,I) - B(a|X,I))\, a^{I-1}\right\}.$$

As $B(\cdot|X,I)$ is continuously differentiable, it follows that

$$\frac{\partial}{\partial a}\left\{(V(\alpha|X,I) - B(a|X,I))\, a^{I-1}\right\}\Big|_{a=\alpha} = 0 \tag{2.3}$$

or equivalently

$$\frac{d}{d\alpha}\left[\alpha^{I-1} B(\alpha|X,I)\right] = (I-1)\,\alpha^{I-2} V(\alpha|X,I)$$
$$\text{with } \frac{d}{d\alpha}\left[\alpha^{I-1} B(\alpha|X,I)\right] = (I-1)\,\alpha^{I-2} V(\alpha|X,I) + \alpha^{I-1}\frac{d}{d\alpha}B(\alpha|X,I).$$

Solving with the initial condition $B(0|X,I) = V(0|X,I)$ and rearranging the equation above gives Proposition 1, which is the cornerstone of our estimation method. From now on $B^{(1)}(\alpha|X,I) = \frac{d}{d\alpha}B(\alpha|X,I)$.

Proposition 1 *Consider a given (X,I), $I \geq 2$, for which $\alpha \in [0,1] \mapsto V(\alpha|X,I)$ is continuously differentiable with a derivative $V^{(1)}(\cdot|X,I) > 0$. Suppose the bids are drawn from the symmetric differential Bayesian Nash equilibrium. Then,*

i. The conditional equilibrium quantile function $B(\cdot|X,I)$ of the I iid optimal bids B_i satisfies

$$B(\alpha|X,I) = \frac{I-1}{\alpha^{I-1}}\int_0^\alpha a^{I-2} V(a|X,I)\, da \text{ with } \lim_{\alpha\downarrow 0} B(\alpha|X,I) = V(0|X,I). \tag{2.4}$$

ii. The bid quantile function $B(\alpha|X,I)$ is continuously differentiable over $[0,1]$ and it holds

$$V(\alpha|X,I) = B(\alpha|X,I) + \frac{\alpha B^{(1)}(\alpha|X,I)}{I-1}. \qquad (2.5)$$

A key feature is the linearity with respect to $V(\cdot|X,I)$ of the private-value to bid quantile functions mapping (2.4), which implies that a private value quantile linear model is mapped into a similar bid linear model, as detailed below for the well-known quantile regression. Proposition 1-(ii) shows that the private-value quantile function is identified from the bid quantile function and its derivative. It is a quantile version of the identification strategy of GPV, which is based on the identity[1]

$$V_i = B_i + \frac{1}{I-1}\frac{G(B_i|X,I)}{g(B_i|X,I)}. \qquad (2.6)$$

Versions of (2.5) with $B^{(1)}(\alpha|X,I)$ changed into $1/g(B(\alpha|X,I)|X,I)$ can be found in Milgrom (2001, Theorem 4.7), Liu and Luo (2014), Liu and Vuong (2016), Luo and Wan (2016), Enache and Florens (2017) and, under risk-aversion, in Guerre et al. (2009) and Campo et al. (2011).

2.2 Private-value quantile regression

The linearity of (2.4) has important model stability implications useful for practical implementation. Consider a private-value quantile given by the quantile-regression specification

$$V(\alpha|X,I) = \gamma_0(\alpha|I) + X'\gamma_1(\alpha|I) = X_1'\gamma(\alpha|I), \quad X_1 = [1, X']'. \qquad (2.7)$$

As a linear regression is often viewed as an alternative to a nonparametric one which is difficult to estimate, this quantile regression is simpler to estimate than a general quantile function which must be estimated nonparametrically. As pointed by a Referee, the quantile

[1] This can be recovered from (2.5) taking $\alpha = A_i$ as $V_i = V(A_i|X,I)$, $B_i = B(A_i|X,I)$ implying that $A_i = G(A_i|X,I)$ and $B^{(1)}(A_i|X,I) = 1/g(B(A_i|X,I)|X,I) = 1/g(B_i|X,I)$.

level α can be viewed as a measure of the bidder efficiency and the slope function $\gamma(\cdot|I)$ indicates how this efficiency affects valuation in the covariate dimension. While more flexible than the homogenized bid specification detailed in Section 2.3.1, the quantile approach only involves a unique signal: in particular, if each slope entries are increasing, then each covariate contribution to the value increases with efficiency. More flexibility is possible with the random coefficient model of Section 2.3.2, which attaches a specific signal to each auction covariate.

Proposition 1-(i) implies that the conditional bid quantile function satisfies,

$$B(\alpha|X, I) = X_1'\beta(\alpha|I) \text{ with } \beta(\alpha|I) = \frac{I-1}{\alpha^{I-1}} \int_0^\alpha a^{I-2}\gamma(a|I)\,da, \tag{2.8}$$

showing that $B(\alpha|X, I)$ belongs to the quantile-regression specification. Hence (2.5) gives

$$\gamma(\alpha|I) = \beta(\alpha|I) + \frac{\alpha\beta^{(1)}(\alpha|I)}{I-1}, \tag{2.9}$$

so that estimating $\gamma(\alpha|I)$ amounts to estimate $\beta(\alpha|I)$ and $\beta^{(1)}(\alpha|I)$.

This approach extends to more flexible nonparametric linear specifications, as developed in Appendix A which considers a sieve extension

$$\begin{cases} V(\alpha|x, I) = P(x)'\gamma(\alpha|I) + \text{approx. error}, \\ B(\alpha|x, I) = P(x)'\frac{I-1}{\alpha^{I-1}}\int_0^\alpha a^{I-2}\gamma(a|I)da + \text{approx. error} \end{cases} \tag{2.10}$$

where $P(\cdot)$ is a localized sieve vector whose dimension grows with a smoothing parameter h. The choice of $P(\cdot)$ can be tailored to cover additivity or less stringent interaction restrictions. As for the quantile-regression estimators proposed below, the sieve approach developed in Appendix A is not affected by asymptotic boundary issues.

2.3 Alternative models and specification testing strategies

2.3.1 Homogenized bids

HHS and Rezende (2008) consider a regression specification

$$V_i = X'\gamma_1 + v_i, \quad i = 1, \ldots, I, \tag{2.11}$$

where the iid v_i, the "homogenized" private values, are independent of X and not centered.[2] The corresponding homogenized-bid quantile-regression specification is the following restriction of (2.7)

$$V(\alpha|X, I) = X'\gamma_1 + v(\alpha|I)$$

where $v(\cdot|I)$ is the quantile function of the v_i's. Since $\frac{I-1}{\alpha^{I-1}}\int_0^\alpha a^{I-2}da = 1$, it follows that the associated bid quantile function is, by (2.4)

$$B(\alpha|X, I) = X'\gamma_1 + b(\alpha|I), \text{ where } b(\alpha|I) = \frac{I-1}{\alpha^{I-1}}\int_0^\alpha a^{I-2}v(a|I)\,da.$$

This gives the bid regression model

$$B_i = X'\gamma_1 + b_i, \quad b_i = b(A_i|I) \text{ and } i = 1, \ldots, I \tag{2.12}$$

where the b_i are the homogenized bids of HHS, which are independent of X but depend upon I. Given a sample $X_\ell, I_\ell, B_{1\ell}, \ldots, B_{I_\ell \ell}$ of $\ell = 1, \ldots, L$ first-price auctions, HSS and Rezende (2008) propose to backup the homogenized bids by regressing the bids $B_{i\ell}$ on $X_{1\ell} = [1, X'_\ell]'$, so that the estimation of the homogenized bids are $\widehat{b}_{i\ell} = B_{i\ell} - X'_\ell\widehat{\gamma}_1$, where $\widehat{\gamma}_1$ is the OLS slope estimator. The pdf of v_i can be estimated applying the GPV two-step method to the homogenized-bid estimates. An important feature of this model is that the dependence of the private values to the covariate is simple enough to allow for accurate estimation of

[2] Centering the v_i's would amount to introduce an intercept parameter γ_0, which would be changed to a new intercept $\beta_0(I)$ when turning to the bid regression when the v_i's are independent of I. In contrast, the bid regression slope is γ_1, therefore unchanged. Hence (2.11) does not include an intercept to better focus on the invariant parameter, the purpose being to estimate γ_1 and the distribution of v_i. Estimating an intercept in the bid regression is however necessary to consistently estimate γ_1 using OLS because the regression error term in (2.12) is not centered.

γ_1. As noted in Paarsch and Hong (2006), a similar two-step procedure applies for the nonparametric regression model $V_i = m(X|I) + v_i$ where the v_i's are independent of X, I, see also Marmer, Shneyerov and Xu (2013b).

However this approach requests independence between the regression error term v_i and the covariate X, an assumption which may be too restrictive in practice as found by Gimenes (2017) and the application below. When $\gamma_1(\cdot)$ is not a constant and $V(\alpha|X,I) = X'\gamma_1(\alpha|I) + v(\alpha|I)$, it holds for $\beta_1(\alpha|I) = \frac{I-1}{\alpha^{I-2}}\int_0^\alpha a^{I-2}\gamma_1(a|I)da$ and the OLS limit $\beta_1(I) = \mathbb{E}[\beta_1(A_i|I)]$ obtained when regressing the bids on the constant and X,

$$B_i = X'\beta_1(I) + b(A_i|X,I) \text{ where } b(A_i|X,I) = b(A_i|I) + X'\left[\beta_1(\alpha|I) - \beta_1(I)\right].$$

As $b(A_i|X,I)$ depends upon X, the homogenized-bid approach does not apply. As explained below, estimating the slope $\gamma_1(\cdot)$ involves nonparametric techniques that cannot deliver the parametric rate feasible in the homogenized-bid model.

2.3.2 Random-coefficient specification

Consider I private values from

$$V_i = X_1'\Gamma_i, \quad i = 1, \ldots, I \tag{2.13}$$

where the random coefficients Γ_i are iid $1 \times (D+1)$ vectors independent of X_1. Since $V_i = \Gamma_{0i} + X'\Gamma_{1i}$, taking Γ_{1i} constant across bidders gives a homogenized-bid specification, which is therefore a particular case of random-coefficients regression. Compared to (2.1) version which involves a unique signal A_i, (2.13) allows for $D+1$ individual signals Γ_{id} which models the impact of the common covariate X_d on the private value V_i. How a quantile approach can be useful is first discussed when Γ_i is drawn from an elliptical distribution.

Elliptical random coefficient. Γ_i is drawn from an elliptical distribution with translation parameter $\gamma(I)$ and symmetric nonnegative dispersion matrix $\Sigma_\Gamma(I)$ if the characteristic function $\mathbb{E}\left[\exp\left(\mathrm{i}t'(\Gamma_i - \gamma(I))\right)\right]$ only depends upon $t'\Sigma_\Gamma(I)t$. Examples include the multivariate normal, lognormal or Student distribution, which can be truncated to satisfy our finite support restriction. A convenient representation of Γ_i involves the Euclidean norm $R_i = \left\| \Sigma_\Gamma^{-1/2}(I)\left(\Gamma_i - \gamma(I)\right)\right\|$ and independent draws $\mathcal{S}_i$ from the uniform distribution over the $D+1$ dimensional unit sphere. Let $\mathcal{C}_i$ be the first coordinate of $\mathcal{S}_i$, noticing that $t'\mathcal{S}_i$ is distributed as $\|t\|\mathcal{C}_i$ for any $(D+1)\times 1$ vector t. Then by Fang, Kotz and Ng (2017, p.29-32), Γ_i and $\gamma(I) + R_i \Sigma_\Gamma^{1/2}(I)\mathcal{S}_i$ have the same distribution, for independent R_i and $\mathcal{S}_i$. It then follows by (2.13), $\stackrel{d}{=}$ indicating random variables with identical distribution

$$V_i \stackrel{d}{=} X_1'\gamma(I) + \left(\Sigma_\Gamma^{1/2}(I)X_1\right)' R_i \mathcal{S}_i \stackrel{d}{=} X_1'\gamma(I) + \left\|\Sigma_\Gamma^{1/2}(I)X_1\right\| R_i\mathcal{C}_i.$$

Hence the quantile specification generated by (2.13) is

$$V(\alpha|X,I) = X_1'\gamma(I) + \left\|\Sigma_\Gamma^{1/2}(I)X_1\right\| v(\alpha|I) \tag{2.14}$$

where the unknown quantile function $v(\alpha|I)$ is the one of $R_i\mathcal{C}_i$ given I. The generated bids have a common quantile function

$$B(\alpha|X,I) = X_1'\gamma(I) + \left\|\Sigma_\Gamma^{1/2}(I)X_1\right\| b(\alpha|I), \quad b(\alpha|I) = \frac{I-1}{\alpha^{I-1}}\int_0^\alpha a^{I-2} v(a|I)da \tag{2.15}$$

by (2.4). Using the normalization $b(1/2|I) = 1$ for identification purpose gives that $B(1/2|X,I) = X_1'\gamma(I) + \left\|\Sigma_\Gamma^{1/2}(I)X_1\right\|$, so that the conditional bid median can be used to identify $\gamma(I)$ and $\Sigma_\Gamma(I)$. Identification of $v(\cdot|I)$ works as in Proposition 1 as $v(\alpha|I) = b(\alpha|I) + \alpha b^{(1)}(\alpha|I)/(I-1)$, observing that $v(\cdot|I)$ identifies the common distribution of the R_i's.

The general case. Hoderlein et al. (2010) propose a nonparametric method that could be used to estimate the distribution of the random slope Γ_i of (2.13) if the private values were observed. This suggests to implement a two-step method using estimated private values. Appendix A proposes a sieve method to estimate $V(\cdot|x,I)$, which is not subject to asymptotic boundary bias. Consider S estimated private values $\widehat{V}(A_s|X_s,I)$ for arbitrary values X_s of the covariate and independent uniform draws A_s, $s=1,\ldots,S$. Assuming that the $X_s/\|X_s\|$ are drawn from the uniform distribution on the unit sphere suggests to estimate the density

$f_\Gamma(\gamma|I)$ of Γ_i given I using in the second step the Hoderlein et al (2010) kernel estimator

$$\widehat{f}_\Gamma(\gamma|I) = \frac{1}{n}\sum_{s=1}^{S} K_{HKM,h}\left(\frac{\widehat{V}(A_s|X_s, I) - X_s'\gamma}{\|X_s\|}\right) \text{ with} \tag{2.16}$$

$$K_{HKM,h}(u) = \frac{1}{(2\pi)^{D+1}}\int_0^{1/h} \cos(tu)t^D\left(1-(ht)^r\right)dt,$$

where $h > 0$ is a bandwidth parameter and $0 < r \leq \infty$.

2.4 Specification testing strategies

The stability of private-value quantile-regression specification allows to use the bid one to test many hypothesis of interest, see Liu and Luo (2017) for a related point of view. This can be useful to obtain better performing tests as the presence of the derivative $\widehat{B}^{(1)}(\alpha|x, I)$ in the implementable private-value expression (2.5) makes its use for testing harder. Examples of tests based on this idea are as follows.

Quantile-regression goodness of fit. There is a recent literature that considers the null hypothesis of correct specification of a quantile-regression model over an subinterval $\mathcal{A}$ of $(0, 1)$. See Escanciano and Velasco (2010), Rothe and Wied (2013), Escanciano and Goh (2014) and the references therein. These three papers propose test statistics of the form $\widehat{T}(\widehat{\beta}(\cdot|I))$, where $\widehat{\beta}(\cdot|I)$ is a quantile-regression estimator which converges to the true slope over $\mathcal{A}$ with a parametric rate, such as the standard quantile-regression estimator or the augmented ones proposed in Section 4. See the Application Section 7 for the $\widehat{T}(\cdot)$ used by Rothe and Wied (2013). Liu and Luo (2017) based an entry exogeneity test on the integral of the squared difference of two quantile estimators, see (2.17) below.

Homogenized bid and elliptical random coefficient. The correct specification of (2.11) or (2.15) can be tested using Rothe and Wied (2013) without any restriction on the quantile alternative. If the alternative is restricted to a quantile-regression model, Koenker and Xiao (2002) or Escanciano and Goh (2014) can be used to test the homogenized-bid null hypothesis, as this specification coincides with the location-shift model considered by these authors. Following Liu and Luo (2017) suggests to consider, for the same null, a test statistic

$$L \int_0^1 \sum_{\ell=1}^{L} \left(X'_{1\ell} \widehat{\beta}_{H_0}(\alpha) - X'_{1\ell} \widehat{\beta}(\alpha) \right)^2 d\alpha \tag{2.17}$$

where L is the number of auctions in the sample, X_ℓ the auction covariate, $\widehat{\beta}(\cdot)$ a quantile-regression slope estimator $\sqrt{L}$-consistent over $[0,1]$ as the one proposed in the next section, and for instance $\widehat{\beta}_{H_0}(\cdot) = [\widehat{\beta}_0(\cdot), \widehat{\beta}_{1,OLS}, \ldots, \widehat{\beta}_{1D,OLS}]'$. Confidence bands can also be used, see Gimenes (2017) and the theory developed in Fan, Guerre and Lazarova (2020).

Exogenous auction format. Let $V_j(\alpha|x, I) = x'\gamma_j(\alpha|I)$ be the private-value quantile function conditionally on participation to an ascending auction ($j = asc$) or a first-price one ($j = fp$). A null hypothesis of interest is exogeneity of the auction format, H_0^F : $V_{fp}(\cdot|\cdot, I) = V_{asc}(\cdot|\cdot, I)$. Gimenes (2017) gives a consistent quantile-regression estimator $\widehat{\gamma}_{asc}(\cdot|I)$ of $\gamma_{asc}(\cdot|I)$ using ascending auction data. It then follows by (2.4) that $\widehat{\beta}_{H_0}(\alpha|I) = (I-1)\alpha^{-(I-1)} \int_0^\alpha a^{I-2} \widehat{\gamma}_{asc}(a|I) da$ is consistent under the null but not the alternative.[3] Then using first-price auction data to compute a test statistic $\widehat{T}(\widehat{\beta}_{H_0}(\cdot|I))$ from Escanciano and Goh (2014) or Rothe and Wied (2013) for an arbitrary alternative, or using Liu and Luo (2017) statistic (2.17) with a quantile-regression alternative, allow to test for auction format exogeneity.

Participation exogeneity. The participation exogeneity null hypothesis states that the private values are independent of the number of bidder conditionally on the covariate H_0^E : $V(\cdot|\cdot, I) = V(\cdot|\cdot)$ for all I, see also Gimenes (2017) for the ascending auction case. Liu and Luo (2017) use an integral version of H_0^E to eliminate the bid quantile derivative in (2.5).

[3] As the standard quantile-regression estimator may not be well-defined for quantile levels near 0, it may be more suitable to use an augmented quantile-regression estimator as in Section 4 to implement Gimenes (2017).

In a quantile-regression setup, Proposition 1 implies under H_0^E,

$$\beta(\alpha|I_2) = \frac{I_2-1}{\alpha^{I_2-1}}\int_0^{\alpha} a^{I_2-2}\left[\beta(a|I_1) + \frac{a\beta^{(1)}(a|I_1)}{I_1-1}\right] da = \beta_{I_1}(\alpha|I_2) \tag{2.18}$$
$$\text{where } \beta_{I_1}(\alpha|I_2) = \frac{I_2-1}{I_1-1}\beta(\alpha|I_1) + \frac{(I_2-1)(I_1-I_2)}{(I_1-1)\alpha^{I_2-1}}\int_0^{\alpha} a^{I_2-2}\beta(a|I_1)da.$$

Then tests for entry exogeneity can be obtained using the same construction than for the auction format exogeneity null, using a sample of first-price auction with I_1 bidders to estimate $\beta_{I_1}(\alpha|I_2)$ and another sample with I_2 bidders to compute a test statistic.

Under participation exogeneity, private value estimates can be averaged over I to improve accuracy. Another important motivation for exogenous participation is risk-aversion estimation, see Guerre, et al. (2009). This approach can be modified to cope with an additional risk-aversion parameter which can be estimated with a parametric rate as shown in Section 5.3.

3 Risk-aversion, expected payoff and other functionals

Many auction parameters of interest can be written using the private-value quantile function or, by (2.5), the bid quantile function and its quantile derivative. We focus here on the conditional and unconditional integral functionals

$$\theta(x) = \int_0^1 \mathcal{F}\left[\alpha, x, B(\alpha|x,I), B^{(1)}(\alpha|x,I); I \in \mathcal{I}\right] d\alpha, \quad \theta = \int_{\mathcal{X}} \theta(x)\, dx \tag{3.1}$$

where $\mathcal{F}(\alpha, x, b_{0I}, b_{1I}; I \in \mathcal{I})$ is a real valued continuous function. Three illustrative examples are as follows.

Example 1: CRRA parameter. For symmetric risk-averse bidders with a concave utility function, the best-response condition (2.3) becomes

$$\frac{\partial}{\partial a}\left\{U\left(V(\alpha|X,I) - B(a|X,I)\right)a^{I-1}\right\}\Big|_{a=\alpha} = 0.$$

Rearranging as in Guerre et al. (2009) yields that $V(\alpha|X,I) = B(\alpha|X,I) + \lambda^{-1}\left(\frac{\alpha B^{(1)}(\alpha|X,I)}{I-1}\right)$ where $\lambda(\cdot) = U(\cdot)/U'(\cdot)$. For risk-averse bidders with a CRRA utility function $U(t) = t^{\nu}$, arguing as for Proposition 1 shows

$$V(\alpha|X,I) = B(\alpha|X,I) + \nu \frac{\alpha B^{(1)}(\alpha|X,I)}{I-1}, \qquad (3.2)$$
$$B(\alpha|X,I) = \frac{\frac{I-1}{\nu}}{\alpha^{\frac{I-1}{\nu}}} \int_0^{\alpha} a^{\frac{I-1}{\nu}-1} V(a|X,I)\, da.$$

These two formulas show that the stability implications of Proposition 1 for linear private-value and bid quantile functions are preserved under CRRA. Assuming as in Guerre et al. (2009) that the number of bidders is exogenous, i.e $V(\alpha|X,I) = V(\alpha|X)$ for all I, gives that the risk-aversion ν satisfies, for any pair $I_0 \neq I_1$

$$\nu = \frac{\theta_n}{\theta_d} = \frac{\int_{\mathcal{X}} \left[\int_0^1 \left(B(\alpha|x,I_1) - B(\alpha|x,I_0)\right)\left(\frac{\alpha B^{(1)}(\alpha|x,I_0)}{I_0-1} - \frac{\alpha B^{(1)}(\alpha|x,I_1)}{I_1-1}\right) d\alpha\right] dx}{\int_{\mathcal{X}} \left[\int_0^1 \left(\frac{\alpha B^{(1)}(\alpha|x,I_0)}{I_0-1} - \frac{\alpha B^{(1)}(\alpha|x,I_1)}{I_1-1}\right)^2 d\alpha\right] dx}, \qquad (3.3)$$

which gives identification of ν. Following Lu and Perrigne (2008), the risk-aversion parameter ν can also be identified combining ascending and first-price auctions data. As seen from Gimenes (2017), the private-value quantile function $V_{asc}(\alpha|X,I)$ can be easily estimated from ascending auctions. Equating $V_{asc}(\alpha|X,I)$ to $V(\alpha|X,I)$ in (3.2) gives that ν satisfies

$$\nu = \frac{\int_{\mathcal{X}} \left[\int_0^1 \left(V_{asc}(\alpha|x,I) - B(\alpha|x,I)\right) \frac{\alpha B^{(1)}(\alpha|x,I)}{I-1} d\alpha\right] dx}{\int_{\mathcal{X}} \left[\int_0^1 \left(\frac{\alpha B^{(1)}(\alpha|x,I)}{I-1}\right)^2 d\alpha\right] dx}. \qquad (3.4)$$

Example 2: Expected revenue. Suppose that the seller decides to reject bids lower than a reserve price R and let $\alpha_R = \alpha_R(X,I)$ be the associated screening level, i.e. $\alpha_R =$

$F(R|X,I)$. For CRRA bidders, the first-price auction seller expected revenue is[4]

$$ER_\nu(\alpha_R|X,I) = \nu \cdot I \cdot V(\alpha_R|X,I) \cdot \alpha_R^{\frac{I-1}{\nu}} \cdot \frac{1-\alpha_R^{\frac{(I-1)(\nu-1)+\nu}{\nu}}}{(I-1)(\nu-1)+\nu} + \frac{I(I-1)}{(I-1)(\nu-1)+\nu} \int_{\alpha_R}^{1} a^{\frac{I-1}{\nu}-1} \frac{1-a^{\frac{(I-1)(\nu-1)+\nu}{\nu}}}{(I-1)(\nu-1)+\nu} V(a|X,I)\, da. \quad (3.5)$$

This expression includes an integral item

$$\theta(X;\alpha_R) = \int_{\alpha_R}^{1} a^{\frac{I-1}{\nu}-1} \frac{1-a^{\frac{(I-1)(\nu-1)+\nu}{\nu}}}{(I-1)(\nu-1)+\nu} V(a|X,I)\, da$$

which can be estimated by plugging in a risk-aversion estimator $\widehat{\nu}$ and an estimator $\widehat{V}(\alpha|X,I)$ of the private-value quantile function, or estimators of the bid quantile function and its derivative by (2.5).[5]

Example 3: Private-value distribution Additional examples of conditional parameter $\theta(\cdot)$ are the private-value conditional cdf and pdf. Note first that (2.1) shows that the conditional private-value cdf is an integral functional of the private-value quantile function

$$F(v|X,I) = \mathbb{E}\left[\mathbb{I}\left[V(A|x,I) \leq v\right] \middle| X,I\right] = \int_0^1 \mathbb{I}\left[V(\alpha|X,I) \leq v\right] d\alpha, \quad A \sim \mathcal{U}_{[0,1]}. \quad (3.6)$$

Dette and Volgushev (2008) have considered a smoothed version $\mathbb{I}_\eta(\cdot)$ of the indicator function

$$F_\eta(v|X,I) = \int_0^1 \mathbb{I}_\eta\left[v - V(\alpha|X,I)\right] d\alpha$$

[4]In the formula below, $\frac{1-a^{\frac{(I-1)(\nu-1)+\nu}{\nu}}}{(I-1)(\nu-1)+\nu}$ is set to its limit $-\frac{I}{I-1}\log a$ when $(I-1)(\nu-1)+\nu$ vanishes, i.e. when $\nu = \frac{I-1}{I}$. It is assumed for the sake of brevity that the seller value for the good is 0. The expected revenue formula for the general case follows from Gimenes (2017).

[5]Under risk-neutrality, integrating by parts gives that

$$\int_{\alpha_R}^{1} B^{(1)}(\alpha|X,I)\,\alpha^{I-1}(1-\alpha)\, d\alpha = B(\alpha_R|X,I)\,\alpha_R^{I-1}(1-\alpha_R) - \int_{\alpha_R}^{1} B(\alpha|X,I)\,\alpha^{I-1}(I-1-I\alpha)\, d\alpha,$$

estimation of $\theta(X;\alpha_R)$ can also be done using only a bid quantile estimator.

where $\mathbb{I}_\eta(t) = \int_{-\infty}^{t/\eta} k(u)\,du$, $k(\cdot)$ being a kernel function and η a bandwidth parameter. Differentiating $F_\eta(v|X,I)$ gives

$$f_\eta(v|X,I) = \frac{1}{\eta}\int_0^1 k\left(\frac{v - V(\alpha|X,I)}{\eta}\right) d\alpha$$

which converges to the private-value pdf when η goes to 0. Note that $F_\eta(v|X,I)$ and $f_\eta(v|X,I)$ can be estimated by plugging in an estimator $\widehat{V}(\alpha|X,I)$ of $V(\alpha|X,I)$. The resulting cdf and pdf estimators inherit of the dimension reduction property of $\widehat{V}(\alpha|X,I)$. As the private-value estimator proposed in the next section is consistent over the whole $[0,1]$, no boundary trimming is needed. This contrasts with the GPV pdf estimator. As noted by Escanciano and Goh (2019) in a general context, the integral in $f_\eta(v|X,I)$ can be replaced by a sample average over iid uniform draws A_s, as used for the density estimator (2.16).

4 Augmented quantile-regression estimation

Proposition 1 suggests to base the estimation of the private-value quantile function on estimations of $B(\alpha|x,I)$ and of its derivative $B^{(1)}(\alpha|x,I)$ with respect to α. The *augmented* methodology applies local polynomial expansion with respect to α for joint estimation of $B(\alpha|x,I)$ and $B^{(1)}(\alpha|x,I)$. To ensure comparability with the auction literature which considers private-value pdf having s continuous derivatives, we assume that the private-value quantile function $V(\alpha|x,I)$ has $s+1$ continuous derivatives with respect to α. As seen from (2.4), this implies that the bid quantile function $B(\alpha|x,I)$ has $s+2$ continuous derivatives with respect to $\alpha > 0$. Let $(X_\ell, I_\ell, B_{1\ell}, \ldots, B_{I_\ell \ell})$, $\ell = 1, \ldots, L$, be an iid first-price auction sample with I_ℓ bids $B_{i\ell}$ and good characteristics X_ℓ.

4.1 Augmented quantile estimation without covariate

Estimation. Assume first that $V(\alpha|X,I) = V(\alpha|I)$ so that $B(\alpha|X,I) = B(\alpha|I)$. Let $\rho_\alpha(\cdot)$ be the check function

$$\rho_\alpha(q) = q\left(\alpha - \mathbb{I}(q \le 0)\right).$$

It is well known that

$$B\left(\alpha|I\right)=\arg\min_{q}\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\rho_{\alpha}\left(B_{i\ell}-q\right)\right],\quad \alpha\in\left(0,1\right).$$

We now exhibit a functional objective function which achieves its minimum at the restriction of $B(\cdot|I)$ over $[\alpha-h,\alpha+h]\cap[0,1]$. It easily follows that, for a non negative kernel function $K\left(\cdot\right)$ with support $[-1,1]$ and a positive bandwidth $h=h_L$,

$$\begin{aligned}&\left\{B\left(\tau|I\right),\tau\in[\alpha-h,\alpha+h]\cap[0,1]\right\}\\&=\arg\min_{q(\cdot)}\int_0^1\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\rho_{a}\left(B_{i\ell}-q\left(a\right)\right)\right]\frac{1}{h}K\left(\frac{a-\alpha}{h}\right)da,\end{aligned}\tag{4.1}$$

where the minimization is performed over the set of functions $q\left(\cdot\right)$ over $[\alpha-h,\alpha+h]\cap[0,1]$. This can be used to estimate the derivative $B^{(1)}\left(\alpha|I\right)$, using minimization over Taylor polynomial of order $s+1$ instead of $q(\cdot)$. A Taylor expansion of order $s+1$ gives

$$\begin{aligned}B\left(a|I\right)&=B\left(\alpha|I\right)+B^{(1)}\left(\alpha|I\right)\left(a-\alpha\right)+\cdots+\frac{B^{(s+1)}\left(\alpha|I\right)}{(s+1)!}\left(a-\alpha\right)^{s+1}+O\left(h^{s+2}\right)\\&=\pi\left(a-\alpha\right)'b(\alpha|I)+O\left(h^{s+2}\right)\text{ where}\\b(\alpha|I)&=\left[B\left(\alpha|I\right),\ldots,B^{(s+1)}(\alpha|I)\right]'\text{ and }\pi\left(t\right)=\left[1,t,\frac{t^2}{2}\ldots,\frac{t^{s+1}}{(s+1)!}\right]'.\end{aligned}$$

The $(s+2)\times1$ vector $b(\cdot|I)$ stacks the successive bid quantile derivatives, and is the parameter to be estimated. Let $b=[\beta_0,\ldots,\beta_{s+1}]'\in\mathbb{R}^{s+2}$ be the generic coefficients of such a Taylor polynomial function. The sample version of the objective function (4.1), restricted to local polynomial functions $\pi(\cdot)'b$ instead of $q(\cdot)$, is

$$\begin{aligned}\widehat{\mathcal{R}}\left(b;\alpha,I\right)&=\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_{\ell}=I\right)\sum_{i=1}^{I_{\ell}}\int_0^1\rho_{a}\left(B_{i\ell}-\pi\left(a-\alpha\right)'b\right)\frac{1}{h}K\left(\frac{a-\alpha}{h}\right)da\\&=\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_{\ell}=I\right)\sum_{i=1}^{I_{\ell}}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\rho_{\alpha+ht}\left(B_{i\ell}-\pi\left(ht\right)'b\right)K\left(t\right)dt.\end{aligned}$$

The *augmented quantile* estimator is $\widehat{b}\,(\alpha|I) = \arg\min_{b\in\mathbb{R}^{s+2}} \widehat{\mathcal{R}}\,(b;\alpha,I)$, $\widehat{\beta}_0\,(\alpha|I)$ and $\widehat{\beta}_1\,(\alpha|I)$ being estimators of $B\,(\alpha|I)$ and its first derivative $B^{(1)}\,(\alpha|I)$, respectively.

Homogenized bid and elliptical random coefficients. A two-step version of the augmented method presented above can be used to estimate the homogenized private-value quantile function $v(\cdot|I)$ from (2.11). Regressing $B_{i\ell}$ on X_ℓ and an intercept for those auctions with $I_\ell = I$ gives a consistent estimator $\widehat{\gamma}_1(I)$ of $\gamma_1(I)$. Let $\widehat{B}_{i\ell} = B_{i\ell} - X_\ell'\widehat{\gamma}_1$ be the estimated homogenized bids. Then replacing $B_{i\ell}$ with $\widehat{B}_{i\ell}$ in the objective function $\widehat{\mathcal{R}}\,(b;\alpha,I)$ gives estimators $\check{\beta}(\cdot|I)$ and $\check{\beta}_1(\cdot|I)$ of the homogenized-bid quantile function and of its first derivative. The resulting estimator of the private-value quantile function is then

$$\widehat{V}(\alpha|X,I) = X'\widehat{\gamma}_1(I) + \check{\beta}(\alpha|I) + \frac{\alpha\check{\beta}_1(\alpha|I)}{I-1}.$$

The elliptical random-coefficient quantile specification (2.15) can be estimated similarly. Studying the asymptotic properties of this two-step procedures is outside the scope of this paper. Bhattacharya (2019) considers a related two-step procedure that can be useful for ascending auctions, where estimating quantile derivative is not needed.

4.2 Augmented quantile-regression

An extension of this procedure is the *augmented quantile-regression* estimator, AQR hereafter, which assumes $V\,(\alpha|x,I) = x_1'\gamma\,(\alpha|I)$, recalling $x_1 = [1,x']'$. Proposition 1-(i) then gives $B\,(\alpha|x,I) = x_1'\beta\,(\alpha|I)$. Define now

$$P\,(x,t) = \pi\,(t)\otimes x_1 = \left[1, x', t, t\cdot x', \ldots, \frac{t^{s+1}}{(s+1)!}, \frac{t^{s+1}}{(s+1)!}\cdot x'\right]' \in \mathbb{R}^{(s+2)(D+1)} \qquad (4.2)$$

$$b(\alpha|I) = \left[\beta(\alpha|I)', \beta^{(1)}(\alpha|I)', \ldots, \beta^{(s+1)}(\alpha|I)'\right]$$

so that the Taylor expansion of $B\,(\alpha|X,I)$ writes

$$B\,(\alpha+ht|x,I) = \sum_{k=0}^{s+1} x_1'\beta^{(k)}(\alpha|I)\frac{(ht)^k}{k!} + O\left(h^{s+2}\right) = P\,(x,ht)'\,b\,(\alpha|I) + O\left(h^{s+2}\right).$$

The corresponding generic parameter is the $(s+2)(D+1)\times 1$ column vector $b = \left[\beta_0', \beta_1', \ldots, \beta_{s+1}'\right]'$ where the β_j are all of dimension $D+1$, and the objective function becomes

$$\begin{aligned}\widehat{\mathcal{R}}\left(b;\alpha,I\right) &= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_\ell = I\right)\sum_{i=1}^{I_\ell}\int_0^1 \rho_a\left(B_{i\ell} - P\left(X_\ell, a-\alpha\right)' b\right)\frac{1}{h}K\left(\frac{a-\alpha}{h}\right)da \\ &= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_\ell = I\right)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell} - P\left(X_\ell, ht\right)' b\right)K\left(t\right)da \end{aligned} \quad (4.3)$$

which accounts for the covariate X_ℓ. The estimation of $b\left(\alpha|I\right)$ is

$$\widehat{b}\left(\alpha|I\right) = \arg\min_{b\in\mathbb{R}^{(s+2)(D+1)}} \widehat{\mathcal{R}}\left(b;\alpha,I\right)$$

and the private-value quantile-regression estimator is

$$\widehat{V}\left(\alpha|x,I\right) = x_1'\widehat{\gamma}\left(\alpha|I\right) \text{ with } \widehat{\gamma}\left(\alpha|I\right) = \widehat{\beta}_0\left(\alpha|I\right) + \frac{\alpha\widehat{\beta}_1\left(\alpha|I\right)}{I-1}.$$

The bid quantile function and its derivatives can be estimated using $\widehat{B}\left(\alpha|x,I\right) = x_1'\widehat{\beta}_0\left(\alpha|I\right)$ and $\widehat{B}^{(1)}\left(\alpha|x,I\right) = x_1'\widehat{\beta}_1\left(\alpha|I\right)$, so that $\widehat{V}(\alpha|x,I) = \widehat{B}(\alpha|x,I) + \frac{\alpha\widehat{B}^{(1)}(\alpha|x,I)}{I-1}$. The rearrangement method of Chernozhukov, Fernández-Val and Gallichon (2010) can be used to obtain increasing quantile estimators.

AQR estimator properties. Bassett and Koenker (1982) report that standard quantile-regression estimators are not defined near the extreme quantile levels $\alpha = 0$ or $\alpha = 1$, mostly because the associated objective function has some flat parts. The AQR is better behaved because the objective function $\widehat{\mathcal{R}}\left(b;\alpha,I\right)$ averages the check function $\rho_a\left(\cdot\right)$ for quantile levels a in $[\alpha-h,\alpha+h]\cap[0,1]$, ensuring that the AQR objective function is not flat for extreme quantile levels, as illustrated in Figure 1.[6]

Therefore the AQR estimator is easier to define for the extreme quantile levels $\alpha = 0$ and $\alpha = 1$ than the standard quantile-regression estimator. This is especially relevant for estimating auction models as the winner is expected to belong to the upper tail as soon

[6] This averaging effect requests that $t \mapsto P\left(X_\ell, ht\right)' b$ is not constant meaning that the derivative components of b should not vanish.

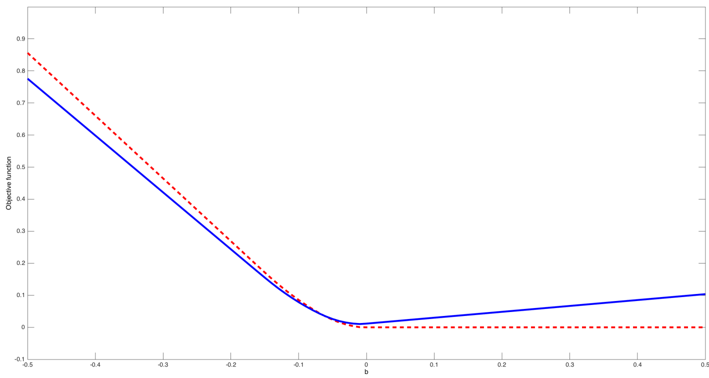

Figure 1: A path of the AQR objective function $\widehat{\mathcal{R}}(b\cdot;1,I)$ (solid line) and of the objective function of the standard quantile-regression estimator (dotted line) when b varies in the direction $[1,\ldots,1]'$.

as the number of bidders is large enough. It also follows from the theoretical study of the objective function $\widehat{\mathcal{R}}(\cdot;\cdot,I)$ that the AQR estimator is uniquely defined for all quantile levels with a probability tending to 1. The bid AQR estimator is also smoother than the standard quantile-regression one, see Figure 6 in the Application Section and Appendix C for a formal argument.

5 Main results

Some additional notations are as follows. Let $S_0 = [1,0,\ldots,0]$ and $S_1 = [0,1,0,\ldots,0]$ be $1\times(s+2)$ selection vectors such that $S_0\pi(t) = 1$, $S_1\pi(t) = t$. Let Id_{D+1} be the $(D+1)\times(D+1)$ identity matrix and set $\mathsf{S}_j = S_j \otimes \mathrm{Id}_{D+1}$, $j = 0,1$, so that $\mathsf{S}_0\widehat{b}(\alpha|I) = \widehat{\beta}_0(\alpha|I)$ and $\mathsf{S}_1\widehat{b}(\alpha|I) = \widehat{\beta}_1(\alpha|I)$ are respectively estimators of $\beta(\alpha)$ and its first derivative $\beta^{(1)}(\alpha)$. $\mathrm{Tr}(\cdot)$ is the trace of a square matrix and ∂_u^n stands for $\frac{\partial^n}{\partial u^n}$. For two sequences $\{a_L\}$ and $\{b_L\}$, $a_L \asymp b_L$ means that both $a_L/b_L = O(1)$ and $b_L/a_L = O(1)$. The norm $\|\cdot\|$ is the Euclidean one, i.e. $\|e\| = (e'e)^{1/2}$. For a matrix A, $\|A\| = \sup_{b:\|b\|=1}\|Ab\|$. Convergence in distribution is denoted as '$\xrightarrow{d}$'.

5.1 Main assumptions

Assumption A *(i) The auction variables* $(I_\ell, X_\ell, V_{i\ell}, B_{i\ell}, i = 1, \ldots, I_\ell)$ *are iid across* ℓ*. The support* $\mathcal{X}$ *of* X_ℓ *given* $I_\ell = I$ *is independent of* I*, compact with non empty interior. The support* $\mathcal{I}$ *of* $I_\ell \geq 2$ *is finite. The matrices* $\mathbb{E}\left[\mathbb{I}\left(I_\ell = I\right) X_{1\ell} X_{1\ell}'\right]$*, where* $X_{1\ell} = [1, X_\ell']'$*, have an inverse for all* I *of* $\mathcal{I}$*.*

(ii) Given $(X_\ell, I_\ell) = (x, I)$*, the* $V_{i\ell}$*,* $i = 1, \ldots, I_\ell$ *are iid with a continuously differentiable conditional quantile function* $V(\alpha|x, I)$ *with* $\min_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x, I) > 0$ *and* $\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x, I) < \infty$*.*

Assumption S *For some* $s \geq 1$ *and each* $I \in \mathcal{I}$*,* $V(\alpha|X, I) = X_1'\gamma(\alpha|I)$ *is as in (2.7), where* $X_1 = [1, X']'$ *and* $\gamma(\cdot|I)$ *is* $(s+1)-$*times continuously differentiable over* $[0, 1]$*.*

Assumption H *The kernel function* $K(\cdot)$ *with support* $(-1, 1)$ *is symmetric, continuously differentiable over the straight line, and strictly positive over* $(-1, 1)$*. The positive bandwidth* h *goes to* 0 *with* $\lim_{L\to\infty} \frac{\log^2 L}{Lh^2} = 0$*.*

Assumption F *For all* x *in* $\mathcal{X}$ *and* α *in* $[0, 1]$*, the function* $\mathcal{F}[\alpha, x, b_{0I}, b_{1I}; I \in \mathcal{I}]$ *in (3.1) is twice differentiable with respect to* b_{0I} *and* b_{1I}*,* I *in* $\mathcal{I}$*. For each* I *in* $\mathcal{I}$*, these partial derivatives of order 1 and 2 are continuous with respect to all the variables.*

Assumption A-(i) is standard. Assumption A-(ii) allows for private values depending on the number of bidders. Recall that

$$V^{(1)}(\alpha|x, I) = \frac{1}{f(V(\alpha|x, I)|x, I)}, \tag{5.1}$$

where $f(v|x, I)$ is the conditional private-value pdf. Hence Assumption A-(ii) amounts to assume that $f(v|x, I)$ is bounded away from 0 and infinity on its support $[V(0|x, I), V(1|x, I)]$ as assumed for instance in Riley and Samuelson (1981), Maskin and Riley (1984) or GPV. The condition $0 < f(v|x, I) < \infty$ is also used for asymptotic normality of quantile-regression estimator, see Koenker (2005). Assumption S is a standard smoothness condition which, by (5.1), parallels GPV who assume that the pdf $f(v|x, I)$ is s-times differentiable.

The bandwidth rate in Assumption H is unusual in kernel or local polynomial nonparametric estimation, where rate conditions as $1/(Lh) = o(1)$ are more common. This is due to a key linearization expansion for $\widehat{V}(\alpha|x, I)$, which holds with an $O_{\mathbb{P}}\left(\log L/(L\sqrt{h})\right)$ error term that must go to 0, see (5.3) and (5.5) in Theorem 2 below.[7]

Assumption F holds for most of the examples of functionals above. A notable exception is the cdf $F(v|x, I)$ in Example 3, which involves an indicator function which is not smooth. However it holds for the smoothed approximation $F_{\eta}(v|x, I)$ of the cdf, although Assumption F implicitly rules out vanishing bandwidth η in Example 3.

5.2 Private value quantile estimation results

The next sections give our theoretical results for integrated mean squared error, uniform consistency and asymptotic distribution of the augmented estimator $\widehat{V}(\cdot|\cdot, I)$. These results are derived using a pseudo-true value framework, in which $\widehat{b}(\cdot|I)$ is viewed as an estimator of the minimizer $\overline{b}(\cdot|I)$ of the population counterpart of $\widehat{\mathcal{R}}(b; \alpha, I)$

$$\overline{b}(\alpha|I) = \arg \min_{b \in \mathbb{R}^{(s+2)(D+1)}} \overline{\mathcal{R}}(b; \alpha, I) \text{ where } \overline{\mathcal{R}}(b; \alpha, I) = \mathbb{E}\left[\widehat{\mathcal{R}}(b; \alpha, I)\right]$$

which asymptotic existence and uniqueness is established for the proofs of our main results. Define accordingly $\overline{\beta}_0(\alpha|I) = \mathsf{S}_0\overline{b}(\alpha|I)$ and $\overline{\beta}_1(\alpha|I) = \mathsf{S}_1\overline{b}(\alpha|I)$ and

$$\overline{V}(\alpha|x, I) = x_1'\left(\overline{\beta}_0 + \frac{\alpha\overline{\beta}_1(\alpha|I)}{I-1}\right). \tag{5.2}$$

The difference $\overline{V}(\alpha|x, I) - V(\alpha|x, I)$ can be interpreted as a bias term.

Because $\widehat{V}(\cdot|\cdot, I)$ is defined in an implicit way via the minimization of the objective function (4.3), its asymptotic study relies on a linearization of $\widehat{b}(\alpha|I) - \overline{b}(\alpha|I)$ which, in a quantile setup, is called a Bahadur expansion, see Theorem D.1 in Appendix D and Koenker (2005, Chap. 4). It is shown that, in a vicinity of $\overline{b}(\alpha|I)$, $b \mapsto \widehat{\mathcal{R}}(b; \alpha, I)$ is twice differentiable with a first derivative $\widehat{\mathcal{R}}^{(1)}(b; \alpha, I)$ satisfying $\mathbb{E}\left[\widehat{\mathcal{R}}^{(1)}\left(\overline{b}(\alpha|I); \alpha, I\right)\right] = 0$, and with a

[7] See also Theorem D.1 in Appendix D, where it is shown more specifically that this bandwidth order is needed for the linearization of $\widehat{B}^{(1)}(\alpha|x, I)$.

Hessian matrix $\overline{\mathcal{R}}^{(2)}\left(\bar{b}(\alpha|I);\alpha,I\right)$ which is invertible. The leading term of $\widehat{b}(\alpha|I)-\bar{b}(\alpha|I)$ is $-\left[\overline{\mathcal{R}}^{(2)}\left(\bar{b}(\alpha|I);\alpha,I\right)\right]^{-1}\widehat{\mathcal{R}}^{(1)}\left(\bar{b}(\alpha|I);\alpha,I\right)$ as shown in Theorem D.1, so that the leading term of $\widehat{V}(\alpha|x,I)$ is

$$\widetilde{V}(\alpha|x,I)=\overline{V}(\alpha|x,I)-x_1'\left(\mathsf{S}_0+\frac{\alpha\mathsf{S}_1}{I-1}\right)\left[\overline{\mathcal{R}}^{(2)}\left(\overline{\beta}(\alpha|I);\alpha,I\right)\right]^{-1}\widehat{\mathcal{R}}^{(1)}\left(\overline{\beta}(\alpha|I);\alpha,I\right), \quad (5.3)$$

see (5.5) below. Because direct computations of the moments of $\widehat{V}(\alpha|x,I)$ are difficult due to its implicit definition and to potential nonlinearities, moments of its linear leading term $\widetilde{V}(\alpha|x,I)$ are used as an approximation.

5.2.1 Integrated mean squared error and uniform consistency rates

Let us first introduce some notations for the integrated mean squared error (IMSE). Let $\Pi^1(\alpha)$ be the second column of the inverse of $\int\pi(t)\pi(t)'K(t)\,dt$, i.e.,

$$\Pi^1(\alpha)=\left(\int\pi(t)\pi(t)'K(t)\,dt\right)^{-1}S_1'$$

and consider the variance terms

$$\begin{aligned}
v^2(\alpha)&=\Pi^1(\alpha)'\int\int\pi(t_1)\pi(t_2)'\min(t_1,t_2)K(t_1)K(t_2)\,dt_1dt_2\Pi^1(\alpha),\\
\Sigma(\alpha|I)&=\frac{\alpha^2v^2(\alpha)}{(I-1)^2}\mathbb{E}^{-1}\left[\frac{X_{1\ell}X_{1\ell}'\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|X_\ell,I_\ell)}\right]\mathbb{E}\left[X_{1\ell}X_{1\ell}'\mathbb{I}(I_\ell=I)\right]\mathbb{E}^{-1}\left[\frac{X_{1\ell}X_{1\ell}'\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|X_\ell,I_\ell)}\right],\\
\Sigma_I&=\int_{\mathcal{X}}\int_0^1 x_1'\Sigma(\alpha|I)\,x_1\,d\alpha dx,
\end{aligned}$$

where $\mathbb{E}^{-1}\left[\frac{X_{1\ell}X_{1\ell}'\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|X_\ell,I_\ell)}\right]$ is the inverse matrix of $\mathbb{E}\left[\frac{X_{1\ell}X_{1\ell}'\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|X_\ell,I_\ell)}\right]$. That $v^2(\alpha)$, and then Σ_I, is strictly positive follows from the proof of Theorem 2 below, see in particular Lemma B.6 in Appendix B. The bias, and integrated squared bias, of the estimator are asymptotically proportional to, respectively[8]

$$\mathsf{Bias}(\alpha|x,I) = \frac{\alpha B^{(s+2)}(\alpha|x,I)}{I-1} S_1 \left(\int \pi(t)\pi(t)' K(t)\,dt\right)^{-1} \int \frac{t^{s+2}\pi(t)}{(s+2)!} K(t)\,dt,$$
$$\mathsf{Bias}_I^2 = \int_{\mathcal{X}} \int_0^1 \mathsf{Bias}^2(\alpha|x,I)\,d\alpha dx.$$

The next Theorem deals with the IMSE of $\widehat{V}(\cdot|\cdot,I)$ and with its difference to its linearization $\widetilde{V}(\cdot|\cdot,I)$ in (5.3).

Theorem 2 *Under Assumptions A, H, S and for $\widetilde{V}(\cdot|\cdot,I)$ as in (5.3), it holds for all I in $\mathcal{I}$*

$$\mathbb{E}\left[\int_{\mathcal{X}} \int_0^1 \left(\widetilde{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx\right] = h^{2(s+1)}\mathsf{Bias}_I^2 + \frac{\Sigma_I}{LIh} + o\left(h^{2(s+1)} + \frac{1}{LIh}\right), \tag{5.4}$$

$$\text{with} \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \sqrt{LIh}\left|\widehat{V}(\alpha|x,I) - \widetilde{V}(\alpha|x,I)\right| = O_{\mathbb{P}}\left(\frac{\log L}{h\sqrt{LI}}\right) = o_{\mathbb{P}}(1). \tag{5.5}$$

It also holds, for each I of $\mathcal{I}$,

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|\widehat{V}(\alpha|x,I) - V(\alpha|x,I)\right| = O_{\mathbb{P}}\left(\sqrt{\frac{\log L}{LIh}}\right) + O(h^{s+1}), \tag{5.6}$$

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|\widehat{B}(\alpha|x,I) - B(\alpha|x,I)\right| = O_{\mathbb{P}}\left(\sqrt{\frac{\log L}{LI}}\right) + o(h^{s+1}). \tag{5.7}$$

Theorem 2 gives the IMSE of the linearization $\widetilde{V}(\cdot|\cdot,I)$ of $\widehat{V}(\cdot|\cdot,I)$ in (5.4). Then (5.5) gives the order of the linearization error in a uniform sense, which is negligible with the order $1/\sqrt{Lh} + O(h^{s+1})$ of the squared root IMSE under Assumption H. The linearization result (5.5) requests $\log L/(h\sqrt{LI}) = o(1)$, which is the main motivation for the unusual rate of bandwidth rate of Assumption H. This condition is driven by the linearization of the bid quantile derivative estimator $\widehat{B}^{(1)}(\cdot|\cdot,I)$.

The bias and variance leading terms in the IMSE expansion (5.4) are from the bid quantile

[8]The expression above depends upon the derivative $\beta^{(s+2)}(\alpha|I)$, which exists by (2.4) for all $\alpha \neq 0$ since $\gamma(\cdot|I)$ is $(s+1)$-th differentiable. Proposition A.1-(iii) in Appendix A shows that $\alpha\beta^{(s+2)}(\alpha|I)$ can be defined over the whole quantile interval $[0,1]$ since $\lim_{\alpha\downarrow 0} \alpha\beta^{(s+2)}(\alpha|I) = 0$.

derivative estimator $\alpha\widehat{B}^{(1)}(\alpha|x,I)/(I-1)$. As

$$B^{(1)}(\alpha|x,I) = \frac{1}{g\left[B(\alpha|x,I)\,|x,I\right]},$$

where $g(\cdot|\cdot)$ is the bid conditional pdf, estimation of this item is similar to estimating a pdf. The rate $1/Lh$ of the variance term $\Sigma_I/(LIh)$ is the rate of a kernel density estimator in the absence of covariate. This is due to the quantile-regression specification. Compared to GPV density estimation rate $1/\sqrt{Lh^{D+1}}$, the rate of the AQR estimator does not suffer from the curse of dimensionality. The order $O(h^{s+1})$ of the bias is given by (2.4), implying that $B^{(1)}(\alpha|x,I)$ has as many derivatives as $V(\alpha|x,I)$, hence the exponent $s+1$.

Minimizing the leading term of the IMSE expansion (5.4) yields the optimal bandwidth

$$h_* = \left(\frac{\Sigma_I}{2(s+1)\,\mathsf{Bias}_I^2}\frac{1}{LI}\right)^{\frac{1}{2s+3}}. \tag{5.8}$$

As in kernel estimation, a pilot bandwidth can be computed using a simple private-value quantile-regression model to proxy Σ_I and Bias_I^2 in a parametric way. The corresponding square root IMSE rate is $L^{\frac{s+1}{2s+3}}$ which corresponds to the optimal minimax rate given in GPV in the absence of covariate, up to a logarithmic term and an exponent $s+1$ due to estimation of the private-value quantile function, instead of s appearing for pdf. In particular, it is $L^{-2/5}$ for $s=1$, with an exponent $2/5=.4$ close to $1/2$ suggesting potential good performances in small samples even in the presence of covariate.

A similar rate can also be derived for the uniform consistency of $\widehat{V}(\cdot|\cdot,I)$ stated in (5.6). Note also that (5.7) shows that the bid quantile estimator $\widehat{B}(\cdot|\cdot,I)$ converges uniformly to $B(\cdot|\cdot,I)$ with a rate which is nearly parametric.[9] All these convergence results take place over the whole $[0,1]\times\mathcal{X}$, meaning that potential boundary biases disappear asymptotically.

[9] The uniform consistency rate in (5.7) includes a bias term $o(h^{s+1})$, which is $O(h^{s+2})$ for $\alpha\neq 0$ because $B(\cdot|x,I)=x_1'\beta(\cdot|I)$ is $(s+2)$-th times continuously differentiable over $(0,1]$. Because $B(\cdot|x,I)$ may have only $(s+1)$ derivatives at $\alpha=0$, the bias order h^{s+2} may not hold uniformly over $[0,1]$.

5.2.2 Bias, variance and Central Limit Theorem

While Theorem 2 reviews the global performance of the AQR estimator, this section details some of its local features, and in particular its upper boundary behavior. Define

$$\Pi_h^1(\alpha) = \left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi(t)\pi(t)' K(t)\, dt \right)^{-1} S_1',$$

$$v_h^2(\alpha) = \Pi_h^1(\alpha)' \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi(t_1)\pi(t_2)' \min(t_1, t_2) K(t_1) K(t_2)\, dt_1 dt_2 \Pi_h^1(\alpha),$$

$$\Sigma_h(\alpha|I) = \frac{\alpha^2 v_h^2(\alpha)}{(I-1)^2} \mathbb{E}^{-1}\left[\frac{X_{1\ell}X_{1\ell}' \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|X_\ell, I_\ell)} \right] \mathbb{E}\left[X_{1\ell}X_{1\ell}' \mathbb{I}(I_\ell = I) \right] \mathbb{E}^{-1}\left[\frac{X_{1\ell}X_{1\ell}' \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|X_\ell, I_\ell)} \right], \tag{5.9}$$

$$\mathsf{Bias}_h(\alpha|x, I) = \frac{\alpha B^{(s+2)}(\alpha|x, I)}{I-1} S_1 \left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi(t)\pi(t)' K(t)\, dt \right)^{-1} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \frac{t^{s+2}\pi(t)}{(s+2)!} K(t)\, dt. \tag{5.10}$$

setting $\mathsf{Bias}_h(0|x, I) = 0$, see Footnote 8. The next Theorem gives some variance and bias expansions and the pointwise asymptotic distribution of the estimator. Recall that, in our pseudo true value framework, the bias of $\widehat{V}(\alpha|x, I)$ is given by $\overline{V}(\alpha|x, I) - V(\alpha|x, I)$. As the variance of $\widehat{V}(\alpha|x, I)$ is difficult to compute due to the implicit definition of this estimator, an expansion for the variance of its leading term $\widetilde{V}(\alpha|x, I)$ is given instead.

Theorem 3 *Under Assumptions A, H, S and for $\overline{V}(\cdot|\cdot, I)$, $\widetilde{V}(\cdot|\cdot, I)$ as in (5.3), (5.2) respectively, it holds for all α in $[0, 1]$ and all x in $\mathcal{X}$,*

$$\overline{V}(\alpha|x, I) = V(\alpha|x, I) + h^{s+1}\mathsf{Bias}_h(\alpha|x, I) + o(h^{s+1}), \tag{5.11}$$

$$\mathrm{Var}\left[\widetilde{V}(\alpha|x, I) \right] = \frac{x_1'\Sigma_h(\alpha|I)x_1 + O(h)}{LIh}, \tag{5.12}$$

with $\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} |\mathsf{Bias}_h(\alpha|x, I)| = O(1)$ *and* $\sup_{\alpha\in[0,1]} \|\Sigma_h(\alpha|I)\| = O(1)$.

If $\alpha \neq 0$, $x_1' \Sigma_h(\alpha|I) x_1$ *stays bounded away from* 0 *for all* x *and*

$$\left(\frac{LIh}{x_1' \Sigma_h(\alpha|I) x_1}\right)^{1/2} \left(\widehat{V}(\alpha|x,I) - \overline{V}(\alpha|x,I)\right)$$

converges in distribution to a standard normal.

These expansions give a better understanding of potential boundary effects affecting $\widehat{V}(\alpha|x,I)$. For $\mathsf{Bias}(\alpha|x,I)$ and $\Sigma(\alpha|I)$ defined before Theorem 2, it holds

$$\mathsf{Bias}_h(\alpha|x,I) = \mathsf{Bias}(\alpha|x,I) \text{ and } \Sigma_h(\alpha|I) = \Sigma(\alpha|I) \text{ for all } \alpha \text{ in } [h, 1-h]$$

since the support of the kernel $K(\cdot)$ is $[-1,1]$. Hence a pointwise optimal bandwidth for central quantile levels is $h_*(\alpha|x,I) = \left(\frac{\Sigma(\alpha|I)}{2(s+1)\mathsf{Bias}^2(\alpha|x,I)}\frac{1}{LI}\right)^{\frac{1}{2s+3}}$, which is obtained by minimizing the leading term of the Mean Squared Error obtained from (5.11) and (5.12).

As the asymptotic bias and standard deviation are proportional to α, more bias and variance are expected for higher quantile levels and smaller I. This follows from the fact that the leading term of $\widehat{V}(\alpha|x,I)$ is $\alpha\widehat{B}^{(1)}(\alpha|x,I)/(I-1)$, which is proportional to $\alpha/(I-1)$. Boundary effects can only occur for quantile levels in $[0,h]$ or $[1-h,1]$, which differ from this respect. As $\widehat{V}(\alpha|x,I) = \widehat{B}(\alpha|x,I) + \alpha\widehat{B}^{(1)}(\alpha|x,I)/(I-1)$, $\widehat{V}(\alpha|x,I)$ is close to $\widehat{B}(\alpha|x,I)$ when α is in $[0,h]$. In particular, $\widehat{V}(0|x,I) = \widehat{B}(0|x,I)$ which converges to $B(0|X,I)$ with the rate $1/\sqrt{LI} + o(h^{s+1})$ at least.

For upper quantile levels α in $[1-h,1]$, the consistency rate of $\widehat{V}(\alpha|x,I)$ is the slower $1/\sqrt{Lh} + O(h^{s+1})$ as for central quantile levels. Bias and variance also involve the matrix

$$\left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi(t)\pi(t)' K(t)\, dt\right)^{-1} = \left(\int_{-1}^{\frac{1-\alpha}{h}} \pi(t)\pi(t)' K(t)\, dt\right)^{-1}$$

for h small enough. This matrix increases with α, so that higher $|\mathsf{Bias}(\alpha|x,I)|$ and $\Sigma_h(\alpha|I)$ can take place near $\alpha = 1$ compared to central quantile levels. See Fan and Gijbels (1996) and the references therein for similar discussions. Accordingly, our simulations show an increase of the bias and variance for α approaching 1, see Figures 2 and 4 in Section 6.2.

5.3 Functional estimation

The plug-in estimators of $\theta(x)$ and θ in (3.1) are

$$\widehat{\theta}(x) = \int_0^1 \mathcal{F}\left[\alpha, x, \widehat{B}(\alpha|x,I), \widehat{B}^{(1)}(\alpha|x,I); I \in \mathcal{I}\right] d\alpha, \quad \widehat{\theta} = \int_{\mathcal{X}} \widehat{\theta}(x)\, dx,$$

with AQR estimators $\widehat{B}(\alpha|x,I)$ and $\widehat{B}^{(1)}(\alpha|x,I)$. Let us now introduce the asymptotic variances of $\widehat{\theta}(x)$ and $\widehat{\theta}$. The variances depend upon the matrices

$$\mathbf{P}(I) = \mathbb{E}\left[\mathbb{I}(I_\ell = I) X_{1\ell} X_{1\ell}'\right], \quad \mathbf{P}_0(\alpha|I) = \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I) X_{1\ell} X_{1\ell}'}{B^{(1)}(\alpha|X_\ell, I_\ell)}\right],$$

and of the functions, recalling b_{0I} and b_{1I} stand for $B(\alpha|x,I)$ and $B^{(1)}(\alpha|x,I)$ respectively,

$$\begin{aligned}
\varphi_{0I}(\alpha,x) &= \partial_{b_{0I}} \mathcal{F}\left[\alpha, x, B(\alpha|x,I), B^{(1)}(\alpha|x,I); I \in \mathcal{I}\right], \\
\varphi_{1I}(\alpha,x) &= \partial_{b_{1I}} \mathcal{F}\left[\alpha, x, B(\alpha|x,I), B^{(1)}(\alpha|x,I); I \in \mathcal{I}\right].
\end{aligned}$$

Let A be a $\mathcal{U}_{[0,1]}$ random variable, $\mathbf{1} = [1, \ldots, 1]'$ a $(D+1) \times 1$ vector, and define, recalling $x_1 = [1, x']'$,

$$\begin{aligned}
\sigma_L^2(x|I) &= \operatorname{Var}\left[\int_0^A \left(\varphi_{0I}(\alpha|x) - \partial_\alpha \varphi_{1I}(\alpha|x)\right) \mathbf{1}'\mathbf{P}_0(\alpha|I)^{-1} \mathbf{P}(I)^{1/2} x_1 d\alpha\right], \\
\sigma_L^2(I) &= \operatorname{Var}\left[\int_{\mathcal{X}} \left\{\int_0^A \left(\varphi_{0I}(\alpha|x) - \partial_\alpha \varphi_{1I}(\alpha|x)\right) \mathbf{1}'\mathbf{P}_0(\alpha|I)^{-1} \mathbf{P}(I)^{1/2} x_1 d\alpha\right\} dx\right], \\
\sigma_L^2(x) &= \sum_{I \in \mathcal{I}} \frac{\sigma_L^2(x|I)}{I}, \quad \sigma_L^2 = \sum_{I \in \mathcal{I}} \frac{\sigma_L^2(I)}{I}.
\end{aligned}$$

The proof of Theorem 4 in Appendix E shows that the asymptotic variances of $\widehat{\theta}(x)$ and $\widehat{\theta}$ are $\sigma_L^2(x)/L$ and σ_L^2/L respectively provided they are bounded away from 0. This holds if

$$\text{for all } x \text{ in } \mathcal{X},\ \varphi_{0I}(\alpha|x) \neq \partial_\alpha \varphi_{1I}(\alpha|x) \text{ for some } (\alpha, I) \text{ of } [0,1] \times \mathcal{I}. \tag{5.13}$$

It may be indeed that $\sigma_L^2(x|I) = 0$ and $\sigma_L^2 = 0$, in which case $\widehat{\theta}(x)$ and $\widehat{\theta}$ can converge to $\theta(x)$ and θ with "superefficient" rates, that is faster than $1/L^{1/2}$. Why it is possible is better understood in our quantile context, through an example of functionals for which (5.13) does not hold. Consider, for some given I_0 of $\mathcal{I}$,

$$\mathcal{F}\left[\alpha, x, B(\alpha|x, I), B^{(1)}(\alpha|x, I); I \in \mathcal{I}\right] = 2B(\alpha|x, I_0) B^{(1)}(\alpha|x, I_0)$$

which gives $(\varphi_{0I_0}(\alpha|x), \varphi_{1I_0}(\alpha|x)) = 2\left(B^{(1)}(\alpha|x, I_0), B(\alpha|x, I_0)\right)$. Hence (5.13) does not hold and $\sigma_L^2(x) = \sigma_L^2 = 0$. Why $\widehat{\theta}(x)$ and $\widehat{\theta}$ can converge with superefficient rates for these functionals is in fact not surprising observing that they estimate

$$\theta(x) = B^2(1|x, I_0) - B^2(0|x, I_0), \quad \theta = \int_{\mathcal{X}} \theta(x)\, dx,$$

respectively. For these examples, the parameters of interest only depend upon extreme quantiles, in which case superefficient estimation is possible, see e.g. Hirano and Porter (2003) and the references therein. The next Theorem establishes the asymptotic normality of $\widehat{\theta}(x)$ and $\widehat{\theta}$.

Theorem 4 *Suppose Assumptions A, H with* $\frac{\log^2 L}{Lh^3} = o(1)$*, Assumption S with* $s \geq 2$ *and (5.13) hold. Then, for all* x *in* $\mathcal{X}$*,* $\sigma_L^2(x)$ *and* σ_L^2 *are bounded away from* 0 *and infinity when* L *grows, and*

$$\frac{\sqrt{L}\left(\widehat{\theta}(x) - \theta(x) - \mathsf{bias}_{L,\theta(x)}\right)}{\sigma_L(x)}, \quad \frac{\sqrt{L}\left(\widehat{\theta} - \theta - \mathsf{bias}_{L,\theta}\right)}{\sigma_L}$$

both converge in distribution to a standard normal, the bias items $\mathsf{bias}_{L,\theta(x)}$ *and* $\mathsf{bias}_{L,\theta}$ *being* $o(h^s)$.

Note that the conditional functional estimator $\widehat{\theta}(x)$ converges with a parametric rate, under a bandwidth condition slightly stronger than in Assumption H. This bandwidth condition corresponds to the fact that the linearization error term $\widehat{B}^{(1)}(\alpha|x, I) - \widetilde{B}(\alpha|x, I)$ is of order $\log L/(Lh^{3/2})$ as guessed from (5.5) and must be $o(1/\sqrt{L})$.

The bias term order $o(h^s)$ is given by the estimation of $B^{(1)}(\alpha|x, I)$, and is of order $O(h^{s+1})$ when $\mathcal{F}(\cdot)$ depends upon $\alpha B^{(1)}(\alpha|x, I)$ as in all the Examples. Let $\mathcal{G}_{b_{1I}}(\cdot)$ be the partial derivative of $\mathcal{F}(\cdot)$ with respect to $\alpha B^{(1)}(\alpha|x, I)$, and $\mathsf{Bias}_h(\alpha|x, I)$ be as in (5.10).

Then

$$\mathsf{bias}_{L,\theta(x)} = h^{s+1}(1+o(1)) \times \sum_{I\in\mathcal{I}} \int_0^1 \mathcal{G}_{b_{1I}}\left[\alpha, x, B(\alpha|x,I), \alpha B^{(1)}(\alpha|x,I); I\in\mathcal{I}\right] \mathsf{Bias}_h(\alpha|x,I)\, d\alpha$$

and $\mathsf{bias}_{L,\theta} = \int_{\mathcal{X}} \mathsf{bias}_{L,\theta(x)} dx$. The estimators $\widehat{\theta}(x)$ or $\widehat{\theta}$ are therefore asymptotically unbiased if $h^{s+1}\sqrt{Lh} = o(1)$ or $h^{s+1}\sqrt{L} = o(1)$ respectively.

Theorem 4 applies to our functional Examples, but the resulting variance can be somehow involved, so that the use of the bootstrap can be preferred as discussed below. We first detail the variance obtained for the cdf estimator of Example 3 to illustrate the influence of the bandwidth η on its variance.

Example 3 (cont'd). For the cdf estimator $\widehat{F}_\eta(v|x,I) = \int_0^1 \mathbb{I}_\eta\left[v - \widehat{V}(\alpha|x,I)\right] d\alpha$,

$$\varphi_{0I}(\alpha|x) = -\frac{1}{\eta} k\left(\frac{v - V(\alpha|x,I)}{\eta}\right), \quad \varphi_{1I}(\alpha|x) = \frac{\alpha}{(I-1)\eta} k\left(\frac{v - V(\alpha|x,I)}{\eta}\right),$$

$$\partial_\alpha \varphi_{1I}(\alpha|x) = \frac{1}{(I-1)\eta} k\left(\frac{v - V(\alpha|x,I)}{\eta}\right) - \frac{\alpha}{(I-1)\eta^2} k^{(1)}\left(\frac{v - V(\alpha|x,I)}{\eta}\right) V^{(1)}(\alpha|x,I).$$

When η goes to 0, the dominant part of the variance is, for inner v, integrating by parts and setting $V_{x,I} = V(A|x,I)$

$$
\begin{aligned}
&\frac{I}{L} \operatorname{Tr}\left\{\operatorname{Var}\left[\left(\int_0^A \partial_\alpha \varphi_{1I}(\alpha|x) \mathbf{P}_0(\alpha|I)^{-1} d\alpha\right) \mathbf{P}(I)^{1/2} x_1\right]\right\} \\
&\quad= \frac{(1+o(1)) I}{L} \operatorname{Tr}\left\{\operatorname{Var}\left[\varphi_{1I}(A|x) \partial_\alpha \left[\mathbf{P}_0(A|I)^{-1}\right] \mathbf{P}(I)^{1/2} x_1\right]\right\} \\
&\quad= \frac{(1+o(1)) I}{(I-1)^2 L} \\
&\qquad \times \operatorname{Tr}\left\{\operatorname{Var}\left[\frac{F(V_{x,I}|x,I)}{f(V_{x,I}|x,I)} \frac{k\left(\frac{v-V_{x,I}}{\eta}\right)}{\eta} \partial_\alpha \left[\mathbf{P}_0(F(V_{x,I}|x,I)|I)^{-1}\right] \mathbf{P}(I)^{1/2} x_1\right]\right\} \\
&\quad= \frac{(1+o(1)) I \int k^2(t)\,dt}{(I-1)^2 L\eta} \left(\frac{F(v|x,I)}{f(v|x,I)}\right)^2 \\
&\qquad \times \operatorname{Tr}\left\{\partial_\alpha \left[\mathbf{P}_0(F(v|x,I)|I)^{-1}\right] \mathbf{P}(I)^{1/2} x_1 x_1' \mathbf{P}(I)^{1/2} \partial_\alpha \left[\mathbf{P}_0(F(v|x,I)|I)^{-1}\right]\right\}.
\end{aligned}
$$

Hence the order of the variance of $\widehat{F}_\eta(v|x,I)$ is $1/(L\eta)$ when η goes to 0. Its bias has two components: the first is $\mathsf{bias}_{L,F_\eta(v|x,I)}$ due to the bias of $\widehat{V}(\alpha|x,I)$ and is of order $O(h^{s+1})$, while the second is $F_\eta(v|x,I) - F(v|x,I) = O(\eta^{s+1})$ if $k(\cdot)$ is a kernel of order s. Further work is needed to determine at which rate η can go to 0 in this heuristic.

Bootstrap inference. Earlier theoretical works considering quantile-regression bootstrap inference are Rao and Zhao (1992), for the weighted bootstrap, and Hahn (1995) for the pairwise bootstrap. See also Liu and Luo (2017) for quantile-based auction testing procedures. For standard and sieve quantile-regression estimators, Belloni et al. (2019) establish consistency of several bootstrap procedures for functionals and uniformly with respect to the quantile levels. It is expected that it carries over to our AQR estimators when bias terms can be neglected, but is out of the scope of the present paper.

6 Simulation experiments

The first simulation experiments compare the AQR estimation method with GPV and its homogenized-bid extension. Other simulations illustrate its performances for estimation of the private-value quantile function, expected seller revenue and optimal reserve price, or risk-aversion parameter. All experiments involve $L = 100$ auctions with $I = 2$ or $I = 3$ bidders, so that small samples of 200 or 300 are considered. In most of the experiments, three auction-specific auction covariates are considered. The number of replications is $1,000$ in all experiments. AQR are computed over the estimation grid $\alpha = 0, 0.01, \ldots, 0.99, 1$. The AQR local polynomial order $s + 1$ is set to 2 and the AQR kernel is the Epanechnikov $K(t) = \frac{3}{4}(1 - t^2)\,\mathbb{I}(t \in [-1, 1])$.

Since the asymptotic bias and variance of $\widehat{V}(\alpha|x, I)$ tend to decrease with I, choosing a small number I of bidders is challenging. Hickman and Hubbard (2015) used 5 bids while $I = 3$ or 5 in Marmer and Shneyerov (2012) and Ma, Marmer and Shneyerov (2019). The number of bids LI ranges from $1,000$ for Hickman and Hubbard (2015) to $4,200$ for Marmer and Shneyerov (2012). In a simulation experiment focused on nonparametric estimation of the utility function of risk-averse bidders, Zincenko (2018) considers $I = 2$ with $L = 300$ and $I = 4$ with $L = 150$. These references do not consider covariate, with the exception of Zincenko (2018) for $L = 900$ auctions with one or two covariates. Therefore our simulation setting correspond to rather demanding small sample situations.

6.1 Comparison with GPV and homogenized-bid

The simulation experiments of this Section makes use of a "trigonometric" quantile function

$$T(\alpha) = \frac{1}{2}\left((\pi + 1)\alpha + \cos(\pi\alpha)\right) \tag{6.1}$$

whose probability density function has a compact support and is bounded away from 0, with a shape similar to a peaked Gaussian one as seen from the left panel of Figure 2.[10]

Comparison with GPV. This experiment considers private values drawn from $T(\cdot)$ and does not include covariate. It compares first the boundary bias corrected GPV two-step

[10] The pdf graph is obtained noting that $T^{(1)}(\alpha) = 1/f(T(\alpha))$, so that the graph $\alpha \in [0, 1] \mapsto (T(\alpha), 1/T^{(1)}(\alpha)$ is the one of the pdf $f(\cdot)$. The associated bid quantile function can be obtained using (2.4), which is used to simulate bids via a quantile transformation of uniform draws, as performed elsewhere in our simulation experiments.

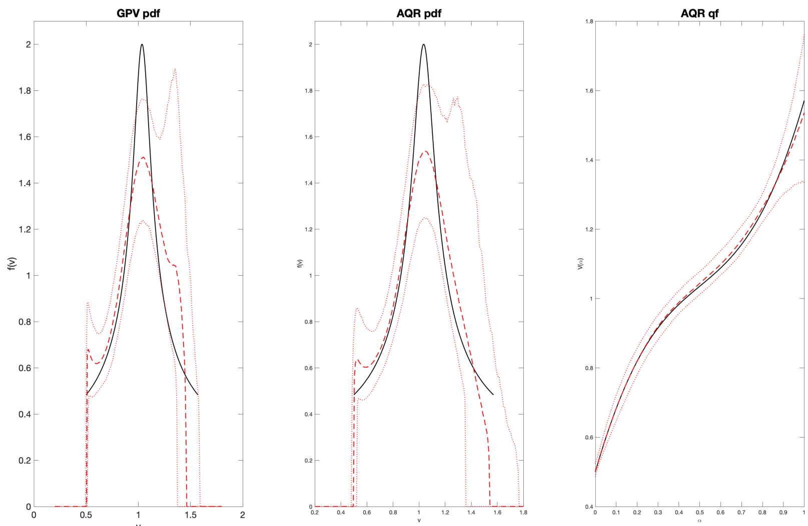

Figure 2: Private value pdf estimation using Hickman and Hubbard (2015) bias corrected version of GPV (left), Private value pdf estimation based upon AQR and Hickman and Hubbard (2015) (center), and AQR $\widehat{V}(\cdot)$ (right) , $L = 100$ and $I = 2$. Black line: true functions. Dashed red and dotted lines: pointwise median and $2.5\% - 97.5\%$ quantiles of the estimated functions across $1,000$ simulations.

pdf estimator with triweight kernel and rule of thumb bandwidth of Hickman and Hubbard (2015) with the AQR pdf estimator

$$\widehat{f}(v) = \int_0^1 \frac{1}{h_{AQR}} \widetilde{K}_{tri} \left(\frac{v - \widehat{V}(\alpha)}{h_{AQR}} \right) d\alpha, \quad h_{AQR} = \frac{3}{(LI)^{1/5}} \left(\int_0^1 \left(\widehat{V}(\alpha) - \int_0^1 \widehat{V}(t)dt \right)^2 d\alpha \right)^{1/2},$$

where the AQR bandwidth is $h = .3$, $I = 2$ and $\widetilde{K}_{tri}(\cdot)$ is the Hickman and Hubbard (2015) boundary bias corrected triweight kernel.

The results are reported in Figure 2, which illustrates how much harder estimating a pdf can be compared to estimating a quantile function. The variance and the bias of the two private-value pdf estimators look much higher, especially just after the density peak. This peak causes a small AQR bias for central quantiles.

Other features are common to the two pdf estimation procedures. The first stage of the GPV procedure is based upon an estimation of the private values from (2.6), which is likely to have more bias and variance for small bids than for large ones. Accordingly, the left panel of Figure 2 reveals that the GPV pdf estimator performs quite well before the density peak, but that both its bias and variance increase for higher private values. The AQR pdf estimator in the center panel looks less affected by these issues. By contrast, the quantile estimation procedure in the right panel of Figure 2 is only affected by a variance increase for upper quantiles as expected.

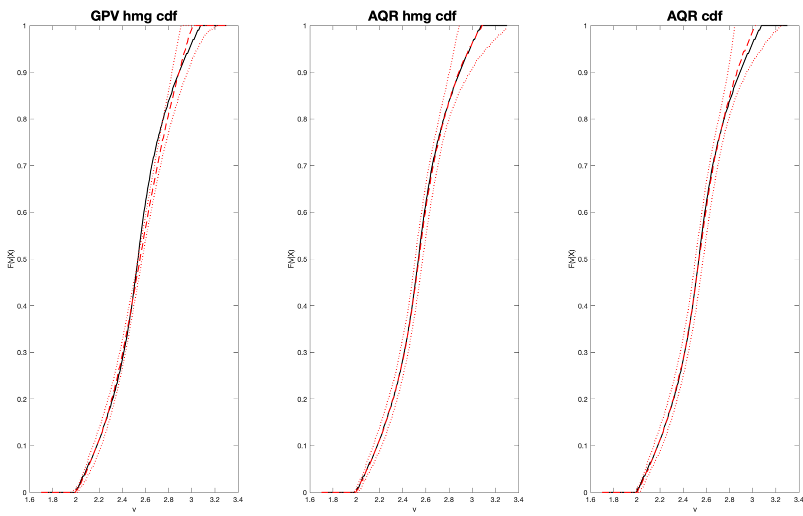

Figure 3: Private value cdf estimation using the homogenized bids and GPV version of Hickmann and Hubbard (2005) (left), AQR with homogenized bids (center) and AQR (right). Black line: true functions. Dashed red and dotted lines: pointwise median and 2.5% − 97.5% quantiles of the estimated functions across 1,000 simulations.

Cdf estimation with homogenized bids and AQR The bids considered in this experiment are associated with private values satisfying

$$V_{i\ell} = X_{1\ell} + X_{2\ell} + X_{3i\ell} + v_{i\ell}$$

where the $X_{j\ell}$'s are independently drawn from the uniform, independent from $v_{i\ell}$, whose quantile function is $T(\cdot)$ in (6.1). The bids are regressed on the covariate and the constant to obtain the homogenized bids $\widehat{b}_{i\ell}$. The latter are used as in the first-stage of Hickman and Hubbard (2015) to obtain homogenized pseudo private values $\widehat{v}_{i\ell}$, and the estimated private-value conditional cdf at $x_1 = x_2 = x_3 = 1/2$ is

$$\widehat{F}(v|x) = \frac{1}{LI} \sum_{\ell}^{L} \sum_{i=1}^{I} \mathbb{I}\left(\frac{1}{2}(\widehat{\beta}_1 + \widehat{\beta}_2 + \widehat{\beta}_3) + \widehat{v}_{i\ell} \leq v\right),$$

where the $\widehat{\beta}_j$ are the OLS slope estimators computed in the homogenized-bid regression. Two AQR conditional estimators are computed. The first uses the homogenized bids and (3.6) while the second is based on the standard AQR $\widehat{V}(\alpha|x, I)$, both using the bandwidth $h = .3$. The performances of these three cdf estimators are reported in Figure 3.

As expected, considering estimation of the private-value cdf gives smaller bias and variance than estimating pdf. All procedures have a similar variability. However the homogenized-bid GPV procedure has a larger bias, dominating its variability in the right centre part of the private-value distribution, than its AQR counterparts. Applying the AQR to the homogenized bids seems to slightly dominate the other AQR procedure.

6.2 Private value and expected revenue

Quantile-regression model and estimation details. The private-value quantile function is given by a quantile-regression model with an intercept and three independent covariates with the uniform distribution over $[0, 1]$,

$$V(\alpha|X) = \gamma_0(\alpha) + \gamma_1(\alpha) X_1 + \gamma_2(\alpha) X_2 + \gamma_3(\alpha) X_3$$

with

$$\gamma_0(\alpha) = 1 + 0.5\exp(5(\alpha - 1)), \quad \gamma_1(\alpha) = 1,$$
$$\gamma_2(\alpha) = 0.5(1 - \exp(-5\alpha)), \quad \gamma_3(\alpha) = 0.8 + 0.15((2\pi + 1)\alpha + \cos(2\pi\alpha)).$$

The coefficient $\gamma_0(\cdot)$ is flat near 0 and strongly increases near 1 while $\gamma_2(\cdot)$ strongly increases near 0 and is flat after. The slope $\gamma_3(\cdot)$ is as the trigonometric quantile function (6.1), but with stronger oscillations which makes it harder to estimate.

The performances of the private-value quantile estimation procedure are evaluated through the individual estimation of each slope function or estimation of $V(\alpha|x)$ when the x_j are set to their median 1/2. The curvature of the expected revenue is mostly due to $\gamma_2(\cdot)$, the other coefficients having a rather flat contribution. The performances of the expected revenue estimation procedure are therefore evaluated removing the intercept, setting x_1 and x_3 to 0 and taking $x_2 = 0.8$. This choice gives a unique optimal reserve price achieved for $\alpha_* = .3$, which is not too close to the boundaries so that the expected revenue function has a substantial concave shape which is suppose to make estimation more difficult. This is also used for evaluating estimation of the optimal reserve price $R_* = .8\gamma_2(\alpha_*)$.

Simulation results. Table 1 summarizes the simulation results for the estimation of the private-value quantile function, the expected revenue and the optimal reserve price. The Bias and square Root Integrated Mean Squared Error (RIMSE) lines for $\widehat{V}(\cdot|\cdot)$ gives the simulation counterparts of, respectively

$$\left(\frac{1}{4}\sum_{j=0}^{3}\int_0^1 \left(\mathbb{E}\left[\widehat{\gamma}_j(\alpha)\right]-\gamma_j(\alpha)\right)^2 d\alpha\right)^{1/2} \text{ and } \left(\frac{1}{4}\sum_{j=0}^{3}\int_0^1 \mathbb{E}\left[\left(\widehat{\gamma}_j(\alpha)-\gamma_j(\alpha)\right)^2\right] d\alpha\right)^{1/2}.$$

The Bias and RIMSE for the expected revenue are computed similarly. Table 1 also gives the Bias and square Root Mean Squared Error (RMSE) of the optimal reserve price estimator. All these quantities are computed for bandwidths $.2, .3, \ldots, .9$.

	h	.2	.3	.4	.5	.6	.7	.8	.9
$\widehat{V}(\cdot\|\cdot)$	Bias	.131	.141	.143	.145	.150	.159	.166	.176
	RIMSE	.433	.386	.355	.332	.322	.309	.303	.305
$\widehat{ER}(\cdot)$	Bias	.036	.044	.049	.050	.051	.049	.047	.045
	RIMSE	.109	.104	.102	.100	.099	.098	.097	.096
$\widehat{R}_*$	Bias	-.036	-.031	-.014	-.002	.009	.022	.037	.043
	RMSE	.129	.099	.075	.067	.062	.064	.066	.066

Table 1: Private value quantile function, expected revenue, and optimal reserve price

Estimation of the private-value slope coefficients seems more sensitive to the bandwidth parameter than the expected revenue or optimal reserve price. It has also a much higher RIMSE. The bandwidth behavior of $\widehat{V}(\alpha|x)$ is also illustrated in Figure 4, which considers

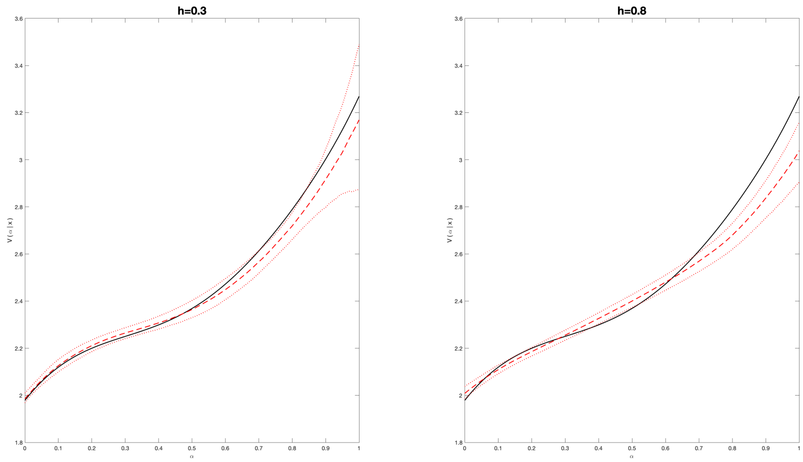

Figure 4: Private value quantile estimation for $h = 0.3$ (left) and $h = 0.8$ (right) for average covariate. True $V(\alpha|x) = \gamma_0(\alpha) + (\gamma_1(\alpha) + \gamma_2(\alpha) + \gamma_3(\alpha))/2$ in black. Dashed red and dotted lines: pointwise median and $2.5\% - 97.5\%$ quantiles of $\widehat{V}(\alpha|x)$ across $1,000$ simulations.

the small bandwidth $h = 0.3$ and the larger $h = 0.8$. As expected from Theorem 3, the dispersion of $\widehat{V}(\alpha|x)$ increases with α and decreases with h, while the bias increases with α and h. Figure 4 also suggests that choosing a large bandwidth as recommended by Table 1 may lead to important bias issues, including underestimating the private-value quantile function for high α. This is mostly due to the slope $\gamma_3(\cdot)$ which is an important source of bias.

This contrasts with estimation of the expected revenue and optimal reserve price, which seems mostly unaffected by the bandwidth. This is partly because the expected revenue depends upon $(1-\alpha)V(\alpha|x)$: multiplying the private-value quantile function by $(1-\alpha)$ mitigates larger bias and variance near the boundary $\alpha = 1$, see also Figure 5. For the considered experiment, the true expected revenue is always in the 95% band of Figure 5 while the true private-value quantile function is out for large α when $h = 0.8$.

6.3 Risk-aversion parameter

Two risk-aversion estimators are considered. The first estimator $\widehat{\nu}_{fp}$ is based upon (3.3) and uses two independent samples of size $L = 100$ with 2 and 3 bidders from the model above,

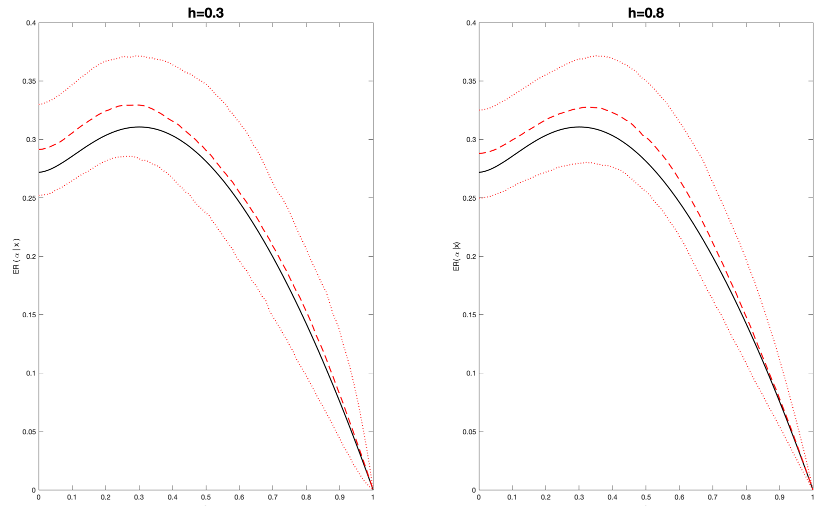

Figure 5: Expected revenue estimation for $h = 0.3$ (left) and $h = 0.8$ (right). True $ER(\alpha|x)$ in black. Dashed red and dotted lines: pointwise median and $2.5\% - 97.5\%$ quantiles of $\widehat{ER}(\alpha|x)$ across $1,000$ simulations.

which corresponds to a CRRA utility function t^{ν} with $\nu = 1$.[11] Integrals with respect to α are computed using Riemann sums whereas integrals with respect to x are replaced with sample means over the two auction samples. The second estimator $\widehat{\nu}_{asc}$ is based upon (3.4) and uses an additional sample of size $L = 100$ of ascending auctions with two bidders. In this case, it is possible to consider various values of ν and the simulation experiment considers the values 0.2, 0.6 and 1. Indeed, if $B(\alpha|X)$ is the first-price auction quantile bid function with $I = 2$, the observed bids drawn from $B(\alpha|X)$ are rationalized by a CRRA utility function t^{ν} if the private-value quantile function is set to

$$V_{\nu}(\alpha|X) = B(\alpha|X, 2) + \nu\alpha B^{(1)}(\alpha|X, 2)$$

provided $V_{\nu}^{(1)}(\cdot|X) > 0$ for all X. As $V_{\nu}^{(1)}(\cdot|\cdot) > 0$ holds in our case, we use $V_{\nu}(\alpha|X)$ to generate two ascending bids for each auction. Following Gimenes (2017), $V_{\nu}(\alpha|X)$ can be estimated from winning bids in these ascending auctions using AQR for quantile level $2\alpha - \alpha^2$.

[11] The optimal bid functions can be computed explicitly under the risk-neutral case $\nu = 1$. Considering other values of ν would request to use numerical computations of the bid functions.

	ν	h	.2	.3	.4	.5	.6	.7	.8	.9
$\widehat{\nu}_{fp}$	1	Bias	-.795	-.564	-.412	-.288	-.178	.-.080	.003	.053
		RMSE	.891	.681	.545	.471	.404	.380	.393	.436
$\widehat{\nu}_{asc}$	1	Bias	-.016	-.019	-.037	-.061	-.085	-.100	-.109	-.111
		RMSE	.240	.247	.248	.254	.260	.267	.276	.282
	.6	Bias	.028	.023	.009	-.008	-.025	-.035	-.040	-.042
		RMSE	.172	.176	.174	.175	.175	.179	.184	.188
	.2	Bias	.088	.083	.075	.066	.058	.053	.052	.053
		RMSE	.135	.133	.126	.122	.117	.116	.116	.118

Table 2: Risk-aversion estimation

Table 2 shows that $\widehat{\nu}_{asc}$ dominates $\widehat{\nu}_{fp}$ in this experiment. While the RMSE and bias of $\widehat{\nu}_{asc}$ do not seem sensitive to h, this is not the case for $\widehat{\nu}_{fp}$ which has a high downward bias, and then RMSE, for small h. Further investigations suggest this is due to an unbalanced variable issue, the difference $\widehat{B}(\alpha|X,3) - \widehat{B}(\alpha|X,2)$ being very smooth while $\alpha\left(\widehat{B}^{(1)}(\alpha|X,3)/2 - \widehat{B}^{(1)}(\alpha|X,2)\right)$ is more erratic, especially when α is close to 1. This issue is addressed in the application by restricting α to $[0,.8]$ for risk-aversion estimation.

7 Timber data application

Timber auctions data have been used in several empirical studies (see Athey and Levin (2001), Athey, Levin and Seira (2011) Li and Zheng (2012), Aradillas-Lopez, Gandhi and Quint (2013) among others). Some other works have investigated risk-aversion in timber auctions (e.g., Lu and Perrigne (2008), Athey and Levin (2001), Campo et al. (2011)). This section uses data from timber auctions run by the US Forest Service (USFS) from Lu and Perrigne (2008) and Campo et al. (2011), which aggregates auctions of 1979 from the states covering the western half of the United States (regions 1–6 as labeled by the USFS). It contains bids and a set of variables characterizing each timber tract, including the estimated volume of the timber measured in thousands of board feet (mbf) and its estimated appraisal value given in dollars per unit of volume. We consider the 107 first-price auctions with two bidders, the first-price auctions with three bidders ($L = 108$) and ascending auctions with two bidders ($L = 241$). The considered covariates are the appraisal value and the timber volume taken in log. AQR is implemented with bandwidth $h = .3$ and the Epanechnikov kernel.

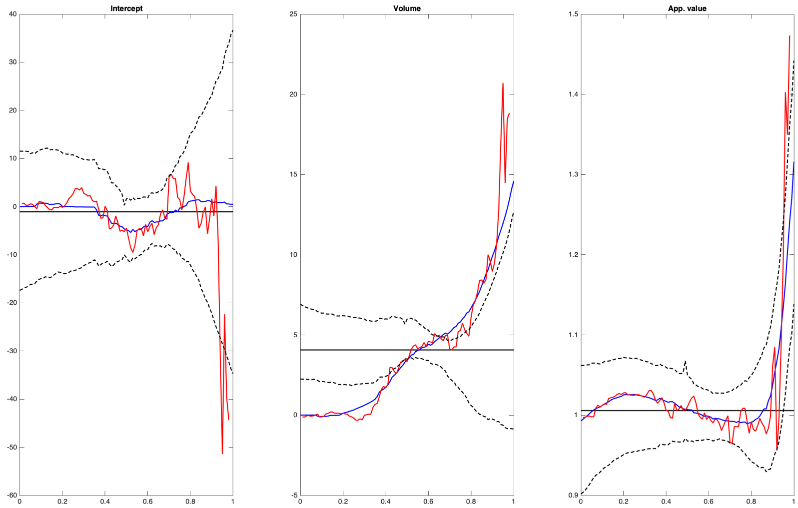

Figure 6: Two bidders first-price auction bid quantile slope coefficients: Intercept (left), volume (center) and appraisal value (right). AQR with $h = .3$ (blue), standard QR (red) and OLS regression (black), and pointwise 90% confidence intervals for the AQR-regression difference (black dashed line) centered at the regression coefficients. A regression or AQR estimated slope coefficients outside the confidence band indicates a potential misspecification of the homogenized-bid regression model.

Pointwise confidence intervals are computed using 10,000 pairwise bootstrap replications. As in the simulation experiments, the CRRA parameter estimator has a high variance, and risk-neutrality cannot be rejected, see Gimenes and Guerre (2019). The rest of the application therefore assumes risk-neutral bidders.

Specification testing. Table 3 reports first the results of Rothe and Wied (2013) test for the four following null hypotheses: correct specification of the quantile-regression (QR), of the homogenized-bid (HHS) model, exogeneity of the auction format (Format), participation exogeneity (Entry). The three last null hypotheses are also tested using quantile-regression coefficients comparison tests. Quantile-regression coefficient test statistics are based upon a discretized version of Liu and Luo (2017) integral statistic (2.17), using Riemann sum over a grid $\alpha = 0, 1/100, \ldots, 1$. For HHS, the intercept of $\widehat{\beta}_{H_0}(\cdot)$ is from the AQR estimator while the slope are OLS. For the Format null hypothesis, $\widehat{\beta}_{H_0}(\cdot) = \alpha^{-1} \int_0^{\alpha} \widehat{\gamma}_{asc}(a|2)da$, $\widehat{\gamma}_{asc}(\cdot|2)$ being an AQR version of Gimenes (2017) using the ascending auction sample with two bidders. For Entry, $\widehat{\beta}_{H_0}(\cdot)$ is an AQR version of (2.18) using first-price auction with three bidders data. The Rothe and Wied (2013) statistic uses the unconstrained and constrained cdf estimators computed from the two bidder sample

$$\widehat{G}(b,x) = \frac{1}{2L}\sum_{\ell=1}^{L}\sum_{i=1}^{2}\mathbb{I}\left(B_{i\ell} \leq b \text{ and } X_\ell \leq x\right),$$

$$\widehat{G}_{H_0}(b,x) = \frac{1}{2L}\sum_{\ell=1}^{L}\sum_{i=1}^{2}\mathbb{I}\left(X_\ell \leq x\right)\widehat{G}\left(B_{i\ell}\,\middle|\,X_\ell, \widehat{\beta}_{H_0}(\cdot)\right) \text{ with}$$

$$\widehat{G}\left(b\,\middle|\,x, \widehat{\beta}_{H_0}(\cdot)\right) = \int_0^1 \mathbb{I}\left[x_1'\widehat{\beta}_{H_0}(\alpha) \leq b\right] d\alpha$$

which are compared using the Cramer-von Mises statistic

$$\frac{1}{2L}\sum_{\ell=1}^{L}\sum_{i=1}^{2}\left(\widehat{G}_{H_0}(B_{i\ell}, X_\ell) - \widehat{G}(B_{i\ell}, X_\ell)\right)^2.$$

The p-values of the tests based upon Rothe and Wied (2013) use 10,000 replications of the

Tests		QR	HHS	Format	Entry
Rothe-Wied (2013)	Stat. value	.022	.084	.031	.031
	p-value	.07	.00	.22	.58
Test stat. (2.17)	Stat. value	x	2.95	0.70	0.42
	p-value	x	.03	.01	.07

Table 3: Specification tests

bootstrap procedure proposed by these authors, while the other p-values are from 10,000 pairwise bootstrap replications. The Rothe and Wied (2013) procedure does not reject the quantile-regression specification at the 5% level. This test also gives very high p-values for the Format and Entry null hypotheses, which both correspond to quantile-regression models, estimated in a different way than from the null hypothesis QR.[12] Both tests reject

[12]Rothe and Wied (2013) testing procedure seems very sensitive to the estimation variance of the considered quantile model. Attempts not reported here show that it also holds for the Escanciano and Goh (2014) bootstrap procedure, which gives smaller p-values. However, it does not include a re-estimation of the quantile-regression specifications, which may underestimate p-values in small samples. While Rothe and Wied (2013) bootstrap combines pairwise bootstrap, which draws auctions with replacement, with a semiparametric one, which draws bids from the considered model, only the semiparametric bootstrap is implemented here.

the homogenized-bid specification at 5% level. The coefficient-based test also rejects at this level exogeneity of the auction format, disagreeing with Rothe and Wied (2013).

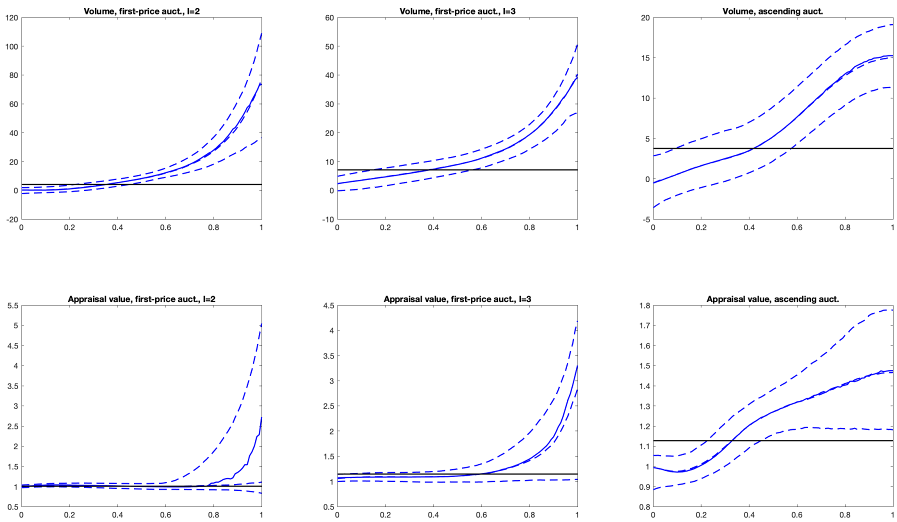

Figure 7: Volume (top) and appraisal value (bottom) estimated private-value slope function for first-price auctions with two bidders (left), three bidders (center) and ascending auctions (right), for $h = .3$. AQR estimation (full line), regression (full straight line) and 5%, 50%, 95% bootstrapped quantile (dashed line).

Bid quantile functions. Table 4 gives the results of bid OLS regressions. The dependent variables are the bids for first-price auctions and the winning bid for the ascending auction. The appraisal value coefficient is close to 1 in all auctions, but is found significantly distinct

Auctions		Intercept	Volume	Appraisal value	R^2
First-price	$I = 2$	-1.06	4.07	1.01	0.77
		(6.67)	(1.12)	(0.04)	
	$I = 3$	-20.79	7.10	1.15	0.70
		(9.55)	(1.34)	(0.06)	
Ascending	$I = 2$	2.76	3.76	1.12	0.67
		(15.05)	(1.85)	(0.06)	

Table 4: Auction bid regressions

at the 5% level when comparing the first-price auction with $I = 2$ with the one with $I = 3$ and the ascending auction. Similarly the volume coefficient of the first-price auction with

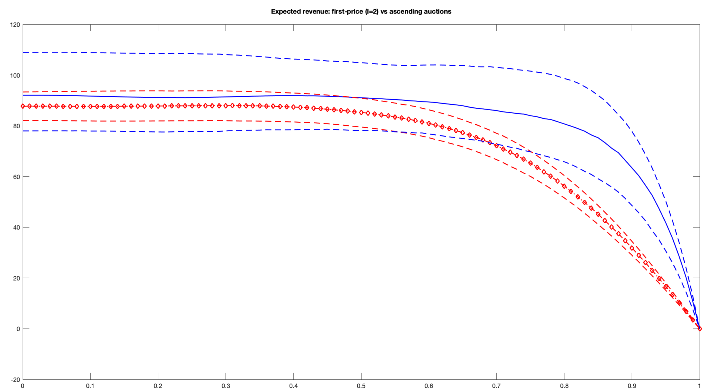

Figure 8: Estimated expected revenue as a function of the screening level for first-price (full line) and ascending (diamond) auctions with two bidders ($h = .4$). Volume and appraisal value set to median of the first-price auctions. $5\% - 95\%$ bootstrap quantiles in dashed lines.

$I = 2$ differs from the one with $I = 3$ at the 10% level. The appraisal value and volume coefficients of the first-price auctions with $I = 2$ and $I = 3$ are statistically distinct at the 5% level. This is not compatible with a homogenized-bid regression model assuming entry exogeneity.

Figure 6 gives the estimated slope for the first-price auction bids with $I = 2$. The volume OLS coefficient is consistently outside the pointwise 90% bootstrap confidence interval of its AQR counterpart. The appraisal value OLS estimate lies outside the AQR confidence intervals for high quantile level in $[.9, 1]$. Figure 6 also reports standard quantile-regression estimators, which exhibit a similar pattern. The intercept function does not look significant. Therefore, the intercept will be kept constant and estimated using OLS in the rest of the application. Comparison of the augmented and standard quantile-regression estimation also shows that the former produces smoother slope coefficients.

Private value quantile function and expected revenue. Figure 7 gives the private-value slope function of the volume and appraisal variables. The shape of the volume slope varies across the type of auctions: while convex and in the $[20, 100]$ range for high α in the first-price case, it is in the $[8, 15]$ range and more oscillating for ascending auctions. This suggests that the private-value distribution and the auction mechanism are not independent, as also reported in Table 3 for the test based upon (2.17).

The appraisal value slope seems statistically different from its OLS counterpart for ascending auctions. For all auctions, the estimated appraisal value slopes start at 1 for α near 0, suggesting that low type bidders do not get added value from the appraisal value. This contrasts with high type bidders with higher α, which markup can be very high, in a significant way for the case of ascending auction. This illustrates the important difference between low type and high type bidders.

A possible discrepancy between first-price and ascending auctions with two bidders also appears in the expected revenue computed for median values of the two explanatory variables, see Figure 8. The ascending auction expected revenue is always below the first-price one. This seems statistically significant for high screening levels. However, this may not be relevant for the seller as the optimal revenue is achieved for a wide range $[0, .5]$ of screening levels over which the two expected revenue curves seem flat.

8 Final remarks

This paper proposes a quantile-regression modeling strategy for first-price auction under the independent private value paradigm, which applies quantile-level local-polynomial to estimate the private-value quantile regression. This new framework can also be used to estimate some private-value random-coefficient models and to test some specification and exogeneity hypotheses of economic interest. This approach is found to work well both in simulations, and in a timber auction application where a strong low type/high type bidder heterogeneity is detected. Another empirical finding is that the seller expected revenue in a median auction is higher in first-price than in ascending auctions, but flat in a large zone around the optimal reserve prices. This suggests that the choice of a reserve price and of an auction mechanism may not be so important, at least for the median auction considered in the application.

Many aspects of the paper deserve further investigations. The estimated constant relative risk-aversion exhibits a quite large variance, suggesting that a better understanding of efficiency issues is needed. Various extensions can also be considered, such as endogenous entry as in Marmer, Shneyerov and Xu (2013a) or Gentry and Li (2014). Our quantile approach can be extended to exchangeable affiliated values as considered in Hubbard, Li and Paarsch (2012), see also Gimenes and Guerre (2020) for the more involved case of interdependent values. The approach of Wei and Carroll (2009) can be used to tackle unobserved heterogeneity as in Krasnokutskaya (2011).

References

[1] Athey, S. & J. Levin (2001). Information and competition in U.S. Forest Service timber auctions. *Journal of Political Economy* **109**, 375–417.

[2] Athey, S., J. Levin & E. Seira (2011). Comparing open and sealed bid auctions: evidence from timber auctions. *The Quarterly Journal of Economics* **126**, 207–257.

[3] Aradillas-Lopez, A., A. Gandhi & D. Quint (2013). Identification and inference in ascending auctions with correlated private values. *Econometrica* **81,** 489–534.

[4] Backus, M. & G. Lewis (2019). Dynamic demand estimation in auction markets. Document paper, NBER.

[5] Bassett, G. & R. Koenker (1982). An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association,* **77**, 407–415.

[6] Belloni, A., V. Chernozhukov, D. Chetverikov & I. Fernandes-Vál (2019). Conditional quantile processes based on series or many regressors. *Journal of Econometrics* **213**, 4–29.

[7] Berry, S., J. Levinsohn & A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica* **63**, 841–890.

[8] Bhattacharya, J. (2019). Quantile regression with generated dependent variable and covariates. Working paper, University of Southampton.

[9] Campo, S., E. Guerre, I. Perrigne & Q. Vuong (2011). Semiparametric estimation of first-price auctions with risk-averse bidders. *Review of Economic Studies,* **78**, 112–147.

[10] Chernozhukov, V., I. Fernández-Val & A. Galichon (2010). Quantile and probability curves without crossing. *Econometrica* **78**, 1093-1125.

[11] Dette, H. & S. Volgushev (2008). Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society: Series B* **70**, 609–627.

[12] Enache, A. & J.P. Florens (2017). A quantile approach for the estimation of first-price private value auction. Working paper, Toulouse School of Economics.

[13] Escanciano, J.C. & S.C. Goh (2014). Specification analysis of linear quantile models. *Journal of Econometrics* **178**, 495–507.

[14] Escanciano, J.C. & S.C. Goh (2019). Quantile-regression inference with adaptive control of size. *Journal of the American Statistical Association* **114**, 1382–1393.

[15] Escanciano, J.C. & C. Velasco (2010). Specification tests of parametric dynamic conditional quantiles. *Journal of Econometrics* **159**, 209–221.

[16] Fan, J. & I. Gijbels (1996). *Local polynomial modeling and its applications.* Chapman and Hall/CRC.

[17] Fan, Y., E. Guerre & S. Lazarova (2020). Local composite quantile regression. Working Paper, University of Washington and QMUL.

[18] Fang, K. T., S. Kotz & K.W. Ng (2017). *Symmetric Multivariate and Related Distributions.* Chapman and Hall/CRC.

[19] Gentry, M. & T. Li (2014). Identification in auctions with selective entry. *Econometrica* **82**, 315–344.

[20] Gimenes, N. (2017). Econometrics of ascending auction by quantile regression. *Review of Economics and Statistics* **99**, 944–953.

[21] Gimenes, N. & E. Guerre (2019). Quantile regression methods for first-price auctions. `arXiv` working paper, first version. `arxiv:1909.05542v1 [econ.EM]`.

[22] Gimenes, N. & E. Guerre (2020). Nonparametric identification of an interdependent value model with buyer covariates from first-price auction bids. *Journal of Econometrics* **219**, 1–18.

[23] Guerre, E., I. Perrigne & Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica* **68**, 525–574.

[24] Guerre, E., I. Perrigne & Q. Vuong (2009). Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions. *Econometrica* **77**, 1193–1227.

[25] Guerre, E. & C. Sabbah (2012). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. *Econometric Theory* **28**, 87–129.

[26] Hahn, J. (1995). Bootstrapping quantile regression estimators. *Econometric Theory* **11**, 105–121.

[27] Haile, P.A., H. Hong & M. Shum (2003). Nonparametric tests for common values in first-price sealed-bid auctions. Cowles Foundation discussion paper.

[28] Hickman, B.R. & T.P. Hubbard (2015). Replacing sample trimming with boundary correction in nonparametric estimation of first-price auctions. *Journal of Applied Econometrics*, **30**, 736-762.

[29] Hirano, K. & J.R. Porter (2003). Asymptotic efficiency in parameter structural models with parameter-dependent support. *Econometrica*, **71**, 1307–1338.

[30] Hoderlein, S., J. Klemelä & E. Mammen (2010). Analyzing the random coefficient model nonparametrically. *Econometric Theory* **26**, 804–837.

[31] Hubbard, T.P., T. Li & H.J. Paarsch (2012). Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. *Journal of Econometrics*, **168**, 4–16.

[32] Koenker, R. (2005). *Quantile regression.* Cambridge University Press.

[33] Koenker, R. & G. Bassett (1978). Regression quantiles. *Econometrica,* **46**, 33–50.

[34] Koenker, R. & Xiao, Z. (2002). Inference on the quantile regression process. *Econometrica,* **70**, 1583–1612.

[35] Krasnokutskaya, E. (2011). Identification and estimation of auctions models with unobserved heterogeneity. *Review of Economic Studies,* **78**, 293–327.

[36] Laffont, J.-J., H. Ossard & Q. Vuong (1995). Econometrics of first-price auctions. *Econometrica* **63**, 953–980.

[37] Li, T., I. Perrigne & Q. Vuong (2003). Semiparametric estimation of the optimal reserve price in first-price Auctions, *Journal of Business & Economic Statistics* **2**1, 53–64.

[38] Li, T. & X. Zheng (2012). Information acquisition and/or bid preparation: A structural analysis of entry and bidding in timber sale auctions. *Journal of Econometrics* **168**, 29–46.

[39] Liu, N. & Y. Luo (2017). A nonparametric test of exogenous participation in first-price auctions. *International Economic Review* **58**, 857–887.

[40] Liu, N. & Q. Vuong (2018). Nonparametric test of monotonicity of bidding strategy in first price auctions. Working paper.

[41] Lu, J. & I. Perrigne (2008). Estimating risk aversion from ascending and sealed-bid auctions: the case of timber auction data. *Journal of Applied Econometrics* **23**, 871–896.

[42] Luo, Y. & Y. Wan (2018). Integrated-quantile-based estimation for first-price auction models. *Journal of Business & Economic Statistics* **36**, 173-180.

[43] Ma, J., Marmer, V., & A. Shneyerov (2019). Inference for first-price auctions with Guerre, Perrigne and Vuong's estimator. *Journal of Econometrics* **211**, 597–538.

[44] Marmer, V., & A. Shneyerov (2012). Quantile-based nonparametric inference for first-price auctions. *Journal of Econometrics* **167**, 345–357.

[45] Marmer, V., A. Shneyerov & P. Xu (2013a). What model for entry in first-price auctions? A nonparametric approach. *The Journal of Econometrics*, **176**, 46–58.

[46] Marmer, V., A. Shneyerov & P. Xu (2013b). What model for entry in first-price auctions? A nonparametric approach. Supplementary material. Website of *The Journal of Econometrics*.

[47] Maskin, E. & J.G. Riley (1984). Optimal auctions with risk averse buyers. *Econometrica* **52,** 1473–1518.

[48] Menzel, K. & P. Morganti (2013). Large sample properties for estimators based on the order statistics approach in auctions. *Quantitative Economics,* **4,** 329–375.

[49] Milgrom, P.R. (2001). *Putting auction theory to work.* Cambridge University Press.

[50] Milgrom, P.R. & R.J. Weber (1982). A theory of auctions and competitive bidding. *Econometrica*, **50**, 1089–1122.

[51] Paarsch, H.J. (1992). Deciding between the common and private values paradigms in empirical models of auctions. *Journal of Econometrics* **51**, 191–215.

[52] Paarsch, H.J. & H. Hong (2006). *An introduction to the structural econometrics of auction data*. MIT Press.

[53] Rao, C.R. & L.C. Zhao (1992). Approximation to the distribution of M-estimates in linear models by randomly weighted bootstrap. *Sankhyā* **54**, Series A, 323–331.

[54] Rezende, L. (2008). Econometrics of auctions by least squares. *Journal of Applied Econometrics* **23**, 925–948.

[55] Rothe, C. & D. Wied (2013). Misspecification testing in a class of conditional distributional models. *Journal of the American Statistical Association* **108**, 314-324.

[56] Wei, Y. & R.J. Carroll (2009). Quantile regression with measurement error. *Journal of the American Statistical Association* **104**, 1129–1143.

[57] Zincenko, F. (2018). Nonparametric estimation of first-price auctions with risk-averse bidders. *Journal of Econometrics* **205**, 303–335.