Quantile Regression Approaches for Auctions

Sanches, Nathalie C. Gimenes Miessi

Quantile Regression Approaches for Auctions

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy (Ph.D.) in Economics

Nathalie C. Gimenes Miessi Sanches

School of Economics and Finance

Queen Mary, University of London

March 2014

For my parents, Jose and Lourdes.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of Queen Mary, University of London, is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it). The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 57,740 words.

Statement of conjoint work

I confirm that Chapter 3 was jointly co-authored with Emmanuel Guerre.

Acknowledgements

I am grateful to Professor Emmanuel Guerre for his help and support throughout my studies. I acknowledge the School of Economics and Finance at Queen Mary, University of London, for generous funding.

Abstract

The goal of this thesis is to propose a new quantile regression approach to identify and estimate the quantiles of the private value conditional distribution in ascending and first price auctions under the Independent Private Value (IPV) paradigm. The quantile regression framework provides a flexible and convenient parametrization of the private value distribution, which is not affected by the curse of dimensionality. The first Chapter of the thesis introduces a quantile regression methodology for ascending auctions. The Chapter focuses on revenue analysis, optimal reservation price and its associated screening level. An empirical application for the USFS timber auctions suggests an optimal reservation price policy with a probability of selling the good as low as 58% for some auctions with two bidders. The second Chapter tries to address this issue by considering a risk averse seller with a CRRA utility function. A numerical exercise based on the USFS timber auctions shows that increasing the CRRA of the sellers is sufficient to give more reasonable policy recommendations and a higher probability of selling the auctioned timber lot. The third Chapter develops a quantile regression methodology for first-price auction. The estimation method combines local polynomial, quantile regression and additive sieve methods. It is shown in addition that the new quantile regression methodology is not subject to boundary issues. The choice of smoothing parameters is also discussed.

Contents

List of Figures

List of Tables

Introduction

Auction is a market institution involving strategic interaction and asymmetric information that can be easily captured by a theoretical model. That is because the primitives of the strategic model, such as the set of players, the rules, the players' objectives and the information structure, can often be well described. By imposing some structure in the theoretical model, the economist is then able to answer relevant questions regarding the players' strategical behaviour and policy recommendations. The cummulative distribution function (c.d.f.) that characterizes the bidders' underlying demand/willingness to pay and the game information structure is however not observed by econometricians and policy makers because bidders do not reveal their own private values. Such a distribution function is therefore the key primitive of the game to be recovered. The goal of the structural analysis is to investigate whether the auction theoretical model can impose some restrictions on the observable data so that the underlying distribution of bidder's private value can be rationalized.

In the past 20 years, several researchers have proposed structural approaches to identify and estimate the private values distribution. The first wave of researches concerns parametric and semiparametric approaches that were proposed to circumvent the complexities in the computation of the equilibrium bidding strategy (See e.g. Paarsch (1992), Donald and Paarsch (1993, 1996), Laffont, Ossard and Vuong (1995)). Some other works have proposed nonparametric approaches to circumvent the misspecification bias that could arise from parametric/semiparametric approaches (See e.g. Guerre, Perrigne and Vuong (2000) and Haile and Tamer (2003)). Chapters 1 and 3 better describe the previous literature and its corresponding drawbacks. The goal of this thesis is to propose a new quantile regression approach to identify and estimate the quantiles of the private value conditional distribution in ascending and first-price auctions under the Independent Private Value (IPV) paradigm. The quantile regression framework provides a flexible and convenient parametrization of the private value distribution, which can cope with misspecifications and is not affected by the curse of dimensionality. The latter can be indeed a potential estimation issue given the importance of controlling for auction characteristics. Quantile has been recently considered in the auction literature, see e.g. Haile, Hong and Shum (2003), Bajari and Hortaçsu (2005), Guerre, Perrigne and Vuong (2009), Marmer and Shneyerov

(2012), Guerre and Sabbah (2012), Enache and Florens (2012) and Zincenko (2013). Quantile levels can be indeed useful for bidders and sellers. Bidders can use a private value quantile level to benchmark their bids in order to achieve a prescribed probability of winning the auction. For the seller, a key decision is the choice of the reservation price. The quantile framework then allows to highlight the screening level implied by the choice of an optimal reservation price, a policy characteristic that has been mostly ignored by previous empirical approaches. As discussed in Chapter 1 and 2, the optimal screening level in timber auctions are usually high, leading to a low probability of trading, which can be inappropriate from a seller's point of view. The focus on the screening level can thus provide the seller a better understanding about how reasonable and appropriate is his optimal reservation price choice.

The first Chapter of this thesis describes a quantile regression methodology for ascending auctions. Starting from the result shown in Athey and Haile (2002) that the c.d.f. of the winning bids can nonparametrically identify the private values distribution, I show that the private value conditional quantiles can then be recovered by a transformation of the winning bid conditional quantile function. A key stability property of the approach proposed is that a linear quantile regression specification for the private value quantiles implies a corresponding linear quantile regression specification for the winning bid quantiles, which then leads to simple estimation methods for private value linear quantile specifications. In this first Chapter, the quantile regression method developed by Koenker and Bassett (1978) is used. The first main advantage of the quantile regression approach in the ascending auction framework is that the private value conditional quantiles can be estimated at a faster parametric rate independently of the covariate dimension. A second advantage is that the entire distribution of private values, including the boundary, can be recovered, which is potentially important for policy recommendations. Several specification test are proposed as a contribution to the empirical auction literature. In the end of the first Chapter, there is an empirical application of the methodology using ascending timber auctions data from the United States Forest Service (USFS).

The second Chapter of this thesis proposes a numerical investigation of the optimal reservation price policy as the seller become more risk averse of not selling the good. This is motivated by the findings in the empirical section of the first Chapter that the optimal screening level when assuming that sellers are risk neutral usually leads to a low probability of selling the auctioned timber lot. This undesirable feature was due to a sharp increase in the private value conditional quantile function, which may also be an indication of strong heterogeneity among the bidders. The numerical experiment proposed shows that considering a constant relative risk averse (CRRA) family of utility functions for the seller is sufficient to achieve a resonable high probability of selling the auctioned good.

The third and last Chapter of this thesis considers a quantile framework for first-price auctions under the IPV paradigm. It is shown that a linear specification for the private value conditional quantile function generates a linear specification for its bids counterpart, which then allows for simple and flexible estimation methods for private value linear quantile specifications. This applies to standard quantile regression models as discussed in Chapter 1 as well as to more flexible additive sieve specifications. Both estimation methods can circumvent the curse of dimensionality and hence are specially convenient to handle a large data set with several auction covariates. The estimation method proposed is a combination of local polynomial, quantile regression and additive sieve methods, which allows to estimate the private value quantile function with the optimal rate of Guerre, Perrigne and Vuong (2000) in the case of only one covariate. It is also considered the case of a higher level of interactions among the variables. The estimation method proposed in this third chapter allows to estimate all the quantile levels of the private value conditional distribution without any boundary issue, that is, all bidders' private values, including the highest ones, can be recovered. This is an advantage in view of important nonparametric contributions for the structural analysis of auctions that suffer from boundary bias.

Chapter 1

Econometrics of Ascending Auctions by Quantile Regression

Abstract

This Chapter suggests an identification and estimation approach based on quantile regression to recover the underlying distribution of bidders' private values in ascending auctions under the IPV paradigm. The quantile regression approach provides a flexible and convenient parametrization of the private values distribution, with an estimation methodology easy to implement and with several specification tests. The quantile framework provides a new focus on the quantile level of the private values distribution and on the seller's optimal screening level, which can be both useful for bidders and seller. The empirical application on timber auctions suggests that using policy recommendations from seller's expected payoff may be sometimes inappropriate from a seller's point of view due to the low probability of selling the good. This seems to be an important issue specially in auctions with strong heterogeneity among the bidders, since the seller has incentive to screen some bidders by setting a high reservation price, which in turn leads to a low probability of selling the good.

1.1 Introduction

In an auction theoretical game, the bidders' private value cumulative distribution function (c.d.f) is a key element for analysing the demand that a seller faces. Because it is required for the computation of the seller's expected payoff, its knowledge is crucial for policy recommendation, as e.g. the optimal reservation price policy. The issue here is that bidders' private values are not observed, thereby their distribution function is unknown for the econometricians and policy makers. In the past 20 years, several structural researches in auction theory have proposed parametric, semiparametric and nonparametric approaches to estimate such a latent distribution function.

The goal of this Chapter is to propose an identification and estimation approach based on quantile regression to recover the private value conditional distribution in ascending auctions. The identification strategy is developed for the Independent Private Values (IPV) setup, but can also be extended for other paradigms. The ascending auction is one of the most common design in practice and is especially suitable for the identification of the private value distribution under the IPV paradigm because, under mild assumptions, the transaction price equals the second-highest private value. As a result, the private value distribution can be nonparametrically identified through the winning bids distribution, as well known from Athey and Haile (2002).

A first wave of researchers has focused on parametric identification and estimation approaches. Paarsch (1992) and Donald and Paarsch (1993, 1996) have proposed to estimate the parameters of the private value distribution via the method of Maximum Likelihood (ML) using the winning bids of a button auction. In an attempt to circumvent the difficulties that arise from ML methods, Laffont, Ossard and Vuong (1995) suggested simulated method of moments based on the revenue equivalence theorem, which can handle a larger class of auction mechanisms and parametric models. A key observation of Rezende (2008) is that Laffont, Ossard and Vuong (1995)'s approach carries over with drastic simplifications for simple semiparametric linear regression models which can still be identified through the revenue equivalence theorem. More recently, taking into consideration endogenous entry and unobserved heterogeneity, Li and Zheng (2009) propose a semiparametric

Bayesian method to estimate the distributions of costs, entry costs and of the unobserved auction heterogeneity. Athey, Levin and Seira (2011) also consider an entry stage and adopt a parametric approach to estimate the distribution of the unobserved heterogeneity. Roberts and Sweeting (2012) investigate parametrically the relative revenue performance of a simultaneous outcry auction and a sequential game when entry is costly and selective.

Some works have proposed nonparametric identification and estimation approaches for ascending auctions in an attempt to circumvent the misspecification bias of parametric ones. Athey and Haile (2002) have shown that the c.d.f of the transaction price can nonparametrically identify the private value distribution. Haile and Tamer (2003) identify bounds on the private value distribution for a model of English auction in which bids are below bidder's value. They also suggest to use a median regression in the estimation as an alternative to the nonparametric estimation method, but not really developed the idea.

Some other nonparametric approaches for first-price sealed-bid auctions have been built on quantiles, starting from the insight of Haile, Hong and Shum (2003) that the quantiles of the private value distribution can be written as a function of the quantiles of the observed bids distribution and density function. Marmer and Shneyerov (2010) have shown that adopting a quantile approach makes the estimation of the private value p.d.f. easier from a technical point of view, whereas Guerre and Sabbah (2012) go one step further by proposing to estimate quantile function instead of p.d.f.. Marmer, Shneyerov and Xu (2011) proposed a nonparametric approach to test for alternative models of entry by exploring variation in the quantiles of the private value distribution due to competition. Enache and Florens (2012) developed a nonparametric approach for third-price auctions under risk aversion.

To my knowledge, quantile approaches have not been systematically and throughly applied to ascending auctions so far. Although nonparametric approaches have the advantages of being flexible in analysing the data at hand since no structure is imposed on that, it has some disadvantages as the curse of dimensionality and the need to choose for a bandwidth parameter. The curse of dimensionality can be indeed a relevant estimation issue in view of important contributions to the empirical auction literature such as Haile and Tamer (2003) and Aradillas-Lopez, Gandhi and Quint (2013), which consider, respectively, 5 and 6 exlanatory variables for a sample size of a few thousands at best. Hastie and Tibshirani (1990) give a convenient example of the curse of dimensionality. Suppose that 1000 points are uniformly distributed over a 5-dimensional unit cube and we wish to construct a cube-shaped neighbourhood containing 10 observations, that is, 1% of the data. The subcube is then required to have length 0.40, on average. If instead of 5 the covariate dimension was set to 1, the length required in a similar exercise would now be 1% of the covariate range, which is much smaller and then local than the 40% of the 5 dimensional case. To improve accuracy of the estimates, is then required to use kernels with very small bandwidths and far more observations. This suggests that the bias in standard nonparametric approaches can be high considering the usual auction sample sizes. By contrast, the quantile regression model used in this Chapter can be in principle estimated with a parametric rate, independently of the dimension of the covariate, and does not involve the choice of a smoothing parameter.

Addressing the curse of dimensionality is an important step to develop a nonparametric framework allowing for many covariates. This is also a first step to better capture unobserved heterogeneity by increasing the number of covariates in a nonparametric framework. Specific quantile techniques have also been developed to deal with omitted variables using instrumental variables, see Chernozhukov and Hansen (2006) among others. Such approaches have natural sieve extensions which could also fit in extensions of the framework studied here. Although not so clearly related with the structure of auction data, quantile panel data estimation techniques has provided important recent development, see eg Canay (2011) and Galvao Jr. (2011). This suggests that nonparametric quantile approach could be much more useful regarding unobserved heterogeneity than other nonparametric methods.

The quantile regression approach is also more flexible than many of the parametric or semiparametric methods mentioned above. Indeed, the model includes functional components that may be helpful to reduce the impact of misspecifications. Compared to the semiparametric regression approach of Rezende (2008), quantile regression is computationally more difficult to perform but delivers an estimation of the full private value distribution, as needed for instance to derive an optimal reservation price. As a consequence, quantile regression is probably better suited for policy recommendations than a simpler regression approach. Also in the context of policy analysis, the quantile regression approach allows to highlight the screening level implied by the choice of a reservation price, a policy characteristic that has been mostly ignored by previous empirical approaches. Based on data from the United States Forest Service (USFS) timber auctions, the empirical application suggests that pratical implementation of the optimal reservation price derived from maximization of the expected payoff may lead to a low probability of selling the auctioned good. This is due to a strong variation observed in the private value conditional quantiles, which may then incentive sellers to screen bidders with low valuation in the auction.

This first Chapter is organized as follows. Section 1.2 describes the ascending auction considered, the quantile regression identification approach and a quantile version for the seller's expected payoff with a corresponding screening level maximizer. Section 1.3 provides the estimation methodology and asymptotic results, allowing for potential misspecifications. Section 1.4 presents some specification tests. Section 1.5 provides an empirical application of the methodology proposed. Finally, section 1.6 concludes the Chapter. The Appendix A of this thesis groups all the proofs of the results achieved in this first Chapter, a description of the random weighting bootstrap method used in two of the hypotheses tests and some Monte Carlo experiments.

1.2 Ascending Auctions in a Quantile Setup

A single and indivisible object with some characteristics $Z \in \mathbb{R}^P$ is auctioned to $I \geq 2$ bidders through an ascending auction. The seller sets a reservation price r prior to the auction that is the minimum price that he would be willing to accept. Both the set of auction covariates $X = (1, Z)$ and the number of actual bidders I participating in the auction are common knowledge. The object is sold to the highest bidder for the price of his last bid, provided that it is at least as high as the reservation price r. Within the IPV paradigm, each bidder $i = 1, \cdots, I$ is assumed to have a private value v_i for the auctioned good, which is not observed by other bidders. The bidder only knows his own private value, but it is common knowledge for bidders and sellers that private values have been identically and independently drawn from a common c.d.f. $F_v(\cdot|X, I)$ conditional upon (X, I), or equivalently, with a conditional quantile function $V(\alpha|X, I)$, $\alpha \in [0, 1]$, defined as

$$V(\alpha|X, I) := \inf \{v : F_v(v|X, I) \geq \alpha\}. \quad (1.2.1)$$

When the private value conditional distribution is absolutely continuous with a probability density function (p.d.f) $f_v(\cdot|X, I)$ positive on its support $[V(0|X, I), V(1|X, I)] \subset \mathbb{R}+$, as considered from now on, $V(\alpha|X, I)$ is the reciprocal function $F_v^{-1}(\alpha|X, I)$.

It is well known that $u_i = F_v(v_i|X, I)$, which can be viewed as the rank of a bidder with private value v_i in the population, is independent of (X, I) with a uniform distribution over $[0, 1]$. The IPV paradigm implies that the ranks u_i, $i = 1, \cdots, I$, are independent. In other words, the dependence between the private values v_i and the auction covariates (X, I) can be fully captured by the nonseparable model

$$v_i = V(u_i|X, I), \quad u_i \overset{iid}{\sim} \mathcal{U}_{[0,1]} \perp (X, I). \quad (1.2.2)$$

Therefore, bidders are identical up to the variable u_i, which represents the bidder ith's position in the private value distribution.

The quantile regression approach, developed by Koenker and Bassett (1978), restricts the quantile representation (1.2.2) to a regression specification, such as

$$\begin{aligned} V(\alpha|X, I) &= h(X\gamma(\alpha|I)) \\ &= h(\gamma_0(\alpha|I) + Z\gamma_1(\alpha|I)), \end{aligned} \quad (1.2.3)$$

where $h(\cdot)$ is a given function, $\gamma_0(\alpha|I)$ the quantile regression intercept and $\gamma_1(\alpha|I)$ the quantile regression slopes. In the basic specification, $h(\cdot)$ is equal to the identity. Note that in (1.2.3), both the intercept and the slope quantile regression coefficients depend upon the rank α of the bidder in the population[1]. Therefore, changes in the conditioning variables not only shift the location of the conditional distribution of v, but may also affect its scale and shape. A shock on the covariate X may affect a bidder with a low rank α in a different way than a bidder with a higher rank. Strong variation of the slope and location coefficients, $\gamma_1(\alpha|I)$ and $\gamma_0(\alpha|I)$, indicates large

[1]Throughout this chapter, the rank of the bidder can be interchangeably represented by α and u.

heterogeneity among the bidders. In the empirical section, the term "heterogeneity" will be used for brevity to indicate such a strong variation[2] in the private value quantiles. As discussed later, taking into consideration such heterogeneity among the bidders may have important implications for both seller and bidders.

We now turn to the assumptions of the model. In the considered ascending auction, bidders raise continuously their prices and drop out of the auction as the prices reach their valuation.

Assumption 1.1. *The transaction price in an auction is the greater of the reservation price and the second-highest bidder's willingness to pay.*

Assumption 1.1 is an assumption on equilibrium play. This assumption was also used in Aradillas-Lopez et al. (2013) and, as noted in Athey and Haile (2002) and Bikhchandani, Haile and Riley (2002), is compatible with the multiple equilibra generated by the ascending auctions. It is for instance the result of the dominant strategy equilibrium of a button[3] auction, which is a stylized version of an ascending auction. Haile and Tamer (2003) use stronger assumptions concerning bidder's behaviour, which determine the joint distribution of all the bids. This is not needed when using only the winning bid. This assumption would also hold approximately in the context of Haile and Tamer (2003) if bidders do not use jump bids at the end of the auction.

The ascending auction format is specially convenient for the identification of the bidders' private value distribution under the IPV paradigm[4] because, under assumption 1.1 and a nonbinding reservation price, the latent private value conditional distribution can be nonparametrically identified from the winning bid conditional distribution. Such a nice feature will be considered in the identification of the model.

[2]Note that heterogeneity and asymmetry are two concepts that should not be confused. Heterogenity is concerned with the variation of $\gamma(\alpha)$ across quantile levels α, while asymmetry implies that different bidders can have different coefficients $\gamma(\alpha)$.

[3]In a button auction, bidders hold down a button to remain active while the price rises continuously, releasing the button to drop out of the game, and the willingness to pay of the losing bidders is learned from their drop out prices.

[4]In the case of affiliated private values, the private value conditional quantiles cannot be nonparametrically point-identified. It is possible however to extend the methodology to identify bounds on the private value conditional quantiles using a strategy similar to Aradillas-Lopez et al. (2013).

The next two assumptions deal with the quantiles of the bidders' private value distribution:

Assumption 1.2. *$V(\alpha|X,I)$ is strictly increasing and continuous on its support $[V(0|X,I), V(1|X,I)]$ for all (X,I).*

Assumption 1.3. *The private value conditional quantiles has a quantile regression specification*

$$V(\alpha|X,I) = h(X\gamma(\alpha|I)), \tag{1.2.4}$$

where $h(\cdot)$ is known by the econometrician.

Assumption 1.2 is usual in the quantile regression literature. Assumption 1.3 imposes correct specification of the private value conditional quantiles as necessary for our identification results. This will be relaxed when studying estimation of the model.

As shown in the next Lemma, the rank α of a bidder in the population has a direct relationship with his probability of winning the auction, so that estimating $V(\alpha|X,I)$ can then be used as a benchmark for bidders interested in increasing their chance of winning the auction[5].

Lemma 1.1. *Under the IPV paradigm and assumptions 1.1-1.2, a bidder with private value $V(\alpha|X,I)$ wins with probability α^{I-1}.*

The proof of Lemma 1.1 is given in Appendix A.1 of this thesis, which also groups the proof of all the results stated in this Chapter.

Define

$$\Psi_I(t) = It^{I-1} - (I-1)t^I.$$

and let $B(\alpha|X,I)$ be the α-quantile of the winning bids conditional distribution given (X,I). It follows from Athey and Haile (2002, equation (5)) that $\Psi_I(F_v(\cdot|X,I))$ is the distribution of the second-highest private value, which is equal to the winning bid. This gives the following cornerstone quantile identification result.

[5] Consider e.g. a procurement auction: if a bidder can estimate and target a low quantile α of the private cost distribution, he can then work on his technological capacity to decrease his production cost and be able to offer a lower price for the buyer, thereby increasing his probability of winning the auction.

Lemma 1.2. *Under IPV and assumptions 1.1-1.2, for each I and $\alpha \in [0,1]$,*

$$V(\alpha|X,I) = B(\Psi_I(\alpha)|X,I). \tag{1.2.5}$$

Lemma 1.2 shows that the private value conditional quantile function can be nonparametrically identified by the conditional quantile of the observed winning bid. Before introducing the quantile regression identification result, we require the following assumption for the auction specific covariate $X = (1, Z)$:

Assumption 1.4. *The auction specific variable, Z, has dimension P, with a compact support in $\mathcal{Z} \subset (0,+\infty)^P$ and a nonempty interior.*

Assumption 1.4 ensures that if $x\gamma_1 = x\gamma_2$, for all $x \in \mathcal{X}$, thus $\gamma_1 = \gamma_2$, and is necessary for the quantile regression identification below.

Lemma 1.3. *Under IPV and assumptions 1.1-1.4,*

(i) There exists, for each $\alpha \in [0,1]$, a vector of coefficients $\beta(\alpha|I)$ such that

$$B(\alpha|X,I) = h(X\beta(\alpha|I));$$

(ii) $\beta(\alpha|I)$ is uniquely defined and satisfies

$$\beta(\Psi_I(\alpha)|I) = \gamma(\alpha|I). \tag{1.2.6}$$

Result (i) in Lemma 1.3 is a stability property of the quantile regression specification, which is a consequence of Lemma 1.2. Indeed, Lemma 1.2 shows that the winning bid quantile function admits the same specification that the one postulated for the private values, but for a transformed quantile level. Lemma 1.3-(ii) gives the identification result of the quantile regression approach. It shows that the coefficient $\gamma(\cdot|I)$ of the private value conditional quantile function is identified through the coefficient $\beta(\cdot|I)$ of the winning bid conditional quantile function, but evaluated at a different quantile level $\Psi_I(\cdot)$.

Dependence upon Bidders Participation

It is assumed for now that the private value distribution is conditional upon the number of bidders I participating in the auction. Although standard in many econometric works as Guerre, Perrigne and Vuong (2000) among others, conditioning on I is not usual in theoretical auction models, see e.g. Krishna (2010). This choice can be however motivated by the three following reasons which consider either unobserved heterogeneity or an entry stage.

A first setting of interest is unobserved heterogeneity. Instead of the observed X, the bidders use an auction characteristic (X, X_u) which includes a component X_u that is not observed by the econometrician. Hence, the private value quantile relevant for policy analysis is $V(\alpha|X, X_u)$, which cannot be estimated without further assumption. It can be for instance assumed that the actual number of bidders depends upon the auction characteristic, that is $I = I(X, X_u)$, in a way that fully captures the impact of the unobserved characteristic, i.e. $V(\cdot|X, X_u) = V(\cdot|X, I(X, X_u))$ so that the conditional quantile $V(\cdot|X, I)$ is fully relevant for policy analysis purposes.

A second motivation is given by the recent econometric literature on entry, see Gentry and Li (2012), Li and Zheng (2009) and Marmer et al. (2011). These models consider a two stage game, where the first stage is entry and the second stage is the auction game. The structural parameter is the joint distribution of the private values and signals given the characteristic X, which is used in the entry stage of the game. The second stage involves an actual number of bidders I, who have decided to participate in the auction, and the conditional quantile $V(\cdot|X, I)$ of private values given X and I. A key contribution of the aforementioned econometric literature is that the structural parameter is identified from I and $V(\cdot|X, I)$, so that estimation of the model can be performed through estimation of the conditional c.d.f. or quantile of private values given X and I.

A third motivation notes that, whereas the reality of the auction is physically clear, the entry stage of the game described above may have a more conceptual nature. Therefore, in some cases, the importance of a entry stage is an assumption that should be tested by investigating whether $V(\cdot|X, I)$ depends indeed upon I. A test built on the null hypothesis of independence, i.e. $V(\cdot|X) = V(\cdot|X, I)$, is proposed in section 1.4 and then applied in the empricial section.

Optimal Reservation Price

Consider a binding reservation price set by the seller, i.e. $r(X,I) \in [V(0|X,I), V(1|X,I)]$. The reservation price thus plays the role of a screening level in the auction since bidders with $V(\alpha_i|X,I) < r(X,I)$ are prevented from participating in the game. Let α_r be the screening level in the private value conditional distribution, i.e. α_r is such that $r(X,I) = V(\alpha_r|X,I)$. It thus represents the percentage of bidders in the population that are not participating in the auction because of a low valuation. Note that the auctioned good will not be sold if all the players have valuation below $r(X,I)$, which implies that the probability of trading is $1-\alpha_r^I$. Therefore, for a given I, the probability of trading decreases with the screening level α_r.

Let the seller's payoff be defined as

$$\pi(r(X,I)) = b_w \mathbb{I}(b_w \geq r(X,I)) + v_0(X)(1-\mathbb{I}(b_w \geq r(X,I))), \tag{1.2.7}$$

where b_w is the winning bid and $v_0(X)$ the seller's private value[6]. In what follows, let $v_0 = v_0(X)$ and $r = r(X,I)$. The following proposition gives a quantile version for the seller's expected payoff, a candidate for the optimal screening level $\alpha_r^* = \alpha_r^*(X,I,v_0)$ and the corresponding optimal reservation price $V(\alpha_r^*|X,I)$. Let $\Pi(\alpha_r|X,I,v_0)$ be the seller's expected payoff given (X,I) when the screening level is α_r.

Proposition 1.1. *Under IPV and assumptions 1.1-1.2,*

(i) the seller's expected payoff is

$$\begin{aligned}\Pi(\alpha_r|X,I,v_0) = v_0\alpha_r^I &+ V(\alpha_r|X,I)\, I\alpha_r^{I-1}(1-\alpha_r) \\ &+ I(I-1)\int_{\alpha_r}^1 V(\alpha|X,I)\,\alpha^{I-2}(1-\alpha)\,d\alpha,\end{aligned} \tag{1.2.8}$$

where v_0 is the seller's private value;

[6]The seller's profit is given by his payoff minus his private value. We will focus our analysis on the seller's payoff.

(ii) The optimal reservation price $V(\alpha_r^*|X,I)$ *satisfies*

$$V(\alpha_r^*|X,I) - V^{(1)}(\alpha_r^*|X,I)(1-\alpha_r^*) = v_0, \tag{1.2.9}$$

where $V^{(1)}(\alpha_r^*|X,I)$ *is the private value quantile density*[7] *function.*

Equation (1.2.8) in Proposition 1.1 gives the seller's expected payoff in a quantile setup. It differs from the expression given in Riley and Samuelson (1981, Proposition 1) because it does not involve the private value conditional density. It is nevertheless convenient for estimation purposes because, in a nonparametric setup, $V(\cdot|X,I)$ can be estimated in a faster rate than $V^{(1)}(\cdot|X,I)$. Equation (2.2.3) is the first-order condition (FOC) associated with the maximization of the seller's expected payoff and represents a quantile version of a well known formula that states that the optimal reservation price satisfies $1 - F_v(r^*|X,I) - (r^* - v_0) f_v(r^*|X,I) = 0$, see Krishna (2010, p.23), Riley and Samuelson (1981, Proposition 3) and Myerson (1981). An interesting implication of (2.2.3) is that when $V(\alpha|X,I)$ does not depend upon I the optimal reservation price $V(\alpha_r^*|X)$ and the optimal screening level α_r^* also do not depend on I, as well known from the aforementioned reference. In addition, since $V^{(1)}(\cdot|X,I) > 0$, it is clear that $\alpha_r^* > \alpha_0$, where $v_0 = V(\alpha_0|X,I)$. That is, the optimal reservation price is above v_0.

Although equation (2.2.3) in Proposition 1.1 gives a closed form to estimate the optimal reservation price $V(\alpha_r^*|X,I)$, it involves an estimation of the quantile density function $V^{(1)}(\cdot|X,I)$. It is, therefore, better to maximize an estimation of $\Pi(\alpha_r|X,I,v_0)$ to get $\widehat{\alpha}_r^*$ than to solve an estimation of the quantile FOC[8]. To estimate the seller's expected payoff, we first estimate $V(\cdot|X,I)$ by using the quantile regression methodology proposed in the next section and then apply numerical integration via a trapezoildal rule over a grid of quantiles $\mathcal{A} = \{\alpha_1, \alpha_2 \cdots, \alpha_K\}$[9] to

[7] The quantile density function is defined as the derivative of the quantile function with respect to α, i.e. $V^{(1)}(\alpha|X,I) = \partial V(\alpha|X,I)/\partial\alpha = 1/f_v(V(\alpha|X,I)|X,I)$.

[8] Li, Perrigne and Vuong (2003) have also used the same strategy to estimate the optimal reservation price in the case of affiliated private values, which includes the IPV as a special case. They define such a strategy as semiparametric, since the optimal reservation price is obtained as the maximizer of an estimated expected payoff in which the distributions and densities were nonparametrically estimated in a first step.

[9] In the empirical application, the grid of prescribed quantiles used for this numerical integration is $\mathcal{A} = \{0.12, 0.14, \cdots, 0.80\}$.

estimate the definite integral in (1.2.8). The optimal screening level $\widehat{\alpha}_r^*$ is then chosen over $\mathcal{A}$ such that $\widehat{\Pi}(\alpha_r|X,I,v_0)$ is maximized, where the latter is an estimation of $\Pi(\alpha_r|X,I,v_0)$. As mentioned above, when $V(\cdot|X) = V(\cdot|X,I)$, the optimal screening level does not depend upon I. In the empirical section, this case is considered and $\widehat{\alpha}_r^*$ is therefore estimated by maximizing an aggregate $\widehat{\Pi}(\alpha_r|X,v_0)$, i.e.

$$\widehat{\alpha}_r^* = \arg\max \sum_{I=2}^{\overline{I}} \widehat{\Pi}(\alpha_r|X,I,v_0) L_I/L, \tag{1.2.10}$$

where $\{2,\cdots,\overline{I}\}$ is the support of I, L_I is the number of auctions in the sample with I bidders competing and L the total sample size. The intuition here is that aggregating over I may potentially give a better estimation $\widehat{\alpha}_r^*$.

A comparison of the quantile regression estimation methodology with a nonparametric alternative is also presented in the end of the empirical section. For the nonparametric approach, the conditional distribution of private values is used instead of their quantiles. Consider that $\mathbb{P}(v_{I:I} < r|X,I) = F_v^I(r|X,I)$, $\mathbb{P}(v_{I-1:I} < r \leq v_{I:I}|X,I) = F_{b_w}(r|X,I) - F_v^I(r|X,I)$ and

$$\mathbb{E}\left[v_{I-1:I}\mathbb{I}(r \leq v_{I-1:I})|X,I\right] = \int_r^{\overline{v}} v f_{b_w}(v|X,I)\,dv,$$

where $F_{b_w}(\cdot|X,I)$ and $f_{b_w}(\cdot|X,I)$ are, respectively, the c.d.f. and p.d.f. of the winning bids conditional on (X,I). From the proof of Proposition 1.1, the seller's expected payoff can be then written as

$$\begin{aligned}\Pi(r|X,I,v_0) = v_0 F_v^I(r|X,I) + r\left(F_{b_w}(r|X,I) - F_v^I(r|X,I)\right) \\ + \int_r^{\overline{v}} v f_{b_w}(v|X,I)\,dv.\end{aligned} \tag{1.2.11}$$

To estimate $\Pi(r|X,I,v_0)$, we first estimate the winning bid conditional c.d.f. and p.d.f., then apply the transformation $\Psi_I^{-1}(\cdot)$ to find the solution of $\widehat{F}_v(\cdot|Z,I) = \Psi_I^{-1}\left(\widehat{F}_{b_w}(\cdot|Z,I)\right)$ and estimate the integral term using numerical integration as above. The winning bid conditional c.d.f. and p.d.f can be nonparametrically estimated[10] via

$$\widehat{F}_{b_w}\left(b_w|Z,I\right)=\frac{\frac{1}{L_I h_Z^P}\sum_{\ell=1}^{L_I}\mathbb{I}\left(I_\ell=I\right)\mathbb{I}\left(b_{w\ell}\leq b_w\right)K\left(\frac{Z-Z_\ell}{h_Z}\right)}{\frac{1}{L_I h_Z^P}\sum_{\ell=1}^{L_I}\mathbb{I}\left(I_\ell=I\right)K\left(\frac{Z-Z_\ell}{h_Z}\right)} \tag{1.2.12}$$

and

$$\widehat{f}_{b_w}\left(b_w|Z,I\right)=\frac{\frac{1}{L_I h_Z^P}\sum_{\ell=1}^{L_I}\mathbb{I}\left(I_\ell=I\right)\mathbb{I}\left(b_{w\ell}\leq b_w\right)K\left(\frac{Z-Z_\ell}{h_Z}\right)}{\frac{1}{L_I h_Z^P}\sum_{\ell=1}^{L_I}\mathbb{I}\left(I_\ell=I\right)K\left(\frac{Z-Z_\ell}{h_Z}\right)}, \tag{1.2.13}$$

where $K\left(u\right)$ is a kernel function and h_Z a vanishing bandwidth. In the estimation, a triweight kernel, $K\left(u\right)=\frac{35}{32}\left(1-u^2\right)^3\mathbb{I}\left(|u|\leq 1\right)$, is considered and the bandwidth parameter is computed using the rule $h_z^*=\sigma_{zz}L_I^{-1/5}$ as suggested by Silverman (1998), where σ_{zz} is the standard deviation of the corresponding variable Z. In the case that $F_v\left(\cdot|X\right)=F_v\left(\cdot|X,I\right)$, the estimated private values conditional distribution is then aggregated over I via

$$\widehat{F}_v\left(\cdot|Z\right)=\sum_{I=2}^{\overline{I}}\widehat{F}_v\left(\cdot|Z,I\right)L_I/L. \tag{1.2.14}$$

Note that it is not necessary to compute $\widehat{f}_v\left(\cdot|Z\right)$ in the approach described above. It is nevertheless possible to estimate the optimal reservation price using

$$1-\widehat{F}_v\left(r^*|Z\right)-\left(r^*-v_0\right)\widehat{f}_v\left(r^*|Z\right)=0,$$

where $\widehat{f}_v\left(\cdot|Z\right)$ is also aggregated over I as in (1.2.14). Lu and Perrigne (2008) have used the same strategy as above to estimate $F_v\left(\cdot|Z,I\right)$ and $f_v\left(\cdot|Z,I\right)$ in their risk aversion analysis. They however assume that $F_v\left(\cdot|X,I\right)$ depend upon I and therefore do not aggregate the c.d.f. and p.d.f. over I.

[10] Recall that $X=(1,Z)$. Therefore, the constant term will not be considered in the nonparametric approach.

1.3 Estimation Methodology and Asymptotic Properties

Consider independent and identically distributed observations $(b_{w_\ell}, Z_\ell, \ell = 1, \ldots, L)$, where b_{w_ℓ} is the winning bid at auction ℓ, and Z_ℓ a specific characteristic of the good auctioned in auction ℓ. Let $X_\ell = (1, Z_\ell) \in \mathcal{X}$ be a row vector of dimension $P + 1$ and $\mathcal{X} = \{1\} \times \mathcal{Z}$. Define L_I as the number of auctions with I players,

$$L_I = \sum_{\ell=1}^{L} \mathbb{I}\left(I_\ell = I\right)$$

for $I \in \mathcal{I}$, where $\mathcal{I}$ is a bounded subset of $\{2, 3, \ldots\}$.

The following assumption concerns the variables in our model:

Assumption 1.5. *The variables $\{I_\ell, X_\ell, v_{i\ell},\ i = 1, 2, \ldots I_\ell, \ell = 1, \ldots, L\}$ are independent and identically distributed. Conditional on (X_ℓ, I_ℓ), the private values $v_{i\ell}$ are independent with common c.d.f. $F_v\left(\cdot|X_\ell, I_\ell\right)$ and a density function $f_v\left(\cdot|X_\ell, I_\ell\right)$ bounded away from zero.*

Assumption 1.5 implies that each auction is independent and that, within an auction, the IPV paradigm holds. Note that we do not assume that I_ℓ is independent upon X_ℓ. Recall that $V\left(\cdot|X_\ell, I_\ell\right) = F_v^{-1}\left(\cdot|X_\ell, I_\ell\right)$.

Assumption 1.3 from the preceding section was considering a correct specification of the private value quantile regression model as necessary for identification. This is not requested in our estimation setup and Assumption 1.6 below, which does not assume a correct specification and considers instead the function $h\left(\cdot\right)$, will be used instead of Assumption 1.3.

Assumption 1.6. *The function $h\left(t\right)$ used in the private values regression model is a given continuous, monotonically increasing and twice differentiable function.*

We are interested in studying two special functions to characterize the bidders' private values: the identity function $h\left(t\right) = t$ and the exponential function $h\left(t\right) = \exp\left(t\right)$.

Consider the winning bid b_{w_ℓ} at the auction ℓ, $\ell = 1, \cdots, L$, which is the amount paid by the winner. From the quantile regression identification result (1.2.6) in Lemma 1.3, $\gamma(\alpha|I)$ can be estimated through

$$\widehat{\gamma}(\alpha|I) = \widehat{\beta}(\Psi_I(\alpha)|I).$$

I propose to estimate the winning bid quantile regression population parameter $\beta(\alpha|I)$ via the method of quantile regression,

$$\widehat{\beta}(\alpha|I) = \arg\min_{\beta\in\Gamma} \frac{1}{L_I}\sum_{\ell=1}^{L_I} \mathbb{I}(I_\ell = I)\,\rho_\alpha(b_{w_\ell} - h(X_\ell\beta))$$

where $\rho_\alpha(u) = u(\alpha - \mathbb{I}(u < 0))$ and Γ is a compact subset of $\mathbb{R}^{P+1}$. In other words,

$$\widehat{\gamma}(\alpha|I) = \arg\min_{\gamma\in\Gamma} \frac{1}{L_I}\sum_{\ell=1}^{L_I} \mathbb{I}(I_\ell = I)\,\rho_{\Psi_I(\alpha)}(b_{w_\ell} - h(X_\ell\gamma)). \tag{1.3.15}$$

In what follows, let

$$Q(\gamma|\alpha, I) = \mathbb{E}\left[\rho_{\Psi_I(\alpha)}(b_w - h(X\gamma))|I\right],$$

$$\widehat{Q}(\gamma|\alpha, I) = \frac{1}{L_I}\sum_{\ell=1}^{L_I} \mathbb{I}(I_\ell = I)\,\rho_{\Psi_I(\alpha)}(b_{w_\ell} - h(X_\ell\gamma))$$

and

$$\widehat{Q}(\widehat{\gamma}|\alpha, I) = \widehat{Q}(\widehat{\gamma}(\alpha|I)|\alpha, I) = \min_{\gamma\in\Gamma} \widehat{Q}(\gamma|\alpha, I) \tag{1.3.16}$$

be the population, the empirical and the optimized quantile regression objective functions. I will denote the first and second derivatives of $Q(\gamma|\alpha, I)$ with respect to γ respectively by $Q_\gamma(\gamma|\alpha, I)$ and $Q_{\gamma\gamma}(\gamma|\alpha, I)$, whereas $h^{(1)}(\cdot)$ and $h^{(2)}(\cdot)$ will be respectively the first and second derivatives of $h(\cdot)$.

In the case of a misspecified model, $\widehat{\gamma}(\alpha|I)$ is expected to converge to a pseudo-true private value quantile regression coefficient defined as

$$\gamma^*(\alpha|I) = \arg\min_{\gamma\in\Gamma} Q(\gamma|\alpha, I), \tag{1.3.17}$$

where the expectation is taken with respect to the true model distribution. In what follows, it is assumed that the considered pseudo-true values are uniquely defined.

Theorem 1.1 gives the asymptotic distribution of the private value quantile regression estimator (1.3.15). Its proof relies on arguments sketched in Newey and McFadden (1994) for estimators optimizing non smooth objective functions.

Theorem 1.1. *Under assumptions 1.1-1.2, 1.4-1.6 and if $\gamma^*(\alpha|I)$ from (1.3.17) is an inner point of Γ,*

$$\sqrt{L_I}\left(\widehat{\gamma}(\alpha|I)-\gamma^*(\alpha|I)\right)\xrightarrow{d}\mathcal{N}\left(0,Q_{\gamma\gamma}^{-1}(\gamma^*|\alpha,I)\,J(\gamma^*|\alpha,I)\,Q_{\gamma\gamma}^{-1}(\gamma^*|\alpha,I)\right),$$

where

$$\begin{aligned}J(\gamma^*|\alpha,I) &= \mathbb{E}\Big[\Big\{F_{b_w}\left(h\left(X\gamma^*(\alpha|I)\right)|X,I\right)-2\Psi_I(\alpha)\,F_{b_w}\left(h\left(X\gamma^*(\alpha|I)\right)|X,I\right)\\ &\qquad +\Psi_I(\alpha)^2\Big\}h^{(1)}\left(X\gamma^*(\alpha|I)\right)^2X'X|I\Big]\end{aligned}$$

and

$$\begin{aligned}Q_{\gamma\gamma}(\gamma^*|\alpha,I) &= \mathbb{E}\left[f_{b_w}\left(h\left(X\gamma^*(\alpha|I)\right)|X,I\right)h^{(1)}\left(X\gamma^*(\alpha|I)\right)^2X'X|I\right]+\\ &\qquad \mathbb{E}\left[\left(F_{b_w}\left(h\left(X\gamma^*(\alpha|I)\right)|X,I\right)-\Psi_I(\alpha)\right)h^{(2)}\left(X\gamma^*(\alpha|I)\right)X'X|I\right],\end{aligned}$$

$F_{b_w}(\cdot|X,I)=\Psi_I\left(F_v(\cdot|X,I)\right)$ and $f_{b_w}(\cdot|X,I)$ being the c.d.f. and p.d.f. of the winning bids given (X,I).

Although $\gamma(\alpha|I)$ is a parameter of the private value distribution, the asymptotic variance of $\widehat{\gamma}(\alpha|I)$ in the Theorem is computed using the winning bids distribution. Note that if the model is correctly specified, then $\gamma^*(\alpha|I)=\gamma(\alpha|I)$ and $F_{b_w}\left(h\left(X\gamma^*(\alpha|I)\right)|X,I\right)=\Psi_I(\alpha)$, so that

$$J(\gamma|\alpha,I)=\Psi_I(\alpha)\left(1-\Psi_I(\alpha)\right)\mathbb{E}\left[h^{(1)}\left(X\gamma(\alpha|I)\right)^2X'X|I\right],$$

and

$$Q_{\gamma\gamma}(\gamma|\alpha,I)=\mathbb{E}\left[f_{b_w}\left(h\left(X\gamma(\alpha|I)\right)|X,I\right)h^{(1)}\left(X\gamma(\alpha|I)\right)^2X'X|I\right].$$

The asymptotic variance of the quantile regression estimator can be estimated using techniques described in Koenker (2005). The applications considered here uses bootstrap inference and for this reason the variance estimation aspects are not detailed.

If the private value distribution is independent upon I given X, then

$$\begin{aligned}\gamma^*(\alpha|I) &= \arg\min_{\gamma\in\Gamma} \mathbb{E}\left[\rho_{\Psi_I(\alpha)}\left(b_w - h\left(X\gamma\right)\right)|I\right]\\ &= \arg\min_{\gamma\in\Gamma} \mathbb{E}\left[\rho_{\Psi_I(\alpha)}\left(b_w - h\left(X\gamma\right)\right)\right]\\ &= \gamma^*(\alpha),\end{aligned}$$

which opens space for a potential improvement in the estimation efficiency since samples with different number of bidders I_ℓ can be pooled, increasing the sample size to $L \geq L_I$. This therefore leads to consider

$$\widehat{\gamma}(\alpha) = \arg\min_{\gamma\in\Gamma} \widehat{Q}(\gamma|\alpha), \tag{1.3.18}$$

where

$$\widehat{Q}(\gamma|\alpha) = \frac{1}{L}\sum_{\ell=1}^{L} \rho_{\Psi_{I_\ell}(\alpha)}\left(b_{w_\ell} - h\left(X_\ell\gamma\right)\right).$$

Let

$$\gamma^*(\alpha) = \arg\min_{\gamma\in\Gamma} Q(\gamma|\alpha) \text{ where } Q(\gamma|\alpha) = \mathbb{E}\left[\rho_{\Psi_I(\alpha)}\left(b_w - h\left(X\gamma\right)\right)\right].$$

The following Corollary gives the asymptotic distribution of the pooled quantile regression estimator.

Corollary 1.1. *Under assumptions 1.1-1.2 and 1.4-1.6,*

$$\sqrt{L}\left(\widehat{\gamma}(\alpha) - \gamma^*(\alpha)\right) \xrightarrow{d} \mathcal{N}\left(0, Q_{\gamma\gamma}^{-1}(\gamma^*|\alpha)\, J(\gamma^*|\alpha)\, Q_{\gamma\gamma}^{-1}(\gamma^*|\alpha)\right),$$

where $\gamma^*(\alpha) = \arg\min_{\gamma\in\Gamma} Q(\gamma|\alpha)$*, with* $Q(\gamma|\alpha) = \mathbb{E}\left[\widehat{Q}(\gamma|\alpha)\right]$*, is the pseudo-true private value quantile regression coefficient,*

$$\begin{aligned} J(\gamma^*|\alpha) \quad = \quad & \mathbb{E}\left[\left\{F_{b_w}\left(h\left(X\gamma^*(\alpha)\right)|X,I\right) - 2\Psi_I(\alpha)\, F_{b_w}\left(h\left(X\gamma^*(\alpha)\right)|X,I\right) + \Psi_I(\alpha)^2\right\}\right.\\ & \left.\times h^{(1)}\left(X\gamma^*(\alpha)\right)^2 X'X\right]\end{aligned}$$

and

$$Q_{\gamma\gamma}(\gamma^*|\alpha) = \mathbb{E}\left[f_{b_w}\left(h\left(X\gamma^*(\alpha)\right)|X,I\right)h^{(1)}\left(X\gamma^*(\alpha)\right)^2 X'X\right] + \\ \mathbb{E}\left[\left(F_{b_w}\left(h\left(X\gamma^*(\alpha)\right)|X,I\right) - \Psi_I(\alpha)\right)h^{(2)}\left(X\gamma^*(\alpha)\right)X'X\right],$$

$F_{b_w}(\cdot|X,I) = \Psi_I(F_v(\cdot|X))$ *and* $f_{b_w}(\cdot|X,I)$ *being the c.d.f. and p.d.f. of the winning bids given* (X,I).

Since L can be substantially larger than L_I, Corollary 1.1 suggests that $\widehat{\gamma}(\alpha)$ can considerably improve on $\widehat{\gamma}(\alpha|I)$ when the private values distribution is independent from I given X.

The regression specification of Rezende (2008)[11] can offer another interesting source of improvement under special conditions. Assume for instance that, for $h(\cdot)$ as in Assumption 1.6,

$$v_{i\ell} = h\left(\gamma_0 + Z_\ell\gamma_1 + \varepsilon_{i\ell}\right), \quad (1.3.19)$$

where the $\varepsilon_{i\ell}$ are i.i.d. and independent of (Z_ℓ, I_ℓ). In this case the conditional quantile function of the private values is

$$V(\alpha|Z) = h\left(\gamma_0(\alpha) + Z\gamma_1\right) \text{ where } \gamma_0(\alpha) = \gamma_0 + F_\varepsilon^{-1}(\alpha). \quad (1.3.20)$$

For such specification, pooling over α can improve the estimation of the slope coefficients γ_1 as proposed by Zou and Yuan (2008) with the Composite Quantile Regression (CQR) estimator

$$\widehat{\gamma}_{\text{CQR}} = \arg\min_{\gamma_{0,1},\ldots,\gamma_{0,K},\gamma_1} \frac{1}{K}\sum_{k=1}^{K}\left\{\frac{1}{L}\sum_{\ell=1}^{L}\rho_{\Psi_{I_\ell}(\alpha_k)}\left(b_{w_\ell} - h\left(\gamma_{0,k} + Z_\ell\gamma_1\right)\right)\right\}, \quad (1.3.21)$$

where $\widehat{\gamma}_{\text{CQR}} = \left(\widehat{\gamma}_0(\alpha_1), \ldots, \widehat{\gamma}_0(\alpha_K), \widehat{\gamma}_1\right)'$ and $\alpha_k \in [0,1]$, $k = 1,\ldots,K$, are some quantile levels. Zou and Yuan (2008) have shown in particular that the CQR estimator can improve on least squares estimation. That estimator is asymptotically normal with a limit variance that can be derived with arguments similar to the ones used in Theorem 1.1. However, implementing this strategy with the considered

[11] For simplicity of notation, it is assumed independence of the private value distribution upon I.

auction dataset has led to a rejection of the correct specification of (1.3.20) (that is, a rejection of $\gamma_1(\alpha) = \gamma_1$ for all retained α_k) so that the limit distribution of the CQR estimator is not detailed here for the sake of brevity.

1.4 Auction Hypotheses Testing

The estimation methodology developed above enables us to investigate some interesting questions for the empirical auction literature that are grouped in the following three hypotheses tests: exclusion participation restriction, functional form of the private values conditional quantile function and constancy in the impact of the auction characteristics across the private values distribution. In what follows, $\mathcal{A} = \{\alpha_1, \cdots, \alpha_K\}$ is a set of prescribed quantile levels used in the considered test statistics.

Testing the Exclusion Participation Restriction

As discussed in section 1.2, testing the condition $V(\cdot|X, I) = V(\cdot|X)$ gives indication about potential unobserved heterogeneity or presence of a entry stage in the auction game. An important policy implication of that condition is that the optimal reservation price is also not dependent upon the number of actual bidders I as usual in the auction literature. In addition, as discussed in the previous section, the exclusion participation restriction enables a potential improvement in the estimation efficiency. The considered null and alternative hypotheses are

$$\begin{aligned} H_0 &: \gamma^*(\alpha|I) = \gamma^*(\alpha) \text{ for all } \alpha \in \mathcal{A} \text{ and } I \in \mathfrak{I} \\ H_1 &: \text{not } H_0. \end{aligned}$$

Note that even if $V(\cdot|X, I)$ differs from $h(X\gamma(\cdot|I))$, i.e. the function $h(\cdot)$ is misspecified, the pseudo-true coefficients still satisfy $\gamma^*(\cdot|I) = \gamma^*(\cdot)$ when $V(\cdot|X, I) = V(\cdot|X)$.

A simple way of testing the null hypothesis above is via a Wald test, jointly for all the coefficients and quantile levels. However, the latter involves standardization of the test-statistic by the variance-covariance matrix of the coefficients, which in turn involves estimation of the unknown density function of the random errors. The estimation of the latter requires either a bandwidth choice (see Powell (1991)) or bootstrap resampling methods (see Buchinsky (1995)). Some preliminary experiments had suggested that an alternative strategy as described below may give better results.

A strategy similar to a maximum likelihood ratio test can be implemented to avoid the estimation of the variance-covariance matrix. Let $\widehat{Q}\left(\widehat{\gamma}|\alpha\right)$ represents the optimized pooled objective function, i.e.

$$\widehat{Q}\left(\widehat{\gamma}|\alpha\right) = \widehat{Q}\left(\widehat{\gamma}\left(\alpha\right)|\alpha\right) = \min_{\gamma\in\Gamma} \widehat{Q}\left(\gamma|\alpha\right)$$

and $\widehat{Q}\left(\widehat{\gamma}|\alpha, I\right)$ the optimized individual objective function as defined in (1.3.16). Under the null hypothesis of independence, $Q\left(\gamma|\alpha\right) = \sum_{I\in\mathcal{I}} Q\left(\gamma|\alpha, I\right)\mathbb{P}\left(I\right)$. This leads to consider the distance metric statistics in the terminology of Newey and McFadden (1994), or the M-statistic in the terminology of Rao and Zhao (1992),

$$M_{\text{Ind}} = \widehat{Q}\left(\widehat{\gamma}|\alpha\right) - \sum_{I\in\mathcal{I}} \widehat{Q}\left(\widehat{\gamma}|\alpha, I\right) L_I/L. \tag{1.4.22}$$

In the application, the critical values and p-values of tests based on (1.4.22) will be calculated using the random weighting bootstrap method proposed by Rao and Zhao (1992), Wang and Zhou (2004) and Zhao, Wu and Yang (2007), which is detailed in the Appendix A.2. Appendix A.3 describes a Monte Carlo experiment of the test of exclusion participation restriction and its corresponding rejection probabilities.

Choice of Specification for the Bidders' Private Values: Linear versus Exponential

As mentioned earlier, we are interested in choosing between a linear or exponential function for the quantile regression specification (1.2.4). While the linear specification is apparently more popular and simpler to estimate, the exponential specification delivers positive private values, which may not be the case of a linear one. Both models may be also misspecified for the data at hand. An additional difficulty is that the retained specification must be valid for several quantile levels.

This section thus proposes a test to investigate which of the two model specifications has a better measure of fit for a prescribed range of quantiles.

There is a extensive literature on testing hypotheses for model selection and goodness of fit. See e.g. Koenker and Machado (1999) for nested quantile regression models, White (1982) for conditional mean analysis, Zheng (1998), Horowitz and Spokoiny (2001) and Whang (2006) for parametric against nonparametrics alternatives.

Considering model selection and non-nested hypothesis testing, Vuong (1989) has proposed likelihood ratio tests using the Kullback-Leibler (1951) information criterion, which measures the distance between a given distribution and the true distribution function. The best model among a collection of competing models is defined to be the one that is closest to the true distribution. The tests are derived for cases in which the models are non-nested, overlapping, or nested and whether both, one, or neither is misspecified. The null hypothesis is the equivalence of the two specifications, i.e. the two specifications are at the same distance of the true distribution function, although dominance can also be considered.

Based on the insights of Vuong (1989), the test proposed in this section leads to compare the population objective function measures of the linear and the exponential models. Let the population objective functions under the previous two models be defined respectively as

$$Q_{\mathrm{L}}\left(\gamma|\alpha, I\right) = \mathbb{E}\left[\rho_{\Psi_{I(\alpha)}}\left(b_w - X\gamma\right)|I\right]$$

and

$$Q_{\mathrm{E}}\left(\gamma|\alpha, I\right) = \mathbb{E}\left[\rho_{\Psi_{I(\alpha)}}\left(b_w - \exp\left(X\gamma\right)\right)|I\right].$$

Define the pseudo-true private values quantile regression coefficients as $\gamma_j^*\left(\alpha|I\right) = \arg\min_{\gamma\in\Gamma} Q_j\left(\gamma|\alpha, I\right)$ and the infimum of the population objective functions as

$$Q_j\left(\gamma_j^*|\alpha, I\right) = Q_j\left(\gamma_j^*\left(\alpha|I\right)|\alpha, I\right) = \inf_{\gamma\in\Gamma} Q_j\left(\gamma|\alpha, I\right),$$

where $j =$ E,L.

The next set of hypotheses considers dominance of the exponential specification:

$$\begin{aligned} H_0 &: Q_{\mathrm{E}}\left(\gamma_{\mathrm{E}}^*|\alpha, I\right) - Q_{\mathrm{L}}\left(\gamma_{\mathrm{L}}^*|\alpha, I\right) \leq 0 \text{ for all } \alpha \in \mathcal{A} \text{ and } I \in \mathfrak{I} \\ H_1 &: \text{not } H_0, \end{aligned}$$

which in turn motivates the following test statistic:

$$\widehat{\text{Vuong}} = \sqrt{L_I} \sup_{\alpha \in \mathcal{A}} \left(\widehat{Q}_{\mathrm{E}}\left(\widehat{\gamma}_{\mathrm{E}}|\alpha, I\right) - \widehat{Q}_{\mathrm{L}}\left(\widehat{\gamma}_{\mathrm{L}}|\alpha, I\right) \right). \tag{1.4.23}$$

Note that exchanging $\widehat{Q}_{\mathrm{E}}\left(\widehat{\gamma}_{\mathrm{E}}|\alpha, I\right)$ and $\widehat{Q}_{\mathrm{L}}\left(\widehat{\gamma}_{\mathrm{L}}|\alpha, I\right)$ in (1.4.23) gives a test statistic for dominance of the linear specification.

In the application, the critical values and p-values of the test based on (1.4.23) will be computed by the pairwise bootstrap method in each original subsample L_I[12], i.e. samples of the (X_ℓ, b_{w_ℓ}) pairs are drawn with replacement from the L_I pairs $\{(X_\ell, b_{w_\ell}) : \ell = 1, \cdots, L_I\}$ of the original subsample, each with probability $1/L_I$. Appendix A.3 describes a Monte Carlo experiment of the specification test above and its rejection probabilities.

Constancy of the Slope Coefficients

In this section, we investigate the impact of changes in the auctioned good characteristics across the entire distribution of private values. Consider the private values conditional quantile specification given in (1.2.3), where the vector $\gamma_1(\alpha|I)$ groups all the slope coefficients, excluding so the intercept. The hypothesis of interest is that $\gamma_1(\alpha|I)$ does not depend upon α, in which case the model (1.2.3) with $h(t) = t$ is the regression model of Rezende (2008).

The null and alternative hypothesis considered in this test are therefore:

$$\begin{aligned} H_0 &: \gamma_1^*(\alpha|I) = \gamma_1^*(I) \text{ for all } \alpha \in \mathcal{A} \text{ and } I \in \mathfrak{I} \\ H_1 &: \text{not } H_0. \end{aligned}$$

[12] This form of bootstrap is also known as (X,Y)-pair bootstrap method and has been widely used in the quantile regression literature.

Define the CQR empirical objective function as

$$\widehat{Q}_{\text{CQR}}\left(\gamma_{\text{CQR}}|I\right) = \frac{1}{K}\sum_{k=1}^{K}\Big\{\frac{1}{L}\sum_{\ell=1}^{L}\rho_{\Psi_{I_\ell}(\alpha_k)}\left(b_{w_\ell} - h\left(\gamma_{0,k} + Z_\ell\gamma_1\right)\right)\Big\}.$$

Under the null hypothesis, the CQR population objective function equals the average of the individual objective functions over the set of prescribed quantiles $\mathcal{A} = \{\alpha_1, \cdots, \alpha_K\}$. The test statistic proposed here will be constructed as in the test of exclusion participation restriction by using the M-statistic defined as follows

$$M_{\text{CQR}} = L\left[\widehat{Q}_{\text{CQR}}\left(\widehat{\gamma}_{\text{CQR}}|I\right) - \frac{1}{K}\sum_{k=1}^{K}\widehat{Q}\left(\widehat{\gamma}|\alpha_k, I\right)\right], \tag{1.4.24}$$

where $\widehat{Q}_{\text{CQR}}\left(\widehat{\gamma}_{\text{CQR}}|I\right)$ is the optimized CQR objective function.

The critical values and p-values of tests based on (1.4.24) will be calculated using the random weighting bootstrap method as described in the Appendix A.2. A Monte Carlo experiment for the test of constancy of slope coefficients above and the corresponding rejection probabilities are given in the Appendix A.3.

1.5 Empirical Application

In this section, we illustrate empirically the methodology proposed in this first chapter using data from ascending timber auctions run by the USFS. Timber auctions data have been used in several empirical studies, see e.g. Baldwin, Marshall and Richard (1997), Haile (2001), Athey and Levin (2001), Athey, Levin and Seira (2011), Li and Zheng (2012), Aradillas-Lopez, Gandhi and Quint (2013), Li and Perrigne (2003) and others. Some other works have investigated risk-aversion on timber auctions, as e.g. Lu and Perrigne (2008), Athey and Levin (2001) and Campo, Guerre, Perrigne and Vuong (2011).

The Timber Auction Data

The data set used here is publicly available on the internet[13]. It aggregates ascending auctions from the states covering the western half of the US (regions 1-6 as labeled by the USFS) occurred in 1979. It contains 472 auctions involving a total of 1175 bids and a set of variables characterizing each timber tract including the estimated volume of the timber measured in thousand of board feet (or mbf) and its estimated appraisal value given in Dollar per unit of volume. The latter measures how much the tract worths at the present time, taking into consideration a combination of factors as volume of each specie, quality and marketable price. Only scaled sales are considered, in which bids are per unit of timber. Therefore, winning bids are given in Dollar per unit of volume.

The government conducts a cruise of the timber tract prior to the auction and publishes a report with the characteristics of the tract uncovered by the cruise. The reservation price is announced prior to the auction and equals the appraisal value of the tract[14]. The auction is conducted in two rounds: in the first round, bidders submit sealed bids[15] that must exceed the appraisal value of the tract to be qualified for the auction; in the second round, bidders compete in an ascending auction.

Table 1.1 gives some descriptive statistics about the dataset. The auctioned tract displays significant heterogeneity in quality and size. The mean of the variables are all above the median, indicating tailored marginal distributions. In particular, the marginal distribution of winning bids is highly positive skewed. In the dataset, the number of actual bidders I takes values 2 and 3, with $I = 2$ more than 50% of the cases. Table 1.2 provides the results of a median regression analysis of the winning bids on the timber appraisal value per mbf and the volume. Lu and Perrigne (2008) performed a mean regression analysis and concluded that the estimated appraisal value and the volume of the timber are the variables that better explain the winning

[13] The same dataset was used by Haile and Tamer (2003), Lu and Perrigne (2008) and Aradillas-Lopez et al. (2013), and it is available at the JAE Data Archieve website: http://qed.econ.queensu.ca/jae/2008-v23.7/lu-perrigne/

[14] It is well known that the screening effect of the appraisal value in timber auctions is almost negligible, being plausible to consider that the reservation price is nonbinding. See e.g. Campo, Guerre, Perrigne and Vuong (2011), Haile and Tamer (2003) and Aradillas-Lopez at al. (2013).

[15] Note that the bids in the first round are not proper bids, but a proposal to be qualified for bidding in the second round.

Table 1.1: Summary Statistics

	Winning Bids	Appraisal Value	Volume	Number of Bidders
Mean	129.17	72.34	3,171.10	2.49
Std. Deviation	119.76	52.92	4,418.20	0.50
25%	59.13	24.47	294	2
50%	113.81	68.38	967	2
75%	172.53	111.25	4,724	3
Min	0.30	0.25	12	2
Max	1,981.50	219.58	24,800	3
Skewness	8.16	0.44	1.99	0.04
Observations	472	472	472	472

bid's variability. Both variables are also highly significant in the median analysis, which can be seen by the coefficient of determination R_α at the median quantile[16]. For this reason, we shall consider a two-dimensional vector of covariates Z grouping both the appraisal value per mbf and the volume of the timber[17].

Table 1.2: Winning Bids

Variable	Coefficient	Std. Deviation	t-value	p-value
Constant	11.1265^{***}	3.8554	2.886	4.08E-03
Appraisal Value	1.2873^{***}	0.0505	25.514	0.000
Volume	0.0029^{***}	0.0008	3.503	0.0005
$R_{0.5}$	0.4159			

Significance level: * 10%, ** 5% and *** 1%.

Results

The choice of the set of prescribed quantiles $\mathcal{A} = \{\alpha_1, \cdots, \alpha_K\}$ used in the tests is an important issue. It would be ideal to estimate all the quantiles of the private values

[16]The coefficient of determination is given by

$$R_\alpha = 1 - \frac{\widehat{Q}\left(\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2}|\alpha\right)}{\widehat{Q}\left(\widehat{\beta_0}|\alpha\right)},$$

and represents how much of the variability in the winning bids is explained by both covariates above at the quantile level α.

[17]We also investigated nonlinearity with respect to the appraisal value and the volume of the timber. However, the results were not statistically significant.

conditional distribution. However, as noticed by Koenker (2005) and also seen from our estimation results, the asymptotic precision of the quantile estimates in general, and the quantile regression estimates in particular, depend on the quantile density function evaluated at the quantile of interest. If the data are sparse at the quantile of interest, then the quantile regression estimates will be less precise. I observed this lack of precision when estimating the quantiles closer to the boundary. For this reason, I restrict to $\mathcal{A} = \{0.12, 0.14, \cdots, 0.80\}$.

The first sealed bid auction stage used to qualify bidders for the ascending auction may select a certain number of bidders for the ascending auction stage, so that it might be interesting to use $V(\cdot|X, I)$, which depends upon I, in the policy analysis to be conducted. Roberts and Sweeting (2012) found evidences for a selective entry when studying timber auctions in California. They consider in the model a significant entry cost affecting bidders participation because bidders are allowed to conduct their own cruise. They argue that conducting a private cruise responds for a large fraction of the bidder's entry cost. Because bidders in our application do not conduct their own cruise, the entry cost might be small compared to the Californian timber auctions. There are nevertheless other costs that may affect bidders participation, such as developing a market studying, preparing the bids and attending to the auction. If indeed the entry costs are not relevant for the bidders' decision in participating, then the optimal reservation price policy could be chosen independently of I, as shown in Proposition 1.1-(ii).

Table 1.3: Test of the Exclusion Participation Restriction

Null Hypothesis	Specification	M-Statistic	p-value[a,b]
$\gamma^*(\alpha\|I) = \gamma^*(\alpha)$ for all $\alpha \in \mathcal{A}$ and $I \in \mathcal{I}$	Linear	653.83	0.3096
	Exponential	720.37	0.4136

[a] The p-value and critical values are computed using the random weighting bootstrap method. More details about the test are given in section 1.4;
[b] The number of bootstrapping replications is 5,000; $\mathcal{A} = \{0.12, 0.14, \cdots, 0.80\}$ and $\mathcal{I} = \{2, 3\}$.

Testing the relevance of a private value quantile function conditional upon bidder's participation can thus be very useful in the analysis that follows. Table 1.3 gives the results of the test of exclusion participation restriction suggested in section 1.4. There is not enough statistical evidences to reject the null hypothesis of independence in both specification models (linear and exponential), implying that a possible improvement in efficiency can be achieved by pooling both samples. In other words, it is possible to rule out unobserved heterogeneity and entry stage affecting bidders participation. Note that an entry model as Levin and Smith (1994), where bidders do not know their private values at the entry stage, would generate private values independent upon I as observed in Table 1.3. However, the bidders' participation decision in Levin and Smith (1994) is random, which may be considered a bit unrealistic.

Table 1.4 gives the results of the choice of specification test described in section 1.4 that investigates which of both regression specifications, linear or exponential, better represents the private value conditional quantiles. In the first row of Table 1.4, the test considers a null hypothesis in favour of the exponential specification, whereas in the second row the null hypothesis is reversed in favour of the linear specification. Both tests conclude that the linear specification dominates the exponential one over the set of prescribed quantiles $\mathcal{A}$. Therefore the linear specification is used in the rest of this section.

Table 1.4: Choice of Specification

Null Hypothesis	Test Statistic	p-value[a,b]
$Q_{\mathrm{E}}(\gamma^*_{\mathrm{E}}\vert\alpha) - Q_{\mathrm{L}}(\gamma^*_{\mathrm{L}}\vert\alpha) \leq 0$ for all $\alpha \in \mathcal{A}$	20,142.55	0.0000
$Q_{\mathrm{L}}(\gamma^*_{\mathrm{L}}\vert\alpha) - Q_{\mathrm{E}}(\gamma^*_{\mathrm{E}}\vert\alpha) \leq 0$ for all $\alpha \in \mathcal{A}$	-3,103.58	0.9938

[a] The p-value and critical values are computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I. Note however that the result of independence given by the test of exclusion participation restriction is considered, so that both samples with $I = \{2, 3\}$ are pooled in this analysis. More details about the test are given in section 1.4;

[b] The number of bootstrapping replications is 5,000; $\mathcal{A} = \{0.12, 0.14, \cdots, 0.80\}$.

Table 1.5 provides the result of testing constancy of the slope coefficients. This test investigates whether changes in the auctioned characteristic Z affect the private values quantiles similarly across the entire distribution. It may also show how heterogeneous[18] are the bidders across the population. The test gives strong statistical evidence to reject the null hypothesis that the slope coefficients are constant across the quantiles $\alpha \in \mathcal{A}$. Therefore, bidders react differently to changes in the quality and size of the timber tract, as also clearly illustrated by Table 1.6

[18]Recall that heterogeneity is the term used in this chapter to represent the different reaction of the bidders to changes in the auctioned characteristics according to their rank in the private values conditional distribution.

and Figures 1.1 and 1.2 below. Given the test result, the pooled quantile regression estimator provides a better characterization of the private value conditional quantiles than the CQR estimator defined in (1.3.21) and the mean analysis suggested by Rezende (2008).

Table 1.5: Constancy of the Slope Coefficients

Null Hypothesis	M-Statistic	p-value[a,b]
$\gamma_1^*(\alpha) = \gamma_1^*$ for all $\alpha \in \mathcal{A}$	124.51	0.0048

[a] The p-value and critical values are computed using the random weighting bootstrap method. Note however that the results obtained in the previous two tests were considered for this analysis. More details about the test are given in section 1.4;
[b] The number of bootstrapping replications is 5,000; $\mathcal{A} = \{0.12, 0.14, \cdots, 0.80\}$.

The next Table and two following Figures describe the private values quantile regression coefficients. The most important variable is the appraisal value, a quality measure released by the seller, which is often interpreted as the seller's private value, see Lu and Perrigne (2008) and Aradillas-Lopez et al. (2013). The associated quantile regression coefficient is given in the second column of Table 1.6 and Figure 1.1 for a median auction, where the volume is 967 thousand of board feet and the appraisal value is about \$68 per thousand of board feet. Note that the coefficient is always significant and larger than 1, suggesting that it acts as a markup indicating how much more the auctioned good appraisal value is valued by the bidders than the seller. The private values can be also interpreted as a measure of how much the bidders would be willing to sell goods made with the timber bought at the auction[19]. This suggests that the higher the bidder's private value, the higher is his efficiency in aggregating value to the timber. The coefficients increase over the quantile levels, suggesting a relative increase in the markup of 75% when comparing bidders in the quantiles $\alpha = 0.10$ and $\alpha = 0.80$ of the private value conditional distribution. This is also evidence that bidders belonging to the upper tail of the private value distribution are more highly affected by changes in the appraisal value than median bidders.

[19] In this interpretation, it is necessary to assume that timber is the most important component of the goods produced by the bidders. This could be however modified to cover other cases where timber would only be a part of these goods.

Table 1.6: Private Value Quantile Regression Estimates

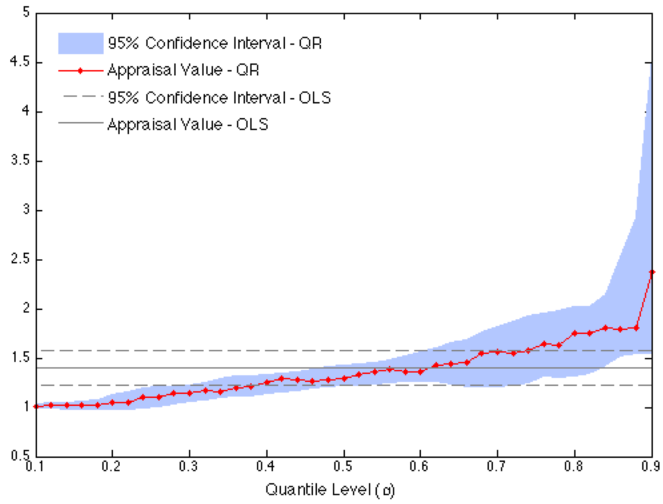
Quantile Level	Intercept	Appraisal Value	Volume
0.1	0.95	1.01	0.0007
	[-0.97,2.28]	[0.99,1.04]	[0.0004,0.0016]
0.2	3.00	1.04	0.0016
	[-0.72,8.49]	[0.99,1.13]	[0.0005,0.0027]
0.3	9.39	1.15	0.0018
	[2.05,15.01]	[1.05,1.22]	[0.0010,0.0033]
0.4	11.77	1.25	0.0034
	[5.32,20.83]	[1.14,1.33]	[0.0013,0.0049]
0.5	21.03	1.29	0.0041
	[10.92,29.24]	[1.22,1.43]	[0.0023,0.0054]
0.6	35.68	1.36	0.0041
	[21.63,45.02]	[1.27,1.56]	[0.0029,0.0055]
0.7	44.28	1.57	0.0045
	[29.94,77.15]	[1.22,1.81]	[0.0029,0.0071]
0.8	67.64	1.75	0.0060
	[32.91,101.02]	[1.31,2.02]	[0.0024,0.0138]
0.9*	72.98	2.37	0.0167
	[12.89,124.22]	[1.54,4.56]	[0.0031,0.0384]

The estimates are for a median auction and were computed using the pooled quantile regression estimator defined in (1.3.18). The 95% confidence interval of the quantile regression estimates in square brackets were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I;

* Note the loss in precision when α gets closer to the upper boundary. This is why such higher quantile have been excluded from the test statistics.

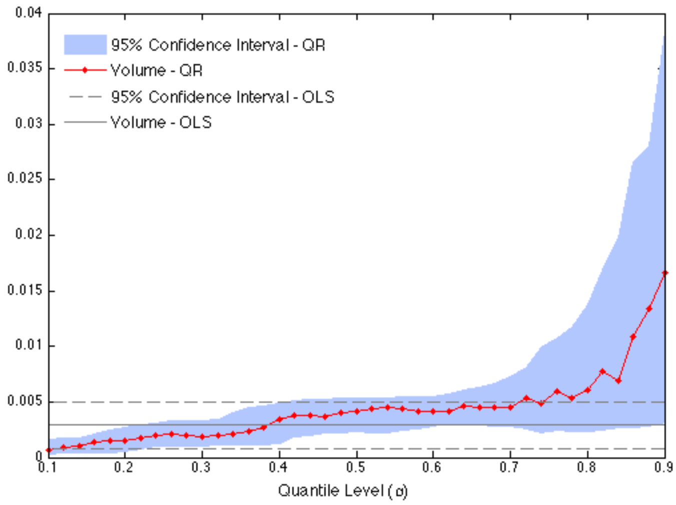
Figure 1.1 shows the quantile regression and the OLS estimates of the appraisal value slopes with their corresponding 95% confidence intervals. Observe how the estimated markup increases over the quantile levels. Figure 1.2 shows the quantile regression and OLS estimates associated with the variable volume and their 95% confidence intervals. Although the coefficients seem to increase with the quantile level in Table 1.6, they lie inside the estimated OLS 95% confidence intervals, suggesting that the volume coefficient may not depend upon α. In the upper tail of the distribution, the coefficients are larger, but the confidence intervals are also wider.

Figure 1.1: Appraisal Value Slope Coefficients



Notes: The 95% confidence intervals for the OLS estimate consider the heteroscedasticity-robust (White) standard errors. The ones for the quantile regression estimates were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

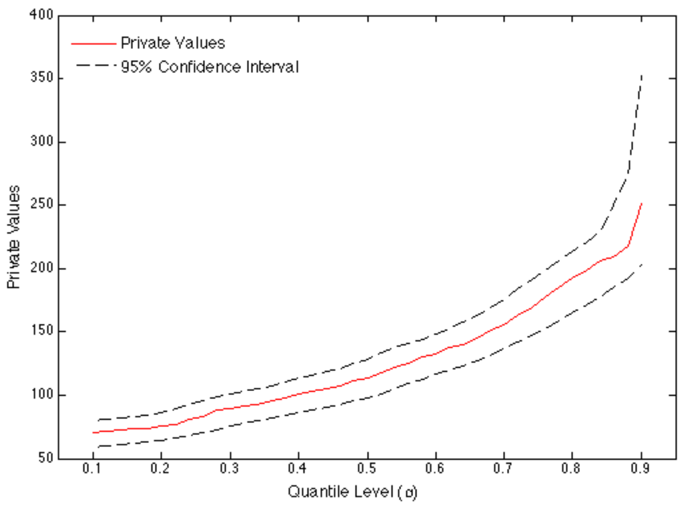
Figure 1.2: Volume Slope Coefficients



Notes: The 95% confidence intervals for the OLS estimate consider the heteroscedasticity-robust (White) standard errors. The ones for the quantile regression estimates were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

In what follows $X(\tau) = (X_1(\tau), X_2(\tau), X_3(\tau))$, where $X_j(\tau)$ is the quantile of order τ of the variable X_j, $j = 1, 2, 3$. With some abuse of terminology, $X(\tau)$ will be called the quantile of order τ of the matrix of auction covariates X, $X(0.50)$ being called its median. Figure 1.3 gives the private value conditional quantile estimates for a median auction and their 95% confidence intervals. Figure 1.4 presents the quantile estimates for several quantile levels of $X(\tau)$, where $\tau = \{0.15, 0.25, 0.50, 0.75, 0.85\}$. In particular, it shows the change in the shape of the private value quantiles due to variations in the quality and size of the timber tract.

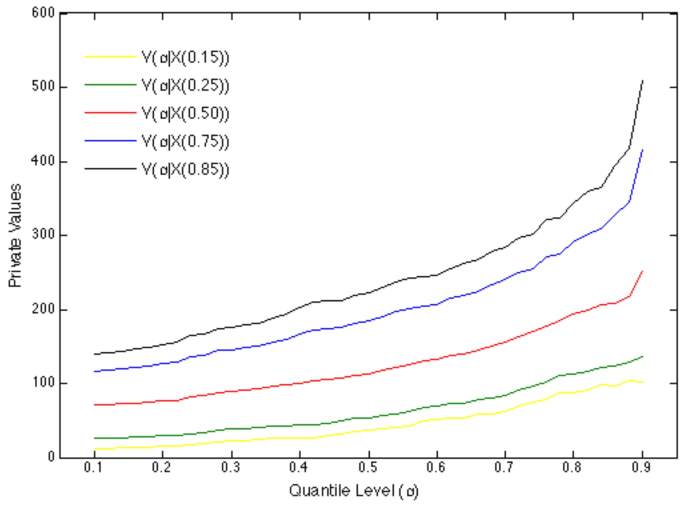
Figure 1.3: Private Values Conditional Quantiles



Notes: The estimates are for a median auction. The 95% confidence intervals were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

Figure 1.4 indeed shows that the auction covariates change significantly the shape of the private value distribution. This effect becomes even clear when comparing a high and a low quantile of the private value conditional distribution. Consider in particular the quantiles $\alpha = 0.12$ and $\alpha = 0.80$ of the private value conditional quantile curves and the quantiles of order $\tau = 0.15, 0.50$ and 0.85 of the auction covariates X, that is, auctions with low, median and high quality and size. The relative increase in the private value is of about 600% in the auctions with low quality and size, whereas it reduces to 172% and 142% in the median and high

Figure 1.4: Private Values Conditional Quantiles and $X(\tau)$

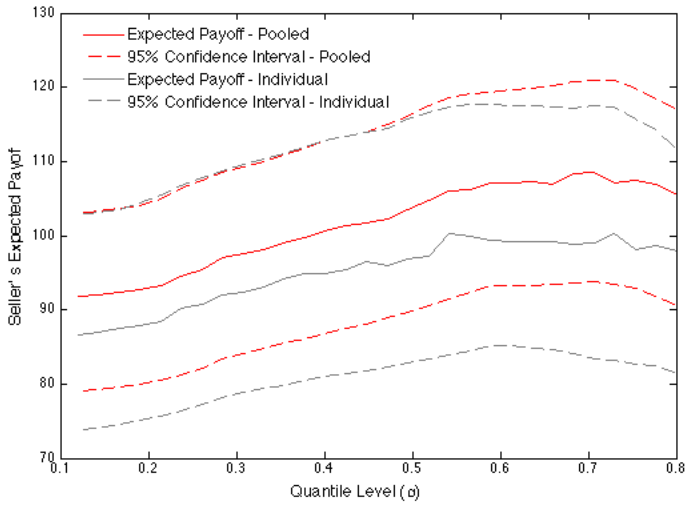


quality and size auctions, respectively. It seems that less productive bidders, i.e. the ones with a low rank α, choose not to participate in auctions with higher quality and size[20].

We now turn to the estimation of the seller's expected payoff and the associated optimal reservation price. The expected payoff of a seller with private values v_0, $\Pi(\alpha_r|X, v_0)$, is estimated using numerical integration via a trapezoidal rule to approximate the definite integral given in Proposition 1.1-(i). Figures 1.5 and 1.6 give the expected payoff of the seller for a median auction with the 95% confidence intervals for $I = 2$ and $I = 3$ bidders, respectively. In both Figures, the estimation procedures given by the pooled and the individual estimators defined respectively in (1.3.18) and (1.3.15) are compared. Observe that the former gives smaller confidence intervals than the latter due to its higher estimation efficiency. Note that the only difference between the seller's expected payoffs estimated via (1.3.18) and (1.3.15) is given by the way the private value conditional quantiles are estimated. In the former, both subsamples with different number of bidders are pooled, whereas in the latter the estimation is done for each I.

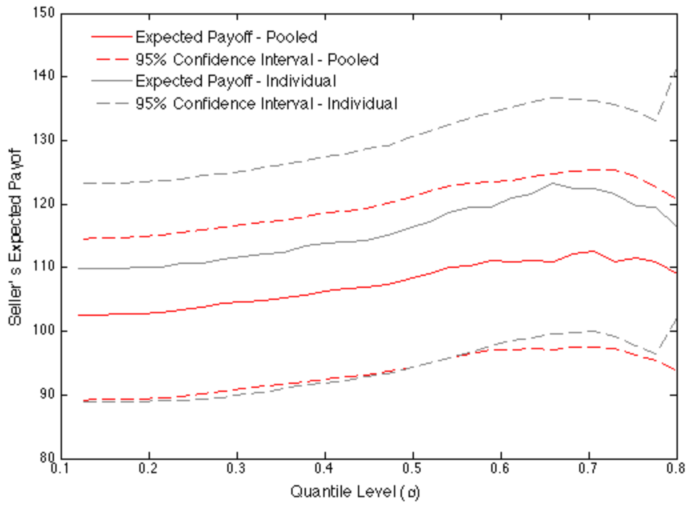
[20]Note that this suggests that the rank of the participant and the number of participants are simultaneously determined, so that quantile regression estimation can be affected by endogeneity. Addressing this issue is outside the scope of this present chapter.

Figure 1.5: Pooled Vs Individual $I = 2$



Notes: The estimates are for a median auction. The seller's private value considered here is $v_0 = AV$, i.e. $v_0 = \$68$ per unit of timber. The 95% confidence intervals were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

Figure 1.6: Pooled Vs Individual $I = 3$



Notes: The estimates are for a median auction. The seller's private value considered here is $v_0 = AV$, i.e. $v_0 = \$68$ per unit of timber. The 95% confidence intervals were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

The sensitivity of the seller's expected payoff and the associated optimal reservation price to choices of the seller's private value v_0 will be now investigated. The choice of the latter seems to be indeed important to determine the optimal screening level policy since it represents the possible gains that the seller may have when selling the good in the outside market. Note that $v_0 = 0$ may represent the case in which the seller has no opportunity to sell the good outside the auction. The most common choice for v_0 is the appraisal value of the timber[21]. The results obtained for $v_0 =$ Appraisal Value (AV) are compared with the case with no outside option $v_0 = 0$. It is also investigated the effect on the seller's optimal behaviour when the quality and size of the auction change.

Table 1.7 gives the optimal screening level α_r^*, the corresponding optimal reservation price $r^* = V(\alpha_r^*|X)$ and the seller optimal expected payoff $\Pi(\alpha_r^*|X, v_0)$ for the auctions with quality and size specified by $\tau = 0.15, 0.50$ and 0.85 and considering both choices of v_0. Recall that the optimal screening level α_r^* is chosen as the maximizer of the seller's expected payoff over $\mathcal{A}$. It is expected that α_r^* that maximize the aggregate expected payoff, as defined in (1.2.10), is the same as the one maximizing the expected payoff computed for each I, i.e. $\alpha_r^* = \alpha_r^*(I)$ for all I, given the independence result obtained in the test of exclusion participation restriction. The estimation indeed confirms that α_r^* is independent upon the number of actual bidders participating in the auction. This will be then reflected in the computation of the corresponding optimal reservation price, which will also not depend upon I. Note however, that α_r^* still depend upon (X, v_0).

A general conclusion from Table 1.7, which is also expected, is that auctions with higher quality and size provide larger expected payoffs for the seller. The optimal screening level in turn reduces when the quality and size of the auction increase. A possible reason for that is the low heterogeneity among bidders observed in better auctions. Recall from Figure 1.4 that auctions with low quality and size show a significant increase in the markup over the timber value. Therefore, as seen from

[21] As mentioned in Aradillas-Lopez et al. (2013), the seller's private value may be even lower than the appraisal value of the timber if exercising an outside option (through, for example, a lump-sum contract) entails additional cost to the seller. It is also possible that v_0 is nevertheless higher than the appraisal value since scaled sales require the timber service to measure the timber actually harvested to calculate the payment. Therefore, by exercising the outside option the seller would avoid those costs.

Table 1.7: Optimal Reservation Price

v_0	0	AV
$\alpha_r^* (v_0, X(0.15))$	0.75	0.75
	[0.52,0.80][a]	[0.56,0.80]
$V(\alpha_r^* \| X(0.15))$	83.08	83.08
	[38.93,104.58]	[45.34,109.03]
$\Pi(\alpha_r^* \| X(0.15), I=2, v_0)$	32.64	39.05
	[26.16,44.78]	[30.26,50.33]
$\Pi(\alpha_r^* \| X(0.15), I=3, v_0)$	38.98	43.80
	[30.62,52.92]	[33.32,56.73]
$\alpha_r^* (v_0, X(0.50))$	0.28	0.71
	[0.12,0.36]	[0.56,0.78]
$V(\alpha_r^* \| X(0.50))$	87.93	159.89
	[67.73,101.96]	[123.40,192.43]
$\Pi(\alpha_r^* \| X(0.50), I=2, v_0)$	91.71	108.49
	[79.17,102.32]	[95.11,122.13]
$\Pi(\alpha_r^* \| X(0.50), I=3, v_0)$	102.93	112.50
	[99,126.57]	[89.63,114.26]
$\alpha_r^* (v_0, X(0.85))$	0.24	0.56
	[0.12,0.42]	[0.47,0.80]
$V(\alpha_r^* \| X(0.85))$	163.65	243.97
	[135.51,213.63]	[220.03,376.25]
$\Pi(\alpha_r^* \| X(0.85), I=2, v_0)$	177.66	203.26
	[163.40,192.73]	[187.71,223.19]
$\Pi(\alpha_r^* \| X(0.85), I=3, v_0)$	196.37	208.61
	[179.89,212.72]	[191.63,231.26]

[a] The 95% confidence intervals in square brackets were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

Proposition 1.2.8, the seller has a stronger incentive to use screening for low quality and size auctioned goods. For these low quality and size auctions, the seller private value does not seem to change reservation price recommendations, which, as the expected payoff, remains constant when v_0 grows from 0 to the appraisal value. This may also be related with the large heterogeneity among the bidders, which resulted in a strong increase in the private value conditional quantiles. Observe as well that the increase in α_r^* when v_0 grows from 0 to the appraisal value is smaller when comparing median auctions with auctions with higher quality and size.

As mentioned in Section 1.2, the probability of trading in the auction with a screening level α_r is $1 - \alpha_r^I$. Table 1.8 groups the probabilities of trading in each of the three kinds of auctions for the two choices of v_0. Note that in the auctions with low quality and size, the probability of selling the good is very low (44% and 58% for $I = 2$ and $I = 3$, respectively). This is because bidders are very heterogeneous and the seller should set a high screening level to avoid low bidders from participating. This somehow carries over for median and higher quality and size auctions when the seller's private value is the appraisal value. Policy recommendations with such low probability of selling may not make sense in practice, especially for goods with a potential high storage cost. Note however that auctions with high quality and size and $I = 3$ seem not to be so much affected by this issue.

Table 1.8: Probability of Trading

		$v_0 = 0$	$v_0 = \text{AV}$
$X\,(0.15)$	$I = 2$	44%	44%
	$I = 3$	58%	58%
$X\,(0.50)$	$I = 2$	92%	50%
	$I = 3$	98%	65%
$X\,(0.85)$	$I = 2$	94%	69%
	$I = 3$	99%	82%

By reducing the seller's private value, the probability of trading increase, but such a consideration is mostly theoretical since it is not possible in practice to change the seller's private value. It nevertheless shows that the seller has a high incentive to decrease the optimal screening level when he faces the case in which there is no trade outside the auction. As can be seen, the practical implementation of the auction theory can be sometimes difficult in the sense that usual choices for the seller's private value may lead to recommendation of mechanisms with very low probability of trading. This may question the relevance of considering expected payoff in the maximization process.

I now show a comparison of the quantile regression estimation approach proposed in this Chapter with a nonparametric alternative. The latter is widely used in the empirical auction literature due to its flexibility in representing the data at hand because no structure is imposed in the analysis. Comparing a nonparametric revenue analysis with a quantile regression one is therefore a way to assess whether the quantile regression is correct from a policy recommendation point of view. The nonparametric estimation is conducted as detailed in the end of section 1.2 and considers a median auction as representative. It is taken into consideration that $F_v(\cdot|X) = F_v(\cdot|X, I)$ due to the exclusion participation restriction test, which implies that $\widehat{F}_v(\cdot|X(0.50), I)$ must be aggregate over $I = 2, 3$ as in (1.2.14). Table 1.9 gives the estimates of the optimal reservation price and optimal expected payoff using both nonparametric and quantile regression approaches. Note that the nonparametric estimation results are less precise than the ones derived from a quantile regression specification. The confidence intervals of the nonparametric estimates are on average 2.58 times larger than the ones obtained in the quantile regression analysis. The two estimation strategies seem to give similar results and there is no reason to reject a quantile regression specification on the ground of policy recommendation.

Table 1.9: Nonparametric vs. Quantile Regression

Optimal Reservation Price		Optimal Expected Payoff[a]			
		$I = 2$		$I = 3$	
Nonparametric	Quantile Regression	Nonparametric	Quantile Regression	Nonparametric	Quantile Regression
189.28	160.28	107.71	108.45	122.86	112.46
[139.47, 298.16][b]	[124.34,194.59]	[87.75,160.34]	[95.45,122.90]	[98.70,178.72]	[99.31,127.34]

[a] For simplicity, it is assumed that $v_0 = AV$.
[b] The 95% confidence intervals in square brackets were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

1.6 Conclusion

In this Chapter I proposed an identification and estimation approach based on quantile regression to recover the bidders' private values conditional distribution. The quantile regression framework provides a flexible and convenient parametrization of the private value distribution, with an estimation methodology easy to implement and with various specification tests that can be derived. The latter includes tests of private values and actual number of bidders dependence, linear versus exponential specifications and constant slope coefficients across quantile levels.

The Chapter shows that a focus on the quantile level of the private values distribution and on the seller's optimal screening level can be both useful for policy recommendations. The former helps bidders to benchmark their bids in order to achieve a desired probability of winning the auction, whereas the latter provides a better understanding of how appropriate is the recommended policy from a seller's point of view. Both focuses above are new in the empirical auction literature.

The empirical application using timber auctions from the USFS shows that policy recommendations should be carefully examined before practical implementation. The screening level associated with the optimal reservation price is usually high, resulting in a low probability of trading. The analysis of the shape of the private value conditional quantile curves suggests that such inappropriate recommendations are due to a sharp increase in the private value conditional quantiles, which may be evidence of large heterogeneity among the bidders. As a consequence, the seller has a strong incentive to screen the bidders by using a high reservation price, leading then to a low probability of selling the auctioned good.

The private values estimated quantile shapes can be genuine but can also be the consequence of a model misspecification. Some other works have also noticed such a high level of the optimal reservation price in timber auctions. Aradillas-Lopez et al. (2013) suggest that neglecting private values affiliation can generate high reservation prices. However, their nonparametric methodology may be affected by the curse of dimensionality. In addition, as noted in Roberts and Sweeting (2012), timber auctions include a preliminary selection that can affect the estimated shape of the private value quantile functions. The strong heterogeneity revealed by the estimation of the private value conditional quantile function can also be an indication of asymmetry. As discussed in Cantillon (2008) and Gavious and Minchunk (2012), sellers facing asymmetry have an incentive to increase competition by increasing reservation prices.

However, analyzing revenue with a risk neutral seller perspective may be not appropriate to address issues such as high reservation prices and low probability of selling the auctioned object. The results given in Hu, Matthews and Zou (2010) regarding risk aversion affecting sellers can be useful to provide more relevant reservation price recommendations. In the next Chapter, I propose a numerical investigation of the variation in the optimal screening level when the seller has a constant relative risk aversion utility function. I then conclude that considering risk averse sellers is indeed sufficient to achieve reasonable policy recomendations.

Chapter 2

To Sell or not to Sell? An Empirical Analysis of the Optimal Reservation Price in Ascending Timber Auctions.

2. To Sell or not to Sell? An Empirical Analysis of the Optimal Reservation Price in Ascending Timber Auctions.

Abstract

This Chapter proposes a numerical investigation of the variation in the optimal reservation price and the associated screening level when the seller risk aversion increases. As noted in the empirical application of the previous Chapter, the optimal screening level when the seller is risk neutral usually leads to a low probability of selling the auctioned timber lot. The numerical experiments of this Chapter then shows that using a higher constant relative risk aversion (CRRA) utility when computing the optimal reservation price increases the probability of trading in comparison with the assumption of risk neutrality.

2.1 Introduction

It is usual in the empirical auction literature to assume that sellers are risk neutral about selling the good (see e.g. Haile and Tamer (2003), Li, Perrigne and Vuong (2003), Li and Perrigne (2003), Lu and Perrigne (2008) and Aradillas-Lopez, Gandhi and Quint (2013)). It is however possible that the costs of holding the auctioned good for a long period makes sellers risk averse of the possibility of not selling at the auction in comparison with an outside option in which the trade is certain but probably at a lower price. As Bulow and Klemperer (2009) say, sellers prefer auctions because it involves bidders competing simultaneously, which may then increase the price at which the object is finally sold[22]. If sellers are risk averse of not trading in the auction, it is thus more likely that an optimal reservation price recommendation that yields a higher probability of selling the good is more appropriate from a seller's point of view.

Chapter 1 in this thesis analysed ascending timber auctions via a quantile regression approach under the Independent Private Value (IPV) framework and a risk neutral seller. The empirical application conducted suggests that the optimal reservation price policy, or equivalently, the optimal screening level was most of the time high, yielding a low probability of selling the good. For instance, in some auctions the optimal reservation price under risk neutrality was about \$83 per unit of volume, which corresponds to the 0.75-quantile of the private values distribution, i.e. 75% of the population of bidders would be out of the game because of a low valuation. Such a recommendation policy yields then a probability of selling the good of about 44% and 58% for 2 and 3 bidders competing, respectively. In the case of timber auctions, the Government sells via an auction the right to clear-cut a timber tract. Not selling may penalize the business activities of the bidders and may have important local economic and political consequences. Not cutting the wood may also have ecological consequences and delaying the trading may not be possible when the forest administration is not in a position to manage forests by itself. Finally, the financial resource may be helpful to ensure a more efficient forest administration. All these consequences would then motivate a risk averse behaviour

[22] Auctions may be inefficient when entry is costly, bidders are asymmetric and/or there might exists a selective process.

by the Government that was not considered in the previous analysis.

The goal of this Chapter is to investigate whether such a high screening level policy can be addressed when a risk aversion parameter affecting the Government's utility in the auction is introduced in the analysis. The econometric motivation behind a concave utility function is that it may smooth out the strong variation observed in the private values quantiles when computing the seller's payoff. If the seller risk aversion decreases with the monetary quantity, he would have a strong incentive to deviate of the situation of no trade. Hu, Matthews and Zou (2010) show that the optimal screening level policy indeed decreases as the sellers becomes more risk averse. The empirical study proposed in this Chapter complements the results in Hu, Matthews and Zou (2010) by showing that risk aversion can drastically increases the probability of selling the good in the case of US forest ascending auctions. The dataset considered here was also used in Haile and Tamer (2003), Lu and Perrigne (2008) and Aradillas-Lopez et al. (2013). A parametric constant relative risk aversion (CRRA) family of utility functions will be used to represent the seller's utility, since it satisfies the seller's deviation incentive mentioned above.

This Chapter is organized as follows. Section 2.2 presents the auction mechanism and a quantile version of the seller's expected payoff. Section 2.3 shows the quantile regression identification and estimation methodology, the assumptions of the model and how to estimate the optimal screening level. Section 2.4 gives the results of an empirical study of the optimal screening level under a risk averse seller. Finally, section 2.5 concludes this chapter.

2.2 Seller's Expected Payoff in a Quantile Setup

A single and indivisible object with some characteristics $Z \in \mathbb{R}^P$ is auctioned to $I \geq 2$ bidders through an ascending auction. The seller announces a reservation price r prior to the auction that is the minimum price that he would be willing to accept. Both the set of auction covariates $X = (1, Z)$ and the number of actual bidders I participating in the auction are common knowledge. The object is sold to the highest bidder for the price of his last bid, provided that it is at least as high as the reservation price r. Within the IPV paradigm, each bidder $i = 1, \cdots, I$ is assumed to have a private value v_i for the auctioned good, which is not observed by other bidders. The bidder only knows his own private value, but it is common knowledge for bidders and sellers that private values have been identically and independently drawn from a common c.d.f. $F_v(\cdot|X)$ conditional upon the auction characteristics[23], or equivalently, with a conditional quantile function $V(\alpha|X)$, $\alpha \in [0,1]$, defined as

$$V(\alpha|X) := \inf\{v : F_v(v|X) \geq \alpha\} = F_v^{-1}(\alpha|X).$$

The rank u_i of a bidder with private value v_i satisfies $u_i = F_v(v_i|X)$, being independent of X and with a uniform distribution over $[0,1]$. The IPV paradigm implies that the ranks u_i, $i = 1, \cdots, I$, are independent. In other words, the dependence between the private values v_i and the auction covariates X can be fully captured by the nonseparable model

$$v_i = V(u_i|X), \quad u_i \overset{iid}{\sim} \mathcal{U}_{[0,1]} \perp X. \tag{2.2.1}$$

Therefore, bidders are identical up to the variable u_i, which represents the bidder ith's position in the private value distribution. Consider the next two assumptions concerning the equilibrium bidding in the ascending auction described above and the private value conditional quantile representation (2.2.1).

Assumption 2.1. *The transaction price in an auction is the greater of the reservation price and the second-highest bidder's willingness to pay.*

Assumption 2.2. $V(\alpha|X)$ *is strictly increasing and continuous on its support* $[V(0|X), V(1|X)]$ *for all* X.

Assumption 2.1 is from Aradillas-Lopez, Gandhi and Quint (2013) and, as noted in Athey and Haile (2002) and Bikhchandani, Haile and Riley (2002), is compatible with the multiple equilibra generated by the ascending auctions. It is for instance the result of the dominant strategy equilibrium of a button[24] auction, which is a stylized version of an ascending auction. This assumption would also hold approximately in

[23] It is assumed that the private values distribution is independent upon the number of actual bidders I competing in the auction, i.e. $F_v(\cdot|X) = F_v(\cdot|X, I)$. This independence assumption is only consistent with a model with no selective entry. See Chapter 1 for a related discussion.

[24] In a button auction, bidders hold down a button to remain active while the price rises continuously, releasing the button to drop out of the game, and the willingness to pay of the losing bidders is learned from their drop out prices.

the context of Haile and Tamer (2003) if bidders do not use jump bids at the end of the auction. Assumption 2.2 is usual in the quantile regression literature.

Consider a binding reservation price $r(X) \in [V(0|X), V(1|X)]$. The latter induces screening in the auction since bidders with $V(\alpha_i|X) < r(X)$ are prevented from participating in the game. Let α_r be the screening level in the private values conditional distribution, such that $V(\alpha_r|X) = r(X)$. It represents the percentage of bidders in the population not participating in the auction because of a low valuation. This gives a probability of not selling the good of $1 - \alpha_r^I$, which implies that, for a given I, the probability of trading decreases with the level of screening α_r. Let $U(\cdot)$ be the seller's Von Neumann-Morgenstern (vNM) utility function, satisfying $U(0) = 0$, $U^{(1)}(\cdot) > 0$ and $U^{(2)}(\cdot) < 0$, where the superscripts (1) and (2) represent the first and second derivatives of $U(\cdot)$.

Let $\pi(r(X))$ denotes the seller's payoff

$$\pi(r(X)) = U(b_w)\,\mathbb{I}(b_w \geq r(X)) + U(v_0(X))\,(1 - \mathbb{I}(b_w \geq r(X))),$$

where $v_0(X) \in [V(0|X), V(1|X)]$ is the seller's private value for the good that will be sold. In what follows, let $v_0 = v_0(X)$ and $r = r(X)$. The following Proposition gives a quantile expression for the seller's expected payoff, a candidate for the optimal screening level $\alpha_r^* = \alpha_r^*(X, v_0)$ and the corresponding optimal reservation price $V(\alpha_r^*|X)$. Let $\Pi(\alpha_r|X, I, v_0)$ be the seller's expected payoff given (X, I, v_0) when the screening level is α_r.

Proposition 2.1. *Suppose the seller has a vNM utility function $U(\cdot)$. Under IPV and assumptions 2.1-2.2,*

(i) the seller's expected payoff is

$$\begin{aligned}\Pi(\alpha_r|X, I, v_0) = U(v_0)\,\alpha_r^I + U(V(\alpha_r|X))\,I\alpha_r^{I-1}(1-\alpha_r)\\ + I(I-1)\int_{\alpha_r}^{1} U(V(\alpha|X))\,\alpha^{I-2}(1-\alpha)\,d\alpha,\end{aligned} \tag{2.2.2}$$

where $v_0 = V(\alpha_0|X)$ is the seller's private value;

(ii) The optimal reservation price $V(\alpha_r^*|X)$ *satisfies*

$$U(V(\alpha_r^*|X)) - U^{(1)}(V(\alpha_r^*|X))\, V^{(1)}(\alpha_r^*|X)(1-\alpha_r^*) = U(v_0), \qquad (2.2.3)$$

where $V^{(1)}(\alpha_r^*|X)$ *is the private value quantile density*[25] *function.*

The proof of Proposition 2.1 is the same as the one in Chapter 1 for Proposition 1.1 since, compared to the risk neutrality case, introducing the utility function only amounts to transform the private value quantile function $V(\alpha|X)$ into $U(V(\alpha|X))$. Equation (2.2.2) in Proposition 2.1 gives the seller's expected payoff in a quantile setup. Equation (2.2.3) is the first-order condition (FOC) associated with the maximization of the seller's expected payoff. The first implication of (2.2.3) is that $\alpha_r^* > \alpha_0$, or equivalently, $V(\alpha_r^*|X) > V(\alpha_0|X)$, since $V^{(1)}(\cdot|X) > 0$ and $U^{(1)}(V(\cdot|X)) > 0$. That is, the seller always benefits from choosing a reservation price above his own private value. The second implication of (2.2.3) is that the optimal screening level is independent upon I if $V(\cdot|X) = V(\cdot|X, I)$, i.e. the seller can choose his optimal screening level policy independently of the number of bidders participating in the auction.

This Chapter considers a CRRA utility function, i.e. $U(c;\theta) = c^\theta$, where $0 < \theta \leq 1$ is the risk aversion parameter and $(1-\theta)/c$ is the corresponding Arrow-Pratt risk aversion measure $-U^{(2)}(c)/U^{(1)}(c)$. The risk neutrality case is given by $\theta = 1$. As a result, the seller's expected payoff with screening level α_r is given by

$$\Pi(\alpha_r|X, I, v_0, \theta) = v_0^\theta \alpha_r^I + V(\alpha_r|X)^\theta I \alpha_r^{I-1}(1-\alpha_r) + I(I-1)\int_{\alpha_r}^1 V(\alpha|X)^\theta \alpha^{I-2}(1-\alpha)\, d\alpha.$$

The CRRA family of utility functions provides a measure of risk aversion decreasing in the monetary quantity c, which can be more useful than a CARA family in the framework considered here because the seller is relatively more risk averse as his gain in the auction approximates zero. Note that, for a given θ, $\partial U(c;\theta)/\partial c = \theta c^{\theta-1} \to \infty$ as $c \to 0$. In other words, when there is no outside option so that $v_0 = 0$, the

[25] The quantile density function is defined as the derivative of the quantile function with respect to α, i.e. $V^{(1)}(\alpha|X) = \partial V(\alpha|X)/\partial\alpha = 1/f_v(V(\alpha|X)|X)$.

seller has a strong incentive to sell the good in the auction since his marginal utility explodes in comparison with the no trade scenario, i.e. he has a strong incentive to deviate from the situation in which there is no trade. On the other hand, the utility implied by a CARA family is pratically the same for all values of c given its exponential behaviour. As a consequence, the seller would not have much incentive to sell the good when $c \to 0$ as in the CRRA case. Hu, Matthews and Zou (2010) show that the optimal reservation price decreases with the Arrow-Pratt measure of risk aversion $-U^{(2)}(c)/U^{(1)}(c)$. As a consequence, the optimal screening level of the CRRA case will increase with the risk aversion parameter θ.

2.3 Quantile Regression Estimation Methodology

This section describes the quantile regression methodology used to identify and estimate the private value conditional quantiles and the corresponding expected payoff maximizer α_r^*. The quantile regression approach, developed by Koenker and Bassett (1978), restricts the quantile representation (2.2.1) to a regression specification, such as

$$\begin{aligned} V(\alpha|X) &= X\gamma(\alpha) \\ &= \gamma_0(\alpha) + Z\gamma_1(\alpha), \end{aligned} \tag{2.3.4}$$

where $\gamma_0(\alpha)$ is the quantile regression intercept and $\gamma_1(\alpha)$ the quantile regression slopes. Note that both the intercept and the slope quantile regression coefficients depend upon the rank α of the bidder in the population. Therefore, changes in the conditioning variables not only shift the location of the conditional distribution of v, but may also affect its scale and shape. A shock on the covariate X may affect bidders differently according to their rank, i.e. a bidder with a low rank α may react in a different way than a bidder with a higher rank. A large variation of the slope and location coefficients, $\gamma_1(\alpha)$ and $\gamma_0(\alpha)$, may indicate strong heterogeneity among the bidders. From now on, the term "heterogeneity" will be used for brevity to represent this strong variation in the private values quantiles[26].

The quantile regression specification (2.3.4) is derived from Chapter 1 in this thesis, where it has been successfully tested against several alternatives. Note that it is assumed a linear specification for the private values conditional quantiles and also that the latter are independent upon the number of actual bidders I competing in the auction, i.e. $V(\cdot|X) = V(\cdot|X, I)$. The identification and estimation strategies are also derived from Chapter 1. As mentioned in the latter, the ascending auction format is specially suitable to identify the private value quantiles under the IPV paradigm since, as shown in Athey and Haile (2002), the c.d.f. of the winning bids can nonparametrically identify the private value distribution. The following assumptions are required for the identification and estimation of $V(\cdot|X)$:

Assumption 2.3. *The private value conditional quantiles has a quantile regression specification such as*

$$V(\alpha|X) = X\gamma(\alpha). \tag{2.3.5}$$

Assumption 2.4. *The auction specific variable, Z, has dimension P, with a compact support in $\mathcal{Z} \subset (0, +\infty)^P$ and a nonempty interior.*

Assumption 2.5. *The variables $\{I_\ell, X_\ell, v_{i\ell},\ i = 1, 2, \ldots I_\ell, \ell = 1, \cdots, L\}$ are independent and identically distributed. Conditional on X_ℓ, the private values $v_{i\ell}$ are independent with common c.d.f. $F_v(\cdot|X_\ell)$ and a density function $f_v(\cdot|X_\ell)$ bounded away from zero.*

Assumptions 2.3 deals with the quantiles of the private value distribution. It imposes a correct specification of the private value conditional quantiles. Assumption 2.4 ensures that if $x\gamma_1 = x\gamma_2$, for all $x \in \mathcal{X}$, thus $\gamma_1 = \gamma_2$. Both assumptions 2.3 and 2.4 are required for the private value quantile identification. Assumption 2.5 is necessary for estimation purposes and implies that each auction is independent and that, within an auction, the IPV paradigm holds.

Consider the winning bid b_{w_ℓ} at the auction ℓ, $\ell = 1, \cdots, L$, which is the amount paid by the winner. Let $B(\alpha|X)$ be the α-quantile of the winning bids conditional

[26]Note that heterogeneity and asymmetry are two concepts that should not be confused. Heterogenity is concerned with the variation of $\gamma(\alpha)$ across quantile levels α, while asymmetry implies that different bidders can have different coefficients $\gamma(\alpha)$.

distribution given the auction characteristics X. It follows from Athey and Haile (2002, equation (5)) that $\Psi_I(F_v(\cdot|X))$ is the distribution of the second highest private value, which is equal to the winning bid. Lemma 1.3 in Chapter 1 establishes that if the private value quantile admits a linear quantile regression representation such as (2.3.5), then the winning bid quantile function is also given by

$$B(\alpha|X) = X\beta(\alpha)$$

with $\beta(\alpha) = \gamma\left(\Psi_I^{-1}(\alpha)\right)$ for $\alpha \in [0,1]$, or equivalently,

$$\gamma(\alpha) = \beta(\Psi_I(\alpha)), \tag{2.3.6}$$

where $\Psi_I(t) = It^{I-1} - (I-1)t^I$. Equation (2.3.6) shows that the private value quantile regression coefficient $\gamma(\cdot)$ can be recovered from the winning bids coefficient $\beta(\cdot)$ but at a different quantile level $\Psi_I(\cdot)$. The coefficient $\gamma(\alpha)$ can thus be estimated via quantile regression method,

$$\widehat{\gamma}(\alpha) = \arg\min_{\gamma\in\Gamma} \frac{1}{L}\sum_{\ell=1}^{L} \rho_{\Psi_{I_\ell}(\alpha)}(b_{w_\ell} - X_\ell\gamma),$$

where $\rho_\alpha(u) = u(\alpha - \mathbb{I}(u<0))$ and Γ is a compact subset of $\mathbb{R}^P$.

I now turn to the estimation of the seller's expected payoff and its maximizer α_r^*. As mentioned in Chapter 1, it is more convenient to find the optimal screening level α_r^* as a maximizer of an estimation of $\Pi(\alpha_r|X, I, v_0, \theta)$ than to use the closed form FOC given in (2.2.3)[27]. This is because the latter requires an estimation of the private value quantile density function, which in turn cannot be estimated at a parametric rate since it would eventually involve a smoothing parameter. On the other hand, the private value conditional quantile can be estimated at a parametric rate due to the quantile regression methodology. The seller's expected payoff $\Pi(\alpha_r|X, I, v_0, \theta)$ will be estimated over a grid of prescribed quantiles $\mathcal{A} = \{0.12, 0.14, \cdots, 0.80\}$. The quantiles closer to the boundary are not considered given a lack of precision in the estimates due to sparsity of the data around this region. To estimate $\Pi(\alpha_r|X, I, v_0, \theta)$, I first estimate $V(\cdot|X)$ via quantile regression and then

[27]See Li, Perrigne and Vuong (2003) for a similar approach.

apply numerical integration using a trapezoildal rule over the grid $\mathcal{A}$ to approximate the definite integral in (2.2.2). Since the optimal screening level does not depend upon I by Proposition 2.1-(ii), the optimal screening level is then chosen over $\mathcal{A}$ such that an aggregate $\widehat{\Pi}\left(\alpha_r|X, v_0, \theta\right)$ is maximized, i.e.

$$\widehat{\alpha}_r^* = \arg\max \sum_{I=2}^{\overline{I}} \widehat{\Pi}\left(\alpha_r|X, I, v_0, \theta\right) L_I/L, \tag{2.3.7}$$

where $\left\{2, \cdots, \overline{I}\right\}$ is the support of I, L_I is the number of auctions in the sample with I bidders competing and L the total sample size. The estimation will be done for each value of θ in order to understand the impact of a risk averse seller on the optimal reservation price policy.

2.4 Empirical Results

This section investigates empirically whether the optimal screening level decreases when the seller becomes more risk averse. It is also investigated the sensitivity of α_r^* to changes in the auction covariates X and the seller's private value v_0. The former seems to play an important role in the recommendation policy, since e.g. goods with better quality may attract stronger bidders to the auction. The seller's private value v_0 represents the price at which the seller would sell the good in the outside market. In particular, $v_0 = 0$ may represent the case in which the seller has no opportunity to sell the good outside the auction.

The data set used here is publicly available on the internet and aggregates ascending auctions from the states covering the western half of the US (regions 1-6 as labeled by the USFS) occurred in 1979. It contains 472 auctions involving a total of 1175 bids and a set of variables characterizing each timber tract including the estimated volume of the timber measured in thousand of board feet (or mbf) and its estimated appraisal value given in Dollar per unit of volume. The latter measures how much the tract worths at the present time, taking into consideration a combination of factors as volume of each specie, quality and marketable price. Only scaled sales are considered, where bids are per unit of timber. Therefore, the winning bids are given in Dollar per unit of volume.

The government announces a reservation price prior to the auction that is equal to the appraisal value of the tract. The auction is conducted in two rounds: in the first round, bidders submit sealed bids[28] that must exceed the appraisal value of the tract to be qualified for the auction; in the second round, bidders compete in an ascending auction. See Lu and Perrigne (2008) and Aradillas-Lopez, Gandhi and Quint (2013) for more details about the dataset.

Table 2.1: Summary Statistics

	Winning Bids	Appraisal Value	Volume	Number of Bidders
Mean	129.17	72.34	3,171.10	2.49
Std. Deviation	119.76	52.92	4,418.20	0.50
25%	59.13	24.47	294	2
50%	113.81	68.38	967	2
75%	172.53	111.25	4,724	3
Min	0.30	0.25	12	2
Max	1,981.50	219.58	24,800	3
Skewness	8.16	0.44	1.99	0.04
Observations	472	472	472	472

Table 2.1 gives some descriptive statistics about the data set. The auctioned tract displays significant heterogeneity in quality and size. The mean of the variables are all above the median, indicating tailored marginal distributions. In particular, the marginal distribution of the winning bids is highly positive skewed. In the data set, the number of actual bidders I takes values 2 and 3, with $I = 2$ more than 50% of the cases. Table 2.2 provides the results of a median regression analysis of the winning bids on the timber appraisal value per mbf and the volume. Both variables are highly significant in the median analysis, which can be seen by the coefficient of determination $R_{0.50}$ at the median quantile[29]. For this reason, a two-dimensional

[28]Note that the bids in the first round are not proper bids, but a proposal to be qualified for bidding in the second round.

[29]The coefficient of determination is given by

$$R_\alpha = 1 - \frac{\widehat{Q}\left(\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2}|\alpha\right)}{\widehat{Q}\left(\widehat{\beta_0}|\alpha\right)},$$

where $\widehat{Q}\left(\widehat{\beta}|\alpha\right)$ is the optimized quantile regression objective function. It represents how much of the variability in the winning bids is explained by both covariates above at the quantile level α.

vector of covariates Z grouping both the appraisal value per mbf and the volume of the timber will be considered.

Table 2.2: Winning Bids

Variable	Coefficient	Std. Deviation	t-value	p-value
Constant	11.1265^{***}	3.8554	2.886	4.08E-03
Appraisal Value	1.2873^{***}	0.0505	25.514	0.000
Volume	0.0029^{***}	0.0008	3.503	0.0005
$R_{0.50}$	0.4159			

Significance level: * 10%, ** 5% and *** 1%.

Table 2.3 gives the estimates of the private values quantile regression coefficients. The most important variable is the appraisal value, a quality measure released by the seller, which is often interpreted as the seller's private value, see Lu and Perrigne (2008) and Aradillas-Lopez, Gandhi and Quint (2013). The associated quantile regression coefficient is given in the second column of the table. The appraisal value coefficient is always significant and larger than 1, suggesting that this coefficient acts as a markup indicating how much more the auctioned good appraisal value is valued by the bidders than the seller. Note that the private value can be interpreted as a measure of how much a bidder would be willing to sell a good made with the timber bought at the auction. This then suggests that the higher the bidder's private value, the higher is his efficiency in aggregating value to the timber. This strong variation in the private values quantile regression coefficients can be also interpreted as heterogeneity among the bidders.

The appraisal value coefficients also suggests that the associated rank of the appraisal value in the private values distribution is likely to be very low. Indeed in 26 out of 472 auctions, the appraisal value rank is in the range of quantiles $\alpha = 0.05$ and $\alpha = 0.08$, while most of the time it is below $V(0.05|X)$. This result is expected given the requirement of the first round of the game. Recall that in the first round, bidders submit sealed bids that must exceed the appraisal value of the tract to be qualified for the auction. Note as well that the appraisal value coefficients increase over the quantile levels, suggesting a relative increase in the markup of 75% when comparing bidders in the quantiles $\alpha = 0.10$ and $\alpha = 0.80$ of the private values conditional distribution. This is also evidence that bidders belonging to the upper tail of the private values distribution are more highly affected by changes in the appraisal value than median bidders[30].

Table 2.3: Private Value Quantile Regression Estimates

Quantile Level	Intercept	Appraisal Value	Volume
0.1	0.95	1.01	0.0007
	[-0.97,2.28]	[0.99,1.04]	[0.0004,0.0016]
0.2	3.00	1.04	0.0016
	[-0.72,8.49]	[0.99,1.13]	[0.0005,0.0027]
0.3	9.39	1.15	0.0018
	[2.05,15.01]	[1.05,1.22]	[0.0010,0.0033]
0.4	11.77	1.25	0.0034
	[5.32,20.83]	[1.14,1.33]	[0.0013,0.0049]
0.5	21.03	1.29	0.0041
	[10.92,29.24]	[1.22,1.43]	[0.0023,0.0054]
0.6	35.68	1.36	0.0041
	[21.63,45.02]	[1.27,1.56]	[0.0029,0.0055]
0.7	44.28	1.57	0.0045
	[29.94,77.15]	[1.22,1.81]	[0.0029,0.0071]
0.8	67.64	1.75	0.0060
	[32.91,101.02]	[1.31,2.02]	[0.0024,0.0138]
0.9*	72.98	2.37	0.0167
	[12.89,124.22]	[1.54,4.56]	[0.0031,0.0384]

The estimates are for a median auction and were computed by pooling the samples with different number of bidders. The 95% confidence interval of the quantile regression estimates in square brackets were computed using a bootstrap procedure that resample with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I;

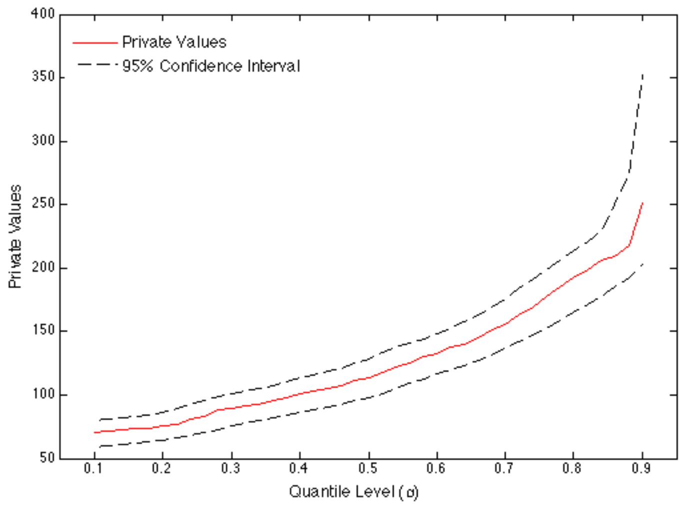
* Note the loss in precision when α gets closer to the upper boundary. This is why such higher quantile have been excluded from our test statistics.

In what follows $X(\tau) = (X_1(\tau), X_2(\tau), X_3(\tau))$, where $X_j(\tau)$ is the quantile of order τ of the variable X_j, $j = 1, 2, 3$. With some abuse of terminology, $X(\tau)$ will be called the quantile of order τ of the vector X, $X(0.50)$ being called its median. Figure 2.1 gives the private value conditional quantile estimates for a median auction and their 95% confidence intervals. Figure 2.2 presents the quantile estimates for

[30] In Chapter 1 of this thesis, the hypothesis of constancy of the slope coefficients across the entire distribution of private values is indeed rejected, which is evidence of strong heterogeneity among the bidders.

three quantile levels of $X(\tau)$, $\tau = \{0.15, 0.50, 0.85\}$, which represent auctions with low, median and high quality and size. In particular, it shows how the shape of the private value conditional quantile curve changes when there is variation in the quality and size of the timber tract.

Figure 2.1: Private Values Conditional Quantiles



Notes: The estimates are for a median auction. The 95% confidence intervals were computed using a bootstrap procedure that resample with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

Figure 2.2: Private Values Conditional Quantiles and $X(\tau)$

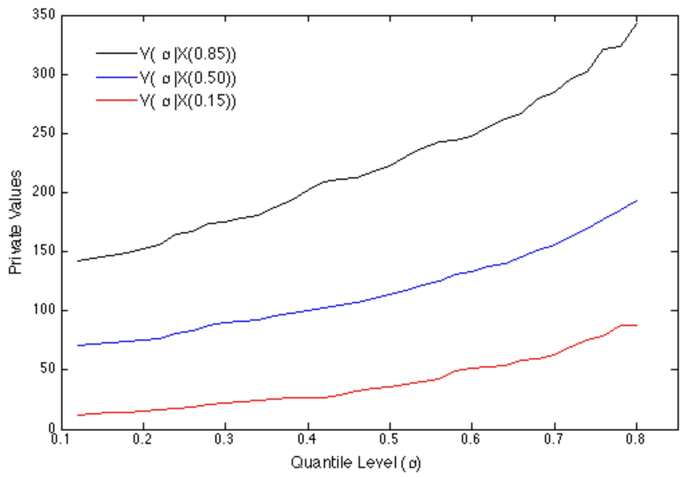


Figure 2.2 indeed shows that the auction covariates change significantly the shape of the private value distribution. This effect becomes even clear when comparing a high and a low quantile of the private value conditional distribution. Consider in particular the quantiles $\alpha = 0.12$ and $\alpha = 0.80$. The relative increase in the private value quantile is about 600% in the auctions with low quality and size, whereas it reduces to 172% and 142% in the median and high quality and size auctions, respectively. This strong variation in the private values seems to affect the optimal screening level policy recommended to the seller, since the latter has a strong incentive to increase the reservation price when bidders are more heterogeneous. Chapter 1 indeed shows that the optimal screening level in auctions with strong heterogeneity is $\alpha_r^* = 0.75$, which yields a probability of trading of 44% and 58% for 2 and 3 bidders, respectively. Note that $I = 2$ in more than half of the sample.

When the seller is risk averse, the concavity of his utility function may smooth out the effect of the strong variation in the private value conditional quantiles. This would then reduce the screening level that maximize the seller's expected payoff. Consider for instance a CRRA $U(\cdot;\theta)$ where $\theta = 0.60$. Figure 2.3 and 2.4 show how such a concavity in the utility function would smooth out the seller's payoff in auctions with low and high quality and size, respectively. The case $\theta = 1$ corresponds to the risk neutrality framework used in Chapter 1. Indeed the strong increase observed in the private value conditional quantile function is neutralized by the concave utility function. This is even more significant in the case of auctions with low quality and size, i.e. the ones with strong heterogeneity among the bidders.

The next three Figures give the optimal screening levels and their 95% confidence intervals for several values of the risk aversion parameter θ and auction quality and size. Each Figure compares the optimal screening level policy according to the choice of the seller's private value: $v_0 =$ Appraisal Value (AV) is shown in red and the case with no outside option $v_0 = 0$ is shown in blue. This is done to understand how the seller's optimal screening level policy changes according to his possible gains in the outside market.

A general conclusion from Figures 2.5, 2.6 and 2.7 is that α_r^* is very sensitive to changes in both (v_0, θ) at auctions with low quality and size, whereas it becomes stable as the quality and size of the auction improve. Consider first the case in which $v_0 =$ AV. Note that the optimal screening level reduces as the quality and

Figure 2.3: Private Values Transformed Quantile Function (X (0.15))

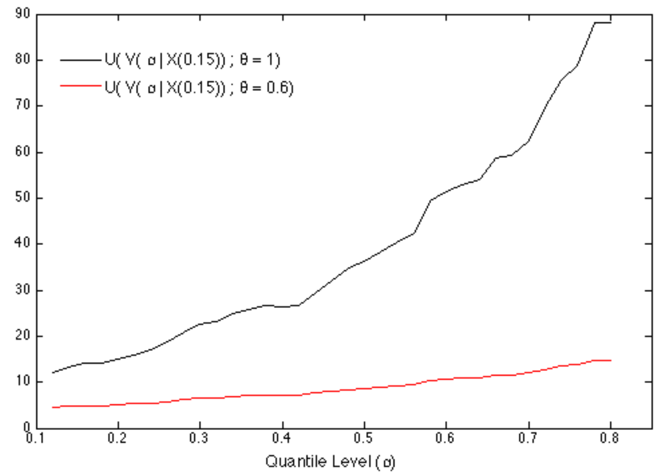
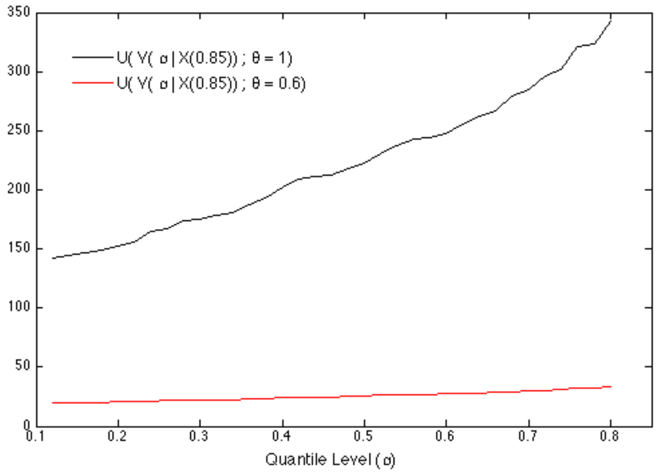


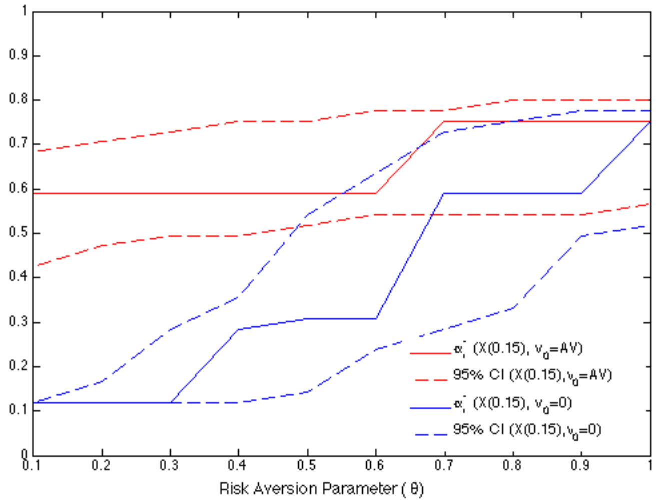
Figure 2.4: Private Values Transformed Quantile Function (X (0.85))



size of the auction increase. This reduction is however more significant when the seller is less risk averse (i.e., θ is closer to 1). This sensitivity to changes in θ can be better understood from Table 2.4. It seems that quality and size do not play a role when the seller becomes more risk averse. Observe that α_r^* remains almost constant when $\theta \leq 0.6$. Note that the optimal screening level reduces significantly when the seller's private value reduces to zero.

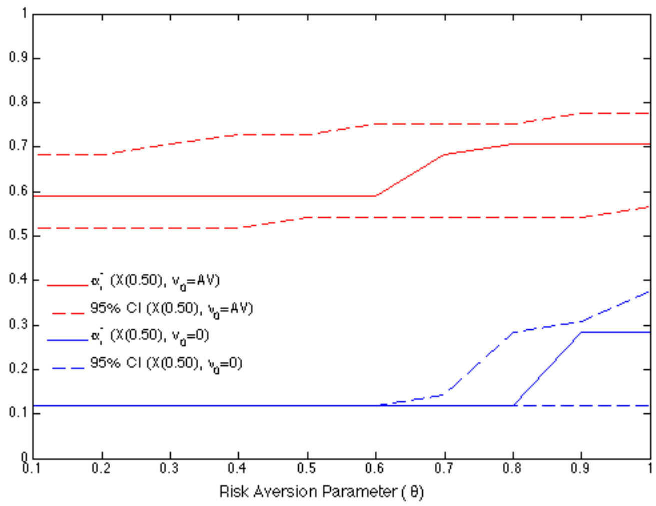
The probability of trading in the auction with a screening level α_r is given by

Figure 2.5: Optimal Screening Level for a Risk Averse Seller ($X(0.15)$)



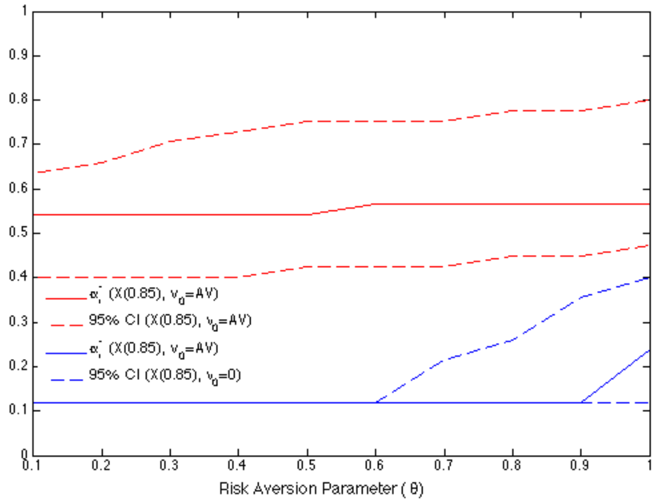
Notes: The estimates are for an auction of low quality and size, i.e. $X(0.15)$. The 95% confidence intervals were computed using a bootstrap procedure that resample with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

Figure 2.6: Optimal Screening Level for a Risk Averse Seller ($X(0.50)$)



Notes: The estimates are for a median auction, i.e. $X(0.50)$. The 95% confidence intervals were computed by resampling with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

Figure 2.7: Optimal Screening Level for a Risk Averse Seller ($X\,(0.85)$)



Notes: The estimates are for an auction of higher quality and size, i.e. $X\,(0.85)$. The 95% confidence intervals were computed using a bootstrap procedure that resample with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

$1-\alpha_r^I$, as mentioned in section 2.2. Figures 2.8, 2.9 and 2.10 show how the probability of trading changes when the quality and size of the auction change and the seller becomes more risk averse for the case in which $v_0 = \text{AV}$. As the optimal screening level decreases when the seller becomes more risk averse, the probability of trading then increases. It seems that a concave utility function for the seller has been indeed sufficient to reduce the optimal screening level, specially in the case in which $I = 3$. The probability of trading is moderately high (ranging e.g. from 65% to 71%) when $I = 2$ and the seller is very risk averse. It however increases to 80% when $I = 3$. The chance of selling the good also increases as the quality and size of the auction improves. Note as well that such a probability remains almost constant over the risk aversion parameter at these better auctions. Therefore, risk aversion does not seem to play a role in determining α_r^* at auctions with high quality and size.

Table 2.4: Optimal Screening Level for a Risk Averse Seller

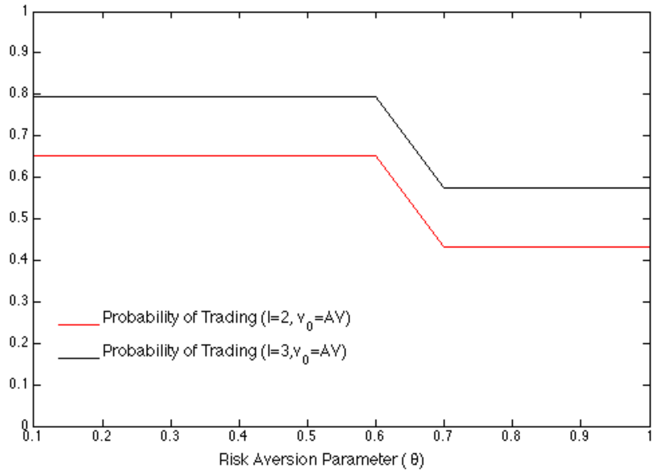
θ	$X\,(0.15)$	$X\,(0.50)$	$X\,(0.85)$
0.1	0.59	0.59	0.54
	[0.42,0.68][b]	[0.52,0.68]	[0.40,0.64]
0.2	0.59	0.59	0.54
	[0.47,0.71]	[0.52,0.68]	[0.40,0.66]
0.3	0.59	0.59	0.54
	[0.49,0.73]	[0.52,0.71]	[0.40,0.71]
0.4	0.59	0.59	0.54
	[0.49,0.75]	[0.52,0.73]	[0.40,0.73]
0.5	0.59	0.59	0.54
	[0.52,0.75]	[0.54,0.73]	[0.42,0.75]
0.6	0.59	0.59	0.56
	[0.54,0.78]	[0.54,0.75]	[0.42,0.75]
0.7	0.75	0.68	0.56
	[0.54,0.78]	[0.54,0.75]	[0.42,0.75]
0.8	0.75	0.71	0.56
	[0.54,0.80]	[0.54,0.75]	[0.45,0.78]
0.9	0.75	0.71	0.56
	[0.54,0.80]	[0.54,0.78]	[0.45,0.78]
1.0	0.75	0.71	0.56
	[0.56,0.80]	[0.56,0.78]	[0.47,0.80]

[a] The estimates cosider the seller's private value equal to the appraisal value of the timber, i.e. v_0 = AV.
[b] The 95% confidence intervals in square brackets were computed using a bootstrap procedure that resample with replacement the (X_ℓ, b_{w_ℓ})-pair in each original subsample L_I.

2.5 Conclusion

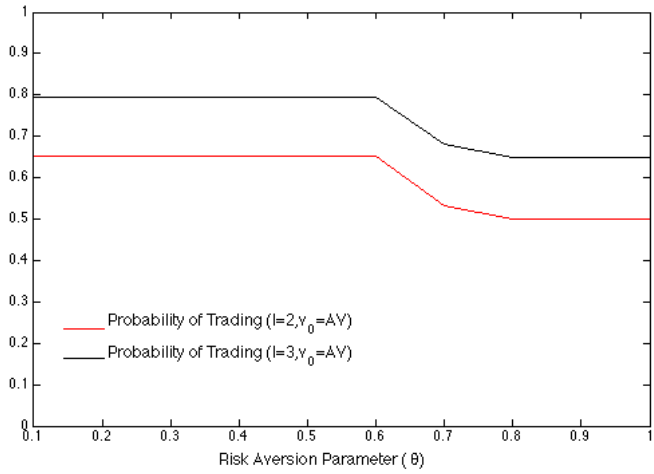
This second chapter of the thesis proposes a numerical investigation of the optimal screening level policy when seller becomes risk averse of not selling the good. As shown by the empirical analysis of Chapter 1, the optimal reservation price policy assuming a risk neutral seller may lead to a low probability of selling the good in ascending timber auctions. This undesirable feature was due to a sharp increase of

Figure 2.8: Probability of Trading for a Risk Averse Seller (X (0.15))



Notes: The estimates are for an auction of low quality and size, i.e. X (0.15).

Figure 2.9: Probability of Trading for a Risk Averse Seller (X (0.50))

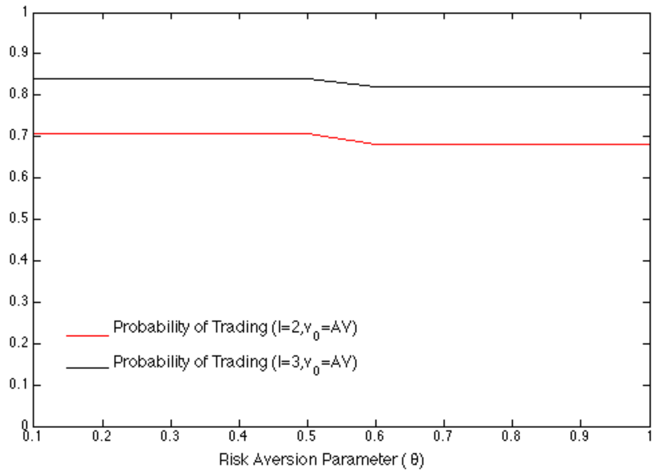


Notes: The estimates are for a median auction, i.e. X (0.50).

the private value quantile function, which leads to a high screening level policy in the risk neutrality case.

The results of Hu, Matthews and Zou (2010) suggest that taking into consideration risk aversion in the seller's utility function may be useful to provide more relevant policy recommendations. This motivates to consider a concave utility

Figure 2.10: Probability of Trading for a Risk Averse Seller (X (0.85))



Notes: The estimates are for an auction of higher quality and size, i.e. X (0.85).

function into the computation of the seller's expected payoff, which may neutralize the effect of the strong variation in the private value quantile function on the optimal screening level policy. This chapter shows that considering a CRRA family of utility functions for the seller can indeed help to address this issue related with the high screening levels since the probability of selling the good increased from 44% to about 80% for very risk averse sellers.

Some authors have previously mentioned that reservation prices were very high in timber auctions. Consider e.g. Aradillas-Lopez et al. (2013), which suggest that the IPV framework may not be appropriate to study ascending timber auctions. They show that this framework may overestimate the optimal reservation price policy for small values of I in comparison with the case in which private values are positively correlated. However, their approach is affected by the curse of dimensionality. Roberts and Sweeting (2012) also argue that the two stage nature of timber auctions may increase the optimal reservation price because the first stage of the game may select a few strong bidders and a bunch of low bidders expecting the absence of the latter. The present chapter proposes an alternative explanation which argues that high reservation prices are due to the shape of the private value quantile function and the assumption of a risk neutral seller. As shown empirically for USFS ascending timber auctions, assuming that the Government has an increasing CRRA utility is sufficient to achieve a reasonable high probability of selling the auctioned good.

Chapter 3

Augmented Quantile Regression for First-Price Auction

Abstract

This third chapter considers a quantile framework for first-price auction under the independent private value paradigm. A key stability property is that a linear specification for the private value conditional quantile function generates a linear specification for the bids one, from which it can be easily identified. This applies in particular for standard quantile regression models but also to more flexible additive sieve specification which are not affected by the curse of dimensionality. A combination of local polynomial and sieve methods allows to estimate the private value quantile function with a fast optimal rate and for all quantile levels in $[0, 1]$ without boundary effects. This allows to estimate the optimal bidding strategy and all bidder's private values near the boundaries with a fast rate. Extensions to binding reservation price and the case where only the winning bid is observed are considered. The choice of the smoothing parameters is also discussed.

3.1 Introduction

Many econometric approaches for auctions have been based on structural models. The equilibrium bidding strategy characterized by Riley and Samuelson (1981) involves numerical complexities because it depends on the latent private values distribution, which is unknown to the econometrician. Donald and Paarsch (1996, 2002) have proposed estimation approaches based on maximum likelihood. Their estimators achieve a rate of convergence faster than the usual parametric one. Laffont, Ossard and Vuong (1995) have proposed a simulation based estimation method allowing for general parametric specifications, see also Li (2005) for an extension to binding reservation price and entry. The semiparametric Bayesian framework of Li and Zheng (2009) allow for unobserved heterogeneity. Semi nonparametric methods were considered in Bierens and Song (2012). These parametric methods can be computationally demanding and are subject to possible misspecifications which may give erroneous policy recommendations. Closer to our approach is the simpler least squares procedure of Rezende (2008).

Another branch of the literature has developed a nonparametric approach. Since the private values are not observed, it may indeed be difficult to choose a parametric model and using a nonparametric method in a first stage may be suitable to avoid strong misspecifications. Guerre, Perrigne and Vuong (2000), hereafter GPV, have shown that, in a first step, it was possible to recover nonparametrically the latent private values from the observed bids. The second step uses the estimated private values to estimate the probability density function (p.d.f) of the true ones. The resulting estimator is computationally simple and rate-optimal under fairly weak smoothness conditions. However, as any other nonparametric methods, this approach is subject to the curse of dimensionality. That such issue is potentially important in empirical approaches can be seen in works as Haile and Tamer (2003), or more recently Aradillas-Lopez, Gandhi and Quint (2013), which in the simpler context of ascending auctions have considered 5 or 6 explanatory variables for at a best a few thousands observations. The parametric approach of Li and Zheng (2009) involve 8 variables and Athey, Levin and Seira (2011) investigates the effects of more than 10 variables. With 2,000 observations and 5 covariates with a uniform distribution over the unit cube, the length of a cube containing an average of 10 observations is 0.34, which is as large as one third of the range of each covariates and suggests that the bias of standard nonparametric procedures can be important for usual auction sample sizes. Because the estimation of the private values involves estimation of a conditional p.d.f, implementing dimension reduction approaches based on additive or single-index restrictions is not straightforward. Another undesirable feature is, as noted in Hickman and Hubbard (2013), the boundary bias which affects the estimation of the private values. This means in particular that it is difficult to recover the privates values of high bids which are especially important since such bidders have a high probability to win the auction. As a consequence empirical applications based on such estimated private values may fail to properly picture the upper tail of the private value distribution, although it is clearly a key component of the model.

The purpose of this third chapter of the thesis is to propose a new quantile approach which can deal with rich set of covariates and are free from asymptotic boundary bias. Quantiles have already been used in econometrics with various purposes. This covers risk aversion identification and estimation as in Campo (2012), Bajari and Hortaçsu (2005) for experimental data, Guerre, Perrigne and Vuong (2009) and Zincenko (2013) for field data. Haile, Hong and Shum (2003) have proposed a quantile based test of the common values hypothesis. Marmer and Shneyerov (2012) build on the insights of Haile et al. (2003) to estimate the p.d.f of the private values through the quantile function of the bids, bypassing the private values estimation step of GPV. Guerre and Sabbah (2012) have proposed to estimate the private value conditional quantile function, which is a weighted sum of the bids conditional quantile function and its derivative with respect to the quantile level as noted in Marmer and Shneyerov (2012). Enache and Florens (2012) have estimated the private values quantile function in a third-price auction model. Marmer, Shneyerov and Xu (2013a) have used a nonparametric estimation of quantiles to test for selective entry as Lee, Song and Whang (2013) who test dominance implications of the independent private value paradigm. Gimenes (2013) has studied quantile regression estimation in ascending auctions as a way to circumvent the curse of dimensionality and the need to choose a bandwidth which affect nonparametric estimation methods. Hence very few works have considered private values quantile function estimation for first-price auctions.

Quantile and quantile levels are useful for bidders and sellers. It is shown that private values quantiles can be used to compute a private value benchmark above which bidders have a prescribed probability to win the auction. Optimal bidding strategies can also be written using bid quantiles which can be estimated from auction data. For the seller, a key decision is the choice of a reservation price. Equating the reservation price with the cost quantile function gives the screening level, which should remain low enough to ensure that bidders will enter the auction and that the auctioned object will be sold. This chapter also shows that the expected utility of the seller can be easily expressed using the private value quantile function.

As noticed in Guerre and Sabbah (2012), a key point for econometric approaches is that the mapping which transforms the private value quantile function into its bid counterpart is linear, see also Gimenes (2013) for ascending auctions. Further implications are that a linear private value quantile specification, as for instance a quantile regression model, is mapped into a linear bid one and that this mapping is one to one due to identification as established in GPV, a key stability property of linear quantile specification. A contribution is to show that this leads to simple estimation methods for private value linear quantile specifications. This includes the quantile regression model of Koenker and Bassett (1978). This specification can be estimated with the optimal rate of GPV holding in the absence of covariate, independently of the number of explanatory used in the quantile regression. As a consequence, quantile regression models are especially convenient to handle large dataset with many covariates. However, as argued in Horowitz and Lee (2005), quantile regression may fit poorly in many applications and more flexible approaches are also needed as a potential alternative. The dimension reduction estimation strategy considered here extends the additive interactive regression (AIR) approach of Andrews and Whang (1990) to quantile models for first-price auctions. This amounts to consider a infinite dimensional sieve quantile regression model constrained by a certain level of interactions. This includes the nonparametric additive quantile regression model of Horowitz and Lee (2005) which corresponds to the case of no interactions. It is shown that the sieve quantile regression estimator achieves the optimal rate of GPV of the case where there is only one covariate, so that this specification can also be used to circumvent the curse of dimensionality. The case of a higher level of interactions is also considered.

A new aspect of the proposed quantile regression model deals with estimation of bid conditional quantile function first derivative with respect to the quantile level. Indeed, it is first shown that the private value conditional quantile function can be written as a sum involving the bid conditional quantile function and its derivative. To estimate the latter the sieve quantile regression objective function is modified by smoothing around the quantile level of interest. Such additional quantile level variation allows to estimate the quantile derivative with respect to the quantile level by standard local polynomial techniques as reviewed in Fan and Gijbels (1996). This is apparently new in both the local polynomial and quantile regression literature and is a key ingredient to obtain an estimation of the conditional private value quantile function which is free of boundary bias. Due to the additional smoothing and derivative estimation, it is refered as an *augmented* quantile approach all along the chapter. Estimating the conditional private value quantile function without asymptotic boundary bias allows to estimate all the bidders' private values. This can be important for applications which, as in Cassola, Hortaçsu and Kastl (2013), implement further statistical analysis using the estimated private values as if they were the true ones.

The rest of the chapter is organized as follows. Section 2 explains identification of the conditional private value quantile function and studies the transformation which maps it into a conditional bid quantile function. This is in particular useful to understand how a sieve expansion for the conditional private value quantile function generates a sieve expansion for the conditional bid quantile function with similar convergence properties, which is important for sieve estimation. Section 3 presents the new quantile regression estimation strategy. The uniform convergence rate, pointwise bias variance expansion and asymptotic normality are obtained for the quantile regression estimation procedure. For the sake of brevity, the results for its nonparametric additive extension focuses on integrated mean square error and uniform consistency rates. Section 4 considers some extensions of interest as estimation of the optimal bidding strategy and the private values, the cases where only the winning bid is observed or where the reservation price is binding. The choice of a bandwidth parameter and of the order of the sieve is also discussed and investigated in a small simulation experiment. Section 5 concludes the chapter and Section 6 gathers the proofs of the theoretical results.

3.2 First price auction and quantile regression models

A single and indivisible object with some characteristic $x \in \mathbb{R}^d$ is auctioned to $I \geq 2$ risk-neutral buyers, with a reservation price r chosen by a risk-neutral seller. The reservation price r is announced prior the auction. The potential number of bidders I, covariates x and reservation price r are known to the seller, the bidders and the econometrician, while the reservation price r may be only known to the seller. The object is sold to the highest bidder who pays his bid to the seller, provided it is at least as high as the reservation price r. In the sealed bids framework considered here, each bids are sealed so that a bidder does not know others' bid when forming his own bid. Within the Independent private value (IPV) paradigm, each potential bidder is assumed to have a private value, which he knows, $V_i \geq 0$, $i = 1, \ldots, I$ for the auctioned object. He does not know the private value of the other bidders. The bidders and the seller know that the private values have been independently drawn from a common distribution given (x, I) with conditional cumulative distribution (c.d.f) $F(\cdot|x, I)$, or equivalently with conditional quantile function $V(\alpha|x, I)$, $\alpha \in [0, 1]$, which is such that

$$V(\alpha|x, I) = \inf \{v \in \mathbb{R} : F(v|x, I) \geq \alpha\}.$$

When the private value conditional distribution is absolutely continuous with a probability density function (p.d.f) $f(\cdot|x, I)$ positive on its support as assumed from now on, $V(\alpha|x, I)$ is the reciprocal function $F^{-1}(\alpha|x, I)$. It is also well-known that $U_i = F(V_i|x, I)$, which can be viewed as the rank of the ith bidder in the private value population, is independent of x and I with a uniform distribution over $[0, 1]$. It follows from the IPV paradigm that the private value ranks $U_i = 1, \ldots, I$ are independent. In other words, the dependence between the private value V_i and the auction covariates x and I is fully captured by the nonseparable model,

$$V_i = V(U_i|x, I), \quad U_i \overset{\text{i.i.d}}{\sim} \mathcal{U}_{[0,1]} \perp (x, I). \tag{3.2.1}$$

As a consequence each bidder are identical up to the variable U_i which describes the ith bidder's rank in the private value distribution. Since the U_i are i.i.d with a known distribution, the bidders are identical ex ante and the game is symmetric.

In a first-price auction, the ith bidder submits a sealed bid B_i and wins the auction if $B_i > \max_{j \neq i} B_j$ and $B_i \geq r$, giving a payoff $V_i - B_i$ while his payoff is 0 otherwise. The unique symmetric differentiable Bayesian Nash equilibrium of the game has been characterized by many authors, including Riley and Samuelson (1981), Maskin and Riley (1984) and Krishna (2002). A key finding is that the optimal bids B_i are given by a strictly increasing and continuous transformation of the private values V_i above r,

$$B_i = s_r\left(V_i; x, I\right), \quad s_r\left(v; x, I\right) = v - \frac{1}{F\left(v|x, I\right)^{I-1}} \int_r^v F^{I-1}\left(t|x, I\right)^{I-1} dt \text{ for } v \geq r, \tag{3.2.2}$$

assuming that the reservation price r lies in the support of the private value conditional distribution. We shall consider here the case of a non binding reservation price $r = 0$. Especially relevant for a quantile approach is the fact that the optimal bidding strategy $s_0\left(\cdot; x, I\right)$ is strictly increasing. Indeed, if $B\left(\alpha|x, I\right)$ denotes the common conditional equilibrium quantile function of the distribution of i.i.d optimal bids B_i, the quantile invariance property implies that,

$$B\left(\alpha|x, I\right) = s_0\left(V\left(\alpha|x, I\right); x, I\right). \tag{3.2.3}$$

The fact that $s_0\left(\cdot; x, I\right)$ is increasing and (3.2.3) have several important implications for the bidder. In the next lemma, $G\left(b|x, I\right)$ stands for the conditional c.d.f. of an optimal bid B_i, i.e. $G\left(\cdot|x, I\right) = B^{-1}\left(\cdot|x, I\right)$.

Lemma 3.1. *Consider a given (x, I). Suppose that the independent private value paradigm holds and that $\alpha \in [0, 1] \mapsto V\left(\alpha|x, I\right)$ is continuously differentiable with a strictly positive α-derivative $V^{(1)}\left(\alpha|x, I\right)$. Then*

i. The private value ranks $U_i = F\left(V_i|x, I\right)$ are identical to the bid ranks $G\left(B_i|x, I\right)$,

$$U_i = F\left(V_i|x, I\right) = G\left(B_i|x, I\right) \text{ for all } i = 1, \ldots, I.$$

ii. The optimal bids are given by,

$$B_i = B\left(U_i|x, I\right) = B\left(F\left(V_i|x, I\right)|x, I\right), \text{ for all } i = 1, \ldots, I. \tag{3.2.4}$$

iii. Suppose that the ith bid is $B(a|x, I)$ for some $a \in [0, 1]$ and assume that the other bidders make optimal bids. Then the probability that the ith bidder wins the auction is a^{I-1}.

Lemma 3.1-(i) states that the rank U_i is an invariant of the model which is a key property and implies (ii) and (iii). Such invariance also holds for many different models involving an increasing strategy. Lemma 3.1-(ii) gives alternative expressions for the optimal bidding strategy $s_0(\cdot|x, I)$ of (3.2.2). In the case where the ith bidder knows his private value rank U_i, an optimal bid can be computed from the conditional bid quantile function as seen from (3.2.4) and, as $B(\cdot|x, I)$, can be estimated from the observed bids. Computing an optimal bid from the private value V_i is more involved due to the private value distribution $F(\cdot|x, I)$.[31] Lemma 3.1-(iii) has important practical implications. Assuming that all bidders use optimal bids and combined with (i), (iii) implies that a bidder with a private value $V(a|x, I)$ has a probability a^{I-1} to win the auction. If a bidder wants to win the auction with a prescribed probability π, he must ensure that his private value is larger or equal than $V\left(\pi^{1/(I-1)}|x, I\right)$. Lemma 3.1 has also some important econometric implications. Since the ranks U_i are an invariant of the model, it can be estimated from the observed bids, so that the private values $V_i = V(U_i|x, I)$ can be estimated from an estimation of the private value conditional quantile function. The expression (3.2.4) implies that an optimal bid can be estimated using an estimation of the conditional quantile bid function $B(\cdot|x, I)$ as soon as the rank U_i is known. When the bidder knows his private value V_i but not his rank U_i, an additional estimation of the private value conditional c.d.f. $F(\cdot|x, I) = V^{-1}(\cdot|x, I)$ is needed to estimate an optimal bid. Such econometric applications of Lemma 3.1 are discussed in Section 3.4.

[31] Interestingly, Lemma 3.1-(i) also holds when the bidders are identically risk averse, in which case the optimal strategy depends upon risk aversion. However, Lemma 3.1-(ii) also holds and shows that an optimal bid can be estimated without estimating the utility function, through an estimation of $B(\cdot|x, I)$ only when using the rank U_i.

From the seller perspective, the knowledge of the conditional function $V(\alpha|x, I)$ is sufficient to compute the expected revenue of a large class of auction mechanisms with a reservation price r as shown by the next lemma. To ensure that the expected revenue equivalence theorem holds, consider an auction mechanism $\mathcal{A}$ such that (i) any bid higher than r can bid the auction; (ii) the highest bid is awarded the auctioned object; (iii) each bidders are treated alike, and (iv) there is a strictly increasing common equilibrium strategy $B_i = b(V_i; x, I)$ for each buyer willing to bid.

Lemma 3.2. *Consider some (x, I) for which $\alpha \in [0,1] \mapsto V(\alpha|x, I)$ is strictly increasing and continuously differentiable. Let $\alpha_r(x, I)$ be the screening level associated with the public reservation price $r(x, I) \in [V(0|x, I), V(1|x, I)]$, that is the unique quantile level such that,*

$$r(x, I) = V(\alpha_r(x, I)|x, I).$$

Suppose that the seller is risk-neutral with a valuation $V_0(x, I)$ of the auctioned good. Then, under the IPV paradigm, the common equilibrium bidding strategy $B_i = b(V_i; x, I)$ for any auction mechanism $\mathcal{A}$ gives an expected utility to the seller of

$$V_0(x, I)\,\alpha_r(x, I)^I + I\left\{ r(x, I)\,\alpha_r(x, I)^{I-1}(1-\alpha_r(x, I)) + (I-1)\int_{\alpha_r(x,I)}^{1} V(\alpha|x, I)\,\alpha^{I-2}(1-\alpha)\,d\alpha \right\}.$$

The expression of the expected utility of the seller differs from the one given in Riley and Samuelson (1981) which involves the conditional p.d.f of the private values, or equivalently the $\alpha-$derivative $V^{(1)}(\alpha|x, I)$, but can easily be recovered from Krishna (2002, p.25). The expression in Lemma 3.2 is more convenient for estimation purpose because, in a nonparametric setup, $V(\alpha|x, I)$ can be estimated with a faster rate than $V^{(1)}(\alpha|x, I)$. The first-order condition associated with the maximization of the expected revenue as a function of the screening level writes, for $V^{(1)}(\alpha|x, I) = \partial V(\alpha|x, I)/\partial\alpha$,

$$(1-\alpha_r(x, I))\,V^{(1)}(\alpha_r(x, I)|x, I) = V(\alpha_r(x, I)|x, I) - V_0(x, I),$$

which since $r(x,I) = V(\alpha_r(x,I)|x,I)$, $\alpha_r(x,I) = F(r(x,I)|x,I)$ and

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

is identical to the first-order condition $r(x,I) - (1 - F(r(x,I)|x,I))/f(r(x,I)|x,I) = V_0(x,I)$ in Krishna (2002, eq. 2.12). When $V(\alpha|x,I) = V(\alpha|x)$ and $V_0(x,I) = V_0(x)$, that is when there is no unobserved heterogeneity, both the optimal screening level α_r^* and reservation price $r^* = V(\alpha_r^*|x)$ does not depend upon the number of bidders I.

How econometric modelling can benefit from a quantile perspective is clarified by the next theorem. Theorem 3.1 has two different purposes. The first part solves the auction game by computing the equilibrium bid quantile function $B(\cdot|x,I)$ as a function of the private value one $V(\cdot|x,I)$. The second part of Theorem 3.1 explains how to identify the private value quantile function $V(\cdot|x,I)$ from the bid one $B(\cdot|x,I)$, see (3.2.6) below. It is therefore the cornerstone of a quantile identification strategy. In Theorem 3.1 and all over the chapter, $V^{(1)}(\alpha|x,I)$ stands for the α-derivative $\partial V(\alpha|x,I)/\partial\alpha$.

Theorem 3.1. *Suppose that the reservation price r is not binding and that $I \geq 2$.*

i. Assume that $V(\alpha|x,I)$ is continuous over $[0,1]$ with $\inf_{\alpha\in[0,1]} V^{(1)}(\alpha|x,I) > 0$. The conditional equilibrium quantile function $B(\cdot|x,I)$ of the I i.i.d optimal bids B_i satisfies,

$$B(\alpha|x,I) = \frac{I-1}{\alpha^{I-1}} \int_0^\alpha t^{I-2} V(t|x,I)\, dt, \tag{3.2.5}$$

which is continuously differentiable over $[0,1]$.

ii. Let $B(\alpha|x,I)$ be a continuously differentiable function with respect to $\alpha \in [0,1]$ and assume it is the common conditional equilibrium quantile function of I i.i.d optimal bids B_i generated from I i.i.d private values drawn from $V(\cdot|x,I)$ as in Part (i). Then it must hold that,

$$V(\alpha|x,I) = B(\alpha|x,I) + \frac{\alpha B^{(1)}(\alpha|x,I)}{I-1}. \tag{3.2.6}$$

As in GPV, equation (3.2.6) in Theorem 3.1 shows that the theoretical auction model imposes some restrictions on the quantile function of the observed bids. Indeed the LHS of (3.2.6) must increase with α.[32] Interestingly, an heuristic argument for proving Theorem 3.1 will derive (3.2.6) in the first place. Consider a bidder with rank α and suppose that he bids $B(a|x,I)$ instead of his optimal bid $B(\alpha|x,I)$. Since the probability of winning with such a bid is a^{I-1}, it follows that his expected payoff satisfies,

$$(V(\alpha|x,I) - B(a|x,I))\, a^{I-1} \leq (V(\alpha|x,I) - B(\alpha|x,I))\, \alpha^{I-1},$$

meaning that the LHS is maximal when $a = \alpha$. The associated first-order condition is,

$$(I-1)(V(\alpha|x,I) - B(\alpha|x,I))\, \alpha^{I-2} - B^{(1)}(\alpha|x,I)\, \alpha^{I-1} = 0,$$

and rearranging gives (3.2.6). Viewing (3.2.6) as a differential equation with initial condition $B(0|x,I) = V(0|x,I)$ and solving gives (3.2.5) in a second step.

Another important implication of Theorem 3.1 comes from the linear functional relationship between the private value and bids quantile functions. Consider a linear model, possibly infinite dimensional, for the conditional private value quantile function,

$$\mathcal{M}_V = \{V(\alpha|x,I)\,; V(\alpha|x,I) = \mathbb{L}(x, \gamma(\alpha|I))\},$$

where, for each x, $\gamma \mapsto \mathbb{L}(x,\gamma)$ is a known linear operator over a finite or infinite linear span containing the admissible parameter γ, which can depend upon I. Then, up to some regularity conditions, (3.2.5) maps $\mathcal{M}_V$ into a linear model $\mathcal{M}_B$ for the conditional bid quantile function,

$$\mathcal{M}_B = \left\{B(\alpha|x,I)\,; B(\alpha|x,I) = \mathbb{L}(x, \beta(\alpha|I)) \text{ with } \beta(\alpha|I) = \frac{I-1}{\alpha^{I-1}} \int_0^\alpha t^{I-2} \gamma(t|I)\, dt\right\},$$

for some admissible $\gamma(\cdot|\cdot)$. Hence linearity of a private value quantile specification is preserved when turning to the quantile function of the observed bids, a key stability property. Reciprocally, given a parameter value $\beta(\alpha|I)$ of the bid linear

[32] Quantile functions which do not satisfy this condition can be derived from the c.d.f. $G(b) = (b/(5-4b))^{1/5}$ and $I = 2$ as in GPV, p. 530.

quantile model such that $B(\alpha|x, I) = \mathbb{L}(x, \beta(\alpha|I))$, (3.2.6) allows to recover the corresponding parameter of the private value linear quantile model,

$$V(\alpha|x, I) = \mathbb{L}(x, \gamma(\alpha|I)) \text{ with } \gamma(\alpha|I) = \beta(\alpha|x, I) + \frac{\alpha\beta^{(1)}(\alpha|x, I)}{I-1}.$$

Hence estimating a linear private value quantile specification is feasible as soon as it is possible to estimate the coefficient $\beta(\alpha|I)$, together with $\beta^{(1)}(\alpha|I)$, in the associated linear bid quantile model.

This includes the quantile regression specification pioneered in Koenker and Bassett (1978) and a sieve additive interactive quantile version of the additive interactive regression of Andrews and Whang (1990). The quantile regression model writes, for the $(d+1) \times 1$ vector $X = (1, x')'$ and a conformable vector parameter $\gamma(\alpha|I) = \left(\gamma_0(\alpha|I), \gamma_1(\alpha|I)'\right)'$,

$$V(\alpha|x, I) = X'\gamma(\alpha|I) = \gamma_0(\alpha|I) + x'\gamma_1(\alpha|I), \quad \text{for all } \alpha \in [0, 1]. \tag{3.2.7}$$

The stability property stated in Theorem 3.1-(i) implies that the conditional bid quantile function satisfies,

$$B(\alpha|x, I) = X'\beta(\alpha|I) \text{ with } \beta(\alpha|I) = \frac{I-1}{\alpha^{I-1}} \int_0^\alpha t^{I-2}\gamma(t|I)\, dt. \tag{3.2.8}$$

When the private value quantile regression specification (3.2.7) is correct, the represention (3.2.1) shows that,

$$V_i = \gamma_0(U_i|I) + x'\gamma_1(U_i|I),$$

showing that the private value response to a change of the covariate can depend upon the rank U_i of the bidder. Setting $\overline{\gamma}(I) = \mathbb{E}[\gamma(U_i|I)|I] = \int_0^1 \gamma(\alpha|I)\, d\alpha$ and $\eta_i = X'(\gamma(U_i|I) - \overline{\gamma}(I))$ gives the linear regression model

$$V_i = X'\overline{\gamma}(I) + \eta_i,$$

where η_i is an heteroskedastic regression error term. When the slope coefficient $\gamma_1(U_i|I)$ is constant so that $\gamma_1(U_i|I) = \overline{\gamma}_1(I)$, $\eta_i = \gamma_0(U_i|I) - \overline{\gamma}_0(I)$ becomes homoskedastic. Such linear regressions were previously considered in Rezende (2008). However Rezende (2008) has focused on the estimation of the slope and intercept coefficients $\overline{\gamma}(I)$. This does not allow without further restrictions for the distribution of η_i to compute the expected utility of the seller in full generality. Indeed, as seen from Lemma 3.2 depends upon an integral of $\gamma(\alpha|I)$ which differs from $\overline{\gamma}(I)$, and $\gamma(\alpha_r(x,I)|I)$ where $\alpha_r(x,I)$ is the screening level associated with the reservation price. Hence focusing on the estimation of $\overline{\gamma}(I)$ is not sufficient to compute an optimal reservation price without further assumption on the regression error term or additional estimation steps. This contrasts with the quantile regression specification (3.2.7) which allows such computation because it is a functional model for the whole quantile function. In addition, as seen from Koenker and Bassett (1978) and in the ideal case of observed private values, this specification can be estimated with a parametric rate although it includes the infinite dimensional parameter $\gamma(\cdot|I)$.

Other specifications considered in the next section cannot be estimated with a parametric rate but are, on one hand more flexible and more robust to misspecifications than quantile regression and, in the other hand, not subject to the curse of dimensionality which plagues fully nonparametric estimation methods for $V(\alpha|x,I)$. Consider the additive quantile specification,

$$\mathfrak{M}_1 = \left\{ V(\alpha|x,I); V(\alpha|x,I) = \sum_{j=1}^{d} V_j(\alpha; x_j, I) \right\},$$

with a suitable convention to ensure identification. Since such quantile specifications are obtained by summing some univariate functions, the effective dimension of this model is 1 and is lower that the dimension d of the model allowing for all quantile functions. Following Andrews and Whang (1990) or Horowitz and Lee (2005) and given a sieve $\{p_k(\cdot), k \geq 1\}$ for univariate functions,[33] $\mathcal{M}_1$ can be embedded in the infinite dimension linear model,

$$\mathcal{M}_1 = \left\{ V(\alpha|x,I); V(\alpha|x,I) = \sum_{j=1}^{d}\sum_{k=1}^{\infty} \gamma_{jk}(\alpha|I)\, p_k(x_j) \right\}.$$

[33]Note that a different basis system could be considered for each variables. Our results can easily extended to consider this setup.

As in Andrews and Whang (1990), the number of covariate interactions can be increased. Allowing for pairwise interactions leads to the linear specification,

$$\mathcal{M}_2 = \left\{ V(\alpha|x, I); V(\alpha|x, I) = \sum_{j=1}^{d} \sum_{k=1}^{\infty} \gamma_{jk}(\alpha|I) p_k(x_j) \right.$$
$$\left. + \sum_{1 \le j_1 < j_2 \le d} \sum_{k_1, k_2 = 1}^{\infty} \gamma_{j_1 j_2 k_1 k_2}(\alpha|I) p_{k_1}(x_{j_1}) p_{k_2}(x_{j_2}) \right\}$$

with effective dimension 2. All these models are a specific case of the infinite linear sieve quantile regression model,

$$\mathcal{M} = \left\{ V(\alpha|x, I); V(\alpha|x, I) = \sum_{k=1}^{\infty} \gamma_k(\alpha|I) P_k(x) \right\},$$

which gives $\mathcal{M}_1$ or $\mathcal{M}_2$ by a suitable choice of the system $\{P_k(\cdot), k \ge 1\}$. Note also that, when $\{P_k(\cdot), k \ge 1\}$ is a sieve allowing approximation of all continuous functions $Q(x)$, $\mathcal{M}$ can be identified with the a non restricted model allowing for all possible interactions between the covariates x_j, $j = 1, \ldots, d$. This can be useful for testing purpose.

However, finding the expression of the conditional bid quantile function generated by $\mathcal{M}$ needs some regularity conditions that are stated in the next lemma. The key conditions is the uniform convergence of the series expansion of $V(\alpha|x, I)$. In what follows, $\mathcal{X}$ stands for the support of the covariate x.

Lemma 3.3. *Assume that, for some continuous $\gamma_k(\cdot|I)$, $k \ge 1$,*

$$V(\alpha|x, I) = \sum_{k=1}^{\infty} \gamma_k(\alpha|I) P_k(x), \tag{3.2.9}$$

converge uniformly over $[0, 1] \times \mathcal{X}$. Then the conditional bid quantile function in

(3.2.6) satisfies,

$$B(\alpha|x,I)=\sum_{k=1}^{\infty}\beta_k(\alpha|I)P_k(x),\ \beta_k(\alpha|I)=\frac{I-1}{\alpha^{I-1}}\int_0^{\alpha}t^{I-2}\gamma_k(t|I)\,dt, \tag{3.2.10}$$

$$V(\alpha|x,I)=\sum_{k=1}^{\infty}\left(\beta_k(\alpha|I)+\frac{\alpha\beta_k^{(1)}(\alpha|I)}{I-1}\right)P_k(x), \tag{3.2.11}$$

where the two series expansion converge uniformly over $[0,1]\times\mathcal{X}$*, and, for all* $k\geq 1$*,* $\beta_k(\alpha|I)$ *is continuously differentiable over* $(0,1]$ *and* $\alpha\beta_k^{(1)}(\alpha|I)$ *is continuous over* $[0,1]$.

Starting from the infinite dimensional private value quantile regression specification (3.2.9), Lemma 3.3 derives the corresponding bids quantile regression specification (3.2.10), and (3.2.11) explains how to recover the conditional private value quantile function from the bid one. Before turning to estimation of the finite or infinite dimensional quantile regression specifications, we complete Theorem 3.1 and Lemma 3.3 by analyzing how the first-price auction model acts on the series convergence rate and smoothness. Let $\mathcal{X}\subset\mathbb{R}^d$ be the support of x, which is assumed to be compact with an nonempty interior. Let s_1, s_2 be two nonegative integer numbers and $\mathcal{P}_{s_1,s_2}\subset\mathcal{M}$ the set of functions $Q(\alpha,x)$ such that, for some integer number $0\leq d_{\mathcal{M}}\leq d$

i. For each K, there exists some real numbers $q_1(\alpha),\ldots,q_K(\alpha)$ which may depend upon K such that

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|Q(\alpha,x)-\sum_{k=1}^{K}q_k(\alpha)P_k(x)\right|=o\left(K^{-\frac{s_2}{d_{\mathcal{M}}}}\right). \tag{3.2.12}$$

ii. The functions $\alpha\in[0,1]\mapsto q_k(\alpha)$ are s_1 times continuously differentiable with, for some $Q^{(s_1)}(\alpha,x)$ continuous over $(\alpha,x)\in[0,1]\times\mathcal{X}$

$$\lim_{K\to\infty}\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|Q^{(s_1)}(\alpha,x)-\sum_{k=1}^{K}q_k^{(s_1)}(\alpha)P_k(x)\right|=0. \tag{3.2.13}$$

As seen from (3.2.12), the parameter s_2 describes the convergence rate of the series expansion $Q(\alpha, x) = \sum_{k=1}^{\infty} q_k(\alpha) P_k(x)$ up to a dimension parameter $d_{\mathcal{M}}$ which corresponds to the number of interactions allowed in $\mathcal{M}$. A limit case is $d_{\mathcal{M}} = 0$, in which case $Q(\alpha, x) = \sum_{k=1}^{K} q_k(\alpha) P_k(x)$ for some K, as for instance when the quantile regression specification (3.2.7) holds. Consider now the case of a univariate x, i.e $d = d_{\mathcal{M}} = 1$ in which case we use sieve $\{p_k(\cdot), k \geq 1\}$ instead of $\{P_k(\cdot), k \geq 1\}$. As surveyed in Chen (2007), the rate condition (3.2.12), $o(K^{-s_2})$, typically holds for functions $Q(\alpha, x)$ such that $x \in \mathcal{X} \longmapsto Q(\alpha, x)$ is s_2 times continuously differentiable and $\{p_k(\cdot), k \geq 1\}$ is a suitable sieve sequence, like Trigonometric or Legendre polynomial functions, B-splines or wavelets. For these examples of sieves given by orthonormal system, $q_k(\alpha)$ is given by the scalar product,

$$q_k(\alpha) = \int_{\mathcal{X}} Q(\alpha, x) p_k(x) dx.$$

When $\alpha \in [0, 1] \mapsto Q(\alpha, x)$ is s_1 times continuously differentiable,

$$q_k^{(s_1)}(\alpha) = \int_{\mathcal{X}} Q^{(s_1)}(\alpha, x) p_k(x) dx,$$

and, as shown by taking $s_2 = 0$ in (3.2.12), (3.2.13) holds provided $Q^{(s_1)}(\alpha, x)$ is continuous over $[0, 1] \times \mathcal{X}$.

Let us now turn to the case of a multivariate x, i.e $d > 1$. In the case of the additive specification $\mathcal{M}_1$ with one interaction, the convergence rate (3.2.12) holds for $d_{\mathcal{M}} = 1$ taking a sieve sequence $\{P_k(\cdot), k \geq 1\}$ which writes $\{p_k(x_1), \ldots, p_k(x_d), k \geq 1\}$ where $\{p_k(\cdot), k \geq 1\}$ is as in the examples of the univariate case. The case of the additive specification $\mathcal{M}_2$ with two interactions is more complicated because the corresponding sieve is

$$\{p_{k_1}(x_{j_1}) p_{k_2}(x_{j_2}), k_1, k_2 \geq 1, 1 \leq j_1 < j_2 \leq d\}.$$

Taking the first K elements of this sieve amounts to take for instance all the product $p_{k_1}(x_{j_1}) p_{k_2}(x_{j_2})$ with $k_1, k_2 \leq K^{1/2}$. This new upper bound affects the convergence rate which is now $o\left(\left(K^{1/2}\right)^{-s_2}\right) = o\left(K^{-s_2/2}\right)$, i.e $d_{\mathcal{M}} = 2$ in (3.2.12). The next Proposition describes in particular the implications of assuming that

$V(\alpha|x,I) \in \mathcal{P}_{s+1,s+2}$ for the convergence rate of the series expansion of $B(\alpha|x,I)$.

Proposition 3.1. *Suppose that $V(\alpha|x,I)$ is a function of $\mathcal{P}_{s+1,s+2}$ for each $I \in \mathcal{I}$ satisfying,*

$$\inf_{(\alpha,x)\in[0,1]\times\mathcal{X}} V^{(1)}(\alpha|x,I) > 0 \text{ and } \sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}} V^{(1)}(\alpha|x,I) < \infty.$$

Then, for $B(\alpha|x,I)$ as in (3.2.5),

i. $\min_{(\alpha,x)\in[0,1]\times\mathcal{X}} \partial B(\alpha|x,I)/\partial\alpha > 0$ *and* $\max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \partial B(\alpha|x,I)/\partial\alpha < \infty$.

ii. $B(\alpha|x,I)$ *is in* $\mathcal{P}_{s+1,s+2}$.

iii. $\alpha B(\alpha|x,I)$ *is in* $\mathcal{P}_{s+2,s+2}$ *and* $\max_{(\alpha,x)\in(0,1]\times\mathcal{X}} \left|\alpha B^{(s+2)}(\alpha|x,I)\right| < \infty$.

Proposition 3.1-(i) implies that the p.d.f $g(b|x,I) = 1/B^{(1)}(B^{-1}(b|x,I)|x,I)$ stays bounded away from 0 and infinity, a property which is important for inference with quantile estimation method. Proposition 3.1-(iii) indicates that the conditional bid function $B(\alpha|x,I)$ and its series coefficients $\beta_k(\alpha|I)$ have one more $\alpha-$derivatives than $V(\alpha|x,I)$ over $(0,1]$.[34] In the case of the quantile regression specification (3.2.7), $V(\alpha|x,I) = X'\gamma(\alpha|I)$ where $\gamma(\alpha|I)$ is $s+1$ times continuously differentiable over $[0,1]$, it implies that $B(\alpha|x,I) = X'\beta(\alpha|I)$ where $\beta(\alpha|I)$ is $s+2$ times continuously differentiable over $(0,1]$ but with a derivative $\beta^{(s+2)}(\alpha|I)$ which can diverge when α goes to 0.

3.3 Augmented quantile regression estimation

The identification Theorem 3.1 gives some guidance for estimating the conditional private value quantile function, which suggests to proceed through an estimation of the bid quantile function $B(\alpha|x,I)$ and its derivative $B^{(1)}(\alpha|x,I) = \partial B(\alpha|x,I)/\partial\alpha$. While there is an important literature on the estimation of a conditional quantile function, estimating $B^{(1)}(\alpha|x,I)$ has received much less attention. The function $B^{(1)}(\alpha|x,I)$ is also known as the sparsity function and appears in the expression of

[34] This is why it is assumed that $V(\alpha|x,I)$ is an element of $\mathcal{P}_{s+1,s+2}$ instead of $\mathcal{P}_{s+2,s+2}$ since it is important the brevity of exposition to have identical indices as in $\mathcal{P}_{s+2,s+2}$.

the asymptotic variance of many quantile estimators. Most of the results reviewed in Koenker (2005) focuses on consistent estimation of $B^{(1)}(\alpha|x,I)$ but do not establish consistency rate or asymptotic normality. Exceptions are Marmer and Shneyerov (2012) and Guerre and Sabbah (2012). However Marmer and Shneyerov (2012) estimator is based on standard kernel density estimation which is not flexible enough to deliver good consistency rates for quantile regression or additive specifications as in the preceding section. The local polynomial quantile approach of Guerre and Sabbah (2012) can be tailored to additive specification but it requests to employ a backfitting algorithm as in Hastie and Tibshirani (1986) which is not straightforward to implement. The estimation method of these papers do not pay attention to boundary bias. The method proposed here copes with boundaries and jointly estimate $B(\alpha|x,I)$ and its derivative $B^{(1)}(\alpha|x,I)$, combining local polynomial technique with a finite or infinite dimensional quantile regression specification.

Definition of the estimators

Consider L i.id. first-price auctions $(I_\ell, x_\ell, B_{i\ell}, i = 1, \ldots, I_\ell)$ and let $X_\ell = (1, x_\ell')'$. Let $\rho_\alpha(u)$ be the usual check function,

$$\rho_\alpha(q) = q(\alpha - \mathbb{I}(q \leq 0)),$$

where $\mathbb{I}(\cdot)$ is the indicator function, i.e $\mathbb{I}(q \leq 0) = 1$ when $q \leq 0$ and 0 otherwise. The intuition of the estimator is better understood at the population level. The conditional quantile function $B(\alpha|x,I)$ satisfies,

$$B(\alpha|\cdot, I) = \arg\min_{Q(x)} \mathbb{E}\left[\mathbb{I}(I_\ell = I)\,\rho_\alpha(B_{i\ell} - Q(x_\ell))\right],$$

where, except for the boundary case with $\alpha \in \{0,1\}$, the argument of the minimum is unique provided $\min_{(\alpha,x)\in[0,1]\times\mathcal{X}} B^{(1)}(\alpha|x,I) > 0$. To estimate $B^{(1)}(\alpha|x,I)$ can be done by introducing local variation of the quantile level in the vicinity of α. Let $K(\cdot)$ be a kernel function with support $[-1,1]$ and $h = h_L \to 0$ be a positive bandwidth parameter. Then it follows from the equation above that,

$$\begin{aligned}
&\{B(a|\cdot,I), a\in[\alpha-h,\alpha+h]\cap[0,1]\}\\
&\quad=\arg\min_{Q(a,x)}\int_0^1\mathbb{E}\left[\mathbb{I}(I_\ell=I)\,\rho_a(B_{i\ell}-Q(a,x_\ell))\right]\frac{1}{h}K\left(\frac{a-\alpha}{h}\right)da, \qquad (3.3.1)
\end{aligned}$$

where the minimization is performed over the set of functions $Q(a,x)$ which are continuous on $[\alpha-h,\alpha+h]\cap[0,1]\times\mathcal{X}$. Note that the argument of the minimum is now unique even when $\alpha=0$ or $\alpha=1$. However, performing a sample version of (3.3.1) is not simple in practice because the functional parameter $Q(a,x)$ lies in a very general set of continuous functions. Local polynomial methods circumvent this issue by performing an approximate minimization of the objective function in (3.3.1) by restricting to Taylor polynomial function instead of functions $Q(a,x)$.

How it works in practice is easily understood in the case of the standard quantile regression specification $V(\alpha|x,I)=X'\gamma(\alpha|I)$ in (3.2.7), in which case $B(\alpha|x,I)=X'\beta(\alpha|I)$ with $\beta(\alpha|I)$ as in (3.2.8). Applying Proposition 3.1, a Taylor expansion of $\beta(a|x,I)$ in the vicinity of $\alpha\in(0,1]$ gives,

$$\beta(a|I)=\beta(\alpha|I)+(a-\alpha)\beta^{(1)}(\alpha|I)+\cdots+\frac{(a-\alpha)^{s+1}}{(s+1)!}\beta^{(s+1)}(\alpha|I)+O\left((a-\alpha)^{s+2}\right).$$

So an approximate minimization of (3.3.1) up to a $O(h^{s+2})$ error uses the Taylor polynomial function, where β_j, $j=0,\ldots,s+1$ are some $(d+1)\times1$ vectors,

$$Q(a,x)=X'\left\{\beta_0+(a-\alpha)\beta_1+\cdots+(a-\alpha)^{s+1}\beta_{s+1}\right\}.$$

The corresponding minimizers $\beta_j^*(\alpha|I)$, $j=0,\ldots,s+1$ of the objective function in (3.3.1) are close to $\beta^{(j)}(\alpha|I)/j!$ up to a $O(h^{s+2-j})$ bias term, as suggested by Fan and Gijbels (1996) and established in the proof section. Note that (3.2.6) shows that it is in fact sufficient that such approximation results holds for $\beta_0^*(\alpha|I)$ and $\beta_1^*(\alpha|I)$. As well known from Fan and Gijbels (1996), it is important to include in the estimation the higher-order Taylor coefficients β_j, $j=2,\ldots,s+1$ to have $\beta_0^*(\alpha|I)=\beta(\alpha|I)+O(h^{s+2})$ and $\beta_1^*(\alpha|I)=\beta^{(1)}(\alpha|I)+O(h^{s+1})$.

Due to the additional smoothing of quantile levels and estimation of the derivatives of the slope coefficient, the corresponding estimator introduced now will be called the *Augmented Quantile Regression* (AQR) estimator. Let a be a real number and define, $\otimes$ being the Kronecker product,

$$\begin{aligned}\pi(a) &= \left[1, a, \ldots, a^{s+1}\right]', \\ X_\ell(a) &= X_\ell \otimes \pi(a) = \left[X_\ell', aX_\ell', \ldots, a^{s+1}X_\ell'\right]', \\ b &= \left[\beta_0', \beta_1', \beta_2', \ldots, \beta_{s+1}'\right]',\end{aligned}$$

where the β_j, $j = 0, \ldots, s+1$, are all $(d+1) \times 1$ so that $X_\ell(a)$ and b are both of dimension $(s+2) \times (d+1)$.[35] The Taylor polynomial function is $X_\ell(a)' b$. Stack the derivatives of the slope coefficients in,

$$b(\alpha|I) = \left(\beta(\alpha|I)', \beta^{(1)}(\alpha|I)', \ldots, (s+1)!\beta^{(s+1)}(\alpha|I)'\right)'. \tag{3.3.2}$$

The AQR estimator of $b(\alpha|I)$ is, up to a suitable rule to break ties,[36]

$$\widehat{b}_{AQR}(\alpha|I) = \arg\min_b \widehat{\mathcal{R}}_{AQR}(b; \alpha, I), \quad \alpha \in [0,1],$$

where, as suggested by (3.3.1), the objective function $\widehat{\mathcal{R}}_{AQR}(b; \alpha, I)$ is defined as

$$\begin{aligned}\widehat{\mathcal{R}}_{AQR}(b; \alpha, I) &= \frac{1}{LIh} \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_0^1 \rho_a\left(B_{i\ell} - X_\ell(a-\alpha)' b\right) K\left(\frac{a-\alpha}{h}\right) da \\ &= \frac{1}{LI} \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell} - X_\ell(ht)' b\right) K(t)\, dt.\end{aligned} \tag{3.3.3}$$

[35] As seen from Guerre and Sabbah (2012), any upper bound for the number of derivatives can be used instead of the exact one $s+1$, so that s needs not to be exactly known because using an upper bound is sufficient to achieve the convergence results stated later. This also holds for the ASQR estimator introduced below.

[36] An important feature of the objective function $\widehat{\mathcal{R}}_{AQR}(b; \alpha, I)$ is that it is still "bowl-shaped" at the boundaries $\alpha = 0, 1$, that is for on $\alpha = 0$, $\lim_{\|b\| \to \infty} \widehat{\mathcal{R}}_{AQR}(b; 0, I) = +\infty$, and that the same rule to break ties can be applied. This is due to smoothing, which implies that $\widehat{\mathcal{R}}_{AQR}(b; 0, I)$ diverges when b does, because all the $\widehat{\mathcal{R}}_{AQR}(b; a, I)$, $a > 0$ even close to 0, diverge. This contrasts with the standard quantile regression estimator. Indeed, since $\rho_0(q) = 0$ for all $q < 0$, the quantile regression objective function is 0 for all β such that $X_\ell{}'\beta \leq B_{i\ell}$ for all i and ℓ, so that it may be difficult to define the quantile regression estimator when $\alpha = 0$.

The corresponding AQR estimator of the private values quantile regression slope $\gamma(\alpha|I)$ and of the private value quantile regression model (3.2.7) are obtained by plugging in $\widehat{b}_{AQR}(\alpha|I)$ in (3.2.6),

$$\widehat{\gamma}_{AQR}(\alpha|I) = \widehat{\beta}^{(0)}_{AQR}(\alpha|I) + \frac{\alpha\widehat{\beta}^{(1)}_{AQR}(\alpha|I)}{I-1}, \quad \widehat{V}_{AQR}(\alpha|x,I) = X'\widehat{\gamma}_{AQR}(\alpha|I), \quad (3.3.4)$$

where $\widehat{\beta}^{(0)}_{AQR}(\alpha|I)$ and $\widehat{\beta}^{(1)}_{AQR}(\alpha|I)$ are the two first vectors of dimension $s+1$ stacked in $\widehat{b}_{AQR}(\alpha|I)$.[37]

The case of additive interactive quantile specification is a natural extension of the AQR procedure which involves a series truncation parameter $K_L \to \infty$ in addition to the bandwidth h. By Lemma 3.3, a series expansion $V(\alpha|x,I) = \sum_{k=1}^{\infty}\gamma_k(\alpha|I)P_k(x)$ for the private value conditional quantile function yields that the bid conditional quantile function can be similarly written as $B(\alpha|x,I) = \sum_{k=1}^{\infty}\beta_k(\alpha|I)P_k(x)$. Instead of estimating the full expansion, focusing on the first coefficients $\left\{\gamma_k(\alpha|I), \beta_k(\alpha|I), \beta_k^{(1)}(\alpha|I), k \le K_L\right\}$ gives a model which is very similar to the first-price auction quantile regression model considered above, up to a dimension K_L which is growing with the sample size.[38] Let

$$\begin{aligned} P(x) &= [P_1(x), \ldots, P_{K_L}(x)]', \\ P(x,a) &= P(x) \otimes \pi(a) = \left[P(x)', aP(x)', \ldots, a^{s+1}P(x)'\right]', \end{aligned}$$

and let b_L be a vector of dimension $(s+2)\times K_L$ which, as b above, stacks subvectors β_j associated with $a^j P(x_\ell, a)$, $j = 0, \ldots, s+1$. The *Augmented Sieve Quantile Regression* estimator of $V(\alpha|x,I)$ writes

$$\widehat{V}_{ASQR}(\alpha|x,I) = P(x)'\left\{\widehat{\beta}^{(0)}_{ASQR}(\alpha|I) + \frac{\alpha\widehat{\beta}^{(1)}_{ASQR}(\alpha|I)}{I-1}\right\}, \quad (3.3.5)$$

[37] Although not explicitely considered here for the sake of brevity, different bandwidths can be used for $\widehat{\beta}^{(0)}_{AQR}(\alpha|I)$ and $\widehat{\beta}^{(1)}_{AQR}(\alpha|I)$, and our theoretical results carry over for such extension under additional bandwidth rate conditions.

[38] In principle it would be possible to consider different series truncation parameters when estimating $B(\alpha|x,I)$ and its derivatives $B^{(1)}(\alpha|x,I)$, putting to 0 the missing $\beta_k(\alpha|I)$ or $\beta_k^{(1)}(\alpha|I)$ in the expression $\gamma_k(\alpha|I) = \beta_k(\alpha|I) + \alpha\beta_k^{(1)}(\alpha|I)/(I-1)$. It is not explicitely considered here for the sake of notation but such extension can be easily studied with the techniques developed here. As well, distinct bandwidths can be used for the estimation of $B(\alpha|x,I)$ and $B^{(1)}(\alpha|x,I)$.

where

$$\widehat{b}_{ASQR}(\alpha|I) = \left[\widehat{\beta}^{(0)}_{ASQR}(\alpha|I)', \widehat{\beta}^{(1)}_{ASQR}(\alpha|I)', \ldots, \widehat{\beta}^{(s+1)}_{ASQR}(\alpha|I)'\right]'$$

satisfies

$$\widehat{b}_{ASQR}(\alpha|I) = \arg\min_{b_L} \widehat{\mathcal{R}}_{ASQR}(b_L; \alpha, I),$$

$$\begin{aligned}\widehat{\mathcal{R}}_{ASQR}(b_L; \alpha, I) &= \frac{1}{LIh}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_0^1 \rho_a\left(B_{i\ell} - P(x_\ell, a-\alpha)' b_L\right) K\left(\frac{a-\alpha}{h}\right) da \\ &= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell} - P(x_\ell, ht)' b_L\right) K(t)\, dt.\end{aligned} \tag{3.3.6}$$

Theoretical estimation results

The main assumptions used in the theoretical results of the chapter are stated below. Assumptions A and S deal with the first-price auction model and the finite or infinite dimensional private value quantile regression specifications. Assumptions R and H concern the estimation method. Recall $a_L \asymp b_L$ 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 $\|e\| = (e'e)^{1/2}$ where the dimension of the column vector e can depend upon the sample size L. For Assumption H-(iii), Recall that the ℓ_1 norm of $e = (e_j)$ is $\sum_j |e_j|$.

Assumption A. *(i) The auction variables $\{I_\ell, x_\ell, V_{i\ell}, B_{i\ell}, i = 1, \ldots, I_\ell\}$ are independent and identically distributed. The common p.d.f $f(x|I)$ of the covariates x_ℓ given $I_\ell = I$ is continuous and bounded away from 0 over its support $\mathcal{X}$, which is a compact subset of $\mathbb{R}^d$, and the actual number of bidders I_ℓ takes values in a bounded set $\mathcal{I}$ of integer numbers larger or equal to 2.*

(ii) Given $(x_\ell, I_\ell) = (x, I)$, the $V_{i\ell}$, $i = 1, \ldots, I_\ell$ are independent and identically distributed with conditional quantile function $V(\alpha|x, I)$, p.d.f $f(v|x, I)$ and c.d.f $F(v|x, I)$ satisfying,

$$\inf_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x, I) > 0 \text{ and } \sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x, I) < \infty.$$

(iii) Given $(x_\ell, I_\ell) = (x, I)$, (3.2.5) holds, that is for all $\alpha \in [0, 1]$ and all

$(x, I) \in \mathcal{X} \times \mathcal{I}$,

$$B(\alpha|x, I) + \frac{\alpha B^{(1)}(\alpha|x, I)}{I-1} = V(\alpha|x, I), \qquad \text{with } B(0|x, I) = V(0|x, I).$$

Assumption S. *For $s \geq 1$ as in Assumption A and each $I \in \mathcal{I}$,*

(i) For the quantile regression model $V(\alpha|x, I) = X'\gamma(\alpha|I)$ as in (3.2.7), the slope coefficient $\gamma(\alpha|I) \in \mathbb{R}^{d+1}$ has $s+1$ bounded derivatives over $[0, 1]$.

(ii) For the sieve quantile regression model (3.2.9), there is a nonnegative number $d_{\mathcal{M}}$ such that $V(\alpha|x, I) \in \mathcal{P}_{s+1,s+2}$.

Assumption A recalls the implications of the IPV paradigm and imposes that bids are determined by the Bayesian Nash equilibrium. Assumption A-(i) assumes that each auctions are independent and identically distributed. The existence of a conditional p.d.f for the covariate x_ℓ is only used for the infinite dimensional quantile regression specification. For a standard quantile regression specification, it is sufficient to assume that the matrix $\mathbb{E}[\mathbb{I}(I_\ell = I) X_\ell X_\ell']$ has an inverse for all $I \in \mathcal{I}$ as recalled in Assumption R-(i) below. Note that, as all along this chapter, private values and number of bidders need not to be independent. A discussion of dependence in relation with an entry stage preliminary to the auction and unobserved heterogenity can be found in Marmer, Shneyerov and Xu (2013a) or Gimenes (2013), who also explains how to improve quantile regression estimation under independence by pooling. Assumption A-(ii) is standard regarding both auction models and quantile regression inference theory. The purpose of Assumption S is first to ensure that the estimated quantile regression specification is the correct one. Misspecified quantile regression specifications could also be considered as in Gimenes (2013) at the price of additional technicalities. Assumption S-(ii) has been discussed in detail prior Proposition 3.1 and concerns both the sieve method together with the smoothness of the conditional private value quantile function.

Assumption R. *(i) The matrix $\mathbb{E}[\mathbb{I}(I_\ell = I) X_\ell X_\ell']$ is full rank for each $I \in \mathcal{I}$.*

(ii) The sieve $\{P_k, k \geq 1\}$ *has the Riesz property, that is there is a constant* $C \geq 1$ *such that for any sequence* $\{\gamma_k, k \geq 1\}$ *of real numbers satisfying* $\sum_{k=1}^{\infty} \gamma_k^2 < \infty$,

$$\frac{1}{C}\left\{\sum_{k=1}^{\infty}\gamma_k^2\right\} \leq \int_{\mathcal{X}}\left(\sum_{k=1}^{\infty}\gamma_k P_k(x)\right)^2 dx \leq C\left\{\sum_{k=1}^{\infty}\gamma_k^2\right\}.$$

Assumption H. *(i) The kernel function* $K(\cdot)$ *is continuous and strictly positive over* $(-1,1)$ *and* $h \to 0$ *with* $\log^3 L/(Lh^2) = o(1)$.

(ii) In addition, for the ASQR estimator, $K_L \asymp h^{-d_{\mathcal{M}}}$, *and some* $\zeta > 0$ *such* $(1+\zeta)\, d_{\mathcal{M}} \leq 2s$,

$$\max_{x \in \mathcal{X}} \|P(x)\| = O\left(K_L^{\frac{1+\zeta}{2}}\right),$$

with $\log^3 L/\left(Lh^{(2+\zeta)d_{\mathcal{M}}+2}\right) = O(1)$.

(iii) $\max_{x\in\mathcal{X}} \|P(x)\| = O\left(K_L^{1/2}\right)$ *and* $\log L/\left(Lh^{4d_{\mathcal{M}}+2}\right) = O(1)$. *The* ℓ_1 *norm of each column of* $\mathbb{E}^{-1}\left[\mathbb{I}(I_\ell = I)\, P(x_\ell)\, P(x_\ell)'\right]$ *is bounded by the same constant* C *for each* L *and* I,

$$\max_{k \leq K_L}\left\{\int_{\mathcal{X}} |P_k(x)|\, dx\right\} \times \max_{x\in\mathcal{X}} \sum_{k=1}^{K_L} |P_k(x)| = O(1), \tag{3.3.7}$$

and for some $p \in (0,1]$, $\overline{K}_{1L}$ *with* $\log \overline{K}_{1L} = O(\log L)$,

$$\|P(x) - P(x')\| \leq \overline{K}_{1L}\, \|x - x'\|^p \ \textit{for all } x, x' \textit{of} \mathcal{X}.$$

Assumption R is a key identification condition which ensures that the coefficients of the private value and bid quantile regression are uniquely defined, in the finite dimensional case for (i) and in the infinite dimensional case for (ii). Assumption R-(ii) and Assumption A-(i) ensures that the eigenvalues of $\mathbb{E}\left[\mathbb{I}(I_\ell = I)\, P(x_\ell)\, P(x_\ell)'\right]$ stays bounded away from 0 and infinity when L increases, which is a key property to ensure that the ASQR estimator has a well-behaved variance. Assumption H-(i,ii) groups some growth conditions for the bandwidth h and series truncation parameter K_L. Part (i) is specific to the standard quantile regression specification

(3.2.7) and the AQR procedure. The rate condition $\log^3 L/\left(Lh^2\right) = o\left(1\right)$ is stronger than the condition $\log L/\left(Lh\right) = o\left(1\right)$ used in Guerre and Sabbah (2012) for local polynomial conditional estimation in a univariate context. This new condition rate is driven by the order of the remainder term in the Bahadur representation and is specific to the augmented procedure, i.e smoothing with respect to quantile level. Part (ii) considers infinite dimensional quantile regression specification and the ASQR procedure. It is possible to obtain general results without $K_L \asymp h^{-d_{\mathcal{M}}}$ but this constraint simplifies the discussion on the sieve truncation K_L and smoothing parameter h. The effective dimension $d_{\mathcal{M}}$ is the number of interactions in the infinite dimensional quantile regression specification (3.2.9) from $\mathcal{M}$. The condition $K_L \asymp h^{-d_{\mathcal{M}}}$ recalls that K_L plays the role of a smoothing parameter associated with a covariate with dimension $d_{\mathcal{M}}$. Setting the exponent ζ to 0 shows that the bandwidth should now satisfies the rate condition $\log^3 L/\left(Lh^{2(d_{\mathcal{M}}+1)}\right) = O\left(1\right)$ which is identical to the one in Assumption H-(i), up to the fact that the effective dimension $d_{\mathcal{M}}$ now appears while it can be set 0 for a standard quantile regression model. The exponent ζ is typical of the sieve $\{P_k\left(\cdot\right), k \geq 1\}$. As reviewed in Chen (2007), $\zeta = 0$ for the (multivariate) trigonometric system, B-spline and wavelets, while $\zeta = 1/2$ for Legendre polynomial functions. While Assumption H-(ii) is used for integrated mean squared error, Assumption H-(iii) is used to derive optimal consistency rates with respect to uniform norm. Such results are difficult to obtain because the ASQR estimator is not constrained to be bounded or smooth, so that it is sometimes necessary to rely on a crude Cauchy-Schwarz inequality bound $\left|P\left(x\right)' e\right| \leq \left\|P\left(x\right)\right\| \left\|e\right\|$ to bound some functional error term $P\left(x\right)' e$, see e.g. Newey (1997). [39] Assumption H-(iii) strengthens the bandwidth rate condition of (ii) to $\log L/\left(Lh^{2(2d_{\mathcal{M}}+1)}\right) = O(1)$. The condition (3.3.7) is used to study a bias term. It greatly restricts the sieve which cannot be the trigonometric or Legendre polynomial functions but holds for B-splines and wavelet functions. To see this, observe that, in the univariate case or of a purely additive model, father wavelet with a compact

[39] Imposing smoothness conditions may help to obtain better bounds. This is the purpose of penalized methods as in Portnoy (1997), Eggermont and LaRiccia (2009) and as also reviewed in Chen (2007). However, these methods consider roughness penalties which depend upon the support $\mathcal{X}$ of the covariates, which needs to be known. Unconstrained methods do not require such a knowledge and may be more appropriate when the support is unknown or can only be poorly estimated.

support or B-spline sieve take the form,

$$P_k(x) = C^{1/2} 2^{K_n/2} \psi\left(C 2^{K_N} x - k\right), \quad x \in \mathbb{R},$$

for some scaling constant C and integrable function $\psi(\cdot)$ with a compact support, so that $\max_{k \leq K_L} \left\{\int_{\mathcal{X}} |P_k(x)|\, dx\right\} = O\left(2^{-K_n/2}\right)$ and $\max_{x \in \mathcal{X}} \sum_{k=1}^{K_L} |P_k(x)| = O\left(2^{K_n/2}\right)$ which gives (3.3.7). The Hölder condition in Assumption H-(iii) is flexible enough to allow for B-splines, and more specifically wavelets which are not always differentiable, see Daubechies (1992). The ℓ_1 condition follows from the property of Gram band matrices.[40]

The two next results states the uniform convergence and asymptotic normality of the AQR procedure when (3.2.7) is correctly specified.

Theorem 3.2. *Suppose the quantile regression specification (3.2.7) is correct. Then under Assumptions A, S-(i) and H-(i),*

$$\sup_{(\alpha,x,I) \in [0,1] \times \mathcal{X} \times \mathcal{I}} \left|\widehat{V}_{AQR}(\alpha|x,I) - V(\alpha|x,I)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh}\right)^{1/2} + h^{s+1}\right).$$

Theorem 3.2 shows that uniform convergence holds uniformly over $[0,1] \times \mathcal{X} \times \mathcal{I}$ so that the AQR procedure is free from asymptotic boundary bias issues. As for all

[40] That splines can achieve optimal uniform consistency rates without a penalization as established below was known since Portnoy (1997) for the univariate case and a single quantile level, and Assumption H-(iii) allows to extend this in particular to the multivariate case and to wavelets, uniform convergence with respect to α of the conditional quantile function and its α-derivative. For localized sieve as the B-splines and compact wavelets considered here, the ℓ_1 condition holds because the entries of $\mathbb{E}\left[\mathbb{I}(I_\ell = I) P(x_\ell) P(x_\ell)'\right]$ are bounded by the ones of $C(\mathrm{Id} + \varrho(J + J'))$ where J is a lower band matrix with an order independent of L and $0 < \varrho < 1$. Then the entries p_{ij} of $\mathbb{E}^{-1}\left[\mathbb{I}(I_\ell = I) P(x_\ell) P(x_\ell)'\right]$ are bounded by the ones of

$$C'\left(\mathrm{Id} + \sum_{n=1}^{\infty} \varrho^n \left(J^n + J^{n'}\right)\right).$$

This implies that $|p_{ij}| \leq C'' \varrho^{|i-j|}$ uniformly in i, j and L, which implies that $\max_L \max_j \sum_i |p_{ij}| < \infty$.

the rates stated here, the rate of Theorem 3.2 is determined by the rate of,

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\frac{\alpha\widehat{\beta}^{(1)}_{AQR}(\alpha|I)}{I-1}-\frac{\alpha\beta^{(1)}(\alpha|I)}{I-1}\right|,$$

since $\widehat{\beta}_{AQR}(\alpha|I)$ converges faster than $\alpha\widehat{\beta}^{(1)}_{AQR}(\alpha|I)$. The optimal bandwidth order is $h \asymp (\log L/L)^{1/(2s+3)}$ which is compatible with Assumption H-(i) for all $s \geq 1$. For such choice of the bandwidth,

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{V}_{AQR}(\alpha|x,I)-V(\alpha|x,I)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{\frac{s+1}{2s+3}}\right),$$

which corresponds to the nonparametric optimal minimax rate in GPV for estimating the cumulative distribution function of the private values in the absence of covariate.[41] For instance assuming that $s = 1$ gives a rate $(\log L/L)^{2/5}$ which is often considered as acceptable for sample size as large as 100.

Theorem 3.2 investigates global convergence and its consistency rate can be due to some specific α or x. The next pointwise asymptotic normality result implies that it is not the case. Let us first introduce some additional notations. Let s_1 be the $1\times(s+2)$ vector $(0,1,0,\ldots,0)$ which is such that $\mathrm{Id}_{d+1}\otimes s_1\widehat{\beta}_{AQR}(\alpha|I) = \widehat{\beta}^{(1)}_{AQR}(\alpha|I)$. Let $\Pi^t_h(\alpha)$ be the second column of the inverse of $\int_{-\alpha/h}^{(1-\alpha)/h}\pi(t)\pi(t)'K(t)\,dt$, i.e,

$$\Pi^t_h(\alpha) = \left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\pi(t)\pi(t)'K(t)\,dt\right)^{-1}s_1',$$

and define the AQR variance quantities,

$$v^2_h(\alpha) = \Pi^t_h(\alpha)'\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\pi(t_1)\pi(t_2)'\min(t_1,t_2)K(t_1)K(t_2)\,dt_1dt_2\Pi^t_h(\alpha),$$

[41] Although GPV focus on the p.d.f for which the optimal minimax rate for uniform convergence is $(\log L/L)^{s/(2s+d+3)}$, their results can easily be extended to the case of the c.d.f for which the minimax rate becomes $(\log L/L)^{(s+1)/(2s+d+3)}$. A simple inversion argument shows that it is also the optimal minimax uniform consistency rate when estimating the conditional quantile function.

$$\Sigma_h(\alpha|I) = \frac{\alpha^2 v_h^2(\alpha)}{(I-1)^2 I}\mathbb{E}^{-1}\left[\frac{X_\ell X_\ell' \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]\mathbb{E}\left[X_\ell X_\ell' \mathbb{I}(I_\ell = I)\right]\mathbb{E}^{-1}\left[\frac{X_\ell X_\ell' \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]. \tag{3.3.8}$$

The bias of the AQR estimator will depend upon,

$$\begin{aligned}\mathsf{Bias}_h(\alpha|I) &= \frac{1}{I-1}s_1\left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi(t)\pi(t)' K(t)\,dt\right)^{-1}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \frac{t^{s+2}\pi(t)}{(s+2)!}K(t)\,dt \\ &\times \mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)X_\ell X_\ell'}{B^{(1)}(\alpha|x_\ell, I)}\right]\mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)X_\ell \alpha B^{(s+2)}(\alpha|x_\ell, I)}{B^{(1)}(\alpha|x_\ell, I)}\right]. \end{aligned}\tag{3.3.9}$$

Theorem 3.3 states the asymptotic normality of $\widehat{\gamma}_{AQR}(\alpha|I)$ and $\widehat{V}_{AQR}(\alpha|x, I)$.

Theorem 3.3. *Suppose the quantile regression specification (3.2.7) is correct. Then under Assumptions A, S-(i) and H-(i) and for all $\alpha \in [0,1]$, all $X = (1, x')'$ with $x \in \mathcal{X}$,*

$$\left(\frac{Lh}{X'\Sigma_h(\alpha|I)X}\right)^{1/2}\left(\widehat{V}_{AQR}(\alpha|x, I) - V(\alpha|x, I) - h^{s+1}X'\mathsf{Bias}_h(\alpha|I) + o\left(h^{s+1}\right)\right) \xrightarrow{d} \mathcal{N}(0,1).$$

Theorem 3.3 implies that

$$\widehat{V}_{AQR}(\alpha|x, I) - V(\alpha|x, I) = O_{\mathbb{P}}\left(\frac{1}{(Lh)^{1/2}} + h^{s+1}\right),$$

and that, except for $\alpha = 0$, this order cannot be improved since $\Sigma(\alpha|I)$ is a full-rank covariance matrix for $\alpha \in (0,1]$. It may however happen that $X'\mathsf{Bias}_h(\alpha|I) = 0$ in which case the order above can be changed into $1/(Lh)^{1/2}$. The only possibility of rate improvement occurs at $\alpha = 0$. The intuition is that the rate of Theorem 3.3 is given by the estimator component $\alpha\widehat{\beta}^{(1)}_{AQR}(\alpha|I)$ which converges with a slower rate than $\widehat{\beta}_{AQR}(\alpha|I)$, and that the former vanishes for $\alpha = 0$. In fact, Chernozhukov (2005) has shown that standard quantile regression estimators can converge faster than $L^{1/2}$ when α goes to 0 or 1. This is also likely to hold here for $\alpha = 0$, but not for $\alpha = 1$ because the asymptotic behavior of $\widehat{\beta}^{(1)}_{AQR}(\alpha|I)$ fundamentally differs from the one of $\widehat{\beta}_{AQR}(\alpha|I)$ near the boundary. Indeed, the estimator $\widehat{\beta}_{AQR}(\alpha|I)$

is close to a vector weighted average of $\left(\mathbb{I}\left(B_{i\ell} \leq B\left(\alpha|x_\ell, I_\ell\right)\right) - \alpha\right)\mathbb{I}\left(I_\ell = I\right)$ with a variance proportional $\alpha\left(1-\alpha\right)$ which goes to 0 when α goes to 0 or 1. By contrast, $\widehat{\beta}^{(1)}_{AQR}\left(\alpha|I\right)$ is somehow similar to the growth rate $\left(\widehat{\beta}_{AQR}\left(\alpha|I\right) - \widehat{\beta}_{AQR}\left(\alpha - h|I\right)\right)/h$ that is to a vector weighted average of variables as,

$$\frac{1}{h}\left(\mathbb{I}\left(B\left(\alpha - h|x_\ell, I_\ell\right) \leq B_{i\ell} \leq B\left(\alpha|x_\ell, I_\ell\right)\right) - h\right)\mathbb{I}\left(I_\ell = I\right),$$

which variance is asymptotically proportional to $1/h$ independently of the quantile level α. This heuristic argument also suggests that $\widehat{V}_{AQR}\left(\alpha_1|x_1, I_1\right), \ldots, \widehat{V}_{AQR}\left(\alpha_J|x_J, I_J\right)$ are asymptotically independent as soon as the α_j, $j = 1, \ldots, J$ are distinct.

Theorem 3.3 can also be used to propose plug in bandwidth choice. Indeed, for α far enough the boundaries so that $\mathsf{Bias}_h(\alpha|I)$ and $\Sigma_h\left(\alpha|I\right)$ do not depend upon h, an optimal bandwidth with respect to the asymptotic MSE of $\widehat{V}_{AQR}\left(\alpha|x, I\right)$ is, provided that $X'\mathsf{Bias}_h(\alpha|I)$ differs from 0,

$$h_{opt} = \left(\frac{X'\Sigma_h\left(\alpha|I\right)X}{2\left(s+1\right)X'\mathsf{Bias}_h(\alpha|I)} \times \frac{1}{L}\right)^{\frac{1}{2s+3}}.$$

Such a bandwidth can easily be estimated using standard pluging techniques. Second order improvement can also be obtained by including the effect of $\widehat{B}_{AQR}\left(\alpha|x, I\right)$ in the asymptotic distribution above.

Although not detailed here, a similar asymptotic normality result also holds for the ASQR estimator, taking into account an additional bias term due to sieve truncation and changing X_ℓ int $P\left(x_\ell\right)$ in the expression of the asymptotic variance (3.3.8). The next theorem describes the global integrated mean square error (IMSE) and uniform convergence rates.

Theorem 3.4. *Suppose the sieve quantile regression specification (3.2.9) is correct. Then under Assumptions A, S-(ii) and H-(i,ii) and for all $I \in \mathcal{I}$,*

$$\int_0^1 \left\{\int_{\mathcal{X}} \left(\widehat{V}_{ASQR}\left(\alpha|x, I\right) - V\left(\alpha|x, I\right)\right)^2 dx\right\} d\alpha = O_{\mathbb{P}}\left(\frac{1}{\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}} + h^{s+1}\right).$$

If Assumption H-(iii) also holds,

$$\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|\widehat{V}_{AQR}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}+h^{s+1}\right).$$

Theorem 3.4 is qualitatively similar to Theorem 3.2, up to the number of interactions $d_{\mathcal{M}}$ of the infinite dimensional additive quantile regression specification (3.2.9). The optimal convergence rates are

$$\left(\frac{1}{L}\right)^{\frac{s+1}{2s+d_{\mathcal{M}}+3}} \text{ for IMSE,} \quad \left(\frac{\log L}{L}\right)^{\frac{s+1}{2s+d_{\mathcal{M}}+3}} \text{ for uniform convergence.}$$

However these rates are not always feasible due to the bandwidth constraints in Assumption H-(ii,iii). The optimal order of the bandwidth is now

$$\left(\frac{1}{L}\right)^{\frac{1}{2s+d_{\mathcal{M}}+3}} \text{ for IMSE,} \quad \left(\frac{\log L}{L}\right)^{\frac{1}{2s+d_{\mathcal{M}}+3}} \text{ for uniform convergence.}$$

These optimal bandwidth rates satisfy Assumption H-(ii,iii) provided,

$$\frac{d_{\mathcal{M}}-1}{2} < s \text{ for IMSE,} \quad \frac{3\left(d_{\mathcal{M}}-1\right)}{2} \leq s \text{ for uniform convergence,}$$

which both holds for $s = 1$ and the infinite dimensional additive quantile regression specification for which $d_{\mathcal{M}} = 1$. The IMSE restriction holds for $s = 1$ and two interactions but not above $d_{\mathcal{M}} \geq 3$. This suggests that the ASQR procedure works well for low dimensional models. For higher number of interactions, the two step procedure of Horowitz and Lee (2005) or the augmented local polynomial quantile estimator of Guerre and Vuong (2013) with backfitting when $d_{\mathcal{M}} < d$ can be useful.

3.4 Extensions, discussions and simulations

Some potentially important extensions deal with imposing no interaction between the rank and the covariates, as in the quantile specifications,

$$
\begin{aligned}
V(\alpha|x, I) &= \gamma_0(\alpha) + x'\gamma_1, \\
V(\alpha|x, I) &= V_0(\alpha) + \sum_{j=1}^{d} V_j(x_j).
\end{aligned}
$$

The former implies that, for some η independent of (x, I), $V = x'\gamma_1 + \eta$ as considered in Rezende (2008) while the latter gives the additive regression model with independent and homoskedastic error $V = \sum_{j=1}^{d} V_j(x_j) + \eta$. Such specifications are important because the components $x'\gamma_1$ of the former and $\sum_{j=1}^{d} V_j(x_j)$ of the latter can be identified even when the number of bidders is not observed. Gimenes (2013) has considered the former quantile regression specification for ascending auctions, as well as the case where the slope coefficient $\gamma(\alpha|I)$ in (3.2.7) does not depend upon I. Her estimation methods can easily be adapted to the first-price auction and sieve expansions. Gimenes (2013) considers also some specification tests for dependence on I and higher order interactions that can be extended to a first-price auction nonparametric setup. Other extensions considered now deal with optimal bidding strategy and private values estimations, the case where only the winning bid is observed and reservation price. The last discussion concerns the choice of the smoothing parameter.

Optimal bidding strategy estimation

When the bidder knows his private value rank u, equation (3.2.4) in Lemma 3.1 shows that the optimal bid is given by $B(u|x, I)$ which can be estimated using $\widehat{B}(u|x, I) = \widehat{B}_{AQR}(u|x, I)$ or $\widehat{B}_{ASQR}(u|x, I)$ depending on the specification of interest. Under the conditions of Theorems 3.2 and 3.4-(ii), it is shown in the proof section that

$$
\max_{(u,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|\widehat{B}(u|x, I) - B(u|x, I)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2} + h^{s+2}\right),
$$

recalling that $d_{\mathcal{M}}$ is the number of interactions in the model with the convention that $d_{\mathcal{M}} = 0$ for a quantile regression. This gives an optimal uniform convergence rate $(\log L/L)^{(s+2)/(2(s+2)+d_{\mathcal{M}})}$ which is faster than the ones derived for the estimation of $V(\alpha|x, I)$ in Theorems 3.2 and 3.4-(ii). For a quantile regression specification with $d_{\mathcal{M}} = 0$, it can be shown that the uniform consistency rate of $\widehat{B}(u|x, I)$ is the parametric rate $1/L^{1/2}$.

When the bidder does not know his rank u, it can be estimated from his private value $v = V(u|x, I)$. For $\widehat{V}(u|x, I) = \widehat{V}_{AQR}(u|x, I)$ or $\widehat{V}_{ASQR}(u|x, I)$, a possible estimator of u is,

$$\widehat{u} = \arg \min_{\alpha \in [0,1]} \left| v - \widehat{V}(\alpha|x, I) \right|,$$

using an appropriate convention to break ties. The estimation of the optimal bid is then $\widehat{B}(\widehat{u}|x, I)$ and the next Lemma describes its uniform consistency rate.

Lemma 3.4. *Assume that Assumption A holds. Let $\widehat{\varrho}$ be the maximum of*

$$\max_{(\alpha,x,I) \in [0,1] \times \mathcal{X} \times \mathcal{I}} \left| \widehat{B}(\alpha|x, I) - B(\alpha|x, I) \right| \text{ and } \max_{(\alpha,x,I) \in [0,1] \times \mathcal{X} \times \mathcal{I}} \left| \widehat{V}(\alpha|x, I) - V(\alpha|x, I) \right|.$$

Then,

$$\max_{(x,I) \in \mathcal{X}} \max_{v \in [V(0|x,I), V(1|x,I)]} \left| \widehat{B}(\widehat{u}|x, I) - B(F(v|x, I)|x, I) \right| = O(\widehat{\varrho}).$$

Under the conditions of Theorems 3.2 and 3.4-(ii), $\widehat{V}(\cdot|\cdot, \cdot)$ has a slower uniform convergence rate than $\widehat{B}(\cdot|\cdot, \cdot)$. It follows that the uniform convergence rate of $\widehat{B}(\widehat{u}|x, I)$ is $(\log L/L)^{\frac{s+1}{2s+d_{\mathcal{M}}+3}}$ at best for the AQR and ASQR estimators.

Private values estimation

As noted by Marmer and Shneyerov (2012), a quantile approach avoids estimating the private values, a step which was a key element of the two step procedure of GPV. However, as shown by Cassola, Hortaçsu and Kastl (2013) in the related context of Treasury auctions, the private values or some equivalent quantities may have a strong economic content especially useful in empirical applications. The quantile regression approach allows to estimate the private values $V_{i\ell}$ from an estimation of the conditional bid and private values quantile function as delivered by the AQR or the ASQR procedure, say $\widehat{B}(\alpha|x,I)$ and $\widehat{V}(\alpha|x,I)$. A key fact from Lemma 3.1 is that the private value and bid ranks of bidder i are identical. In the case of the private value $V_{i\ell}$ of bidder i in auction ℓ, the rank $U_{i\ell}$ can be estimated from (3.2.4) using,

$$\widehat{U}_{i\ell} = \arg\min_{\alpha\in[0,1]} \left| B_{i\ell} - \widehat{B}(\alpha|x_\ell, I_\ell) \right|,$$

using an appropriate convention to break ties. Then (3.2.1) suggests the estimated private values,

$$\widehat{V}_{i\ell} = \widehat{V}\left(\widehat{U}_{i\ell}|x_\ell, I_\ell\right).$$

The next Lemma gives the convergence rate of $\widehat{V}_{i\ell}$.

Lemma 3.5. *Assume that Assumption A holds. Let $\widehat{\varrho}$ be the maximum of*

$$\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left| \widehat{B}(\alpha|x,I) - B(\alpha|x,I) \right| \quad and \quad \max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left| \widehat{V}(\alpha|x,I) - V(\alpha|x,I) \right|.$$

Then,

$$\max_{\ell=1,\ldots,L} \max_{i=1,\ldots,I_\ell} \left| \widehat{V}_{i\ell} - V_{i\ell} \right| = O(\widehat{\varrho}).$$

For the AQR and ASQR procedures, the rate $\widehat{\varrho}$ is the uniform convergence rate $\widehat{V}(\alpha|x,I)$, which is respectively given by Theorems 3.2 and 3.4, since the uniform convergence rate $\widehat{V}(\alpha|x,I)$ is the one of $\alpha\widehat{B}^{(1)}(\alpha|x,I)$ which is slower than uniform convergence rate of $\widehat{B}(\alpha|x,I)$. Compared to GPV, the proposed private value estimation procedure is free of boundary issues so that all the private values can be recovered asymptotically. The intermediary results used to establish these two theorems can be used to obtain a more precise uniform stochastic expansion for $\widehat{V}_{i\ell} - V_{i\ell}$. Such expansions can be useful to study estimators based on $\widehat{V}_{i\ell}$ as in the two step estimation studied in Newey and McFadden (1994). Indeed, using the rate of Lemma 3.5 may give too conservative rates of convergence for such procedure.

Winning bids

It has been considered so far that all the bids of the auction were observed but the quantile regression approach easily extends to the case where only the winning bid is observed. The next Proposition shows that the conditional private value quantile function can be identified from the one of the winning bid. This follows from Theorem 3.1 and the fact that the conditional bid quantile function $B(\cdot|x,I)$ is easily recovered from the winning bid one $B_{\max}(\cdot|x,I)$ since $B_{\max}(\alpha|x,I) = B\left(\alpha^{1/I}|x,I\right)$

Proposition 3.2. *Assume that Assumption A holds. Then*

i. The conditional equilibrium quantile function $B_{\max}(\cdot|x,I)$ of the winning bid satisfies

$$B_{\max}(\alpha|x,I) = B\left(\alpha^{1/I}|x,I\right) \text{ with } B(\alpha|x,I) = \frac{I-1}{\alpha^{I-1}} \int_0^{\alpha} t^{I-2} V(t|x,I)\,dt,$$

and $B_{\max}\left(\alpha^I|x,I\right)$ is continuously differentiable over $[0,1]$.

ii. Let $B_{\max}(\alpha|x,I)$ be as in (i). Then it must hold that,

$$V(\alpha|x,I) = B_{\max}\left(\alpha^I|x,I\right) + \frac{\alpha}{I-1}\frac{\partial}{\partial\alpha}\left[B_{\max}\left(\alpha^I|x,I\right)\right].$$

Proposition 3.2-(i) implies that linear quantile regression specification enjoys the same stability property than discussed after Theorem 3.1. Proposition 3.2-(ii) indicates how to implement the AQR or ASQR procedure to the winning bid. The key difference with the case where all the bids are observed is that the quantile α^I should be used instead of α in the estimation procedure. It follows that Theorems 3.2-3.4 can easily be modified to cover this observational scheme. The estimation procedure can be modified in a similar way to cover the case where any prescribed subset of the ordered bids is observed.

Reservation price

A binding reservation price changes the bidders' optimal strategy and introduces censorship. Let $r = r(x,I)$ be the reservation price chosen by the seller. The associated screening level is the rank $\alpha_r = \alpha_r(x,I)$ solving $V(\alpha|x,I) = r(x,I)$, or equivalently $\alpha_r(x,I) = F(r(x,I)|x,I)$. Bidders ranked below $\alpha_r(x,I)$ will not bid and a bidder with rank $\alpha_r(x,I)$ will bid $r(x,I)$. Consequently the optimal bidding strategy $B_r(\alpha|x,I)$ solves (3.2.6) for $\alpha \geq \alpha_r(x,I)$ with the initial condition $B_r(\alpha_r(x,I)|x,I) = r(x,I)$. Solving gives, for all $\alpha \geq \alpha_r(x,I)$,

$$\begin{aligned} B_r(\alpha|x,I) &= \frac{1}{\alpha^{I-1}}\left\{ r(x,I)\,\alpha_r(x,I)^{I-1} + (I-1)\int_{\alpha_r(x,I)}^{\alpha} t^{I-2}V(t|x,I)\,dt \right\} \\ &= \left(\frac{\alpha_r(x,I)}{\alpha}\right)^{I-1}\left(r(x,I) - B(\alpha_r(x,I)|x,I)\right) + B(\alpha|x,I), \end{aligned}$$

where $B(\cdot|x,I)$ is the optimal bidding strategy in the absence of reservation price defined in (3.2.5). Since $B(\alpha_r(x,I)|x,I) < r(x,I)$, $B_r(\alpha|x,I) - B(\alpha|x,I) > 0$ and decreases with α.

Three econometric issues must be addressed here. First, (3.2.6) together with the new constraint $B_r(\alpha_r(x,I)|x,I) = r(x,I)$ implies that $B_r^{(1)}(\alpha_r(x,I)|x,I) = 0$, so that quantile estimation may have non standard asymptotic properties. Second, the fact that $B_r(\alpha|x,I)$ depends upon $r(x,I)$ and $\alpha_r(x,I)$ implies that quantile linear specifications are not as stable as for the case where the reservation price is nonbinding. Third some bids are not observed. To address the first issue can be done as in GPV by considering the transformed observations $B_{\dagger i} = (B_i - r(x,I))^{1/2}$, with conditional quantile function,

$$B_{\dagger r}(\alpha|x,I) = (B_r(\alpha|x,I) - r(x,I))^{1/2}, \; \alpha > \alpha_r(x,I).$$

In order to extend this function for $\alpha \leq \alpha_r(x,I)$, set $B_{\dagger i} = 0$ for all those buyers who did not bid, so that the model is somehow analogous to the censored quantile regression in Powell (1986).

It is assumed here that $r(x,I)$, x and the number of potential bidders I are observed, together with the number I^* of actual bidders and their bids B_i. The next proposition is a key result to design an estimation strategy when the reservation price is binding.

Proposition 3.3. *Assume that Assumption A and that the reservation price is now binding, $r(x,I) \in (V(0|x,I), V(1|x,I))$. Then given x and I,*

i. The number of observed bids I^* has a binomial distribution with parameter $(I, 1-\alpha_r(x,I))$ with $\alpha_r(x,I) = 1 - \mathbb{E}[I^*|x,I]/I$.

ii. The equilibrium conditional quantile function $B_{\dagger r}(\alpha|x,I)$ satisfies, for all $\alpha > \alpha_r(x,I)$,

$$B_{\dagger r}(\alpha|x,I) = \left(r(x,I) \left(\left(\frac{\alpha_r(x,I)}{\alpha} \right)^{I-1} - 1 \right) + \frac{I-1}{\alpha^{I-1}} \int_{\alpha_r(x,I)}^{\alpha} t^{I-2} V(t|x,I)\, dt \right)^{1/2}.$$

iii. The private value conditional quantile function is given by, for all $\alpha > \alpha_r(x,I)$,

$$V(\alpha|x,I) = r(x,I) + B_{\dagger r}^2(\alpha|x,I) + \frac{\alpha}{I-1} \frac{\partial}{\partial\alpha} \left[B_{\dagger r}^2(\alpha|x,I) \right].$$

Proposition 3.3 shows that identification and estimation of the model are straightforward when $\alpha_r(x,I) = \alpha_r(I)$, including the case $\alpha_r(x,I) = \alpha_r$ where the reservation price is given by the α_rth quantile of the private value conditional distribution. Part (i) implies that $\alpha_r(I)$ can be estimated using $\widehat{\alpha}_r(I) = 1 - \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) I_\ell^* / \left(I \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) \right)$ which is $L^{1/2}$-consistent. Assuming that the quantile regression specification $V(\alpha|x,I) = X'\gamma(\alpha|I)$ holds, part (ii) implies that

$$B_{\dagger r}(\alpha|x,I) = \left(r(x,I) \left(\left(\frac{\alpha_r(I)}{\alpha} \right)^{I-1} - 1 \right) + X'\beta(\alpha|I) \right)^{1/2},$$

with

$$\beta(\alpha|I) = \frac{I-1}{\alpha^{I-1}} \int_{\alpha_r(I)}^{\alpha} t^{I-2} \gamma(t|I)\, dt \text{ for all } \alpha > \alpha_r(I). \tag{3.4.1}$$

Since $r(x,I)$ is known, $r(x,I) \left((\alpha_r(I)/\alpha)^{I-1} - 1 \right)$ can be estimating leading to consider the estimated quantile model,

$$\widehat{Q}_\beta(\alpha, x, I) = \left(r(x,I) \left(\left(\frac{\widehat{\alpha}_r(I)}{\alpha} \right)^{I-1} - 1 \right) + X'\beta(\alpha|I) \right)^{1/2}.$$

Combining the augmented methodology, nonlinear quantile regression and Powell (1986) to account for "censored" bids[42] will deliver estimation of $\beta(\alpha|I)$ and

[42] Alternatively it is possible to restrict to the observed bids and to change the quantile functions into quantile functions given that the $B_{i\ell}$ are such that $V_{i\ell} \geq r(x_\ell, I_\ell)$ to account for selection.

$\beta^{(1)}(\alpha|I)$ which can be used as in (3.3.4) to estimate $\gamma(\alpha|I)$ for all $\alpha \geq \alpha_r(I)$. Indeed, (3.4.1) gives again that (3.3.4) holds, leading to an estimator $\widehat{\gamma}(\alpha|I) = \widehat{\beta}(\alpha|I) + \alpha\widehat{\beta}^{(1)}(\alpha|I)/(I-1)$.[43]

The case where $\alpha_r(x,I)$ depends upon x[44] is more challenging. However the whole conditional quantile function $V(\alpha|x,I) = X'\gamma(\alpha|I)$, $\alpha \in [0,1]$, can be identified provided $\inf_{x\in\mathcal{X}} \alpha_r(x,I) = 0$ for all $I \in \mathcal{I}$. The issue here is that the linear specification is not stable since changing $\alpha_r(I)$ into $\alpha_r(x,I)$ in (3.4.1) would give a slope coefficient $\beta(\alpha|I)$ which depends upon x. It is however possible to estimate $\gamma(\alpha|I)$ by selecting those x such that $\alpha_r(x,I)$ is close to some prescribed deterministic a provided an estimator of $\alpha_r(x,I)$ is available. Proposition 3.3-(i) suggests a nonparametric estimation but using the fact that $V(\alpha|x,I) = X'\gamma(\alpha|I)$ may improve this estimator in the favorable case where the whole quantile regression can be recovered. An iterative algorithm uses an additional smoothing parameter $1/J$ where $J = J_L$ is an integer number which diverges with the sample size, a preliminary estimation $\widehat{\alpha}_r^{(0)}(x,I)$ of $\alpha_r(x,I)$, and given the updated $\widehat{\alpha}_r^{(k)}(x,I)$, a partition of the sample according (x_ℓ, I_ℓ),[45]

$$\mathcal{L}_j^{(k)} = \left\{\ell = 1,\ldots,L; \frac{j-1}{J} \leq \widehat{\alpha}_r^{(k)}(x_\ell, I_\ell) < \frac{j}{J}\right\}, \quad j = 1,\ldots,J.$$

The algorithm decomposes into the following steps.

Step 1: a first estimation. For each j, estimate $\beta(\alpha|I)$, $\beta^{(1)}(\alpha|I)$ and $\gamma(\alpha|I)$ using the modified AQR procedure detailed above for $\alpha_r(x,I) = j/J$ and the subsample $\mathcal{L}_j^{(k)}$, for all $\alpha \geq j/J$. Let $\widehat{\gamma}^{j,k}(\alpha|I)$ be the resulting estimator.

Step 2: aggregation. Step 1 gives several estimation of $\gamma(\alpha|I)$, which can be improved using weighted estimators as

$$\widehat{\gamma}_{(k)}(\alpha|I) = \frac{\sum_{j=1}^{J} \widehat{\gamma}^{j,k}(\alpha|I)\,\mathbb{I}\left(\frac{j}{J} \leq \alpha\right)}{\sum_{j=1}^{J} \mathbb{I}\left(\frac{j}{J} \leq \alpha\right)}.$$

[43] Using a quantile specification $Q_{\beta,\alpha_r(I)}(\alpha,x,I)$ where $\alpha_r(I)$ becomes a parameter to be estimated is also feasible in principle.

[44] This can be in principle tested nonparametrically using Proposition 3.3-(i).

[45] The partition could be replaced by a kernel function as seen from the details of the algorithm. Using a kernel instead of a partition may be useful to decrease bias in the same way that a kernel p.d.f estimator improves on histogram.

Define also $\widehat{V}_{(k)}(\alpha|x,I) = X'\widehat{\gamma}_{(k)}(\alpha|I)$.

Step 3: updating $\widehat{\alpha}_r^{(k)}(x,I)$. For each (x_ℓ, I_ℓ), define

$$\widehat{\alpha}_r^{(k+1)}(x_\ell, I_\ell) = \arg\min_{\alpha} \left| \widehat{V}_{(k)}(\alpha|x,I) - r_\ell \right|.$$

Stop if the updated $\widehat{\alpha}_r^{(k+1)}(x_\ell, I_\ell)$ are close to the previous one and if the slope coefficient $\widehat{\gamma}_{(k)}(\alpha|I)$'s do not differ too much from the ones of previous iterations.

The updating Step 3 aims to improve the initial nonparametric estimation $\widehat{\alpha}_r^{(0)}(x,I)$ suggested by Proposition 3.3, which may be inefficient if the dimension of the covariate is large so that the resulting estimation of the slope $\gamma(\alpha|I)$ would be poor. Using an unconstrained estimator like $\widehat{\alpha}_r^{(0)}(x,I)$ may also give a poor idea of where the estimation performs well and where it does not. The aggregation step 2 is also important to recover fast estimation rates.

Choice of the smoothing parameters

Theorem 3.3 gives some asymptotic expressions for the bias and variance which can be used to propose a plug in bandwidth choice. Such bandwidth choices are however difficult to justify theoretically because they rely on an estimation of the bias which is hard to perform because it depends on higher order derivatives which cannot be estimated with a good rate.[46] Before proposing an alternative bandwidth choice for the AQR procedure, let us detail the integrated MSE approximation which follows

[46] Such plug in procedures are affected by an important contradiction: the focus should be on a bias corrected version of the initial estimator if estimating the bias was easily feasible.

from Theorem 3.3. Define, where s_1 is the $1 \times (s+2)$ vector $(0,1,0,\ldots,0)$,

$$\begin{aligned}
M_1(\alpha) &= \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I) X_\ell \alpha B^{(s+2)}(\alpha|x_\ell, I)}{B^{(1)}(\alpha|x_\ell, I)}\right], \quad M_2(\alpha) = \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I) X_\ell X_\ell'}{B^{(1)}(\alpha|x_\ell, I)}\right], \\
M &= \mathbb{E}\left[\mathbb{I}(I_\ell = I) X_\ell X_\ell'\right], \\
\Pi_1 &= \int \frac{t^{s+2}\pi(t)}{(s+2)!} K(t)\, dt, \quad \Pi_2 = \int \pi(t)\pi(t)' K(t)\, dt, \\
v^2 &= s_1 \Pi_2^{-1} \int\int \pi(t_1)\pi(t_2)' \min(t_1, t_2) K(t_1) K(t_2)\, dt_1 dt_2 s_1'.
\end{aligned}$$

The formulas of the asymptotic bias and variance in (3.3.9) and (3.3.8) and Theorem 3.3 suggest that the weighted integrated MSE,

$$\int_{\mathcal{X}} \int_0^1 \left(\widehat{V}_{AQR}(\alpha|x, I) - V(\alpha|x, I)\right)^2 \frac{d\alpha}{B(\alpha|x, I)} f(x, I)\, dx$$

has an expectation which leading term is,

$$\begin{aligned}
&\frac{v^2}{LIh} \int_0^1 \operatorname{Tr}\left(M_2(\alpha)^{-1} M\right) \frac{\alpha^2}{(I-1)^2} d\alpha \\
&\quad + h^{2(s+1)} \frac{(s_1 \Pi_1')^2 s_1 \Pi_2^{-1} s_1'}{(I-1)^2} \int_0^1 M_1(\alpha)' M_2(\alpha)^{-1} M_1(\alpha)\, d\alpha. \qquad (3.4.2)
\end{aligned}$$

The first item in (3.4.2) is a variance term which can be estimated while the second item is a bias term which involves the high-order derivative $B^{(s+2)}(\alpha|x_\ell, I)$ through $M_1(\alpha)$. The bandwidth miminizing the asymptotic leading term (3.4.2) of the IMSE is,

$$h_{opt} = \left(\frac{v^2 \int_0^1 \operatorname{Tr}\left(M_2(\alpha)^{-1} M\right) \alpha^2 d\alpha}{2(s+1) \times s_1 \Pi_1' \Pi_2^{-1} \Pi_1 s_1' \int_0^1 M_1(\alpha)' M_2(\alpha)^{-1} M_1(\alpha)\, d\alpha} \frac{1}{LI}\right)^{\frac{1}{2s+3}}. \qquad (3.4.3)$$

The key point is that it is possible to estimate the bias term $\int_0^1 M_1(\alpha)' M_2(\alpha)^{-1} M_1(\alpha)\, d\alpha$ up to scale and translation without estimating $B^{(s+2)}(\alpha|x_\ell, I)$. Define,

$$K_a\left(\frac{a-\alpha}{h}\right) = \frac{K\left(\frac{a-\alpha}{h}\right)}{\int K\left(\frac{a-\alpha}{h}\right) d\alpha},$$

$$\widehat{\boldsymbol{B}}(h;\alpha,I)=\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell=I)\sum_{i=1}^{I_\ell}\int_0^1 a^2\rho_a\left(B_{i\ell}-X_\ell(a-\alpha)'\widehat{\beta}_{AQR}(\alpha|I)\right)K_a\left(\frac{a-\alpha}{h}\right)\frac{da}{h}.$$

The kernel renormalisation is necessary to ensure that the leading term of $\int_0^1\widehat{\boldsymbol{B}}(h;\alpha,I)\,d\alpha$ is independent of h as implied by the next Proposition, and gives a simpler expression for the term in front of $h^{2(s+2)}$.

Proposition 3.4. *Suppose Assumptions A, S-(i), R-(i) and H-(i) hold, that the kernel $K(\cdot)$ is symmetric with $\int K(t)\,dt=1$ and $\lim_{L\to\infty}h\log^{(s+2)/4}L=0$. Then,*

$$\begin{aligned}\int_0^1\widehat{\boldsymbol{B}}(h;\alpha,I)\,d\alpha&=\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell=I)\sum_{i=1}^{I_\ell}\int_0^1\alpha^2\rho_\alpha\left(B_{i\ell}-B(\alpha|x_\ell,I_\ell)\right)d\alpha\\&+\frac{h^{2(s+2)}\Pi_1'\Pi_2^{-1}\Pi_1}{2}\int_0^1 M_1(\alpha)'M_2(\alpha)^{-1}M_1(\alpha)\,d\alpha\\&+o_{\mathbb{P}}\left(h^{2(s+2)}+\frac{1}{Lh}\right).\end{aligned}$$

The expansion of Proposition 3.4 does not allow for a direct estimation of the optimal bandwidth derived from (3.4.2) due to the power $h^{2(s+2)}$ instead of $h^{2(s+1)}$ in front of the bias term and a proportionality factor $\Pi_1'\Pi_2^{-1}\Pi_1/2I$ instead of $s_1\Pi_1'\Pi_2^{-1}\Pi_1 s_1'/(I-1)$. However building on correction ideas from Müller, Stadtmüller and Schmitt (1987) or Fan and Gijbels (1996, Sections 4.5 and 4.9) allows to propose an indirect estimator of the optimal bandwidth of (3.4.2). Let $\widehat{\sigma}^2$ be an estimator of $v^2\int_0^1\operatorname{Tr}\left(M_2(\alpha)^{-1}M\right)\alpha^2d\alpha$ and consider the pilot bandwidth,

$$\widehat{h}_0=\arg\min_h\left\{2\int_0^1\widehat{\boldsymbol{B}}(h;\alpha,I)\,d\alpha+\frac{\widehat{\sigma}^2}{Lh}\right\}.\tag{3.4.4}$$

Then Proposition 3.4 suggests that $\widehat{h}_0/h_0\xrightarrow{\mathbb{P}}1$ where,[47]

$$h_0=\left(\frac{v^2\int_0^1\operatorname{Tr}\left(M_2(\alpha)^{-1}M\right)\alpha^2d\alpha}{2(s+2)\times\Pi_1'\Pi_2^{-1}\Pi_1\int_0^1 M_1(\alpha)'M_2(\alpha)^{-1}M_1(\alpha)\,d\alpha}\frac{1}{LI}\right)^{\frac{1}{2s+5}}.$$

[47] Establishing such a result would request a uniform in bandwidth version of Proposition 3.4 which can be obtained with arguments similar to the ones used in Guerre and Sabbah (2012).

The pilot bandwidth h_0 differs from the optimal bandwidth h_{opt} of (3.4.3) through an exponent term and a proportionality factor. But it should be noted that h_{opt} is a known function of h_0 and the smoothness index s. This suggests, following Müller et al. (1987), to propose the following estimation of h_{opt},

$$\widehat{h}_{opt} = \left(\frac{(s+2) \times \Pi_1' \Pi_2^{-1} \Pi_1}{(s+1) \times (s_1 \Pi_1')^2 \, s_1 \Pi_2^{-1} s_1'} \widehat{h}_0^{2s+5} \right)^{\frac{1}{2s+3}} . \tag{3.4.5}$$

Alternative bandwidth choices to (3.4.5) can be derived from Rice (1986) or Charnigo, Hall and Srinivasab (2011), which propose bandwidth choices for estimation of the derivative of a regression function, including cross validation algorithm. Cross validation for local polynomial quantile estimation has been introduced in Abberger (1998), see also Li, Lin and Racine (2013) and Reiss and Huang (2012) for a review in the spline setup. Horowitz and Lee (2005) have used an Schwarz information criterion for the choice of the sieve truncation order K, see also Koenker, Ng and Portnoy (1994).

The ASQR case is more difficult due to the sieve truncation bias. However it is likely that Proposition 3.4 can be extended to cover this case, and that a minimization of a modified version of the objective function in (3.4.4) can deliver a suitable choice of K_L. The adaptive procedure of Goldenshluger and Lepski (2011) can also be useful here.

3.5 Conclusion

Starting with the standard quantile regression of Koenker and Bassett (1978), this third chapter has studied estimation of flexible additive interactive sieve quantile specifications which can be used with a rich set of covariates without being subject to the curse of dimensionality. A simple modification of the quantile regression objective function allows to estimate the first derivative of the conditional bid quantile function with respect to the quantile level. The conditional private value quantile function is recovered by summing the estimation of this derivative and of the conditional bid quantile function. The resulting estimator is free of asymptotic boundary bias, as necessary to estimate private values in the upper tail of the distribution or to estimate corresponding optimal bids. Hence the characteristics of potential winners can be better discovered using such quantile approaches than using other nonparametric methods which are subject to boundary bias. A small simulation study shows that the proposed private value quantile regression estimator using a data-driven bandwidth performs well even with small sample size of 100 bids.

Various extensions can be considered. Automatic bandwidth and sieve order choices were briefly considered but need deeper investigations. When the reservation price is binding, the first-price auction model involves censoring, and a dependence to the censoring quantile level which was not apparently considered elsewhere. Another issue is models with interactions higher than pairwise, for which the proposed approach works but under stronger conditions than for additive models, an issue that can probably be addressed with augmented local polynomial estimation of the bids quantile function and backfitting, or a two step procedure building on the proposed sieve approach as in Horowitz and Lee (2005). The proposed approach can also be implemented with alternative dimension reduction techniques such as single index quantile modelling as recently studied by Kong and Xia (2012) and suggested in an auction context by Marmer, Shneyerov and Xu (2013a,b). The fact that the augmented methodology is free from boundary effects can be useful for risk aversion estimation where the lower tail matters, see respectively Guerre et al. (2009) and Guerre and Vuong (2013) for risk aversion identification and estimation.

Appendix

Appendix A

Appendix for Chapter 1

A.1 Proofs

Proof of Lemma 1.1:

From the definition (1.2.1), the α-quantile of the private values conditional distribution must satisfies

$$F_v\left(V\left(\alpha|X,I\right)|X,I\right)=\alpha.$$

Under the IPV paradigm, a bidder i with private value v_i wins with probability $F_v^{I-1}\left(v_i|X,I\right)$. Therefore, in a quantile setup, we have that

$$F_v^{I-1}\left(V\left(\alpha|X,I\right)|X,I\right)=\alpha^{I-1},$$

which gives the result. □

Proof of Lemma 1.2:

From the definitions of quantiles and $\Psi_I\left(\cdot\right)$, the α-quantile of the winning bids c.d.f. must satisfies $\Psi_I\left(F_v\left(B\left(\alpha|X,I\right)|X,I\right)\right)=\alpha$. Because $\Psi_I\left(\cdot\right)$ is strictly increasing[48] in $[0,1]$, the winning bids conditional quantile can be written as

$$B\left(\alpha|X,I\right)=F_v^{-1}\left(\Psi_I^{-1}\left(\alpha\right)|X,I\right),$$

where $\Psi_I^{-1}\left(\cdot\right)$ is the inverse of $\Psi_I\left(\cdot\right)$. Hence, from (1.2.1) we find (1.2.5). □

Proof of Lemma 1.3:

If the private values conditional quantile has a quantile regression specification as in (1.2.4), by Lemma 1.2 there exists a vector of coefficient $\beta\left(\alpha|I\right)$, for each $\alpha\in[0,1]$, such that the winning bids conditional quantile satisfies $B\left(\alpha|X,I\right)=h\left(X\beta\left(\alpha|I\right)\right)$.

[48] The first derivative of $\Psi_I\left(t\right)$ is always positive for all $t\in(0,1)$:

$$\Psi_I^{(1)}\left(t\right)=\frac{\partial\Psi_I\left(t\right)}{\partial t}=I\left(I-1\right)t^{I-2}\left(1-t\right).$$

This is a stability property of the quantile regression approach. Since $h(\cdot)$ is continuous and strictly increasing, the candidate $\gamma(\cdot|I)$ must solve

$$X\gamma(\alpha|I) = X\beta(\Psi_I(\alpha)|I)$$

for all $\mathcal{X} = \{1\} \times \mathcal{Z}$. Since $\mathcal{Z}$ has a nonempty interior by Assumption 1.4, this equation has a unique solution and $\gamma(.|I)$ is identified for all I. $\square$

Proof of Proposition 1.1:

Consider the seller's payoff defined in (1.2.7). Under assumption 1.1, the seller possible payoffs are

$$\pi(r) = \begin{cases} v_0 & \text{if } v_{I:I} < r, \\ r & \text{if } v_{I-1:I} < r \leq v_{I:I}, \\ v_{I-1:I} & \text{if } r \leq v_{I-1:I}. \end{cases},$$

where $v_{1:I} \leq \cdots \leq v_{I:I}$ are the ordered private values.

We now rewrite $\pi(r)$ using quantiles and the ordered ranks

$$\alpha_{1:I} \leq \cdots \leq \alpha_{I:I}, \quad v_{i:I} = V(\alpha_{i:I}) \text{ for } i = 1, \ldots, I.$$

Recall that the non ordered α_i are i.i.d. $\mathcal{U}_{[0,1]}$ random variables. Because $V(\cdot|X, I)$ is strictly increasing by assumption 1.2,

$$\pi(\alpha_r|X, I) = \begin{cases} v_0, & \text{if } \alpha_{I:I} < \alpha_r, \\ V(\alpha_r|X, I), & \text{if } \alpha_{I-1:I} < \alpha_r \leq \alpha_{I:I}, \\ V(\alpha_{I-1:I}|X, I), & \text{if } \alpha_r \leq \alpha_{I-1:I}, \end{cases}$$

where α_r is the level of screening and $V(\alpha_r|X, I)$ the reservation price. It follows that the seller's expected payoff is

$$\begin{aligned} \Pi(\alpha_r|X, I, v_0) = v_0 \mathbb{P}(\alpha_{I:I} < \alpha_r|I) + V(\alpha_r|X, I)\, \mathbb{P}(\alpha_{I-1:I} < \alpha_r \leq \alpha_{I:I}|I) \\ + \mathbb{E}\left[V(\alpha_{I-1:I}|X, I)\, \mathbb{I}(\alpha_r \leq \alpha_{I-1:I})\,|X, I\right]. \end{aligned} \tag{A.1.1}$$

Observe that:

(i) $\mathbb{P}\left(\alpha_{I:I} < \alpha_r | I\right) = \alpha_r^I$;

(ii) Since $\sum_{i=1}^{I} \mathbb{I}\left(\alpha_i \leq \alpha_r\right)$ is a binomial distribution with parameters I and α_r,

$$\mathbb{P}\left(\alpha_{I-1:I} < \alpha_r \leq \alpha_{I:I} | I\right) = \mathbb{P}\left(\sum_{i=1}^{I} \mathbb{I}\left(\alpha_i \leq \alpha_r\right) = I - 1 | I\right) = I\alpha_r^{I-1}\left(1 - \alpha_r\right);$$

(iii) Since the c.d.f of $\alpha_{I-1:I}$ given I is $\Psi_I\left(\alpha\right)$ with p.d.f. $I\left(I-1\right)\alpha^{I-2}\left(1-\alpha\right)$, $\alpha \in [0,1]$,

$$\mathbb{E}\left[V\left(\alpha_{I-1:I}|X, I\right)\mathbb{I}\left(\alpha_r \leq \alpha_{I-1:I}\right)|X, I\right] = I\left(I-1\right)\int_{\alpha_r}^{1} V\left(\alpha|X, I\right)\alpha^{I-2}\left(1-\alpha\right)d\alpha.$$

Therefore, substituting (i)-(iii) into (A.1.1) gives equation (1.2.8), which is a quantile version of the seller's expected payoff.

Now, if $\Pi\left(\alpha_r|X, I, v_0\right)$ achieves a maximum for $\alpha_r^* = \alpha_r^*\left(X, I, v_0\right) \in [0,1]$, α_r^* must satisfy a first order condition $\frac{\partial}{\partial\alpha}\left\{\Pi\left(\alpha_r^*|X, I, v_0\right)\right\} = 0$. An expression for the first derivative of $\Pi\left(\alpha_r^*|X, I, v_0\right)$ with respect to the rank α is

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\left\{\Pi\left(\alpha|X, I, v_0\right)\right\} &= v_0 I\alpha^{I-1} + V^{(1)}\left(\alpha|X, I\right)I\alpha^{I-1}\left(1-\alpha\right) \\
&\quad + V\left(\alpha|X, I\right)\left\{I\left(I-1\right)\alpha^{I-2}\left(1-\alpha\right) - I\alpha^{I-1}\right\} \\
&\quad - V\left(\alpha|X, I\right)I\left(I-1\right)\alpha^{I-2}\left(1-\alpha\right) \\
&= I\alpha^{I-1}\left\{v_0 + V^{(1)}\left(\alpha|X, I\right)\left(1-\alpha\right) - V\left(\alpha|X, I\right)\right\}.
\end{aligned}$$

Hence, a candidate for the optimal level of screening α_r^* is given by

$$V\left(\alpha_r^*|X, I\right) - V^{(1)}\left(\alpha_r^*|X, I\right)\left(1-\alpha_r^*\right) = v_0.$$

□

Proof of Theorem 1.1 and Corollary 1.1:

We just detail the proof of the Theorem since the proof of the Corollary follows the same steps. Changes needed for the proof of the Corollary are discussed at the end of the section.

We show first that the private values quantile regression estimator in (1.3.15) is consistent. Since $-1 \leq \partial \rho_{\Psi_I(\alpha)}(t)/\partial t \leq 1$, the Taylor inequality gives

$$\left|\rho_{\Psi_I(\alpha)}(b_{w_\ell} - h(X_\ell\gamma)) - \rho_{\Psi_I(\alpha)}(b_{w_\ell})\right| \leq |h(X_\ell\gamma)| \leq C_0,$$

because $X_\ell\gamma \in \mathcal{X}\times\Gamma$ compact, and for this reason stays bounded from above. Hence,

$$\sup_{\gamma\in\Gamma}\left|\widehat{Q}(\gamma|\alpha,I) - Q(\gamma|\alpha,I)\right| \xrightarrow{\mathbb{P}} 0,$$

by the Uniform Law of Large Numbers in Newey and McFadden (1994, Lemma 2.4), where $Q(\gamma|\alpha,I)$ and $\widehat{Q}(\gamma|\alpha,I)$ are defined in section 1.3. Therefore, by (ii) in Lemma 1.3 and Newey and McFadden (1994, Theorem 2.1), $\gamma^*(\alpha|I)$ is uniquely identified and $\widehat{\gamma}(\alpha|I) \xrightarrow{\mathbb{P}} \gamma^*(\alpha|I)$.

For the sake of brevity, we remove the dependence on α and I, that is $\widehat{Q}(\gamma|\alpha,I)$, $Q(\gamma|\alpha,I)$, $\widehat{\gamma}(\alpha|I)$ and $\gamma^*(\alpha|I)$ become $\widehat{Q}(\gamma)$, $Q(\gamma)$, $\widehat{\gamma}$ and γ^*. Observe that the first and second-order derivatives of $Q(\gamma)$ are

$$\begin{aligned}
Q_\gamma(\gamma) &= \mathbb{E}\left[\{\mathbb{I}(b_w < h(X\gamma)) - \Psi_I(\alpha)\} h^{(1)}(X\gamma) X'|I\right] \\
&= \mathbb{E}\left[\{\mathbb{P}(b_w < h(X\gamma)|X,I) - \Psi_I(\alpha)\} h^{(1)}(X\gamma) X'|I\right] \\
&= \mathbb{E}\left[\{F_{b_w}(h(X\gamma)|X,I) - \Psi_I(\alpha)\} h^{(1)}(X\gamma) X'|I\right], \\
Q_{\gamma\gamma}(\gamma) &= \mathbb{E}\left[f_{b_w}(h(X\gamma)|X,I) h^{(1)}(X\gamma) X'X|I\right] \\
&\quad + \mathbb{E}\left[\{F_{b_w}(h(X\gamma)|X,I) - \Psi_I(\alpha)\} h^{(2)}(X\gamma) X'X|I\right],
\end{aligned}$$

as stated in the Theorem. Define

$$\begin{aligned}
u_\ell &= b_{w_\ell} - h(X_\ell\gamma^*), \\
\widehat{W} &= \frac{1}{L_I}\sum_{\ell=1}^{L} h^{(1)}(X_\ell\gamma^*)\,\mathbb{I}(I_\ell = I)\left(\mathbb{I}(u_\ell < 0) - \Psi_I(\alpha)\right) X_\ell, \\
\widehat{R}(\gamma) &= \frac{\sqrt{L_I}\left\{\widehat{Q}(\gamma) - \widehat{Q}(\gamma^*) - (\gamma-\gamma^*)'\widehat{W} - (Q(\gamma) - Q(\gamma^*))\right\}}{\|\gamma - \gamma^*\|},
\end{aligned}$$

where $\|\cdot\|$ is the usual Euclidean norm. Since γ^* is an interior point of Γ and because

$\widehat{\gamma}$ is consistent, the asymptotic normality of the Theorem follows from

$$\sqrt{L_I}\widehat{W} \xrightarrow{d} \mathcal{N}\left(0, J\left(\gamma^*|\alpha, I\right)\right), \tag{A.1.2}$$

$$\sup_{\gamma\in\Gamma;\|\gamma-\gamma^*\|\leq\epsilon_L} \left|\frac{\widehat{R}\left(\gamma\right)}{1+\sqrt{L_I}\|\gamma-\gamma^*\|}\right| = o_{\mathbb{P}}\left(1\right) \text{ for any } \epsilon_L = o\left(1\right), \tag{A.1.3}$$

see Theorem 7.1 in Newey and Mc Fadden (1994). We first establish (A.1.2). Note that the summands in $\widehat{W}$ are centered since

$$\mathbb{E}\left[\left(\mathbb{I}\left(u<0\right)-\Psi_I\left(\alpha\right)\right)h^{(1)}\left(X\gamma^*\right)X'|I\right] = Q_\gamma\left(\gamma^*\right) = 0,$$

because γ^* in the interior of Γ minimizes $Q\left(\cdot\right)$, and with a variance given I which is

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathbb{I}\left(u<0\right)-\Psi_I\left(\alpha\right)\right)^2 h^{(1)}\left(X\gamma^*\right)^2 X'X|I\right] = \\
&\mathbb{E}\left[\left(\mathbb{P}\left(u<0|X,I\right)-2\Psi_I\left(\alpha\right)\mathbb{P}\left(u<0|X,I\right)+\Psi_I\left(\alpha\right)^2\right)h^{(1)}\left(X\gamma^*\right)^2 X'X|I\right] \\
&= J\left(\gamma^*|\alpha, I\right),
\end{aligned}$$

since $P\left(u<0|X,I\right) = F_{b_w}\left(h\left(X\gamma^*\right)|X,I\right)$. Hence (A.1.2) follows from the Central Limit Theorem.

The rest of the proof consists then in showing (A.1.3). We first introduce some useful notations. Note that $Q_{\gamma\gamma}\left(\gamma\right) = Q^1_{\gamma\gamma}\left(\gamma\right) + Q^2_{\gamma\gamma}\left(\gamma\right)$ with

$$\begin{aligned}
Q^1_{\gamma\gamma}\left(\gamma\right) &= \mathbb{E}\left[\left\{F_{b_w}\left(h\left(X\gamma\right)|X,I\right)-\Psi_I\left(\alpha\right)\right\}h^{(2)}\left(X\gamma\right)X'X|I\right], \\
Q^2_{\gamma\gamma}\left(\gamma\right) &= \mathbb{E}\left[f_{b_w}\left(h\left(X\gamma\right)|X,I\right)h^{(1)}\left(X\gamma\right)^2 X'X|I\right].
\end{aligned}$$

Write

$$\begin{aligned}
\widehat{Q}\left(\gamma\right)-\widehat{Q}\left(\gamma^*\right) &= \frac{1}{L_I}\sum_{\ell=1}^{L}\mathbb{I}\left(I_\ell = I\right)\left\{\rho_{\Psi_I(\alpha)}\left(u_\ell - v_\ell\right)-\rho_{\Psi_I(\alpha)}\left(u_\ell\right)\right\} \text{ where,} \\
v_\ell &= v_\ell\left(\delta\right) = h\left(X_\ell\gamma\right)-h\left(X_\ell\gamma^*\right), \text{ where } \delta = \gamma-\gamma^*.
\end{aligned}$$

By an identity from Knight (1998),

$$\rho_{\Psi_I(\alpha)}(u_\ell - v_\ell) - \rho_{\Psi_I(\alpha)}(u_\ell) = v_\ell \{\mathbb{I}(u_\ell < 0) - \Psi_I(\alpha)\} + \int_0^{v_\ell} \{\mathbb{I}(u_\ell \le s) - \mathbb{I}(u_\ell \le 0)\} ds.$$

Hence, for $\delta = \gamma - \gamma^*$,

$$\widehat{Q}(\gamma) - \widehat{Q}(\gamma^*) - (\gamma - \gamma^*)' \widehat{W} - (Q(\gamma) - Q(\gamma^*)) = \sum_{j=1}^{5} \overline{\Delta}_j(\delta),$$

where $\overline{\Delta}_j(\delta) = \sum_{\ell=1}^{L} \Delta_{j\ell}(\delta) / L_I$ with

$$\begin{aligned}
\Delta_{1\ell}(\delta) &= \mathbb{I}(I_\ell = I) \left[\left(v_\ell - h^{(1)}(X_\ell \gamma^*) X_\ell \delta \right) \{\mathbb{I}(u_\ell < 0) - \Psi_I(\alpha)\} \right. \\
&\quad \left. - \mathbb{E}\left[\left(v_\ell - h^{(1)}(X_\ell \gamma^*) X_\ell \delta \right) \{\mathbb{I}(u_\ell < 0) - \Psi_I(\alpha)\} | I_\ell \right] \right], \\
\Delta_{2\ell}(\delta) &= \mathbb{I}(I_\ell = I) \left[\mathbb{E}\left[\left(v_\ell - h^{(1)}(X_\ell \gamma^*) X_\ell \delta \right) \{\mathbb{I}(u_\ell < 0) - \Psi_I(\alpha)\} | I_\ell \right] \right. \\
&\quad \left. - \frac{\delta' Q^1_{\gamma\gamma}(\gamma^*) \delta}{2} \right], \\
\Delta_{3\ell}(\delta) &= \mathbb{I}(I_\ell = I) \left[\mathbb{E}\left[\int_0^{v_\ell} \{\mathbb{I}(u_\ell \le s) - \mathbb{I}(u_\ell \le 0)\} ds | I_\ell \right] - \frac{\delta' Q^2_{\gamma\gamma}(\gamma^*) \delta}{2} \right], \\
\Delta_{4\ell}(\delta) &= \mathbb{I}(I_\ell = I) \left[\int_0^{v_\ell} \{\mathbb{I}(u_\ell \le s) - \mathbb{I}(u_\ell \le 0)\} ds \right. \\
&\quad \left. - \mathbb{E}\left[\int_0^{v_\ell} \{\mathbb{I}(u_\ell \le s) - \mathbb{I}(u_\ell \le 0)\} | I_\ell \right] \right], \\
\Delta_{5\ell}(\delta) &= \mathbb{I}(I_\ell = I) \left[Q(\gamma) - Q(\gamma^*) - \frac{\delta' Q^1_{\gamma\gamma}(\gamma^*) \delta}{2} - \frac{\delta' Q^2_{\gamma\gamma}(\gamma^*) \delta}{2} \right].
\end{aligned}$$

Now, since $v_\ell = h(X_\ell(\gamma^* + \delta)) - h(X_\ell \gamma)$ satisfies,

$$\begin{aligned}
v_\ell &= X_\ell \delta \int_0^1 h^{(1)}(X_\ell(\gamma^* + t\delta))(1 - t) dt, \\
v_\ell - h^{(1)}(X_\ell \gamma^*) X_\ell \delta &= \frac{\delta' X_\ell' X_\ell \delta}{2} \int_0^1 h^{(2)}(X_\ell(\gamma^* + t\delta))(1 - t) dt,
\end{aligned}$$

we have

$$\begin{aligned}\frac{\partial}{\partial\delta}\left[\frac{v_\ell - h^{(1)}\left(X_\ell\gamma^*\right)X_\ell\delta}{\|\delta\|}\right] &= \frac{v_\ell}{\|\delta\|} - \frac{\delta\left(v_\ell - h^{(1)}\left(X_\ell\gamma^*\right)X_\ell\delta\right)}{\|\delta\|^2} \\ &= \frac{\delta}{\|\delta\|}X_\ell\int_0^1 h^{(1)}\left(X_\ell\left(\gamma^*+t\delta\right)\right)(1-t)\,dt - \\ &\qquad \delta\frac{\delta' X_\ell' X_\ell\delta}{2\|\delta\|^2}\int_0^1 h^{(2)}\left(X_\ell\left(\gamma^*+t\delta\right)\right)(1-t)\,dt.\end{aligned}$$

This first differential is bounded when $\|\delta\| \leq 1$. Hence the set of functions $\{\Delta_{1\ell}(\delta)/\|\delta\|, \|\delta\| \leq 1\}$ can be covered with a number less than $O\left(\epsilon^{-C}\right)$ of functional brackets with length ϵ. To see that the same holds for $\{\Delta_{4\ell}(\delta)/\|\delta\|, \|\delta\| \leq 1\}$, observe

$$\begin{aligned}&\frac{1}{\|\delta\|}\int_0^{v_\ell}\left\{\mathbb{I}\left(u_\ell \leq s\right) - \mathbb{I}\left(u_\ell \leq 0\right)\right\}ds = \frac{v_\ell}{\|\delta\|}\int_0^1\left\{\mathbb{I}\left(u_\ell \leq s + tv_\ell\right) - \mathbb{I}\left(u_\ell \leq 0\right)\right\}dt \\ &= \left(X_\ell\frac{\delta}{\|\delta\|}\right)\times\left(\int_0^1 h^{(1)}\left(X_\ell\left(\gamma^*+t\delta\right)\right)(1-t)\,dt\right) \\ &\qquad\times\left(\int_0^1\left\{\mathbb{I}\left(u_\ell \leq s + tv_\ell\right) - \mathbb{I}\left(u_\ell \leq 0\right)\right\}dt\right).\end{aligned}$$

Each of the three functions of δ in this product can be covered with $O\left(\epsilon^{-C}\right)$ of functional brackets with length ϵ, implying that the same covering property holds for $\{\Delta_{4\ell}(\delta)/\|\delta\|, \|\delta\| \leq 1\}$. Hence, since $\{\Delta_{j\ell}(\delta)/\|\delta\|, \|\delta\| \leq \epsilon_L\}$ admits an envelope $\Delta_j(X_\ell)$ satisfying $E\left[\mathbf{\Delta}_j^2(X_\ell)\right] = o(1)$, $j = 1, 4$, Lemma 19.38 in van der Vaart (1998) gives that, with $\Gamma_\delta = \{\delta; \gamma^* + \delta \in \Gamma\}$

$$\mathbb{E}\left[\max_{\delta\in\Gamma_\delta;\|\delta\|\leq\epsilon_L}\left|\sqrt{L_I}\frac{\overline{\Delta}_j(\delta)}{\|\delta\|}\right| | I_\ell, 1 \leq \ell \leq L\right] = o_{\mathbb{P}}(1), \quad j = 1, 4.$$

This gives, by the Markov inequality,

$$\max_{\delta\in\Gamma_\delta;\|\delta\|\leq\epsilon_L}\left|\sqrt{L_I}\frac{\overline{\Delta}_j(\delta)}{\|\delta\|}\right| = o_{\mathbb{P}}(1), \quad j = 1, 4. \tag{A.1.4}$$

Now elementary expansions give, for the items in $\Delta_{j\ell}(\delta)$, $j = 2, 3, 5$, uniformly in ℓ since X_ℓ lies in a compact set,

$$
\begin{aligned}
&\mathbb{I}\left(I_\ell = I\right)\mathbb{E}\left[\left(v_\ell - h^{(1)}\left(X_\ell\gamma^*\right)X_\ell\delta\right)\left\{\mathbb{I}\left(u_\ell < 0\right) - \Psi_I\left(\alpha\right)\right\}|I_\ell\right] = \\
&\mathbb{I}\left(I_\ell = I\right)\mathbb{E}\left[\left(\frac{h^{(2)}\left(X_\ell\gamma^*\right)\delta' X_\ell X_\ell \delta}{2} + o\left(\|\delta\|^2\right)\right)\left\{\mathbb{P}\left(u_\ell < 0|X_\ell, I_\ell\right) - \Psi_I\left(\alpha\right)\right\}|I_\ell\right] \\
&\quad = \mathbb{I}\left(I_\ell = I\right)\frac{\delta' Q^1_{\gamma\gamma}\left(\gamma^*\right)\delta}{2} + o\left(\|\delta\|^2\right), \\
&\mathbb{I}\left(I_\ell = I\right)\mathbb{E}\left[\int_0^{v_\ell}\left\{\mathbb{I}\left(u_\ell \le s\right) - \mathbb{I}\left(u_\ell \le 0\right)\right\}ds|I_\ell\right] = \\
&\mathbb{I}\left(I_\ell = I\right)\mathbb{E}\left[\int_0^{v_\ell}\left\{\mathbb{P}\left(u_\ell \le s|X_\ell, I_\ell\right) - \mathbb{P}\left(u_\ell \le 0|X_\ell, I_\ell\right)\right\}ds|I_\ell\right] = \\
&\mathbb{I}\left(I_\ell = I\right)\mathbb{E}\left[\int_0^{v_\ell}\left\{f_{b_w}\left(h\left(X_\ell\gamma^*\right)|X_\ell, I_\ell\right) + o\left(1\right)\right\}sds|I_\ell\right] = \\
&\mathbb{I}\left(I_\ell = I\right)\mathbb{E}\left[\left\{f_{b_w}\left(h\left(X_\ell\gamma^*\right)|X_\ell, I_\ell\right) + o\left(1\right)\right\}\frac{v_\ell^2}{2}|I_\ell\right] = \\
&\mathbb{I}\left(I_\ell = I\right)\mathbb{E}\left[f_{b_w}\left(h\left(X_\ell\gamma^*\right)|X_\ell, I_\ell\right)\frac{\left(h^{(1)}\left(X_\ell\gamma^*\right)X_\ell\right)^2}{2}|I_\ell\right] + o\left(\|\delta\|^2\right) \\
&\quad = \mathbb{I}\left(I_\ell = I\right)\frac{\delta' Q^2_{\gamma\gamma}\left(\gamma^*\right)\delta}{2} + o\left(\|\delta\|^2\right), \\
&Q\left(\gamma\right) - Q\left(\gamma^*\right) - \frac{\delta' Q^1_{\gamma\gamma}\left(\gamma^*\right)\delta}{2} - \frac{\delta' Q^2_{\gamma\gamma}\left(\gamma^*\right)\delta}{2} = o\left(\|\delta\|^2\right).
\end{aligned}
$$

This gives by definition of $\overline{\Delta}_j\left(\delta\right)$, $j = 2, 3, 5$,

$$
\max_{\delta\in\Gamma_\delta;\|\delta\|\le\epsilon_L}\left|\frac{\overline{\Delta}_j\left(\delta\right)}{\|\delta\|^2}\right| = o_{\mathbb{P}}\left(1\right), \quad j = 2, 3, 5. \tag{A.1.5}
$$

Therefore (A.1.4), (A.1.5), $\gamma - \gamma^* = \delta$, and $\widehat{R}\left(\gamma\right) = \sqrt{L_I}\sum_{j=1}^5 \overline{\Delta}_j\left(\delta\right)$ imply

$$
\begin{aligned}
\sup_{\gamma\in\Gamma;\|\gamma-\gamma^*\|\le\epsilon_L}\left|\frac{\widehat{R}\left(\gamma\right)}{1 + \sqrt{L_I}\|\gamma - \gamma^*\|}\right| &\le 2\max_{j=1,4}\max_{\delta\in\Gamma_\delta;\|\delta\|\le\epsilon_L}\left|\sqrt{L_I}\frac{\overline{\Delta}_j\left(\delta\right)}{\|\delta\|}\right| \\
&\quad +3\max_{j=1,3,5}\max_{\delta\in\Gamma_\delta;\|\delta\|\le\epsilon_L}\left|\frac{\overline{\Delta}_j\left(\delta\right)}{\|\delta\|^2}\right| \\
&= o_{\mathbb{P}}\left(1\right).
\end{aligned}
$$

Hence (A.1.3) is true.

For Corrollary 1.1, the first derivative of $\widehat{Q}^0(\gamma) = \widehat{Q}(\gamma|\alpha)$ is

$$\widehat{Q}^0_\gamma(\gamma) = \frac{1}{L}\sum_{\ell=1}^{L}\left(\mathbb{I}\left(b_{w_\ell} \leq h\left(X_\ell\gamma\right)\right) - \Psi_{I_\ell}(\alpha)\right) h^{(1)}\left(X_\ell\gamma\right) X'_\ell.$$

Let $\gamma^* = \gamma(\alpha)$. As seen from the proof of Theorem 1.1, the matrix $J(\gamma^*|\alpha)$ is the variance of $\widehat{W}^0 = L^{1/2}\widehat{Q}^0_\gamma(\gamma^*)$ and because $\mathbb{E}\left[\widehat{W}^0\right] = 0$ asymptotic normality follows. Hence $J(\gamma^*|\alpha)$ is equal to

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathbb{I}\left(b_w \leq h\left(X\gamma^*\right)\right) - 2\mathbb{I}\left(b_w \leq h\left(X\gamma^*\right)\right)\Psi_I(\alpha) + \Psi_I^2(\alpha)\right) h^{(1)}\left(X\gamma^*\right) X'X\right] = \\
&\mathbb{E}\left[\left(\mathbb{P}\left(b_w \leq h\left(X\gamma^*\right)|X, I\right) - 2\mathbb{P}\left(b_w \leq h\left(X\gamma^*\right)|X, I\right)\Psi_I(\alpha) + \Psi_I^2(\alpha)\right)\right. \\
&\qquad\qquad\qquad\qquad \left.\times h^{(1)}\left(X\gamma^*\right) X'X\right],
\end{aligned}$$

which gives the expression of the Corollary since $\mathbb{P}\left(b_w \leq h\left(X\gamma^*\right)|X, I\right) = F_{b_w}\left(h\left(X\gamma^*\right)|X, I\right)$. Now, similar computations give that

$$\begin{aligned}
Q^0_{\gamma\gamma}(\gamma^*) &= \mathbb{E}\left[\widehat{Q}^0_{\gamma\gamma}(\gamma^*)\right] \\
&= \mathbb{E}\left[\left(f_{b_w}\left(h\left(X\gamma^*\right)|X, I\right) - \Psi_{I_\ell}(\alpha)\right) h^{(1)}\left(X\gamma^*\right) X'X\right] \\
&\quad + \mathbb{E}\left[\left(F_{b_w}\left(h\left(X\gamma^*\right)|X, I\right) - \Psi_{I_\ell}(\alpha)\right) h^{(2)}\left(X\gamma^*\right) X'X\right].
\end{aligned}$$

□

A.2 Random Weighting Bootstrap Method

This section describes the random weighting bootstrap method used in the tests of exclusion participation restriction and constancy of the slope coefficients. Let $\widehat{Q}_{H_0}(\widehat{\gamma}_{H_0}|\alpha, I)$ and $\widehat{Q}_{H_1}(\widehat{\gamma}_{H_1}|\alpha, I)$ be the optimized value of the quantile regression objective function under the null and alternative hypothesis, respectively. The M-test statistic defined in Rao and Zhao (1992), Wang and Zhou (2004) and Zhao, Wu and Yang (2007) is

$$M = L_I\left[\widehat{Q}_{H_0}(\widehat{\gamma}_{H_0}|\alpha, I) - \widehat{Q}_{H_1}(\widehat{\gamma}_{H_1}|\alpha, I)\right]. \tag{A.2.6}$$

Let π_ℓ be i.i.d. random weights multinomially distributed with parameters $(L_I, 1/L_I)$. Define the random weighting empirical objective function under the null and alternative hypotheses as $\widehat{Q}_{H_0}(\gamma_{H_0}|\alpha, I, \pi)$ and $\widehat{Q}_{H_1}(\gamma_{H_1}|\alpha, I, \pi)$, where

$$\widehat{Q}_{H_j}\left(\gamma_{H_j}|\alpha, I, \pi\right) = \frac{1}{L_I}\sum_{\ell=1}^{L_I} \mathbb{I}\left(I_\ell = I\right)\pi_\ell \rho_{\Psi_I(\alpha)}\left(b_{w_\ell} - h\left(X_\ell\gamma\right)\right), \tag{A.2.7}$$

for $j = \{0, 1\}$. The aforementioned reference suggests to random weighting bootstrap the M-statistic (A.2.6) by using the random weights π_ℓ. Let b be the index for a draw π^b of the weights. For each $j = \{0, 1\}$, define

$$\widehat{\gamma}^b_{H_j}(\alpha|I) = \arg\min_{\gamma_{H_j}\in\Gamma} \widehat{Q}_{H_j}\left(\gamma_{H_j}|\alpha, I, \pi^b\right).$$

Then, the bth draw of the bootstrapped statistic is

$$\begin{aligned} M^b = L_I &\left[\widehat{Q}_{H_0}\left(\widehat{\gamma}^b_{H_0}|\alpha, I, \pi^b\right) - \widehat{Q}_{H_1}\left(\widehat{\gamma}^b_{H_1}|\alpha, I, \pi^b\right)\right] \\ &\qquad - L_I\left[\widehat{Q}_{H_0}\left(\widehat{\gamma}_{H_0}|\alpha, I, \pi^b\right) - \widehat{Q}_{H_1}\left(\widehat{\gamma}_{H_1}|\alpha, I, \pi^b\right)\right] \end{aligned}$$

where $\widehat{\gamma}_{H_j}$ are the estimators obtained from the initial population.

The null hypothesis H_0 is rejected at a significance level τ if M is larger than the sample $(1-\tau)$-quantile of the bootstrapped statistic $M^1, \cdots, M^B$ and accepted otherwise, where B is the number of bootstrap replications.

A.3 Monte Carlo Experiments

This section presents some Monte Carlo experiments used to illustrate the hypotheses tests and their performance in terms of size. For all the experiments below, let $\mathcal{A} = \{0.50, 0.60, 0.70, 0.80\}$.

Exclusion Participation Restriction.

Data Generating Process. We generate two samples of auction covariates and error terms. One sample corresponds to L_2 auctions with $I = 2$ actual bidders and the other to L_3 auctions with $I = 3$, where $L_2 = L_3 = 100$. For the auction covariates, we generate L_I random Normal $(0,1)$ variables and replicate it I times. The $I \times L_I$ error terms are randomly generated considering three different parametric distributions: Uniform $(0,1)$, Standard Exponential and Standard Normal. The bidders' private values are generated via two possible regression specifications: linear $v_{i\ell} = 1 + X_{1\ell} + X_{2\ell} + u_{i\ell}$ and exponential $v_{i\ell} = \exp(0.3 + 0.3X_{1\ell} + 0.3X_{2\ell} + u_{i\ell})$[49]. The winning bid is given by the second-highest private value: $b_{w_\ell} = v_{I-1:I,\ell}$.

Estimation and Test. We use the winning bids to estimate $\gamma(\alpha|I=2)$ and $\gamma(\alpha|I=3)$, via the individual quantile regression estimator (1.3.15), and $\gamma(\alpha)$, via the pooled quantile regression estimator (1.3.18), for all $\alpha \in \mathcal{A}$. We then compute

$$M = \widehat{Q}(\widehat{\gamma}|\alpha) - \sum_{I\in\{2,3\}} \widehat{Q}(\widehat{\gamma}(\alpha|I)|\alpha, I)\, L_I/L$$

for each quantile level α and sum over $\mathcal{A}$.

Bootstrap. We generate L_2 and L_3 i.i.d. random weights π multinomially distributed with parameters $(L_I, 1/L_I)$ and estimate $\widehat{Q}\left(\widehat{\gamma}^b(\alpha)|\alpha,\pi^b\right)$ and $\widehat{Q}\left(\widehat{\gamma}^b(\alpha|I)|\alpha,I,\pi^b\right)$ for each bootstrap draw b, where the former corresponds to the pooled optimized objective function and the latter to the individual one. We then compute the random weighting bootstrap statistic

$$\begin{aligned} M^b = &\Big[\widehat{Q}\left(\widehat{\gamma}^b(\alpha)|\alpha,\pi^b\right) - \sum_{I\in\{2,3\}} \widehat{Q}\left(\widehat{\gamma}^b(\alpha|I)|\alpha,I,\pi^b\right) L_I/L\Big] - \\ &\qquad \Big[\widehat{Q}\left(\widehat{\gamma}(\alpha)|\alpha,\pi^b\right) - \sum_{I\in\{2,3\}} \widehat{Q}\left(\widehat{\gamma}(\alpha|I)|\alpha,I,\pi^b\right) L_I/L\Big], \end{aligned}$$

where $\widehat{\gamma}(\alpha)$ and $\widehat{\gamma}(\alpha|I)$ are the estimates of the original population, and sum over the quantile levels. The number of bootstrap replications is $B = 500$.

Bootstrap Critical Values. We choose a significance level τ for the test. The bootstrap critical value $c^*(\tau)$ is an approximation of the $c(\tau)$ critical value of the test statistic M and is given by the $(1-\tau)$th quantile of the empirical distribution of the bootstrapped statistic.

Rejection Rule. If $M < c^*(\tau)$, then we do not reject the null hypothesis at $\tau\%$ significance level.

[49]The values of the intercept and slope coefficients are not an important issue and were chosen simply to speed up the estimation process.

Table A.1: Rejection Probability - Linear Specification

Nominal Level	Std. Normal	Uniform $(0,1)$	Std. Exponential
1%	1.06%	1.18%	0.87%
5%	5.09%	5.47%	3.62%
10%	9.84%	10.84%	7.30%

The number of bootstrap replications is $B = 500$; The experiment was repeated $N = 10,000$ times; The set of prescribed quantiles is $\mathcal{A} = \{0.50, 0.60, 0.70, 0.80\}$.

Table A.2: Rejection Probability - Exponential Specification

Nominal Level	Std. Normal	Uniform $(0,1)$	Std. Exponential
1%	1.32%	1.62%	1.17%
5%	5.14%	6.16%	4.61%
10%	9.94%	11.54%	9.34%

The number of bootstrap replications is $B = 500$; The experiment was repeated $N = 10,000$ times; The set of prescribed quantiles is $\mathcal{A} = \{0.50, 0.60, 0.70, 0.80\}$.

Rejection Probability. We repeat the procedure above $N = 10,000$ times and compute how many times the null hypothesis is rejected.

The results from the Monte Carlo experiment above are given in Tables A.1 and A.2. The rejection probabilities are very close to the nominal levels in both cases. It is slightly undersized in the case in which the errors follow a standard exponential distribution and the model has a linear specification.

Choice of Specification

Data Generating Process. As in the previous experiment, we generate two samples of auction covariates and error term with L_2 and L_3 auctions, where $L_2 = L_3 = 200$. The auction covariates are random Standard Normal variables that are replicated I times. For the error terms, we consider the distributions: Standard Normal, Uniform $(0, 1)$ and Standard Exponential. The bidders' private values are generated via two possible regression specifications: linear $v^{\mathrm{L}}_{i\ell} = \gamma_0 + \gamma_1 Z_{1\ell} + u_{i\ell}$ and exponential $v^{\mathrm{E}}_{i\ell} = \exp(\gamma_0 + \gamma_1 Z_{1\ell} + u_{i\ell})$, where $(\gamma_0, \gamma_1) = (1, 1)$ in the case the errors are normally distributed and $(\gamma_0, \gamma_1) = (0, 0.5)$ when the errors are either uniformly or exponentially distributed[50]. Under the null hypothesis, the

[50]The values of the intercept and slope parameters were chosen according to the distribution of the errors in an attempt to find a DGP that satisfies the equivalence property for a given quantile

exponential specification is equivalent to the linear specification for a given $\alpha_0 \in \mathcal{A}$ and dominates the latter for all $\alpha \in \mathcal{A} \backslash \{\alpha_0\}$. In order to have equivalence at $\alpha_0 \in \mathcal{A}$, we create a private value $v_{i\ell}$ that is a linear combination of both private values above: $v_{i\ell} = \lambda v_{i\ell}^{\mathrm{L}} + (1-\lambda) v_{i\ell}^{\mathrm{E}}$. To find the weight λ that satisfies the equivalence property, we split $\lambda \in [0,1]$ into $1,000$ equal spaces and compute the relative difference between both optimized objective functions,

$$\Delta(\alpha, \lambda_t) = \frac{\left(\widehat{Q}_{\mathrm{E}}(\widehat{\gamma}_{\mathrm{E}}|\alpha, \lambda_t) - \widehat{Q}_{\mathrm{L}}(\widehat{\gamma}_{\mathrm{L}}|\alpha, \lambda_t)\right)}{\widehat{Q}_{\mathrm{L}}(\widehat{\gamma}_{\mathrm{L}}|\alpha, \lambda_t)},$$

for $t = \{1, \cdots, 1,000\}$. We choose λ^* across t that minimizes $\Delta(\alpha, \lambda_t)$ for a given α. The equivalence property must hold for a given $\alpha_0 \in \mathcal{A}$ and the exponential specification should be better than the linear one for all $\alpha \in \mathcal{A} \backslash \{\alpha_0\}$. Therefore, we choose $\lambda_{\max}^*$ such that $\sup_{\alpha \in \mathcal{A}} \Delta(\alpha, \lambda^*) = 0$. The winning bid is given by the second-highest private value: $b_{w_\ell} = v_{I-1:I,\ell}$.

Estimation and Test. We generate a DGP using $\lambda_{\max}^*$ and compute the test statistic $\widehat{\text{Vuong}}$ as in (1.4.23).

Bootstrap. We resample with replacement the (X_ℓ, b_{w_ℓ})-pair from their joint sample distribution in each original subsample L_I. The number of bootstrap replications is $B = 500$.

Bootstrapping the Test Statistic. For each bootstrap replication b, we compute the recentered bootstrapped test statistic $\widehat{\text{Vuong}}^b$ as

$$\widehat{\text{Vuong}}^b = \sup_{\alpha \in \mathcal{A}} \sqrt{L} \left\{ \left(\widehat{Q}_{\mathrm{E}}(\widehat{\gamma}_{\mathrm{E}}^b|\alpha) - \widehat{Q}_{\mathrm{L}}(\widehat{\gamma}_{\mathrm{L}}^b|\alpha) \right) - \left(\widehat{Q}_{\mathrm{E}}(\widehat{\gamma}_{\mathrm{E}}|\alpha) - \widehat{Q}_{\mathrm{L}}(\widehat{\gamma}_{\mathrm{L}}|\alpha) \right) \right\},$$

where $\widehat{\gamma}_{\mathrm{E}}$ and $\widehat{\gamma}_{\mathrm{L}}$ are the estimates of the original population generated by $\lambda_{\max}^*$.

The bootstrap critical values, rejection rule and rejection probability are computed as in Appendix C.1. In Table A.3, we present the rejection probabilities and the values for $\lambda_{\max}^*$ used in each experiment. The test performs well although a bit undersized. That is because we use the supremum over the quantiles in the formula, which means that the rejection probability could be higher for a given quantile level.

level.

Table A.3: Rejection Probability - Test of Specification

Nominal Level	Std. Normal	Uniform $(0,1)$	Std. Exponential
1%	2.23%	0.94%	0.91%
5%	5.67%	4.50%	4.22%
10%	8.31%	9.04%	8.22%
$\lambda^*_{\max}$	0.6547	0.5846	0.5095

The number of bootstrap replications is $B = 500$; The experiment was repeated $N = 10,000$ times; The set of prescribed quantiles is $\mathcal{A} = \{0.50, 0.60, 0.70, 0.80\}$; It is assumed independence of the private values distribution upon I in this Monte Carlo experiment.

Constancy of the Slope Coefficients

Data Generating Process. As in the previous two experiments, we generate two samples of auction covariates and error terms with L_2 and L_3 auctions, where $L_2 = L_3 = 100$. For the auction covariates, we generate L_I random Standard Normal variables and replicate it I times. The $I \times L_I$ error terms are randomly generated considering the parametric distributions: Uniform $(0,1)$, Standard Exponential and Standard Normal. The private values are generated using a linear regression model due to the results of the empirical application, i.e. $v_{i\ell} = 0 + Z_\ell + u_{i\ell}$. The winning bid is given by the second-highest private value: $b_{w_\ell} = v_{I-1:I,\ell}$.

Estimation and Test. We use the winning bids to estimate (γ_{0_k}, γ_1), $k = 1, \cdots, K$, via the CQR estimator defined in (1.3.21), and $\gamma(\alpha)$ via the pooled quantile regression estimator (1.3.18). Under the null hypothesis, the CQR population objective function equals the average of the individual objective functions over $\mathcal{A}$. We then compute the M-statistic

$$M = L \left[\widehat{Q}_{\mathrm{CQR}}\left(\widehat{\gamma}_{\mathrm{CQR}}\right) - \frac{1}{K} \sum_{k=1}^{K} \widehat{Q}\left(\widehat{\gamma}|\alpha_k\right) \right],$$

where $K = 4$ since $\mathcal{A} = \{0.50, 0.60, 0.70, 0.80\}$.

Bootstrap. We generate L_2 and L_3 i.i.d. random weights π multinomially distributed with parameters $(L_I, 1/L_I)$ and estimate $\widehat{Q}_{\mathrm{CQR}}\left(\widehat{\gamma}^b_{\mathrm{CQR}}|\pi^b\right)$ and $\widehat{Q}\left(\widehat{\gamma}^b\left(\alpha_k\right)|\alpha_k, \pi^b\right)$ at each bootstrap draw b and quantile level α_k. We then compute

Table A.4: Rejection Probability - Constancy of the Slope Coefficients

Nominal Level	Std. Normal	Uniform $(0,1)$	Std. Exponential
1%	1.25%	1.15%	1.09%
5%	6.25%	6.73%	5.59%
10%	11.74%	12.27%	10.95%

The number of bootstrap replications is $B = 500$; The experiment was repeated $N = 10,000$ times; The set of prescribed quantiles is $\mathcal{A} = \{0.50, 0.60, 0.70, 0.80\}$; It is assumed independence of the private values distribution upon I in this Monte Carlo experiment.

the random weight bootstrap statistic as

$$M^b = L\Big[\widehat{Q}_{\text{CQR}}\left(\widehat{\gamma}^b_{\text{CQR}}|\pi^b\right) - \frac{1}{K}\sum_{k=1}^{K}\widehat{Q}\left(\widehat{\gamma}^b\left(\alpha_k\right)|\alpha_k,\pi^b\right)\Big] - L\left[\widehat{Q}_{\text{CQR}}\left(\widehat{\gamma}_{\text{CQR}}|\pi^b\right) - \frac{1}{K}\sum_{k=1}^{K}\widehat{Q}\left(\widehat{\gamma}\left(\alpha_k\right)|\alpha_k,\pi^b\right)\right],$$

where $\widehat{\gamma}_{\text{CQR}}$ and $\widehat{\gamma}\left(\alpha_k\right)$ are the estimates of the original population. The number of bootstrap replications is $B = 500$.

The bootstrap critical values, rejection rule and rejection probability are computed as in Appendix C.1. Table A.4 gives the rejection probabilities of the test. The test performs well although a bit oversized, which is due to the tolerance level asigned for the convergence of the algorithm.

Appendix B

Appendix for Chapter 3

Proofs

In all this proof section, $g(b|x,I)$ and $G(b|x,I)$ are respectively the conditional p.d.f and c.d.f of the bids $B_{i\ell}$ given $(x_\ell, I_\ell) = (x, I)$, so that

$$B(\alpha|x,I) = G^{-1}(\alpha|x,I), \quad B^{(1)}(\alpha|x,I) = \frac{1}{g(B(\alpha|x,I)|x,I)}$$

will be often used.

Proofs of the results in Sections 3.2 and 3.4

This subsection groups the proofs of the results of Sections 3.2 and 3.4 with the exception of the proof of Proposition 3.4 which is established with our main estimation results, see Section B below.

Proof of Lemma 3.1. If $\alpha \in [0,1] \mapsto V(\alpha|x,I)$ is continuous and strictly increasing, the private value rank U_i in (3.2.1) is uniquely defined. Recall also that Lemma A2 in GPV yields that the continuously differentiable $s_0(\cdot|x,I)$ has a derivative $s_0^{(1)}(v|x,I) > 0$ over $[V(0|x,I), V(1|x,I)]$ so that $B(\cdot|x,I)$ in (3.2.3) is continuous strictly increasing. $B_i = s_0(V_i|x,I)$ and $V_i = V(U_i|x,I)$ give, for $s_0(v|x,I)$ as in (3.2.2),

$$B_i = s_0(V(U_i|x,I)|x,I).$$

This implies that the optimal bids satisfies $B_i = B(U_i|x,I)$ by (3.2.3), hence (ii) is proved. Since the continuous strictly increasing $B(\cdot|x,I)$ has a reciprocal function $B^{-1}(\cdot|x,I)$ and $U_i = F(V_i|x,I)$ is uniform over $[0,1]$, $B^{-1}(\cdot|x,I) = G(\cdot|x,I)$ and $B_i = B(U_i|x,I)$ gives $U_i = G(B_i|x,I)$, hence (i). For (iii), $B(a|x,I)$ is a winning bid if and only if $B(a|x,I) > \max_{1\leq j\neq i\leq I} B_j$ so that the probability of interest is, since $B(\cdot|x,I)$ is continuous and strictly increasing and $B_j = B(U_j|x,I)$ with i.i.d. uniform U_j given (x,I),

$$\begin{aligned}\mathbb{P}\left(B(a|x,I) > \max_{1\leq j\neq i\leq I} B_j|x,I\right) &= \mathbb{P}\left(B(a|x,I) > \max_{1\leq j\neq i\leq I} B(U_j|x,I)|x,I\right) \\ &= \mathbb{P}\left(a > \max_{1\leq j\neq i\leq I} U_j\right) = a^{I-1}. \qquad \square\end{aligned}$$

Proof of Lemma 3.2. Drop the dependence upon x and I for the sake of brevity. By the expected revenue theorem, it is sufficient to compute the expected revenue for a second price auction where the paid price is the maximum of r and second highest bid, assuming that the largest bid is higher than r since there is no transaction otherwise. For such auction, a dominant strategy is to bid his private value. Let $V_{(1)} = V\left(U_{(1)}\right) \leq \cdots \leq V_{(I)} = V\left(U_{(I)}\right)$ be the ordered private values, so that the expected revenue is, since $V(\cdot)$ is continuous and strictly increasing,

$$
\begin{aligned}
&V_0(x, I)\,\mathbb{P}\left(V_{(I)} \leq r\right) + \mathbb{E}\left[r\mathbb{I}\left(V_{(I-1)} \leq r \leq V_{(I)}\right) + V_{(I-1)}\mathbb{I}\left(r \leq V_{(I-1)}\right)\right] \\
&\quad = V_0(x, I)\,\alpha_r^I + r\mathbb{P}\left(\sum_{i=1}^{I} \mathbb{I}\left(U_i \leq \alpha_r\right) = I - 1\right) + \mathbb{E}\left[V\left(U_{(I-1)}\right)\mathbb{I}\left(\alpha_r \leq U_{(I-1)}\right)\right].
\end{aligned}
$$

The probability above is $I\alpha_r^{I-1}(1-\alpha_r)$ by definition of the binomial distribution with parameter (I, α_r). As recalled in Athey and Haile (2002, eq. 5), the c.d.f of $U_{(I-2)}$ is $\Psi(\alpha) = I\alpha^{I-1} - (I-1)\alpha^I$, with density $\psi(\alpha) = I(I-1)\alpha^{I-2} - I(I-1)\alpha^{I-1} = I(I-1)\alpha^{I-2}(1-\alpha)$, so that

$$
\mathbb{E}\left[V\left(U_{(I-1)}\right)\mathbb{I}\left(r \leq V\left(U_{(I-1)}\right)\right)\right] = I(I-1)\int_{\alpha_r}^{1} V(\alpha)\,\alpha^{I-2}(1-\alpha)\,d\alpha,
$$

which establishes the claim of the Lemma. □

Proof of Theorem 3.1. Drop the dependence on x and I. We first prove part (ii). As shown in GPV and provided $f(v)$ is continuous with $\inf_{v\in[V(0),V(1)]} f(v) > 0$ as ensured by the condition on $V(\cdot)$, an optimal bid b computed from a private value v satisfies $v = \xi(b)$ where the continuously differentiable and strictly increasing $\xi(\cdot)$ is the inverse of $s_0(\cdot)$ and can be written,

$$
\xi(b) = b + \frac{1}{I-1}\frac{G(b)}{g(b)}.
$$

Note that $G(b) = F\left(s_0^{-1}(b)\right)$ is one to one with $\inf_{b\in[B(0),B(1)]} g(b) > 0$ since $s_0(\cdot)$ is continuously differentiable with $\inf_{v\in[V(0),V(1)]} s_0^{(1)}(v) > 0$, so that $G(B(\alpha)) = \alpha$ and $B^{(1)}(\alpha) = 1/g(B(\alpha))$. Hence since $B(\alpha)$ is an optimal bid for a private value

$V(\alpha)$ by (3.2.3),

$$V(\alpha) = B(\alpha) + \frac{1}{I-1}\frac{G(B(\alpha))}{g(B(\alpha))} = B(\alpha) + \frac{\alpha B^{(1)}(\alpha)}{I-1},$$

which is (3.2.6).

Consider now part (i). Since $s_0(V(0)) = V(0)$, the initial condition $B(0) = V(0)$ holds. Write, for $\alpha \in (0,1]$, $B(\alpha) = C(\alpha)/\alpha^{I-1}$, so that $B^{(1)}(\alpha) = C^{(1)}(\alpha)/\alpha^{I-1} - (I-1)B(\alpha)/\alpha$. Substituting in (3.2.6) gives that $C(\cdot)$ satisfies $C^{(1)}(\alpha) = (I-1)\alpha^{I-2}V(\alpha)$, implying that

$$C(\alpha) = C(0) + (I-1)\int_0^\alpha t^{I-2}V(t)\,dt.$$

Since continuity of $V(\cdot)$ gives $\lim_{\alpha\to 0}\alpha^{-(I-1)}(I-1)\int_0^\alpha t^{I-2}V(t)\,dt = V(0)$, $B(\alpha) = \lim_{\alpha\to 0} C(\alpha)/\alpha^{I-1} = B(0) = V(0)$ implies,

$$B(\alpha) = \frac{I-1}{\alpha^{I-1}}\int_0^\alpha t^{I-2}V(t)\,dt,$$

showing that the claim of part (i) is true. □

Proof of Lemma 3.3. Observe that,

$$\frac{I-1}{\alpha^{I-1}}\int_0^\alpha t^{I-2}V(t|x,I)\,dt = (I-1)\int_0^1 u^{I-2}V(\alpha u|x,I)\,dt,$$
$$\frac{I-1}{\alpha^{I-1}}\int_0^\alpha t^{I-2}\gamma_k(t|I)\,dt = (I-1)\int_0^1 u^{I-2}\gamma_k(\alpha u|I)\,dt,$$

which implies that $\sum_{k=1}^\infty \beta_k(\alpha|I)P_k(x)$ converges uniformly over $[0,1]\times\mathcal{X}$ and is equal to $B(\alpha|x,I)$ as defined in (3.2.5). For $B^{(1)}(\alpha|x,I)$, integrating by parts gives

$$\beta_k(\alpha|I) = \frac{1}{\alpha^{I-1}}\int_0^\alpha \gamma_k(t|I)\,d\left[t^{I-1}\right] = \gamma_k(\alpha|I) - \frac{1}{\alpha^{I-1}}\int_0^\alpha t^{I-1}\gamma_k^{(1)}(t|I)\,dt.$$

It follows that,

$$\beta_k^{(1)}(\alpha|I) = \gamma_k^{(1)}(\alpha|I) - \gamma_k^{(1)}(\alpha|I) + \frac{I-1}{\alpha^I}\int_0^\alpha t^{I-1}\gamma_k^{(1)}(t|I)\,dt = (I-1)\int_0^1 u^{I-1}\gamma_k^{(1)}(\alpha u|I)\,dt,$$

which implies that, as $\sum_{k=1}^{\infty} \gamma_k^{(1)}(\alpha|I) P_k(x)$, $\sum_{k=1}^{\infty} \beta_k^{(1)}(\alpha|I) P_k(x)$ converges uniformly over $[0,1] \times \mathcal{X}$ and is equal to the α-derivative of $\sum_{k=1}^{\infty} \beta_k(\alpha|I) P_k(x) = B(\alpha|x, I)$. This also gives that $\beta_k(\alpha|I) + \alpha\beta_k^{(1)}(\alpha|I)/(I-1) = \gamma_k(\alpha|I)$. □

Proof of Proposition 3.1. By (3.2.5), $B(\alpha|x, I) = (I-1)\int_0^1 u^{I-2} V(\alpha u|x, I)\, du$, so that $B^{(1)}(\alpha|x, I) = (I-1)\int_0^1 u^{I-1} V^{(1)}(\alpha u|x, I)\, du$ which gives (i). For (ii), (3.2.10) and $\beta_k(\alpha|I) = (I-1)\int_0^1 u^{I-2}\gamma_k(\alpha u|I)\, du$ implies,

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| B(\alpha|x, I) - \sum_{k=1}^{K} \beta_k(\alpha|I) P_k(x) \right|$$
$$\leq \max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| V(\alpha|x, I) - \sum_{k=1}^{K} \gamma_k(\alpha|I) P_k(x) \right| = o\left(K^{-\frac{s+2}{d_{\mathcal{M}}}}\right).$$

This also implies that $B^{(s+1)}(\alpha|x, I) = (I-1)\int_0^1 u^{I-1} V^{(s+1)}(\alpha u|x, I)\, du$ is continuous over $[0,1] \times \mathcal{X}$ as is $V^{(s+1)}(\alpha|x, I)$, that

$$\beta_k^{(s+1)}(\alpha|I) = (I-1)\int_0^1 u^{I+s-1}\gamma_k^{(s+1)}(\alpha u|I)\, du,$$

and then,

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| B^{(s+1)}(\alpha|x, I) - \sum_{k=1}^{K} \beta_k^{(s+1)}(\alpha|I) P_k(x) \right|$$
$$\leq \max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| V^{(s+1)}(\alpha|x, I) - \sum_{k=1}^{K} \gamma_k^{(s+1)}(\alpha|I) P_k(x) \right| \underset{K\to\infty}{\longrightarrow} 0.$$

For (iii), it follows from (ii) that,

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| \alpha B(\alpha|x, I) - \sum_{k=1}^{K} \alpha\beta_k(\alpha|I) P_k(x) \right| = o\left(K^{-\frac{s+2}{d_{\mathcal{M}}}}\right).$$

Moreover, for all $\alpha \in [0,1]$,

$$B^{(s+1)}(\alpha|x,I) = (I-1)\int_0^1 u^{I-1+s} V^{(s+1)}(\alpha u|x,I)\,du = \frac{I-1}{\alpha^{I+s}}\int_0^\alpha t^{I-1+s} V^{(s+1)}(t|x,I)\,dt,$$

and then, for $\alpha \in (0,1]$,

$$\begin{aligned} B^{(s+2)}(\alpha|x,I) &= -\frac{(I-1)(I+s)}{\alpha^{I+s+1}}\int_0^\alpha t^{I-1+s} V^{(s+1)}(t|x,I)\,dt + \frac{(I-1)V^{(s+1)}(\alpha|x,I)}{\alpha} \\ &= -\frac{(I+s)}{\alpha} B^{(s+1)}(\alpha|x,I) + \frac{(I-1)V^{(s+1)}(\alpha|x,I)}{\alpha}. \end{aligned}$$

Hence, it follows from the L'Hospital rule that, for all $\alpha \in [0,1]$,

$$\begin{aligned} &\frac{\partial^{s+2}}{\partial \alpha^{s+2}}\left[\alpha B(\alpha|x,I)\right] = \alpha B^{(s+2)}(\alpha|x,I) + (s+2) B^{(s+1)}(\alpha|x,I) \\ &= (2-I) B^{(s+1)}(\alpha|x,I) + (I-1) V^{(s+1)}(\alpha|x,I), \\ &\frac{\partial^{s+2}}{\partial \alpha^{s+2}}\left[\alpha \beta_k(\alpha|I)\right] = (2-I)\beta_k^{(s+1)}(\alpha|I) + (I-1)\gamma_k^{(s+1)}(\alpha|I). \end{aligned}$$

This shows $\alpha B(\alpha|x,I) \in \mathcal{P}_{s+2,s+1}$ since $B(\alpha|x,I)$ and $V(\alpha|x,I)$ are both in $\mathcal{P}_{s+1,s+1}$. $\square$

Proofs of Lemmas 3.4 and 3.5. Since the two proofs are similar, we just detail the one of Lemma 3.5. Let

$$\widehat{\varrho}_A = \max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|\widehat{A}(\alpha|x,I) - A(\alpha|x,I)\right|, \quad A \in \{B, V\}.$$

Then, since $\widehat{U}_{i\ell} = \arg\min_{\alpha\in[0,1]}\left|B_{i\ell} - \widehat{B}(\alpha|x_\ell, I_\ell)\right|$ and $B_{i\ell} = B(U_{i\ell}|x_\ell, I_\ell)$ where

$B(\cdot|x_\ell, I_\ell)$ is increasing,

$$\begin{aligned}
&\left\{\left|\widehat{U}_{i\ell} - U_{i\ell}\right| \geq t\right\} \\
&\subset \left\{\min_{\alpha\in[0,1];|U_{i\ell}-\alpha|\geq t} \left|B_{i\ell} - \widehat{B}(\alpha|x_\ell, I_\ell)\right| \leq \min_{\alpha\in[0,1];|U_{i\ell}-\alpha|<t} \left|B_{i\ell} - \widehat{B}(\alpha|x_\ell, I_\ell)\right|\right\} \\
&\subset \left\{\min_{\alpha\in[0,1];|U_{i\ell}-\alpha|\geq t} \left|B_{i\ell} - \widehat{B}(\alpha|x_\ell, I_\ell)\right| \leq \left|B(U_{i\ell}|x_\ell, I_\ell) - \widehat{B}(U_{i\ell}|x_\ell, I_\ell)\right|\right\} \\
&\subset \left\{\min_{\alpha\in[0,1];|U_{i\ell}-\alpha|\geq t} \left|B(U_{i\ell}|x_\ell, I_\ell) - B(\alpha|x_\ell, I_\ell)\right| - \widehat{\varrho}_B \leq \widehat{\varrho}_B\right\} \\
&\subset \left\{t \times \min_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} B^{(1)}(\alpha|x, I) \leq 2\widehat{\varrho}_B\right\},
\end{aligned}$$

and Proposition 3.1 implies that, for $C > 2/\min_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} B^{(1)}(\alpha|x, I)$,

$$\max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|\widehat{U}_{i\ell} - U_{i\ell}\right| \leq C\widehat{\varrho}_B.$$

Hence, since $\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x, I) < \infty$,

$$\begin{aligned}
&\max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|\widehat{V}_{i\ell} - V_{i\ell}\right| = \max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|\widehat{V}\left(\widehat{U}_{i\ell}|x_\ell, I_\ell\right) - V(U_{i\ell}|x_\ell, I_\ell)\right| \\
&\leq \max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{V}(\alpha|x, I) - V(\alpha|x, I)\right| + \max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|V\left(\widehat{U}_{i\ell}|x_\ell, I_\ell\right) - V(U_{i\ell}|x_\ell, I_\ell)\right| \\
&\leq \widehat{\varrho}_V + C\widehat{\varrho}_B. \qquad \square
\end{aligned}$$

Proof of Proposition 3.2. This follows from similar steps than the proof of Theorem 3.1. $\square$

Proof of Proposition 3.3. (i) is true because all bidders with a private value larger than $r(x, I)$ make a bid, so that the conditional distribution I^* is a binomial distribution with parameters I and $1 - F(r(x, I)|x, I) = 1 - \alpha_r$, and $\mathbb{E}[I^*|x, I] = I(1 - \alpha_r)$. (ii) follows from solving (3.2.6) with the initial condition $B(\alpha_r(x, I)|x, I) = r(x, I)$ as in Theorem 3.1. Dropping the dependence on x and I, and setting $B(\alpha) = C(\alpha)/\alpha^{I-1}$ now gives

$$C(\alpha) = C(\alpha_r) + (I - 1)\int_{\alpha_r}^{\alpha} t^{I-2}V(t)\,dt.$$

The initial condition implies that $C(\alpha_r) = \alpha_r^{I-1} r$, which gives the desired solution. (iii) follows from (ii). $\square$

Bias of the AQR and ASQR estimators

We start with additional notations used all along the proof section and some preliminary lemmas. Let $\mathcal{R}_{AQR}(b;\alpha,I)$ and $\mathcal{R}_{ASQR}(b_L;\alpha,I)$ be the population counterparts of $\widehat{\mathcal{R}}_{AQR}(b;\alpha,I)$ and $\widehat{\mathcal{R}}_{ASQR}(b_L;\alpha,I)$, that is, setting $b_L = b$ for the sake of brevity,

$$\mathcal{R}_{AQR}(b;\alpha,I) = \mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell} - X_\ell(ht)' b\right) K(t)\,dt\right],$$
$$\mathcal{R}_{ASQR}(b;\alpha,I) = \mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell} - P(x_\ell,ht)' b\right) K(t)\,dt\right],$$

where U is a $\mathcal{U}_{[0,1]}$ random variable independent of (x_ℓ, I_ℓ). These objective functions are very similar so that it is sufficient in many cases to study $\mathcal{R}_{ASQR}(b;\alpha,I)$. Since

$$\frac{\partial \rho_{\alpha+ht}\left(B(U|x_\ell,I) - P(x_\ell,ht)' b\right)}{\partial b} = \left\{\mathbb{I}\left(B_{i\ell} \le P(x_\ell,ht)' b\right) - (\alpha+ht)\right\} P(x_\ell,ht),$$

almost everywhere, the Lebesgue differentiation theorem yields that the partial derivative of $\mathcal{R}_{ASQR}(b;\alpha,I)$ with respect to b is

$$\mathcal{R}^{(1)}_{ASQR}(b;\alpha,I) = \mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left\{G\left(P(x_\ell,ht)' b | x_\ell, I_\ell\right) - (\alpha+ht)\right\} P(x_\ell,ht) K(t)\,dt\right],$$

where $G(\cdot|x,I) = B^{-1}(\cdot|x,I)$ is the conditional c.d.f of $B_{i\ell}$ given $(x_\ell, I_\ell) = (x,I)$. By Proposition 3.1 and $s+1 \ge 1$, $G(\cdot|x,I)$ has a p.d.f $g(\cdot|x,I)$ satisfying,

$$g(y|x,I) = \frac{1}{\frac{\partial B}{\partial \alpha}(G(y|x,I),x,I)}, \quad y \in [B(0|x,I), B(1|x,I)].$$

Hence $\mathcal{R}_{ASQR}(b;\alpha,I)$ has a second-order partial derivative with respect to b, which is,

$$\mathcal{R}^{(2)}_{ASQR}(b;\alpha,I)=\mathbb{E}\left[\mathbb{I}(I_\ell=I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} g\left(P(x_\ell,ht)'b|x_\ell,I_\ell\right)P(x_\ell,ht)P(x_\ell,ht)'K(t)\,dt\right],$$

and $\mathcal{R}^{(2)}_{AQR}(b;\alpha,I)$ has a similar expression. However these matrices depends upon $h\to 0$ and will not have an inverse. To avoid such a drawback, define

$$\mathsf{b}=\left[\beta_{0,1},\ldots,\beta_{0,k},h\beta_{1,1},\ldots,h\beta_{1,k},\ldots,h^{s+1}\beta_{1,1},\ldots,h^{s+1}\beta_{1,k}\right], \tag{B.0.1}$$

with $k=d+1$ for AQR and $k=K_L$ for ASQR, so that $P(x_\ell,ht)'b=P(x_\ell,t)'\mathsf{b}$. Define also

$$\mathsf{R}_{AQR}(\mathsf{b};\alpha,I)=\mathcal{R}_{AQR}(b;\alpha,I),\quad \mathsf{R}_{ASQR}(\mathsf{b};\alpha,I)=\mathcal{R}_{ASQR}(b;\alpha,I),$$

so that,

$$\mathsf{R}^{(2)}_{ASQR}(\mathsf{b};\alpha,I)=\mathbb{E}\left[\mathbb{I}(I_\ell=I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} g\left(P(x_\ell,t)'\mathsf{b}|x_\ell,I_\ell\right)P(x_\ell,t)P(x_\ell,t)'K(t)\,dt\right],$$

the expression of $\mathsf{R}^{(2)}_{AQR}(\mathsf{b};\alpha,I)$ being very similar. The next Lemma gives conditions ensuring that $\mathsf{R}^{(2)}_{ASQR}(\mathsf{b};\alpha,I)$ is definite positive.

Lemma B.1. *Suppose that Assumptions A, H, R hold and that $h\in[0,1/2]$. Then there are some constants $\Lambda_{AQR}>0$ and $\Lambda_{ASQR}>0$ such that, for all L,*

i. For all conformable vector v, all $\alpha\in[0,1]$, $I\in\mathcal{I}$ and b,

$$v'\mathsf{R}^{(2)}_{AQR}(\mathsf{b};\alpha,I)v\le\Lambda_{AQR}\times(v'v),\quad v'\mathsf{R}^{(2)}_{ASQR}(\mathsf{b};\alpha,I)v\le\Lambda_{ASQR}\times(v'v);$$

ii. Suppose that for all $x\in\mathcal{X}$, all $t\in\left[\frac{1}{4},\frac{3}{4}\right]$ if $\alpha\in\left[0,\frac{1}{2}\right]$ or all $t\in\left[-\frac{3}{8},-\frac{1}{4}\right]$ if $\alpha\in\left(\frac{1}{2},1\right]$, $X(x,t)'\mathsf{b}$ (respectively $P(x,t)'\mathsf{b}$) is in $[B(0|x,I),B(1|x,I)]$. Then, for all $\alpha\in[0,1]$, $I\in\mathcal{I}$ and all such b, it holds that for all conformable *vector v,*

$$v'\mathsf{R}^{(2)}_{AQR}(\mathsf{b};\alpha,I)\,v \geq \frac{v'v}{\Lambda_{AQR}}, \quad \textit{(respectively } v'\mathsf{R}^{(2)}_{ASQR}(\mathsf{b};\alpha,I)\,v \geq \frac{v'v}{\Lambda_{ASQR}}\textit{).}$$

Proof of Lemma B.1. It is sufficient to study the ASQR case and $\alpha \in [0,1/2]$. Let $\preceq$ be the usual order for symmetric matrices. Then Assumption A, the expression of $g(\cdot|\cdot,\cdot)$ and Proposition 3.1-(i) give $\mathsf{R}^{(2)}_{ASQR}(\mathsf{b};\alpha,I) \preceq C\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} P(x,t)\,P(x,t)'\,K(t)\,dtdx$. But

$$\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} P(x,t)\,P(x,t)'\,K(t)\,dtdx = \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} P(x)\,P(x)'\,dx \otimes \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi(t)\,\pi(t)'\,K(t)\,dt,$$

where the eigenvalues of $\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} P(x)\,P(x)'\,dx$ are in a compact set for all L by Assumption R-(ii). This gives (i). For (ii), consider the case where $\alpha \in [0,1/2]$. Then Assumption A, the expression of $g(\cdot|\cdot,\cdot)$ and Proposition 3.1-(i) yield for those b,

$$\begin{aligned}\mathsf{R}^{(2)}_{ASQR}(\mathsf{b};\alpha,I) \succeq & \min_{(x,I)\in\mathcal{X}\times\mathcal{I}}\ \min_{y\in[B(0|x,I),B(1|x,I)]}\{g(y|x,I)\,f(x|I)\,\mathbb{P}(I_\ell=I)\} \\ & \times \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} P(x)\,P(x)'\,dx \otimes \int_{\frac{1}{4}}^{\frac{3}{4}} \pi(t)\,\pi(t)'\,K(t)\,dt,\end{aligned}$$

which gives the desired result since the two matrices in the tensor product are symmetric and strictly definite positive. $\square$

Since

$$\rho_{\alpha+ht}\left(B_{i\ell} - P(x_\ell,ht)'\,b\right) \geq \min(\alpha+ht, 1-(\alpha+ht))\left|B_{i\ell} - P(x_\ell,ht)'\,b\right|,$$

it follows, under Assumption H, that for all b,

$$\begin{aligned}\mathcal{R}_{ASQR}(b;\alpha,I) &= \mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell} - P(x_\ell,ht)'b\right)K(t)\,dt\right] \\ &\geq Ch\mathbb{E}\left[\mathbb{I}(I_\ell = I)\left|B_{i\ell} - P(x_\ell,ht)'b\right|\right] \\ &\geq Ch\left\{\mathbb{E}\left[\mathbb{I}(I_\ell = I)\left|P(x_\ell,ht)'b\right|\right] - \mathbb{E}\left[\mathbb{I}(I_\ell = I)\left|B_{i\ell}\right|\right]\right\}.\end{aligned}$$

Because $b \rightarrow \mathbb{E}\left[\mathbb{I}(I_\ell = I)\left|P(x_\ell,ht)'b\right|\right]$ is a norm under Assumption A, $\lim_{\|b\|\to\infty}\mathcal{R}_{ASQR}(b;\alpha,I) = +\infty$, and the convex and continuous $b \rightarrow \mathcal{R}_{ASQR}(b;\alpha,I)$ is bowl-shaped, and so is also $\mathcal{R}_{AQR}(b;\alpha,I)$. Therefore, $\mathcal{R}_{AQR}(b;\alpha,I)$ and $\mathcal{R}_{ASQR}(b;\alpha,I)$ both have a minimizer, say for all $\alpha \in [0,1]$,

$$\begin{aligned}b_{AQR}(\alpha|I) &= \arg\min_b \mathcal{R}_{AQR}(b;\alpha,I), \\ b_{ASQR}(\alpha|I) &= \arg\min_b \mathcal{R}_{ASQR}(b;\alpha,I).\end{aligned}$$

It will be shown below that these minimizers are both unique. Recall that,

$$\begin{aligned}b_{AQR}(\alpha|I) &= \left[\beta_{0,AQR}(\alpha|I)',\ldots,\beta'_{s+1,AQR}(\alpha|I)\right]', \quad \beta_{0,AQR}(\alpha|I),\ \beta_{1,AQR}(\alpha|I) \in \mathbb{R}^{d+1}, \\ \mathsf{b}_{AQR}(\alpha|I) &= \mathrm{Diag}(\pi(h)) \otimes \mathrm{Id}_{p+1}\, b_{AQR}(\alpha|I), \\ \mathsf{b}_{0,AQR}(\alpha|I) &= \beta_{0,AQR}(\alpha|I),\ \mathsf{b}_{1,AQR}(\alpha|I) = h\beta_{1,AQR}(\alpha|I), \qquad\qquad (\text{B.0.2}) \\ b_{ASQR}(\alpha|I) &= \left[\beta_{0,ASQR}(\alpha|I)',\ldots,\beta'_{s+1,ASQR}(\alpha|I)\right]', \quad \beta_{0,ASQR}(\alpha|I), \beta_{1,ASQR}(\alpha|I) \in \mathbb{R}^{K_L}, \\ \mathsf{b}_{SAQR}(\alpha|I) &= \mathrm{Diag}(\pi(h)) \otimes \mathrm{Id}_{K_L}\, b_{ASQR}(\alpha|I), \\ \mathsf{b}_{0,ASQR}(\alpha|I) &= \beta_{0,ASQR}(\alpha|I),\ \mathsf{b}_{1,ASQR}(\alpha|I) = h\beta_{1,ASQR}(\alpha|I).\end{aligned}$$

Following (3.2.5) in Theorem 3.1, (3.3.4) and (3.3.5), define,

$$\begin{aligned}\gamma_{AQR}(\alpha|I) &= \beta_{0,AQR}(\alpha|I) + \frac{\alpha\beta_{1,AQR}(\alpha|I)}{I-1}, \\ \gamma_{ASQR}(\alpha|I) &= \beta_{0,ASQR}(\alpha|I) + \frac{\alpha\beta_{1,ASQR}(\alpha|I)}{I-1}, \\ V_{AQR}(\alpha|x,I) &= X'\gamma_{AQR}(\alpha|I), \quad V_{ASQR}(\alpha|x,I) = P(x)'\gamma_{ASQR}(\alpha|I).\end{aligned}$$

The study of the bias $V_{AQR}(\alpha|x,I) - V(\alpha|x,I)$ and $V_{ASQR}(\alpha|x,I) - V(\alpha|x,I)$

requires some specific techniques to cope with the fact that $B^{(s+2)}(\alpha|x,I)$ may be unbounded in the vicinity of 0. It is based on the following Lemma which is a consequence of the Kantorovitch-Newton Theorem, see e.g Gragg and Tapia (1974). In this lemma $\|\cdot\|$ stands for the Euclidean norm over $\mathbb{R}^D$ or for the associate operator norm.

Lemma B.2. *Let $\mathcal{F}(\cdot):\mathbb{R}^D\to\mathbb{R}$ be a function. Suppose that there is a $\mathbf{x}^*\in\mathbb{R}^D$ and some real numbers $\epsilon>0$ and $C_0>0$ such that $\mathcal{F}(\cdot)$ is twice differentiable on $\mathcal{B}(\mathbf{x}^*,2C_0\epsilon)=\left\{x\in\mathbb{R}^D;\|x-\mathbf{x}^*\|<2C_0\epsilon\right\}$. If, in addition,*

i. $\left\|\mathcal{F}^{(1)}(\mathbf{x}^*)\right\|\leq\epsilon$ *and* $\left\|\left[\mathcal{F}^{(2)}(\mathbf{x}^*)\right]^{-1}\right\|\leq C_0$;

ii. There is a $C_1>0$ such that $\left\|\mathcal{F}^{(2)}(x)-\mathcal{F}^{(2)}(x')\right\|\leq C_1\|x-x'\|$ for all $x,x'\in\mathcal{B}(\mathbf{x}^*,2C_0\epsilon)$;

iii. $C_0^2C_1\epsilon\leq 1/2$.

Then there is a unique $\mathbf{x}$ such that $\|\mathbf{x}-\mathbf{x}^*\|<2C_0\epsilon$ and $\mathcal{F}^{(1)}(\mathbf{x})=0$.

In the next theorem, $\mathsf{Bias}_h(\alpha|x,I)$ is from (3.3.9).

Theorem B.1. *Suppose that Assumptions A and H-(i) hold. Then,*

i. If the quantile regression model in Assumption S-(i) is the true one, Assumption R-(i) holds, $b_{AQR}(\alpha|I)$ is unique and

$$\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\|V_{AQR}(\alpha|x,I)-V(\alpha|x,I)\|=O\left(h^{s+1}\right).$$

Moreover, $V_{AQR}(\alpha|x,I)=V(\alpha|x,I)+h^{s+1}X'\mathsf{Bias}_h(\alpha|x,I)+o\left(h^{s+1}\right)$ uniformly over $[0,1]\times\mathcal{X}$.

ii. For the sieve quantile regression model (3.2.9) with $K_L\to 0$ and under Assumptions S-(ii), R-(ii),

$$\max_{x\in\mathcal{X}}\left\{\sum_{k=1}^{K_L}P_k^2(x)\right\}=O\left(\frac{h^2}{\left(h^{s+1}+K_L^{-(s+2)/d_{\mathcal{M}}}\right)^2}\right)\ \textit{and}\ K_L^{-(s+2)/d_{\mathcal{M}}}=o(h),\tag{B.0.3}$$

$b_{ASQR}(\alpha|I)$ is unique and

$$\max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\int_{\mathcal{X}}\left(V_{ASQR}(\alpha|x,I)-V(\alpha|x,I)\right)^2dx=O\left(h^{s+1}+\frac{K_L^{-(s+2)/d_{\mathcal{M}}}}{h}\right);$$

iii. If in addition Assumption H-(iii) holds,

$$\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|V_{ASQR}(\alpha|x,I)-V(\alpha|x,I)\right|=O\left(h^{s+1}+\frac{K_L^{-(s+2)/d_{\mathcal{M}}}}{h}\right).$$

Note that (B.0.3) follows from Assumption H-(ii) since, by $K_L \asymp h^{-d_{\mathcal{M}}}$, $K_L^{-(s+2)/d_{\mathcal{M}}} \asymp h^{s+2} = o(h)$ and, because $(1+\zeta)d_{\mathcal{M}}/2 \leq s$, $\max_{x\in\mathcal{X}}\|P(x)\| = O\left(K_L^{(1+\zeta)/2}\right) = O\left(h^{-(1+\zeta)d_{\mathcal{M}}/2}\right) = O(h^{-s})$.

Proof of Theorem B.1. Consider (ii) and (iii), the proof of (i) being similar as detailed below. The proof works by establishing that there is a solution of the first-order condition in a open ball where $\mathsf{R}_{ASQR}(\mathsf{b};\alpha,I)$ is strictly convex by checking the conditions of Lemma B.2, which will also gives the rate stated in the Theorem. Define, for $\beta_k(\alpha|I)$ as in (3.2.10) and $a+ht\in[0,1]$,

$$\mathsf{b}=\left[\beta_{0,1},\ldots,\beta_{0,K_L},h\beta_{1,1},\ldots,h\beta_{1,K_L},\ldots,h^{s+1}\beta_{1,1},\ldots,h^{s+1}\beta_{1,K_L}\right],$$
$$\mathsf{b}(\alpha|I)=\left[\beta_1(\alpha|I),\ldots,\beta_{K_L}(\alpha|I),\ldots,\frac{h^{s+1}}{(s+1)!}\beta_1^{(s+1)}(\alpha|I),\ldots,\frac{h^{s+1}}{(s+1)!}\beta_{K_L}^{(s+1)}(\alpha|I)\right].$$

Abbreviate $\mathsf{R}_{ASQR}(\mathsf{b};\alpha,I)$ into $\mathsf{R}(\mathsf{b};\alpha,I)$. Note that,

$$\begin{aligned}B(\alpha+ht|x,I)-P(x,t)'\mathsf{b}(\alpha|I)&=\sum_{k=1}^{K_L}\left(\beta_k(\alpha+ht|I)-\sum_{j=0}^{s+1}\frac{(ht)^j}{j!}\beta_k^{(j)}(\alpha|I)\right)P_k(x)\\&\quad+\sum_{k=K_L+1}^{+\infty}\beta_k(\alpha+ht|I)P_k(x).\end{aligned}$$

Assumption S-(ii), Proposition 3.1 and a Taylor expansion give

$$\max\left|P(x,t)'\mathsf{b}(\alpha|I)-B(\alpha+ht|x,I)\right|=o\left(h^{s+1}+K_L^{-(s+2)/d_{\mathcal{M}}}\right), \tag{B.0.4}$$

where the maximum is over those $(\alpha, t, x, I) \in [0,1] \times [-1,1] \times \mathcal{X} \times \mathcal{I}$ with $a + ht \in [0,1]$. This, together with the expression of $\mathsf{R}^{(1)}(\mathsf{b}; \alpha, I)$ and Assumption R-(ii) give,

$$\max_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left\| \mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right); \alpha, I\right) \right\| = \epsilon_L \text{ with } \epsilon_L = o\left(h^{s+1} + K_L^{-(s+2)/d_{\mathcal{M}}}\right). \tag{B.0.5}$$

To see this, observe that

$$\left\| \mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right); \alpha, I\right) \right\| = \max_{\theta;\theta'\theta=1} \left| \theta' \mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right); \alpha, I\right) \right|. \tag{B.0.6}$$

But uniformly in $\alpha \in [0,1]$, setting $\pi(t) = [1, \ldots, t^{s+1}]'$ and by Assumption R-(ii), (B.0.4),

$$\begin{aligned}
&\left| \theta' \mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right); \alpha, I\right) \right| \\
&= \mathbb{E}\left[\mathbb{I}\left(I_\ell = I\right) \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left\{ G\left(P\left(x_\ell, t\right) \mathsf{b}\left(\alpha|I\right) | x_\ell, I_\ell\right) - G\left(B\left(\alpha + ht|x, I\right) | x_\ell, I_\ell\right) \right\} \right. \\
&\qquad \left. \theta'\left(P\left(x_\ell\right) \otimes \pi\left(t\right)\right) K\left(t\right) dt \right] \\
&\leq C\epsilon_L \mathbb{E}^{1/2}\left[\int_{-1}^{1} \left(\theta'\left(P\left(x_\ell\right) \otimes \pi\left(t\right)\right)\right)^2 dt \right] \leq C\epsilon_L \left(\theta'\theta\right)^{1/2}.
\end{aligned}$$

Consider without loss of generality the case where $\alpha \in [0, 1/2]$, $\alpha \in (1/2, 1]$ being similar. Since $B\left(\alpha + ht|x, I\right) = B\left(\alpha|x, I\right) + O\left(h\right)$ uniformly in α, x, I and $t \in [-1,1]$, $\partial B\left(\alpha|x, I\right)/\partial\alpha$ is bounded away from 0 and $K_L^{-(s+2)/d_{\mathcal{M}}} = o\left(h\right)$, (B.0.4) gives that

$$P\left(x, t\right)' \mathsf{b}\left(\alpha|I\right) \in \left[B\left(0|x, I\right), B\left(1|x, I\right)\right] \text{ for all } t \in \left[\frac{1}{4}, \frac{3}{4}\right],$$

all $\alpha \in [0, 1/2]$, x and I and L large enough. It then follows by Lemma B.1 that, for some $C_0 > 0$ and all L large enough,

$$\max_{(\alpha,I)\in[0,1/2]\times\mathcal{I}} \left\| \left[\mathsf{R}^{(2)}\left(\mathsf{b}\left(\alpha|I\right); \alpha, I\right)\right]^{-1} \right\| \leq C_0,$$

recalling that $\|\cdot\|$ stands for the spectral norm in the equation above. Hence condition (i) in Lemma B.2 holds provided L is large enough, as assumed from now on. Let $\mathcal{B} = \mathcal{B}\left(\mathsf{b}\left(\alpha|I\right), 2C_0\epsilon_L\right)$. For each $\mathsf{b} \in \mathcal{B}$,

$$\begin{aligned}\left|P\left(x,t\right)'\mathsf{b} - B\left(\alpha+ht|x,I\right)\right| &\le \left|P\left(x,t\right)'\mathsf{b}\left(\alpha|I\right) - B\left(\alpha+ht|x,I\right)\right| \\ &\quad + \left\|P\left(x,t\right)\right\|\left\|\mathsf{b}-\mathsf{b}\left(\alpha|I\right)\right\| \\ &= O\left(\left\{1+\max_{x\in\mathcal{X}}\left\|P\left(x\right)\right\|\right\}\epsilon_L\right) = o\left(h\right),\end{aligned}$$

uniformly by (B.0.3), so that

$$\max_{\mathsf{b}\in\mathcal{B}}\max\left|P\left(x,t\right)'\mathsf{b} - B\left(\alpha+ht|x,I\right)\right| = o\left(h\right), \tag{B.0.7}$$

where the second max is as in (B.0.4). For condition (ii), recall that, for $\pi\left(t\right) = \left[1,\ldots,t^{s+1}\right]'$,

$$\mathsf{R}^{(2)}\left(\mathsf{b};\alpha,I\right) = \mathbb{E}\left[\mathbb{I}\left(I_\ell=I\right)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} g\left(P\left(x_\ell,t\right)'\mathsf{b}|x_\ell,I_\ell\right)P\left(x_\ell\right)P\left(x_\ell\right)'\otimes\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt\right].$$

To deal with the discontinuities of $y \to g\left(y|x_\ell,I_\ell\right)$, let $\delta_t\left(\cdot\right)$ be the Dirac mass at t. By Jones (2008, Theorem 7.40), the generalized derivative $\mathsf{R}^{(3)}\left(\mathsf{b};\alpha,I\right)$ of $\mathsf{b}\to\mathsf{R}^{(2)}\left(\mathsf{b};\alpha,I\right)$ is the linear function taking value in the space of squared matrices with dimension $\left(s+2\right)K_L$, that is for any vector Δ in $\mathbb{R}^{(s+2)K_L}$

$$\begin{aligned}&\mathsf{R}^{(3)}\left(\mathsf{b};\alpha,I\right)\left(\Delta\right) \\ &= \mathbb{E}\Bigg[\mathbb{I}\left(I_\ell=I\right)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left(P\left(x_\ell,t\right)'\Delta\right)g^{(1)}\left(P\left(x_\ell,t\right)'\mathsf{b}|x_\ell,I_\ell\right) \\ &\qquad \times P\left(x_\ell\right)P\left(x_\ell\right)'\otimes\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt\Bigg] \\ &+ \mathbb{E}\Bigg[\mathbb{I}\left(I_\ell=I\right)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left(P\left(x_\ell,t\right)'\Delta\right)\delta_{V(1|x_\ell,I_\ell)}\left(P\left(x_\ell,t\right)'\mathsf{b}\right)g\left(V\left(1|x_\ell,I_\ell\right)|x_\ell,I_\ell\right) \\ &\qquad P\left(x_\ell\right)P\left(x_\ell\right)'\otimes\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt\Bigg] \\ &- \mathbb{E}\Bigg[\mathbb{I}\left(I_\ell=I\right)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left(P\left(x_\ell,t\right)'\Delta\right)\delta_{V(0|x_\ell,I_\ell)}\left(P\left(x_\ell,t\right)'\mathsf{b}\right)g\left(V\left(0|x_\ell,I_\ell\right)|x_\ell,I_\ell\right) \\ &\qquad P\left(x_\ell\right)P\left(x_\ell\right)'\otimes\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt\Bigg].\end{aligned}$$

The items involving a Dirac mass can be written as, focusing on the term with $\delta_{V(1|x_\ell,I_\ell)}(\cdot)$ without loss of generality,

$$\mathbb{E}\left[\mathbb{I}(I_\ell = I) \sum_{t\in\left[-\frac{\alpha}{h},\frac{1-\alpha}{h}\right];P(x_\ell,t)'\mathsf{b}=V(1|x_\ell,I_\ell)} \left(P(x_\ell,t)'\Delta\right) g\left(V(1|x_\ell,I_\ell)|x_\ell,I_\ell\right) \right.$$
$$\left. P(x_\ell)P(x_\ell)' \otimes \pi(t)\pi(t)'K(t)\right].$$

Because $B(\alpha+ht|x,I)$ is strictly increasing in $t\in[-\alpha/h,(1-\alpha)/h]$ by Proposition 3.1-(i) and by (B.0.7), the equation $P(x_\ell,t)'\mathsf{b} = V(1|x_\ell,I_\ell)$ has at most $s+1$ solutions for all $\mathsf{b}\in\mathcal{B}$. Since the matrix $\pi(t)\pi(t)'K(t)$ has a finite dimension and bounded entries, it follows from the boundedness of $g(\cdot|\cdot,\cdot)$ and Assumption H that

$$\mathsf{R}^{(3)}(\mathsf{b};\alpha,I)(\Delta) \preceq C\max_{x\in\mathcal{X}}\|P(x)\|\,\|\Delta\|\,\mathbb{E}\left[P(x_\ell)P(x_\ell)'\right]\otimes \mathrm{Id}_{s+2}. \tag{B.0.8}$$

Assumption R-(ii) implies that the operator norm of $\mathsf{R}^{(3)}(\mathsf{b};\alpha,I)$ is bounded by $C\max_{x\in\mathcal{X}}\|P(x)\|$ uniformly in $\mathsf{b}\in\mathcal{B}$ and $\alpha\in[0,1]$. Hence there is a constant c_1 such that, for all $\alpha\in[0,1]$,

$$\left\|\mathsf{R}^{(2)}(\mathsf{b}_1;\alpha,I)-\mathsf{R}^{(2)}(\mathsf{b}_0;\alpha,I)\right\| \le c_1\max_{x\in\mathcal{X}}\|P(x)\|\,\|\mathsf{b}_1-\mathsf{b}_0\| \text{ for all } \mathsf{b}_0,\,\mathsf{b}_1\in\mathcal{B}.$$

Now, since (B.0.3) ensures that $C_0^2 c_1\max_{x\in\mathcal{X}}\|P(x)\|\epsilon_L\to 0$, it follows by Lemma B.2 that, for L large enough and all $\alpha\in[0,1]$, $\mathsf{R}(\cdot;\alpha,I)$ has a unique minimizer $\mathsf{b}_{ASQR}(\alpha|I)$ over $\mathcal{B}$, and therefore of the all space since $\mathsf{R}(\cdot;\alpha,I)$ is convex. Moreover, by definition of $\mathcal{B}=\mathcal{B}(\mathsf{b}(\alpha|I),2C_0\epsilon_L)$,

$$\max_{\alpha\in[0,1]}\|\mathsf{b}_{ASQR}(\alpha|I)-\mathsf{b}(\alpha|I)\| \le 2C_0\epsilon_L. \tag{B.0.9}$$

Recall that $V_{ASQR}(\alpha|x,I)=B_{ASQR}(\alpha|x,I)+\alpha B^{(1)}_{ASQR}(\alpha|x,I)/(I-1)$ with

$$B_{ASQR}(\alpha|x,I)=P(x)'\mathsf{b}_{0,ASQR}(\alpha|I),\quad B^{(1)}_{ASQR}(\alpha|x,I)=h^{-1}P(x)'\mathsf{b}_{1,ASQR}(\alpha|I).$$

The definition of $\mathsf{b}(\alpha|I)$ and (B.0.4), (B.0.7), Assumptions S-(ii) and R-(ii) give that,

$$\max_{\alpha\in[0,1]}\int_{\mathcal{X}}\left(B_{ASQR}\left(\alpha|x,I\right)-B\left(\alpha|x,I\right)\right)^{2}dx=O\left(\epsilon_{L}^{2}\right),$$
$$\max_{(\alpha+ht,x)\in[0,1]\times\mathcal{X}}\left|P\left(x,t\right)'\mathsf{b}_{ASQR}\left(\alpha|I\right)-B\left(\alpha+ht|x,I\right)\right|=O\left(\epsilon_{L}\max_{x\in}\left\|P\left(x\right)\right\|\right)=o\left(h\right). \tag{B.0.10}$$

We now study $\int_{\mathcal{X}}\left(\alpha B_{ASQR}^{(1)}\left(\alpha|x,I\right)-\alpha B^{(1)}\left(\alpha|x,I\right)\right)^{2}dx$. $\mathsf{b}_{ASQR}\left(\alpha|I\right)=\arg\max_{\mathsf{b}}\mathsf{R}\left(\cdot;\alpha,I\right)$ satisfies the first-order condition $\mathsf{R}^{(1)}\left(\mathsf{b}_{ASQR}\left(\alpha|I\right);\alpha,I\right)=0$ which can be rearranged as follows. Define

$$\mathfrak{g}\left(\alpha|t,x,I\right)=\int_{0}^{1}g\left(P\left(x,t\right)'\mathsf{b}_{ASQR}\left(\alpha|I\right)+u\left(B\left(\alpha+ht|x,I\right)-P\left(x,t\right)'\mathsf{b}_{ASQR}\left(\alpha|I\right)\right)\right)du.$$

Arguing as in the proof of Lemma B.1 gives that $\min_{(\alpha+ht,x)\in[0,1]\times\mathcal{X}}\mathfrak{g}\left(\alpha|t,x,I\right)\geq C>0$ for L large enough. Since $\alpha+ht=G\left(B\left(\alpha+ht|x,I\right)|x,I\right)$, $\mathsf{R}^{(1)}\left(\mathsf{b}_{ASQR}\left(\alpha|I\right);\alpha,I\right)=0$ and the definition of $\mathfrak{g}\left(\alpha|t,x,I\right)$ gives, for any $0<\eta<1$,

$$\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\mathfrak{g}\left(\alpha|t,x_{\ell},I\right)P\left(x_{\ell},t\right)P\left(x_{\ell},t\right)'K\left(t\right)dt\right]\left\{\alpha\mathsf{b}_{ASQR}\left(\alpha|I\right)-\alpha\mathsf{b}\left(\alpha|I\right)\right\}$$
$$=\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\int_{-\frac{\alpha}{(1+\eta)h}}^{\frac{1+\alpha}{h}}\mathfrak{g}\left(\alpha|t,x_{\ell},I\right)P\left(x_{\ell},t\right)\left\{\alpha B\left(\alpha+ht|x_{\ell},I\right)-P\left(x_{\ell},t\right)'\alpha\mathsf{b}\left(\alpha|I\right)\right\}K\left(t\right)dt\right] \tag{B.0.11}$$
$$+\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\int_{-\frac{\alpha}{h}}^{-\frac{\alpha}{(1+\eta)h}}\mathfrak{g}\left(\alpha|t,x_{\ell},I\right)P\left(x_{\ell},t\right)\left\{\alpha B\left(\alpha+ht|x_{\ell},I\right)-P\left(x_{\ell},t\right)'\alpha\mathsf{b}\left(\alpha|I\right)\right\}K\left(t\right)dt\right] \tag{B.0.12}$$

For (B.0.12), observe that this item differs fom 0 only when $\alpha\leq h$. Hence a Taylor expansion gives that, by (B.0.9) and arguing as for (B.0.6),

$$\left\|\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\int_{-\frac{\alpha}{h}}^{-\frac{\alpha}{(1+\eta)h}}\mathfrak{g}\left(\alpha|t,x_{\ell},I\right)P\left(x_{\ell},t\right)\left\{\alpha B\left(\alpha+ht|x_{\ell},I\right)-P\left(x_{\ell},t\right)'\alpha\mathsf{b}\left(\alpha|I\right)\right\}K\left(t\right)dt\right]\right\|$$
$$\leq C\eta hO\left(\epsilon_{L}\right)=\eta\left\{o\left(h^{s+2}\right)+o\left(hK_{L}^{-(s+2)/d_{\mathcal{M}}}\right)\right\}.$$

For (B.0.11), note that $\alpha + ht \geq \eta\alpha/(1+\eta)$ when $t \in [-\alpha/((1+\eta)h), (1-\alpha)/h]$. Assumption S-(ii), a corresponding Taylor expansion of order $s+2$ and Proposition 3.1 give

$$
\begin{aligned}
&\max_{(\alpha,x)\in(0,1]\times\mathcal{X}} \max_{t\in\left[-\frac{\alpha}{(1+\eta)h},\frac{1-\alpha}{h}\right]} \left|\alpha B(\alpha+ht|x,I) - P(x,t)'\alpha\mathsf{b}(\alpha|I)\right| \\
&\leq Ch^{s+2} \max_{(a,x)\in[0,1]\times\mathcal{X}} \left|\frac{1+\eta}{\eta} aB^{(s+2)}(a|x,I)\right| + O\left(K_L^{-(s+2)/d_{\mathcal{M}}}\right) \\
&= \eta^{-1}O\left(h^{s+2} + K_L^{-(s+2)/d_{\mathcal{M}}}\right). \qquad \text{(B.0.13)}
\end{aligned}
$$

Hence $\min_{(\alpha+ht,x)\in[0,1]\times\mathcal{X}} \mathfrak{g}(\alpha|t,x,I) \geq C$ and Assumption R-(ii) give, arguing as for (B.0.6),

$$
\max_{\alpha\in[0,1]} \|\alpha\mathsf{b}_{ASQR}(\alpha|I) - \alpha\mathsf{b}(\alpha|I)\| = O\left(h^{s+2} + K_L^{-(s+2)/d_{\mathcal{M}}}\right), \qquad \text{(B.0.14)}
$$

so that, since $B^{(1)}_{ASQR}(\alpha|x,I) = P(x)'\,\mathsf{b}_{1,ASQR}(\alpha|I)/h$ and by Assumption R-(ii)

$$
\max_{\alpha\in[0,1]} \int_{\mathcal{X}} \left(\alpha B^{(1)}_{ASQR}(\alpha|x,I) - \alpha B^{(1)}(\alpha|x,I)\right)^2 dx = O\left(h^{s+1} + \frac{K_L^{-(s+2)/d_{\mathcal{M}}}}{h}\right),
$$

which, together with the equation above (B.0.10), implies the first order stated in Theorem B.1-(ii). The uniform rate of Theorem B.1-(i) follows from (B.0.9) and (B.0.14) since the power of K_L can be removed for the ASR case and the Euclidean norm operates over a space with finite dimension in the case of the quantile regression specification (3.2.7).Moreover, since $V(\alpha|x,I) = X\gamma(\alpha|I)$ is $(s+1)$ times continuously differentiable, it follows from (B.0.11), a suitable choice of $\eta \to 0$ and Proposition 3.1 that, since $g(B(\alpha|x,I)|x,I) = 1/B^{(1)}(\alpha|x,I)$ and for all $\alpha \in [0,1]$

$$
\begin{aligned}
\alpha\mathsf{b}_{AQR}(\alpha|I) - \alpha\mathsf{b}(\alpha|I) = &\left(\mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)X_\ell X_\ell'}{B^{(1)}(\alpha|x_\ell,I)} \otimes \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi(t)\pi(t)'K(t)\,dt\right] + o(1)\right)^{-1} \\
&\times \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)X_\ell \alpha B^{(s+2)}(\alpha|x_\ell,I)}{B^{(1)}(\alpha|x_\ell,I)} \otimes \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \frac{t^{s+2}\pi(t)}{(s+2)!}K(t)\,dt\right] h^{s+2} + o\left(h^{s+2}\right),
\end{aligned}
$$
(B.0.15)

which gives the second expansion in Theorem B.1-(i) since $h\beta_1 = S_1\mathsf{b}$ and because the estimation of $\alpha B^{(1)}(\alpha|x,I)$ gives the dominant bias term.

Consider now $\max|V_{ASQR}(\alpha|x,I) - V(\alpha|x,I)|$ for which we only detail how to bound $\max\left|\alpha B^{(1)}_{ASQR}(\alpha|x,I) - \alpha B^{(1)}(\alpha|x,I)\right|$ since the study of $\max|B_{ASQR}(\alpha|x,I) - B(\alpha|x,I)|$ can be similarly done. Let $A = [A_1, \dots A_{J_1}]$ be a a $J_2 \times J_1$ matrix with columns A_j, $j = 1, \dots, J_1$, $|A_j|_1$ the associated ℓ_1 norm and $|A|_{1,\infty} = \max_{j \le J_1} |A_j|_1$, a, b some conformable vectors so that,

$$|a'Ab| = \left|\sum_{j=1}^{J} b_j a' A_j\right| \le |b|_1 |A|_{1,\infty} \max_{j \le J_2} |a_j| .$$

Let S_1 be the selection matrix such that $S_1\mathsf{b} = \mathsf{b}_1$. (B.0.11), (B.0.13) with $\min_{(\alpha+ht,x)\in[0,1]\times\mathcal{X}} \mathfrak{g}(\alpha|t,x,I) \ge C$, $P(x,t) = P(x) \otimes \pi(t)$ give by Assumptions A and H-(iii) and the bound above,

$$\begin{aligned}
&\left|\alpha\left(B^{(1)}_{ASQR}(\alpha|x,I) - B^{(1)}(\alpha|x,I)\right)\right| \\
&= h^{-1}\left|P(x)' S_1 \mathbb{E}^{-1}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \mathfrak{g}(\alpha|t,x_\ell,I) P(x_\ell,t) P(x_\ell,t)' K(t)\,dt\right]\right. \\
&\times \mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \mathfrak{g}(\alpha|t,x_\ell,I) P(x_\ell,t)\right. \\
&\quad \times \left.\left\{\alpha B(\alpha+ht|x_\ell,I) - P(x_\ell,t)'\alpha\mathsf{b}(\alpha|I)\right\} K(t)\,dt\right] \\
&\le C\left(\frac{h^{s+2} + K_L^{-\frac{s+2}{d_{\mathcal{M}}}}}{h}\right)\left(\sum_{k=1}^{K_L} |P_k(x)|\right)\left|\mathbb{E}^{-1}\left[P(x_\ell) P(x_\ell)'\right]\right|_{1,\infty} \max_{k \le K_L}\left|\int_{\mathcal{X}} |P_k(x)|\,dx\right| \\
&\le C\left(h^{s+1} + \frac{K_L^{-\frac{s+2}{d_{\mathcal{M}}}}}{h}\right),
\end{aligned}$$

uniformly in α and x. This ends the proof of (iii). $\square$

Bahadur representation for the AQR and ASQR estimators

The notations $\mathsf{b}_{AQR}(\alpha|I)$, $\mathsf{b}_{ASQR}(\alpha|I)$, $\mathsf{b}(\alpha|I)$, b from (B.0.1), which renormalizes the parameter β, and $P(x,t)$, $\mathsf{R}_{AQR}(\mathsf{b};\alpha,I)$, $\mathsf{R}_{ASQR}(\mathsf{b};\alpha,I)$ from the proof of

Proposition B.1, will be used here. Define also $\widehat{\mathsf{R}}_{AQR}(\mathsf{b};\alpha,I)$, $\widehat{\mathsf{R}}_{ASQR}(\mathsf{b};\alpha,I)$ as the corresponding $\widehat{\mathcal{R}}_{AQR}(\mathsf{b};\alpha,I)$, $\widehat{\mathcal{R}}_{ASQR}(\mathsf{b};\alpha,I)$ and

$$\widehat{\mathsf{b}}_{AQR}(\alpha|I) = \arg\min_{\mathsf{b}} \widehat{\mathsf{R}}_{AQR}(\mathsf{b};\alpha,I), \quad \widehat{\mathsf{b}}_{ASQR}(\alpha|I) = \arg\min_{\mathsf{b}} \widehat{\mathsf{R}}_{ASQR}(\mathsf{b};\alpha,I),$$
$$\widehat{\mathsf{e}}_{AQR}(\alpha|I) = -\left(\mathsf{R}^{(2)}_{AQR}(\mathsf{b}_{AQR}(\alpha|I);\alpha,I)\right)^{-1}\widehat{\mathsf{R}}^{(1)}_{AQR}(\mathsf{b}_{AQR}(\alpha|I);\alpha,I),$$
$$\widehat{\mathsf{e}}_{ASQR}(\alpha|I) = -\left(\mathsf{R}^{(2)}_{ASQR}(\mathsf{b}_{ASQR}(\alpha|I);\alpha,I)\right)^{-1}\widehat{\mathsf{R}}^{(1)}_{ASQR}(\mathsf{b}_{ASQR}(\alpha|I);\alpha,I),$$

$$\widehat{\mathsf{d}}_{AQR}(\alpha|I) = \widehat{\mathsf{b}}_{AQR}(\alpha|I) - \mathsf{b}_{AQR}(\alpha|I) - \widehat{\mathsf{e}}_{AQR}(\alpha|I), \tag{B.0.16}$$
$$\widehat{\mathsf{d}}_{ASQR}(\alpha|I) = \widehat{\mathsf{b}}_{ASQR}(\alpha|I) - \mathsf{b}_{ASQR}(\alpha|I) - \widehat{\mathsf{e}}_{ASQR}(\alpha|I). \tag{B.0.17}$$

The aim of this section is to study the Bahadur remainder terms $\widehat{\mathsf{d}}_{AQR}(\alpha|I)$, $\widehat{\mathsf{d}}_{ASQR}(\alpha|I)$ and the score terms $\widehat{\mathsf{R}}^{(1)}_{AQR}(\mathsf{b}_{AQR}(\alpha|I);\alpha,I)$, $\widehat{\mathsf{R}}^{(1)}_{ASQR}(\mathsf{b}_{AQR}(\alpha|I);\alpha,I)$, that is to establish that,

Theorem B.2. *Suppose that Assumptions A and H-(i) hold and let $\widehat{\mathsf{d}}_{AQR}(\alpha|I)$, $\widehat{\mathsf{d}}_{ASQR}(\alpha|I)$ be as in (B.0.16) and (B.0.17). Then,*

i. If the quantile regression model in Assumption S-(i) is the true one and under Assumption R-(i),

$$\max_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left\|\widehat{\mathsf{d}}_{AQR}(\alpha|I)\right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{3/4}\right);$$

ii. For the sieve quantile regression model (3.2.9) and under Assumptions S-(ii), R-(ii), (H)-(ii), and for $\overline{P}_L = \max_{x\in\mathcal{X}}\|P(x)\|$, it holds that,

$$\max_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left\|\widehat{\mathsf{d}}_{ASQR}(\alpha|I)\right\| = O_{\mathbb{P}}\left(\overline{P}_L^{1/2}\left(\frac{K_L\log L}{L}\right)^{3/4}\right) + O_{\mathbb{P}}\left(\frac{K_L\overline{P}_L\log L}{L}\right).$$

Horowitz and Lee (2005) have derived a similar expansion for a fixed α in the case where $\overline{P}_L = O\left(K_L^{1/2}\right)$ in their Theorem 2. The rate of Theorem B.2-(ii) is slightly better since instead of the order $\overline{P}_L^{1/2}(K_L\log L/L)^{3/4}$ they obtained

$\overline{P}_L^{3/2}\left(K_L \log L/L\right)^{3/4}$ up to a logarithmic term. Because K_L plays a role similar to a bandwidth term $1/h^d$, the rate of Theorem B.2-(ii) is slightly worse than the one derived in Guerre and Sabbah (2012) for local polynomial conditional quantile due to the item $\overline{P}_L^{1/2}$ in front of $\left(K_L \log L/L\right)^{3/4}$. This term is used as a bound for $P\left(x,t\right)'\theta$ for all $(x,t) \in \mathcal{X} \times [-1,1]$ and θ in the unit sphere of dimension $K_L\left(s+2\right)$, where θ plays the role of the normalized estimated parameter. A better bound can therefore be obtained by imposing some smoothness constraints on θ, for instance by imposing that the $x-$derivatives of $P\left(x,t\right)'\theta$ must stay bounded in quadratic mean or by using a corresponding penalty term, see e.g, Chen (2007) eq. (2.18). That sieve estimators must be constrained to satisfy some smoothness conditions is a key difference with kernel smoothing where this holds asymptotically by uniform convergence of derivatives of the kernel estimator.

The next lemma derives the order of the Euclidean norm of $\widehat{\mathsf{R}}_{AQR}^{(1)}\left(\mathsf{b}_{AQR}\left(\alpha|I\right);\alpha,I\right)$ and $\widehat{\mathsf{R}}_{ASQR}^{(1)}\left(\mathsf{b}_{ASQR}\left(\alpha|I\right);\alpha,I\right)$. Let $\preceq$ be the usual order for symmetric matrices, i.e $A \preceq B$ if $B-A$ is semidefinite positive. The proofs of Lemma B.3, Propositions B.1 and B.2 are given in the last subsection B of this proof section.

Lemma B.3. *Suppose that Assumptions A and H-(i) hold. Then,*

i. If the quantile regression model in Assumption S-(i) is the true one and under Assumption R-(i),

$$\operatorname{Var}\left(\widehat{\mathsf{R}}_{AQR}^{(1)}\left(\mathsf{b}_{AQR}\left(\alpha|I\right);\alpha,I\right)\right) \preceq CL^{-1/2}\,\mathrm{Id}_{(s+2)(d+1)},$$

for all $(\alpha,I) \in [0,1] \times \mathcal{I}$ *and,*

$$\mathbb{E}\left[\max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|\widehat{\mathsf{R}}_{AQR}^{(1)}\left(\mathsf{b}_{AQR}\left(\alpha|I\right);\alpha,I\right)\right\|\right] \leq C\left(\frac{\log L}{L}\right)^{1/2};$$

ii. For the sieve quantile regression model (3.2.9) and under Assumptions S-(ii), R-(ii),

$$\operatorname{Var}\left(\widehat{\mathsf{R}}_{AQR}^{(1)}\left(\mathsf{b}_{AQR}\left(\alpha|I\right);\alpha,I\right)\right) \preceq CL^{-1/2}\,\mathrm{Id}_{K_L(s+2)},$$

for all $(\alpha,I) \in [0,1] \times \mathcal{I}$ *. If (B.0.3) is true and for* $\overline{P}_L = \max_{x\in\mathcal{X}}\left\|P\left(x\right)\right\|$*, it* *holds that,*

$$\mathbb{E}\left[\max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|\widehat{\mathsf{R}}^{(1)}_{ASQR}\left(\mathsf{b}_{ASQR}\left(\alpha|I\right);\alpha,I\right)\right\|\right]\leq C\left(\left(\frac{K_L\log L}{L}\right)^{1/2}+\frac{K_L\overline{P}_L\log L}{L}\right).$$

We now introduce some linearization remainders for $\widehat{\mathsf{R}}_{AQR}(\cdot;\alpha,I)$ and $\widehat{\mathsf{R}}_{ASQR}(\cdot;\alpha,I)$ which are useful to study $\widehat{\mathsf{d}}_{AQR}(\alpha|I)$ and $\widehat{\mathsf{d}}_{ASQR}(\alpha|I)$. Observe that,

$$\widehat{\mathsf{d}}_{AQR}\left(\alpha|I\right)=\arg\min_{\mathsf{d}}\widehat{\mathsf{R}}_{AQR}\left(\mathsf{b}_{AQR}\left(\alpha|I\right)+\widehat{\mathsf{e}}_{AQR}\left(\alpha|I\right)+\mathsf{d};\alpha,I\right).$$

By Lemmas B.3 and B.1, $\widehat{\mathsf{e}}_{AQR}(\alpha|I)$ goes to 0. This leads to consider the following remainder terms, where e is a new variable which would be set equal to $\widehat{\mathsf{e}}_{AQR}(\alpha|I)$ or $\widehat{\mathsf{e}}_{ASQR}(\alpha|I)$ accordingly,

$$\begin{aligned}\widehat{\mathsf{O}}_{AQR}\left(\mathsf{d},\mathsf{e};\alpha,I\right)&=\widehat{\mathsf{R}}_{AQR}\left(\mathsf{b}_{AQR}\left(\alpha|I\right)+\mathsf{e}+\mathsf{d};\alpha,I\right)-\widehat{\mathsf{R}}_{AQR}\left(\mathsf{b}_{AQR}\left(\alpha|I\right)+\mathsf{e};\alpha,I\right)\\&\quad-\widehat{\mathsf{R}}^{(1)}_{AQR}\left(\mathsf{b}_{AQR}\left(\alpha|I\right);\alpha,I\right)'\mathsf{d},\\\widehat{\mathsf{O}}_{1,AQR}\left(\mathsf{d},\mathsf{e};\alpha,I\right)&=\widehat{\mathsf{O}}_{AQR}\left(\mathsf{d},\mathsf{e};\alpha,I\right)-\mathbb{E}\left[\widehat{\mathsf{O}}_{AQR}\left(\mathsf{d},\mathsf{e};\alpha,I\right)\right],\\\mathsf{O}_{1,AQR}\left(\mathsf{d},\mathsf{e};\alpha,I\right)&=\mathbb{E}\left[\widehat{\mathsf{O}}_{AQR}\left(\mathsf{d},\mathsf{e};\alpha,I\right)\right],\end{aligned}$$

$\widehat{\mathsf{O}}_{ASQR}$, $\widehat{\mathsf{O}}_{1,ASQR}$ and $\mathsf{O}_{1,ASQR}$ being defined as $\widehat{\mathsf{O}}_{AQR}$, $\widehat{\mathsf{O}}_{1,AQR}$ and $\mathsf{O}_{1,AQR}$.

Proposition B.1. *Let $\mathcal{B}(0,\varrho)$ be the Euclidean closed ball of center 0 and radius ϱ, and consider $\varrho_d,\varrho_e>0$ which can depend upon L but must stay bounded away from infinity. Suppose that Assumptions A and H-(i) hold. Then,*

i. If the quantile regression model in Assumption S-(i) is the true one and under Assumption R-(i),

$$\begin{aligned}&\mathbb{E}\left[\max_{(\alpha,\mathsf{d},\mathsf{e},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{B}(0,\varrho_e)\times\mathcal{I}}\left|\widehat{\mathsf{O}}_{1,AQR}\left(\mathsf{d},\mathsf{e};\alpha,I\right)\right|\right]\\&\qquad\leq C\left(\frac{\log L}{L}\right)^{1/2}\varrho_d\left(\left(\varrho_d+\varrho_e\right)^{1/2}+\left(\frac{\log L}{L}\right)^{1/2}\right);\end{aligned}$$

ii. For the sieve quantile regression model (3.2.9) and under Assumptions S-(ii), R-(ii), (H)-(ii), and for $\overline{P}_L = \max_{x \in \mathcal{X}} \|P(x)\|$*, it holds that,*

$$\mathbb{E}\left[\max_{(\alpha,\mathsf{d},\mathsf{e},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{B}(0,\varrho_e)\times\mathcal{I}} \left|\widehat{\mathsf{O}}_{1,ASQR}(\mathsf{d},\mathsf{e};\alpha,I)\right|\right]$$
$$\leq C\left(\frac{K_L\overline{P}_L\log L}{L}\right)^{1/2}\varrho_d\left((\varrho_d+\varrho_e)^{1/2}+\left(\frac{K_L\overline{P}_L\log L}{L}\right)^{1/2}\right).$$

Proposition B.2. *Let* $\mathcal{B}(0,\varrho)$ *be the Euclidean closed ball of center* 0 *and radius* ϱ*, and consider* $\varrho_d, \varrho_e > 0$ *which can depend upon* L *but must stay bounded away from infinity. Suppose that Assumptions A and H-(i) hold. Then,*

i. If the quantile regression model in Assumption S-(i) is the true one and under Assumption R-(i),

$$\max_{(\alpha,\mathsf{d},\mathsf{e},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{B}(0,\varrho_e)\times\mathcal{I}} \left|\mathsf{O}_{1,AQR}(\mathsf{d},\mathsf{e};\alpha,I) - \frac{\mathsf{d}'\mathsf{R}^{(2)}_{AQR}(\mathsf{b}_{AQR}(\alpha|I);\alpha,I)(\mathsf{d}+2\mathsf{e})}{2}\right|$$
$$= O\left((\varrho_d+\varrho_e)^2\varrho_d\right);$$

ii. For the sieve quantile regression model (3.2.9) and under Assumptions S-(ii), R-(ii), (H)-(ii), and for $\overline{P}_L = \max_{x \in \mathcal{X}} \|P(x)\|$*, it holds that,*

$$\max_{(\alpha,\mathsf{d},\mathsf{e},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{B}(0,\varrho_e)\times\mathcal{I}} \left|\mathsf{O}_{1,ASQR}(\mathsf{d},\mathsf{e};\alpha,I) - \frac{\mathsf{d}'\mathsf{R}^{(2)}_{ASQR}(\mathsf{b}_{ASQR}(\alpha|I);\alpha,I)(\mathsf{d}+2\mathsf{e})}{2}\right|$$
$$= O\left(\overline{P}_L(\varrho_d+\varrho_e)^2\varrho_d\right).$$

Proof of Theorem B.2. It is sufficient to study the ASQR case and for the sake of brevity the subscript ASQR is removed from the notations. Define for some C_d, $C_e \geq 1$ to be chosen large enough,

$$\varrho_e = C_e\left(\frac{K_L\log L}{L}\right)^{1/2}$$
$$\varrho_d = \left(\frac{K_L\log L}{L}\right)^{1/2} r_d \text{ with } r_d = C_d\max\left\{\overline{P}_L^{1/2}\left(\frac{K_L\log L}{L}\right)^{1/4}, \overline{P}_L\left(\frac{K_L\log L}{L}\right)^{1/2}\right\}.$$

Under Assumption H-(ii), ϱ_e, ϱ_d and r_d go to 0 when $L \to \infty$. Observe that Lemmas B.1 and B.3 give under Assumption H-(ii), by the Markov inequality and for $\underline{L}$ large enough,

$$\lim_{C_e \to \infty} \sup_{L \geq \underline{L}} \mathbb{P}\left(\max_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left\| \widehat{\mathsf{e}}\left(\alpha|I\right) \right\| \geq \varrho_e \right) = 0. \tag{B.0.18}$$

Define

$$\begin{aligned}
\widehat{\mathsf{R}}_0\left(\mathsf{d},\mathsf{e}\right) &= \widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)'\mathsf{d} + \frac{\mathsf{d}'\mathsf{R}^{(2)}\left(\mathsf{b}\left(\alpha|I_\ell\right);\alpha,I\right)\left(\mathsf{d}+2\mathsf{e}\right)}{2},\\
\widehat{\mathsf{R}}_c\left(\mathsf{d},\mathsf{e}\right) &= \widehat{\mathsf{R}}\left(\mathsf{b}\left(\alpha|I\right)+\mathsf{e}+\mathsf{d};\alpha,I\right) - \widehat{\mathsf{R}}\left(\mathsf{b}\left(\alpha|I\right)+\mathsf{e};\alpha,I\right),\\
\widehat{\mathsf{O}}_2\left(\mathsf{d},\mathsf{e}\right) &= \widehat{\mathsf{R}}_c\left(\mathsf{b}\left(\alpha|I\right)+\mathsf{e}+\mathsf{d};\alpha,I\right) - \widehat{\mathsf{R}}_0\left(\mathsf{d},\mathsf{e}\right).
\end{aligned}$$

Propositions B.1 and B.2 give, under Assumption H-(ii)

$$\mathbb{E}\left[\max_{(\alpha,\mathsf{d},\mathsf{e},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{B}(0,\varrho_e)\times\mathcal{I}} \left|\widehat{\mathsf{O}}_2\left(\mathsf{d},\mathsf{e}\right)\right|\right] \leq C\left\{\overline{P}_L\varrho_e^2\varrho_d + \left(\frac{K_L\overline{P}_L\log L}{L}\right)^{1/2}\varrho_e\varrho_d\right\}. \tag{B.0.19}$$

Recall also that,

$$\widehat{\mathsf{d}}\left(\alpha|I\right) = \arg\min_{\mathsf{d}} \widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right).$$

Let also $\Lambda = \Lambda_{ASQR}$ be as in Lemma B.1. Consider some arbitrary $\epsilon > 0$.

Step 1: Comparing $\max\left|\widehat{\mathsf{O}}_2\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\right|$ *and* ϱ_d^2. For C_e large enough and by (B.0.18) and (B.0.19),

$$\begin{aligned}
&\mathbb{P}\left(\max_{(\alpha,\mathsf{d},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{I}} \left|\widehat{\mathsf{O}}_2\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\right| \geq \frac{\varrho_d^2}{\Lambda}\right)\\
&\quad\leq \frac{\epsilon}{2} + \mathbb{P}\left(\max_{(\alpha,\mathsf{d},\mathsf{e},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{B}(0,\varrho_e)\times\mathcal{I}} \left|\widehat{\mathsf{O}}_2\left(\mathsf{d},\mathsf{e}\right)\right| \geq \frac{\varrho_d^2}{\Lambda}\right) \leq \frac{\epsilon}{2} + C\frac{\Lambda}{\varrho_d^2}\frac{C_e\varrho_d^2}{C_d},
\end{aligned}$$

which in turn can be made arbitrarily small by increasing C_d.[51] Hence

$$\mathbb{P}\left(\max_{(\alpha,\mathsf{d},I)\in[0,1]\times\mathcal{B}(0,\varrho_d)\times\mathcal{I}} \left|\widehat{\mathsf{O}}_2\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\right| \geq \frac{\varrho_d^2}{\Lambda}\right) \leq \epsilon. \tag{B.0.20}$$

[51] Take for instance $C_d = C_e^2$ here.

Step 2: order of $\sup_{\alpha\in[0,1]}\left\|\widehat{\mathsf{d}}\left(\alpha|I\right)\right\|$. Note that

$$\left\{\left\|\widehat{\mathsf{d}}\left(\alpha|I\right)\right\|\geq\varrho_d\right\}\subset\left\{\min_{\mathsf{d};\|\mathsf{d}\|\geq\varrho_d}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\leq\min_{\mathsf{d}\in\mathcal{B}(0,\varrho_d)}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\right\}$$

We first recall a convexity argument due to Pollard (1991) which allows to change $\{\mathsf{d};\|\mathsf{d}\|\geq\varrho_d\}$ into $\mathcal{S}\left(\varrho_d\right)=\{\mathsf{d};\|\mathsf{d}\|=\varrho_d\}$ in the inclusion above. Recall that

$$\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)=\widehat{\mathsf{R}}_0\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)+\widehat{\mathsf{O}}_2\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right),$$

where both $\mathsf{d}\longmapsto\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)$, $\widehat{\mathsf{R}}_0\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)$ are convex with, by definition of $\widehat{\mathsf{e}}\left(\alpha|I\right)$,

$$\begin{aligned}\widehat{\mathsf{R}}_0\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)&=\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)'\mathsf{d}+\frac{\mathsf{d}'\mathsf{R}^{(2)}\left(\mathsf{b}\left(\alpha|I_\ell\right);\alpha,I\right)\left(\mathsf{d}+2\widehat{\mathsf{e}}\left(\alpha|I\right)\right)}{2}\\&=\frac{\mathsf{d}'\mathsf{R}^{(2)}\left(\mathsf{b}\left(\alpha|I_\ell\right);\alpha,I\right)\mathsf{d}}{2}.\end{aligned}$$

Consider a δ with $\|\delta\|=1$ and $\mathsf{d}=\varrho\delta$ for some $\varrho\geq\varrho_d$. Since $\widehat{\mathsf{R}}_c\left(\mathsf{0},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)=0$, the convexity inequality implies

$$\begin{aligned}\frac{\varrho_d}{\varrho}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)&=\frac{\varrho_d}{\varrho}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)+\left(1-\frac{\varrho_d}{\varrho}\right)\widehat{\mathsf{R}}_c\left(\mathsf{0},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\\&\geq\widehat{\mathsf{R}}_c\left(\frac{\varrho_d}{\varrho}\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)=\widehat{\mathsf{R}}_c\left(\varrho_d\delta,\widehat{\mathsf{e}}\left(\alpha|I\right)\right).\end{aligned}$$

Since $\min_{\mathsf{d}\in\mathcal{B}(0,\varrho_d)}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\leq\widehat{\mathsf{R}}_c\left(\mathsf{0},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)=0$ this implies

$$\begin{aligned}&\left\{\min_{\mathsf{d};\|\mathsf{d}\|\geq\varrho_d}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\leq\min_{\mathsf{d}\in\mathcal{B}(0,\varrho_d)}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\right\}\\&\subset\left\{\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\leq 0\text{ for all }\mathsf{d}\text{ with }\|\mathsf{d}\|\geq\varrho_d\right\}\\&\subset\left\{\widehat{\mathsf{R}}_c\left(\varrho_d\delta,\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\leq 0\text{ for all }\delta\text{ with }\|\delta\|=1\right\}=\left\{\min_{\mathsf{d}\in\mathcal{S}(\varrho_d)}\widehat{\mathsf{R}}_c\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\leq 0\right\}\\&\subset\left\{\min_{\mathsf{d}\in\mathcal{S}(\varrho_d)}\widehat{\mathsf{R}}_0\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\leq\max_{\mathsf{d}\in\mathcal{B}(0,\varrho_d)}\left|\widehat{\mathsf{O}}_2\left(\mathsf{d},\widehat{\mathsf{e}}\left(\alpha|I\right)\right)\right|\right\}.\end{aligned}$$

It then follows that, by definition of $\widehat{\mathsf{R}}_0(\mathsf{d}, \widehat{\mathsf{e}}(\alpha|I))$ and Lemma B.1,

$$\begin{aligned}
&\left\{\sup_{\alpha\in[0,1]}\left\|\widehat{\mathsf{d}}(\alpha|I)\right\| \geq \varrho_d\right\} = \bigcup_{\alpha\in[0,1]}\left\{\left\|\widehat{\mathsf{d}}(\alpha|I)\right\| \geq \varrho_d\right\} \\
&\subset \left\{\min_{(\alpha,\mathsf{d})\in[0,1]\times\mathcal{S}(\varrho_d)} \widehat{\mathsf{R}}_0(\mathsf{d}, \widehat{\mathsf{e}}(\alpha|I)) \leq \max_{(\alpha,\mathsf{d})\in[0,1]\times\mathcal{B}(0,\varrho_d)}\left|\widehat{\mathsf{O}}_2(\mathsf{d},\widehat{\mathsf{e}}(\alpha|I))\right|\right\} \\
&\subset \left\{\min_{(\alpha,\mathsf{d})\in[0,1]\times\mathcal{S}(\varrho_d)} \frac{\mathsf{d}'\mathsf{R}^{(2)}(\mathsf{b}(\alpha|I_\ell);\alpha, I)\,\mathsf{d}}{2} \leq \max_{(\alpha,\mathsf{d})\in[0,1]\times\mathcal{B}(0,\varrho_d)}\left|\widehat{\mathsf{O}}_2(\mathsf{d},\widehat{\mathsf{e}}(\alpha|I))\right|\right\} \\
&\subset \left\{\frac{\varrho_d^2}{2\Lambda} \leq \max_{(\alpha,\mathsf{d})\in[0,1]\times\mathcal{B}(0,\varrho_d)}\left|\widehat{\mathsf{O}}_2(\mathsf{d},\widehat{\mathsf{e}}(\alpha|I))\right|\right\}.
\end{aligned}$$

Now (B.0.20) gives that for any $\epsilon > 0$, there are some large enough C_e and C_d such that

$$\mathbb{P}\left(\sup_{\alpha\in[0,1]}\left\|\widehat{\mathsf{d}}(\alpha|I)\right\| \geq \varrho_d\right) \leq \epsilon,$$

which shows that the Theorem is proved since $\mathcal{I}$ is finite under Assumption A. $\square$

Main estimation theorems and Proposition 3.4

The end of this section groups the proofs of the main estimation theorems and some additional intermediary Lemmas are now stated. Lemma B.3 gives an upper bound for order of the score vectors $\widehat{\mathsf{R}}^{(1)}_{AQR}(\mathsf{b}_{AQR}(\alpha|I);\alpha, I)$ and $\widehat{\mathsf{R}}^{(1)}_{ASQR}(\mathsf{b}_{ASQR}(\alpha|I);\alpha, I)$ which can be improved for some linear combinations corresponding to estimation of the derivatives of $\beta_{AQR}(\alpha|I)$ and in particular $h\beta^{(1)}_{AQR}(\alpha|I)$ as considered now. Let S_{AQR} be as S_1 from Theorem B.1-(i), such that, using the notations of (3.3.2), $S_{1,AQR}b(\alpha|I) = \beta^{(1)}_{AQR}(\alpha|I)$ and define $S_{1,ASQR}$ for the $ASQR$ case. The next lemma completes Lemma B.3 by studying the order of $S_{1,AQR}\widehat{\mathsf{e}}_{AQR}(\alpha|I)$ and $S_{1,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)$ which are the stochastic error error terms when estimating $h\beta^{(1)}_{AQR}(\alpha|I)$ and $h\beta^{(1)}_{ASQR}(\alpha|I)$ as seen from Theorem B.2. Recall that $\preceq$ is the usual order for symmetric matrices.

Lemma B.4. *Suppose that Assumptions A and H-(i) hold. Then,*

i. If the quantile regression model in Assumption S-(i) is the true one and under

Assumption R-(i),

$$\operatorname{Var}\left(S_{1,AQR}\widehat{\mathsf{e}}_{AQR}(\alpha|I)\right) = \frac{h}{LI}v_h^2(\alpha)\,\mathbb{E}^{-1}\left[\frac{X_\ell X_\ell'\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]\mathbb{E}\left[X_\ell X_\ell'\mathbb{I}(I_\ell = I)\right]$$
$$\times\,\mathbb{E}^{-1}\left[\frac{X_\ell X_\ell'\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] + o(h),$$

uniformly in $(\alpha, I) \in [0,1] \times \mathcal{I}$ *and,*

$$\mathbb{E}\left[\max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|S_{1,AQR}\widehat{\mathsf{e}}_{AQR}(\alpha|I)\right\|\right] \leq Ch^{1/2}\left(\frac{\log L}{L}\right)^{1/2};$$

ii. For the sieve quantile regression model (3.2.9) and under Assumptions S-(ii) with $s \geq 1$*, R-(ii), (B.0.3), and* $\overline{P}_L = \max_{x\in\mathcal{X}}\|P(x)\|$*, it holds that,*

$$\operatorname{Var}\left(S_{1,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right) = \frac{h}{LI}v_h^2(\alpha)\,\mathbb{E}^{-1}\left[\frac{P(x_\ell)P(x_\ell)'\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]$$
$$\times\,\mathbb{E}\left[P(x_\ell)P(x_\ell)'\mathbb{I}(I_\ell = I)\right]\mathbb{E}^{-1}\left[\frac{P(x_\ell)P(x_\ell)'\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] + o(h),$$

uniformly in $(\alpha, I) \in [0,1] \times \mathcal{I}$ *and,*

$$\mathbb{E}\left[\max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|S_{1,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right\|\right] \leq C\left(h^{1/2}\left(\frac{K_L\log L}{L}\right)^{1/2} + \frac{K_L\overline{P}_L\log L}{L}\right).$$

In the next lemma, $S_{0,ASQR} = \mathrm{Id}_{K_L} \otimes [1, 0, \ldots, 0]$ be the $K_L \times (K_L(s+2))$ selection matrix such that $S_{0,AQR}b(\alpha|I) = \beta_{AQR}(\alpha|I)$.

Lemma B.5. *Consider the sieve quantile regression model (3.2.9). Then under Assumptions A and H-(i,ii), S-(ii) with* $s \geq 1$*, R-(ii), Then, uniformly in* $(\alpha, x) \in [0,1] \times \mathcal{X}$*,*

$$\operatorname{Var}\left(P(x)'S_{0,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right) = O\left(\|P(x)\|^2\right),$$

and

$$\operatorname{Var}\left(P'(x)\, S_{1,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right) = hv_h(\alpha)\, P'(x)\, \mathbb{E}^{-1}\left[\frac{P(x_\ell)\, P(x_\ell)'\, \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]$$
$$\times\, \mathbb{E}\left[P(x_\ell)\, P(x_\ell)'\, \mathbb{I}(I_\ell = I)\right] \mathbb{E}^{-1}\left[\frac{P(x_\ell)\, P(x_\ell)'\, \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] P(x) + o(h)$$
$$= O\left(h\,\|P(x)\|^2\right).$$

If moreover Assumption H-(iii) holds,

$$\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\frac{P(x)'}{\|P(x)\|+1} S_{0,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{1/2}\right),$$
$$\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\frac{P(x)'}{\|P(x)\|+1} S_{1,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right| = O_{\mathbb{P}}\left(h^{1/2}\left(\frac{\log L}{L}\right)^{1/2}\right).$$

Proof of Theorem 3.2. Recall that s_1 is the column vector $[0, 1, 0, \ldots, 0]$ of dimension $s + 2$ and let $s_0 = [1, 0, \ldots, 0]$, $S_0 = \mathrm{Id}_{p+1} \otimes s_0$, $S_1 = \mathrm{Id}_{p+1} \otimes s_1$ so that for $\widehat{\mathsf{b}}_{AQR}(\alpha|I)$ as in Theorem B.2 and $\mathsf{b}_{AQR}(\alpha|I)$ as in (B.0.2),

$$\widehat{V}_{AQR}(\alpha|x, I) = X'\left(S_0 + \frac{\alpha S_1}{h(I-1)}\right)\widehat{\mathsf{b}}_{AQR}(\alpha|I),$$
$$V_{AQR}(\alpha|x, I) = X'\left(S_0 + \frac{\alpha S_1}{h(I-1)}\right)\mathsf{b}_{AQR}(\alpha|I).$$

It follows from Theorems B.1-(i) and B.2-(i) that, since $\mathcal{X}$ is compact by Assumption

A-(i),

$$\begin{aligned}
&\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{V}_{AQR}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)\right|\\
&\quad\leq\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{V}_{AQR}\left(\alpha|x,I\right)-V_{AQR}\left(\alpha|x,I\right)\right|+\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|V_{AQR}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)\right|\\
&\quad\leq C\left(\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|S_0\widehat{\mathsf{e}}_{AQR}\left(\alpha|I\right)\right|+h^{-1}\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|S_1\widehat{\mathsf{e}}_{AQR}\left(\alpha|I\right)\right|\right)\\
&\quad+O_{\mathbb{P}}\left(\frac{1}{h}\left(\frac{\log L}{L}\right)^{3/4}\right)+O\left(h^{s+1}\right).
\end{aligned}$$

By Lemmas B.3-(i), B.1 and definition of $\widehat{\mathsf{e}}_{AQR}\left(\alpha|I\right)$, Lemma B.4-(i),

$$\begin{aligned}
\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|S_0\widehat{\mathsf{e}}_{AQR}\left(\alpha|I\right)\right|&=O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{1/2}\right),\\
h^{-1}\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|S_1\widehat{\mathsf{e}}_{AQR}\left(\alpha|I\right)\right|&=O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh}\right)^{1/2}\right).
\end{aligned}$$

Substituting gives

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{V}_{AQR}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)\right|=O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh}\right)^{1/2}\left(1+\left(\frac{\log L}{Lh^2}\right)^{1/4}\right)\right)+O\left(h^{s+1}\right),$$

which gives the desired result since $Lh^2/\log L\to\infty$ by Assumption H-(i). $\square$

Proof of Theorem 3.3. By Theorem B.2-(i), Lemmas B.3-(i) and B.1,

$$\begin{aligned}
&(Lh)^{1/2}\left(\widehat{V}_{AQR}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)-\frac{\alpha S_1\widehat{\mathsf{e}}_{AQR}\left(\alpha|I\right)}{h\left(I-1\right)}-\left(V_{AQR}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)\right)\right)\\
&\quad=(Lh)^{1/2}\left\{O_{\mathbb{P}}\left(\frac{1}{L^{1/2}}\right)+O_{\mathbb{P}}\left(\frac{1}{h}\left(\frac{\log L}{L}\right)^{3/4}\right)\right\}\\
&\quad=O_{\mathbb{P}}\left(h^{1/2}\right)+O_{\mathbb{P}}\left(\left(\frac{\log^3 L}{Lh^2}\right)^{1/4}\right)=o_{\mathbb{P}}\left(1\right),
\end{aligned}$$

by Assumption H-(i), with $V_{AQR}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)=h^{s+1}X'\mathsf{Bias}_h\left(\alpha|I\right)+o\left(h^{s+1}\right)$ by Theorem B.1-(i). Recall that (B.0.5), Lemma B.4-(i), the expression of

$(L/h)^{1/2}\,\widehat{\mathsf{e}}_{AQR}(\alpha|I)$ give that this vector is a triangular average of weighted centered binomial variables which cannot degenerate. The Lindebergh Central Limit Theorem together with Theorem B.1-(i) and Lemma B.4-(i) gives that

$$(Lh)^{1/2}\,\Sigma_h^{-1/2}(\alpha|I)\,\frac{\alpha S_1\widehat{\mathsf{e}}_{AQR}(\alpha|I)}{h(I-1)} \xrightarrow{d} \mathcal{N}(0,\mathrm{Id}). \qquad \square$$

Proof of Theorem 3.4. Redefine S_0 and S_1 as $S_0 = \mathrm{Id}_{K_L}\otimes s_0$, $S_1 = \mathrm{Id}_{K_L}\otimes s_1$ and recall that

$$\widehat{V}_{ASQR}(\alpha|x,I) = P(x)'\left(S_0 + \frac{\alpha S_1}{h(I-1)}\right)\widehat{\mathsf{b}}_{ASQR}(\alpha|I),$$
$$V_{ASQR}(\alpha|x,I) = P(x)'\left(S_0 + \frac{\alpha S_1}{h(I-1)}\right)\mathsf{b}_{ASQR}(\alpha|I).$$

For $\widehat{\mathsf{e}}_{ASQR}(\alpha|I)$ and $\widehat{\mathsf{d}}_{ASQR}(\alpha|I)$ as in (B.0.17), define

$$\widehat{E}_{ASQR}(\alpha|x,I) = P(x)'\left(S_0 + \frac{\alpha S_1}{h(I-1)}\right)\widehat{\mathsf{e}}_{ASQR}(\alpha|I),$$
$$\widehat{D}_{ASQR}(\alpha|x,I) = P(x)'\left(S_0 + \frac{\alpha S_1}{h(I-1)}\right)\widehat{\mathsf{d}}_{ASQR}(\alpha|I),$$

so that

$$\widehat{V}_{ASQR}(\alpha|x,I) - V(\alpha|x,I) = V_{ASQR}(\alpha|x,I) - V(\alpha|x,I) + \widehat{E}_{ASQR}(\alpha|x,I) + \widehat{D}_{ASQR}(\alpha|x,I).$$

By Theorem B.1-(ii) and $K_L \asymp h^{-d_{\mathcal{M}}}$, Assumption Riesz-(ii),

$$\left(\int_0^1\int_{\mathcal{X}}\left(\widehat{V}_{ASQR}(\alpha|x,I) - V(\alpha|x,I)\right)^2 dx\,d\alpha\right)^{1/2}$$
$$= O\left(h^{s+1}\right) + O\left(\int_0^1 \left\{\left\|S_0\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right\| + h^{-1}\left\|S_0\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right\|\right\} d\alpha\right)$$
$$+ O\left(\max_{\alpha\in[0,1]}\left\{\left\|S_0\widehat{\mathsf{d}}_{ASQR}(\alpha|I)\right\| + h^{-1}\left\|S_0\widehat{\mathsf{d}}_{ASQR}(\alpha|I)\right\|\right\}\right)$$

Theorem B.2-(ii) gives, for $\overline{P}_L = \max_{x\in\mathcal{X}}\|P(x)\| \asymp h^{-(1+\zeta)d_{\mathcal{M}}/2}$ and since

$K_L\overline{P}_L^2 \log L/L = \log L/\left(Lh^{(2+\zeta)d_{\mathcal{M}}}\right) \to 0$ by Assumption H-(ii)

$$\begin{aligned}
&\max_{\alpha\in[0,1]}\left\{\left\|S_0\widehat{\mathsf{d}}_{ASQR}(\alpha|I)\right\| + h^{-1}\left\|S_0\widehat{\mathsf{d}}_{ASQR}(\alpha|I)\right\|\right\}\\
&= h^{-1}O_{\mathbb{P}}\left\{\overline{P}_L^{1/2}\left(\frac{K_L\log L}{L}\right)^{3/4}\left(1+\left(\frac{K_L\overline{P}_L^2\log L}{L}\right)^{1/4}\right)\right\}\\
&= h^{-1}O_{\mathbb{P}}\left(\overline{P}_L^{1/2}\left(\frac{K_L\log L}{L}\right)^{3/4}\right).
\end{aligned}$$

By Lemmas B.4-(ii), B.3-(ii) and B.1,

$$\begin{aligned}
&\int_0^1\left\{\left\|S_0\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right\| + h^{-1}\left\|S_0\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right\|\right\}d\alpha\\
&= O_{\mathbb{P}}\left(\left(\int_0^1\mathbb{E}\left[\left(h^{-1}\left\|S_0\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right\|\right)^2\right]d\alpha\right)^{1/2}\right) = O_{\mathbb{P}}\left(\left(\frac{K_L}{Lh}\right)^{1/2}\right).
\end{aligned}$$

Substituting gives, since $\left(\overline{P}_L^2 K_L\log^3 L\right)/\left(Lh^2\right) = \log^3 L/\left(Lh^{(2+\zeta)d_{\mathcal{M}}+2}\right) = O(1)$

$$\begin{aligned}
&\left(\int_0^1\int_{\mathcal{X}}\left(\widehat{V}_{ASQR}(\alpha|x,I) - V(\alpha|x,I)\right)^2 dx d\alpha\right)^{1/2}\\
&= O\left(h^{s+1}\right) + O_{\mathbb{P}}\left(\left(\frac{K_L}{Lh}\right)^{1/2}\left(1+\left(\frac{\overline{P}_L^2K_L\log^3 L}{Lh^2}\right)^{1/4}\right)\right)\\
&= O\left(h^{s+1}\right) + O_{\mathbb{P}}\left(\left(\frac{K_L}{Lh}\right)^{1/2}\right) = O\left(h^{s+1}\right) + O_{\mathbb{P}}\left(\frac{1}{\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}}\right),
\end{aligned}$$

which gives the desired result for the integrated mean squared error.

For the uniform norm, observe first that, since $\overline{P}_L = O\left(K_L^{1/2}\right)$ by Assumption

H-(ii),

$$
\begin{aligned}
\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|\widehat{D}_{ASQR}(\alpha|x,I)\right| &= O\left(\max_{x\in\mathcal{X}}\|P(x)\| \times h^{-1} \max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|S_0\widehat{\mathsf{d}}_{ASQR}(\alpha|I)\right\|\right)\\
&= h^{-1}O_{\mathbb{P}}\left(\overline{P}_L^{3/2}\left(\frac{K_L\log L}{L}\right)^{3/4}\right)\\
&= h^{-1}O_{\mathbb{P}}\left(\left(\frac{K_L^2\log L}{L}\right)^{3/4}\right).
\end{aligned}
$$

Consider now the uniform convergence result, which will use the following normalization of $\widehat{V}_{ASQR}(\alpha|x,I)$, $V_{ASQR}(\alpha|x,I)$ and $V(\alpha|x,I)$,

$$
\begin{aligned}
\widehat{N}(\alpha|x,I) &= \frac{\widehat{V}_{ASQR}(\alpha|x,I)}{\|P(x)\|+1} = \frac{P(x)'}{\|P(x)\|+1}\left(S_0+\frac{\alpha S_1}{h(I-1)}\right)\widehat{\mathsf{b}}_{ASQR}(\alpha|I),\\
\overline{N}(\alpha|x,I) &= \frac{V_{ASQR}(\alpha|x,I)}{\|P(x)\|+1} = \frac{P(x)'}{\|P(x)\|+1}\left(S_0+\frac{\alpha S_1}{h(I-1)}\right)\mathsf{b}_{ASQR}(\alpha|I),
\end{aligned}
$$

keeping in mind that $\|P(x)/(1+\|P(x)\|)\| \leq 1$. Theorem B.1-(iii) and $K_L \asymp h^{-d_{\mathcal{M}}}$, the Cauchy-Schwarz inequality, Lemma B.5-(ii) and Theorem B.2-(ii), $\overline{P}_L = O\left(K_L^{1/2}\right)$ by Assumption H-(iii) give,

$$
\begin{aligned}
&\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{V}_{ASQR}(\alpha|x,I) - V(\alpha|x,I)\right|\\
&\quad\leq \max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|V_{AQR}(\alpha|x,I)-V(\alpha|x,I)\right| + \left(1+\overline{P}_L\right)\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{N}(\alpha|x,I)-\overline{N}(\alpha|x,\right.\\
&\quad\leq O\left(h^{s+1}\right) + \left(1+\overline{P}_L\right)\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\frac{P(x)'}{\|P(x)\|+1}\left(S_0+\frac{\alpha S_1}{h(I-1)}\right)\left(\widehat{\mathsf{e}}_{ASQR}(\alpha|I)-\mathsf{e}_{ASQR}(\right.\right.\\
&\quad\quad+\left(1+\overline{P}_L\right)\max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|\left(S_0+\frac{\alpha S_1}{h(I-1)}\right)\left(\widehat{\mathsf{d}}_{ASQR}(\alpha|I)-\mathsf{d}_{ASQR}(\alpha|I)\right)\right\|\\
&\quad\leq O\left(h^{s+1}\right) + \left(1+\overline{P}_L\right)\left\{O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh}\right)^{1/2}\right) + h^{-1}O_{\mathbb{P}}\left(\overline{P}_L^{1/2}\left(\frac{K_L\log L}{L}\right)^{3/4}\right)\right\}\\
&\quad= O\left(h^{s+1}\right) + O_{\mathbb{P}}\left(\left(\frac{K_L\log L}{Lh}\right)^{1/2}\right)\left\{1+O_{\mathbb{P}}\left(\left(\frac{K_L^4\log L}{Lh^2}\right)^{1/4}\right)\right\}\\
&\quad= O\left(h^{s+1}\right) + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)\left\{1+O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{4d_{\mathcal{M}}+2}}\right)^{1/4}\right)\right\},
\end{aligned}
$$

which gives the desired result since $\log L / \left(Lh^{4d_{\mathcal{M}}+2}\right) = O(1)$. $\square$

Proof of Proposition 3.4. Observe first that, for $\widehat{\mathsf{b}}_{AQR}(\alpha|I)$ computed from $\widehat{\beta}_{AQR}(\alpha|I)$ as in (B.0.1),

$$\widehat{\boldsymbol{B}}(h;\alpha,I) = \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\widetilde{\rho}_{\alpha+ht}\left(B_{i\ell} - X_\ell(t)'\widehat{\mathsf{b}}_{AQR}(\alpha|I)\right)K_{\alpha+ht}(t)\,dt.$$

Let $\mathsf{b}_{AQR}(\alpha|I)$ is as defined in (B.0.2) and $\widehat{\mathsf{d}}_{AQR}(\alpha|I)$, $\widehat{\mathsf{e}}_{AQR}(\alpha|I)$ as in (B.0.16) so that $\widehat{\mathsf{b}}_{AQR}(\alpha|I) = \mathsf{b}_{AQR}(\alpha|I) + \widehat{\Delta}(\alpha|I)$, $\widehat{\Delta}(\alpha|I) = \widehat{\mathsf{d}}_{AQR}(\alpha|I) + \widehat{\mathsf{e}}_{AQR}(\alpha|I)$. Define, for $\rho_\alpha^{(1)}(q) = \alpha - \mathbb{I}(q \leq 0)$,

$$\begin{aligned}
\widehat{\boldsymbol{B}}_0(h;\alpha,I) &= \frac{1}{L}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\widetilde{\rho}_{\alpha+ht}\left(B_{i\ell} - B(\alpha+ht|x_\ell,I)\right)K_{\alpha+ht}(t)\,dt,\\
\widehat{\boldsymbol{B}}_0^{(1)}(h;\alpha,I) &= -\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\widetilde{\rho}^{(1)}_{\alpha+ht}\left(B_{i\ell} - B(\alpha+ht|x_\ell,I)\right)\\
&\qquad\times\left(X_\ell(t)\,\mathsf{b}_{AQR}(\alpha|I) - B(\alpha+ht|x_\ell,I)\right)K_{\alpha+ht}(t)\,dt,\\
\widehat{\boldsymbol{B}}_1^{(1)}(h;\alpha,I) &= -\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\widetilde{\rho}^{(1)}_{\alpha+ht}\left(B_{i\ell} - B(\alpha+ht|x_\ell,I)\right)\\
&\qquad\times X_\ell(t)'K_{\alpha+ht}(t)\,dt\,\widehat{\Delta}(\alpha|I),\\
\boldsymbol{B}_0^{(2)}(h;\alpha,I) &= \mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\frac{\left(X_\ell(t)\,\mathsf{b}_{AQR}(\alpha|I) - B(\alpha+ht|x_\ell,I)\right)^2}{B^{(1)}(\alpha+ht|x_\ell,I)}\right.\\
&\qquad\left.(\alpha+ht)^2K_{\alpha+ht}(t)\,dt\right],\\
\boldsymbol{B}_1^{(2)}(h;\alpha,I) &= \widehat{\Delta}(\alpha|I)'\\
&\qquad\times\mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\frac{X_\ell(t)X_\ell(t)'}{B^{(1)}(\alpha+ht|x_\ell,I)}(\alpha+ht)^2K_{\alpha+ht}(t)\,dt\right]\widehat{\Delta}(\alpha|I).
\end{aligned}$$

Repeating the arguments in Propositions B.1 and B.2 gives, uniformly in

$\alpha \in [0,1]$.

$$\int_0^1 \widehat{\boldsymbol{B}}(h;\alpha,I)\,d\alpha = \int_0^1 \widehat{\boldsymbol{B}}_0(h;\alpha,I)\,d\alpha + \frac{1}{2}\int_0^1 \boldsymbol{B}_0^{(2)}(h;\alpha,I)\,d\alpha$$
$$+ \int_0^1 \left\{ \widehat{\boldsymbol{B}}_0^{(1)}(h;\alpha,I)\,d\alpha + \frac{1}{2}\boldsymbol{B}_1^{(2)}(h;\alpha,I) \right\} d\alpha + O\left(\int_0^1 \left(\boldsymbol{B}_0^{(2)}(h;\alpha,I)\,\boldsymbol{B}_1^{(2)}(h;\alpha,I) \right)^{1/2} d\alpha \right)$$
$$+ O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{1/2} \max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left\{ \left\| \widehat{\Delta}(\alpha|I) \right\|^{3/2} + \left\| \alpha\left(\mathsf{b}_{AQR}(\alpha|I) - \mathsf{b}(\alpha|I)\right) \right\|^{3/2} \right\} \right).$$

Repeating the arguments in Lemma B.4 gives, using (B.0.14),

$$\max_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left| \widehat{\boldsymbol{B}}_0^{(1)}(h;\alpha,I) \right| = O_{\mathbb{P}}\left(\frac{h^{s+2}}{L^{1/2}}\right) = O_{\mathbb{P}}\left(\frac{h^{s+2}\times h^{1/4}}{L^{1/2}\times h^{1/4}}\right)$$
$$= O_{\mathbb{P}}\left(h^{2(s+2)+\frac{1}{2}} + \frac{1}{Lh^{1/2}} \right) = o_{\mathbb{P}}\left(h^{2(s+2)} + \frac{1}{Lh} \right),$$

since $2|ab| \leq a^2 + b^2$. Theorem B.2-(i) and (B.0.14) also yield,

$$\max_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left| \boldsymbol{B}_1^{(2)}(h;\alpha,I) \right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{3/2} \right) = o_{\mathbb{P}}\left(h^{2(s+2)} + \frac{1}{Lh} \right),$$

$$\int_0^1 \left(\boldsymbol{B}_0^{(2)}(h;\alpha,I)\,\boldsymbol{B}_1^{(2)}(h;\alpha,I) \right)^{1/2} d\alpha = O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{3/4} h^{s+2} \right)$$
$$= O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{1+\frac{1}{4}} + \left(\frac{\log L}{L}\right)^{\frac{1}{4}} h^{2(s+2)} \right) = o_{\mathbb{P}}\left(h^{2(s+2)} + \frac{1}{Lh} \right),$$

$$
\begin{aligned}
&\left(\frac{\log L}{L}\right)^{1/2} \max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left\{\left\|\widehat{\Delta}\left(\alpha|I\right)\right\|^{3/2} + \left\|\alpha\left(\mathsf{b}_{AQR}\left(\alpha|I\right) - \mathsf{b}\left(\alpha|I\right)\right)\right\|^{3/2}\right\} \\
&= \left(\frac{\log L}{L}\right)^{1/2} \left\{O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{3/4}\right) + O\left(h^{\frac{3}{2}(s+2)}\right)\right\} \\
&= O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{5/4}\right) + O\left(\frac{\log L}{L} h^{\frac{1}{4}(s+2)}\right) + O\left(h^{2(s+2)} h^{\frac{1}{4}(s+2)}\right) \\
&= o_{\mathbb{P}}\left(h^{2(s+2)} + \frac{1}{Lh}\right).
\end{aligned}
$$

Let

$$
M_1\left(\alpha\right) = \mathbb{E}\left[\frac{\mathbb{I}\left(I_\ell = I\right) X_\ell \alpha B^{(s+2)}\left(\alpha|x_\ell, I\right)}{B^{(1)}\left(\alpha|x_\ell, I\right)}\right], \quad M_2\left(\alpha\right) = \mathbb{E}\left[\frac{\mathbb{I}\left(I_\ell = I\right) X_\ell X_\ell'}{B^{(1)}\left(\alpha|x_\ell, I\right)}\right],
$$

$$
\Pi_1\left(\alpha, h\right) = \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \frac{t^{s+2}\pi\left(t\right)}{(s+2)!} K\left(t\right) dt, \quad \Pi_2\left(\alpha, h\right) = \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi\left(t\right)\pi\left(t\right)' K\left(t\right) dt.
$$

Observe that except for $\alpha = 0, 1$,

$$
\begin{aligned}
&\lim_{h\to 0} K_{\alpha+ht}\left(t\right) = K\left(t\right), \ \lim_{h\to 0} \Pi_1\left(\alpha, h\right) = \int \frac{t^{s+2}\pi\left(t\right)}{(s+2)!} K\left(t\right) dt = \Pi_1, \\
&\lim_{h\to 0} \Pi_2\left(\alpha, h\right) = \int \pi\left(t\right)\pi\left(t\right)' K\left(t\right) dt = \Pi_2.
\end{aligned}
$$

A Taylor expansion, (B.0.15) and arguing as for (B.0.12), and the dominated convergence theorem give,

$$\begin{aligned}
&\int_0^1 \boldsymbol{B}_0^{(2)}(h;\alpha,I)\,d\alpha \\
&= \mathbb{E}\left[\mathbb{I}(I_\ell = I)\int_0^1\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \frac{\left(X_\ell(t)\,\mathsf{b}_{AQR}(\alpha|I) - B(\alpha+ht|x_\ell,I)\right)^2}{B^{(1)}(\alpha+ht|x_\ell,I)}(\alpha+ht)^2 K_{\alpha+ht}(t)\,dtd\alpha\right] \\
&= \int_0^1\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left(\alpha\left(\mathsf{b}_{AQR}(\alpha|I) - \mathsf{b}(\alpha|I)\right)\right)'\left(M_2(\alpha+ht)\otimes\Pi_2(\alpha,h)\right) \\
&\qquad \times\left(\alpha\left(\mathsf{b}_{AQR}(\alpha|I) - \mathsf{b}(\alpha|I)\right)\right)K_{\alpha+ht}(t)\,dtd\alpha + o\left(h^{2(s+2)}\right) \\
&= h^{2(s+2)}\int_0^1\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left(M_1(\alpha)\otimes\Pi_1(\alpha,h)\right)'\left(M_2(\alpha)\otimes\Pi_2(\alpha,h)\right)^{-1}\left(M_2(\alpha+ht)\otimes\Pi_2(\alpha,h)\right) \\
&\qquad \times\left(M_2(\alpha)\otimes\Pi_2(\alpha,h)\right)^{-1}\left(M_1(\alpha)\otimes\Pi_1(\alpha,h)\right)K_{\alpha+ht}(t)\,dtd\alpha + o\left(h^{2(s+2)}\right) \\
&= h^{2(s+2)}\int_0^1 M_1(\alpha)' M_2(\alpha)^{-1} M_1(\alpha)\,d\alpha \times \Pi_1'\Pi_2^{-1}\Pi_1 + o\left(h^{2(s+2)}\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
\int_0^1 \widehat{\boldsymbol{B}}(h;\alpha,I)\,d\alpha = \int_0^1 \widehat{\boldsymbol{B}}_0(h;\alpha,I)\,d\alpha + \frac{h^{2(s+2)}\Pi_1'\Pi_2^{-1}\Pi_1}{2I}\int_0^1 M_1(\alpha)' M_2(\alpha)^{-1} M_1(\alpha)\,d\alpha \\
+ o_{\mathbb{P}}\left(h^{2(s+2)} + \frac{1}{Lh}\right).
\end{aligned}$$

Observe now that,

$$\begin{aligned}
&\int_0^1 \widehat{\boldsymbol{B}}_0(h;\alpha,I)\,d\alpha \\
&= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_0^1\int_0^1 \rho_a\left(B_{i\ell} - B(a|x_\ell,I)\right)K_a\left(\frac{a-\alpha}{h}\right)da\frac{d\alpha}{h} \\
&= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_0^1 \rho_a\left(B_{i\ell} - B(a|x_\ell,I)\right)\left\{\frac{\int K\left(\frac{a-\alpha}{h}\right)\frac{d\alpha}{h}}{\int K\left(\frac{a-\alpha}{h}\right)\frac{d\alpha}{h}}\right\}da \\
&= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_0^1 \rho_a\left(B_{i\ell} - B(a|x_\ell,I)\right)da,
\end{aligned}$$

which ends the proof of the Proposition. $\square$

Proof of intermediary results

We now give the proof of Lemma B.3, B.4, Propositions B.1 and B.2, for which it is sufficient to consider the ASQR case since taking constant K_L and $\max_{x\in\mathcal{X}}\|P(x)\|$ gives the results stated in the AQR case. The proofs of this section heavily relies on the following maximal inequality from Massart (2007, Theorem 6.8 and Corollary 6.9). When applying this bound, the variable ξ_ℓ will stand for the auction observation $(I_\ell, x_\ell, B_{i\ell}, i = 1, \ldots, I_\ell)$ or some function of this variable.

Theorem B.3. *Let $\{\xi_\ell, \ell \geq 1\}$ be a sequence of independent variables and let $\mathcal{F}$ be a separable class of real functions of ξ_ℓ. Assume that (i) there exists some positive numbers σ and b such that for all integer numbers $k \geq 2$ and any $f(\cdot) \in \mathcal{F}$,*

$$\frac{1}{L}\sum_{\ell=1}^{L}\mathbb{E}\left[|f(\xi_\ell)|^k\right] \leq \frac{k!}{2}\sigma^2 M^{k-2};$$

(ii) for any $\epsilon > 0$, there exists some finite set $\mathbf{F}_\epsilon$ of $\exp(\mathbf{H}(\epsilon))$ brackets which union covers $\mathcal{F}$ and such that, for any $\left[\underline{f}, \overline{f}\right] \in \mathbf{F}_\epsilon$,

$$\frac{1}{L}\sum_{\ell=1}^{L}\mathbb{E}\left[\left|\overline{f}(\xi_\ell) - \underline{f}(\xi_\ell)\right|^k\right] \leq \frac{k!}{2}\epsilon^2 M^{k-2}.$$

Then, for

$$\mathbf{E} = 27L^{1/2}\int_0^\sigma \min{}^{1/2}(H(t), L)\,dt + 2(M+\sigma)H(\sigma),$$

it holds that,

$$\mathbb{E}\left[\sup_{f\in\mathcal{F}}\left|\sum_{\ell=1}^{L}\left(f(\xi_\ell) - \mathbb{E}(f(\xi_\ell))\right)\right|\right] \leq \mathbf{E}.$$

The upper bound for the mean in Theorem B.3 implies that, for any $e \in (0,1]$,

$$\sup_{f\in\mathcal{F}}\left|\sum_{\ell=1}^{L}\left(f(\xi_\ell) - \mathbb{E}(f(\xi_\ell))\right)\right| = O_{\mathbb{P}}(\mathbf{E}),$$

where σ, b and therefore $\mathbf{E}$ can depend upon L. In this statement and when the ξ_ℓ

are i.i.d, σ^2 and M can be viewed as, respectively upper bounds for $\mathrm{Var}\left(f\left(\xi_\ell\right)\right)$ and $\left|f\left(\xi_\ell\right)\right|$ valid for all $f(\cdot) \in \mathcal{F} \bigcup \cup_{\epsilon\in(0,e\sigma]} \mathbf{F}_\epsilon$.

Note that it is sufficient to establish the intermediary results for the ASQR case since it implies the results for the AQR case by taking constant K_L and $\max_{x\in\mathcal{X}} \|P(x)\|$.

Proof of Lemma B.3. It is sufficient to consider $\widehat{\mathsf{R}}^{(1)}_{ASQR}\left(\mathsf{b}_{ASQR}\left(\alpha|I\right);\alpha,I\right)$ and a fixed I since $\mathcal{I}$ is finite. Abbreviate $\mathsf{b}_{ASQR}\left(\alpha|I\right)$ into $\mathsf{b}\left(\alpha|I\right)$ and $\widehat{\mathsf{R}}^{(1)}_{ASQR}\left(\mathsf{b};\alpha,I\right)$, $\mathsf{R}^{(1)}_{ASQR}\left(\mathsf{b};\alpha,I\right)$ into $\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b};\alpha,I\right)$, $\mathsf{R}^{(1)}\left(\mathsf{b};\alpha,I\right)$. Recall also that $\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)$ is centered since $\mathsf{b}\left(\alpha|I\right) = \arg\min_{\mathsf{b}} \mathsf{R}\left(\mathsf{b};\alpha,I\right)$ satisfies a first-order condition so that,

$$\mathbb{E}\left[\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)\right] = \mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right) = 0.$$

Define,

$$\widehat{\mathsf{r}}\left(\theta;\alpha,I\right) = \widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)'\theta, \quad \mathcal{S} = \left\{\theta;\theta'\theta = 1\right\},$$

which is such that $\left\|\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)\right\| = \max_{\theta\in\mathcal{S}}\left|\widehat{\mathsf{r}}\left(\theta;\alpha,I\right)\right|$. It is then sufficient to bound the exceedance probability of

$$\max_{(\alpha,\theta)\in[0,1]\times\mathcal{S}}\left|\widehat{\mathsf{r}}\left(\theta;\alpha,I\right)\right|.$$

Define, using the notations of Theorem B.1,

$$f\left(\xi_\ell;\alpha,\theta\right) = \frac{\mathbb{I}\left(I_\ell = I\right)}{I}\sum_{i=1}^{I}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left\{\mathbb{I}\left(B_{i\ell} \leq P\left(x_\ell,t\right)'\mathsf{b}\left(\alpha|I\right)\right) - \left(\alpha + ht\right)\right\}$$
$$\times\, \theta' P\left(x_\ell\right)\otimes\pi\left(t\right)K\left(t\right)dt,$$

which is such that $\widehat{\mathsf{r}}\left(\theta;\alpha,I\right) = \sum_{\ell=1}^{L} f\left(\xi_\ell;\alpha,\theta\right)/L$, with,

$$\max_{\ell\geq 1}\sup_{(\alpha,\theta)\in[0,1]\times\mathcal{S}}\left|f\left(\xi_\ell;\alpha,\theta\right)\right| \leq C\max_{x\in\mathcal{X}}\left\|P\left(x\right)\right\| \leq C\overline{P}_L.$$

For the variance of $f\left(\xi_\ell;\alpha,\theta\right)$, observe that, under Assumptions H, A and by

(B.0.10),

$$
\begin{aligned}
&\operatorname{Var}\left(If\left(\xi_{\ell};\alpha,\theta\right)\right)=\mathbb{E}\left[\mathbb{E}\left[f^{2}\left(\xi_{\ell};\alpha,\theta\right)|x_{\ell},I_{\ell}\right]\right]\\
&\leq CI\mathbb{E}\left[\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\mathbb{E}\left[\left\{\mathbb{I}\left(B_{i\ell}\leq P\left(x_{\ell},t\right)'\mathsf{b}\left(\alpha|I\right)\right)-\left(\alpha+ht\right)\right\}^{2}|x_{\ell},I_{\ell}=I\right]\right.\\
&\left.\times\left(\theta'P\left(x_{\ell}\right)\otimes\pi\left(t\right)\right)^{2}K\left(t\right)dt\right]\\
&\leq C\mathbb{E}\left[\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left\{\mathbb{P}\left(B_{i\ell}\leq P\left(x_{\ell},t\right)'\mathsf{b}\left(\alpha|I\right)|x_{\ell},I\right)-2\left(\alpha+ht\right)\mathbb{P}\left(B_{i\ell}\leq P\left(x_{\ell},t\right)'\mathsf{b}\left(\alpha|I\right)|x_{\ell},I\right)\right.\right.\\
&\left.\left.+\left(\alpha+ht\right)^{2}\right\}\left(\theta'P\left(x_{\ell}\right)\otimes\pi\left(t\right)\right)^{2}K\left(t\right)dt\right]\\
&=2C\mathbb{E}\left[\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left\{\left(\alpha+ht\right)\left(1-\alpha-ht\right)+o\left(h\right)\right\}\left(\theta'P\left(x_{\ell}\right)\otimes\pi\left(t\right)\right)^{2}K\left(t\right)dt\right]\\
&=C\left\{\alpha\left(1-\alpha\right)\right\}+O\left(h\right)\leq\sigma^{2}.
\end{aligned}
$$

This bound hold for all α and I and implies the variance result of the Lemma.

To find some brackets as in Theorem B.3, we first bound the differential $\mathsf{b}^{(1)}\left(\alpha|I\right)=d\mathsf{b}\left(\alpha|I\right)/d\alpha$. The implicit function theorem gives since $\mathsf{b}\left(\alpha|I\right)$ is the unique solution of $\mathsf{R}^{(1)}\left(\mathsf{b};\alpha,I\right)=0$,

$$
\mathsf{b}^{(1)}\left(\alpha|I\right)=-\left(\mathsf{R}^{(2)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)\right)^{-1}\frac{\partial\mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)}{\partial\alpha},
$$

where

$$
\begin{aligned}
&\frac{\partial\mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)}{\partial\alpha}\\
&=\frac{1}{h}\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\left\{G\left(P\left(x_{\ell},-\frac{\alpha}{h}\right)\mathsf{b}\left(\alpha|I\right)|x_{\ell},I_{\ell}\right)-\alpha\right\}\left(P\left(x_{\ell}\right)\otimes\pi\left(-\frac{\alpha}{h}\right)\right)K\left(-\frac{\alpha}{h}\right)\right]\\
&-\frac{1}{h}\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\left\{G\left(P\left(x_{\ell},\frac{1-\alpha}{h}\right)\mathsf{b}\left(\alpha|I\right)|x_{\ell},I_{\ell}\right)-1-\alpha\right\}\right.\\
&\qquad\left.\times\left(P\left(x_{\ell}\right)\otimes\pi\left(\frac{1-\alpha}{h}\right)\right)K\left(\frac{1-\alpha}{h}\right)\right]\\
&-\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}P\left(x_{\ell}\right)\otimes\pi\left(t\right)K\left(t\right)dt\right].
\end{aligned}
$$

By (B.0.10) and Assumption R-(ii),

$$\max_{\alpha\in[0,1]}\left\|\frac{\partial\mathsf{R}^{(1)}\left(\mathsf{b}\left(\alpha|I\right);\alpha,I\right)}{\partial\alpha}\right\|\leq C\left\|\mathbb{E}\left[P\left(x_\ell\right)\right]\right\|=C\max_{\theta\in\mathcal{S}}\left|\mathbb{E}\left[P\left(x_\ell\right)'\theta\right]\right|$$
$$\leq C\max_{\theta\in\mathcal{S}}\left|\mathbb{E}^{\frac{1}{2}}\left[\left(P\left(x_\ell\right)'\theta\right)^2\right]\right|\leq C,$$

and then, by Lemma B.1 and for all h small enough,

$$\max_{\alpha\in[0,1]}\left\|\mathsf{b}^{(1)}\left(\alpha|I\right)\right\|\leq C_1. \tag{B.0.21}$$

Let $D_L=(s+2)K_L+1$ be the dimension of $(\alpha,\theta')'$, $|(\alpha,\theta')|_2=|\alpha|+\|\theta\|$ and

$$\Delta=\frac{h\epsilon^2}{2C_2\left(1+\max_{x\in\mathcal{X}}\|P(x)\|\right)}.$$

Let $\left\{\left(\alpha_j,\theta_j'\right)',j\in\mathcal{J}_\epsilon\right\}\subset[0,1]\times\mathcal{S}$ be such that for any $(\alpha,\theta')'\in[0,1]\times\mathcal{S}$ there is a $\left(\alpha_j,\theta_j'\right)$ such that $\left|(\alpha,\theta')-\left(\alpha_j,\theta_j'\right)\right|_2\leq\Delta$. Define

$$\begin{aligned}\underline{f}_j\left(\xi_\ell\right)=&\min_{\alpha\in[0,1],\theta';|(\alpha-\alpha_j,\theta'-\theta_j)|_2\leq\Delta}\left\{\mathbb{I}\left(I_\ell=I\right)\sum_{i=1}^{I}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\right.\\&\left\{\mathbb{I}\left(B_{i\ell}\leq P\left(x_\ell,t\right)'\mathsf{b}^{(1)}\left(\alpha_j|I\right)+C_1\left(\alpha-\alpha_j\right)\left\|P\left(x_\ell\right)\right\|\right)-\left(\alpha+ht\right)\right\}\\&\left.\theta'P\left(x_\ell\right)\otimes\pi\left(t\right)K\left(t\right)dt\right\},\\\overline{f_j}\left(\xi_\ell\right)=&\max_{\alpha\in[0,1],\theta';|(\alpha-\alpha_j,\theta'-\theta_j)|_2\leq\Delta}\left\{\mathbb{I}\left(I_\ell=I\right)\sum_{i=1}^{I}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\right.\\&\left\{\mathbb{I}\left(B_{i\ell}\leq P\left(x_\ell,t\right)'\mathsf{b}^{(1)}\left(\alpha_j|I\right)+C_1\left(\alpha-\alpha_j\right)\left\|P\left(x_\ell\right)\right\|\right)-\left(\alpha+ht\right)\right\}\\&\left.\theta'P\left(x_\ell\right)\otimes\pi\left(t\right)K\left(t\right)dt\right\}\end{aligned}$$

where the min and the max are achieved by some $(\alpha,\theta')'$ such that $|(\alpha-\alpha_j,\theta'-\theta_j)|_2=\Delta$. Since $[0,1]\times\mathcal{S}\subset\cup_{j\in\mathcal{J}_\epsilon}\left\{(\alpha,\theta');|(\alpha-\alpha_j,\theta'-\theta_j)|_2\leq\Delta\right\}$, it follows that by (B.0.21),

$$\mathcal{F}:=\left\{f\left(\cdot\right)=f\left(\cdot,\alpha,\theta\right),(\alpha,\theta')'\in[0,1]\times\mathcal{S}\right\}\subset\cup_{j\in\mathcal{J}_\epsilon}\left[\underline{f}_j,\overline{f_j}\right].$$

Observe now that,

$$
\begin{aligned}
\left|\overline{f_j}\left(\xi_\ell\right)-\underline{f}_j\left(\xi_\ell\right)\right| &\leq \left|\overline{f_j}\left(\xi_\ell\right)-f\left(\xi_\ell,\alpha_j,\theta_j\right)\right|+\left|\underline{f}_j\left(\xi_\ell\right)-f\left(\xi_\ell,\alpha_j,\theta_j\right)\right| \\
&\leq C\left|P\left(x_\ell\right)\theta_j'\right|\left\{\int_{\frac{1-\alpha_j-\Delta}{h}}^{\frac{1-\alpha_j+\Delta}{h}} K(t)\,dt+\int_{\frac{-\alpha_j-\Delta}{h}}^{\frac{-\alpha_j+\Delta}{h}} K(t)\,dt+\right. \\
&\qquad \left.+\,\mathbb{I}\left(\left|B_{i\ell}-P\left(x_\ell,t\right)'\mathsf{b}^{(1)}\left(\alpha_j|I\right)\right|\leq C_1\Delta\left\|P\left(x_\ell\right)\right\|\right)\right\} \\
&\quad +C\left\|P\left(x_\ell\right)\right\|\Delta
\end{aligned}
$$

This gives, taking $M=C\overline{P}_L$ which is also a bound for $|f(\xi_\ell,\alpha,\theta)|$, by Proposition 3.1 which implies that the p.d.f $g(\cdot|\cdot,\cdot)$ of $B_{i\ell}$ given (x_ℓ,I_ℓ) is bounded, Assumptions A and R-(ii),

$$
\begin{aligned}
&\mathbb{E}\left[\left|\overline{f_j}\left(\xi_\ell\right)-\underline{f}_j\left(\xi_\ell\right)\right|^k\right]\leq M^{k-2}\mathbb{E}\left[\left|\overline{f_j}\left(\xi_\ell\right)-\underline{f}_j\left(\xi_\ell\right)\right|^2\right] \\
&\leq CM^{k-2}\left\{\mathbb{E}\left[\left|P\left(x_\ell\right)\theta_j'\right|^2\right]\frac{\Delta^2}{h^2}+\mathbb{E}\left[\left|P\left(x_\ell\right)\theta_j'\right|^2\left\|P\left(x_\ell\right)\right\|\right]\Delta+\mathbb{E}\left[\left\|P\left(x_\ell\right)\right\|^2\right]\Delta^2\right\} \\
&\leq \frac{C\left(2\epsilon+1\right)}{C_2}M^{k-2}\epsilon,
\end{aligned}
$$

which can make smaller that $M^{k-2}\epsilon$ for all $\epsilon\in(0,C]$ by increasing C_2. Before applying Theorem B.3, observe that van de Geer (1999, p.20) gives

$$
\mathbf{H}(\epsilon)=\log\left(\operatorname{Card}\mathfrak{J}_\epsilon\right)=CD_L\log\left(\frac{1}{\Delta}\right)\leq CK_L\left(\log\frac{1}{\epsilon}+\log L\right).
$$

Note that, for L large enough,

$$
\begin{aligned}
&\int_0^\sigma \mathbf{H}^{1/2}(t)\,dt\leq CK_L^{1/2}\int_0^\sigma\left\{\log^{1/2}\frac{1}{t}+\log^{1/2}L\right\}dt \\
&\leq CK_L^{1/2}\sigma\left\{\log^{1/2}L+\left(\int_0^\sigma\log\frac{1}{t}dt\right)^{1/2}\right\}=CK_L^{1/2}\sigma\left\{\log^{1/2}L-\left.t\left(\log t+1\right)\right]_0^\sigma\right\} \\
&\leq CK_L^{1/2}\log^{1/2}L.
\end{aligned}
$$

This gives for $\mathbf{E}$ as in Theorem B.3,

$$\mathbf{E} \leq C\left(\left(LK_L \log L\right)^{1/2} + K_L \overline{P}_L \log L\right).$$

Hence Theorem B.3 gives for $\widehat{\mathsf{r}}(\theta;\alpha,I) = \sum_{\ell=1}^{L} f(\xi_\ell;\alpha,\theta)/L$,

$$\mathbb{E}\left[\max_{(\alpha,\theta)\in[0,1]\times\mathcal{S}} \left|\widehat{\mathsf{r}}(\theta;\alpha,I)\right|\right] \leq C\left(\left(\frac{K_L \log L}{L}\right)^{1/2} + \frac{K_L \overline{P}_L \log L}{L}\right),$$

which gives the desired result. □

Proof of Lemma B.4. It is sufficient to consider the ASQR case. Following the steps of the proof of Lemma B.3 gives, since $\max_{\alpha\in[0,1]} v_h^2(\alpha) = O(1)$ and by (B.0.8), the bound for the expectation of the maximum of $S_{1,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)$ provided the variance bound holds as established now. Remove for the sake of brevity the subscript ASQR and the variables α and I from the notations, so that

$$\widehat{\mathsf{e}} = \widehat{\mathsf{e}}_{ASQR}(\alpha|I) = -\left(\mathsf{R}^{(2)}(\mathsf{b})\right)^{-1} \widehat{\mathsf{R}}^{(1)}(\mathsf{b}),$$

with $\mathsf{b} = \mathsf{b}_{ASQR}(\alpha|I)$ and $\mathsf{R}^{(2)}(\cdot) = \mathsf{R}^{(2)}_{ASQR}(\cdot|\alpha,I)$, $\widehat{\mathsf{R}}^{(1)}(\cdot) = \widehat{\mathsf{R}}^{(1)}_{ASQR}(\cdot|\alpha,I)$. Let

$$\widetilde{\mathsf{R}}^{(2)}(\mathsf{b}) = \mathbb{E}\left[\mathbb{I}(I_\ell = I)\, P(x_\ell)\, P(x_\ell)' \otimes \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \frac{\pi(t)\,\pi'(t)\,K(t)}{B^{(1)}(\alpha+ht|x_\ell,I)} dt\right],$$

and let $\|\cdot\|$ stands for the Euclidean or operator norm. (B.0.9), (B.0.8) and the Taylor inequality of order $s+1$ gives, since $s \geq 1$ and by (B.0.3),

$$\max_{\alpha\in[0,1]} \left\|\widetilde{\mathsf{R}}^{(2)}(\mathsf{b}) - \mathsf{R}^{(2)}(\mathsf{b})\right\| = o\left(\overline{P}_L\left(h^{s+1} + K_L^{-\frac{s+2}{d_{\mathcal{M}}}}\right)\right) + o(h^s) = o(h).$$

Define now,

$$\begin{aligned}
\mathbf{P} &= \mathbb{E}\left[\mathbb{I}\left(I_\ell = I\right) P\left(x_\ell\right) P\left(x_\ell\right)'\right], \\
\mathbf{P}_0 &= \mathbb{E}\left[\frac{\mathbb{I}\left(I_\ell = I\right) P\left(x_\ell\right) P\left(x_\ell\right)'}{B^{(1)}\left(\alpha | x_\ell, I\right)}\right], \\
\mathbf{P}_1 &= -\mathbb{E}\left[\mathbb{I}\left(I_\ell = I\right) P\left(x_\ell\right) P\left(x_\ell\right)' \frac{B^{(2)}\left(\alpha | x_\ell, I\right)}{\left(B^{(1)}\left(\alpha | x_\ell, I\right)\right)^2}\right], \\
\mathbf{\Pi}_0 &= \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi\left(t\right) \pi'\left(t\right) K\left(t\right) dt, \quad \pi_0 = \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \pi\left(t\right) \pi'\left(t\right) K\left(t\right) dt, \\
\mathbf{\Pi}_1 &= \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} t\pi\left(t\right) \pi'\left(t\right) K\left(t\right) dt, \quad \pi_1 = \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} t\pi\left(t\right) K\left(t\right) dt, \\
\mathbf{\Pi}_m &= \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \min\left(t_1, t_2\right) \pi\left(t_1\right) \pi'\left(t_2\right) K\left(t_1\right) K\left(t_2\right) dt_1 dt_2.
\end{aligned}$$

Then Proposition 3.1 and a standard Taylor expansion give

$$\max_{\alpha \in [0,1]} \left\| \widetilde{\mathsf{R}}^{(2)}\left(\mathsf{b}\right) - \mathbf{P}_0 \otimes \mathbf{\Pi}_0 - h\mathbf{P}_1 \otimes \mathbf{\Pi}_1 \right\| = o\left(h\right),$$

so that

$$\max_{\alpha \in [0,1]} \left\| \mathsf{R}^{(2)}\left(\mathsf{b}\right) - \mathbf{P}_0 \otimes \mathbf{\Pi}_0 - h\mathbf{P}_1 \otimes \mathbf{\Pi}_1 \right\| = o\left(h\right). \tag{B.0.22}$$

Consider now $\mathsf{V}\left(\mathsf{b}\right) = \mathrm{Var}\left(\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}\right)\right)$. Elementary manipulations and (B.0.10) give, with $\min\left(t_1, t_2\right) = t_1 \wedge t_2$ and uniform remainder terms with respect to $\|\cdot\|$,

$$\begin{aligned}
\mathsf{V}\left(\mathsf{b}\right) = I^{-1}\mathbb{E}\Bigg[&\mathbb{I}\left(I_\ell = I\right) P\left(x_\ell\right) P\left(x_\ell\right)' \otimes \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left\{G\left(P\left(x_\ell, t_1\right)' \mathsf{b} \wedge P\left(x_\ell, t_2\right)' \mathsf{b} | x_\ell, I_\ell\right)\right. \\
&- 2\left(\alpha + ht_1\right) G\left(P\left(x_\ell, t_2\right)' \mathsf{b} | x_\ell, I_\ell\right) + \left(\alpha + ht_1\right)\left(\alpha + ht_2\right)\big\} \\
&\times \pi\left(t_1\right) \pi\left(t_2\right)' K\left(t_1\right) K\left(t_2\right) dt_1 dt_2\Bigg] \\
&= I^{-1}\mathbf{P} \otimes \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left\{\left(\alpha + ht_1\right) \wedge \left(\alpha + ht_2\right) - \left(\alpha + ht_1\right)\left(\alpha + ht_2\right)\right\} \\
&\times \pi\left(t_1\right) \pi\left(t_2\right)' K\left(t_1\right) K\left(t_2\right) dt_1 dt_2 + o\left(h\right) \\
&= \frac{\alpha\left(1 - \alpha\right)}{I} \mathbf{P} \otimes \pi_0 \pi_0' + h\left\{\mathbf{P} \otimes \mathbf{\Pi}_m - \alpha\mathbf{P} \otimes \left(\pi_1 \pi_0' + \pi_0 \pi_1'\right)\right\} + o\left(h\right).
\end{aligned}$$

Hence an elementary expansion gives, uniformly in $\alpha \in [0,1]$, $\operatorname{Var}(\widehat{\mathsf{e}}) = \mathsf{V}_e + o(h)$ with

$$
\begin{aligned}
\mathsf{V}_e &= \alpha(1-\alpha)\mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1} \times \mathbf{P} \otimes \pi_0\pi_0' \times \mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1} \\
&- h\alpha(1-\alpha)\left(\mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1}\right)^2 \times \mathbf{P}_1 \otimes \mathbf{\Pi}_1 \times \mathbf{P} \otimes \pi_0\pi_0' \times \mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1} \\
&- h\alpha(1-\alpha)\mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1} \times \mathbf{P} \otimes \pi_0\pi_0' \times \mathbf{P}_1 \otimes \mathbf{\Pi}_1 \times \left(\mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1}\right)^2 \\
&+ h\mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1} \times \{\mathbf{P} \otimes \mathbf{\Pi}_m - \alpha\mathbf{P} \otimes (\pi_1\pi_0' + \pi_0\pi_1')\} \times \mathbf{P}_0^{-1} \otimes \mathbf{\Pi}_0^{-1} \\
&= \alpha(1-\alpha)\left(\mathbf{P}_0^{-1} \times \mathbf{P} \times \mathbf{P}_0^{-1}\right) \otimes \left(\mathbf{\Pi}_0^{-1} \times \pi_0\pi_0' \times \mathbf{\Pi}_0^{-1}\right) \\
&- h\alpha(1-\alpha)\left(\mathbf{P}_0^{-2} \times \mathbf{P}_1 \times \mathbf{P} \times \mathbf{P}_0^{-1}\right) \otimes \left(\mathbf{\Pi}_0^{-2} \times \mathbf{\Pi}_1 \times \pi_0\pi_0' \times \mathbf{\Pi}_0^{-1}\right) \\
&- h\alpha(1-\alpha)\left(\mathbf{P}_0^{-1} \times \mathbf{P} \times \mathbf{P}_1 \times \mathbf{P}_0^{-2}\right) \otimes \left(\mathbf{\Pi}_0^{-1} \times \pi_0\pi_0' \times \mathbf{\Pi}_1 \times \mathbf{\Pi}_0^{-2}\right) \\
&+ h\left(\mathbf{P}_0^{-1} \times \mathbf{P} \times \mathbf{P}_0^{-1}\right) \otimes \left(\mathbf{\Pi}_0^{-1} \times \mathbf{\Pi}_m \times \mathbf{\Pi}_0^{-1} - \alpha\mathbf{\Pi}_0^{-1} \times (\pi_1\pi_0' + \pi_0\pi_1') \times \mathbf{\Pi}_0^{-1}\right)
\end{aligned}
$$

Observe now that

$$S_{1,ASQR} = \mathrm{Id}_{K_L} \otimes s_1' \text{ with } s_1 = [0,1,0,\ldots,0]' \in \mathbb{R}^{s+2}.$$

Let $s_0 = [1,0,0,\ldots,0]' \in \mathbb{R}^{s+2}$. The key point here is that $s_0's_1 = s_1's_0 = 0$ gives

$$\mathbf{\Pi}_0^{-1}\pi_0 = s_0 \text{ so that } \pi_0'\mathbf{\Pi}_0^{-1}s_1 = s_1'\mathbf{\Pi}_0^{-1}\pi_0 = 0.$$

Because $\mathbf{\Pi}_0^{-1}s_1 = \Pi_h^t(\alpha)$, the expression of V_e gives, uniformly in $\alpha \in [0,1]$,

$$
\begin{aligned}
&\operatorname{Var}(D_{1,ASQR}\widehat{\mathsf{e}}) = I^{-1}(\mathrm{Id}_{K_L} \otimes s_1') \times \mathsf{V}_e \times (\mathrm{Id}_{K_L} \otimes s_1) + o(h) \\
&= I^{-1}h\left(\mathbf{P}_0^{-1} \times \mathbf{P} \times \mathbf{P}_0^{-1}\right) \otimes s_1'\mathbf{\Pi}_0^{-1} \times \mathbf{\Pi}_m \times \mathbf{\Pi}_0^{-1}s_1 + o(h) \\
&= h\frac{v_h^2(\alpha)}{I}\left(\mathbf{P}_0^{-1} \times \mathbf{P} \times \mathbf{P}_0^{-1}\right) + o(h) \\
&= h\frac{v_h^2(\alpha)}{I}\mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)P(x_\ell)P(x_\ell)'}{B^{(1)}(\alpha|x_\ell, I)}\right]\mathbb{E}^{-1}\left[\mathbb{I}(I_\ell = I)P(x_\ell)P(x_\ell)'\right] \\
&\qquad \times \mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)P(x_\ell)P(x_\ell)'}{B^{(1)}(\alpha|x_\ell, I)}\right] + o(h),
\end{aligned}
$$

as stated in the Lemma. $\square$

Proof of Proposition B.1. It is sufficient to study $\widehat{\mathsf{O}}_{1,ASQR}(\mathsf{d},\mathsf{e};\alpha,I)$

and $\widehat{\mathsf{O}}_{ASQR}(\mathsf{d},\mathsf{e};\alpha,I)$ which are abbreviated into $\widehat{\mathsf{O}}_1(\mathsf{d},\mathsf{e})$ and $\widehat{\mathsf{O}}(\mathsf{d},\mathsf{e})$, as $\widehat{\mathsf{R}}_{ASQR}(\mathsf{b}_{ASQR}(\alpha|I)+\mathsf{e};\alpha,I)$ which will be denoted $\widehat{\mathsf{R}}_0(\mathsf{e})$. Observe that,

$$\begin{aligned}\widehat{\mathsf{O}}(\mathsf{d},\mathsf{e}) &= \int_0^1 \left\{\widehat{\mathsf{R}}_0^{(1)}(\mathsf{e}+u\mathsf{d}) - \widehat{\mathsf{R}}_0^{(1)}(\mathsf{e})\right\}' du\mathsf{d} + \left\{\widehat{\mathsf{R}}_0^{(1)}(\mathsf{e}) - \widehat{\mathsf{R}}_0^{(1)}(0)\right\}' \mathsf{d} \\ &:= \frac{1}{L}\sum_{\ell=1}^{L}\left(f(\xi_\ell;\tau) + g(\xi_\ell;\tau)\right),\end{aligned}$$

where τ stands for α, e, and d and,

$$\begin{aligned}f(\xi_\ell;\tau) &= \int_0^1 \left\{\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \mathbb{I}_\ell(t,u;\tau)\,\mathsf{d}'P(x_\ell,t)\,K(t)\,dt\right\} du\mathbb{I}(I_\ell = I), \\ \mathbb{I}_\ell(t,u;\tau) &= \sum_{i=1}^{I_\ell}\left\{\mathbb{I}\left(B_{i\ell} \le P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell)+\mathsf{e}+u\mathsf{d})\right) - \mathbb{I}\left(B_{i\ell} \le P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell)+\mathsf{e})\right)\right\} \\ &= \operatorname{sgn}\left(P(x_\ell,t)'\mathsf{d}\right)\sum_{i=1}^{I_\ell}\mathbb{I}\left(B_{i\ell} - P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell)+\mathsf{e}) \text{ between } 0 \text{ and } uP(x_\ell,t)'\mathsf{d}\right), \\ g(\xi_\ell;\tau) &= \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \sum_{i=1}^{I_\ell}\left\{\mathbb{I}\left(B_{i\ell} \le P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell)+\mathsf{e})\right) - \mathbb{I}\left(B_{i\ell} \le P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell))\right)\right\} \\ &\quad \times \mathsf{d}'P(x_\ell,t)\,K(t)\,dt\mathbb{I}(I_\ell = I).\end{aligned}$$

It then follows that, under Assumptions H-(i), R-(ii) and because the conditional p.d.f of $B_{i\ell}$ is bounded,

$$\begin{aligned}\operatorname{Var}(f(\xi_\ell;\tau)) &\le CI\int_0^1\left\{\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \mathbb{E}\left[\left(\mathbb{I}_\ell(t,u;\tau)\,\mathsf{d}'P(x_\ell,t)\,K(t)\right)^2\right]dt\right\}du \\ &\le C\int_0^1\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \mathbb{E}\left[\mathbb{P}\left(B_{i\ell} - P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell)+\mathsf{e}) \text{ between } 0 \text{ and } P(x_\ell,t)'\mathsf{d}|x_\ell,I_\ell\right)\right. \\ &\qquad \left.\left(\mathsf{d}'P(x_\ell,t)\,K(t)\right)^2\right]dtdu \\ &\le C\mathbb{E}\left[\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left(\mathsf{d}'P(x_\ell,t)\right)^3K^2(t)\,dt\right] \le C\|\mathsf{d}\|\max_{x\in\mathcal{X}}\|P(x)\|\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\mathbb{E}\left[\left(\mathsf{d}'P(x_\ell,t)\right)^2\right]K^2(t)\,dt \\ &\le C\|\mathsf{d}\|^3\max_{x\in\mathcal{X}}\|P(x)\| \le C\overline{P}_L\varrho_d^3,\end{aligned}$$

while $|f(\xi_\ell;\tau)| \le C\overline{P}_L \varrho_d$ uniformly. Similarly

$$\operatorname{Var}\left(g(\xi_\ell;\tau)\right) \le C\overline{P}_L \varrho_d^2 \varrho_e, \quad \max_\tau |g(\xi_\ell;\tau)| \le C\overline{P}_L \varrho_d.$$

Then arguing as in the proof of Lemma B.3 gives

$$\mathbb{E}\left[\max_\tau \left|\frac{1}{L}\sum_{\ell=1}^{L} f(\xi_\ell;\tau)\right|\right] \le C\left(\left(\frac{K_L \overline{P}_L \varrho_d^3 \log L}{L}\right)^{1/2} + \frac{K_L \overline{P}_L \varrho_d \log L}{L}\right),$$
$$\mathbb{E}\left[\max_\tau \left|\frac{1}{L}\sum_{\ell=1}^{L} g(\xi_\ell;\tau)\right|\right] \le C\left(\left(\frac{K_L \overline{P}_L \varrho_d^2 \varrho_e \log L}{L}\right)^{1/2} + \frac{K_L \overline{P}_L \varrho_d \log L}{L}\right),$$

which shows that the Proposition is proved. □

Proof of Proposition B.2. The notations of the proof of Proposition B.1 are used and the proof is detailed for the ASQR case. Observe first that,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{I}_\ell(t,u;\tau)\,|x_\ell, I_\ell = I\right] \\
&\quad = I\left\{G\left(B_{i\ell} \le P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell) + \mathsf{e}{+}u\mathsf{d})\,|x_\ell, I\right) - G\left(B_{i\ell} \le P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell) + \mathsf{e})\,|x_\ell, I\right)\right\} \\
&\quad = I\int_0^1 g\left(B_{i\ell} \le P(x_\ell,t)'(\mathsf{b}(\alpha|I_\ell) + \mathsf{e}{+}vu\mathsf{d})\,|x_\ell, I\right) uP(x_\ell,t)'\,\mathsf{d}dv.
\end{aligned}$$

It then follows that

$$\begin{aligned}
\mathsf{O}_1(\mathsf{d},\mathsf{e};\alpha,I) &= \mathsf{d}'\int_0^1\int_0^1 \mathsf{R}^{(2)}(\mathsf{b}(\alpha|I_\ell) + \mathsf{e}{+}vu\mathsf{d};\alpha,I)\,vdvdu\mathsf{d} \\
&\quad + \mathsf{d}'\int_0^1 \mathsf{R}^{(2)}(\mathsf{b}(\alpha|I_\ell) + u\mathsf{e};\alpha,I)\,dv\mathsf{e}.
\end{aligned}$$

Hence (B.0.8) gives,

$$\left| \mathsf{O}_1(\mathsf{d},\mathsf{e};\alpha,I) - \frac{\mathsf{d}'\mathsf{R}^{(2)}(\mathsf{b}(\alpha|I_\ell);\alpha,I)(\mathsf{d}+2\mathsf{e})}{2} \right|$$
$$\leq \left\| \int_0^1 \int_0^1 \left\{ \mathsf{R}^{(2)}(\mathsf{b}(\alpha|I_\ell)+\mathsf{e}+vu\mathsf{d};\alpha,I) - \mathsf{R}^{(2)}(\mathsf{b}_{ASQR}(\alpha|I_\ell);\alpha,I) \right\} vdvdu \right\| \varrho_d^2$$
$$+ \left\| \int_0^1 \left\{ \mathsf{R}^{(2)}(\mathsf{b}(\alpha|I_\ell)+u\mathsf{e};\alpha,I) - \mathsf{R}^{(2)}(\mathsf{b}(\alpha|I_\ell);\alpha,I) \right\} du \right\| \varrho_d\varrho_e$$
$$\leq C \max_{x\in\mathcal{X}} \|P(x)\| \left(\varrho_d^2(\varrho_d+\varrho_e) + \varrho_d\varrho_e^2 \right) \leq C\overline{P}_L (\varrho_{d+}\varrho_e)^2 \varrho_d,$$

uniformly with respect to $(\alpha, \mathsf{d}, \mathsf{e}, I)$ in $[0,1] \times \mathcal{B}(0,\varrho_d) \times \mathcal{B}(0,\varrho_e) \times \mathcal{I}$. □

Proof of Lemma B.5. The variance results follows from Lemma B.3-(ii) and Lemma B.4-(ii) together with Assumption R-(ii) and the fact that $B^{(1)}(\alpha|x,I)$ is bounded away from 0 and infinity by Proposition 3.1. For the maximum order, consider the first one and define with $S_0 = S_{0,ASQR}$,

$$\widehat{\mathsf{r}}_0(\alpha,x) = \frac{P(x)'}{1+\|P(x)\|} S_0 \widehat{\mathsf{e}}_{ASQR}(\alpha|I) := \frac{1}{L}\sum_{\ell=1}^{L} f_0(\xi_\ell,\theta),$$

with $\theta = (\alpha, x')'$ and,

$$f_0(\xi_\ell,\theta) = -\frac{\mathbb{I}(I_\ell = I) P(x)'}{1+\|P(x)\|} S_0 \mathsf{R}^{(2)}(\mathsf{b}(\alpha|I);\alpha,I)^{-1}$$
$$\times \sum_{i=1}^{I_\ell} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left\{ \mathbb{I}\left(B_{i\ell} \leq P(x_\ell,t)' \mathsf{b}(\alpha|I) \right) - (\alpha+ht) \right\} P(x_\ell) K(t)\, dt,$$

which is centered by (B.0.5) and satisfies, since $\mathsf{R}^{(2)}(\mathsf{b}(\alpha|I);\alpha,I)^{-1} \preceq C\,\mathrm{Id}$,

$$\mathrm{Var}(f_0(\xi_\ell,\theta)) \leq C\frac{\|P(x)\|^2}{(1+\|P(x)\|)^2} \leq C,$$
$$|f_0(\xi_\ell,\theta)| \leq C\frac{\|P(x)\|\,\|P(x_\ell)\|}{1+\|P(x)\|} \leq C\max_{x\in\mathcal{X}}\|P(x)\| = O\left(K_L^{1/2}\right).$$

Under Assumption H-(iii), $P(x)/(1+\|P(x)\|)$ is a Hölder function with exponent p and constant K_{2n} satisfying $\log K_{2n} = O(\log L)$. By arguments as in the proof of

Lemmas B.3 and B.4, the bracketing entropy $\mathbf{H}(\epsilon)$ to cover $\{f(\xi,\theta), \theta \in [0,1] \times \mathcal{X}\}$ in Theorem B.3 satisfies $\mathbf{H}(\epsilon) \leq C \log(L/\epsilon)$. Hence using the bound for $\int_0^C \mathbf{H}^{1/2}(t)\, dt$ of Lemma B.3 gives

$$\mathbb{E}\left[\max_{\theta \in [0,1]\times\mathcal{X}} \left|\widehat{\mathsf{r}}_0(\alpha, x)\right|\right] \leq C\left(\left(\frac{\log L}{L}\right)^{1/2} + \frac{K_L^{1/2}\log L}{L}\right) \leq C\left(\frac{\log L}{L}\right)^{1/2},$$

since $K_L = O(\log L/L)$ under Assumption H-(ii). The Markov inequality then gives

$$\begin{aligned}\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|P'(x)\, S_{0,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I)\right| &\leq \left(1 + \max_{x\in\mathcal{X}} \|P(x)\|\right) \max_{\theta\in[0,1]\times\mathcal{X}} \left|\widehat{\mathsf{r}}_0(\alpha, x)\right| \\ &= O_{\mathbb{P}}\left(\left(\frac{K_L \log L}{L}\right)^{1/2}\right).\end{aligned}$$

The proof of the second bound proceeds similarly starting from

$$\widehat{\mathsf{r}}_1(\alpha, x) = \frac{P'(x)}{1 + \|P(x)\|} S_{1,ASQR}\widehat{\mathsf{e}}_{ASQR}(\alpha|I).$$

Arguing as above gives

$$\mathbb{E}\left[\max_{\theta\in[0,1]\times\mathcal{X}} \left|\widehat{\mathsf{r}}_1(\alpha, x)\right|\right] \leq C\left(h^{1/2}\left(\frac{\log L}{L}\right)^{1/2} + \frac{K_L^{1/2}\log L}{L}\right) \leq Ch^{1/2}\left(\frac{\log L}{L}\right)^{1/2},$$

since $K_L = O(Lh/\log L)$. The bound above implies the last bound of Lemma B.5.□

Bibliography

Abberger, K. (1998). Cross validation in nonparametric quantile regression. *Allgemeines statistisches Archiv,* **82**, 146-161.

Andrews, D.W.K & Y.J. Wang (1990). Additive interactive regression models: circumventing the curse of dimensionality. *Econometric Theory* **6**, 466–479.

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

Athey, S. & P. Haile (2002). Identification of standard auction models. *Econometrica,* **70**, 2107-2140.

Athey, S. & P.A. Haile (2007). Nonparametric approaches to auctions. *Handbook of Econometrics,* Vol. **6a**, Chap. 60, 3847-3965.

Athey, S., J. Levin (2001). Information and competition in U.S. forest service timber auctions. *Journal of Political Economy,* **109,** 375-417.

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

Bajari, P. & A. Hortaçsu (2005). Are structural estimates of auction models reasonable? Evidence from experimental data. *The Journal of Political Economy*, **113**, 703-741.

Baldwin, L., R. Marshall & J.F. Richard (1997). Bidder collusion at Forest Service Timber Sales. *Journal of Political Economy,* **105**, 657-699.

Bierens, H.J. & H. Song (2006). Nonparametric identification of first-price auction models with unbounded values and observed auction-specific heterogeneity. Working Paper, Pennsylvania State University.

Bikhchandani, S., P.A. Haile & J.G. Riley (2002). Symmetric separating equilibria in English auctions. *Games and Economic Behavior,* **38**, 19-27.

Buchinsky, M. (1995). Estimating the asymptotic covariance matrix for quantile regression models: a Monte Carlo study. *Journal of Econometrics,* **68**, 303-338.

Bulow, J. & P. Klemperer (2009). Why do sellers (usually) prefer auctions. *American Economic Review,* **99**, 1544-1575.

Campo, S. (2012). Risk aversion and asymmetry in procurement auctions: identification, estimation and application to construction procurements. *Journal of Econometrics*, **168**, 96-107.

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.

Canay, I.A.(2011). A simple approach to quantile regression for panel data. *The Econometrics Journal,* **14,** 368-386.

Cantillon, E.(2008). The effect of bidders' asymmetries on expected revenue in auctions. *Games and Economic Behavior*, **62,** 1-25.

Cassola, N., A. Hortaçsu & J. Kastl (2013). The 2007 subprime market crisis through the lens of European Central Bank auctions for short-term funds. *Econometrica* **81**, 1309-1345.

Charnigo, R., B. Hall & C. Srinivasan (2013). A generalized C_p Criterion for derivative estimation. *Technometrics*, **53**, 238-253.

Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. Chap. 76 in *Handbook of Econometrics,* Vol. **6B**. Elsevier.

Chernozhukov, V. (2005). Extremal quantile regression. *The Annals of*

Statistics **13**, 806-839.

Chernozhukov, V. & C. Hansen(2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics,* **132**, 491-525.

Daubechies, I. (1992). *Ten lectures on wavelets.* SIAM.

Donald, S. & H. Paarsch (1993). Piecewise pseudo-maximum likelihood estimation in empirical models of auctions. *International Economic Review,* **34,** 121-148.

Donald, S. & H. Paarsch (1996). Identification, estimation, and testing in parametric empirical models of auctions within the independent private values paradigm. *Econometric Theory,* **12,** 517-567.

Donald, S. & H. Paarsch (2002). Superconsistent estimation and inference in structural econometric models using extreme order statistics. *Journal of Econometrics,* **109,** 305-340.

Eggermont, P.P.B. & V.N. LaRiccia (2009). *Maximum penalized likelihood estimation. Volume II: Regression.* Springer.

Enache, A. & J.P Florens (2012). Identification and estimation in a third-price auction model. Working Paper, Paris School of Economics.

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

Galvao Jr., A. (2011). Quantile regression for dynamic panel data with fixed effects. *Journal of Econometrics,* **164,** 142-157.

Gavious, A. & Y. Minchuk (2012). A note on the effect of asymmetry on revenue in second-price auctions. *International Game Theory Review,* **14,** Issue 03.

Gentry, M. & T. Li (2012). Identification in auctions with selective entry. Working Paper. Revised and resubmitted to *Econometrica.*

Goldenshluger, A. & O. Lepski (2011). Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality. *The Annals of Statistics*, **39**, 1608-1632.

Gragg, W.B. & R.A. Tapia (1974). Optimal error bounds for the Newton-Kantorovich Theorem. *SIAM Journal on Numerical Analysis* **11**, 10–13.

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

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

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.

Guerre, E. & C. Sabbah (2013). Work in Progress.

Guerre, E. & Q. Vuong (2013). Work in Progress.

Haile, P.A. (2001). Auctions with resale markets: an application to U.S. Forest Service Timber Sales. *American Economic Review,* **91**, 399-427.

Haile, P.A. & E. Tamer (2003). Inference with an incomplete model of English auctions. *Journal of Political Economy,* **111**, 1-51.

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

Hastie, T. & R. Tibshirani (1985). Generalised additive models. *Statistical Science* **1**, 297-310.

Hickman, B.R., T.P. Hubbard & Y. Sağlam (2012). Structural econometric methods in auctions: a guide to the literature. *Journal of Econometric Methods,* **1**, 67-106.

Horowitz, J.L. & S. Lee (2005). Nonparametric estimation of an additive quantile regression model. *Journal of the American Statistical Association* **100**, 1238-1249.

HOROWITZ, J. & V. SPOKOINY (2001). An adaptive, rate-optimal test of a para- metric mean-regression model against a nonparametric alternative. *Econometrica,* **69**, 599-631.

HU, A., S.A. MATTHEWS & L. ZOU (2010). Risk aversion and optimal reserve prices in first- and second-price auctions. *Journal of Economic Theory,* **145**, 1188-1202.

JONES, D.S. (2008). *The theory of generalised functions.* Cambridge University Press.

KNIGHT, K.(1998). Limiting distributions for L_1 regression estimators under general conditions. *Annals of Statistics,* **26**, 755-770.

KOENKER, R. (2005). *Quantile regression.* Cambridge University Press.

KOENKER, R. & G. BASSETT (1978). Regression quantiles. *Econometrica,* **46**, 33-50.

KOENKER, R. & J. MACHADO (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association,* **94,** 1296-1310.

KOENKER, R., P. NG & S. PORTNOY (1994). Quantile smoothing splines. *Biometrika,* **81**, 673-680.

KRISHNA, V. (2009). *Auction Theory.* 2nd edition, Academic Press.

KULLBACK, S. & R.A. LEIBLER (1951). On information and sufficiency. *Annals of Mathematical Statistics,* **22**, 79-86.

LAFFONT, J.J., H. OSSARD & Q. VUONG (1995). Econometrics of first-price auctions. *Econometrica* **63**, 953-980.

LEVIN, D. & J. SMITH (1994). Equilibrium in auctions with entry. *The American Economic Review,* **84,** 585-599.

LI, Q., J. LIN & J.S. RACINE (2013). Optimal bandwidth selection for nonparametric conditional distribution and quantile functions. *Journal of Business and Economic Statistics,* **31**, 57-65.

Li, T. (2005). Econometrics of first-price auctions with entry and binding reservation price. *Journal of Econometrics,* **126**, 173-200.

Li, T. & I. Perrigne (2003). Timber sale auctions with random reserve prices. *The Review of Economics and Statistics,* **85**, 189-200.

Li, T., I. Perrigne & Q. Vuong (2003). Semiparametric estimation of the optimal reserve price in first-price auctions. *Journal of Business and Economic Statistics,* **21**, 53-64.

Li, T. & X. Zheng (2009). Entry and competition effects in first-price auctions: theory and evidence from procurement auctions. *Review of Economic Studies,* **76**, 1397-1429.

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.

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.

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

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

Myerson, R.(1981). Optimal auction design. *Mathematics of Operations Research,* **6**, 58-63.

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

Massart, P. (2007). *Concentration inequalities and model selection.* Lectures Notes in Mathematics 1986. Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003, Jean Picard (ed.). Springer-Verlag.

Müller, H.-G., U. Stadtmüller & T. Schmitt (1987). Bandwdth choice and confidence intervals for derivatives of noisy data. *Biometrika*, **74**, **7**, 743-749.

Newey, W.K. (1997). Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics* **79**, 147-168.

Newey, W. & D. McFadden (1994). Large sample estimation and hypothesis testing. Chap. 36 In *Handbook of Econometrics*, Vol. **4**. Elsevier.

Paarsch, H. (1992). Deciding between the common and private value paradigms in empirical models of auctions. *Journal of Econometrics,* **51**, 191-215.

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory*, **7**, 186-199.

Portnoy, S. (1997). Local asymptotics for quantile smoothing splines. *The Annals of Statistics,* **25**, 414-434.

Powell, J.L. (1986). Censored regression quantiles. *Journal of Econometrics* **32**, 143-155.

Powell, J.L. (1991). Estimation of monotonic regression models under quantile restrictions, in *Nonparametric and Semiparametric Methods in Econometrics,* (ed. by W. Barnett, J.L. Powell and G. Tauchen), Cambridge: Cambridge University Press.

Pratt, J. (1964). Risk aversion in the small and in the large. *Econometrica,* **32**, 122-136.

Rao, C.R., & L.C. Zhao (1992). Approximation to the distribution of M-estimates in linear models by randomly weighted bootstrap. *Sankhya* **54**, 323-331.

Reiss, P.T. & L. Huang (2012). Smoothness selection for penalized quantile regression splines. *The International Journal of Biostatistics,* **8**, Issue 1, Article 10.

Rezende, L. (2008). Econometrics of auctions by least squares. *Journal of Applied Econometrics* **23**, 925-948.

RICE, J.A. (1986). Bandwidth choice for differentation. *Journal of Multivariate Analysis*, **19**, 251-264.

RILEY, J. & W. F. SAMUELSON (1981). Optimal auctions. *The American Economic Review,* **71,** 381-392.

ROBERTS, J. & A. SWEETING (2012). When Should Sellers Use Auctions? Working Paper, Duke University.

SILVERMAN, B.W. (1998). Density estimation for statistics and data analysis. London: Chapman & Hall/CRC. Monographs on Statistics & Applied Probability.

VAN DE GEER, S. (1999). *Empirical processes in M-estimation.* Cambridge University Press.

VAN DER VAART, A.W. (1998). Asymptotic Statistics. Cambridge University Press, Cambridge.

VUONG, Q. (1989). Likelihood ratio test for model selection and non-nested hypotheses. *Econometrica,* **57**, 307-333.

WANG, X.M. & W. ZHOU (2004). Bootstrap approximation to the distribution of M-estimates in a linear model. *Acta Mathematica Sinica,* **20**, 93-104.

WHANG, Y.J.(2006). Consistent specification testing for quantile regression models, in *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, ed. by D. Corbae, S. Durlauf, and B. Hansen, Cambridge University Press.

WHITE, H. (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica,* **50**, 1-25.

ZHAO, L., X. WU & Y. YANG (2007). Approximation by random weighting method for M-test in linear models. *Science in China Series A: Mathematics,* **50**, 87-99.

ZHENG, J.(1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory,* **14**, 123-138.

ZINCENKO, F. (2013). Nonparametric sieve estimation in first-price auctions with risk-averse bidders. Document paper, UCLA.

ZOU, H. & M. YUAN (2008). Composite quantile regression and the oracle model selection theory. *The Annals of Statistics,* **36**, 1108-1126.