Quantile methods for first-price auction: A signal approach

Nathalie Gimenes
Department of Economics
University of São Paulo
Brazil

Emmanuel Guerre
School of Economics and Finance
Queen Mary University of London
Great Britain

September 2016 - (First version: November 2013)
Preliminary version - Do not quote

Abstract

This paper considers a quantile signal 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. The choice of the smoothing parameters is also discussed. Extensions to interdependent values including bidder specific variables are also possible under some functional restrictions, which tie up the bidder covariate and signal. The identification of this new model is established and some estimation methods are suggested.

JEL: C14, L70

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

A previous version of this paper has been circulated under the title "Quantile regression methods for first-price auction:a signal approach". The authors acknowledge useful discussions and comments from Christian Bontemps, Xiaohong Chen, Pierre Dubois, Yanqin Fan, Jean-Pierre Florens, Eric Gautier, Phil Haile, Xavier d'Haultfoeuille, Asen Ivanov, Efang Kong, Hongjun Li, Qi Li, Nianqing Liu, Gauthier Lanot, Simon Lee, Aniol Llorente-Saguer, Adam McCloskey, Vadim Marmer, Nour Meddahi, Lars Neisheim, Andriy Norets, Elena Pastorino, Aureo de Paula, Isabelle Perrigne, Martin Pesendorfer, Eric Renault, Christoph Roth, Camille Sabbah, Suzanne Schennach, Art Shneyerov, Kevin Song, Quang Vuong, Martin Weidner, Daniel Wilhem, Pai Xu, the participants of AMES2014, LAMES2014, 2014 Non and Semiparametric methods mini conference in York, 2014 Canadian Econometrics Study Group in Vancouver, AMES2016, the 6$^{\text{th}}$ Shanghai Econometrics Workshop, the 2$^{\text{nd}}$ Dongbei Econometrics Workshop, the Séminaire Malinvaud, Queen Mary Economic Reading Group, New York University Econometrics Lunch, Kent, Brown, UBC, Yale, Glasgow, Umeå, CEMMAP, HKU, NUS, CUEB and TSE seminars. All remaining errors are our responsibility. Both authors would like to thank the School of Economics and Finance, Queen Mary University of London, for generous funding.

1 Introduction

1.1 Contents of the paper

This work aims to introduce new quantile based identification and estimation techniques tailored for the econometrics of first-price auction within a signal framework. The proposed approach builds on the monotonicity of bidding strategies. For instance, in the symmetric signal framework of Milgrom and Weber (1981), the optimal unique bidding strategy is given by an increasing function of the bidder's signal as established in McAdams (2007). Because the information content of the signal is unchanged after a one to one transformation, each bidder's signal can be normalized to be a uniform random variable. Since the bids are equal to the strategy taken at each bidder signal, it is easily seen that, under this signal normalization, the probability that a bid exceeds b is the inverse of the strategy taken at b. As a consequence, the bidding strategy is identical to the bid quantile function, which is identified from the observation of the bids. The signal of a bidder can also be recovered by applying the bid cumulative distribution (cdf hereafter) to his bid. Such elementary consequences of the monotonicity of the bidding strategy suggests that quantile techniques can play an important role in the econometrics of auctions, as already observed among others by Marmer and Shneyerov (2012) for symmetric independent private value and Hong and Shum (2002) for the common value Wilson model. The approach proposed here better explores the implications of optimal bidding for the bid quantile function. Indeed, the bid quantile function can replaced the bidding strategy in the best response condition, which characterizes optimal bidding and which has been used intensively in the econometrics of auction, see Athey and Haile (2007) for a survey of such developments.

1.1.1 The symmetric independent private value case

Under risk neutrality and independent private value (IPV), the expected profit of a bid is the difference of the bidder private value to the bid time the probability of winning. Under symmetry, the probability of winning is very simple and is equal to the bid quantile level at a power given by the number of opponents minus 1. In this expression, the private value can be written as the private value quantile function taken at the signal, which is also the optimal bid quantile for bids drawn from the Bayesian Nash equilibrium. The associated first order condition (FOC hereafter) give a linear differential equation which identifies the private value quantile function as a simple linear combination of the bid quantile function and its derivative times the quantile level. This contrasts with the nonlinear identification of the private value distribution of Guerre, Perrigne and Vuong (2000, GPV hereafter). Furthermore, solving this linear differential equation shows that the bid quantile function is a linear functional of the private value quantile function. It follows that postulating a linear specification for the private value quantile function implies that the bid quantile function belongs to a similar linear specification, as noticed by Haile, Hong and Shum (2003, HHS hereafter) or Rezende (2008) for the particular case of a linear regression specification. This suggests that quantile linear specification derived from the popular quantile regression of Koenker and Bassett

(1978) can play a central role in the econometrics of auctions. As noted in Gimenes (2016), a simple regression model may be indeed insufficient to capture interactions between the signal and good covariate in a standard timber ascending auction dataset. Building on the nonparametric additive quantile model of Horowitz and Lee (2005) and on quantile versions of the additive interactive regression of Andrews and Whang (1990), the paper proposes a full scale of quantile specifications, ranging from additive to fully saturated ones. This family of specifications is expected to be useful in small sample with many covariates. A new local polynomial estimation method, called *Augmented Quantile Regression* (AQR) later on, is proposed to jointly estimate a quantile regression model jointly with its derivatives, as requested to estimate the private value quantile function. This local polynomial methodology is combined with sieve methods to estimate additive interactive specification of the private value quantile function. The AQR methodology is not affected by boundary bias delivering an asymptotically unbiased estimation of the upper part of private value quantile function, as desirable to estimate the value of the winner. Bandwidth choices are also simpler than in GPV.

1.1.2 The heterogeneous interdependent value case

A bidder's interdependent value can depend upon the signals of the other bidders. As known since Laffont and Vuong (1996) in the symmetric case, this feature considerably complicates identification as first-price auction bids can be rationalized with an IPV model, so that common value models are not nonparametrically identified.

Our approach differs and makes identification possible by introducing bidder specific observed by the econometrician as in Somaini (2015). The considered parameters are the signal distribution and the valuation functions. As in the private value case, the signals can be easily recovered assuming strictly increasing strategies. The signal distribution can be identified, up to censoring issues due to potential aggressive bidding crowding out bidders with low signal. The focus is on the identification of the valuation function of a specific bidder, say bidder 1, which is supposed to use a best response strategy. The other bidders do not need to have a valuation function or to bid optimally, but it is assumed that the bidder characteristic partial derivatives of the strategies satisfy a rank condition stating that bidder asymmetry is strong enough. The first bidder valuation function may depend upon all bidder characteristic provided it interacts with the bidder signal in a multiplicative way, a functional restriction which yields identification. Such a restriction holds for instance in some simple auction with resale. Quantile based estimation procedures are also proposed and can be simple to implement for linear valuation functions.

1.2 Relation with existing literature

1.2.1 Quantile approaches under IPV

HHS were probably the first to use quantile estimation in an auction setup, to test common versus private value. The independent private value part of the paper is probably more

related to Marmer and Shneyerov (2012), who have proposed the first nonparametric quantile based framework for the estimation of the private value probability density function (pdf). However focusing on the private value quantile function instead of the pdf can simplify estimation procedure as proposed in Guerre and Sabbah (2012), see also Luo and Wan (2016) for an increasing version of the private value quantile estimator of these authors. Marmer, Shneyerov and Xu (2013a) and Liu and Luo (2014) have used a nonparametric estimation of quantiles to test for selective entry. Liu and Vuong (2016) also use a quantile approach to test monotonicity of bidding strategy.

Guerre, Perrigne and Vuong (2009), Campo, Guerre, Perrigne and Vuong (2011) and Zincenko (2013) use a quantile approach to identify and estimate risk aversion in first price auctions. Lee, Song and Whang (2014) have proposed to test some related inequalities using quantile. See also Bajari and Hortaçsu (2005) for experimental data. For ascending auction, Menzel and Morganti (2013) developed an approach based on the order statistic, which is the collection of the sample quantile. The quantile regression specification considered here has been considered in Gimenes (2016), and gave a good fit for the considered application to timber auction. Enache and Florens (2012) have developed a quantile framework for a third-price auction model.

1.2.2 Dimension reduction

Many auction samples include many covariate and have a sample size which is not compatible with a fully nonparametric approach. Haile and Tamer (2003) or Aradillas-Lopez, Gandhi and Quint (2013) have considered auction samples with 5 or 6 explanatory variables for at 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. A fully nonparametric approach does not seem reasonable in this setup.

To address this issue, HHS have introduced a bid homogenization technique which has been implemented in many applications. This amounts to consider a linear regression model with independent error terms for the private value, see also Rezende (2008). In this model, the uniform signal is a normalization of the regression error term and the signal cannot interact with the covariate. This feature is potentially restrictive for applications. In the case of ascending auction, Gimenes (2016) has estimated a quantile regression specification, which allows for interactions between the covariate and the signal. These interactions were found significant for a popular timber auction dataset. As detailed in Section 2.1.1, the bid homogenization technique does not apply in this case.

An alternative parsimonious specification is the regression single-index model proposed in Paarsch and Hong (2006) for private value. However, as for the standard regression model, a single index specification does not allow for signal-covariate interaction. Marmer, Shneyerov and Xu (2013b) have recently proposed a quantile single-index model which allows for such interactions. Although this approach is promising, its estimation looks at first sight more involved that the estimation procedures considered here.

1.2.3 Interdependent value

The case of interdependent and common values has attracted some attention. This setup considerably differs from the IPV case, with much less recommendations from economic theory. Structural econometric approaches are therefore especially appealing by making counterfactual analysis possible. It can therefore suggest answers that are not available from the theory to the decision maker regarding for instance auction design. See Athey and Haile (2007), Hendricks and Porter (2007) and the references therein for a review of the economic and econometric literature.

A first issue addressed in the literature is testing for common value. See among others the seminal contribution of Paarsch (1992) and HHS. Probably due to the negative identification results of Laffont and Vuong (1996) for symmetric bidders, the estimation literature is not so developed. An important exception is Février (2008) with a common value first-price auction model which is identified imposing that the support of the signal distribution depends upon the common value.[1] Other estimation results consider the Wilson model, where the signal is a noisy observation of the common value. Li, Perrigne and Vuong (2000, LPV afterwards) consider a semiparametric restriction which allows to recover the distributions of the common value and of the noise using deconvolution techniques. Hong and Shum (2002) have proposed a parametric quantile approach for the common value Wilson model. Interestingly, their estimation procedure is based upon a nonlinear quantile regression based upon the bid quantile function derived from the model via a simulation approach.

Somaini (2015) has pioneered an alternative approach, which considers observed bidder specific covariate, an extra source of variation which can help identification, as already noticed in Athey and Haile (2002, section 3.3.2) for affiliated values in second-price auction. As in this paper, the primitives of Somaini (2015) are the signal distribution and in the conditional expectation of the value given all the signals, called the *valuation function* hereafter.[2] To achieve identification, Somaini (2015) considers an exclusion restriction for the valuation functions of each bidder, which must only depend upon the bidder specific covariate. The approach of this paper differs by considering a different identification restriction, which may yield identification over a larger set of signals as discussed in Section 4.2.1. The proof technique of this identification result is apparently new. It proceeds by finding an (integro) differential system solved by the parameters of the valuation function and establishes identification by checking uniqueness of the solution of this system. It is also more flexible and allows to estimate the valuation function of a specific bidder without assumptions on the valuation functions of the other bidders.

[1] As in this paper, Février (2008) departs from the usual signal standardization detailed in Athey and Haile (2007). His standardized signal is equal to the bid.

[2] This differs from the primitives of LPV and Hong and Shum (2002), which are the common value and noise distribution. If it is possible to recover such primitives from the expected common value under some suitable restriction is an open issue.

1.2.4 Statistical aspects

Another important econometric property of the AQR procedure is consistency in the upper and lower tails of the distribution. As noted in Hickman and Hubbard (2015), the kernel procedure of GPV does not deliver proper estimation in the tails. This may be problematic since the winning bid is very likely to come from the upper tail of the distribution when the number of bidders is large. The local polynomial nature of the AQR procedure addresses this issue and allows for consistent estimation in the upper and lower quantile tails. As a by-product of this result, all the private values can be consistently estimated, an important feature for applications based on such estimation as Cassola, Hortaçsu and Kastl (2013). Another issue addressed by the AQR procedure is the lack of a clear cut bandwidth choice for GPV, see Henderson, List, Millimet, Parmeter and Price (2012). It is worth mentioning that these two contributions are achieved here in the presence of covariate.

1.3 Organization of the paper

The two next sections deal with the symmetric IPV case. Section 2 introduces our simple quantile identification method and establishes the stability of linear quantile specification, i.e. that a linear specification for the private value quantile function generates a similar one for the bid quantile function. This leads to introduce parsimonious additive interactive specification which can be estimated using sieve. Section 3 introduces the AQR estimation procedure and derives some consistency rates, Mean Square Error expansion and CLT for the estimation of the private value quantile function. In particular considering a quantile regression specification gives a private value quantile estimator which behaves well in a simulation experiment, for a sample as small as 100 observations. Section 4 considers the interdependent value case. It first recasts the HHS framework using bid quantile functions and introducing bidder specific covariates. The new mixed signal value specification is then introduced and its identification is established. Some suggestions for the estimation of the valuation function then follows. Section 6 concludes the paper and proofs are gathered in two appendices.

2 First price auction and quantile specification

A single and indivisible object with some characteristic $x \in \mathbb{R}^d$ is auctioned to $I \geq 2$ buyers. The potential number of bidders I and x are known to the bidders and the econometrician. The object is sold to the highest bidder who pays his bid B_i to the seller. In the sealed bids framework considered here, bids are sealed so that a bidder does not know others' bid when forming his own bid. Each potential bidder is assumed to have a private value $V_i \geq 0$, $i = 1, \ldots, I$ for the auctioned object. A buyer knows his private value but not the private value of the other bidders. Under the independent private value (IPV) paradigm, the private values are independently drawn from a common distribution given (x, I) with cdf $F(\cdot|x, I)$, or equivalently with conditional quantile function $V(\alpha|x, I)$, $\alpha \in [0, 1]$, which

is the generalized inverse of $F(\cdot|x,I)$,

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

The joint distribution of the private values are known to the bidders. When the private value conditional distribution is absolutely continuous with a pdf $f(\cdot|x,I)$ positive on its support as assumed from now on, $V(\alpha|x,I)$ is the standard inverse function $F^{-1}(\alpha|x,I)$. From now on, $V^{(1)}(\alpha|x,I)$ stands for the α-derivative $\partial V(\alpha|x,I)/\partial\alpha$. Since

$$V^{(1)}(\alpha|x,I)=\frac{1}{f(V(\alpha|x,I)|x,I)}, \tag{2.1}$$

assuming $V^{(1)}(\alpha|x,I)$ exists and is finite for all $\alpha\in[0,1]$ amounts to assume that $f(v|x,I)$ is bounded away from 0 and infinity on its support $[V(0|x,I),V(1|x,I)]$ as assumed for instance in Riley and Samuelson (1981), Maskin and Riley (1984) or GPV.

It is well-known that $A_i=F(V_i|x,I)$, which can be viewed as the rank of the ith bidder in the private value population, has uniform distribution over $[0,1]$ and is independent of x and I. It also follows from the IPV paradigm that the private value ranks $A_i=1,\ldots,I$ are independent. 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,

$$V_i=V(A_i|x,I), \quad A_i\overset{\text{iid}}{\sim}\mathcal{U}_{[0,1]}\perp(x,I), \tag{2.2}$$

where the functional parameter of interest for the econometrician is the private value conditional quantile function. This quantile representation of the values suggests, following Milgrom and Weber (1982), to interpret the private value rank A_i as a signal, see also Krishna (2002).

The case where the bids are given by an increasing function of the private values has attracted considerable attention. Maskin and Riley (1984) have shown that Bayesian Nash Equilibrium bids of symmetric risk averse or risk neutral bidders must increase with the private values under the IPV paradigm, more precisely $B_i=\sigma(V_i;x,I)$ for an increasing bid function $\sigma(\cdot;x,I)$. The next Lemma recalls some important properties of this case which are useful for econometric identification in a quantile approach. In what follows, $G(\cdot|x,I)$ and $g(\cdot|x,I)$ are the conditional cdf and pdf of the bids and $B(\cdot|x,I)$ is the conditional bid function..

Lemma 1 *Suppose that the independent private value paradigm holds and that, for all x, I, $\alpha\in[0,1]\mapsto V(\alpha|x,I)$ is continuously differentiable with a strictly positive $V^{(1)}(\cdot|x,I)$ and that,*

$$B_i=\sigma(V_i;x,I),\ \textit{for all } i=1,\ldots,I, \tag{2.3}$$

for a strictly increasing continuous strategy $v\in[V(0|x,I),V(1|x,I)]\mapsto\sigma(v;x,I)$. Then

*i. **[Signal identification]** The conditional private value ranks $A_i=F(V_i|x,I)$ are iden-*

tical to the conditional bid ranks $G(B_i|x,I)$,

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

ii. ***[Identification of the signal bid function** $\sigma[V(\alpha|x,I)|x,I]$] The bids are given by*

$$B_i = B(A_i|x,I), \text{ for } i = 1,\ldots,I. \tag{2.4}$$

as a function of the signal A_i, and as $B_i = B[F(V_i|x,I)|x,I]$ as a function of the private values.

iii. ***[Probability of winning]** Suppose bidder i bid is $s_i(a|x,z,I)$ while his signal A_i is equal to α. Then the probability that bidder i wins the auction given $A_i = \alpha$ and (x,z,I) is a^{I-1}.*

The rank invariance stated in Lemma 1-(i) is a well-known consequence of the monotone strategy assumption. It also shows that the signals are identified from the bids in a constructive way since it is sufficient to estimate $G(\cdot|\cdot,\cdot)$ to estimate the signals from the bids, the covariate and the number of bidder. When the private value quantile function is known, this can be used to recover the private values via an estimation of the signals and (2.2). Lemma 1-(ii) directly follows from the rank invariance identity. The expression (2.4) shows that the bid quantile function identifies the signal bid function which computes the bids from the signals, that is $\sigma[V(\alpha|x,I)|x,I]$, even when the bid function $\sigma(\cdot|\cdot)$ and the private value distribution are not identified.[3]. This interpretation of the bid quantile function as a signal bid function is a first indication of the relevance of a quantile approach.

The simple expression of the probability of winning obtained in Lemma 1-(iii) is a key ingredient to identify the private value quantile function assuming that the bids are given by the Bayesian Nash equilibrium and that the bidders are risk-neutral, as detailed now. As shown in Maskin and Riley (1984), the Bayesian Nash equilibrium bid function $\sigma(\cdot|\cdot)$ is strictly increasing and differentiable. This carries over the bid quantile function since $B(\alpha|x,I) = \sigma[V(\alpha|x,I)|x,I]$. Suppose now that a bidder bids $B(a|x,I)$ while his signal is α. Then Lemma 1-(ii,iii) gives that his expected payoff is

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

which derivative with respect to a is

$$\begin{aligned}&(V(\alpha|x,I) - B(a|x,I))(I-1)\,a^{I-2} - B^{(1)}(a|x,I)\,a^{I-1}\\&= (I-1)\,a^{I-2}\left(V(\alpha|x,I) - B(a|x,I) - \frac{aB^{(1)}(a|x,I)}{I-1}\right)\end{aligned}$$

[3]As for instance in the nonparametric risk aversion setup considered in Campo, Guerre, Perrigne and Vuong (2011).

where $B^{(1)}(\alpha|x,I) = \partial B(\alpha|x,I)/\partial\alpha$. But since the optimal bid is $B(\alpha|x,I)$, it holds

$$\alpha = \arg\max_{a\in[0,1]} \left(V(\alpha|x,I) - B(a|x,I)\right) a^{I-1}$$

which gives a first order condition which implies for all α in $[0,1]$

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

Hence the private value quantile function is easily identified from the bid quantile function and its derivative. GPV identification is based on a similar equation derived from a first-order condition for the optimal bid function $\sigma(\cdot|x,I)$ which allows to compute the private value as a function of the bid and its distribution. When deriving the identification equation of GPV, Milgrom (2001, Theorem 4.7) derives a version of (2.5) with $1/g\left[B(\alpha|x,I)|x,I\right]$ instead of $B^{(1)}(\alpha|x,I)$, see also Marmer and Shneyerov (2012) and, HHS in a common value context. Marmer and Shneyerov (2012) use this version of (2.5) to obtain an expression of the private value p.d.f. depending upon $B(\cdot|x,I)$, $G(\cdot|x,I)$, $g(\cdot|x,I)$ and its derivative. See also Marmer, Sheneyrov and Xu (2013a) and Liu and Yao (2014) for a similar approach and other applications. Guerre and Sabbah (2012) focus on the estimation of the bid quantile function and on its derivative $B^{(1)}(\alpha|x,I)$ suggesting (2.5) can be used to estimate the private value quantile function. The next section elaborates on (2.5) to propose parsimonious nonparametric specifications which are not subject to the curse of dimensionality.

2.1 Stability of quantile linear specifications

The identification equation (2.5) is a differentiable equation which can be solved using the fact that the private value and bid distributions has the same lower bound by Maskin and Riley (1984), which gives the initial condition $B(0|x,I) = V(0|x,I)$. The expression of the corresponding bid function is given in the next Proposition.

Proposition 2 *Consider a given (x,I), $I \geq 2$, for which $\alpha \in [0,1] \mapsto V(\alpha|x,I)$ is continuously differentiable with a strictly positive $V^{(1)}(\cdot|x,I)$. Then,*

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

$$B(\alpha|x,I) = \frac{I-1}{\alpha^{I-1}} \int_0^\alpha a^{I-2} V(a|x,I)\, da, \tag{2.6}$$

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 iid optimal bids

B_i generated from I iid private values drawn from $V(\cdot|x,I)$. Then it must hold that,

$$V(\alpha|x,I) = B(\alpha|x,I) + \frac{\alpha B^{(1)}(\alpha|x,I)}{I-1}. \tag{2.7}$$

Equation (2.6) in Proposition 2-(i) expresses the bid quantile function as an integral of the private value quantile function. This expression can be used to find a parametric family for the bid quantile function or its p.d.f. generated by a parametric model for the private value distribution, in which case quantile methods from Koenker and Bassett (1978) can be used to estimate such specification. A more general key econometric insight behind Proposition 2-(i) is that the bid quantile function in (2.6) is a linear transformation of the private value quantile function. As a consequence, (2.6) maps a linear private value quantile specification into a linear bid quantile specification, which is a subset of the considered linear private value quantile specification in the cases detailed below. A well known example of such situation is the bid homogenization technique from HHS: if the private value distribution is given by a linear regression specification with independent error term, then the bid distribution is given by the same linear model with a different error terms, see also Rezende (2008). Proposition 2-(i) extends this stability result to the case of quantile regression specification or nonparametric additive interactive quantile considered below.

Proposition 2-(ii) recasts the quantile identification equation (2.5), which can now be viewed as the reciprocal of the linear operator in (2.6). Note that the map $B(\alpha|x,I) + \alpha B^{(1)}(\alpha|x,I)/(I-1)$ in (2.7) may not be increasing in α when the bids are not from the Bayesian Nash equilibrium, see GPV for some examples. This property can be used to investigate the validity of Bayesian Nash equilibrium optimal bidding as assumed here. The equation (2.6) will be used in Section 3 to estimate the examples of linear quantile specification introduced now.

2.1.1 Quantile regression and bids homogenization

Consider a d dimensional covariate x. The private value quantile regression model is

$$V(\alpha|x,I) = \gamma_0(\alpha|I) + x'\gamma_1(\alpha|I) = [1,x']\,\gamma(\alpha|I), \quad \text{for all } \alpha \in [0,1]. \tag{2.8}$$

For a fixed α, this specification is parametric and can be estimated with a parametric rate $n^{1/2}$ in the ideal case of available private values, n being the sample size. However the slope $\gamma(\alpha|I)$ is a function of the quantile level α, showing that this model has some nonparametric features which can make it flexible enough to fit many datasets, as found in particular in Gimenes (2016) for timber ascending auctions. Proposition 2-(i) implies that the conditional bid quantile function satisfies,

$$B(\alpha|x,I) = [1,x']\,\beta(\alpha|I) \text{ with } \beta(\alpha|I) = \frac{I-1}{\alpha^{I-1}}\int_0^\alpha t^{I-2}\gamma(t|I)\,dt, \tag{2.9}$$

showing $B(\alpha|x,I)$ belongs to the quantile regression specification. It follows from (2.7) or from the expression of $\beta(\alpha|I)$ that,

$$\gamma(\alpha|I) = \beta(\alpha|I) + \frac{\alpha\beta^{(1)}(\alpha|I)}{I-1}, \tag{2.10}$$

so that $\gamma(\alpha|I)$ can easily be estimated from an estimation of $\beta(\alpha|I)$ and $\beta^{(1)}(\alpha|I)$.

The private value quantile regression model includes as a special case the regression model of HHS and Rezende (2008). To see this, observe that for $A_i = F(V_i|x,I)$ as in (2.2) by Lemma 1-(i),

$$\begin{aligned} V_i &= [1,x']\,\gamma(I) + v_i, \quad \gamma(I) = \mathbb{E}\left[\gamma(A_i|I)\,|x,I\right] = \int_0^1 \gamma(t|I)\,dt, \\ B_i &= [1,x']\,\beta(I) + b_i, \quad \beta(I) = \mathbb{E}\left[\beta(A_i|I)\,|x,I\right] = \int_0^1 \beta(t|I)\,dt, \end{aligned}$$

where

$$v_i = [1,x']\left(\gamma(A_i|I) - \gamma(I)\right) \text{ and } b_i = [1,x']\left(\beta(A_i|I) - \beta(I)\right),$$

which shows that the private value quantile regression model generates a regression model with heteroscedastic error terms v_i. The regression model of HHS and Rezende (2008) corresponds to an homoscedastic term v_i, that is a constant slope $\gamma_1(\cdot|I) = \gamma_1(I)$. In this case it holds that $\gamma_1(I) = \beta_1(I)$ so that $\gamma_1(I)$ can be estimated by regressing the bids on the covariate as needed in the bid homogenization technique of HHS or in the identification strategy of Rezende (2008). But this identity is not robust and may not hold under heteroscedasticity. Indeed, by (2.10), definition of $\gamma_1(I)$ and $\beta_1(I)$ and integrating by parts,

$$\begin{aligned} \gamma_1(I) &= \int_0^1 \left(\beta_1(t|I) + \frac{t\beta_1^{(1)}(t|I)}{I-1}\right) dt = \beta_1(I) + \frac{1}{I-1}\int_0^1 t\beta_1^{(1)}(t|I)\,dt \\ &= \beta_1(I) + \frac{1}{I-1}\left(\beta_1(1|I) - \int_0^1 \beta_1(t|I)\,dt\right) \end{aligned}$$

which differs from $\beta_1(I)$ for instance when some of the entries of $\beta_1(\cdot|I)$ are strictly increasing. An empirical application where this arises can be found in Gimenes (2016), suggesting that the bid homogenization technique should be applied with care.

2.1.2 Additive interactive quantile specification and sieve

The private value quantile regression model (2.8) assumes linearity of the private value quantile function with respect to the covariate x. This may be too strong in some cases but may be relaxed using a quantile nonparametric additive specification, which was considered

in Horowitz and Lee (2005). Recall that $x = (x_1, \ldots, x_d)$ and consider the additive quantile function

$$V(\alpha|x, I) = \sum_{j=1}^{d} V_j(\alpha; x_j, I) \tag{2.11}$$

where each functions $V_j(\alpha; x_j, I)$ is specific to the entry x_j. The functions $V_j(\alpha; x_j, I)$ are not necessarily linear and will be estimated nonparametrically. Since such quantile specifications are obtained by summing some univariate functions, the effective dimension involved in the nonparametric dimension of this model is 1 because it can be estimated with the same rate than a nonparametric model with a unique covariate as shown in Horowitz and Lee (2005). This parsimonious model can be generalized as in the additive interactive regression model of Andrews and Whang (1990) to allow for more covariate interactions. This gives the additive interactive quantile specification with $d_{\mathcal{M}}$ interactions

$$V(\alpha|x, I) = \sum_{D=1}^{d_{\mathcal{M}}} \sum_{1 \leq j_1 < \cdots < j_D \leq d} V_{j_1 \ldots j_D}(\alpha; x_{j_1}, \ldots x_{j_D}, I) \tag{2.12}$$

where each functions $V_{j_1 \ldots j_D}(\alpha; x_{j_1}, \ldots x_{j_D}, I)$ can now depend upon D entries of x with $D \leq d_{\mathcal{M}} \leq d$. When $d_{\mathcal{M}} = 1$, (2.12) is identical with the additive quantile specification (2.11) and taking $d_{\mathcal{M}}$ greater or equal to 2 gives non additive specification that allows pairwise or higher order interactions. Setting $d_{\mathcal{M}}$ equal to the dimension d of the covariate gives the general quantile specification. As seen from Andrews and Whang (1990) for the regression case, such specification can be estimated with the same rate than a function of $d_{\mathcal{M}}$ variables, so that $d_{\mathcal{M}}$ can be viewed as the effective dimension of this model.

The stability property in Proposition 2-(i) ensures that a private value quantile specification with $d_{\mathcal{M}}$ interaction will generate a bid quantile specification with the same number of interactions: if (2.12) holds, then the bid quantile function satisfies

$$B(\alpha|x, I) = \sum_{D=1}^{d_{\mathcal{M}}} \sum_{1 \leq j_1 < \cdots < j_D \leq d} B_{j_1 \ldots j_D}(\alpha; x_{j_1}, \ldots x_{j_D}, I)$$

and the private values components of the specification can be recovered using

$$V_{j_1 \ldots j_D}(\alpha; x_{j_1}, \ldots x_{j_D}, I) = B_{j_1 \ldots j_D}(\alpha; x_{j_1}, \ldots x_{j_D}, I) + \frac{\alpha}{I-1} B^{(1)}_{j_1 \ldots j_D}(\alpha; x_{j_1}, \ldots x_{j_D}, I)$$

by Proposition 2-(ii).

In the regression case, Andrews and Whang (1990) have proposed to estimate additive interactive specification using a linear sieve approach, see also Horowitz and Lee (2005) for

the case of a quantile additive specification. This amounts to consider the model

$$\mathcal{M} = \left\{ V(\alpha|x,I) = \lim_{K\to\infty} \sum_{k=1}^{K} \gamma_k(\alpha|I) P_k(x) \right\} \tag{2.13}$$

where the sieve $\{P_k(x), 1 \le k \le K\}$ is a family of functions $P_k(\cdot) = P_{kK}(\cdot)$ and $\gamma_k(\alpha|I) = \gamma_{kK}(\alpha|I)$ are the parameters to be estimated and depend upon the quantile level α. The choice of the sieve $\{P_k(x), 1 \le k \le K\}$ should depend upon the order of interactions $d_{\mathcal{M}}$ in the model. In what follows, a number $d_{\mathcal{M}} = 0$ of interactions will be used for the quantile regression model (2.8), in which case $\{P_k(x), 1 \le k \le K\}$ can be set to $[1, x']'$ for all K, so that this simple specification becomes a particular case of a more general linear sieve approach. The model $\mathcal{M}$ can therefore be viewed as a sieve extension of the quantile regression, a *sieve quantile regression*. It follows from (2.6) in Proposition 2-(i) that, provided the limit in (2.13) holds uniformly with respect to α,

$$B(\alpha|x,I) = \lim_{K\to\infty} \sum_{k=1}^{K} \beta_k(\alpha|I) P_k(x), \; \beta_k(\alpha|I) = \frac{I-1}{\alpha^{I-1}} \int_0^{\alpha} t^{I-2} \gamma_k(t|I)\, dt, \tag{2.14}$$

$$V(\alpha|x,I) = \lim_{K\to\infty} \sum_{k=1}^{K} \left(\beta_k(\alpha|I) + \frac{\alpha \beta_k^{(1)}(\alpha|I)}{I-1} \right) P_k(x), \tag{2.15}$$

where (2.15) follows from Proposition 2-(ii). Hence estimating the private value sieve quantile regression can proceed from estimating the coefficients of the bid sieve quantile regression in (2.14) and their first derivatives.

For a general number of interactions, it will be assumed later on that the sieve $\{P_k(x), 1 \le k \le K\}$ is a localized one, that is the support of each $P_k(x)$ shrinks with K and only a finite number of $P_k(x)$'s have an overlapping support. Following Chen (2007), such sieve can be defined using products. Consider a univariate function $p(t)$ with compact support and define, for a bandwidth h and an integer q,

$$p_q(t) = h^{-1/2} p\left(\frac{t-qh}{h}\right)$$

which depends upon on h in an implicit way, and where the normalization with $h^{-1/2}$ ensures that the mean square norm $\int p_q^2(t)\, dt = \int p^2(t)\, dt$ for all q. A simple choice of $p(\cdot)$ is the indicator function of $[0,1]$ which corresponds to regressogram estimation. However, regressogram methods have poor approximation properties and better choices of $p(\cdot)$ are a Cardinal B-spline or father wavelet, see Chen (2007). For the additive quantile specification (2.11) with $d_{\mathcal{M}} = 1$, a choice of sieve $\{P_k(x), 1 \le k \le K\}$ is a reordering of

$$p_q(x_j), \quad q = 1, \ldots, Q_j, \quad j = 1, \ldots, d$$

where the Q_j's are such the supports of the $p_q(x_j)$ cover the compact support of the entry

x_j, implying that $Q_j = O(1/h)$ and then $K = O(1/h)$. For pairwise interactions ($d_{\mathcal{M}} = 2$), a possible $\{P_k(x), 1 \leq k \leq K\}$ is obtained by selecting those product functions

$$p_{q_1}(x_{j_1})\, p_{q_2}(x_{j_2})$$

to obtain a cover of the support of (x_{j_1}, x_{j_2}) for all pair $1 \leq j_1 < j_2 \leq d$, which gives a sieve $\{P_k(x), 1 \leq k \leq K\}$ with $K = O(1/h^2)$. For a general number $d_{\mathcal{M}}$ of interactions, the sieve $\{P_k(x), 1 \leq k \leq K\}$ is obtained by selecting similarly the product functions

$$p_{q_1}(x_{j_1}) \times \cdots \times p_{q_{d_{\mathcal{M}}}}\left(x_{j_{d_{\mathcal{M}}}}\right) = h^{-d_{\mathcal{M}}/2} \prod_{k=1}^{d_{\mathcal{M}}} p\left(\frac{x_{j_k} - q_k h}{h}\right) \tag{2.16}$$

for all possible $d_{\mathcal{M}}$ indexes $1 \leq j_1 < \cdots < j_{d_{\mathcal{M}}} \leq d$.

2.2 Sieve expansion convergence rates

This section is mostly specific to the sieve quantile regression specification (2.13) and briefly considers some technical aspects of sieve approximation, which will be important to obtain estimation rates for the sieve quantile regression estimation proposed later on. It is indeed important for such results that a convergence rate for the sieve approximation (2.13) is available. Since estimating the private value quantile function builds on a local polynomial estimation method to estimate the derivatives $\beta_k^{(1)}(\cdot|I)$, this needs to be combined with smoothness assumptions for the sieve coefficients $\gamma_k(\cdot|I)$. In a second step, the implications for the bid sieve quantile expansion (2.14) will be given. Sieve satisfying the following approximation property will be considered.

Approximation property S. *For each function $V(\alpha; x)$, $(s+1)$th continuously differentiable over $[0,1] \times X$, there exists some coefficients $\gamma(\cdot)$, $(s+1)$th continuously differentiable over $[0,1]$, such that*

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| V(\alpha; x) - \sum_{k=1}^{K} \gamma_k(\alpha) P_k(x) \right| = o\left(K^{-\frac{s+1}{d_{\mathcal{M}}}}\right), \tag{2.17}$$

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| \frac{\partial^p V(\alpha; x)}{\partial \alpha^p} - \sum_{k=1}^{K} \gamma_k^{(p)}(\alpha) P_k(x) \right| = o(1), \quad p = 1, \ldots, s+1. \tag{2.18}$$

The localized product sieve (2.16) approximation properties are very similar to the ones of kernel estimators, with an approximation error of order h^{s+1} for functions $(s+1)$ times differentiable, and an approximation error of order h^{s+1-p} for partial derivatives of order p. The order $K^{-(s+1)/d_{\mathcal{M}}}$ in (2.17) replaces the order h^{s+1} for the product sieve (2.16) with $d_{\mathcal{M}}$

interactions since K is of order $h^{-d_{\mathcal{M}}}$. A weaker version of (2.17) as

$$\sup_{x \in \mathcal{X}} \left| V(\alpha|x,I) - \sum_{k=1}^{K} \gamma_k(\alpha|I) P_k(x) \right| = O\left(K^{-\frac{s+1}{d_{\mathcal{M}}}}\right)$$

holds for well-chosen spline or wavelet sieve detailed below as surveyed in Chen (2007, p.5573), provided that the partial derivatives of $V(\alpha|x,I)$ of order $s+1$ with respect to x are bounded. Assuming in addition that these partial derivatives are continuous with respect to α and x over $[0,1] \times \mathcal{X}$ will give the stronger condition (2.17), as seen from Schumaker (2007, Theorem 12.8) and the proof of Härdle, Kerkyacharian, Picard and Tsybakov (1998, Theorem 8.1), for splines and wavelets chosen as follows.

- An example of spline satisfying (2.17) is given by choosing the function $p(\cdot)$ in (2.16) as the equispaced knots cardinal B-spline with an order larger than $s+2$,

  $$p(t) = \frac{1}{(r-1)!} \sum_{j=0}^{r} (-1)^j \frac{r!}{j!\,(r-j)!} \left[\max(0, x-j)\right]^{r-1}, \quad r \geq s+2,$$

  see Schumaker (2007, (4.47) and Theorem 12.8) and Chen (2007, p.5577). Other examples of splines satisfying (2.17) use general knots as considered in Schumaker (2007).

- A wavelet example of sieve satisfying (2.17) is given by choosing the function $p(\cdot)$ in (2.16) as a father wavelet of order $s+1$, such that $\{p(t-j), j = -\infty, \ldots, \infty\}$ is an orthonormal system and $\int t^r p(t)\,dt = 0$ for $r = 1, \ldots, s+1$, see Härdle et al. (1998), Chen (2007) and the references therein, in particular Daubechies (1992). In this case the bandwidth h is chosen as a negative power of 2.

A result as (2.18) similarly follows from the continuity of $V^{(p)}(\alpha|x,I)$ with respect to x and α and the expression of the sieve coefficients that can be used to approximate $V^{(p)}(\alpha|x,I)$. In the wavelet example, these sieve coefficients are

$$\int_{\mathcal{X}} V^{(p)}(\alpha|x,I) P_k(x)\,dx$$

and are therefore equal to $\gamma_k^{(p)}(\alpha|I)$ by the Dominated Convergence Theorem. A similar result holds for cardinal B-splines as seen from formulas (4.84) and (4.85) in Schumaker (2007). Another example of sieve such that (2.17) and (2.18) hold for smooth enough functions is given by piecewise polynomials over bins of size h, for which the sieve coefficients are given by the partial derivatives of $V(\alpha|x,I)$ with respect to x at the center of each bins.

The next Proposition will be used to study the private value sieve quantile estimators introduced in the next section. It establishes some sieve approximation properties for the bid quantile function. As discussed above, this amounts to study the smoothness properties of $B(\alpha|x,I)$ given the ones of $V(\alpha|x,I)$ as done in Proposition 1 of GPV using the bid and private value cdf instead quantiles. As in GPV, Proposition 3-(i) shows that the bid quantile function is slightly smoother than the private value one, so that, as shown in Proposition 3-(iii), the derivative of the bid quantile function can be approximated with the same rate than the private value quantile function. Proposition 3-(i) is also useful for the quantile regression specification (2.8).

Proposition 3 *Assume the approximation property S holds. Suppose that $V(\alpha|x,I)$ is a $(s+1)$th continuously differentiable function over $[0,1]\times\mathcal{X}$ satisfying,*

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

Then, for $B(\alpha|x,I)$ as in (2.6) and sieve coefficients $\{\gamma_k(\alpha|I), 1\le k\le K\}$ of $V(\alpha|x,I)$ as in Property S

i. $\min_{(\alpha,x)\in[0,1]\times\mathcal{X}} B^{(1)}(\alpha|x,I) > 0$, $\max_{(\alpha,x)\in[0,1]\times\mathcal{X}} B^{(1)}(\alpha|x,I) < \infty$ *and* $B(\alpha|x,I)$ *is* $(s+2)$*th continuously differentiable over* $(0,1]$ *with*

$$\lim_{\alpha\to 0}\sup_{(x,I)\in\mathcal{X}\times\mathcal{I}}\left|\alpha B^{(s+2)}(\alpha|x,I)\right| = 0.$$

ii. The coefficients $\{\beta_k(\alpha|I), 1\le k\le K\}$ from (2.14) are $(s+1)$th continuously differentiable and satisfy

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|B(\alpha|x,I)-\sum_{k=1}^{K}\beta_k(\alpha|I)P_k(x)\right| = o\left(K^{-\frac{s+1}{d_{\mathcal{M}}}}\right),$$
$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|B^{(p)}(\alpha|x,I)-\sum_{k=1}^{K}\beta_k^{(p)}(\alpha)P_k(x)\right| = o(1), \quad p=1,\ldots,s+1.$$

iii. Moreover $\alpha\beta_k^{(1)}(\alpha)=(I-1)[\gamma_k(\alpha|I)-\beta_k(\alpha)]$ and is therefore $(s+1)$th continuously differentiable for all $1\le k\le K$. In addition

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\alpha B^{(1)}(\alpha|x,I)-\sum_{k=1}^{K}\alpha\beta_k^{(1)}(\alpha|x,I)P_k(x)\right| = o\left(K^{-\frac{s+1}{d_{\mathcal{M}}}}\right),$$
$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{\partial^p\left[\alpha B^{(1)}(\alpha|x,I)\right]}{\partial\alpha^p}-\sum_{k=1}^{K}\frac{\partial^p\left[\alpha\beta_k^{(1)}(\alpha|x,I)\right]}{\partial\alpha^p}P_k(x)\right| = o(1), \quad p=1,\ldots,s+1.$$

3 Augmented sieve quantile regression estimation

Proposition 2 gives some guidance for estimating the conditional private value quantile function from an estimation of the coefficients of a linear expansion of $B(\alpha|x,I)$ and its derivative $B^{(1)}(\alpha|x,I)$. While there is an important literature on the estimation of a conditional quantile function, estimating the first derivative of a quantile function has received much less attention. The *augmented* methodology proposed here combines local polynomial and sieve techniques, with a local polynomial part for estimating derivatives and a sieve part for the covariate as in Horowitz and Lee (2005) or Andrews and Whang (1990). The main theoretical results in Section 3.2 focus on estimation of the conditional bid and private values quantile functions. Section 3.2.3 deals with estimation of the private values and of the optimal bidding strategy as a function of the private value. The choice of smoothing parameters is discussed in Section 5.

3.1 Definition of the estimators

Consider L iid. first-price auctions $(I_\ell, x_\ell, B_{i\ell}, i = 1, \ldots, I_\ell)$. To introduce our estimation strategy, assume first that the private values and the bids do not depend upon x_ℓ given I_ℓ, $V(\alpha|x,I) = V(\alpha|I)$ and $B(\alpha|x,I) = B(\alpha|I)$. Let $\rho_\alpha(u)$ be the check function,

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

$\mathbb{I}(\cdot)$ being the indicator function, $\mathbb{I}(q \le 0) = 1$ for $q \le 0$ and 0 otherwise. It is well known that,

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

where $\alpha = 0$ and $\alpha = 1$ are excluded due to multiple minimizers. Estimating the derivative $B^{(1)}(\alpha|I)$ can be done by introducing local variation of the quantile level in the vicinity of α. Let $K(\cdot) \ge 0$ be a kernel function with support $[-1,1]$ and $h = h_L$ be a positive bandwidth parameter going to 0 with the sample size. Then it follows that

$$\begin{aligned}&\{B(a|I), a \in [\alpha - h, \alpha + h] \cap [0,1]\} \\ &= \arg\min_{q(a)} \int_0^1 \mathbb{E}\left[\mathbb{I}(I_\ell = I)\rho_a(B_{i\ell} - q(a))\right]\frac{1}{h}K\left(\frac{a-\alpha}{h}\right)da, \qquad (3.1)\end{aligned}$$

where the minimization is performed over the set of functions $q(a)$ which are continuous on $[\alpha - h, \alpha + h] \cap [0,1]$. Instead of a minimization over the set of all continuous functions, it is sufficient to consider minimization over a set of polynomial functions. Indeed, a good polynomial approximation of $B(a|I)$ over $[\alpha - h, \alpha + h]$ is given by the Taylor expansion

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

where the order $s+1$ and the remainder term $O\left(h^{s+2}\right)$ for α in $(0,1]$ follow from Proposition 3. Let $b=\left(\beta_0,\ldots,\beta_{s+1}\right)'$ be the generic coefficients of such a polynomial function and

$$\pi\left(a\right)=\left[1,a,\frac{a^2}{2}\ldots,\frac{a^{s+1}}{(s+1)!}\right]'.$$

The sample version of the objective function (3.1) restricted to polynomial functions is

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

The *augmented quantile* estimator is

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

$\widehat{\beta}_0\left(\alpha|I\right)$ and $\widehat{\beta}_1\left(\alpha|I\right)$ being estimators of $B\left(\alpha|I\right)$ and its first derivative $B^{(1)}\left(\alpha|I\right)$, respectively. The estimator of the private value quantile is[4]

$$\widehat{V}\left(\alpha|I\right)=\widehat{\beta}_0\left(\alpha|I\right)+\frac{\alpha\widehat{\beta}_1\left(\alpha|I\right)}{I-1}.$$

As detailed below, $\widehat{V}\left(\alpha|I\right)$ is defined for all quantile levels in $[0,1]$.[5]

[4]When the private value distribution does not depend upon I, the bid quantile functions $B\left(\cdot|I\right)$ are such that the derivatives

$$\frac{\partial^j}{\partial\alpha^j}\left[B\left(\alpha|I\right)+\frac{\alpha B^{(1)}\left(\alpha|I\right)}{I-1}\right]=\left(1+\frac{j}{I-1}\right)B^{(j)}\left(\alpha|I\right)+\frac{\alpha B^{(j+1)}\left(\alpha|I\right)}{I-1}$$

do not depend upon I as they are equal to $V^{(j)}\left(\alpha|I\right)=V^{(j)}\left(\alpha\right)$, $j=0,\ldots,s+1$. These constraints can be used to estimate $V\left(\alpha\right)$ using the parameters $\gamma=\left(\gamma_0,\ldots\gamma_s\right)$, $\delta=\left(\delta_2,\ldots,\delta_{\overline{I}}\right)$ where γ_j is for $V^{(j)}\left(\alpha\right)$ and δ_I for the derivatives $B^{(s+1)}\left(\alpha|I\right)$, $I=2,\ldots,\overline{I}$ and $b_I\left(\gamma,\delta\right)=\left[b_{0,I},\ldots,b_{s,I},\delta_I\right]'$ with $b_{s,I}=\left(1+\frac{s}{I-1}\right)^{-1}\left(\gamma_s-\frac{\alpha}{I-1}\delta_I\right)$ and the $b_{j,I}$'s are computed recursively using

$$b_{j,I}=\left(1+\frac{j}{I-1}\right)^{-1}\left(\gamma_j-\frac{\alpha}{I-1}b_{j+1,I}\right),\ j=0,\ldots,s.$$

The estimator of $V\left(\alpha\right)$ is $\widehat{\gamma}_0$ where $\left(\widehat{\gamma},\widehat{\delta}\right)=\arg\min_{\gamma,\delta}\sum_{I=2}^{\overline{I}}\widehat{\mathcal{R}}\left(b_I\left(\gamma,\delta\right);\alpha,I\right)$.

[5]Although not considered here, the augmented quantile estimation procedure can be used to estimate the p.d.f. $f\left(v|I\right)$ of the private value using $f\left(v|I\right)=1/V^{(1)}\left[F\left(v|I\right)|I\right]$. An estimator for $F\left(\cdot|I\right)$ is $\widehat{V}^{-1}\left(\cdot|I\right)$. Set $\widehat{V}^{(1)}\left(\alpha|I\right)=\widehat{\beta}_1\left(\alpha|I\right)+\alpha\widehat{\beta}_2\left(\alpha|I\right)/\left(I-1\right)$ and $\widehat{f}\left(v|I\right)=1/\widehat{V}^{(1)}\left[\widehat{F}\left(v|I\right)|I\right]$. This p.d.f. estimator can

A first extension of this procedure is the *augmented quantile regression* estimator, AQR hereafter, which considers the private quantile regression specification

$$V(\alpha|x, I) = [1, x'] \gamma(\alpha|I).$$

In this case, the bid quantile function satisfies $B(\alpha|x, I) = [1, x'] \beta(\alpha|I)$ by (2.9) with $\gamma(\alpha|I) = \beta(\alpha|I) + \alpha\beta^{(1)}(\alpha|I)/(I-1)$ by (2.10). Define now the parameter

$$b = \left[\beta_0', \beta_1', \ldots, \beta_{s+1}'\right]$$

where all the β_j have the same dimension $d+1$ and

$$P(x, t) = \pi(t) \otimes [1, x']'$$

which is such that the Taylor expansion of $B(\alpha|x, I)$ writes

$$B(\alpha + ht|x, I) = P(x, ht)' b(\alpha|I) + O\left(h^{s+2}\right)$$

where $b(\alpha|I)$ stacks $\beta(\alpha|I)$ and its successive derivatives $\beta^{(1)}(\alpha|I), \ldots, \beta^{(s+1)}(\alpha|I)$. The objective function of the AQR estimation procedure becomes

$$\begin{aligned}
\widehat{\mathcal{R}}(b; \alpha, I) &= \frac{1}{LI} \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_0^1 \rho_a\left(B_{i\ell} - P(x_\ell, a-\alpha)' b\right) \frac{1}{h} K\left(\frac{a-\alpha}{h}\right) da \\
&= \frac{1}{LI} \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell} - P(x_\ell, ht)' b\right) K(t)\, da \qquad (3.2)
\end{aligned}$$

which accounts for the covariate x_ℓ. The estimation of $b(\alpha|I)$ is $\widehat{b}(\alpha|I) = \arg\min_b \widehat{\mathcal{R}}(b; \alpha, I)$ and the AQR private value quantile regression estimator is

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

The bid quantile function can be estimated using $\widehat{B}(\alpha|x, I) = [1, x']' \widehat{\beta}_0(\alpha|I)$.

The second extension is the *augmented sieve quantile regression* (ASQR) procedure which considers an interactive specification with $d_{\mathcal{M}}$ interactions

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

where $\gamma_k(\alpha|I)$ and $P_k(x)$ can depend upon the truncation index K and $P_k(x)$ only depend

account for covariates by using the AQR and ASQR procedures introduced below.

upon at most $d_{\mathcal{M}}$ entries of x. Let K_L be a truncation parameter which diverges and will be taken of order $h^{-d_{\mathcal{M}}}$ later on and focus on the parameter

$$\gamma(\alpha|I) = [\gamma_1(\alpha|I), \ldots, \gamma_{K_L}(\alpha|I)]'$$

of dimension K_L. Stack the $P_k(x)$, $1 \leq k \leq K_L$, into a vector $P(x)$ and observe that estimating $P(x)' \gamma(\alpha|I)$ and its derivative can be done implementing an AQR procedure using the covariate $P(x)$ instead of $[1, x']'$. The corresponding truncated specification for the bid quantile specification is $P(x)' \beta(\alpha|I)$ where $\gamma(\alpha|I) = \beta(\alpha|I) + \alpha\beta^{(1)}(\alpha|I)/(I-1)$. Redefine $P(x,t)$ as

$$P(x,t) = \pi(t) \otimes P(x)'$$

and stacks the successive derivative $\beta^{(j)}(\alpha|I)$ into a vector $b(\alpha|I)$ of dimension $K_L(s+2)$. Define $\widehat{\mathcal{R}}(b;\alpha,I)$ as in the AQR case. The ASQR private value quantile estimator is, for $\widehat{b}(\alpha|I) = \arg\min_b \widehat{\mathcal{R}}(b;\alpha,I)$,

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

The ASQR bid quantile estimator is $P(x)' \widehat{\beta}_0(\alpha|I)$.

As mentioned earlier, the augmented quantile estimators are defined for all quantile levels α including the boundaries $\alpha = 0$ and $\alpha = 1$. Indeed the objective functions $\widehat{\mathcal{R}}(\cdot;\alpha,I)$ of the AQR and ASQR procedures are based on a smoothing of the check function $\rho_a(\cdot)$ for quantile levels a in a vicinity of α. As a consequence the shape of $b \mapsto \widehat{\mathcal{R}}(\cdot;0,I)$ or $\widehat{\mathcal{R}}(\cdot;1,I)$ is driven by the shape of the items $\rho_{\alpha+ht}\left(B_{i\ell} - P(x_\ell, ht)' b\right)$, $\alpha = 0$ and 1 for all t in $[-1,1]$ such that $\alpha + ht$ lies in $[0,1]$. But when $\alpha + ht$ differs from 0 or 1, the sum of the items $\rho_{\alpha+ht}\left(B_{i\ell} - P(x_\ell, ht)' b\right)$ tends to be bowl shaped around a unique minimizer. Hence summing over the observations gives objective functions $\widehat{\mathcal{R}}(\cdot;\alpha,I)$ with convex paths and a unique minimizer with a probability tending to 1 when the sample size grows, as illustrated by the next figure for $\alpha = 1$. Therefore the AQR and ASQR estimators are asymptotically well defined for the extreme quantile levels.

This contrasts with the objective function of the standard quantile regression estimator for $\alpha = 0$ or 1 which contains some flat parts. As seen in Bassett and Koenker (1982), there is no unique quantile regression estimator for the extreme quantile $\alpha = 0$ and $\alpha = 1$ in the standard approach. This extends to other quantile levels since the shape of the objective function $\widehat{\mathcal{R}}(\cdot;\alpha,I)$ ensures that the AQR and ASQR estimators $\widehat{\beta}(\alpha|I)$ is unique for all α in $[0,1]$, with a probability tending to 1 with the sample size.[6] While the linear programming algorithms for the standard quantile regression estimator in Koenker (2005) do not seem to apply here, the AQR and ASQR estimators can be computed using simple modifications of Majorize-Minimize (MM) algorithm of Hunter and Lange (2000) to account for the presence of an integral in (3.2).

[6]See the discussion following Theorem B.8 in the proof section for a formal argument.

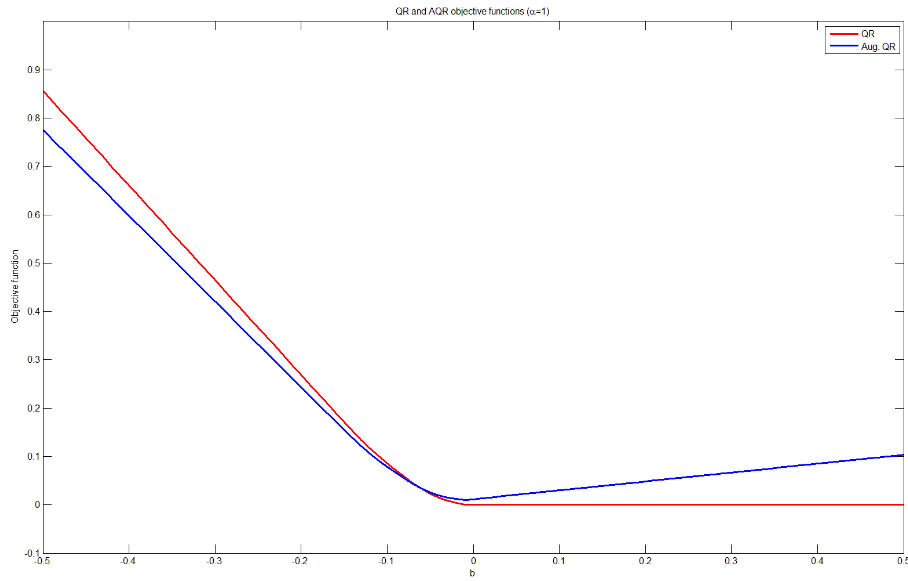


Figure 1: A path of the objective function $\widehat{\mathcal{R}}(\cdot; 1, I)$ (blue) of the augmented quantile regression estimator and of the objective function of the standard quantile regression estimator for $\alpha = 1$ (red).

3.2 Main estimation results

The main assumptions are stated below. Assumptions A and S deal with the first-price auction model and the econometric 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. The notations $a \vee b$ and $a \wedge b$ are used instead of $\max(a, b)$ and $\min(a, b)$.

Assumption A *(i) The auction variables $\{I_\ell, x_\ell, V_{i\ell}, B_{i\ell}, i = 1, \ldots, I_\ell\}$ are iid.. The p.d.f $f(x|I)$ of the covariates x_ℓ given $I_\ell = I$ is continuous and bounded away from 0 over its bounded support $\mathcal{X}$, with a non empty interior and which does not depend upon I. The actual number of bidders I_ℓ belongs to a finite 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 iid. with a conditional quantile function $V(\alpha|x, I)$, which is continuously differentiable over $[0, 1] \times \mathcal{X}$ with

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

(iii) (2.6) holds with $B(0|x, I) = V(0|x, I)$ for all $(x, I) \in \mathcal{X} \times \mathcal{I}$.

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

(i) For the quantile regression model $V(\alpha|x, I) = X'\gamma(\alpha|I)$ as in (2.8), 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 (2.13) and $d_{\mathcal{M}} \in (0, d]$, the sieve satisfies the approximation property S.

Assumption A recalls the implications of Bayesian Nash equilibrium bidding under the independent private value paradigm. In Assumption A-(i), 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 paper, 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 heterogeneity can be found in Marmer, Shneyerov and Xu (2013a). Assumption A-(ii) is standard regarding both auction models and quantile regression inference theory, and Assumption A-(iii) imposes a conditional bid function compatible with the Bayesian Nash equilibrium. Assumption S has been discussed in detail prior Proposition 3. In the ASQR case, it imposes implicitly that the private value quantile function is additive with $d_{\mathcal{M}}$ interactions.

The next set of assumptions deals with sieve, kernel and bandwidth choices.

Assumption R *In the AQR case the matrices $\mathbb{E}[\mathbb{I}(I_\ell = I) X_\ell X_\ell']$, I in $\mathcal{I}$, are full rank and in the ASQR case (i) The eigenvalues of the Gram matrix $\int_{\mathcal{X}} P(x) P'(x)\, dx$ stay bounded away from 0 and infinity when the dimension K_L of $P(\cdot)$ increases and*

$$\max_{x \in \mathcal{X}} \|P(x)\| = O\left(K_L^{1/2}\right).$$

(ii) The sieve $\{P_k, 1 \leq k \leq K_L\}$ is composed with localized functions, in the sense there is a $c > 0$ such that $P_{k_1}(\cdot) P_{k_1}(\cdot) = 0$ as soon as $|k_2 - k_1| > c/2$ with

$$\max_{k \leq K_L} \left\{ \int_{\mathcal{X}} |P_k(x)|\, dx \right\} = O\left(K_L^{-1/2}\right).$$

(iii) For some $\eta \in (0,1]$ and $\overline{K}_{1L}$ with $\log \overline{K}_{1L} = O(\log L)$, it holds that

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

Assumption H *The kernel function $K(\cdot)$ with support $(-1,1)$ is continuously differentiable over the straight line, and strictly positive over $(-1,1)$. The positive bandwidth h goes to 0 with*

$$\lim_{L \to \infty} \frac{\log L}{L h^{2(d_{\mathcal{M}}+1)}} = 0.$$

For the ASQR estimator, $K_L \asymp h^{-d_{\mathcal{M}}}$.

Assumption R first assumes that the matrices $\mathbb{E}\left[\mathbb{I}\left(I_\ell = I\right) P\left(x_\ell\right) P\left(x_\ell\right)'\right]$ for the AQR case and $\int_{\mathcal{X}} P\left(x\right) P'\left(x\right) dx$ for the ASQR case are well conditioned, so that the coefficients of the quantile specifications (2.8) or (2.13) are identified. The rest of Assumption R recalls some properties of the localized sieve (2.16) which are such that

$$\max_{x\in\mathcal{X}} \left\|P\left(x\right)\right\| = O\left(h^{-d_{\mathcal{M}}/2}\right), \quad \max_{k\leq K_L}\left\{\int_{\mathcal{X}} \left|P_k\left(x\right)\right| dx\right\} = O\left(h^{-d_{\mathcal{M}}/2}\right)$$

which, when K_L is proportional to $h^{-1/d_{\mathcal{M}}}$ as assumed later on, gives the orders of Assumption R-(i,ii). The disjoint support condition in Assumption R-(ii) amounts to assume that the function $p\left(\cdot\right)$ in the localized product sieve (2.16) has a compact support.[7] Assumption R-(iii) holds when the bandwidth h of the sieve (2.16) decreases with a polynomial rate and provided $p\left(\cdot\right)$ is Hölder with exponent η. This allows for cardinal B-splines, and wavelets which are not always differentiable, see Daubechies (1992). Assumption H restricts the rate at which the bandwidth can go to 0. In the AQR case, it writes $\lim_{L\to\infty} \log L/\left(Lh^2\right) = 0$ which is slightly more restrictive than the condition $\lim_{L\to\infty} \log L/\left(Lh\right) = 0$ used in nonparametric estimation.

All the theoretical results are stated in a unified framework which does not distinguish the augmented and augmented sieve quantile regression cases, except through the model dimension $d_{\mathcal{M}}$. The case $d_{\mathcal{M}} = 0$ corresponds to the private value quantile regression specification (2.8), whereas other values indicates a nonparametric additive interactive specification with $d_{\mathcal{M}} \geq 1$ interactions.

3.2.1 Bias variance decomposition of the IMSE

The next Theorem presents an asymptotic expansion of the integrated mean squared error (IMSE) of the private value quantile function estimator, which uses some additional notations introduced now. Recall that $K_L = d+1$ in the AQR case. Let s_1 be the $1 \times (s+2)$ selection vector $(0,1,0,\ldots,0)$, which is such that $\mathrm{Id}_{K_L} \otimes s_1 \widehat{\beta}\left(\alpha|I\right) = \widehat{\beta}_1\left(\alpha|I\right)$ is the estimator of sieve coefficient derivative $\beta^{(1)}\left(\alpha\right)$. Let $\Pi^1\left(\alpha\right)$ be the second column of the inverse of $\int \pi\left(t\right)\pi\left(t\right)' K\left(t\right) dt$, i.e.,

$$\Pi^1\left(\alpha\right) = \left(\int \pi\left(t\right)\pi\left(t\right)' K\left(t\right) dt\right)^{-1} s_1'$$

[7]This can be easily checked in the univariate case, as in the multivariate case assuming that the multi index product sieve (2.16) has been properly reordered. The compact support condition for $p\left(\cdot\right)$ can be weakened.

and define, recalling $P(x_\ell) = [1, x_\ell']'$ in the AQR case,

$$v^2(\alpha) = \Pi^1(\alpha)' \int\int \pi(t_1)\pi(t_2)' \min(t_1, t_2) K(t_1) K(t_2)\, dt_1 dt_2 \Pi^1(\alpha),$$

$$\begin{aligned}
\Sigma(\alpha|I) &= \frac{\alpha^2 v^2(\alpha)}{(I-1)^2} \mathbb{E}^{-1}\left[\frac{P(x_\ell)P(x_\ell)'\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] \\
&\quad \times \mathbb{E}\left[P(x_\ell)P(x_\ell)'\mathbb{I}(I_\ell = I)\right]\mathbb{E}^{-1}\left[\frac{P(x_\ell)P(x_\ell)'\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right], \\
\Sigma_{IL} &= \int_{\mathcal{X}}\int_0^1 P(x)'\Sigma(\alpha|I)P(x)\, d\alpha dx.
\end{aligned}$$

That $v^2(\alpha)$, and then Σ_{IL}, is strictly positive follows from the proof of Theorem 4 below, see in particular Lemma B.5 in Appendix B. The bias of the estimator will depend upon

$$\begin{aligned}
\mathsf{Bias}(\alpha|I) &= \frac{1}{I-1}s_1\left(\int \pi(t)\pi(t)'K(t)\,dt\right)^{-1}\int \frac{t^{s+2}\pi(t)}{(s+2)!}K(t)\,dt \\
&\times \mathbb{E}^{-1}\left[\frac{P(x_\ell)P(x_\ell)'\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]\mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)P(x_\ell)\alpha B^{(s+2)}(\alpha|x_\ell, I_\ell)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right], \\
\mathsf{Bias}^2_{IL} &= \int_{\mathcal{X}}\int_0^1 \left(P(x)'\mathsf{Bias}(\alpha|I)\right)^2 d\alpha dx.
\end{aligned}$$

Theorem 4 *Suppose that the private value conditional quantile function $V(\cdot|\cdot)$ is a quantile regression (2.8) or a sieve quantile regression (2.13) with $d_{\mathcal{M}}$ interactions. Then under Assumptions A, H, S and R-(i,ii) with $s \geq d_{\mathcal{M}}/2$, there exists a stochastic approximation $\widetilde{V}(\alpha|x, I)$ of $\widehat{V}(\alpha|x, I)$ such that*

$$\begin{aligned}
\mathbb{E}\left[\int_{\mathcal{X}}\int_0^1 \left(\widetilde{V}(\alpha|x, I) - V(\alpha|x, I)\right)^2 d\alpha dx\right] &= h^{2(s+1)}\mathsf{Bias}^2_{IL} + \frac{\Sigma_{IL}}{LIh^{d_{\mathcal{M}}+1}} \\
&\quad + o\left(h^{2(s+1)} + \frac{1}{Lh^{d_{\mathcal{M}}+1}}\right)
\end{aligned}$$

where $\mathsf{Bias}^{-2}_{IL} = O(1)$, $\Sigma_{IL} = O(1)$ and

$$\int_{\mathcal{X}}\int_0^1 \left(\widehat{V}(\alpha|x, I) - \widetilde{V}(\alpha|x, I)\right)^2 d\alpha dx = O_{\mathbb{P}}\left(\frac{\log^2 L}{L^2 h^{2(d_{\mathcal{M}}+1)+d_{\mathcal{M}}\vee 1}}\right) = o_{\mathbb{P}}\left(\frac{1}{Lh^{d_{\mathcal{M}}+1}}\right). \quad (3.3)$$

The quantile estimator $\widehat{V}(\alpha|x, I)$ is not explicit and nonlinear and attempting a direct computations of its IMSE is a heavy task. Theorem 4 proceeds by replacing $\widehat{V}(\alpha|x, I)$ by an approximation $\widetilde{V}(\alpha|x, I)$ which writes as a sum of variables, so that its IMSE can be easily computed. The IMSE of $\widetilde{V}(\alpha|x, I)$ is a suitable substitute of the IMSE of $\widehat{V}(\alpha|x, I)$ as

suggested by (3.3), a Bahadur representation result established in Theorem B.9 in Appendix B which shows that the difference $\widehat{V}(\alpha|x,I) - \widetilde{V}(\alpha|x,I)$ is negligible with respect to the IMSE of $\widetilde{V}(\alpha|x,I)$. Note that this holds over the full range $[0,1]$ of quantile levels α, suggesting that the augmented estimation procedure performs well near the boundaries. The bias variance decomposition of the IMSE is driven by the estimation of $\alpha B^{(1)}(\alpha|x,I)$ in $V(\alpha|x,I) = B(\alpha|x,I) + \alpha B^{(1)}(\alpha|x,I)/(I-1)$, a function which is $(s+1)$th continuously differentiable by Proposition 3 and which gives the squared bias term $h^{2(s+1)}\mathsf{Bias}_{IL}^2$. The bias component due to the estimation of $B(\alpha|x,I)$ is of the negligible order h^{s+2} except perhaps over a small vicinity of 0 where it is $o(h^{s+1})$. The estimation of $\alpha B^{(1)}(\alpha|x,I)/(I-1)$ also contributes to the IMSE through its asymptotic variance $\Sigma_{IL}/\left(LIh^{d_{\mathcal{M}}+1}\right)$, which is similar to the asymptotic variance obtained for kernel estimation of a conditional p.d.f with $d_{\mathcal{M}}$ covariates. Indeed, the bid quantile derivative is homogeneous to a conditional p.d.f. since

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

where $g(\cdot|\cdot)$ is the bid conditional p.d.f and where the conditional quantile function can be estimated with a faster asymptotic variance of order $1/Lh^{d_{\mathcal{M}}}$ since it is the inverse of the conditional bid cdf $G(\cdot|\cdot)$. Note that the asymptotic variance term $\Sigma_{IL}/\left(LIh^{d_{\mathcal{M}}+1}\right)$ depends upon the number of interactions $d_{\mathcal{M}}$ and not the dimension of the covariate d. Hence Theorem 4 illustrates the dimension reduction features of the AQR and AQSR procedures. In particular, the variance term is of order $1/(Lh)$ in the AQR case independently of the dimension of the covariate d, which therefore can be large.

Maximizing the leading term of the IMSE yields the optimal bandwidth

$$h_* = \left(\frac{(d_{\mathcal{M}}+1)\Sigma_{IL}}{2(s+1)\mathsf{Bias}_{IL}^2}\frac{1}{LI}\right)^{\frac{1}{2s+d_{\mathcal{M}}+3}}. \tag{3.4}$$

As in simple standard bandwidth choice for kernel estimation, a pilot bandwidth can be computed using a simple private value quantile regression model to estimate Σ_{IL} and Bias_{IL}^2 in a parametric way as implemented in Section 5 below. The corresponding IMSE rate is

$$L^{\frac{s+1}{2s+d_{\mathcal{M}}+3}}$$

which depends upon the number of interactions $d_{\mathcal{M}}$ but not on the dimension d of the covariate. In the AQR case with $d_{\mathcal{M}} = 0$, the IMSE rate $L^{\frac{s+1}{2s+3}}$ is as expected the optimal rate for estimating the marginal p.d.f. of a real random variable. For $s = 1$, it is equal to $L^{2/5}$ independently of the dimension d of the covariate, which is close of $L^{1/2}$. The IMSE rate decreases with the complexity of the model as measured by the degree of interactions $d_{\mathcal{M}}$.

Two assumptions limit the use of the optimal bandwidth (3.4). First, Theorem 4 assumes $s \geq d_{\mathcal{M}}/2$ but this condition is only binding for a number of interactions $d_{\mathcal{M}}$ larger than 3 since $s \geq 1$ under Assumption S. The second potentially binding assumption is the

bandwidth rate of Assumption H, but it only requires $2\left(d_{\mathcal{M}}+1\right)/\left(2s+d_{\mathcal{M}}+3\right)<1$ which boils down to the less stringent $s+1>d_{\mathcal{M}}/2$. This contrasts with the smoothness condition $s>0$ used for sieve regression estimator in Belloni, Chernozhukov, Chetverikov and Kato (2015) and Chen and Christensen (2015). The stronger condition $s\geq d_{\mathcal{M}}/2$ is due to the nonlinear nature of the sieve quantile regression estimator. Studying its asymptotic bias and Bahadur linear representation involves the Hessian of the population objective function $\mathbb{E}\left[\widehat{\mathcal{R}}\left(\cdot;\alpha,I\right)\right]$, see Lemma B.2, Theorems B.8 and B.9 where the inverse of the Hessian plays a crucial role, in Appendix B. Due to smoothing, the Hessian exists and has an inverse in a $h^{d_{\mathcal{M}}/2+1}$ vicinity of the derivatives of the "true" sieve quantile coefficients, normalized by h at the power of the derivatives order. But the bias of the ASQR normalized derivatives is of order $o\left(h^{s+1}\right)$ and having this bias term negligible with respect to the radius $h^{d_{\mathcal{M}}/2+1}$ of the vicinity ensuring a regular Hessian gives the restriction $s\geq d_{\mathcal{M}}/2$ used in Theorem 4.

3.2.2 Central limit theorem and uniform consistency

The simple additive structure of the private value ASQR estimators allows for IMSE bias variance decomposition as in Theorem 4. This section similarly establishes a Central Limit Theorem for $\widehat{V}\left(\alpha|x,I\right)$. Theorem 5 is also useful to better understand the pointwise properties of $\widehat{V}\left(\alpha|x,I\right)$, especially near the upper boundary $\alpha=1$. Theorem 6 obtains its uniform convergence rates, a result allowing comparison with GPV. Theorem 6 establishes the uniform convergence of $\widehat{B}\left(\alpha|x,I\right)$ as the estimated bid quantile function will be used to recover bidder's signals and private values.

Let s_1 be the selection vector defined earlier and

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

$$v_h^2\left(\alpha\right)=\Pi_h^1\left(\alpha\right)'\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\pi\left(t_1\right)\pi\left(t_2\right)'\min\left(t_1,t_2\right)K\left(t_1\right)K\left(t_2\right)dt_1dt_2\Pi_h^1\left(\alpha\right),$$

$$\begin{aligned}\Sigma_h\left(\alpha|I\right)&=\frac{\alpha^2v_h^2\left(\alpha\right)}{\left(I-1\right)^2}\mathbb{E}^{-1}\left[\frac{P\left(x_\ell\right)P\left(x_\ell\right)'\mathbb{I}\left(I_\ell=I\right)}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]\\&\times\mathbb{E}\left[P\left(x_\ell\right)P\left(x_\ell\right)'\mathbb{I}\left(I_\ell=I\right)\right]\mathbb{E}^{-1}\left[\frac{P\left(x_\ell\right)P\left(x_\ell\right)'\mathbb{I}\left(I_\ell=I\right)}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right],\end{aligned}\qquad(3.5)$$

$$\begin{aligned}\mathsf{Bias}_h(\alpha|I)&=\frac{1}{I-1}s_1\left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt\right)^{-1}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\frac{t^{s+2}\pi\left(t\right)}{(s+2)!}K\left(t\right)dt\\&\times\mathbb{E}^{-1}\left[\frac{P\left(x_\ell\right)P\left(x_\ell\right)'\mathbb{I}\left(I_\ell=I\right)}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]\mathbb{E}\left[\frac{\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)\alpha B^{(s+2)}\left(\alpha|x_\ell,I\right)}{B^{(1)}\left(\alpha|x_\ell,I\right)}\right].\end{aligned}\qquad(3.6)$$

Theorem 5 *Suppose that the private value conditional quantile function $V(\cdot|\cdot)$ is a quantile regression (2.8) or a sieve quantile regression (2.13) with $d_{\mathcal{M}}$ interactions. Then under Assumptions A, H, S and R-(i,ii) with $s \geq d_{\mathcal{M}}/2$ and*

$$\frac{\log^2 L}{Lh^{2d_{\mathcal{M}}+1+(d_{\mathcal{M}}\vee 1)}} = o(1),$$

it holds for L large enough that $P(x)' \Sigma_h(\alpha|I) P(x) \neq 0$ for all α in $(0,1]$ all x in $\mathcal{X}$, $\max_{x\in\mathcal{X}} P(x)' \Sigma_h(\alpha|I) P(x) = O\left(h^{-d_{\mathcal{M}}}\right)$ and

$$\left(\frac{LIh}{P(x)' \Sigma_h(\alpha|I) P(x)}\right)^{1/2} \left(\widehat{V}(\alpha|x,I) - V(\alpha|x,I) - h^{s+1} P(x)' \mathsf{Bias}_h(\alpha|I) + o\left(h^{s+1}\right)\right)$$

converges in distribution to a standard normal.

While Theorem 4 analyzes the global performance of the private value conditional quantile estimator over the whole range $[0,1]$ of quantile levels, Theorem 5 holds for any quantile level with the exception of $\alpha = 0$, which is such that $\widehat{V}(0|x,I) = \widehat{B}(0|x,I)$ has a smaller variance of order $1/\left(Lh^{d_{\mathcal{M}}}\right)$. For other quantile levels the private value conditional quantile estimator depends upon $\widehat{B}^{(1)}(\alpha|x,I)$ so that the asymptotic variance of $\widehat{V}(\alpha|x,I)$ has a larger order $1/\left(Lh^{d_{\mathcal{M}}+1}\right)$ which also holds in Theorem 4. As also seen from Theorem 6 below, the private value conditional quantile estimator is consistent for all quantile levels as expected from its local polynomial construction. Therefore the potential boundary effects only appear through the bias and variance factors $\mathsf{Bias}_h(\alpha|I)$ and $\Sigma_h(\alpha|I)$. Since the support of the kernel is $[-1,1]$, it holds that

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

where $\mathsf{Bias}(\alpha|I)$ and $\Sigma(\alpha|I)$ are defined before Theorem 4, allowing in principle to implement simple choice of a pointwise optimal bandwidth for quantile levels well inside $[0,1]$. When α lies in $(0,h]$ or $[1-h,1]$, the bias and variance factors depend upon h. It is commonly believed that the variance factor is inflated near the boundaries whereas there is no clear guideline for the bias factor, see Fan and Gijbels (1996) and the references therein.

Theorem 5 holds under a stronger bandwidth condition than Theorem 4. This is due to a Bahadur remainder term combined with the different behavior of $\int_{\mathcal{X}} P(x) P(x)' dx$, with bounded entries by Assumption R-(i), which appears in the IMSE when studying this remainder term and its Theorem 5 counterpart $P(x) P(x)'$, with entries of order $h^{-d_{\mathcal{M}}/2}$ by Assumption R-(i). Since the optimal bandwidths of Theorems 4 and 5 have the same order $L^{-1/(2s+d_{\mathcal{M}}+3)}$, applying Theorem 5 with such bandwidths necessitates a larger smoothness index s. In addition to $s \geq d_{\mathcal{M}}/2$, it must now hold that $s > d_{\mathcal{M}} - 1$. Note that it is not really binding for models with small number $d_{\mathcal{M}}$ of interactions. For larger $d_{\mathcal{M}}$ it is stronger than the condition $s \geq d_{\mathcal{M}}/4$ used in Belloni et al. (2015) and Chen and Christensen (2015) for uniform consistency.

The next Theorem deals with uniform consistency of the ASQR procedure.

Theorem 6 *Suppose that the private value conditional quantile function $V(\cdot|\cdot)$ is a quantile regression (2.8) or a sieve quantile regression (2.13) with $d_{\mathcal{M}}$ interactions. Then under Assumptions A, H, S and R with $s \geq d_{\mathcal{M}}/2$ and*

$$\frac{\log L}{Lh^{2d_{\mathcal{M}}+1+(d_{\mathcal{M}}\vee 1)}} = O(1),$$

it holds

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

The bandwidth condition used in Theorem 6 is similar to the one of Theorem 5 and allows an optimal bandwidth of order $(\log L/L)^{1/(2d_{\mathcal{M}}+s+3)}$ provided the smoothness s satisfies

$$s \geq \max\left(\frac{d_{\mathcal{M}}}{2}, d_{\mathcal{M}} - 1\right).$$

Under this condition the uniform consistency rate of the private value conditional quantile estimator is

$$\left(\frac{\log L}{L}\right)^{\frac{s+1}{2s+d_{\mathcal{M}}+3}}$$

which coincides with the GPV optimal minimax uniform consistency rate for the estimation of the private value conditional cdf in the presence of $d_{\mathcal{M}}$ covariates.[8] Theorem 6 also includes a uniform consistency rate for the bid conditional quantile function estimator which will be used to estimate the bidders' signals and private values.

3.2.3 Private values estimation

The private values estimation proposed here builds on Lemma 1-(i) which shows that the private value and bid ranks of bidder i are identical. The signal $A_{i\ell}$ can be estimated by matching the estimated conditional bid quantile function with the observed bid,

$$\widehat{A}_{i\ell} = \arg\min_{\alpha\in[0,1]} \left|B_{i\ell} - \widehat{B}(\alpha|x_\ell, I_\ell)\right| \tag{3.7}$$

[8] GPV consider the pdf but the rate for cdf or quantile can be derived similarly.

using an appropriate convention to break ties. Then (2.2) suggests the estimated private values,

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

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

Corollary 7 *Under the conditions of Theorem 6*

$$\max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|\widehat{V}_{i\ell} - V_{i\ell}\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2} + h^{s+1}\right).$$

The proposed private value estimation procedure is free of boundary issues so that all the private values can be recovered asymptotically. This contrasts with the kernel private value estimation procedure of GPV.

The optimal private value strategy can be estimated in a similar way. By Lemma 1-(ii)

$$s\left(v|x,I\right) = B\left[F\left(v|x,I\right)|x,I\right]$$

An estimator of the private value cdf. $F\left(v|x,I\right)$ is,

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

and a estimator of the optimal private value bidding strategy is

$$\widehat{s}\left(v|x,I\right) = \widehat{B}\left[\widehat{F}\left(v|x,I\right)|x,I\right].$$

Arguing as in the proof of Corollary 7 shows that the estimated private value strategy converges uniformly with the same rate than the estimated private values. This holds over the private value support which is unknown but can be estimated using $\widehat{V}\left(0|x,I\right)$ and $\widehat{V}\left(1|x,I\right)$.

4 Extension to heterogenous interdependent value

This section considers an extension inspired by Milgrom and Weber (1982) where bidder's valuations can depend upon the vector of signals $A = (A_1, \ldots, A_I)'$, being therefore potentially unknown to bidders at the time of the auction.[9] As noted in Laffont and Vuong (1996), this may lead to a model which cannot be identified from the bids. Identification will be achieved here assuming that a bidder characteristic z_i is observed and under some functional restrictions detailed in Section 4.2.2. For the sake of brevity, the analysis will be restricted to

[9] Milgrom and Weber (1982) consider valuations depending on A and on an additional signal A_0, which is unknown to all bidders. As noted in Somaini (2015), ignoring A_0 amounts here to replace the valuations by their conditional expectation given A.

univariate z_i and the good covariate x will be dropped out. The signal vector A is assumed to be independent of $z = (z_1, \ldots, z_I)$, the individual signal A_i being possibly dependent but with a marginal distribution uniform over $[0, 1]$. It is assumed that bidder's identities are observed. Indeed, due to the presence of the covariate z, this framework is asymmetric and bidding strategies should depend upon bidder's identities,

$$B_i = s_i (A_i; z), \; i = 1, \ldots, I. \tag{4.1}$$

The observations consist on the bids B_i and the covariate z_i. The first parameter of interest is the signal distribution. The second parameter of interest are the bidder's valuation function. The proposed extension focuses on a particular bidder, say bidder 1, whose valuation for the auctioned good is

$$U_1 = U_1 (A; z).$$

Assume in a first step that the strategy functions $s_i (\cdot; z)$ are continuous and increasing over $[0, 1]$ and satisfy the terminal condition $s_1 (1; z) = \cdots = s_I (1; z)$, as established in Lizerri and Persico (2002) for Bayesian Nash equilibrium bids. Note that the initial bids $s_i (0; z)$ may differ, in which case a bid $s_i (A_i; z)$ has no chance to win the auction if the signal A_i is too low, i.e. smaller than the signal threshold

$$\underline{\alpha}_i (z) = \max \left\{ \alpha \in [0, 1]; s_i (\alpha; z) \leq \max_{1 \leq j \neq i \leq I} s_i (0; z) \right\}, \tag{4.2}$$

which are well-defined under the terminal condition. Note that there is always at least one bidder such that $\underline{\alpha}_i (z) = 0$, so that the set $\mathcal{I}$ of bidders with $\underline{\alpha}_i (z) = 0$ is not empty, being identical to the set of bidders i such that $s_i (0; z) = \max_{1 \leq j \leq I} s_j (0; z)$. In what follows, it is said that there is *aggressive bidding* if the set $\{1, \ldots, I\} \setminus \mathcal{I}$ of potentially dominated bidders with $\underline{\alpha}_i (z) > 0$ is not empty, or equivalently if the set $\mathcal{I}$ of potentially aggressive bidders is a strict subset of $\{1, \ldots, I\}$. When there is no aggressive bidding

$$\underline{\alpha}_i (z) = 0 \text{ for all } i = 1, \ldots, I$$

and all the bidders share the same initial bid $s_i (0; z)$. In what follows, $B_i (\cdot|z)$, $G_i (\cdot|z)$ and $g_i (\cdot|z)$ are respectively the conditional bid quantile, cumulative distribution and probability density functions. The bidder specific covariate z_i takes value in $(0, \overline{z}]$ with $0 < \overline{z} \leq \infty$ and the support of z is $\mathcal{Z} = (0, \overline{z}]^I$. Our framework relies on the following high-level assumptions.

Assumption DV *The signal vector A is independent of the bidders characteristic z and the support of its distribution is $[0, 1]^I$. The signal vector A has a pdf which is strictly positive and continuously differentiable over $[0, 1]^I$. The valuation functions $U_i (A; z)$ are positive over $[0, 1]^I \times \mathcal{Z}$ for all i. For each z in $\mathcal{Z}$ and for each $i = 1, \ldots, I$, $\underline{\alpha}_i (z) < 1$, $s_i (\cdot; z)$ is continuous and increasing on $[\underline{\alpha}_i (z), 1]$, with $s_i (\alpha; z) \leq s_i (\underline{\alpha}_i (z); z)$ on $[0, \underline{\alpha}_i (z)]$ and $s_1 (1; z) = \cdots = s_I (1; z)$. It may also hold in addition that*

(i) For each z in $\mathcal{Z}$ and for each $i = 1, \ldots, I$, $s_i (\cdot; z)$ is continuous and strictly increasing on $[\underline{\alpha}_i (z), 1]$.

(ii) For each z in $\mathcal{Z}$ and for each $i = 1, \ldots, I$, $s_i(\cdot; z)$ is twice continuously differentiable on $[\underline{\alpha}_i(z), 1]$ with $\frac{\partial}{\partial \alpha} s_i(\cdot; z) > 0$ over $[\underline{\alpha}_i(z), 1]$.[10]

(iii) For each z in $\mathcal{Z}$, the bidding strategy $s_1(\cdot; z)$ satisfies the best response condition

$$s_1(\alpha; z) \in \arg\max_b \mathbb{E}\left[(U_1(A; z) - b)\, \mathbb{I}\left\{b \geq \max_{2 \leq i \leq I} B_i\right\} | A_1 = \alpha, z\right] \tag{4.3}$$

for all α in $[\underline{\alpha}_i(z), 1]$.

Assumption DV retains some key features of optimal Bayesian Nash Equilibrium bidding, which are sufficient to identify the signal distribution and the valuation function $U_1(A; z)$ of the first bidder. In particular, the best response condition (4.3) in DV-(iii) only assumes that the first bidder behaves strategically whereas optimal Bayesian Nash Equilibrium bidding would request that all bids satisfy a best response condition. Identifying the valuation functions of the other bidders, as necessary for instance to compute the outcomes of other types of auctions under an equilibrium assumption, would request for instance that all bids satisfy similar best response conditions.

Consider now the monotonicity and smoothness of the bidding strategies. For affiliated signal and valuation functions $U_i(A; z)$ strictly increasing with respect to the private signal A_i and increasing with respect to the other A_j, Reny and Zamir (2004) have established existence, but not uniqueness, of a Bayesian Nash equilibrium with increasing bidding strategies. In such setup, Assumption DV supposes that the bidders select such an equilibrium. The terminal condition $s_1(1; z) = \cdots = s_I(1; z)$ holds under the Bayesian Nash equilibrium as established by Lizzeri and Persico (2000, Step 2).[11] In the two bidder case, Lizzeri and Persico (2000, Appendix) have studied uniqueness, strict monotonicity and smoothness of the optimal bidding strategies. As discussed after equation (4.4), the best response condition (4.3) does not provide any information about the bidding strategy for $\alpha < \underline{\alpha}_1(z)$. Assuming strictly increasing strategies for $\alpha < \underline{\alpha}_i(z)$ when $\underline{\alpha}_i(z) > 0$ can therefore be highly unrealistic. This will somehow limit the possibility to identify the signal distribution over its whole support $[0, 1]$.

4.1 Signal distribution, winning probability and bidding strategies

Compared to the independent private value paradigm, the interdependent value model introduces an additional parameter, the joint signal distribution. However, the signals can be identified as in the independent private value case thanks to the monotonicity of the bidding strategy. This ensures identification of the joint signal distribution on a proper subset of its support as established in Lemma 8. The most important difference between the independent private and interdependent value setups lies in the difficulty of identifying the valuation

[10] Derivatives at $\underline{\alpha}_i(z)$ are derivatives coming from the right. Our results also hold if the right first derivative vanishes at $\underline{\alpha}_i(z) = 0$ with a right second derivative $\frac{\partial^2}{\partial \alpha^2} s_i(\underline{\alpha}_i(z); z) > 0$.

[11] Lizzeri and Persico (2000) consider the two bidders case but this terminal condition result can easily be extended to the case of a general number I of bidders, see Lemma A.1 in Appendix A.

function, see the discussion following Lemma 9.

4.1.1 Signal identification through bid quantiles

The next Lemma parallels Lemma 1 of the private value case. The main difference between these two lemmas is due to the possible existence of dominated bidders, with a signal threshold $\underline{\alpha}_i(z)$ strictly above 0. In this case, it may not be possible to recover signals A_i smaller than $\underline{\alpha}_i(z)$. It is therefore important to show that the $\underline{\alpha}_i(\cdot)$'s are identified as follows from Lemma 8-(i).

Lemma 8 *Suppose Assumption DV-(i) holds.*

i. ***[Signal threshold identification]*** *For each z in $\mathcal{Z}$ and $i = 1, \ldots, I$, the signal threshold $\underline{\alpha}_i(z)$ is equal to*

$$G_i\left[\max_{1 \leq j \leq I} B_j(0|z)|z\right]$$

ii. ***[Signal identification]*** *For each $i = 1, \ldots, I$, the signals A_i satisfy*

$$A_i = G_i(B_i|z) \text{ provided } B_i \geq B_i[\underline{\alpha}_i(z)|z]$$

where z in the equation above stands for the random bidder characteristic vector which has generated the bids. The signal joint distribution is therefore identified over

$$\bigcup_{z \in \mathcal{Z}} [\underline{\alpha}_1(z), 1] \times \cdots \times [\underline{\alpha}_I(z), 1].$$

iii. ***[Signal bid function identification]*** *For each z in $\mathcal{Z}$ and $i = 1, \ldots, I$, the signal bid function satisfies*

$$s_i(\alpha; z) = B_i(\alpha|z) \text{ for } \alpha \text{ in } [\underline{\alpha}_i(z), 1].$$

iv. ***[Winning probability identification]*** *Suppose bidder 1 bid is $s_1(a; z)$ while his signal A_1 is equal to α. Then the probability $\omega(a|\alpha, x, z)$ that bidder 1 wins the auction given $A_1 = \alpha$ and z is identified provided α is in $[\underline{\alpha}_1(z), 1]$ and is equal to*

$$\omega(a|\alpha, z) = \mathbb{P}\left[B_1(a|z) > \max_{2 \leq j \leq I} B_j | A_1 = \alpha, z\right]$$

As for Lemma 1, the proof of Lemma 8 works by showing that the bidding strategies are identical to the conditional bid quantile. The main difference between this two results is that it now only holds for high enough signals, see Lemma 8-(iii), due to potential aggressive bidding. The most important consequence is that the signal distribution may be nonparametrically identified over a subset of $[0, 1]^I$ only, as established in Lemma 8-(ii). However

the set over which identification holds is large enough to identify many parametric models for the signal distribution, as for instance the copula model employed by Hubbard, Li and Paarsch (2012) or the Gaussian factor model of Somaini (2015). A simple condition ensuring nonparametric identification over the full support $[0,1]^I$ is that there is no aggressive bidding for some value z_0 of the bidder specific covariate, in which case the set of Lemma 8-(ii) is equal to $[0,1]^I$ as $\underline{\alpha}_i(z_0) = 0$ for all bidders i.

Comparing the expressions for the winning probability in Lemmas 1-(iii) and 8-(iv) illustrates the difference between the symmetric independent private and asymmetric interdependent value setups. While the winning probability of the symmetric independent private value is explicit and simple, the winning probability in Lemma 8-(iv) is more involved due to asymmetric strategies and signal dependence. It is however identified and can be estimated with standard kernel methods using an estimation of the signal A_1 as in (3.7).[12] The winning probability is useful to compute the conditional bid function $B_1(\cdot|\cdot)$ as a functional of the valuation function $U_1(\cdot|\cdot)$ and the signal distribution under the best response condition (4.3), as detailed in the next section to parallel Proposition 2.

4.1.2 Bid quantile and valuation functions

An important difficulty already noted in HHS is that the best response condition only identifies a derivative of a conditional average $\overline{U}_1(a|\alpha,z)$ of $U_1(A;z)$ over the event that bidder 1 wins, given $A_1 = \alpha \geq \underline{\alpha}_1(z)$ and z. To see this, observe that the expected profit writes, setting $b = s_1(a;z)$,

$$\mathbb{E}\left[U_1(A;z)\,\mathbb{I}\left\{B_1(a;z) \geq \max_{2\leq i\leq I} B_i\right\}|A_1 = \alpha, z\right] - s_1(a;z)\,\omega(a|\alpha,z).$$

Define

$$\overline{U}_1(a|\alpha,z) = \mathbb{E}\left[U_1(A;z)\,\mathbb{I}\left\{s_1(a;z) \geq \max_{2\leq i\leq I} B_i\right\}|A_1 = \alpha, z\right]$$

and observe that (4.3) gives

$$\alpha = \arg\min_{a\in[0,1]} \left\{\overline{U}_1(a|\alpha,z) - s_1(a;z)\,\omega(a|\alpha,z)\right\}.$$

This gives the first order condition

$$\left.\frac{\partial \overline{U}_1(a|\alpha,z)}{\partial a}\right|_{a=\alpha} - s_1(\alpha;z)\left.\frac{\partial \omega(a|\alpha,z)}{\partial a}\right|_{a=\alpha} - s_1^{(1)}(\alpha;z)\,\omega(\alpha|\alpha,z) = 0. \tag{4.4}$$

[12] Alternatively, the conditioning event $A_1 = \alpha$ can be equivalently written $B_1 = B_1(\alpha|z)$ which suggests using kernel weights $K\left[\left(B_{1\ell} - \widehat{B}_1(\alpha|z_\ell)\right)/h, (z_\ell - z)/h\right]$ instead of $K\left[\left(\widehat{A}_{1\ell} - \alpha\right)/h, (z_\ell - z)/h\right]$. Note that the dimension of the conditioning variable is $I+1$, which is potentially large. Furthermore, bid homogeneization specifications seem much less effective when applied with bidder specific covariate than for good specific one. It is however possible to assume that one of the z_i is constant.

Suppose first $\alpha < \underline{\alpha}_1(z)$ and $s_1(\alpha; z) < s_1(\underline{\alpha}_1(z); z)$, so that bidder 1 loses the auction with probability 1 as there is an opponent i_* with $B_{i_*} \geq s_{i_*}(0; z) > s_1(\alpha; z)$. Hence by definition of $\overline{U}_1(\cdot|\alpha, z)$ and $\omega(\cdot|\alpha, z)$

$$\left.\frac{\partial \overline{U}_1(a|\alpha, z)}{\partial a}\right|_{a=\alpha} = \left.\frac{\partial \omega(a|\alpha, z)}{\partial a}\right|_{a=\alpha} = \omega(\alpha|\alpha, z) = 0 \text{ for } \alpha < \underline{\alpha}_1(z)$$

and (4.4) holds for any differentiable strategy satisfying $s_1(\alpha; z) < s_1(\underline{\alpha}_1(z); z)$. Hence best response strategies are not uniquely defined when $\alpha < \underline{\alpha}_1(z)$. For $\alpha \geq \underline{\alpha}_1(z)$ define

$$\Omega(\alpha|z) = \frac{\omega(\alpha|\alpha, z)}{\left.\frac{\partial \omega(a|\alpha,z)}{\partial a}\right|_{a=\alpha}}, \quad U_1(\alpha|z) = \frac{\left.\frac{\partial \overline{U}_1(a|\alpha,z)}{\partial a}\right|_{a=\alpha}}{\left.\frac{\partial \omega(a|\alpha,z)}{\partial a}\right|_{a=\alpha}}.$$

As shown in the proof of Lemma 9, these quantities are well defined for $\underline{\alpha}_1(z)$. Since, by Lemma 8-(iii), $s_1(\alpha; z) = B_1(\alpha|z)$ when $\alpha \geq \underline{\alpha}_1(z)$, (4.5) in Lemma 9 shows that $U_1(\alpha|z)$ is identified for α in $[\underline{\alpha}_1(z), 1]$, a preliminary step for the identification of the valuation function $U_1(A; z)$.

Lemma 9 *Under Assumption DV-(ii,iii), it holds for each z of $\mathcal{Z}$ and all all α in $[\underline{\alpha}_1(z), 1]$*

$$U_1(\alpha|z) = B_1(\alpha|z) + B_1^{(1)}(\alpha|z)\,\Omega(\alpha|z), \tag{4.5}$$

$$B_1(\alpha|z) = \exp\left(\int_\alpha^1 \frac{dt}{\Omega(t|z)}\right) \int_{\underline{\alpha}_1(z)}^{\alpha} \frac{U_1(a|z)}{\Omega(a|z)} \exp\left(-\int_a^1 \frac{dt}{\Omega(t|z)}\right) da, \tag{4.6}$$

with $B_1(\underline{\alpha}_1(z)|z) = U_1(\underline{\alpha}_1(z)|z)$.

The identity (4.5) extends (2.7), with $U_1(\alpha|z)$ and $\Omega(\alpha|z)$ being respectively the private value quantile and $\alpha/(I-1)$ for symmetric independent private values. HHS and LPV have used specific versions of (), respectively for testing common versus private values or identifying some common value models with symmetric homogeneous bidders. The equation (4.6) is obtained solving the differential equation (4.5) with the initial condition $B_1(\underline{\alpha}_1(z)|z) = U_1(\underline{\alpha}_1(z)|z)$ implied by (4.5) and $\Omega(\underline{\alpha}_1(z)|z) = 0$. The identity (4.6) parallels (2.6) but is not as useful due to the presence of the covariate z. For instance, if $U_1(A; z) = \sum_{i=1}^I z_i \gamma_i(A_i)$, (4.6) yields that $B_1(\alpha|z)$ writes $\sum_{i=1}^I z_i \gamma_i(\alpha|z)$, a specification which is not identified without further restriction.[13]

Lemma 9 however suggests a general matching procedure to estimate an identified valuation function model from an estimation of $B_1(\alpha|z)$, $B_1^{(1)}(\alpha|z)$ and $\Omega(\alpha|z)$. Indeed, for each

[13]For instance, if $I = 2$, setting $\gamma_1'(\alpha|z) = z_2\gamma_2(\alpha|z)/z_1$ and $\gamma_2'(\alpha|z) = z_1\gamma_1(\alpha|z)/z_2$ gives $z_1\gamma_1(\alpha|z) + z_2\gamma_2(\alpha|z) = z_1\gamma_1'(\alpha|z) + z_2\gamma_2'(\alpha|z)$ although $(\gamma_1'(\alpha|z), \gamma_2'(\alpha|z))$ generally differs from $(\gamma_1(\alpha|z), \gamma_2(\alpha|z))$. Interestingly, Graham, Hahn, Poirier and Powell (2015) consider a similar quantile specification in a different panel context which allows for identification.

$U_1(A;z)$ of the identified model, $U_1(\alpha|z)$ can be estimated and matched with the estimated $B_1(\alpha|z)+B_1^{(1)}(\alpha|z)\,\Omega(\alpha|z)$ to get an estimation of $U_1(A;z)$ through (4.5). Likewise, the integral in (4.6) can estimated and matched with the estimated bid quantile function $B_1(\alpha|z)$. Establishing the identification of a particular model can suggest alternative estimation procedures. For asymmetric private value function $U_1(A;z)=U_1(A_1;z)$, $U_1(\alpha|z)=U_1(\alpha;z)$ so that the value function is identified. The interdependent case where $U_1(A;z)$ may depend upon the signals of other bidders is more complicated and some restrictions yielding identification are proposed in the next section.

4.2 The valuation function

The equation (4.5) suggests to identify the valuation function from $U(\alpha|z)$. As shown by Laffont and Vuong (1996) this may not be feasible. To get a simple intuition of the identification issue and to understand how observed bidder covariate can help, it is useful to consider the two bidder case as in Section 4.2.1. Section 4.2.2 introduces a nonparametric model of valuation functions which is shown to be identified.

4.2.1 Identification issues

Consider a valuation function $U_1(A_1,A_2;z)$ and let $c(\cdot|\alpha)$ be the conditional pdf of the second bidder signal A_2 given that the first bidder signal A_1 is equal to α, assuming $\alpha\geq\underline{\alpha}_1(z)$. Then the conditional expectation $U_1(\alpha|z)$ of $U_1(A_1,A_2;z)$ on the event that a bid $B_1(a|z)$ wins is

$$\begin{aligned} U_1(a|\alpha,z) &= \int U_1(\alpha,t;z)\,\mathbb{I}\left[B_1(a|z)\geq B_2(t|z)\,dt\right]c(t|\alpha)\,dt \\ &= \int_0^{G_{12}(a|z)} U_1(\alpha,t;z)\,c(t|\alpha)\,dt \end{aligned}$$

where $G_{12}(a|z)=G_2[B_1(a|z)\,|z]$, recalling that $G_2(\cdot|z)$ is the cdf of B_2 given z. Hence

$$\frac{\partial U_1(a|\alpha,z)}{\partial a}=\frac{\partial}{a}\left[\int_0^{G_{12}(a|z)}c(t|\alpha)\,dt\right]=g_{12}(a|z)\,U_1(\alpha,G_{12}(a|z)\,;z)\,c(G_{12}(a|z)\,|\alpha)$$

where $g_{12}(a|z)$ is the derivative of $G_{12}(a|z)$ which is well-defined for $a\geq\underline{\alpha}_1(z)$ by Assumption DV-(ii) and Lemma 8-(iii). Similarly

$$\frac{\partial\omega(a|\alpha,z)}{\partial a}=g_{12}(a|z)\,c(G_{12}(a|z)\,|\alpha)$$

with $c(\cdot|\alpha) > 0$ and $g_{12}(a|z) > 0$ for a in $[\underline{\alpha}_1(z), 1]$ under Assumption DV-(ii). This gives that the identified $U_1(\alpha|z)$ is equal to

$$U_1(\alpha|z) = \frac{\left.\frac{\partial \overline{U}_1(a|\alpha,z)}{\partial a}\right|_{a=\alpha}}{\left.\frac{\partial \omega(a|\alpha,z)}{\partial a}\right|_{a=\alpha}} = U_1(\alpha, G_{12}(\alpha|z); z). \tag{4.7}$$

In the absence of z and if the bidders use the same strategy, say $B_1(\alpha)$,

$$G_{12}(\alpha) = G_1(B_1(\alpha)) = \alpha$$

and the function $U_1(\alpha|z)$ in (4.7) is equal to $U_1(\alpha, \alpha)$, from which it is difficult to recover the valuation function $U_1(\alpha_1, \alpha_2)$ without further information restriction as for instance assuming a private value specification $U_1(\alpha_1, \alpha_2) = U_1(\alpha_1)$. See Laffont and Vuong (1996) for a rigorous derivation of this negative identification result. Somaini (2015) proposes to use the variations of $G_{12}(\alpha|z)$ to identify the valuation function. He considers a valuation function $U_1(\alpha_1, \alpha_2; z_1)$, which depends upon z_1 only, in which case $U_1(\alpha_1|z)$ is equal to $U_1(\alpha_1, \alpha_2; z_1)$ by (4.7) setting $\alpha = \alpha_1$, fixing z_1 and choosing z_2 such that $\alpha_2 = G_{12}(\alpha_1|z_1, z_2)$. However, if bidders are close to symmetry, $G_{12}(\alpha|z)$ stays close to α and identification of $U_1(\alpha_1, \alpha_2; z_1)$ will only occur in a narrow zone around the diagonal $\alpha_1 = \alpha_2$. This holds even under asymmetry when $\alpha_1 = 1$ since $G_{12}(1|z) = 1$ for all z by the terminal bid condition $B_1(1|z) = B_2(1|z)$, so that this approach will only identify the valuation function near the diagonal $\alpha_1 = \alpha_2$ when α_1 is large. More generally, a lack of variation in $G_{12}(\alpha|z)$ will result in a poor identification of the valuation function. As seen from the next section, further functional restrictions may help to address these two pitfalls.

4.2.2 The mixed signal model

Our approach tries to make a better use of the variation of z by viewing each z_i as an observed component of the signal of bidder i which can be paired with the private A_i into a more relevant "mixed" signal $V_i(A_i; z_i)$. This leads to the *mixed signal model*

$$U_1(A; z) = \Phi[V_1(A_1; z_1), \ldots, V_I(A_I; z_I)] \tag{4.8}$$

where $\Phi(\cdot)$ is unknown and is a first parameter of interest. It will be assumed here that the mixed signals $V_i(A_i; z_i)$ satisfy a multiplicative decomposition

$$V_i(A_i; z_i) = \gamma_i(A_i) z_i,\ i = 1, \ldots, I \tag{4.9}$$

for some unknown nonnegative $\gamma_i(\cdot)$ to be also identified. This choice for the $V_i(A_i; z_i)$ forbids to have a valuation function only depending on A, i.e. $U_1(A; z) = U_1(A)$, which would not be identified by Laffont and Vuong (1996). Other functional forms can be considered for such a purpose but may not be as convenient. It will also be assumed in a first step that $\gamma_i(1) = 1$ for all i to ensure identification of the functional parameters in (4.8) and (4.9) but alternative

conditions will be considered below.[14] Note that this specification nests the asymmetric private value model, which corresponds to a function $\Phi(\cdot)$ satisfying $\Phi(v_1,\ldots,v_I)=\Phi(v_1)$, a restriction that can be tested from an estimation of $\Phi(\cdot)$. Some examples are as follows.

- **Additive valuation model**. Each bidder observes a component $\gamma_i(A_i)z_i$ of the total valuation of the auctioned good

$$U_1(A;z)=\sum_{i=1}^{I}\pi_i\gamma_i(A_i)z_i \text{ with } \pi_i\geq 0$$

  in which case $\Phi(v_1,\ldots,v_I)=\sum_{i=1}^{I}\pi_i v_i$. Since $\gamma_i(1)=1$, the bidders are initially given equal weights (up to z_i) and the weights π_i can be used to overweight or underweight some bidders. Note that those $\gamma_i(\cdot)$ associated with a vanishing π_i cannot be statistically identified. This specification can be useful for mineral rights auctions or more generally when bidders have specific expertise area, or when the valuation has a reputation components.

- **Auction with resale.** Suppose each bidder is in contact with a final buyer to whom he can sell the good at a price $\gamma_i(A_i)z_i$ if he wins the auction. However the winner may be also in position to sell to other final buyers. Suppose bidder 1 can sell to final buyers in a subset in a subset $\mathcal{I}$ of $\{1,\ldots,I\}$. If the prices $\gamma_i(A_i)z_i$ are revealed after the auction, bidder 1 may get

$$U_1(A;z)=\max_{i\in\mathcal{I}}\{\gamma_i(A_i)z_i\}$$

  in which case $\Phi(v_1,\ldots,v_I)=\max_{i\in\mathcal{I}}\{v_i\}$. If the prices $\gamma_i(A_i)z_i$ are not known, bidder 1 may organize a second price auction to allocate the good when he wins, in which case $\Phi(v_1,\ldots,v_I)$ will be the second largest v_i for i in $\mathcal{I}$. Other $\Phi(\cdot)$ would correspond to other resale mechanisms.

4.2.3 Identification of mixed signal valuation functions

The two bidder case. The two bidder case is helpful to understand how to identify the mixed signal model. For this specification, the identified function $U_1(\alpha|z)$ from (4.7) writes,

[14] Other nornalizations as $\gamma_i(1/2)=1$ can be considered but $\gamma_i(1)=1$ is less restrictive as the $\gamma_i(\cdot)$ will be identified over $\bigcup_{z\in\mathcal{Z}}[\underline{\alpha}_i(z),1]$. Using $\gamma_i(1/2)=1$ would request that $\underline{\alpha}_i(z)\geq 1/2$ for some z and would reduce the identification set to $\bigcup_{z\in\mathcal{Z},\underline{\alpha}_i(z)\geq 1/2}[\underline{\alpha}_i(z),1]$. This would also affect identification of the function $\Phi(\cdot)$. However, when there is no aggressive bidding so that $\underline{\alpha}_i(z)=0$ for all i and z, it is possible to consider the identification condition $\gamma_i(0)=1$ as done below.

for $\alpha \geq \underline{\alpha}_1(z)$,[15]

$$U_1(\alpha|z) = \Phi\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right] \tag{4.10}$$

where $G_{12}(\alpha|z) = B_2^{-1}\left[B_1(\alpha|z)|z\right]$ is such that $G_{12}(1|z) = 1$ as $B_1(1|z) = B_2(1|z)$ by the terminal bidding condition in Assumption DV and Lemma 8-(iii) which identifies bidding strategies with bid quantile functions. Since $\gamma_1(1) = \gamma_2(1) = 1$ it follows that $\Phi(\cdot)$ is identified as

$$U_1(1|z) = \Phi\left[\gamma_1(1) z_1, \gamma_2\left[G_{12}(1|z)\right] z_2\right] = \Phi(z_1, z_2).$$

In the private value case $\Phi(z_1, z_2) = \Phi(z_1)$, $\gamma_1(\cdot)$ is identified over $\bigcup_{z\in\mathcal{Z}}\left[\underline{\alpha}_1(z), 1\right]$, assuming for instance that $\Phi(\cdot)$ is strictly increasing. The coefficient $\gamma_2(\cdot)$ cannot be identified but this is not relevant since the valuation function of bidder 1 does not depend upon $\gamma_2(\cdot)$.

The interdependent value case of a function $\Phi(z_1, z_2)$ which depends both z_1 and z_2 is more involved. It would be tempting to take $z_1 = 1$ and $z_2 = 0$ and to invert

$$\Phi\left[\gamma_1(\alpha), 0\right] = U_1(\alpha|1, 0)$$

to identify $\gamma_1(\cdot)$ over $\left[\underline{\alpha}_1\left[(1,0)\right], 1\right]$. This would however identify $\gamma_1(\cdot)$ over a smaller set than $\bigcup_{z\in\mathcal{Z}}\left[\underline{\alpha}_1(z), 1\right]$. Taking $z_2 = 0$ may also be an issue in view of the form of the mixed signal $V_2(\alpha; z) = \gamma_2(\alpha) z_2$, which vanishes for $z_2 = 0$. To see this, consider the Resale Example and assume bidder 2 valuation is the private value $V_2(A_2; z) = \gamma_2(A_2) z_2$. If $z_2 = 0$, bidder 2 will bid 0 whatever A_2 is, a bid function would violate our identifying assumption, which supposes that the bidding strategies of all bidders are strictly increasing for high signals.[16]

An alternative identification approach is to show that $\left(\gamma_1(\cdot), \gamma_2\left[G_{12}(\cdot|z)\right]\right)$ solves a differential system which has a unique solution given the initial condition

$$\left(\gamma_1(1), \gamma_2\left[G_{12}(1|z)\right]\right) = (1, 1).$$

Differentiating (4.10) with respect to α yields

$$\begin{aligned}\Phi_{z_1}&\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right] z_1 \frac{d\gamma_1(\alpha)}{d\alpha} \\ &+ \Phi_{z_2}\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right] z_2 \frac{\partial\left\{\gamma_2\left[G_{12}(\alpha|z)\right]\right\}}{\partial\alpha} = \frac{\partial U_1(\alpha|z)}{\partial\alpha}\end{aligned}$$

[15] The expression of $U_1(\alpha|z)$ has the flavour of a multi-index model but the presence of the transformation $G_{12}(\alpha|z)$ does not allow to apply standard average derivative technique to identify $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ in a simple way.

[16] Bidder 1 may bid just over 0 whatever his signal is, so both bidders can use constant strategies. This would forbid identification of the signal distribution. Note that Assumption DV forbids $z_1 = 0$ and $z_2 = 0$, but that an assumption rejecting the existence of a z such that $U_i(A; z) = 0$ for all A is implicit here.

where $\Phi_{z_i}(t_1,t_2) = \frac{\partial \Phi(t_1,t_2)}{\partial z_i}$. Before differentiating (4.10) with respect to z_2, observe

$$\begin{aligned}
\frac{\partial \left\{\gamma_2 \left[G_{12}(\alpha|z)\right]\right\}}{\partial z_2} &= \gamma_2^{(1)}\left[G_{12}(\alpha|z)\right]\frac{\partial G_{12}(\alpha|z)}{\partial z_2} = \gamma_2^{(1)}\left[G_{12}(\alpha|z)\right] g_{12}(\alpha|z)\frac{\frac{\partial G_{12}(\alpha|z)}{\partial z_2}}{g_{12}(\alpha|z)} \\
&= \frac{\partial \left\{\gamma_2 \left[G_{12}(\alpha|z)\right]\right\}}{\partial \alpha}\frac{\frac{\partial G_{12}(\alpha|z)}{\partial z_2}}{g_{12}(\alpha|z)}.
\end{aligned} \tag{4.11}$$

Hence

$$\begin{aligned}
&\Phi_{z_2}\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right]\gamma_2\left[G_{12}(\alpha|z)\right] \\
&\quad + \Phi_{z_2}\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right] z_2 \frac{\frac{\partial G_{12}(\alpha|z)}{\partial z_2}}{g_{12}(\alpha|z)}\frac{\partial\left\{\gamma_2\left[G_{12}(\alpha|z)\right]\right\}}{\partial\alpha} = \frac{\partial U_1(\alpha|z)}{\partial z_2}.
\end{aligned}$$

Let

$$\begin{aligned}
\mathbf{D}[\Phi,\gamma](\alpha|z) &= \begin{bmatrix} \Phi_{z_1}\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right] & 0 \\ 0 & \Phi_{z_2}\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right] \end{bmatrix}, \\
\mathbf{G}_2(\alpha|z) &= \begin{bmatrix} 1 & 1 \\ 0 & \frac{\partial G_{12}(\alpha|z)}{\partial z_2} \end{bmatrix}, \\
\mathbf{\Psi}[\Phi,\gamma](\alpha|z) &= \begin{bmatrix} \frac{\partial U_1(\alpha|z)}{\partial \alpha} \\ g_{12}(\alpha|z)\left\{\frac{\partial U_1(\alpha|z)}{\partial z_2} - \Phi_{z_2}\left[\gamma_1(\alpha) z_1, \gamma_2\left[G_{12}(\alpha|z)\right] z_2\right]\gamma_2\left[G_{12}(\alpha|z)\right]\right\} \end{bmatrix}
\end{aligned}$$

recalling that $U_1(\cdot|\cdot)$ and $\Phi(\cdot)$ are identified at this stage. Combining the two differential equations derived from (4.10) shows that $\gamma_1(\cdot)$ and $\gamma_2[G_{12}(\cdot|z)]$ solves the differential system

$$\begin{bmatrix} \frac{d}{d\alpha}\gamma_1(\alpha) \\ \frac{d}{d\alpha}\left\{\gamma_2\left[G_{12}(\alpha|z)\right]\right\} \end{bmatrix} = \left\{\mathbf{D}[\Phi,\gamma](\alpha|z)\,\mathbf{G}_2(\alpha|z)\right\}^{-1}\mathbf{\Psi}[\Phi,\gamma](\alpha|z) \tag{4.12}$$

α is in $[\underline{\alpha}_1(z), 1]$, with the initial condition $(\gamma_1(1), \gamma_2[G_{12}(1|z)]) = (1,1)$. Standard rank and smoothness conditions on $\mathbf{D}[\Phi,\gamma](\alpha|z)\,\mathbf{G}_2(\alpha|z)$ will ensure that its unique solution identifies $\gamma_1(\cdot)$ on $\bigcup_{z\in\mathcal{Z}}[\underline{\alpha}_1(z), 1]$ and $\gamma_2(\cdot)$ on $\bigcup_{z\in\mathcal{Z}}[\underline{\alpha}_2(z), 1]$ since $\underline{\alpha}_2(z) = G_{12}(\underline{\alpha}_1(z)|z)$.

The three bidder case. The case of a larger number of bidders is more difficult and cannot be dealt with standard differential system results. This is because the identified $U_1(\alpha|z)$ is a multiple integral of order $I-2$. To see this, consider the three bidder case for which $U_1(a|\alpha,z)$ writes

$$\int_0^{G_{12}(a|z)}\int_0^{G_{13}(a|z)} \Phi\left(z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3(t_3)\right) c(t_2,t_3|\alpha)\, dt_2 dt_3$$

where $c(\cdot,\cdot|\alpha)$ stands now for the pdf of (A_2, A_3) given $A_1 = \alpha$. The function $U_1(\alpha|z)$ multiplied by the identified $\left.\frac{\partial\omega(a|\alpha,z)}{\partial a}\right|_{a=\alpha}$ is

$$\begin{aligned} & g_{12}(\alpha|z)\int_0^{G_{13}(\alpha|z)} \Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2\left[G_{12}(\alpha|z)\right], z_3\gamma_3(t_3)\right] c\left[G_{12}(\alpha|z), t_3|\alpha\right] dt_3 \\ & + g_{13}(\alpha|z)\int_0^{G_{12}(\alpha|z)} \Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3\left[G_{13}(\alpha|z)\right]\right] c\left[t_2, G_{13}(\alpha|z)|\alpha\right] dt_2. \end{aligned} \tag{4.13}$$

The expression (4.13) is not as convenient as (4.7) to identify $\Phi(\cdot)$, due to the two integrals.[17] In the absence of aggressive bidding, that is if

$$\underline{\alpha}_i(z) = 0 \text{ for } i = 1,2,3 \text{ and all } z \text{ of } \mathcal{Z},$$

it holds under standard continuity conditions that

$$\lim_{\alpha\downarrow 0} U_1(\alpha|z) = \Phi\left[z_1\gamma_1(0), z_2\gamma_2(0), z_3\gamma_3(0)\right] \tag{4.14}$$

as seen from (4.13), $G_{1i}(\underline{\alpha}_1(z)|z) = 0$ for $i = 2,3$ and recalling

$$\begin{aligned} \left.\frac{\partial\omega(a|\alpha,z)}{\partial a}\right|_{a=\alpha} &= g_{12}(\alpha|z)\int_0^{G_{13}(\alpha|z)} c\left[G_{12}(\alpha|z), t_3|\alpha\right] dt_3 \\ &+ g_{13}(\alpha|z)\int_0^{G_{12}(\alpha|z)} c\left[t_2, G_{13}(\alpha|z)|\alpha\right] dt_2. \end{aligned}$$

Hence it is possible to identify $\Phi(\cdot)$ up to scale through $U_1(0|z)$ provided $\gamma_i(0) \neq 0$ for all i. Recovering the $\gamma_i(\cdot)$'s can be done differentiating (4.13) with respect to α, z_2 and z_3 to obtain an integro-differential system which can play the role of (4.12).

The general case. Let us now introduce some rank and smoothness conditions ensuring identification of the mixed signal specification (4.8-4.9). As above, $\Phi_{z_i}(z) = \frac{\partial}{\partial z_i}\Phi(z)$.

Assumption MSM. *The valuation function of bidder 1 is given by (4.8), with a twice continuously differentiable $\Phi(\cdot)$ over $\overline{\mathcal{Z}}$, and (4.9) with continuously differentiable $\gamma_i(\cdot)$ taking value in $(0,1]$. It also holds that:*

(i) There is a subset $\mathcal{I}$ of $\{1,\ldots,I\}$ such that $\Phi(z_1,\ldots,z_I) = \Phi[z_i, i\in\mathcal{I}]$ for all z of $\mathcal{Z}$. For all i of $\mathcal{I}$ $\Phi_{z_i}(z) \neq 0$ except for a finite number of z of $\overline{\mathcal{Z}}$ and these functions are Lipschitz.

(ii) The joint signal distribution is identified over $[0,1]^I$. For all i,j in $\mathcal{I}$, all z of $\mathcal{Z}$

[17] Proceeding as in the two bidder case by recovering $\Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2\left[G_{12}(\alpha|z)\right], z_3\gamma_3\left[G_{13}(\alpha|z)\right]\right]$ from (4.13) does not seem feasible. A similar problem is to recover $f(\alpha,a)$ from $\int_0^\alpha f(\alpha,t)\,dt$, $\alpha\in[0,1]$, which is impossible as $g(\alpha,t) = f(\alpha,t) + th'(t) + h(t) - h(\alpha)$ is such that $\int_0^\alpha g(\alpha,t)\,dt = \int_0^\alpha f(\alpha,t)\,dt$ but $g(\alpha,\alpha) \neq f(\alpha,\alpha)$ provided $\alpha h'(\alpha) \neq 0$.

and all $\alpha > \underline{\alpha}_1(z)$, $G_i[B_1(\alpha|z)|z]$ is differentiable with respect to z_j in $(0, \overline{z}]$. Moreover, for all z of $\mathcal{Z}$,

$$\det\left[\frac{\partial G_i[B_1(\alpha|z)|z]}{\partial z_j}, \quad i,j \in \mathcal{I}\backslash\{1\}\right] = 0$$

for only a finite number of $\alpha > \underline{\alpha}_1(z)$.

(iii) For all i,j in $\mathcal{I}$, all z of $\mathcal{Z}$ and all $\alpha > \underline{\alpha}_1(z)$, $g_{1i}(\alpha|z) = \frac{\partial}{\partial\alpha}\{G_i[B_1(\alpha|z)|z]\}$ is continuously differentiable with respect to α and z.

It is important to note that the valuation functions of bidders $i = 2, \ldots, I$ is left unspecified. These bidders can have private value $V_i(A_i; z)$, which are not restricted by (4.8-4.9) and can be estimated using ASQR techniques and an estimation of $\Omega_i(\alpha|z)$ by (4.5) as $U_i(\alpha|z) = V_i(A_i; z)$. The other main condition in Assumption MSM is a full-rank condition for the matrix with entries $\partial G_{1i}(\alpha|z)/\partial z_j$, $2 \leq i, j \leq I$ which must hold for almost all α, see (ii). Note that this condition is in principle testable from the data. This can be weakened assuming that the full rank condition holds over a non empty subset $\mathcal{Z}_0$ of $\mathcal{Z}$ which, in practice, can be estimated from the data.

Assumption MSM-(i) puts some smoothness restrictions on $\Phi(\cdot)$, which holds for instance if this function is twice continuously differentiable. This is however for technical reasons and can be relaxed by smoothing the identified function $U_1(\alpha|z)$ with respect to z before differentiating with respect to z. Note that Assumption MSM allows for functions $\Phi(\cdot)$ and coefficients $\gamma_i(\cdot)$ which are not necessarily increasing. This contrasts with the Bayesian Nash Equilibrium existence results of Lizzeri and Persico (2000) or Reny and Zamir (2004) which typically assume that these functions are increasing. The next Theorem establishes identification of the first bidder mixed signal valuation function from the observations of all the bids of a first-price auction.

Theorem 10 *Suppose that Assumption DV-(ii,iii) and MSM-(i,ii) are true. It holds that:*

i. Suppose $I = 2$ and that the coefficients $\gamma_i(\cdot)$ satisfy the identification restriction $\gamma_i(1) = 1$, $i = 1, \ldots, I$. Then $\Phi(\cdot)$ is identified over $\mathcal{Z}$ and, for all i in $\mathcal{I}$, $\gamma_i(\cdot)$ is identified over $\bigcup_{z\in\mathcal{Z}}[\underline{\alpha}_i(z), 1]$.

ii. Suppose $I \geq 3$, that $\underline{\alpha}_i(z) = 0$ for all z of $\mathcal{Z}$ and all $i = 1, \ldots, I$, that the coefficients $\gamma_i(\cdot)$ satisfy the identification restriction $\gamma_i(0) = 1$, $i = 1, \ldots, I$, and that Assumption MSM-(iii) holds. Then $\Phi(\cdot)$ is identified over $\mathcal{Z}$ and, for all i in $\mathcal{I}$, $\gamma_i(\cdot)$ is identified over $[0, 1]$.

Theorem 10 allows for aggressive bidding in the two bidder case which is simpler. In the case of a larger number of bidders, the identification of $\Phi(\cdot)$ is achieved assuming non aggressive bidding and using a signal vector $A = 0$. Compared to Somaini (2015), these two

identification results yield identification of the valuation function $U_1(A; z)$ on a larger set of signals.[18]

The proofs of Theorem 10 (i) and (ii) are somehow similar, establishing first identification of $\Phi(\cdot)$ and, in a second step, differentiating $U_1(\alpha|z)$ or $\left.\frac{\partial\omega(a|\alpha,z)}{\partial a}\right|_{a=\alpha} U_1(\alpha|z)$ with respect to α and z_i, $i \geq 2$ to find an (integro)differential system characterizing the coefficients $\gamma_i(\cdot)$ in a unique way, as explained in the heuristic exposition above. This last step only involves one value of z, suggesting that the $\gamma_i(\cdot)$ are overidentified. Allowing for $\gamma_i(\cdot)$ depending upon z would still allow to identify $\Phi(\cdot)$ under the condition of Theorem 10. Differentiating with respect to α and z would give a partial (integro)differential system which may be more difficult to study.

4.3 Estimation

Estimating a mixed signal specification can be done following the identification strategy of Theorem 10, estimating $\Phi(\cdot)$ from an estimation of $U_1(\alpha|z)$ based on (4.5) in Lemma 9

$$\widehat{U}_1(\alpha|z) = \widehat{B}_1(\alpha|z) + \widehat{B}_1^{(1)}(\alpha|z)\,\widehat{\Omega}(\alpha|z)$$

using unconstrained ASQR estimators to get consistency of $\widehat{U}_1(\alpha|z)$ at extreme quantiles. If $I = 2$ and under the identification condition of Theorem 10-(i), an estimator of $\Phi(\cdot)$ is $\widehat{U}_1(1|z)$. Under the conditions of Theorem 10-(ii), estimators of $\Phi(\cdot)$ are $\widehat{U}_1(0|z)$ or $\widehat{B}_1(0|z)$ which potentially converges with a faster rate. In the case of a parametric $\Phi_\theta(\cdot)$ as in the additive valuation example, the parameter θ can be estimated matching $\Phi_\theta(\cdot)$ with the nonparametric $\widehat{\Phi}(\cdot)$, for instance minimizing $\int_{\mathcal{Z}} \left(\Phi_\theta(z) - \widehat{\Phi}(z)\right)^2 dz$ with respect to θ to obtain an improved estimator of $\Phi(\cdot)$. Nonparametric significance techniques can be used to estimate the set $\mathcal{I}$ of signals A_i appearing in the valuation function $U_1(A; z)$ of the first bidder.

Estimation of the coefficients $\gamma_i(\cdot)$ can be done matching $\widehat{U}_1(\alpha|z)$ with an estimation $\widehat{U}_1\left(\alpha|z, \widehat{\Phi}, \gamma\right)$ of $U_1(\alpha|z)$ treating $\widehat{\Phi}(\cdot)$ and the $\gamma_i(\cdot)$ as true values. For instance, if $I = 2$, (4.10) suggests

$$\widehat{U}_1\left(\alpha|z, \widehat{\Phi}, \gamma\right) = \widehat{\Phi}\left[\gamma_1(\alpha)\, z_1, \gamma_2\left(\widehat{G}_{12}(\alpha|z)\right)\right]$$

[18] Consider the two bidder case for the sake of brevity and $U_1(\alpha_1, \alpha_2; z) = U_1(\alpha_1, \alpha_2; z_1)$ as in Somaini (2015). Then (4.7) shows that, for each z_1, $U_1(\alpha_1, \alpha_2; z_1)$ is identified over

$$\left\{(\alpha_1, \alpha_2); \exists z_2 \in (0, \overline{z}] \text{ such that } \alpha_1 \geq \underline{\alpha}_1(z_1, z_2) \text{ and } \alpha_2 = G_{12}(\alpha_1|z_1, z_2)\right\}.$$

See Somaini (2015) for the general case. In particular α_2 must be equal to 1 if $\alpha_1 = 1$ as $G_{12}(1|z) = 1$ for all z due to the terminal bidding condition. Since $G_{12}(\alpha_1|z) \geq \underline{\alpha}_2(z)$ when $\alpha_1 \geq \underline{\alpha}_1(z)$, the identification set of Somaini (2015) is strictly smaller than the identification set $[\min_{z\in\mathcal{Z}} \underline{\alpha}_1(z), 1] \times [\min_{z\in\mathcal{Z}} \underline{\alpha}_2(z), 1]$ of Theorem 10-(i).

where $\widehat{G}_{12}(\alpha|z) = \widehat{G}_2\left[\widehat{B}_1(\alpha|z)|z\right]$, $\widehat{G}_2(b|z)$ being an estimator of the conditional cdf of the bids of the second bidder. An example of estimators for $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ is the minimizer of

$$\int_{\mathcal{Z}} \int_{\underline{\widehat{\alpha}}_1(z)}^{1} \left(\widehat{U}_1\left(\alpha|z, \widehat{\Phi}, \gamma\right) - \widehat{U}_1(\alpha|z)\right)^2 d\alpha dz$$

over a sieve Γ_n satisfying the identification restriction used to construct $\widehat{\Phi}(\cdot)$, that is $\gamma_1(1) = \gamma_2(1) = 1$ if $\widehat{\Phi}(z) = \widehat{U}_1(1|z)$ or $\gamma_1(0) = \gamma_2(0) = 1$ if $\Phi(\cdot)$ is estimated using the lower extreme quantile. The case of a larger number of bidders complicates the choice of the estimator $\widehat{U}_1\left(\alpha|z, \widehat{\Phi}, \gamma\right)$ but this can be done using extensions of (4.13) valid for a general I, noting that $U_1(\alpha|z)$ can be written as a combination of conditional expectations given z, $B_1 = B_1(\alpha|z)$ and $B_i = B_1$, $i = 2, \ldots, I$. Such estimation procedure can be easily to implement when $\widehat{\Phi}(\cdot)$ is linear as in the additive valuation example. Note that iterative updating is possible.

5 Simulation experiments

This section reports the results of a simulation experiment for the AQR estimation under symmetric IPV. It also describes some implementation details such as bandwidth choice and computation algorithm.

5.1 The model

The considered private value quantile regression model is

$$\begin{aligned} &V(\alpha|x) = \gamma_0(\alpha) + x_1 + \gamma_2(\alpha)x_2, \\ &\quad \gamma_2(\alpha) = 0.4(1 - \exp(-6\alpha)), \quad \gamma_1(\alpha) = 1 \quad, \gamma_0(\alpha) = -0.1\log(1 - (1 - 1/e)\alpha). \end{aligned}$$

The quantile function $\gamma_0(\alpha)$ is the one of a exponential distribution with scale parameter 0.1, truncated over $[0, 0.1]$. The slope coefficient $\gamma_2(\alpha)$ is such that the variable x_2 does not affect the bidders with a low α but increases with α. By contrast, the slope coefficient of x_1 is equal to 1 independently of α. The covariates x_1 and x_2 are two independent uniform variables. As shown by (2.6) in Proposition 2, the corresponding bid quantile regression model is given by

$$B(\alpha|x, I) = \beta_0(\alpha|I) + x_1 + \beta_2(\alpha|I)x_2, \quad \beta_j(\alpha|I) = \frac{I-1}{\alpha^{I-1}} \int_0^\alpha t^{I-2}\gamma_j(t)\,dt, \; j = 0, 2$$

where the coefficient of x_1 is also equal to 1. The simulation experiment considers two numbers of bidders, $I = 2$ and $I = 5$. Since

$$\widehat{\gamma}_j(\alpha|I) = \widehat{\beta}(\alpha|I) + \frac{\alpha\widehat{\beta}^{(1)}(\alpha|I)}{I-1}$$

a larger number of bidders may reduce the contribution of the term $\widehat{\beta}^{(1)}(\alpha|I)$, which converges with a slow rate, to the estimation of the private value slope coefficients.

The simulation uses simulated bids

$$B_{i\ell} = B(A_{i\ell}|x_\ell, I), \quad i = 1, \ldots, I,$$

where the independent $A_{i\ell}$ have a common uniform distribution over $[0,1]$. This simulation step uses an explicit computation of bid slope coefficients

$$\beta_0(\alpha|I) = 0.1b_0\left(\left(1 - \frac{1}{e}\right)\alpha|I\right) \text{ and } \beta_2(\alpha|I) = 0.4b_2(6\alpha|I),$$

with

$$\begin{aligned}
b_0(\alpha|I=2) &= \frac{(1-\alpha)\log(1-\alpha)+\alpha}{\alpha}, \\
b_0(\alpha|I=5) &= \frac{(12-12\alpha^4)\log(1-\alpha)+3\alpha^4+8\alpha^3+6\alpha^2+12\alpha}{12\alpha^4}
\end{aligned}$$

and

$$\begin{aligned}
b_2(\alpha|I=2) &= \frac{\exp(-\alpha)-1+\alpha}{\alpha}, \\
b_2(\alpha|I=5) &= \frac{(4\alpha^3+12\alpha^2+24\alpha+24)\exp(-\alpha)+\alpha^4-24}{\alpha^4}.
\end{aligned}$$

In each simulations, the total number of bids is set to 100, which is 10 time less than in the simulation experiment of GPV, which did not include covariate. Hence there are 50 auctions when $I = 2$ and 20 for $I = 5$. Each simulation experiments make use of 10,000 replications.

5.2 AQR computation details

The AQR procedure is implemented setting $s+1 = 2$, meaning that up to the second derivative of the $\beta_j(\alpha|I)$, $j = 0, 1, 2$ are estimated. The kernel function is $K(a) = (1-a^2)\mathbb{I}(a \in [-1,1])$. The integral in the AQR objective function is replaced with a Riemann sum using a discretization of $[-1,1]$ in 1,000 points. The AQR estimator is computed using the Majorize-Minimize (MM) algorithm of Hunter and Lange (2000). The AQR estimator is computed for $\alpha = 0, 1/100, 2/100, \ldots, 1$. For $\alpha = 1/2$, the MM algorithm is initialized with the median of a pilot model described below. For $\alpha_k > 1/2$ ($\alpha_k < 1/2$), the MM algorithm is initialized

with the AQR estimator computed for α_{k-1} (α_{k+1} respectively). The bandwidth is chosen according to the pilot model introduced now.

5.2.1 A pilot model

The initialization of the MM algorithm and the bandwidth choice both rely on the pseudo regression model

$$B_{i\ell} = \beta_0 + \beta_1 x_{1\ell} + \beta_2 x_{2\ell} + e_{i\ell},\ i = 1, \ldots, I,\ \ \ell = 1, \ldots, L. \tag{5.1}$$

where the disturbance error term $e_{i\ell}$ has an exponential distribution with scaling parameter λ_1 truncated over $[0, \lambda_2]$, with a quantile function equal to

$$Q_\lambda(\alpha) = -\frac{1}{\lambda_1} \log\left(1 - \left(1 - \exp\left(-\lambda_1 \lambda_2\right)\right)\alpha\right).$$

The parameters of this pseudo model can be estimated using the OLS slope coefficients $\widehat{b}_0$, $\widehat{b}_1$, $\widehat{b}_2$ and the OLS residuals $\widehat{\varepsilon}_{i\ell}$. Estimators for β_1 and β_2 are $\widehat{b}_1$ and $\widehat{b}_2$ respectively. A natural estimator of λ_2 is the range

$$\widehat{\lambda}_2 = \max_{1 \le \ell \le L, 1 \le i \le I} \widehat{\varepsilon}_{i\ell} - \min_{1 \le \ell \le L, 1 \le i \le I} \widehat{\varepsilon}_{i\ell}.$$

β_0 and the disturbance terms $e_{i\ell}$ are estimated using

$$\widehat{\beta}_0 = \widehat{b}_0 + \min_{1 \le \ell \le L, 1 \le i \le I} \widehat{\varepsilon}_{i\ell}, \quad \widehat{e}_{i\ell} = \widehat{\varepsilon}_{i\ell} - \min_{1 \le \ell \le L, 1 \le i \le I} \widehat{\varepsilon}_{i\ell}.$$

The estimation of λ_1 is more difficult. The proposed estimator uses the formula

$$\lambda_1 = \frac{1}{Q_\lambda(0.95) - Q_\lambda(0.05)} \log\left(\frac{1 - \left(1 - \exp\left(-\lambda_1\lambda_2\right)\right)0.05}{1 - \left(1 - \exp\left(-\lambda_1\lambda_2\right)\right)0.95}\right)$$

in a recursive way. An estimator $\widehat{\lambda}_1$ is the limit of the sequence $\widehat{\lambda}_{1,k}$ with

$$\widehat{\lambda}_{1,k+1} = \frac{1}{\widehat{Q}(0.95) - \widehat{Q}(0.05)} \log\left(\frac{1 - \left(1 - \exp\left(-\widehat{\lambda}_{1,k}\widehat{\lambda}_2\right)\right)0.05}{1 - \left(1 - \exp\left(-\widehat{\lambda}_{1,k}\widehat{\lambda}_2\right)\right)0.95}\right)$$

taking $\widehat{\lambda}_{1,0} = 1/\left(\widehat{Q}(0.95) - \widehat{Q}(0.05)\right)$ where $\widehat{Q}(0.95)$ and $\widehat{Q}(0.05)$ are respectively the 95% and 5% sample quantiles of the estimated disturbance terms $\widehat{e}_{i\ell}$. In the experiment below, this algorithm is iterated one hundred times to compute $\widehat{\lambda}_1$. This gives a pseudo median

$$-\frac{1}{\widehat{\lambda}_1} \log\left(1 - \frac{1}{2}\left(1 - \exp\left(-\widehat{\lambda}_1\widehat{\lambda}_2\right)\right)\right)$$

which is used to initialize the MM algorithm of the AQR estimator.

5.2.2 Bandwidth choice

A pilot bandwidth can be proposed assuming in a first step that the pseudo regression model (5.1) and using a modified Theorem 4 to propose an expansion for the modified IMSE

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

where $f(x|I)$ is the pdf of x_ℓ given the number of bidders is equal to I. Recall that $x_\ell = (x_{1\ell}, x_{2\ell})'$, $X_\ell = (1, x_{1\ell}, x_{2\ell})'$ and define, s_1 being the $1 \times (s+2)$ vector $(0, 1, 0, \ldots, 0)$ and $\mathbb{E}_I [\cdot]$ standing for an expectation computed given that the number of bidders is I,

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

The expressions of the asymptotic bias and variance in (3.6) and (3.5) and Theorem 5 suggest that the modified IMSE leading term is

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

Minimizing this leading term gives the infeasible optimal bandwidth

$$\left(\frac{v^2 \int_0^1 \operatorname{Tr} \left(M_2 (\alpha)^{-1} M_2 M_2 (\alpha)^{-1} M_2 \right) \alpha^2 d\alpha}{2 (s+1) \left(s_1' \Pi_2^{-1} \Pi_1 \right)^2 \int_0^1 \left(M_1 (\alpha)' M_2 (\alpha)^{-1} M_2 M_2 (\alpha)^{-1} M_1 (\alpha)' \right) d\alpha} \right)^{\frac{1}{2s+3}}.$$

If (5.1) holds, this expression simplifies as follows because the conditional quantile function of the observations is $B(\cdot|x, I) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + Q_\lambda (\cdot)$ where $Q_\lambda (\cdot)$ does not depend

upon X and can therefore taken out of expectation. This gives

$$M_1 = \mathbb{E}_I [X_\ell] \frac{\alpha Q_\lambda^{(s+2)}(\alpha)}{Q_\lambda^{(1)}(\alpha)} = M_1 \frac{\alpha Q_\lambda^{(s+2)}(\alpha)}{Q_\lambda^{(1)}(\alpha)},$$
$$M_2(\alpha, I) = \frac{\mathbb{E}_I [X_\ell X_\ell']}{Q_\lambda^{(1)}(\alpha)} = \frac{M_2}{Q_\lambda^{(1)}(\alpha)}$$

with

$$Q_\lambda^{(p)}(\alpha) = \frac{p!\,(1-\exp(-\lambda_1\lambda_2))^p}{\lambda_1} \frac{1}{(1-(1-\exp(-\lambda_1\lambda_2))\,\alpha)^p}, \quad p = 1, s+2.$$

Further elementary but cumbersome calculations yield that

$$\begin{aligned}
C_v &= \int_0^1 \operatorname{Tr}\left(M_2(\alpha)^{-1} M_2 M_2(\alpha)^{-1} M_2\right)\alpha^2 d\alpha = \int_0^1 \left(\alpha Q_\lambda^{(1)}(\alpha)\right)^2 d\alpha \\
&= \frac{\exp(\lambda_1\lambda_2) - \exp(-\lambda_1\lambda_2) + 2\lambda_1\lambda_2}{\lambda_1^2(1-\exp(-\lambda_1\lambda_2))}, \\
C_b &= \int_0^1 \left(M_1(\alpha)' M_2(\alpha)^{-1} M_2 M_2(\alpha)^{-1} M_1(\alpha)'\right) d\alpha = M_1' M_2^{-1} M_1' \int_0^1 \left(\alpha Q_\lambda^{(s+2)}(\alpha)\right)^2 d\alpha \\
&= \frac{((s+2)!^2(1-\exp(-\lambda_1\lambda_2)))^{2s+1}}{\lambda_1^2} \left(\frac{\exp((2s+3)\lambda_1\lambda_2) - 1}{2s+3}\right. \\
&\qquad \left. - \frac{\exp((2s+4)\lambda_1\lambda_2) - 1}{s+2} + \frac{\exp((2s+5)\lambda_1\lambda_2) - 1}{2s+5}\right)
\end{aligned}$$

which can be estimated plugging in $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$. This allows to compute the pilot bandwidth

$$\widehat{h} = \left(\frac{\widehat{C}_v}{2(s+1)\,\widehat{C}_b\left(s_1'\Pi_2^{-1}\Pi_1\right)^2} \frac{1}{\widehat{M_1'}\widehat{M_2}^{-1}\widehat{M_1}} \frac{1}{LI}\right)^{\frac{1}{2s+3}},$$

$\widehat{M_1} = \sum_{\ell=1}^L X_\ell/(LI)$ and $\widehat{M_2} = \sum_{\ell=1}^L X_\ell X_\ell'/(LI)$, which can be used in a first step. In practice, it may give a too small and variable bandwidth. In the simulation experiment with $I = 5$ the bandwidth 5% and 95% quantiles were 0.0063 and 0.1038 respectively, with a minimum 0.00003 and maximum 0.4749. Hence $\widehat{h}$ has been truncated to a minimum value of 0.05 and a maximal value of 0.3.

5.3 Simulation results

The two next figures summarize the results of 10,000 replications of the private value AQR estimator for twenty auctions with five bidders, while the third figure considers samples with fifty auctions and two bidders. In all figures, the solid line is the truth, the dashed line is the

median of the estimation and the two dotted lines are the individual 5% and 95% quantiles across simulations.

All these graphs suggest that the bias of the AQR estimator is very small, with a maximum of 0.07 in Figure 4 at the extreme upper quantile $\alpha = 1$. Figure 2 considers the AQR estimator of the coefficients $\gamma_j(\cdot)$, for which the individual 5%-95% estimation quantile differences all have a similar shape, decreasing from a maximum around 0.2 at $\alpha = 0$ to a minimum around 0.1 for $a = 0.5, 0.6$, and with a sharp increase starting at $\alpha = 0.95$ to reach 0.25 or 0.30 at $\alpha = 1$. This suggests a larger estimation variance at the tails of a distribution, but with a very reasonable magnitude in view of the small sample size of 100 bids.

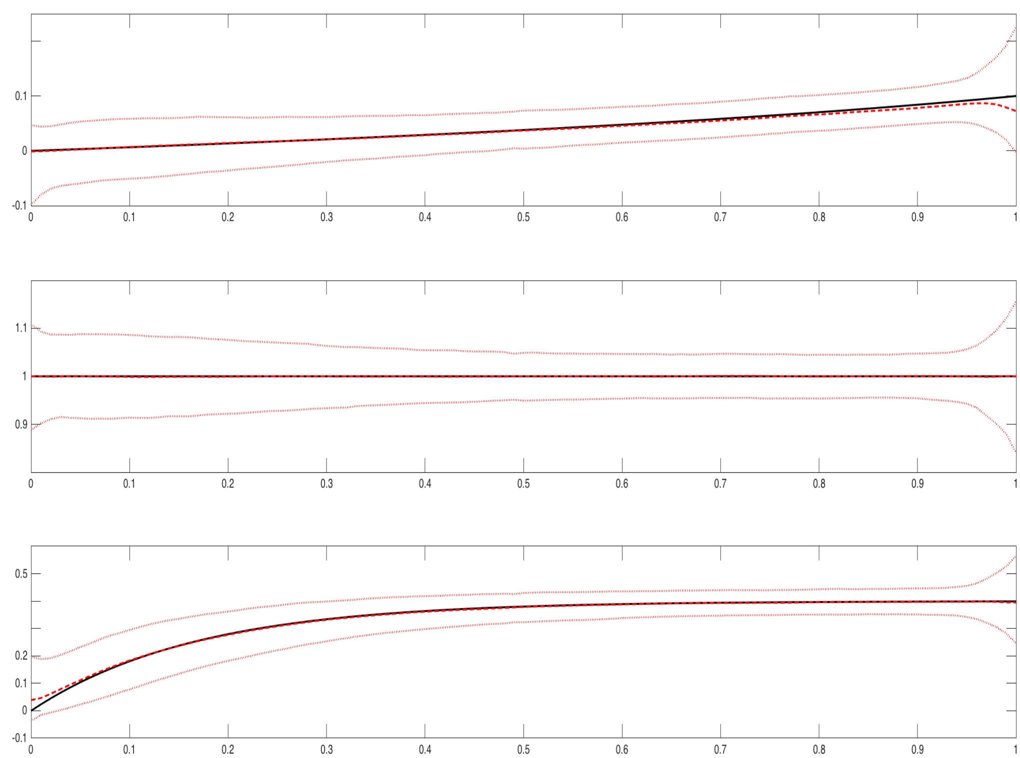


Figure 2: AQR estimation of $\gamma_0(\alpha)$ (bottom), $\gamma_1(\alpha)$ (middle) and $\gamma_2(\alpha)$ (top) from 10,000 replications of a sample with $L = 20$, $I = 5$. The black line is the true $\gamma(\alpha)$, the dashed line is the median of the estimation across 10,000 replications, the dotted lines are the individual 5% and 95% quantile of the estimation for each quantile levels.

Figure 3 considers the private value quantile function $V(\cdot|x)$ for low quality good with $x_1 = x_2 = 0.1$, average good ($x_1 = x_2 = 0.5$) and high quality good ($x_1 = x_2 = 0.9$).

The qualitative features of Figures 2 and 3 are very similar, with less dispersion for the estimation of quantile functions than for the quantile regression slope functions. In particular the estimation of private value quantile function for the average good looks very good.

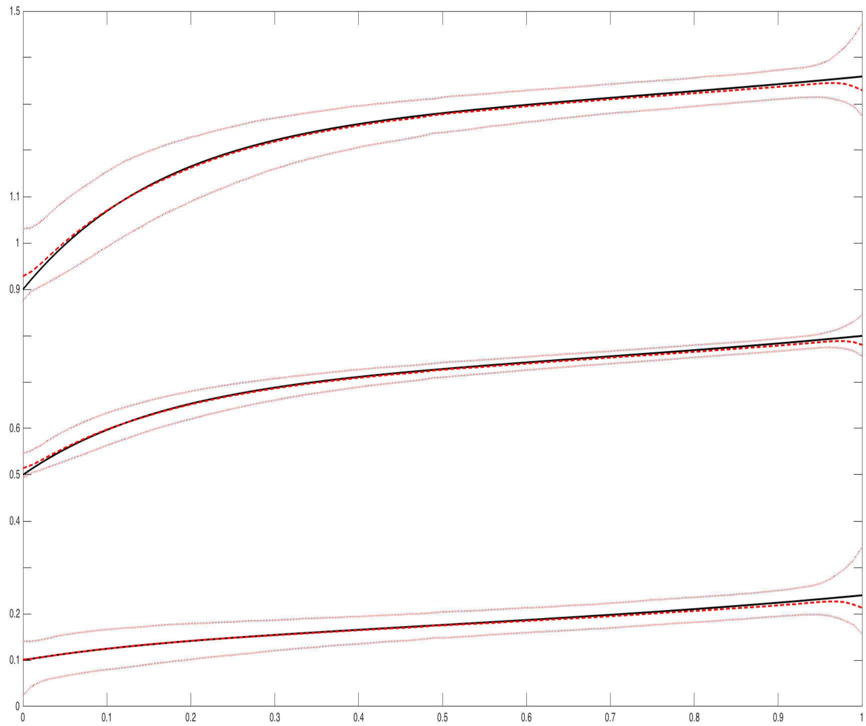


Figure 3: AQR estimation of $V(\alpha|x)$ from 10,000 replications of a sample with $L = 20$, $I = 5$. Top $x_1 = x_2 = 0.2$, middle $x_1 = x_2 = 0.5$ and bottom $x_1 = x_2 = 0.8$. The black line is the true $V(\alpha|x)$, the dashed line is the median of the estimation across the 10,000 replications, the dotted lines are the individual 5% and 95% quantiles of the estimation for each quantile levels.

Figure 3 can be compared with the simulations results of GPV, who considered 200 auctions with 5 bidders without covariate. Their Figure 2 deals with pdf estimation, which has a larger bias in the tails and much larger variance in the center of the distribution, with an individual 5%-95% estimation quantile difference of 0.4 despite a ten time larger number of auctions L. Their Figure 1 reports simulation results for the inverse of the bidding strategy,

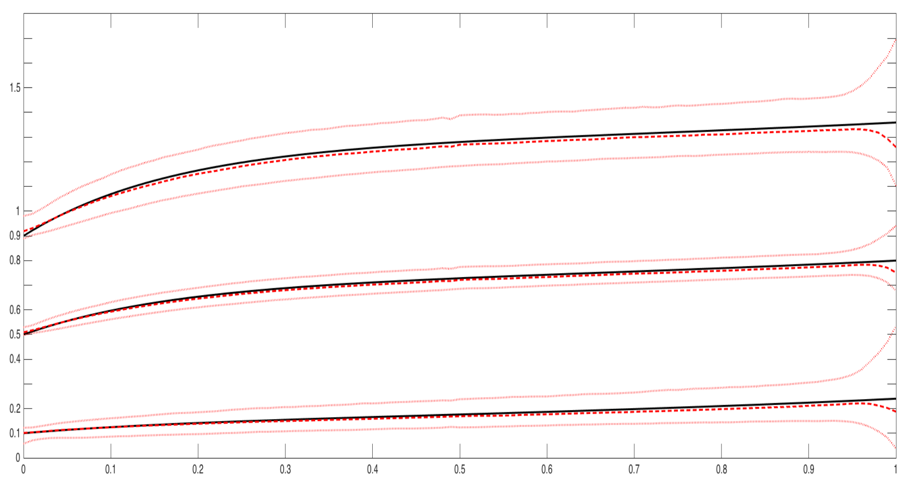
$$v = b + \frac{1}{I-1}\frac{G(b)}{g(b)} = \xi(b).$$

Changing b into $B(\alpha|I)$ gives $V(\alpha|I) = B(\alpha|I) + \alpha B^{(1)}(\alpha|I)/(I-1)$, so that estimating the private value quantile function with the AQR procedure is very similar to estimating $\xi(\cdot)$ as in GPV. Figure 1 in GPV shows a huge increase in the bias and the variance of their estimation of $\xi(b)$ when b is in the upper tail of the bid distribution. In comparison, the behavior of the private value AQR estimator $\widehat{V}(\alpha|x)$ in Figure 3 looks very good for $\alpha = 1$, and, for the average good with $x_1 = x_2 = 0.5$ and other quantile levels, is comparable to the behavior of the GPV estimator of $\xi(b)$ in the middle of the bid distribution.

Figure 4 reports the performance of the private value AQR estimator $\widehat{V}(\alpha|x, I)$ for a lower number of bidders $I = 2$ but still 100 bids. Since

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

having less bidders is expected to increase the contribution of $\widehat{B}^{(1)}(\alpha|x, I)$ to the quantile estimation, so that $\widehat{V}(\alpha|x, I)$ should not performed as well as for a large I because $\widehat{B}^{(1)}(\alpha|x, I)$ converges slower than $\widehat{B}(\alpha|x, I)$.



Indeed Figure 4 suggests that $\widehat{V}(\alpha|x, 2)$ has a bigger variance than $\widehat{V}(\alpha|x, 5)$, especially when $\alpha = 1$. The impact of the number of bidders can also be seen from the behavior of the $5\% - 95\%$ estimator quantile, which increases with α. This is expected as $\widehat{V}(\alpha|x, I)$ does not strongly depend upon the slowly converging $\widehat{B}^{(1)}(\alpha|x, I)$ when α is close to 0, due to the multiplicative factor α in front of $\widehat{B}^{(1)}(\alpha|x, I)$.

6 Conclusion

When signals are normalized to have a uniform marginal distribution, increasing bidding strategies are equal to bid quantile functions. In the case of symmetric independent private value and first-price auction, the bid quantile function is a one to one linear functional of the private value quantile function and its derivative with respect to the quantile level. This implies that a linear specification for the private value quantile function generates a bid one in a similar model. This paper proposes to use sieve interactive versions of the quantile regression Koenker and Bassett (1978) as a model for the private value quantile function. This is expected to provide flexible first-price auction specifications which can be applied for samples including many covariates characterizing the auctioned good,but with a reasonable number of observations. An additional feature of the proposed quantile approach is a good behavior in the extreme quantiles, as necessary to understand formation of winning bids. A simulation experiment illustrates the good behavior of the private value quantile regression model in small samples with one or few hundreds of bids.

The case of possibly asymmetric interdependent value is much more difficult, in particular because it is difficult to obtain simple expressions for the probability of winning the auction and bidding strategies. Assuming strictly increasing strategies allows to identify the joint distribution of the normalized signals, up to possible censoring due to aggressive bidders. As known from Laffont and Vuong (1996), identifying valuation functions is more difficult. Somaini (2015) has shown that the presence of observed bidder specific characteristics affecting bidding strategies can help to recover identification. The paper considers identification of the valuation function of a specific bidder, assuming that the other bidders have strictly increasing strategies and that the bidder of interest uses a best response strategy. The valuation function of interest must depend upon multiplicative functions tying up each bidder signal with his characteristic. This restriction, which covers simple case of auction with resale, allows to identify the valuation function of interest. The bid quantile estimation method developed for private value can be useful to estimate this valuation function.

Many important issues have not be addressed here. The variance performance of the private value quantile estimator can probably be improved by reweighting the observations as suggested in Koenker (2005) for standard quantile regression. Quantile techniques for censored observations can be useful in the presence of a reserve price or when entry decision matters. The Wei and Carroll (2009) procedure to estimate quantile regression with omitted variables can be adapted to cope with unobserved heterogeneity as in Krasnokutskaya (2012).

References

[1] Andrews, D.W.K & Y.J. Whang (1990). Additive interactive regression models: circumventing the curse of dimensionality. *Econometric Theory* **6**, 466–479.

[2] Athey, S. & P.A. Haile (2002). Identification of standard auction models. *Econometrica,* **70**, 2107–2140.

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

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

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

[6] Bajari, P. & A. Hortaçsu (2005). Are structural estimates of auction models reasonable? Evidence from experimental data. *The Journal of Political Economy*, **113**, 703–741.

[7] Belloni, A., V. Chernozhukov, D. Chetverikov & K. Kato (2015). Some new asymptotic theory for least squares series: Pointwise and uniform results. *The Journal of Econometrics*, **141**, 84–108.

[8] 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.

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

[10] Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. Chap. 76 in *Handbook of Econometrics,* Vol. **6B**. Elsevier.

[11] Chen, X. & T.M. Christensen (2014). Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak condition. *The Journal of Econometrics*, **188**, 447–465.

[12] Daubechies, I. (1992). *Ten lectures on wavelets.* SIAM.

[13] Enache, A. & J.P. Florens (2012). Identification and estimation in a third-price auction model. Working Paper, Paris School of Economics.

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

[15] Gimenes, N. (2016). Econometrics of ascending auction by quantile regression. Working paper, University of São Paulo.

[16] Graham, B.S., J. Hahn, A. Poirier & J.L. Powell (2015). A quantile correlated random coefficients panel data model. Working paper.

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

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

[19] 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.

[20] Härdle, W., G. Kerkyacharian, D. Picard & A. Tsybakov (1998). *Wavelets, approximation and statistical applications.* Springer.

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

[22] Haile, P.A. & E. Tamer (2003). Inference with an incomplete model of English auctions. *Journal of Political Economy,* **111**, 1–51.

[23] Henderson, D.J., J.A. List, D.L. Millimet, C.F. Parmeter, M.K. Price (2012). Empirical implementation of nonparametric first-price auction models. *Journal of Econometrics,* **168**, 17-28.

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

[25] Hong, H. & M. Shum (2002). Increasing competition and the winner's curse: evidence from procurement. *Review of Economic Studies*, **69**, 871-898

[26] Horowitz, J.L. & S. Lee (2005). Nonparametric estimation of an additive quantile regression model. *Journal of the American Statistical Association* **100**, 1238–1249.

[27] Hubbard, T.P., T. Li & H.J. Paarsch (2015). Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. *Journal of Econometrics*, **168**, 4–16.

[28] Hunter, D.R. & K. Lange (2000). Quantile regression via an MM algorithm. *Journal of Computational and Graphical Statistics* **9**, 60–77.

[29] Koenker, R. (1984). A note on L-estimates for linear models. *Statistics and Probability Letters* **2**, 323–325.

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

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

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

[33] Krishna, V. (2009). *Auction Theory.* 2nd edition, Academic Press.

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

[35] Laffont, J.-J. & Q. Vuong (1996). Structural Analysis of Auction Data. *The American Economic Review. Papers and Proceedings* **86**, 414–420.

[36] Lee, S., K. Song & Y.-J. Whang (2014). Testing for a general class of functional inequalities. Working Paper, CEMMAP.

[37] Liu, N. & Q. Vuong (2016). Nonparametric test of monotonicity of bidding strategy in first price auctions. Document paper, SHUFE.

[38] Li, T., I. Perrigne & Q. Vuong (2000). Conditionally independent private information in OCS wildcat auctions. *Journal of Econometrics,* **98**, 129–161.

[39] 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.

[40] Liu, N. & Y. Luo (2014). A nonparametric test of exogenous participation in first-price auctions. *Journal of Business and Economic Statistics*, forthcoming.

[41] Luo, Y. & Y. Wan (2016). Integrated-Quantile-Based Estimation for First-Price Auction Models. Working paper, University of Toronto.

[42] Lizzeri, A. & N. Persico (2000). Uniqueness and existence of equilibrium in auctions with a reserve price. *Games and Economic Behavior* **30**, 83–114.

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

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

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

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

[47] McAdams, D. (2007). Uniqueness in symmetric first-price auctions with affiliation. *Journal of Economic Theory* **136**, 144–166.

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

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

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

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

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

[53] RENY, P.J. & S. ZAMIR (2004). On the existence of pure strategy monotone equilibria in asymmetric first-price auctions. *Econometrica* **72**, 1105–1125.

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

[55] RILEY, J. & W. F. SAMUELSON (1981). Optimal auctions. *The American Economic Review,* **71,** 381–392.

[56] SCHUMAKER, L.L. (2007). *Spline functions: basic theory.* Cambridge University Press.

[57] SOMAINI, P. (2015). Identification in auction models with interdependent costs, MIT.

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

[59] ZINCENKO, F. (2013). Nonparametric sieve estimation in first-price auctions with risk-averse bidders. Document paper, UCLA.

Appendix A: Proofs of the identification results

In all this proof section, $g(b|x,I)$ and $G(b|x,I)$ are respectively the conditional p.d.f and cdf 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.

A.1 Proofs of the results in Sections 2 and 4

This subsection groups the proofs of the results of Sections 2 and 4.

Proof of Lemmas 1 and 8. Consider first Lemma 1. If $\alpha \in [0,1] \mapsto V(\alpha|x,I)$ is continuous and strictly increasing, the private value rank A_i in (2.2) is uniquely defined and is equal to $F(V_i|x,I)$. That $v \in [V(0|x,I), V(1|x,I)] \mapsto \sigma(v;x,I)$ is continuous strictly increasing implies that,

$$G(b|x,I) = F\left(\sigma^{-1}(b;x,I)|x,I\right),$$

for all $b \in [B(0|x,I), B(1|x,I)] = [\sigma(V(0|x,I);x,I), \sigma(V(1|x,I);x,I)]$. Hence (2.3) gives $G(B_i|x,I) = F(\sigma^{-1}(B_i;x,I)|x,I) = F(V_i|x,I) = A_i$, which is (i). This implies

$$B_i = G^{-1}(A_i|x,I) = B(A_i|x,I),$$

since the expression of $G(b|x,I)$ implies that $B(\alpha|x,I)$ is uniquely defined. It also follows that $B_i = B(F(V_i|x,I)|x,I)$, which ends the proof of (ii). 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 $B_j = B(A_j|x,I)$ with iid $\mathcal{U}_{[0,1]}$ A_j given (x,I),

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

Consider now Lemma 8 and (iii). Since $s_i(\alpha;z)$ is continuous, smaller than $s(\underline{\alpha}_i(z);z)$ for $\alpha \leq \underline{\alpha}_i(z)$ and strictly increasing for $a \geq \underline{\alpha}_i(z)$, it holds for any b in $(s(\underline{\alpha}_i(z);z), s(1;z)]$,

$$\mathbb{P}(B_i \leq b|z) = \mathbb{P}(s_i(A_i;z) \leq b|z) = \mathbb{P}\left(A_i \leq s^{-1}(b;z)|z\right) = s^{-1}(b;z)$$

since A_i has a uniform distribution over $[0,1]$ and $\underline{\alpha}_i(z) \leq s^{-1}(b;z) \leq 1$. Hence $s_i(\alpha;z) =$

$B_i(\alpha|z)$ over $[\underline{\alpha}_i(z), 1]$. (ii) easily follows. For (i), let i_* be one of the most aggressive bidders, $i_* \in \arg\max_{i=1,\ldots,I} s_i(0;z)$ so that $\underline{\alpha}_{i_*}(z) = 0$ and $s_{i_*}(\alpha;z) = B_{i_*}(\alpha|z)$ for all α in $[0,1]$ since $s_{i_*}(\cdot;z)$ is continuous strictly increasing by Assumption DV-(i). Then (4.2) gives

$$\underline{\alpha}_i(z) = \max\{\alpha \in [0,1], s_i(\alpha;z) \leq B_{i_*}(0|z)\}$$

which by continuity gives $s_i(\underline{\alpha}_i(z); z) = B_{i_*}(0|z)$ and then $B_i(\underline{\alpha}_i(z)|z) = B_{i_*}(0|z)$. Since $B(\cdot|z)$ is strictly increasing at $\underline{\alpha}_i(z)$ by Assumption DV-(i) and (iii),

$$\underline{\alpha}_i(z) = G_i[B_{i_*}(0|z)|z] = G_i\left[\max_{1\leq j\leq I} B_j(0|z)|z\right] = \max_{1\leq j\leq I} G_i[B_j(0|z)|z].$$

For (iv), observe that for $a \geq \underline{\alpha}_i(z)$ and setting $\max_{\varnothing} = -\infty$,

$$\omega(a|\alpha,z) = \mathbb{P}\left(B_1(a|z) > \max_{2\leq j\leq I; A_j \geq \underline{\alpha}_i(z)} B_j(A_j|z)|A_i = \alpha, z\right)$$

which is identified by (ii), the $B_i(\cdot|x,z)$, $i = 1,\ldots,I$ being also identified. □

Proof of Proposition 2. From Maskin and Riley (1984),

$$B_i = \sigma(V_i; x, I) \text{ with the initial condition } V(0|x,I) = \sigma[V(0|x,I); x, I]$$

for a strictly increasing and continuously differentiable $\sigma(\cdot;x,I)$. Proposition 2-(ii) is (2.5) and (i) follows by solving the differential equation (2.5) with the initial condition above as in the proof of Lemma 9. □

Proof of Proposition 3. By (2.6), $B(\alpha|x,I) = (I-1)\int_0^1 u^{I-2}V(\alpha u|x,I)\,du$, so that $B^{(1)}(\alpha|x,I) = (I-1)\int_0^1 u^{I-1}V^{(1)}(\alpha u|x,I)\,du$ which implies the two first statements in (i) about lower and upper bounds for $B^{(1)}(\alpha|x,I)$ and that $B(\cdot|\cdot,I)$ is $(s+1)$th continuously differentiable. That $B(\cdot|x,I)$ is $(s+2)$th continuously differentiable over $(0,1]$ follows from its integral expression (2.6). Observe now that for $p = 1,\ldots,s+2$

$$\frac{\partial^p[\alpha B(\alpha|x,I)]}{\partial\alpha^p} = \alpha B^{(p)}(\alpha|x,I) + pB^{(p-1)}(\alpha|x,I)$$

with, for $p = 1, \ldots, s+1$

$$B^{(p)}(\alpha|x,I) = (I-1)\int_0^1 u^{I-2+p} V^{(p)}(\alpha u|x,I)\,du = \frac{I-1}{\alpha^{I-1+p}}\int_0^\alpha t^{I-2+p} V^{(p)}(t|x,I)\,dt$$

$$\begin{aligned} B^{(p+1)}(\alpha|x,I) &= -\frac{(I-1)(I-1+p)}{\alpha^{I+p}}\int_0^\alpha t^{I-2+p} V^{(p)}(t|x,I)\,dt + \frac{(I-1)V^{(p)}(\alpha|x,I)}{\alpha} \\ &= -\frac{I-1+p}{\alpha} B^{(p)}(\alpha|x,I) + \frac{(I-1)V^{(p)}(\alpha|x,I)}{\alpha}. \end{aligned}$$

Hence, when α goes to 0

$$\begin{aligned} &\alpha B^{(s+2)}(\alpha|x,I) = -(I+s)B^{(s+1)}(0|x,I) + (I-1)V^{(s+1)}(0|x,I) + o(1) \\ &\quad = -(I+s)(I-1)\int_0^1 u^{I+s-1} V^{(s+1)}(0|x,I)\,du + (I-1)V^{(s+1)}(0|x,I) + o(1) \\ &\quad = o(1) \end{aligned}$$

uniformly on x.

For (ii), consider a sequence of $\{\gamma_k(\alpha|I), k \leq K\}$ approximating $V(\alpha|x,I)$ and its derivatives as in Property S. For $\{\beta_k(\alpha|I), k \leq K\}$ as in (2.14)

$$\beta_k^{(p)}(\alpha|I) = (I-1)\int_0^1 u^{I+p-2}\gamma_k^{(p)}(\alpha u|I)\,du, \quad p = 0, \ldots, s+1$$

and

$$\begin{aligned} &\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| B^{(p)}(\alpha|x,I) - \sum_{k=1}^K \beta_k^{(p)}(\alpha|I) P_k(x) \right| \\ &\quad = \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| (I-1)\int_0^1 u^{I+p-2}\left(V^{(p)}(\alpha u|x,I) - \sum_{k=1}^K \gamma_k^{(p)}(\alpha u|I) P_k(x) \right) du \right| \\ &\quad \leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| V^{(p)}(\alpha|x,I) - \sum_{k=1}^K \gamma_k^{(p)}(\alpha|I) P_k(x) \right| \end{aligned}$$

which gives the sieve approximation result for $B(\alpha|x,I)$ in (ii). Now, for $\alpha B^{(1)}(\alpha|x,I)$, observe that $\alpha B^{(1)}(\alpha|x,I) = (I-1)[V(\alpha|x,I) - B(\alpha|x,I)]$ and

$$\begin{aligned} \alpha\beta_k^{(1)}(\alpha|I) &= \alpha \times \left(-\frac{(I-1)^2}{\alpha^I}\int_0^1 t^{I-2}\gamma_k(t|I)\,dt + \frac{I-1}{\alpha}\gamma_k(\alpha|I) \right) \\ &= (I-1)[\gamma_k(\alpha|I) - \beta_k(\alpha|I)]. \end{aligned}$$

It follows

$$
\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{\partial^p\left[\alpha B^{(1)}(\alpha|x,I)\right]}{\partial\alpha^p}-\sum_{k=1}^{K}\frac{\partial^p\left[\alpha\beta_k^{(1)}(\alpha|I)\right]}{\partial\alpha^p}P_k(x)\right| \\
&\leq (I-1)\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|V^{(p)}(\alpha|x,I)-\sum_{k=1}^{K}\gamma_k^{(p)}(\alpha|I)P_k(x)\right| \\
&+(I-1)\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|B^{(p)}(\alpha|x,I)-\sum_{k=1}^{K}\beta_k^{(p)}(\alpha|I)P_k(x)\right| \\
&\leq 2(I-1)\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|V^{(p)}(\alpha|x,I)-\sum_{k=1}^{K}\gamma_k^{(p)}(\alpha|I)P_k(x)\right|
\end{aligned}
$$

which gives the approximation result for $\alpha B^{(1)}(\alpha|x,I)$ in (iii). $\square$

Proof of Lemma 9. The next Lemma is a preparatory one, which first extends the Lizerri and Persico result (2000) on terminal bids to the case of an arbitrary number of bidders, a result assumed in Assumption DV but established here when the bids satisfy a best response condition. It also studies the existence of $\Omega(\alpha|z)$ when α decreases to $\underline{\alpha}_1(z)$.

Lemma A.1 *Suppose Assumption DV-(i) holds. Then*

i. If all the bidders have a valuation U_i and the strategies $s_i(\cdot;z)$ satisfy a best response condition as (4.3), $s_1(1;z)=\cdots s_I(1;z)$ and

$$B_1(1|z)=\cdots=B_I(1|z) \text{ for each } z \text{ in } \mathcal{Z},$$

ii. It holds $B_1[\underline{\alpha}_1(z)|z]=\cdots=B_I[\underline{\alpha}_I(z)|z]$ for each z in $\mathcal{Z}$.

iii. If Assumption DV-(ii) holds, $\Omega(\alpha|z)$ and $U_1(\alpha|z)$ are well defined for α in $(\underline{\alpha}_1(z),1]$ with $\lim_{\alpha\downarrow\underline{\alpha}_1(z)}\Omega(\alpha|z)=0$, $\lim_{\alpha\downarrow\underline{\alpha}_1(z)}U_1(\alpha|z)=U_1(\underline{\alpha}_1(z)|z)$ and

$$\lim_{\alpha\downarrow\underline{\alpha}_1(z)}\frac{\Omega(\alpha|z)}{\alpha-\underline{\alpha}_1(z)}\in(0,\infty).$$

Proof of Lemma A.1. For the sake of brevity, remove z from the various functions. For (i) suppose for instance $s_I(1)\geq s_{I-1}(1)\geq\cdots\geq s_1(1)$. Assume $s_I(1)>s_{I-1}(1)$. Then $\omega_I(1|1)=1$ but, for any $\epsilon>0$ such $1-\epsilon\geq\underline{\alpha}_I$ $s_I(1-\epsilon)>s_{I-1}(1)$, $\omega_I(1-\epsilon|1)=1$ with $\overline{U}_I(1-\epsilon|1)<\overline{U}_I(1|1)$ so that bidding $s_I(1-\epsilon)$ instead of $s_I(1)$ would give a better expected profit, contradicting the fact that $s_I(1)$ is a best response. Hence $s_I(1)=s_{I-1}(1)$. But, arguing as above, if $s_I(1)=s_{I-1}(1)>s_{I-2}(1)$, slightly decreasing the two bids $s_I(1)=s_{I-1}(1)$ simultaneously would increase the expected profit of the two top tied bidders, another

contradiction. Hence $s_I(1) = s_{I-1}(1) = s_{I-2}(1)$. Iterating gives $s_1(1) = \cdots = s_I(1)$ and then $B_1(1) = \cdots = B_I(1)$ by Lemma 8-(iii). (ii) follows from $\underline{\alpha}_i = G_i(B_{i_*}(0))$ where i_* is an aggressive bidder, $B_{i_*}(0) = \max_{1\leq i\leq I} B_i(0)$.

Consider now (iii). Recall $G_{1i}(a) = G_i(B_1(a))$ is continuously differentiable over $[\underline{\alpha}_1, 1]$ with $\frac{\partial}{\partial\alpha} G_{1i}(\alpha) > 0$ by Assumption DV-(ii) and

$$\begin{aligned}\omega(a|\alpha) &= \int \mathbb{I}\left(B_1(a) \geq \max_{2\leq i\leq I} B_i(\alpha_i)\right) c(\alpha_2, \ldots, \alpha_I|\alpha) \prod_{i=2}^{I} d\alpha_i \\ &= \int_0^{G_{12}(a)} \cdots \int_0^{G_{1I}(a)} c(\alpha_2, \ldots, \alpha_I|\alpha) \prod_{i=2}^{I} d\alpha_i\end{aligned}$$

Hence $\Omega_1(\alpha)$ is well defined for α in $(\underline{\alpha}_1, 1]$ since $c(\cdot|\cdot)$ is positive. The limit for $\Omega(\alpha)$ when α decreases to $\underline{\alpha}_1$ follows from the limit of $\Omega(\alpha)/(\alpha - \underline{\alpha}_1)$. For the limit of $U_1(\alpha)$, recall

$$\begin{aligned}\overline{U}_1(a|\alpha) &= \int_0^{G_{12}(a)} \cdots \int_0^{G_{1I}(a)} U_1(\alpha, \alpha_2, \ldots, \alpha_I) c(\alpha_2, \ldots, \alpha_I|\alpha) \prod_{i=2}^{I} d\alpha_i \\ &= \int_0^{G_{12}(a)} \cdots \int_0^{G_{1I}(a)} U_1(\alpha, \alpha_{-1}) c(\alpha_{-1}|\alpha) d\alpha_{-1}\end{aligned}$$

where $\alpha_{-1} = (\alpha_2, \ldots, \alpha_I)$ and $d\alpha_{-1} = \prod_{i=2}^{I} d\alpha_i$. Define similarly

$$\begin{aligned}\alpha_{-1,j} &= (\alpha_2, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_I), \\ \alpha_{-1,2,j} &= (\alpha_3, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_I), \\ d\alpha_{-1,j} &= \prod_{i=2, i\neq j}^{I} d\alpha_i, \quad d\alpha_{-1,2,j} = \prod_{i=3, i\neq j}^{I} d\alpha_i, \\ U_1(\alpha, G_j B_1(a), \alpha_{-1,j}) &= U_1(\alpha, \alpha_2, \ldots, \alpha_{j-1}, G_j B_1(a), \alpha_{j+1}, \ldots, \alpha_I), \\ c(G_j B_1(a), \alpha_{-1,j}|\alpha) &= c(\alpha_2, \ldots, \alpha_{j-1}, G_j B_1(a), \alpha_{j+1}, \ldots, \alpha_I|\alpha), \\ c(G_2 B_1(t|z), G_j B_1(a), \alpha_{-1,2,j}|\alpha) &= c(G_2 B_1(t|z), \alpha_3, \ldots, \alpha_{j-1}, G_j B_1(a), \alpha_{j+1}, \ldots, \alpha_I|\alpha).\end{aligned}$$

This gives for $a \geq \underline{\alpha}_1$

$$\begin{aligned}\frac{\partial \overline{U}_1(a|\alpha)}{\partial a} &= \sum_{j=2}^{I} \frac{\partial G_{1j}(a)}{\partial a} \int \prod_{2\leq k\neq j\leq I} \mathbb{I}[\alpha_k \leq G_{1k}(a)] \\ &\qquad \times U_1(\alpha, G_{1j}(a), \alpha_{-1,j}) c(G_{1j}(a), \alpha_{-1,j}|\alpha) d\alpha_{-1,j}, \\ \frac{\partial \omega(a|\alpha)}{\partial a} &= \sum_{j=2}^{I} \frac{\partial G_{1j}(a)}{\partial a} \int \prod_{2\leq k\neq j\leq I} \mathbb{I}[\alpha_k \leq G_{1k}(a)] c(G_{1j}(a), \alpha_{-1,j}|\alpha) d\alpha_{-1,j}.\end{aligned}$$

Let $n_1 - 1 \geq 0$ be the number of aggressive bidders $j \geq 2$, i.e. the number of j in $\{2, \ldots, I\}$ such that $B_j(0) = \min_{1\leq} B_i(0)$. For some positive constants C,

$$
\begin{aligned}
\frac{\partial \overline{U}_1(a|\alpha)}{\partial a} &= C_U (a - \underline{\alpha}_1)^{n_1 - 1} (1 + o(1)), \\
\omega(a|\alpha) &= C_\omega (a - \underline{\alpha}_1)^{n_1} (1 + o(1)) \\
\frac{\partial \omega(a|\alpha)}{\partial a} &= C'_\omega (a - \underline{\alpha}_1)^{n_1 - 1} (1 + o(1))
\end{aligned}
$$

when a and α decrease to $\underline{\alpha}_1$. This gives the limits stated in the Lemma for $U_1(\alpha|z)$ and $\Omega_1(\alpha|z)$. $\square$

Proof of Lemma 9. Remove the dependence upon z. The existence of $U_1(\cdot)$ and $\Omega(\cdot)$ over $[\underline{\alpha}_1, 1]$ with $\underline{\alpha}_1 < 1$ follows from Lemma A.1. Then (**??**) gives (4.5)

$$
U_1(\alpha) = B_1(\alpha) + B_1^{(1)}(\alpha) \Omega(\alpha), \quad \alpha \in [\underline{\alpha}_1, 1]
$$

with $B_1(\underline{\alpha}_1) = U_1(\underline{\alpha}_1)$ since $\Omega(\underline{\alpha}_1) = 0$ as established in Lemma A.1 and by Assumption DV-(ii) that ensures that $B_1^{(1)}(\underline{\alpha}_1)$ is finite. Hence Lemma 9-(i) holds. We now solve (4.5) to establish (4.6). Define

$$
\Psi(\alpha) = \exp\left(\int_\alpha^1 \frac{dt}{\Omega(t)} \right)
$$

which is such that $\Psi^{(1)}(\alpha) = -\Psi(\alpha) / \Omega(\alpha)$ and $\Psi(\alpha) \sim C' (\alpha - \underline{\alpha}_1)^{-1/C}$ when α goes to $\underline{\alpha}_1$ by Lemma A.1, where $C > 0$ is such that $\Omega(\alpha) \sim C(\alpha - \underline{\alpha}_1)$. Suppose that $B_1(\alpha) = \Psi(\alpha) b(\alpha)$ is a solution of (4.5). Since

$$
B_1^{(1)}(\alpha) = \Psi^{(1)}(\alpha) b(\alpha) + \Psi(\alpha) b^{(1)}(\alpha) = -\Psi(\alpha) b(\alpha) / \Omega(\alpha) + \Psi(\alpha) b^{(1)}(\alpha)
$$

$b(\cdot)$ must be such

$$
\Omega(\alpha) \Psi(\alpha) b^{(1)}(\alpha) = U_1(\alpha), \text{ so that } b(\alpha) = C' + \int_{\underline{\alpha}_1}^\alpha \frac{U_1(t)}{\Omega(t) \Psi(t)} dt
$$

and then

$$
B_1(\alpha) = \Psi(\alpha) \times \left(C' + \int_{\underline{\alpha}_1}^\alpha \frac{U_1(t)}{\Omega(t) \Psi(t)} dt \right).
$$

To show that $C' = 0$ observe that $\lim_{\alpha \downarrow \underline{\alpha}_1} \Psi(\alpha) = +\infty$, so that $C' \neq 0$ would give an infinite $B_1(\underline{\alpha}_1)$. Now, by Lemma A.1-(iii),

$$\begin{aligned}\lim_{\alpha \downarrow \underline{\alpha}_1} \Psi(\alpha) \int_{\underline{\alpha}_i}^{\alpha} \frac{U_1(t)}{\Omega_1(t)\Psi(t)} dt &= \lim_{\alpha \downarrow \underline{\alpha}_1} C'(\alpha - \underline{\alpha}_1)^{-1/C} \int_{\underline{\alpha}_1}^{\alpha} \frac{U_1(\underline{\alpha}_1)\, dt}{C'C(t - \underline{\alpha}_1)^{1-1/C}} \\ &= \lim_{\alpha \downarrow \underline{\alpha}_1} \left\{ (\alpha - \underline{\alpha}_1)^{-1/C} U_1(\underline{\alpha}_1) \left. (t - \underline{\alpha}_1)^{1/C} \right]_{\underline{\alpha}_1}^{\alpha} \right\} = U_1(\underline{\alpha}_1)\end{aligned}$$

which end the proof of the Lemma. $\square$

Proof of Theorem 10 Only the proof of (ii) is detailed, the proof of (i) being simpler. For the sake of notation, assume $I = 3$, the case of a larger number of bidders being similar. Recall that $G_{1i}(\alpha|z) = G_i[B_1(\alpha|z)|z]$. The proof follows by differentiating (4.13) with respect to α, z_2 and z_3 as possible under Assumption MSM-(i,ii,iii). In what follows

$$\Psi(\alpha|z) = \left. \frac{\partial \overline{U}_1(a|\alpha, z)}{\partial a} \right|_{a=\alpha}$$

is the identified expression in (4.13). Recall also

$$\begin{aligned} G_{ij}(\alpha|z) = G_j[B_i(\alpha|z)|z], \quad g_{ij}(\alpha|z) = \frac{\partial G_j B_i(\alpha|z)}{\partial \alpha}, \\ c_{ij}(t_k|\alpha, z) = c(t_k|A_i = \alpha, A_j = G_{ij}(\alpha|z)). \end{aligned}$$

Recall also that $\Phi(\cdot)$ is identified over $\mathcal{Z} = (0, \overline{z}]^3$ by (4.14), $\gamma_i(0) = 1$ for $i = 1, 2, 3$, and that $z_i \gamma_i[G_{1i}(\alpha|z)]$ belongs to $(0, \overline{z}]$ as the $\gamma_i(\cdot)$'s are valued in $(0, 1]$. Differentiating $\Psi(\alpha|z)$

with respect to α in $[0,1]$ gives

$$\begin{aligned}
&\left\{g_{12}(\alpha|z)\int_0^{G_{13}(\alpha|z)}\Phi_{z_1}\left[z_1\gamma_1(\alpha), z_2\gamma_2\left[G_{12}(\alpha|z)\right], z_3\gamma_3(t_3)\right]c_{12}(t_3|\alpha,z)\,dt_3\right.\\
&\quad\left.+g_{13}(\alpha|z)\int_0^{G_{12}(\alpha|z)}\Phi_{z_1}\left[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3\left[G_{13}(\alpha|z)\right]\right]c_{13}(t_2|\alpha,z)\,dt_2\right\}z_1\gamma_1^{(1)}(\alpha)\\
&\quad+g_{12}(\alpha|z)\int_0^{G_{13}(\alpha|z)}\Phi_{z_2}\left[z_1\gamma_1(\alpha), z_2\gamma_2\left[G_{12}(\alpha|z)\right], z_3\gamma_3(t_3)\right]c_{12}(t_3|\alpha,z)\,dt_3\\
&\qquad\times z_2\frac{\partial\left\{\gamma_2\left[G_{12}(\alpha|z)\right]\right\}}{\partial\alpha}\\
&\quad+g_{13}(\alpha|z)\int_0^{G_{12}(\alpha|z)}\Phi_{z_3}\left[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3\left[G_{13}(\alpha|z)\right]\right]c_{13}(t_2|\alpha,z)\,dt_2\\
&\qquad\times z_3\frac{\partial\left\{\gamma_3\left[G_{13}(\alpha|z)\right]\right\}}{\partial\alpha}\\
&\qquad=\mathbf{e}_\alpha[\gamma](\alpha|z)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{e}_\alpha[\gamma](\alpha|z) &= \frac{\partial\Psi(\alpha|z)}{\partial\alpha}\\
&-\frac{\partial g_{12}(\alpha|z)}{\partial\alpha}\int_0^{G_{13}(\alpha|z)}\Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2\left[G_{12}(\alpha|z)\right], z_3\gamma_3(t_3)\right]c_{12}(t_3|\alpha,z)\,dt_3\\
&-\frac{\partial g_{13}(\alpha|z)}{\partial\alpha}\int_0^{G_{12}(\alpha|z)}\Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3\left[G_{13}(\alpha|z)\right]\right]c_{13}(t_2|\alpha,z)\,dt_2\\
&-2g_{12}(\alpha|z)\,g_{13}(\alpha|z)\,\Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2\left[G_{12}(\alpha|z)\right], z_3\left[G_{13}(\alpha|z)\right]\right]c\left[G_{12}(\alpha|z), G_{13}(\alpha|z)|\alpha\right]\\
&-g_{12}(\alpha|z)\int_0^{G_{13}(\alpha|z)}\Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2\left[G_{12}(\alpha|z)\right], z_3\gamma_3(t_3)\right]\frac{\partial c_{12}(t_3|\alpha,z)}{\partial\alpha}dt_3\\
&-g_{13}(\alpha|z)\int_0^{G_{12}(\alpha|z)}\Phi\left[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3\left[G_{13}(\alpha|z)\right]\right]\frac{\partial c_{13}(t_2|\alpha,z)}{\partial\alpha}dt_2.
\end{aligned}$$

Differentiating $\Psi\left(\alpha|z\right)$ with respect to z_2 and z_3 give, respectively,

$$
\begin{aligned}
&\frac{\partial G_{12}\left(\alpha|z\right)}{\partial z_2}\int_0^{G_{13}(\alpha|z)}\Phi_{z_2}\left[z_1\gamma_1\left(\alpha\right),z_2\gamma_2\left[G_{12}\left(\alpha|z\right)\right],z_3\gamma_3\left(t_3\right)\right]c_{12}\left(t_3|\alpha,z\right)dt_3\\
&\qquad\times z_2\frac{\partial\left\{\gamma_2\left[G_{12}\left(\alpha|z\right)\right]\right\}}{\partial\alpha}\\
&+\frac{\partial G_{13}\left(\alpha|z\right)}{\partial z_2}\int_0^{G_{12}(\alpha|z)}\Phi_{z_3}\left[z_1\gamma_1\left(\alpha\right),z_2\gamma_2\left(t_2\right),z_3\gamma_3\left[G_{13}\left(\alpha|z\right)\right]\right]c_{13}\left(t_2|\alpha,z\right)dt_2\\
&\qquad\times z_3\frac{\partial\left\{\gamma_3\left[G_{13}\left(\alpha|z\right)\right]\right\}}{\partial\alpha}\\
&\qquad=\mathbf{e}_{z_2}\left[\gamma\right]\left(\alpha|z\right),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial G_{12}\left(\alpha|z\right)}{\partial z_3}\int_0^{G_{13}(\alpha|z)}\Phi_{z_2}\left[z_1\gamma_1\left(\alpha\right),z_2\gamma_2\left[G_{12}\left(\alpha|z\right)\right],z_3\gamma_3\left(t_3\right)\right]c_{12}\left(t_3|\alpha,z\right)dt_3\\
&\qquad\times z_2\frac{\partial\left\{\gamma_2\left[G_{12}\left(\alpha|z\right)\right]\right\}}{\partial\alpha}\\
&+\frac{\partial G_{13}\left(\alpha|z\right)}{\partial z_3}\int_0^{G_{12}(\alpha|z)}\Phi_{z_3}\left[z_1\gamma_1\left(\alpha\right),z_2\gamma_2\left(t_2\right),z_3\gamma_3\left[G_{13}\left(\alpha|z\right)\right]\right]c_{13}\left(t_2|\alpha,z\right)dt_2\\
&\qquad\times z_3\frac{\partial\left\{\gamma_3\left[G_{13}\left(\alpha|z\right)\right]\right\}}{\partial\alpha}\\
&\qquad=\mathbf{e}_{z_3}\left[\gamma\right]\left(\alpha|z\right)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{e}_{z_2}[\gamma](\alpha|z) = {}& \frac{\partial \Psi(\alpha|z)}{\partial z_2} \\
& - g_{12}(\alpha|z)\,\gamma_2[G_{12}(\alpha|z)] \int_0^{G_{13}(\alpha|z)} \Phi_{z_2}[z_1\gamma_1(\alpha), z_2\gamma_2[G_{12}(\alpha|z)], z_3\gamma_3(t_3)]\, c_{12}(t_3|\alpha,z)\, dt_3 \\
& - \frac{\partial g_{12}(\alpha|z)}{\partial z_2} \int_0^{G_{13}(\alpha|z)} \Phi[z_1\gamma_1(\alpha), z_2\gamma_2[G_{12}(\alpha|z)], z_3\gamma_3(t_3)]\, c_{12}(t_3|\alpha,z)\, dt_3 \\
& - \frac{\partial g_{13}(\alpha|z)}{\partial z_2} \int_0^{G_{12}(\alpha|z)} \Phi[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3[G_{13}(\alpha|z)]]\, c_{13}(t_2|\alpha,z)\, dt_2 \\
& - \left(g_{12}(\alpha|z) \frac{\partial G_{13}(\alpha|z)}{\partial z_2} + g_{13}(\alpha|z) \frac{\partial G_{12}(\alpha|z)}{\partial z_2} \right) \\
& \qquad \times \Phi[z_1\gamma_1(\alpha), z_2\gamma_2[G_{12}(\alpha|z)], z_3[G_{13}(\alpha|z)]]\, c[G_{12}(\alpha|z), G_{13}(\alpha|z)|\alpha] \\
& - g_{12}(\alpha|z) \int_0^{G_{13}(\alpha|z)} \Phi[z_1\gamma_1(\alpha), z_2\gamma_2[G_{12}(\alpha|z)], z_3\gamma_3(t_3)] \frac{\partial c_{12}(t_3|\alpha,z)}{\partial z_2} dt_3 \\
& - g_{13}(\alpha|z) \int_0^{G_{12}(\alpha|z)} \Phi[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3[G_{13}(\alpha|z)]] \frac{\partial c_{13}(t_2|\alpha,z)}{\partial z_2} dt_2
\end{aligned}
$$

$\mathbf{e}_{z_3}[\gamma]$ having a similar expression.

These three integro-differential equations can be stacked to obtain an expression similar to (4.12). Fix now a z in $\mathcal{Z}$. Let $\mathbf{e}[\gamma] = [\mathbf{e}_\alpha[\gamma], \mathbf{e}_{z_2}[\gamma], \mathbf{e}_{z_3}[\gamma]]'$. Consider the 3×3 matrices $\mathbf{D}[\gamma](\alpha|z)$ and $\mathbf{G}(\alpha|z)$ with

$$
\begin{aligned}
\mathbf{D}_{\alpha\alpha}[\gamma](\alpha|z) &= z_1 g_{12}(\alpha|z) \int_0^{G_{13}(\alpha|z)} \Phi_{z_1}[z_1\gamma_1(\alpha), z_2\gamma_2[G_{12}(\alpha|z)], z_3\gamma_3(t_3)]\, c_{12}(t_3|\alpha,z)\, dt_3 \\
&\quad + z_1 g_{13}(\alpha|z) \int_0^{G_{12}(\alpha|z)} \Phi_{z_1}[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3[G_{13}(\alpha|z)]]\, c_{13}(t_2|\alpha,z)\, dt_2, \\
\mathbf{D}_{z_2z_2}[\gamma](\alpha|z) &= z_2 \int_0^{G_{13}(\alpha|z)} \Phi_{z_2}[z_1\gamma_1(\alpha), z_2\gamma_2[G_{12}(\alpha|z)], z_3\gamma_3(t_3)]\, c_{12}(t_3|\alpha,z)\, dt_3, \\
\mathbf{D}_{z_3z_3}[\gamma](\alpha|z) &= z_3 \int_0^{G_{12}(\alpha|z)} \Phi_{z_3}[z_1\gamma_1(\alpha), z_2\gamma_2(t_2), z_3\gamma_3[G_{13}(\alpha|z)]]\, c_{13}(t_2|\alpha,z)\, dt_2,
\end{aligned}
$$

$\mathbf{D}[\gamma](\alpha|z)$ being diagonal and

$$
\mathbf{G}(\alpha|z) = \begin{bmatrix} 1 & g_{12}(\alpha|z) & g_{13}(\alpha|z) \\ 0 & \frac{\partial G_{12}(\alpha|z)}{\partial z_2} & \frac{\partial G_{13}(\alpha|z)}{\partial z_2} \\ 0 & \frac{\partial G_{12}(\alpha|z)}{\partial z_3} & \frac{\partial G_{13}(\alpha|z)}{\partial z_3} \end{bmatrix}.
$$

Then the three last equations above write

$$\mathbf{G}(\alpha|z)\,\mathbf{D}[\gamma](\alpha|z)\begin{bmatrix}\gamma_1^{(1)}(\alpha)\\ \frac{\partial\{\gamma_2[G_{12}(\alpha|z)]\}}{\partial\alpha}\\ \frac{\partial\{\gamma_3[G_{13}(\alpha|z)]\}}{\partial\alpha}\end{bmatrix}=\mathbf{e}[\gamma](\alpha|z) \tag{A.1.1}$$

so that for all α in $[0,1]$

$$\begin{bmatrix}\gamma_1(\alpha)\\ \gamma_2[G_{12}(\alpha|z)]\\ \gamma_3[G_{13}(\alpha|z)]\end{bmatrix}=\int_0^\alpha\{\mathbf{D}[\gamma](t|z)\}^{-1}\mathbf{G}(t|z)^{-1}\mathbf{e}[\gamma](t|z)\,dt=\mathbf{E}_\gamma[\gamma](\alpha|z)$$

assuming $\Phi_{z_2}(\cdot)$ and $\Phi_{z_3}(\cdot)$ do not vanish to ensure existence of the inverse matrix $\{\mathbf{D}[\gamma](\cdot|z)\}^{-1}$ and ignoring that $\mathbf{G}(1|z)$ is not full rank as $G_{ij}(1|z)=1$ for all z. These issues are addressed using a regularized version of the operator $\mathbf{E}_\gamma[\gamma]$. Let $\tau(\cdot)$ be a continuously differentiable function such that $\tau(x)=x$ when x belongs to $[-1,1]$ and $\tau(x)=0$ when $|x|>2$ with $\sup_{x\in\mathbb{R}}\left|\tau^{(1)}(x)\right|<\infty$. For $\epsilon>0$, set

$$\tau_\epsilon(x)=\frac{\tau(\epsilon x)}{\epsilon}$$

which is equal to x for $|x|\leq 1/\epsilon$ and 0 for $|x|>2/\epsilon$, with $\sup_{x\in\mathbb{R}}|\tau_\epsilon(x)|\leq C/\epsilon$ and $\sup_{x\in\mathbb{R}}\left|\tau_\epsilon^{(1)}(x)\right|\leq C$. Let $T(\cdot)\in[0,1]$ be a continuously differentiable function such that $T(x)=x$ when x is in $[0,1]$ and $T(x)=0$ if $x\leq -1$ or $x\geq 2$. For a matrix or a vector $[x_{ij}]$, set $\tau_\epsilon([x_{ij}])=[\tau_\epsilon(x_{ij})]$ and $T([x_{ij}])=T[\tau_\epsilon(x_{ij})]$. The regularized $\mathbf{E}_\gamma[\cdot]$ is

$$\mathbf{E}_\gamma^\epsilon[\zeta](\alpha|z)=T\left\{\int_0^\alpha\tau_\epsilon\left[\{\mathbf{D}[\zeta](t|z)\}^{-1}\mathbf{G}(t|z)^{-1}\mathbf{e}[\zeta](t|z)\right]dt\right\}.$$

The role of the transformation $T(\cdot)$ is to ensure that the entries of $\mathbf{E}_\gamma^\epsilon[\zeta](\alpha|z)$ are in $[0,1]$. The transformation $\tau_\epsilon(\cdot)$ forces the non identified $\gamma_i(\cdot)$ (such that i is not in $\mathcal{I}$) to be equal to 0 and allows for matrices $\mathbf{D}[\gamma](t|z)$ or $\mathbf{G}(t|z)$ which are not full rank everywhere as permitted by Assumption MSM-(i,ii). Note however that, for ϵ small enough

$$\tau_\epsilon\left[\{\mathbf{D}[\gamma](t|z)\}^{-1}\mathbf{G}(t|z)^{-1}\mathbf{e}[\gamma](t|z)\right]=\{\mathbf{D}[\gamma](t|z)\}^{-1}\mathbf{G}(t|z)^{-1}\mathbf{e}[\gamma](t|z)$$

for all t in $[0,1]$ by (A.1.1) due to differentiability of the $\gamma_i[G_{1i}(\alpha|z)]$ over $[0,1]$. Consider now such ϵ. It follows

$$\begin{bmatrix}\gamma_1(\alpha)\\ \gamma_2[G_{12}(\alpha|z)]\\ \gamma_3[G_{13}(\alpha|z)]\end{bmatrix}=\mathbf{E}_\gamma^\epsilon[\gamma](\alpha|z)\ \text{ for all } \alpha \text{ in } [0,1], \tag{A.1.2}$$

setting from now on $\gamma_i(\cdot)=0$ when i is not in $\not{\mathcal{I}}$.

The rest of proof shows that the $\gamma_i(\cdot)$ are identified through a fixed point argument. Let $\gamma^z(\alpha) = [\gamma_1(\alpha), \gamma_2[G_{12}(\alpha|z_0)], \gamma_3[G_{13}(\alpha|z_0)]]'$. Suppose that there exists another differentiable $\gamma_a^z(\cdot)$ with $\gamma_a^z(0) = [0,0,0]$ satisfying (A.1.2) so that

$$\gamma^z(\alpha) - \gamma_a^z(\alpha) = \mathbf{E}_\gamma^\epsilon[\gamma](\alpha|z) - \mathbf{E}_\gamma^\epsilon[\gamma_a](\alpha|z) \tag{A.1.3}$$

for all α in $[0,1]$. It follows by definition of $\mathbf{E}_\gamma^\epsilon$ that

$$\|\gamma^z(\alpha) - \gamma_a^z(\alpha)\| = \left\|\mathbf{E}_\gamma^\epsilon[\gamma](\alpha|z) - \mathbf{E}_\gamma^\epsilon[\gamma_a](\alpha|z)\right\| \leq \mathbf{C}\alpha$$

where the constant $\mathbf{C}$ depends on $\mathbf{E}_\gamma^\epsilon$, $\gamma^z(\cdot)$ and $\gamma_a^z(\cdot)$ but does not depend upon α. Substituting this bound in the fixed point condition (A.1.3) gives, after k iterations and for the same constant $\mathbf{C}$,

$$\|\gamma^z(\alpha) - \gamma_a^z(\alpha)\| \leq \frac{(\mathbf{C}\alpha)^k}{k!} \leq \frac{\mathbf{C}^k}{k!} \underset{k\to\infty}{\to} 0$$

which gives $\gamma^z(\alpha) = \gamma_a^z(\alpha)$ over $[0,1]$, showing that the $\gamma_i(\cdot)$'s are identified for i in $\mathcal{I}$.

Appendix B: Proofs of intermediary results

B.1 Notations and objective function smoothness

We start with additional notations used all along the proof section and some preliminary lemmas which are established in Appendix B. In what follows

$$P(x) = \begin{cases} [1, x']' & \text{in the AQR case } (K_L = d+1) \\ [P_1(x), \ldots, P_{K_L}(x)]' & \text{in the ASQR case} \end{cases}$$

allowing an unified treatment of the two estimators, although the proof focus is on the more difficult ASQR case. Recall that $\|P(x)\| = \left(P(x)' P(x)\right)^{1/2}$ is the standard Euclidean norm and that, under Assumptions R-(i) and H-(ii),

$$\max_{x \in \mathcal{X}} \|P(x)\| = O\left(K_L^{1/2}\right) = O\left(h^{-d_{\mathcal{M}}/2}\right), \quad \max_{(x,t) \in \mathcal{X} \times [-1,1]} \|P(x,t)\| = O\left(h^{-d_{\mathcal{M}}/2}\right),$$

with $d_{\mathcal{M}} = 0$ in the AQR case. Recall that

$$P(x, ht) = \pi(ht) \otimes P(x), \quad \pi(ht)' = \left[1, ht, \ldots, \frac{(ht)^{s+1}}{(s+1)!}\right]$$

so that the "design" matrix $\mathbb{E}\left[P(x_\ell, ht) P(x_\ell, ht)'\right]$ degenerates asymptotically. To avoid this, consider the change of parameters $\mathsf{b} = Hb$ with $H = \operatorname{Diag}(\pi(h)) \otimes \operatorname{Id}_{K_L}$,

$$\mathsf{b} = \Bigg[\underbrace{\beta_{0,1}, \ldots, \beta_{0,K_L}}_{\mathsf{b}_0' = \beta_0'}, \underbrace{h\beta_{1,1}, \ldots, h\beta_{1,K_L}}_{\mathsf{b}_1' = h\beta_1'}, \ldots, \underbrace{h^{s+1}\beta_{s+1,1}, \ldots, h^{s+1}\beta_{s+1,K_L}}_{\mathsf{b}_{s+1}' = h^{s+1}\beta_{s+1}}\Bigg] \tag{B.1}$$

so that $P(x_\ell, ht)' \beta = P(x_\ell, t)' \mathsf{b}$. Define accordingly

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

Note that $\mathsf{b} \to \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{a+ht}\left(B_{i\ell} - P\left(x_\ell, t\right)' \mathsf{b}\right) K\left(t\right) dt$ is convex as an integral of convex functions. It follows that $\widehat{\mathsf{R}}\left(\mathsf{b}; \alpha, I\right)$ and $\overline{\mathsf{R}}\left(\mathsf{b}; \alpha, I\right)$ have minimizers,

$$\begin{aligned}
\widehat{\mathsf{b}}\left(\alpha|I\right) &= \arg\min_{\mathsf{b}} \widehat{\mathsf{R}}\left(\mathsf{b}; \alpha, I\right) = H\widehat{\beta}\left(\alpha|I\right), \\
\overline{\mathsf{b}}\left(\alpha|I\right) &= \arg\min_{\mathsf{b}} \overline{\mathsf{R}}\left(\mathsf{b}; \alpha, I\right),
\end{aligned}$$

which uniqueness will be established in the next section. Set $\overline{b}\left(\alpha|I\right) = H^{-1}\overline{\mathsf{b}}\left(\alpha|I\right)$ recalling $\overline{b}\left(\alpha|I\right) = \left[\overline{\beta}_0\left(\alpha|I\right)', \ldots, \overline{\beta}_{s+1}'\left(\alpha|I\right)\right]'$ and define $\overline{B}\left(\alpha|x, I\right) = P\left(x\right)'\overline{\beta}_0\left(\alpha|I\right)$,

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

By Proposition 3 and its proof, there exists some $\beta^*\left(\cdot|\cdot\right)$ grouping the entries in (2.14) such that

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|P\left(x\right)\beta^*\left(\alpha|I\right) - B\left(\alpha|x, I\right)\right| = o\left(K_L^{-\frac{s+1}{d_{\mathcal{M}}}}\right) = o\left(h^{s+1}\right).$$

Let $b^*\left(\cdot|\cdot\right)$ and $\mathsf{b}^*\left(\cdot|\cdot\right) = Hb^*\left(\cdot|\cdot\right)$ with

$$\beta^*\left(\alpha|I\right)' = \left[\beta_0^*\left(\alpha|I\right)', \beta_1^*\left(\alpha|I\right)', \ldots, \beta_{s+1}^*\left(\alpha|I\right)'\right],$$

$\beta_p^*\left(\alpha|I\right) = \left[\beta_k^{(p)}\left(\alpha|I\right), 1 \le k \le K_L\right]$ as in (2.14), $p = 0, \ldots, s+1$.

The next notations deal with the differentiability of the objective functions $\widehat{\mathsf{R}}\left(\cdot; \alpha, I\right)$. Since

$$\frac{\partial\rho_{\alpha+ht}\left(B - P\left(x_\ell, t\right)'\mathsf{b}\right)}{\partial\mathsf{b}''} = \left\{\mathbb{I}\left(B_{i\ell} \le P\left(x_\ell, t\right)'\mathsf{b}\right) - \left(\alpha + ht\right)\right\} P\left(x_\ell, t\right),$$

almost everywhere, it follows that $\widehat{\mathsf{R}}\left(\cdot; \alpha, I\right)$ is differentiable with

$$\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}; \alpha, I\right) = \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_\ell = I\right)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left\{\mathbb{I}\left(B_{i\ell} \le P\left(x_\ell, t\right)'\mathsf{b}\right) - \left(\alpha + ht\right)\right\} P\left(x_\ell, t\right) K\left(t\right) dt$$

and $\overline{\mathsf{R}}^{(1)}\left(\mathsf{b}; \alpha, I\right) = \mathbb{E}\left[\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b}; \alpha, I\right)\right]$ by the Dominated Convergence Theorem. When $\mathsf{b} = \mathsf{b}^*\left(\alpha|I\right)$, $P\left(x, t\right)'\mathsf{b}^*\left(\alpha|I\right) = P\left(x, ht\right)'\beta^*\left(\alpha|I\right)$ is close to $B\left(\alpha + ht|x, I\right)$, which inverse as a function of t in

$$\mathcal{I}_{\alpha,h} = \left[\underline{I}_{\alpha,h}, \overline{I}_{\alpha,h}\right] = \left[-\min\left(1, \frac{\alpha}{h}\right), \min\left(1, \frac{1-\alpha}{h}\right)\right] = [-1, 1] \cap \left[-\frac{\alpha}{h}, \frac{1-\alpha}{h}\right]$$

is

$$\frac{G\left(u|x, I\right) - \alpha}{h}, \quad u \in \left[B\left(\alpha + h\underline{I}_{\alpha,h}|x, I\right), B\left(\alpha + h\overline{I}_{\alpha,h}|x, I\right)\right].$$

When h is small enough, it will be shown in the proof of Lemma B.1 below that

$$\begin{aligned}\frac{\partial}{\partial t}\left[P(x,ht)'b^{*}(\alpha|I)\right] &= h\left[\pi^{(1)}(ht)\otimes P(x)\right]'b^{*}(\alpha|I) \\ &= hP(x)'\beta_1^{*}(\alpha|I)+O\left(h^{2}\right)\end{aligned}$$

uniformly since $\pi^{(1)}(ht)' = [0,1,ht,\ldots,(ht)^{s}/s!]$ and that $P(x)'\beta_1^{*}(\alpha|I)$ converges uniformly to $B^{(1)}(\alpha|x,I)$ when K_L diverges and is therefore positive, so that $P(x,t)'\mathsf{b}^{*}(\alpha|I)$ is an increasing function of t in $\mathcal{I}_{\alpha,h}$ for h small enough. Since $\max_{(x,t)\in\mathcal{X}\times[-1,1]}\|P(x,t)\| = O\left(h^{-d_{\mathcal{M}}/2}\right)$, $t\rightarrow P(x,t)'\mathsf{b}$ is also strictly increasing provided b is close enough to $\mathsf{b}^{*}(\alpha|I)$. In such case, it is convenient to redefine $P(x,t)'\mathsf{b}$ as follows[1]

$$\Psi(t|x,\mathsf{b}) = \begin{cases} P\left(x,\overline{I}_{\alpha,h}\right)'\mathsf{b} & t>\overline{I}_{\alpha,h} \\ P(x,t)'\mathsf{b} & t\in\mathcal{I}_{\alpha,h} \\ P\left(x,\underline{I}_{\alpha,h}\right)'\mathsf{b} & t<\underline{I}_{\alpha,h}\end{cases}.$$

When $\Psi(\cdot|x,\mathsf{b})$ has an inverse, define

$$\Phi(u|x,\mathsf{b}) = \begin{cases}\alpha+h\overline{I}_{\alpha,h} & u>\Psi\left(\overline{I}_{\alpha,h}|x,\mathsf{b}\right) \\ \alpha+h\Psi^{-1}(u|x,\mathsf{b}) & u\in\Psi\left(\mathcal{I}_{\alpha,h}|x,\mathsf{b}\right) \\ \alpha+h\underline{I}_{\alpha,h} & u<\Psi\left(\underline{I}_{\alpha,h}|x,\mathsf{b}\right)\end{cases},$$

$$\Delta(u|x,\mathsf{b}) = \frac{\Phi(u|x,\mathsf{b})-\alpha}{h} = \begin{cases}\overline{I}_{\alpha,h} & u>\Psi\left(\overline{I}_{\alpha,h}|x,\mathsf{b}\right) \\ \Psi^{-1}(u|x,\mathsf{b}) & u\in\Psi\left(\mathcal{I}_{\alpha,h}|x,\mathsf{b}\right) \\ \underline{I}_{\alpha,h} & u<\Psi\left(\underline{I}_{\alpha,h}|x,\mathsf{b}\right)\end{cases},$$

which is such that, as seen above, the central part of $\Phi(u|x,\mathsf{b}^{*}(\alpha|I))$ is close to $G(u|x,I)$ when u is in $\Psi(\mathcal{I}_{\alpha,h}|x,\mathsf{b})$. Observe now that, provided $\Psi(\cdot|x,\mathsf{b})$ is increasing and since the support of $K(\cdot)$ is $[-1,1]$

$$\begin{aligned}&\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\left\{\mathbb{I}\left(B_{i\ell}\leq\Psi(t|x_\ell,\mathsf{b})\right)-(\alpha+ht)\right\}P(x_\ell,t)K(t)\,dt \\ &= \int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\left\{\mathbb{I}\left(\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b})-\alpha}{h}\leq t\right)-(\alpha+ht)\right\}P(x_\ell,t)K(t)\,dt \\ &= \int_{\frac{\Phi(\mathsf{B}_{i\ell}|\mathsf{x}_\ell,\mathsf{b})-\alpha}{\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}P(x_\ell,t)K(t)\,dt-\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}(\alpha+ht)P(x_\ell,t)K(t)\,dt\end{aligned}$$

[1]In principle $\Psi(\cdot|\cdot)$ should be denoted $\Psi_{\alpha,h}(\cdot|\cdot)$ to acknowledge that its definition depends upon α and h. Instead, t is restricted to lie in $\mathcal{I}_{\alpha,h}$ in the sequel. The same comment applies for the functions $\Psi(\cdot|\cdot)$ and $\Delta(\cdot|\cdot)$ introduced below.

which is differentiable with respect to b, with for $B_{i\ell}$ in $\Psi\left(\mathcal{I}_{\alpha,h}|x,\mathsf{b}\right)$

$$\frac{\partial\Phi\left(B_{i\ell}|x_\ell,\mathsf{b}\right)}{\partial\mathsf{b}'}=-\frac{P\left(x,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)}{\Psi^{(1)}\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)/h}\mathbb{I}\left[B_{i\ell}\in\Psi\left(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}\right)\right].$$

Hence, for h small enough and for b in the vicinity of $\mathsf{b}^*\left(\alpha|I\right)$, $\widehat{\mathsf{R}}\left(\mathsf{b};\alpha,I\right)$ and $\overline{\mathsf{R}}\left(\mathsf{b};\alpha,I\right)$ are twice continuously differentiable with,

$$\begin{aligned}\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)&=\frac{1}{LIh}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\mathbb{I}\left[B_{i\ell}\in\Psi\left(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}\right),I_\ell=I\right]\\&\qquad\frac{P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)'}{\Psi^{(1)}\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)/h}K\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right),\\\overline{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)&=\mathbb{E}\left[\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)\right].\end{aligned}$$

The next lemma details some properties of the functions $\Psi\left(\cdot|x,\mathsf{b}\right)$ and $\Phi\left(\cdot|x,\mathsf{b}\right)$ that were briefly sketched above. Define

$$\begin{aligned}\mathcal{BI}_{\alpha,h}&=\left\{\mathsf{b};\min_{(t,x)\in\mathcal{I}_{\alpha,h}\times\mathcal{X}}\frac{\partial\Psi\left(t|x,\mathsf{b}\right)}{\partial t}>0\right\},\\\underline{\mathcal{BI}}_{\alpha,h}&=\left\{\mathsf{b};\min_{(t,x)\in\mathcal{I}_{\alpha,h}\times\mathcal{X}}\frac{\partial\Psi\left(t|x,\mathsf{b}\right)}{\partial t}>h/\underline{f},\max_{p=1,\ldots,s+1}\left(\frac{\max_{x\in\mathcal{X}}\left|P\left(x\right)'\mathsf{b}_p\right|}{h}\right)<\overline{f}\right\},\end{aligned}$$

recalling that $\mathsf{b}=\left[\mathsf{b}_0',\ldots,\mathsf{b}_{s+1}'\right]'$ and where $\underline{f}$ and $\overline{f}$ will be taken large enough. While $\mathcal{BI}_{\alpha,h}$ is used to bound the first derivative of $\Psi\left(\cdot|x,\mathsf{b}\right)$ away from 0, $\underline{\mathcal{BI}}_{\alpha,h}$ is used to bound the successive derivatives $\Psi^{(p)}\left(\cdot|x,\mathsf{b}\right)$, $p=1,\ldots,s+1$, away from infinity. As made possible by Lemma B.1-(i), below, an Euclidean ball $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),Ch^{d_{\mathcal{M}}/2+1}\right)$ with a small enough constant $C>0$ will be considered instead of the sets $\mathcal{BI}_{\alpha,h}$ and $\underline{\mathcal{BI}}_{\alpha,h}$.

Lemma B.1 *Suppose Assumptions A and S hold with* $\max_{x\in\mathcal{X}}\|P(x)\|=O\left(K_L^{1/2}\right)$, $K_L=h^{1/d_{\mathcal{M}}}$ *that* $\underline{f}$ *and* $\overline{f}$ *are large enough. Then, h small enough and all I in* $\mathcal{I}$,

i. $\mathsf{b}^*\left(\alpha|I\right)$ *belongs to* $\underline{\mathcal{BI}}_{\alpha,h}\subset\mathcal{BI}_{\alpha,h}$ *and for C small enough* $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),Ch^{d_{\mathcal{M}}/2+1}\right)$ *is a subset of* $\underline{\mathcal{BI}}_{\alpha,h}$, *for all* α *in* $[0,1]$.

ii. For all b *in* $\mathcal{BI}_{\alpha,h}$ *and all u in* $\Psi\left(\mathcal{I}_{\alpha,h}|x,\mathsf{b}\right)$

$$\begin{aligned}\frac{\partial\Phi\left(u|x,\mathsf{b}\right)}{\partial\mathsf{b}'}&=-\frac{P\left(x,\Delta\left(u|x,\mathsf{b}\right)\right)}{\Psi\left(\Delta\left(u|x,\mathsf{b}\right)|x,\mathsf{b}\right)/h},\\\frac{\partial\Phi\left(u|x,\mathsf{b}\right)}{\partial u}&=\frac{1}{\Psi\left(\Delta\left(u|x,\mathsf{b}\right)|x,\mathsf{b}\right)/h}.\end{aligned}$$

iii. It holds that

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)-B\left(\alpha+ht|x,I\right)\right|=o\left(h^{s+1}\right),$$

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\left(B\left(\alpha+ht|x,I\right)-\Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)\right)\right.$$
$$\left.-\frac{(ht)^{s+2}}{(s+2)!}\alpha B^{(s+2)}\left(\alpha|I\right)\right|=o\left(h^{s+2}\right),$$

and, recalling $\mathsf{b}_1^*\left(\alpha|I\right)=h\beta_1^*\left(\alpha|I\right)$

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\alpha\beta_1^*\left(\alpha|I\right)-\alpha B^{(1)}\left(\alpha|x,I\right)\right|=o\left(h^{s+1}\right),$$
$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|\Phi\left(u|x,\mathsf{b}^*\left(\alpha|I\right)\right)-G\left(u|x,I\right)\right|=o\left(h^{s+1}\right).$$

iv. There is a $C>0$ *such that for any* b_0 *and* b_1 *in* $\underline{\mathcal{BI}}_{\alpha,h}$ *and all* α *in* $[0,1]$

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\Psi\left(t|x,\mathsf{b}_1\right)-\Psi\left(t|x,\mathsf{b}_0\right)\right|,$$
$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_0\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]}\left|\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)\right|,$$
$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_0\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]}\left|\frac{\partial\Phi}{\partial u}\left(u|x,\mathsf{b}_1\right)-\frac{\partial\Phi}{\partial u}\left(u|x,\mathsf{b}_0\right)\right|,$$
$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_0\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]}\left|\Psi^{(1)}\left(\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right)-\Psi^{(1)}\left(\Delta\left(u|x_\ell,\mathsf{b}_0\right)|x_\ell,\mathsf{b}_0\right)\right|,$$

are all smaller or equal to $Ch^{-d_{\mathcal{M}}/2}\left\|\mathsf{b}_1-\mathsf{b}_0\right\|$.

Let $\Omega_h\left(\alpha\right)$, $\Omega\left(0\right)$, $\Omega\left(1\right)$, $\Omega=\Omega\left(0\right)+\Omega\left(1\right)$ and $\Omega_{1h}\left(\alpha\right)$ be the $(s+2)\times(s+2)$ matrices

$$\Omega_h\left(\alpha\right)=\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt=\left[\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}t^{p_1+p_2}K\left(t\right)dt,0\leq p_1,p_2\leq s+1\right],$$
$$\Omega\left(0\right)=\int_{-1}^{0}\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt,\quad\Omega\left(1\right)=\int_{0}^{1}\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt,$$
$$\Omega_{1h}\left(\alpha\right)=\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}t\pi\left(t\right)\pi\left(t\right)'K\left(t\right)dt,$$

While $\Omega_h\left(\alpha\right)\preceq\Omega$ for all α and h, it holds that for h small enough $\Omega_h\left(\alpha\right)\succeq\Omega\left(0\right)$ for all α in $[0,1/2]$ and $\Omega_h\left(\alpha\right)\succeq\Omega\left(1\right)$ for all α in $[1/2,1]$.

Lemma B.2 *Suppose Assumptions A, R-(i) and S hold, that* $\underline{f}$ *and* $\overline{f}$ *are large enough. Then, for* $K_L^{-1/d_{\mathcal{M}}} = O(h)$*,* h *small enough, all* I *in* $\mathcal{I}$*, and any* $C > 0$ *small enough, (i) It holds that* $\overline{\mathsf{R}}^{(2)}(\cdot;\alpha,I)$ *is continuously differentiable over* $\mathcal{B}\left(\mathsf{b}^*(\alpha|I), Ch^{d_{\mathcal{M}}/2+1}\right)$ *with*

$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}_1,\mathsf{b}_0\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)} \frac{\left\|\overline{\mathsf{R}}^{(2)}(\mathsf{b}_1;\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b}_0;\alpha,I)\right\|}{\|\mathsf{b}_1-\mathsf{b}_0\| / (\alpha(1-\alpha)+h)} = O\left(h^{-d_{\mathcal{M}}/2}\right).$$

(ii) The eigenvalues of $\overline{\mathsf{R}}^{(2)}[\mathsf{b}^*(\alpha|I);\alpha,I]$ *belongs to* $[1/C, C]$ *for a large enough* C*, for all* α *in* $[0,1]$ *and* h *small enough with*

$$\begin{aligned}\max_{\alpha\in[0,1]} \Bigg\| \overline{\mathsf{R}}^{(2)}[\mathsf{b}^*(\alpha|I);\alpha,I] - \Omega_h(\alpha)\otimes\mathbb{E}\left[\frac{\mathbb{I}(I_\ell=1)P(x_\ell)P(x_\ell)'}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right] \\ +\Omega_{1h}(\alpha)\otimes\mathbb{E}\left[\frac{\mathbb{I}(I_\ell=1)B^{(2)}(\alpha|x_\ell,I_\ell)P(x_\ell)P(x_\ell)'}{\left(B^{(1)}(\alpha|x_\ell,I_\ell)\right)^2}\right]\Bigg\| = o(h).\end{aligned}$$

Lemma B.2-(i) yields, for any $C > 0$,

$$\begin{aligned}\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}(\mathsf{b}^*(\alpha|I),Ch^{s+1})} \left\|\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b}^*(\alpha|I);\alpha,I)\right\| &= O\left(h^{s-d_{\mathcal{M}}/2}\right) \\ \text{if } h^s &= o\left(h^{d_{\mathcal{M}}/2}\right), \\ \max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),C\left(\frac{\log L}{L}(\alpha(1-\alpha)+h)\right)^{1/2}\right)} \frac{\left\|\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b}^*(\alpha|I);\alpha,I)\right\|}{\left(\frac{\log L}{L(\alpha(1-\alpha)+h)}\right)^{1/2}} &= O\left(h^{-d_{\mathcal{M}}/2}\right) \\ \text{if } \left(\frac{\log L}{L}\right)^{1/2} &= o\left(h^{d_{\mathcal{M}}/2+1}\right).\end{aligned}$$

It then follows that the eigenvalues of $\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)$ stays bounded away from 0 and infinity uniformly in α and in b in the two neighborhoods considered above, under the corresponding bandwidth assumption.

The two next Lemmas study the first and second derivatives of $\widehat{\mathsf{R}}(\cdot;\alpha,I)$ in a shrinking vicinity of $\mathsf{b}^*(\alpha|I)$. In particular, Lemma B.3 implies that $\widehat{\mathsf{R}}(\cdot;\alpha,I)$ is strictly convex over such a vicinity with a probability tending to 1.

Lemma B.3 *Suppose Assumptions A, R-(i,ii) and S hold, and* $\log L/\left(Lh^{d_{\mathcal{M}}+1}\right) = o(1)$*. Then, for any* $C > 0$ *small enough,*

$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)} \left\|\widehat{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)\right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)$$

Lemma B.4 *Suppose Assumptions A, R-(i,ii) and S hold, and* $\log L / \left(Lh^{d_{\mathcal{M}}+1}\right) = o(1)$. *Then, for any* $C > 0$,

$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)} \left\| \frac{\widehat{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)}{(h+\alpha(1-\alpha))^{1/2}} \right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\right).$$

Since $\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) = 0$ and assuming $h^{s+1} = O\left(h^{d_{\mathcal{M}}/2+1}\right)$, $\sup_{\alpha\in[0,1]} \left\|\overline{\mathsf{b}}(\alpha|I) - \mathsf{b}^*(\alpha|I)\right\| = o\left(h^{s+1}\right)$ as established in (B.4), it holds that

$$\max_{\alpha\in[0,1]} \left\| \frac{\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)}{(h+\alpha(1-\alpha))^{1/2}} \right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\right).$$

The next Lemma studies the leading term $\widehat{\mathsf{e}}(\alpha|I)$ of $\widehat{\mathsf{b}}(\alpha|I) - \overline{\mathsf{b}}(\alpha|I)$,

$$\widehat{\mathsf{e}}(\alpha|I) = -\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]^{-1} \widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)$$

see Theorem B.9 below. Note that $\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)$ is not necessarily defined and invertible unless $h^{s+1} = O\left(h^{d_{\mathcal{M}}/2+1}\right)$ and $\sup_{\alpha\in[0,1]} \left\|\overline{\mathsf{b}}(\alpha|I) - \mathsf{b}^*(\alpha|I)\right\| = o\left(h^{s+1}\right)$ as therefore assumed and established in the proof of Theorem B.8 below, see (B.4).

Lemma B.5 *Suppose Assumptions A, R and S-(i,ii) hold, and* $1/\left(Lh^{d_{\mathcal{M}}+1}\right) = o(1)$, $s \geq d_{\mathcal{M}}/2$ *and* $\sup_{\alpha\in[0,1]} \left\|\overline{\mathsf{b}}(\alpha|I) - \mathsf{b}^*(\alpha|I)\right\| = o\left(h^{s+1}\right)$. *Then (i) uniformly in* (α, x) *in* $[0,1]\times\mathcal{X}$

$$\operatorname{Var}\left[P(x)'\widehat{\mathsf{e}}_0(\alpha|I)\right] = O\left(\frac{1}{Lh^{d_{\mathcal{M}}}}\right)$$

and $\operatorname{Var}\left[P(x)'\widehat{\mathsf{e}}_1(\alpha|I)/h\right] = O\left(\frac{1}{Lh^{d_{\mathcal{M}}+1}}\right)$ *with* $\operatorname{Var}\left[\widehat{\mathsf{e}}_1(\alpha|I)/h\right]$ *having the expansion*

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

(ii) If Assumption S-(iii) also holds

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|P(x)'\widehat{\mathsf{e}}_0(\alpha|I)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\right),$$

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|P(x)'\frac{\widehat{\mathsf{e}}_1(\alpha|I)}{h}\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right).$$

B.2 Asymptotic bias

The study of the bias $\overline{V}(\alpha|x,I) - V(\alpha|x,I)$ and $\overline{B}(\alpha|x,I) - B(\alpha|x,I)$ is based on the following Lemma which is a consequence of the Kantorovitch-Newton Theorem, see e.g. Gragg and Tapia (1974).

Lemma B.6 *Let $\mathcal{F}(\cdot):\mathbb{R}^D \rightarrow \mathbb{R}$ 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. $C_0^2 C_1 \epsilon \leq 1/2$.

Then there is a unique $\overline{\mathbf{x}}$ such that $\|\overline{\mathbf{x}} - \mathbf{x}^*\| < 2C_0\epsilon$ and $\mathcal{F}^{(1)}(\overline{\mathbf{x}}) = 0$.

The next lemma, established in Appendix B, will be used at the end of the proof of Theorem B.8 below.

Lemma B.7 *Suppose Assumptions A, S and R-(ii). Then the ℓ_1 norm of the columns of the matrix*

$$A_{\alpha,h} = \mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P(x_\ell,t)\,P(x_\ell,t)'\,K(t)\,dt}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right]$$

are bounded independently of L and α. That is, if $A_{\alpha,h} = \left[A_{\alpha,h}(j_1,j_2), 1 \leq j_1, j_2 \leq (s+1)K_L\right]$,

$$\max_L \max_{\alpha\in[0,1]} \max_{1\leq j_1\leq(s+1)K_L} \sum_{j_2=1}^{(s+1)K_L} |A_{\alpha,h}(j_1,j_2)| < \infty.$$

In the next theorem,

$$\begin{aligned}\mathsf{bias}_h(\alpha|I) = \mathbb{E}^{-1}&\left[\frac{\mathbb{I}(I_\ell = I)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P(x_\ell,t)\,P(x_\ell,t)'\,K(t)\,dt}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right]\\ \times\,\mathbb{E}&\left[\frac{\mathbb{I}(I_\ell = I)\,B^{(s+2)}(\alpha|x_\ell,I_\ell)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} t^{s+2}P(x_\ell,t)\,K(t)\,dt}{(s+2)!\,B^{(1)}(\alpha|x_\ell,I_\ell)}\right],\end{aligned}$$

and

$$\mathsf{bias}_h(\alpha|I) = \left[\mathsf{bias}_{0h}(\alpha|I)', \ldots, \mathsf{bias}_{s+1,h}(\alpha|I)'\right]$$

where the subvectors $\mathsf{bias}_{ph}(\alpha|I)$ are of dimension K_L. While $\mathsf{bias}_h(\alpha|I)$ may not exist for $\alpha = 0$, the function $\mathsf{Bias}_h(\alpha|I) = \alpha\mathsf{bias}_h(\alpha|I)$ in (3.6) can be set to 0 when $\alpha = 0$ by Proposition 3-(i).

Theorem B.8 *Suppose that Assumptions A, H-(i) and R hold with* $h = O\left(K_L^{-1/d_{\mathcal{M}}}\right)$ *and* $s \geq d_{\mathcal{M}}/2$. *Then, for* h *small enough* $\overline{\mathsf{b}}(\alpha|I) = \arg\min_{\mathsf{b}} \overline{\mathsf{R}}(\mathsf{b};\alpha,I)$ *is unique for all* α *in* $[0,1]$ *and*

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|\overline{V}(\alpha|x,I) - V(\alpha|x,I) - \frac{h^{s+1}P(x)'\alpha\mathsf{bias}_{1h}(\alpha|I)}{I-1}\right| = o\left(h^{s+1}\right)$$

with $\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|P(x)'\alpha\mathsf{bias}_{1h}(\alpha|I)\right| = O(1)$.

Moreover $\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|\overline{B}(\alpha|x,I) - B(\alpha|x,I)\right| = o\left(h^{s+1}\right)$,

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|\alpha\left(\overline{B}(\alpha|x,I) - B(\alpha|x,I) - h^{s+2}P(x)'\mathsf{bias}_{0h}(\alpha|I)\right)\right| = o\left(h^{s+2}\right)$$

with $\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} \left|P(x)'\alpha\mathsf{bias}_{0h}(\alpha|I)\right| = O(1)$. *Hence uniformly over* x *and* α *in any compact subset of* $(0,1]$,

$$\overline{B}(\alpha|x,I) = B(\alpha|x,I) + h^{s+2}P(x)'\mathsf{bias}_{0h}(\alpha|I) + o\left(h^{s+2}\right).$$

The proof of Theorem B.8 establishes that $\sup_{\alpha\in[0,1]} \left\|\overline{\mathsf{b}}(\alpha|I) - \mathsf{b}^*(\alpha|I)\right\| = o\left(h^{s+1}\right)$, see (B.4), an intermediary result which will be used all along the proof. If $d_{\mathcal{M}}/2 \leq s$, $\log L/\left(Lh^{d_{\mathcal{M}}+1}\right) = o(1)$ and by Lemma B.3 and a second order Taylor expansion

$$\sup_{\alpha\in[0,1]} \sup_{\mathsf{b}\in\mathcal{B}\left(\overline{\mathsf{b}}(\alpha|I),Ch^{s+1}\right)} \left|h^{-2(s+1)}\left\{\widehat{\mathsf{R}}(\mathsf{b};\alpha,I) - \widehat{\mathsf{R}}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) - \left(\mathsf{b}-\overline{\mathsf{b}}(\alpha|I)\right)'\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right\}\right.$$
$$\left. -\frac{h^{-2(s+1)}}{2}\left(\mathsf{b}-\overline{\mathsf{b}}(\alpha|I)\right)'\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\left(\mathsf{b}-\overline{\mathsf{b}}(\alpha|I)\right)\right| = o_{\mathbb{P}}(1).$$

Then by Lemma B.2 and the Argmax Theorem $\widehat{\mathsf{R}}(\cdot;\alpha,I)$ has a unique minimizer over $\mathsf{b}\in\mathcal{B}\left(\overline{\mathsf{b}}(\alpha|I), Ch^{s+1}\right)$ for each α, with a probability tending to 1. Since $\widehat{\mathsf{R}}(\cdot;\alpha,I)$ is convex a local minimum is also a global one. This implies that the AQR or ASQR estimators $\widehat{b}(\alpha|I) = H^{-1}\widehat{\mathsf{b}}(\alpha|I)$ are unique for all α in $[0,1]$ with a probability tending to 1.

Proof of Theorem B.8. 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 $\overline{\mathsf{R}}(\mathsf{b};\alpha,I)$ is strictly convex by checking the conditions of Lemma B.6, which will also gives the rate stated in the Theorem and the uniqueness of $\overline{\mathsf{b}}(\alpha|I)$. It is first claimed that

$$\max_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|\overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^{*}\left(\alpha|I\right);\alpha,I\right)\right\|=\epsilon_{L}\text{ with}\tag{B.2}$$
$$\epsilon_{L}=O\left(\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\Psi\left(t|x,\mathsf{b}^{*}\left(\alpha|I\right)\right)-B\left(\alpha+ht|x,I\right)\right|\right)=o\left(h^{s+1}\right),$$

where $\epsilon_L=o\left(h^{s+1}\right)$ follows from Lemma B.1-(iii). To see that (B.2) holds, observe that

$$\left\|\overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^{*}\left(\alpha|I\right);\alpha,I\right)\right\|=\max_{\theta;\theta'\theta=1}\left|\theta'\overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^{*}\left(\alpha|I\right);\alpha,I\right)\right|.\tag{B.3}$$

But uniformly in $\alpha\in[0,1]$ and by Assumption R-(i), Lemma B.1-(iii),

$$\begin{aligned}
&\left|\theta'\overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^{*}\left(\alpha|I\right);\alpha,I\right)\right|\\
&=\mathbb{E}\left[\mathbb{I}\left(I_{\ell}=I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left\{G\left(P\left(x_{\ell},t\right)\mathsf{b}^{*}\left(\alpha|I\right)|x_{\ell},I_{\ell}\right)-G\left(B\left(\alpha+ht|x,I\right)|x_{\ell},I_{\ell}\right)\right\}\right.\\
&\qquad\left.\theta'\left(P\left(x_{\ell}\right)\otimes\pi\left(t\right)\right)K\left(t\right)dt\right]\\
&\leq C\epsilon_{L}\mathbb{E}^{1/2}\left[\int_{-1}^{1}\left(\theta'\left(P\left(x_{\ell}\right)\otimes\pi\left(t\right)\right)\right)^{2}dt\right]\leq C\epsilon_{L}\left(\theta'\theta\right)^{1/2}=C\epsilon_{L}.
\end{aligned}$$

Hence (B.2) holds, which is the first part of Condition (i) in Lemma B.6. The second part of Condition (i) follows from Lemma B.2-(ii) which ensures that there is a $C_0>0$ such that, for L large enough,

$$\sup_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|\left[\overline{\mathsf{R}}^{(2)}\left(\mathsf{b}^{*}\left(\alpha|I\right);\alpha,I\right)\right]^{-1}\right\|\leq C_{0}$$

Note that $s\geq d_{\mathcal{M}}/2$ and $\epsilon_L=o\left(h^{s+1}\right)$ gives that

$$\mathcal{B}\left(\mathsf{b}^{*}\left(\alpha|I\right),2C_{0}\epsilon_{L}\right)\subset\mathcal{B}\left(\mathsf{b}^{*}\left(\alpha|I\right),Ch^{d_{\mathcal{M}}/2+1}\right)$$

for all $C_0,C>0$ provided L is large enough, for all α and all I. Condition (ii) in Lemma B.6 follows from Lemma B.2-(i) which ensures that for $C_{1L}=O\left(h^{d_{\mathcal{M}}/2+1}\right)$,

$$\left\|\overline{\mathsf{R}}^{(2)}\left(\mathsf{b}_{1};\alpha,I\right)-\overline{\mathsf{R}}^{(2)}\left(\mathsf{b}_{0};\alpha,I\right)\right\|\leq C_{1L}\left\|\mathsf{b}_{1}-\mathsf{b}_{0}\right\|$$

for all b_1, b_0 in $\mathcal{B}\left(\mathsf{b}^{*}\left(\alpha|I\right),2C_{0}\epsilon_{L}\right)$ and all α, I. For condition (iii) in Lemma B.6, $\epsilon_L=o\left(h^{s+1}\right)$ and $s\geq d_{\mathcal{M}}/2$ implies $C_0^2C_{1L}\epsilon_L=o\left(h^{s-d_{\mathcal{M}}/2}\right)=o(1)<1/2$ for L large enough. Hence Lemma B.6 ensures that, for L large enough, all α and all I, there is a unique $\overline{\mathsf{b}}(\alpha|I)$

in $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right), 2C_0\epsilon_L\right)$ such that

$$\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right); \alpha, I\right) = 0$$

and is therefore the unique minimizer of $\overline{\mathsf{R}}\left(\cdot;\alpha,I\right)$ over $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right), 2C_0\epsilon_L\right)$. Since the convex function $\overline{\mathsf{R}}\left(\cdot;\alpha,I\right)$ cannot have several local minimizers, $\overline{\mathsf{b}}\left(\alpha|I\right)$ is also the unique global minimizer of $\overline{\mathsf{R}}\left(\cdot;\alpha,I\right)$. Since $\epsilon_L = o\left(h^{s+1}\right)$, it follows that

$$\sup_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left\|\overline{\mathsf{b}}\left(\alpha|I\right) - \mathsf{b}^*\left(\alpha|I\right)\right\| = o\left(h^{s+1}\right). \tag{B.4}$$

Consider now $\alpha\overline{\mathsf{b}}\left(\alpha|I\right) - \alpha\mathsf{b}^*\left(\alpha|I\right)$. Define

$$\overline{g}\left(\alpha|t,x,I\right) = \int_0^1 g\left(\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right) + u\left(B\left(\alpha+ht|x,I\right) - \Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)\right)|t,x,I\right) du$$

which is such that, uniformly in α in $[3h,1]$, x in $\mathcal{X}$ and t in $[-1,3/4]$

$$\begin{aligned}\overline{g}\left(\alpha|t,x,I\right) &= \int_0^1 g\left(\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right) + u\left(B\left(\alpha+ht|x,I\right) - \Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right)\right)|t,x,I\right) du \\ &= \int_0^1 g\left(B\left(\alpha+ht|x,I\right) + o\left(h^{s+1-d_{\mathcal{M}}/2}\right)|t,x,I\right) du \\ &\geq \left(1+o\left(1\right)\right) \max_{y\in[B(2h|x,I),B(1-2h|x,I)]} g\left(y|x,I\right) \geq C'' > 0\end{aligned}$$

by Lemma B.1-(iii,iv), (B.4), $o\left(h^{s+1-d_{\mathcal{M}}/2}\right) = o\left(h\right)$ and Proposition 3-(i). Now $\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right) = 0$ gives

$$\begin{aligned}0 &= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \left\{G\left[\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right)|x,I\right] - \left(\alpha+ht\right)\right\} P\left(x,t\right) K\left(t\right) dt\right) f\left(x,I\right) dx \\ &= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \left\{G\left[\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right)|x,I\right] - G\left[B\left(\alpha+ht|x,I\right)|x,I\right]\right\} P\left(x,t\right) K\left(t\right) dt\right) f\left(x,I\right) dx \\ &= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \overline{g}\left(\alpha|t,x,I\right)\left\{\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right) - B\left(\alpha+ht|x,I\right)\right\} P\left(x,t\right) K\left(t\right) dt\right) f\left(x,I\right) dx \\ &= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \overline{g}\left(\alpha|t,x,I\right)\left\{\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right) - \Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)\right\} P\left(x,t\right) K\left(t\right) dt\right) f\left(x,I\right) dx \\ &+ \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \overline{g}\left(\alpha|t,x,I\right)\left\{\Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right) - B\left(\alpha+ht|x,I\right)\right\} P\left(x,t\right) K\left(t\right) dt\right) f\left(x,I\right) dx.\end{aligned}$$

Since $\left\{\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right)-\Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)\right\}P\left(x,t\right)=P\left(x,t\right)P\left(x,t\right)'\left(\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^*\left(\alpha|I\right)\right)$, by Assumption R-(i), and because $\overline{g}\left(\alpha|t,x,I\right)$, $f\left(x,I\right)$ are bounded away from 0 and infinity

$$\begin{aligned}\alpha\left(\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^*\left(\alpha|I\right)\right)&=\left[\int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\overline{g}\left(\alpha|t,x,I\right)P\left(x,t\right)P\left(x,t\right)'K\left(t\right)dt\right)f\left(x,I\right)dx\right]^{-1}\\&\times\int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\overline{g}\left(\alpha|t,x,I\right)\left\{\frac{\left(ht\right)^{s+2}}{\left(s+2\right)!}\alpha B^{(s+2)}\left(\alpha|I\right)+o\left(h^{s+2}\right)\right\}P\left(x,t\right)K\left(t\right)dt\right)f\left(x,I\right)dx\end{aligned}$$

uniformly in α in $[0,1]$ by Lemma B.1-(iii). By Assumption R-(ii) which implies in particular $\left\|\int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left|P\left(x,t\right)\right|K\left(t\right)dt\right)dx\right\|=O\left(1\right)$, it follows

$$\begin{aligned}&\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^*\left(\alpha|I\right)\\&\quad=o\left(h^{s+1}\right)\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_\ell=I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}P\left(x_\ell,t\right)P\left(x_\ell,t\right)'K\left(t\right)dt}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left|P\left(x_\ell,t\right)\right|K\left(t\right)dt\right],\end{aligned}$$

$$\begin{aligned}&\alpha\left(\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^*\left(\alpha|I\right)\right)=h^{s+2}\alpha\mathsf{bias}_h\left(\alpha|I\right)\\&\quad+o\left(h^{s+2}\right)\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_\ell=I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}P\left(x_\ell,t\right)P\left(x_\ell,t\right)'K\left(t\right)dt}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left|P\left(x_\ell,t\right)\right|K\left(t\right)dt\right],\end{aligned}\tag{B.5}$$

uniformly over $[0,1]$. Let

$$A=A_{\alpha,h}=\left[A_1,\ldots A_{J_L}\right]=\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_\ell=I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}P\left(x_\ell,t\right)P\left(x_\ell,t\right)'K\left(t\right)dt}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]$$

be a a $J_L\times J_L$ matrix with columns A_j, $j=1,\ldots,J_1$, $\left|A_j\right|_1$ the associated ℓ_1 norm and $\left|A\right|_{1,\infty}=\max_{j\leq J_L}\left|A_j\right|_1$, S a selection matrix which selects some columns of A, a, b some conformable vectors and $\left|a\right|_\infty$ the largest entry of a.

$$\left|a'SAb\right|=\left|\sum_j b_j a'\left[SA\right]_j\right|\leq\sum_j\left|b_j\right|\max_j\left|a'\left[SA\right]_j\right|\leq\left|b\right|_1\left|A\right|_{1,\infty}\left|a\right|_\infty.$$

This gives, since $\max_{\alpha,L}|A|_{1,\infty}<\infty$ by Lemma B.7 and by Assumption R-(ii),

$$\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P'(x)\,S\mathsf{bias}_h(\alpha|I)\right| \\
&\qquad\le C\left(\max_{x\in\mathcal{X}}\sum_{k=1}^{K_L}|P_k(x)|\right)\times\max_{1\le k\le K_L}\int|P_k(x)|\,dx=O(1),\\
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P'(x)\,SA\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}|P(x_\ell,t)|\,K(t)\,dt\right]\right|\\
&\qquad\le C\left(\max_{x\in\mathcal{X}}\|P(x)\|\right)\times\max_{1\le k\le K_L}\int|P_k(x)|\,dx=O(1).
\end{aligned}$$

Let S_0 and S_1 be the selection matrices $S_0\mathsf{b}=\beta_0$ and $S_1\mathsf{b}=h\beta_1$, so that $\overline{B}(\alpha|x,I)=P'(x)\,S_0\overline{\mathsf{b}}(\alpha|I)$ and $\overline{B}^{(1)}(\alpha|x,I)=P'(x)\,S_1\overline{\mathsf{b}}(\alpha|I)/h$. Then (B.4), (B.5), Lemma B.1-(iii) and the above imply

$$\begin{aligned}
\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\overline{B}(\alpha|x,I)-B(\alpha|x,I)\right| &\le \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P'(x)\,S_0\left(\overline{\mathsf{b}}(\alpha|I)-\mathsf{b}^*(\alpha|I)\right)\right|\\
&\quad+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\Psi\left(0|x,\mathsf{b}^*(\alpha|I)\right)-B(\alpha|x,I)\right|\\
&=o\left(h^{s+1}\right),
\end{aligned}$$

$$\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\alpha\left(\overline{B}^{(1)}(\alpha|x,I)-B^{(1)}(\alpha|x,I)\right)-h^{s+1}P'(x)\,\alpha S_1\mathsf{bias}_h(\alpha|I)\right|\\
&=\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\frac{1}{h}\left|\alpha P'(x)\,S_1\left(\overline{\mathsf{b}}(\alpha|I)-\mathsf{b}^*(\alpha|I)-h^{s+2}P'(x)\,\mathsf{bias}_h(\alpha|I)\right)\right|\\
&+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\frac{1}{h}\left|\alpha\left(P'(x)\,\mathsf{b}_1^*(\alpha|I)-hB^{(1)}(\alpha|x,I)\right)\right|\\
&=o\left(h^{s+1}\right).
\end{aligned}$$

This ends the proof of the Theorem since $\overline{V}(\alpha|x,I)=\overline{B}(\alpha|x,I)+\alpha\overline{B}^{(1)}(\alpha|x,I)/(I-1)$.□

B.3 Bahadur representation

Let $\widehat{\mathsf{e}}(\alpha|I)$ be a candidate linearization leading term for $\widehat{\mathsf{b}}(\alpha|I) - \overline{\mathsf{b}}(\alpha|I)$ and $\widehat{\mathsf{d}}(\alpha|I)$ the associate linearization error term, or Bahadur remainder term,

$$\widehat{\mathsf{e}}(\alpha|I) = -\left(\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right)^{-1}\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right), \tag{B.6}$$

$$\widehat{\mathsf{d}}(\alpha|I) = \widehat{\mathsf{b}}(\alpha|I) - \overline{\mathsf{b}}(\alpha|I) - \widehat{\mathsf{e}}(\alpha|I). \tag{B.7}$$

This section goal is to study the magnitude of $\widehat{\mathsf{d}}(\alpha|I)$ and, in the ASQR case, the magnitude of $P'(x)\widehat{\mathsf{d}}_0(\alpha|I)$ and $P'(x)\widehat{\mathsf{d}}_1(\alpha|I)/h$.

Theorem B.9 *Suppose Assumptions A, R-(i,ii) and S hold, $s \geq d_{\mathcal{M}}/2$ and*

$$\frac{\log L}{Lh^{2(d_{\mathcal{M}}+1)}} = o(1).$$

Then

$$\max_{\alpha\in[0,1]}\left\|\frac{Lh^{d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)/2}}{(h+\alpha(1-\alpha))^{1/2}\log L}\left\{\widehat{\mathsf{b}}(\alpha|I) - \overline{\mathsf{b}}(\alpha|I)\right.\right.$$
$$\left.\left.+\left(\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right)^{-1}\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right\}\right\| = O_{\mathbb{P}}(1)$$

with a diverging normalization term $Lh^{d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)/2}/\log L$. Moreover, for $\widehat{\mathsf{d}}(\alpha|I)$ as in (B.7),

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}\left\|P'(x)\widehat{\mathsf{d}}_0(\alpha|I)\right\| = O_{\mathbb{P}}\left(\frac{h^{1/2}\log L}{\left(Lh^{2d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)}\right)^{1/2}}\right),$$

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}\left\|P'(x)\frac{\widehat{\mathsf{d}}_1(\alpha|I)}{h}\right\| = O_{\mathbb{P}}\left(\frac{\log L}{\left(Lh^{2d_{\mathcal{M}}+1+(d_{\mathcal{M}}\vee 1)}\right)^{1/2}}\right).$$

Proof of Theorem B.9. We first introduce some renormalizations. Let, for $\widehat{\mathsf{e}}(\alpha|I)$ as in (B.6),

$$\varrho_{\alpha L} = \frac{(h+\alpha(1-\alpha))^{1/2}\log L}{Lh^{d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)/2}},$$
$$\widehat{\boldsymbol{R}}(d;\alpha,I) = \widehat{\mathsf{R}}\left(\overline{\mathsf{b}}(\alpha|I)+\widehat{\mathsf{e}}(\alpha|I)+\varrho_{\alpha L}d;\alpha,I\right) - \widehat{\mathsf{R}}\left(\overline{\mathsf{b}}(\alpha|I)+\widehat{\mathsf{e}}(\alpha|I);\alpha,I\right),$$

which is such that $\varrho_{\alpha L} = o(1)$ by $\log L/\left(Lh^{2(d_{\mathcal{M}}+1)}\right) = o(1)$

$$\frac{\widehat{\mathsf{d}}(\alpha|I)}{\varrho_{\alpha L}} = \arg\min_{d}\widehat{\boldsymbol{R}}(d;\alpha,I).$$

It follows that,

$$\left\{\sup_{\alpha\in[0,1]}\left\|\frac{\widehat{\mathsf{d}}\left(\alpha|I\right)}{\varrho_{\alpha L}}\right\|\geq t\right\}=\bigcup_{\alpha\in[0,1]}\left\{\left\|\frac{\widehat{\mathsf{d}}\left(\alpha|I\right)}{\varrho_{\alpha L}}\right\|\geq t\right\}$$
$$\subset\bigcup_{\alpha\in[0,1]}\left\{\inf_{\|d\|\geq t}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)\leq\inf_{\|d\|\leq t}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)\right\}\subset\bigcup_{\alpha\in[0,1]}\left\{\inf_{\|d\|\geq t}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)\leq 0\right\}$$

since $\inf_{\|d\|\leq t}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)\leq\widehat{\boldsymbol{R}}\left(0;\alpha,I\right)=0$. The next step uses a convexity argument that can be found in Pollard (1991). For any d with $\|d\|\geq t$

$$\begin{aligned}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)&=\frac{\|d\|}{t}\left\{\frac{t}{\|d\|}\widehat{\boldsymbol{R}}\left(\|d\|\frac{d}{\|d\|};\alpha,I\right)+\left(1-\frac{t}{\|d\|}\right)\widehat{\boldsymbol{R}}\left(0;\alpha,I\right)\right\}\\&\geq\frac{\|d\|}{t}\widehat{\boldsymbol{R}}\left(t\frac{d}{\|d\|};\alpha,I\right)\end{aligned}$$

so that $\inf_{\|d\|\geq t}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)\leq 0$ implies $\inf_{\|d\|=t}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)\leq 0$ and them

$$\left\{\sup_{\alpha\in[0,1]}\left\|\frac{\widehat{\mathsf{d}}\left(\alpha|I\right)}{\varrho_{\alpha L}}\right\|\geq t\right\}\subset\left\{\inf_{\alpha\in[0,1]}\inf_{\|d\|=t}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)\leq 0\right\}. \tag{B.8}$$

Thus it is sufficient to consider those d with $\|d\|=t$. The expression of $\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)$ gives, using two Taylor expansions with integral remainder,

$$\begin{aligned}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)&=\varrho_{\alpha L}d'\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right)+\widehat{\mathsf{e}}\left(\alpha|I\right);\alpha,I\right)\\&+\varrho_{\alpha L}^{2}d'\left[\int_0^1\widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right)+\widehat{\mathsf{e}}\left(\alpha|I\right)+u\varrho_{\alpha L}d;\alpha,I\right)(1-u)\,du\right]d'\\&=\varrho_{\alpha L}d'\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)\\&+\varrho_{\alpha L}d'\left[\int_0^1\widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right)+u\widehat{\mathsf{e}}\left(\alpha|I\right);\alpha,I\right)du\right]\widehat{\mathsf{e}}\left(\alpha|I\right)\\&+\varrho_{\alpha L}^{2}d'\left[\int_0^1\widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right)+\widehat{\mathsf{e}}\left(\alpha|I\right)+u\varrho_{\alpha L}d;\alpha,I\right)(1-u)\,du\right]d'.\end{aligned}$$

Since $\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)+\widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)\widehat{\mathsf{e}}\left(\alpha|I\right)=0$ by (B.6), it follows that

$$\begin{aligned}\widehat{\boldsymbol{R}}\left(d;\alpha,I\right)=\varrho_{\alpha L}d'&\left[\left\{\int_0^1\widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right)+u\widehat{\mathsf{e}}\left(\alpha|I\right);\alpha,I\right)-\widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)\right\}du\right]\widehat{\mathsf{e}}\left(\alpha|I\right)\\&+\varrho_{\alpha L}^{2}d'\left[\int_0^1\widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right)+\widehat{\mathsf{e}}\left(\alpha|I\right)+u\varrho_{\alpha L}d;\alpha,I\right)(1-u)\,du\right]d'.\end{aligned}$$

Lemma B.4 and (B.4) with $s \geq d_{\mathcal{M}}/2$, $\log L / \left(Lh^{2(d_{\mathcal{M}}+1)}\right) = o(1)$, Lemma B.2-(ii) give

$$\sup_{\alpha\in[0,1]} \left\| \frac{\widehat{\mathsf{e}}(\alpha|I)}{(h+\alpha(1-\alpha))^{1/2}} \right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\right) = o_{\mathbb{P}}\left(h^{d_{\mathcal{M}}/2+1}\right).$$

Lemmas B.3 and B.2-(i) then imply for the first item in $\widehat{\boldsymbol{R}}(d;\alpha,I)$, uniformly in α and d with $\|d\| = t$,

$$\begin{aligned}
&\left|\varrho_{\alpha L} d' \left[\left\{\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + u\widehat{\mathsf{e}}(\alpha|I);\alpha,I\right) - \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right\} du\right] \widehat{\mathsf{e}}(\alpha|I)\right| \\
&\quad = \left|\varrho_{\alpha L} d' \left[\int_0^1 \left\{\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + u\widehat{\mathsf{e}}(\alpha|I);\alpha,I\right) - \overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right.\right.\right. \\
&\qquad\qquad \left.\left.\left. + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)\right\} du\right] \widehat{\mathsf{e}}(\alpha|I)\right| \\
&\quad = \left|\varrho_{\alpha L} d' \left[O_{\mathbb{P}}\left(h^{-d_{\mathcal{M}}/2}\right)\|\widehat{\mathsf{e}}(\alpha|I)\| + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)\right] \widehat{\mathsf{e}}(\alpha|I)\right| \\
&\quad = t\left|\varrho_{\alpha L}\left[O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{2d_{\mathcal{M}}}}\right)^{1/2}\right) + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)\right] O_{\mathbb{P}}\left(\left(\frac{(h+\alpha(1-\alpha))\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\right)\right| \\
&\quad = t\varrho_{\alpha L} O_{\mathbb{P}}\left(\frac{(h+\alpha(1-\alpha))^{1/2}\log L}{Lh^{d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)/2}}\right) = t\varrho_{\alpha L}^2 O_{\mathbb{P}}(1).
\end{aligned}$$

Observe that the condition $\log L / \left(Lh^{2(d_{\mathcal{M}}+1)}\right) = o(1)$ implies

$$\frac{\log L}{Lh^{d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)}} = o(1) \text{ and them } \varrho_{\alpha L} = o\left(\left(\frac{(h+\alpha(1-\alpha))\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\right).$$

Lemmas B.3 and B.2 then imply for the second item in $\widehat{\boldsymbol{R}}(d;\alpha,I)$, uniformly in α and d with $\|d\| = t$,

$$\begin{aligned}
&\varrho_{\alpha L}^2 d' \left[\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I) + u\varrho_{\alpha L} d;\alpha,I\right)(1-u)\,du\right] d' \\
&\quad = \varrho_{\alpha L}^2 d' \left[\int_0^1 \left\{\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I) + u\varrho_{\alpha L} d;\alpha,I\right) + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)\right\}(1-u)\,du\right] d' \\
&\quad = \varrho_{\alpha L}^2 d' \left[\int_0^1 \left\{\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I)\right) + tO_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{2d_{\mathcal{M}}}}\right)^{1/2}\right) + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)\right\}(1-u)\,du\right] d' \\
&\quad \geq C\varrho_{\alpha L}^2 t^2 \left(1 + to_{\mathbb{P}}(1)\right).
\end{aligned}$$

Now (B.8) gives, with $O_{\mathbb{P}}(1)$ and $o_{\mathbb{P}}(1)$ which are uniform in α,

$$\begin{aligned}\mathbb{P}\left(\sup_{\alpha\in[0,1]}\left\|\frac{\widehat{\mathsf{d}}(\alpha|I)}{\varrho_{\alpha L}}\right\|\geq t\right) &\leq \mathbb{P}\left(\inf_{\alpha\in[0,1]}\left\{C\varrho_{\alpha L}^2 t^2(1+to_{\mathbb{P}}(1))+t\varrho_{\alpha L}^2 O_{\mathbb{P}}(1)\right\}\leq 0\right)\\ &= \mathbb{P}\left(Ct(1+to_{\mathbb{P}}(1))+O_{\mathbb{P}}(1)\leq 0\right)\\ &\leq \mathbb{P}\left(t(1+to_{\mathbb{P}}(1))\leq |O_{\mathbb{P}}(1)|\right)\end{aligned}$$

which can be made as small as needed asymptotically by increasing t. This gives the first result of the Theorem. For the second and third, observe that $\max_{\alpha\in[0,1]}\varrho_{\alpha L}=\log L/Lh^{d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)/2}$ so that, uniformly in α and x,

$$\begin{aligned}\left|\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}P(x)'\widehat{\mathsf{d}}_0(\alpha|I)\right| &= (Lh)^{1/2}h^{d_{\mathcal{M}}/2}\max_{x\in\mathcal{X}}\|P(x)\|\left\|\widehat{\mathsf{d}}(\alpha|I)\right\|\\ &= O_{\mathbb{P}}\left((Lh)^{1/2}\varrho_{\alpha L}\right)=O_{\mathbb{P}}\left(\frac{h^{1/2}\log L}{\left(Lh^{2d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)}\right)^{1/2}}\right),\\ \left|\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}P(x)'\frac{\widehat{\mathsf{d}}_1(\alpha|I)}{h}\right| &= O_{\mathbb{P}}\left(\left(\frac{L}{h}\right)^{1/2}\varrho_{\alpha L}\right)=O_{\mathbb{P}}\left(\frac{\log L}{\left(Lh^{2d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)+1}\right)^{1/2}}\right).\end{aligned}$$

This ends the proof of the Theorem. □

B.4 Proof of main estimation theorems

Proof of Theorem 4. Recall that s_1 is the row vector $[0,1,0,\ldots,0]$ of dimension $s+2$ and let $s_0=[1,0,\ldots,0]$, $S_0=s_0\otimes\mathrm{Id}_{K_L}$, $S_1=s_1\otimes\mathrm{Id}_{K_L}$ so that $\widehat{\beta}_j(\alpha|I)=S_j\widehat{\beta}(\alpha|I)$, $j=0,1$ and

$$\begin{aligned}\widehat{V}(\alpha|x,I) &= P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{b}}(\alpha|I),\\ \overline{V}(\alpha|x,I) &= P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\overline{\mathsf{b}}(\alpha|I)\end{aligned}$$

Define, for $\widehat{\mathsf{e}}(\alpha|I)$ as in (B.6)

$$\widetilde{V}(\alpha|x,I)=\overline{V}(\alpha|x,I)+P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{e}}(\alpha|I)$$

which is such, for $\widehat{\mathsf{d}}(\alpha|I)$ as in (B.7),

$$\widehat{V}(\alpha|x,I)-\widetilde{V}(\alpha|x,I)=P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{d}}(\alpha|I).$$

As the eigenvalues of $\int_{\mathcal{X}} P(x) P(x)' dx$ are bounded away from infinity under Assumption R-(i)

$$\int_{\mathcal{X}} \int_0^1 \left(\widehat{V}(\alpha|x,I) - \widetilde{V}(\alpha|x,I)\right)^2 d\alpha dx = \frac{O\left(\sup_{\alpha\in[0,1]} \left\|\widehat{\mathsf{d}}(\alpha|I)\right\|^2\right)}{h^2}$$
$$= O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1+(d_{\mathcal{M}}\vee 1)/2}}\right)^2\right)$$

by Theorem B.9, which gives (3.3) since, by Assumption H,

$$\frac{Lh^{d_{\mathcal{M}}+1}}{\log L}\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1+(d_{\mathcal{M}}\vee 1)/2}}\right)^2 = \frac{\log L}{Lh^{d_{\mathcal{M}}+1+(d_{\mathcal{M}}\vee 1)}} = o\left(\frac{\log L}{Lh^{2(d_{\mathcal{M}}+1)}}\right).$$

That $\mathsf{bias}_{IL}^2 = O(1)$ and $\Sigma_{IL} = O(1)$ similarly follow from Assumption R-(i) and Proposition 3-(i).

It holds since $\mathbb{E}[\widehat{\mathsf{e}}(\alpha|I)] = \overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) = 0$ for all α in $[0,1]$

$$\mathbb{E}\left[\int_{\mathcal{X}}\int_0^1 \left(\widetilde{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx\right] = \int_{\mathcal{X}}\int_0^1 \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx$$
$$+ \int_{\mathcal{X}}\int_0^1 \mathbb{E}\left[\left(P(x)'\left[S_0 + \frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{e}}(\alpha|I)\right)^2\right] d\alpha dx.$$

For the bias part, Theorem B.8 gives

$$\int_{\mathcal{X}}\int_0^1 \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx = \int_{\mathcal{X}}\int_0^1 \left(\frac{h^{s+1}P(x)'\alpha\mathsf{bias}_{1h}(\alpha|I)}{I-1} + o\left(h^{s+1}\right)\right)^2 d\alpha dx$$
$$= h^{2(s+1)}\int_{\mathcal{X}}\int_0^1 \left(\frac{P(x)'\alpha\mathsf{bias}_{1h}(\alpha|I)}{I-1}\right)^2 d\alpha dx + o\left(h^{2(s+1)}\right),$$

Since $\alpha\mathsf{bias}_{1h}(\alpha|I)/(I-1)$ differs from $\mathsf{bias}(\alpha|I)$ for α in $[0,h]$ or $[1-h,1]$, it follows

$$\int_{\mathcal{X}}\int_0^1 \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx = h^{2(s+1)}\int_{\mathcal{X}}\int_0^1 \left(P(x)'\mathsf{bias}(\alpha|I)\right)^2 d\alpha dx + o\left(h^{2(s+1)}\right)$$
$$= h^{2(s+1)}\mathsf{bias}_{IL}^2 + o\left(h^{2(s+1)}\right).$$

Arguing similarly with Lemma B.5-(i) yields

$$
\begin{aligned}
&\int_{\mathcal{X}} \int_0^1 \mathbb{E}\left[\left(P(x)'\left[S_0 + \frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{e}}(\alpha|I)\right)^2\right] d\alpha dx \\
&= \int_{\mathcal{X}} \int_0^1 \mathbb{E}\left[\left(\left[\frac{P(x)' \alpha \widehat{\mathsf{e}}_1(\alpha|I)}{h(I-1)}\right]\right)^2\right] d\alpha dx + O\left(\frac{1}{Lh^{d_{\mathcal{M}}}}\right) \\
&= \frac{\sigma_{LI}^2}{LIh^{d_{\mathcal{M}}+1}} + o\left(\frac{1}{Lh^{d_{\mathcal{M}}+1}}\right).
\end{aligned}
$$

Substituting in the bias-variance decomposition of the integrated mean squared error ends the proof of the Theorem. □

Proof of Theorem 5. Assumption R-(i) and Proposition 3-(i) imply that $P(x)' \Sigma_h(\alpha|I) P(x) = 0$ holds only if $P(x) = 0$, which is impossible in the AQR case. But, in the ASQR case, if $P(x) = 0$ for some $x \in \mathcal{X}$ and all K_L large enough, the approximation property S cannot hold, contradicting Assumption S-(ii). Assumptions R-(i), H and Proposition 3-(i) imply

$$
\max_{x \in \mathcal{X}} \left(P(x)' \Sigma_h(\alpha|I) P(x)\right) = O\left(\max_{x \in \mathcal{X}} \|P(x)\|^2\right) = O\left(h^{-d_{\mathcal{M}}}\right).
$$

By Theorem B.9, Lemma B.5, Assumptions R-(i), H, and using the same notations than in the proof of Theorem 4

$$
\begin{aligned}
&\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}\left(\widehat{V}(\alpha|x,I) - V(\alpha|x,I) - \frac{P'(x)\alpha S_1 \widehat{\mathsf{e}}(\alpha|I)}{h(I-1)} - \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right)\right) \\
&= \left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}\left\{P'(x)\widehat{\mathsf{e}}_0(\alpha|I) + P'(x)\left[S_0 + \frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{d}}(\alpha|I)\right\} \\
&= \left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}\left\{O_{\mathbb{P}}\left(\frac{1}{\left(Lh^{d_{\mathcal{M}}}\right)^{1/2}}\right) + O\left(\frac{\left\|P(x)'\widehat{\mathsf{d}}(\alpha|I)\right\|}{h}\right)\right\} \\
&= O_{\mathbb{P}}\left(h^{1/2} + \left(\frac{\log^2 L}{Lh^{2d_{\mathcal{M}}-1+(d_{\mathcal{M}}\vee 1)}}\right)^{1/2}\right) = o_{\mathbb{P}}(1).
\end{aligned}
$$

Since $\overline{V}(\alpha|x,I) - V(\alpha|x,I) = h^{s+1}P(x)' \mathsf{Bias}_h(\alpha|I) + o\left(h^{s+1}\right)$, it remains to show that

$$
\left(\frac{LIh}{P(x)' \Sigma_h(\alpha|I) P(x)}\right)^{1/2} \frac{\alpha P(x)' S_1 \widehat{\mathsf{e}}(\alpha|I)}{h(I-1)} \xrightarrow{d} \mathcal{N}(0,1).
$$

Write

$$\left(\frac{LIh}{P(x)'\Sigma_h(\alpha|I)P(x)}\right)^{1/2}\frac{\alpha P(x)' S_1\widehat{\mathsf{e}}(\alpha|I)}{h(I-1)}=\sum_{\ell=1}^{L} r_\ell(\alpha|x,I)$$

with $r_\ell(\alpha|x,I)=\mathbb{I}(I_\ell=I)\sum_{i=1}^{I_\ell} r_{i\ell}(\alpha|x,I)$ and

$$\begin{aligned}r_{i\ell}(\alpha|x,I)=&\left(\frac{\alpha^2}{LIh(I-1)^2}\right)^{1/2}\frac{P(x)'}{\left(P(x)'\Sigma_h(\alpha|I)P(x)\right)^{1/2}}S_1\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|x,I);\alpha,I\right)\right]^{-1}\\&\times\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left\{\mathbb{I}\left(B_{i\ell}\le P(x_\ell,t)\overline{\mathsf{b}}(\alpha|x,I)\right)-(\alpha+ht)\right\}P(x_\ell,t)K(t)\,dt.\end{aligned}$$

Since $\mathbb{E}[r_\ell(\alpha|x,I)]=0$ and $\max_{1\le\ell\le L}|\mathrm{Var}(r_\ell(\alpha|x,I))-1|=o(1)$, it is sufficient to show that $\max_{1\le\ell\le L}|\mathbb{E}[r_\ell^3(\alpha|x,I)]|=o(1)$ holds, see e.g. Theorem <19> p.179 in Pollard (2002). But Assumption R-(i) and Proposition 3-(i), Lemma B.2 and (B.4),

$$|r_{i\ell}(\alpha|x,I)|\le\frac{C}{(Lh)^{1/2}}\frac{\|P(x)\|}{\|P(x)\|}\times\max_{x\in\mathcal{X}}\|P(x)\|=O\left(\frac{1}{(Lh^{d_{\mathcal{M}}+1})^{1/2}}\right).$$

It follows that by Assumption H

$$\begin{aligned}\max_{1\le\ell\le L}\left|\mathbb{E}\left[r_\ell^3(\alpha|x,I)\right]\right|&\le I\max_{1\le\ell\le L,1\le i\le I_\ell}|r_{i\ell}(\alpha|x,I)|\max_{1\le\ell\le L}\left|\mathbb{E}\left[r_\ell^2(\alpha|x,I)\right]\right|\\&=O\left(\frac{1}{(Lh^{d_{\mathcal{M}}+1})^{1/2}}\right)=o(1).\end{aligned}$$

This ends the proof of the Theorem. □

Proof of Theorem 6. By Theorems B.8 and B.9, Lemma B.5 and using the notations of the proof of Theorem 4

$$\begin{aligned}&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\widehat{B}(\alpha|x,I)-B(\alpha|x,I)\right|\\&\le\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P(x)'S_0\left[\widehat{\mathsf{b}}(\alpha|I)-\overline{\mathsf{b}}(\alpha|I)\right]\right|+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\overline{B}(\alpha|x,I)-B(\alpha|x,I)\right|\\&\le\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P(x)'\widehat{\mathsf{e}}_0(\alpha|I)\right|+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left\|P(x)'\widehat{\mathsf{d}}_0(\alpha|I)\right\|+o\left(h^{s+1}\right)\\&=O_{\mathbb{P}}\left[\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\left\{1+\left(\frac{\log L}{Lh^{2d_{\mathcal{M}}+(d_{\mathcal{M}}\vee 1)}}\right)^{1/2}\right\}\right]+o\left(h^{s+1}\right)\\&=O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}\right)+o\left(h^{s+1}\right)\end{aligned}$$

$$
\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|\widehat{V}(\alpha|x,I) - V(\alpha|x,I)\right| \\
&\quad\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|P(x)'\left(S_0+\frac{\alpha}{h}S_1\right)\left[\widehat{\mathsf{b}}(\alpha|I)-\overline{\mathsf{b}}(\alpha|I)\right]\right| + \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\overline{V}(\alpha|x,I)-V(\alpha|x,I)\right| \\
&\quad\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|P(x)'\widehat{\mathsf{e}}_0(\alpha|I)\right| + \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P(x)'\frac{\widehat{\mathsf{e}}_1(\alpha|I)}{h}\right| \\
&\qquad + \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left\|P(x)'\left(\widehat{\mathsf{d}}_0+\alpha\frac{\widehat{\mathsf{d}}_1(\alpha|I)}{h}\right)\right\| + O\left(h^{s+1}\right) \\
&\quad= O_{\mathbb{P}}\left[\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\left\{1+\left(\frac{\log L}{Lh^{2d_{\mathcal{M}}+1+(d_{\mathcal{M}}\vee 1)}}\right)^{1/2}\right\}\right] + O\left(h^{s+1}\right) \\
&\quad= O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right) + O\left(h^{s+1}\right).
\end{aligned}
$$

This end the proof of the Theorem. $\square$

Proof of Corollary 7. Let

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

Then, since $\widehat{A}_{i\ell} = \arg\min_{\alpha\in[0,1]}\left|B_{i\ell}-\widehat{B}(\alpha|x_\ell,I_\ell)\right|$ and $B_{i\ell} = B(A_{i\ell}|x_\ell,I_\ell)$ where $B(\cdot|x_\ell,I_\ell)$ is increasing,

$$
\begin{aligned}
&\left\{\left|\widehat{A}_{i\ell}-A_{i\ell}\right|\geq t\right\} \\
&\subset\left\{\min_{\alpha\in[0,1];|A_{i\ell}-\alpha|\geq t}\left|B_{i\ell}-\widehat{B}(\alpha|x_\ell,I_\ell)\right| \leq \min_{\alpha\in[0,1];|A_{i\ell}-\alpha|<t}\left|B_{i\ell}-\widehat{B}(\alpha|x_\ell,I_\ell)\right|\right\} \\
&\subset\left\{\min_{\alpha\in[0,1];|A_{i\ell}-\alpha|\geq t}\left|B_{i\ell}-\widehat{B}(\alpha|x_\ell,I_\ell)\right| \leq \left|B(A_{i\ell}|x_\ell,I_\ell)-\widehat{B}(A_{i\ell}|x_\ell,I_\ell)\right|\right\} \\
&\subset\left\{\min_{\alpha\in[0,1];|A_{i\ell}-\alpha|\geq t}\left|B(A_{i\ell}|x_\ell,I_\ell)-B(\alpha|x_\ell,I_\ell)\right| - \widehat{\varrho}_B \leq \widehat{\varrho}_B\right\} \\
&\subset\left\{t\times\min_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} B^{(1)}(\alpha|x,I)\leq 2\widehat{\varrho}_B\right\},
\end{aligned}
$$

and Proposition 3 implies that, for $C > 2/\min_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} B^{(1)}(\alpha|x,I)$,

$$
\max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|\widehat{A}_{i\ell}-A_{i\ell}\right|\leq C\widehat{\varrho}_B.
$$

Hence, since $\max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x,I) < \infty$,

$$\begin{aligned}
&\max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|\widehat{V}_{i\ell}-V_{i\ell}\right| = \max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|\widehat{V}\left(\widehat{A}_{i\ell}|x_\ell,I_\ell\right)-V\left(A_{i\ell}|x_\ell,I_\ell\right)\right| \\
&\leq \max_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\widehat{V}(\alpha|x,I)-V(\alpha|x,I)\right| + \max_{\ell=1,\ldots,L}\max_{i=1,\ldots,I_\ell}\left|V\left(\widehat{A}_{i\ell}|x_\ell,I_\ell\right)-V\left(A_{i\ell}|x_\ell,I_\ell\right)\right| \\
&\leq \widehat{\varrho}_V + C\widehat{\varrho}_B.
\end{aligned}$$

This gives the rate stated in the Corollary. □

References

[1] Gragg, W.B. & R.A. Tapia (1974). Optimal error bounds for the Newton-Kantorovich Theorem. *SIAM Journal on Numerical Analysis* **11**, 10–13.

[2] Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory*, **7**, 186–199.

[3] Pollard, D. (2002). *A user's guide to measure theoretic probability.* Cambridge University Press.

Appendix C: Proofs of intermediary estimation results

C.1 Lemmas B.1, B.2 and B.7

Proof of Lemma B.1. Consider the harder $ASQR$ case. (i) It holds that, for $\beta_k(\cdot|\cdot)$ as in (2.14),

$$\begin{aligned}
&B(\alpha+ht|x,I) - P(x,t)' \mathsf{b}^*(\alpha|I) \\
&\quad = B(\alpha+ht|x,I) - \sum_{k=1}^{K_L} P_k(x)\beta_k(\alpha+ht|I) \\
&\quad + \sum_{k=1}^{K_L} P_k(x)\beta_k(\alpha+ht|I) - \sum_{k=1}^{K_L} P_k(x)\sum_{p=0}^{s+1}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I) \\
&\quad = B(\alpha+ht|x,I) - \sum_{k=1}^{K_L} P_k(x)\beta_k(\alpha+ht|I) \\
&\quad + \sum_{k=1}^{K_L} P_k(x)\left(\beta_k(\alpha+ht|I) - \sum_{p=0}^{s}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I)\right) - \frac{(ht)^{s+1}}{(s+1)!}\sum_{k=1}^{K_L} P_k(x)\beta_k^{(s+1)}(\alpha|I).
\end{aligned}$$

A Taylor expansion with integral remainder gives

$$\beta_k(\alpha+ht|I) - \sum_{p=0}^{s}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I) = \frac{(ht)^{s+1}}{s!}\int_0^1 \beta_k^{(s+1)}(\alpha+uht|I)(1-u)^s\,du$$

so that

$$\begin{aligned}
&B(\alpha+ht|x,I) - P(x,t)' \mathsf{b}^*(\alpha|I) \\
&\quad = B(\alpha+ht|x,I) - \sum_{k=1}^{K_L} P_k(x)\beta_k(\alpha+ht|I) \\
&\quad + \frac{(ht)^{s+1}}{s!}\int_0^1 \left\{\sum_{k=1}^{K_L} P_k(x)\beta_k^{(s+1)}(\alpha+uht|I) - B^{(s+1)}(\alpha+uht|I)\right\}(1-u)^s\,du \\
&\quad + \frac{(ht)^{s+1}}{s!}\int_0^1 \left\{B^{(s+1)}(\alpha+uht|x,I) - B^{(s+1)}(\alpha|x,I)\right\}(1-u)^s\,du \\
&\quad + \frac{(ht)^{s+1}}{(s+1)!}\left\{B^{(s+1)}(\alpha|x,I) - \sum_{k=1}^{K_L} P_k(x)\beta_k^{(s+1)}(\alpha|x,I)\right\}.
\end{aligned}$$

Hence since $B^{(s+1)}(\alpha|x,I)$ is continuous, by Property S and Proposition 3

$$\begin{aligned}\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|B(\alpha+ht|x,I)-P(x,t)\,\mathsf{b}^*(\alpha|I)\right| &= o\left(h^{s+1}\right)+o\left(K_L^{-\frac{s+1}{d_{\mathcal{M}}}}\right)\\ &= o\left(h^{s+1}\right) \qquad \text{(C.1)}\end{aligned}$$

since $K_L^{-1/d_{\mathcal{M}}}=O(h)$. Observe also that, uniformly in α, x and t as above,

$$\begin{aligned}\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}^*(\alpha|I)\right] &= \sum_{p=1}^{s+1}h^p\frac{t^{p-1}}{(p-1)!}\sum_{k=1}^{K_L}P_k(x)\,\beta_k^{(p)}(\alpha|I)\\ &= h\left(B^{(1)}(\alpha|x,I)+o(1)\right)+h^2\left(\sum_{p=2}^{s+1}h^{p-2}\frac{t^{p-1}}{(p-1)!}B^{(p)}(\alpha|x,I)+o(1)\right)\\ &= hB^{(1)}(\alpha|x,I)+o(h)\end{aligned}$$

by Property S, which also gives,

$$\begin{aligned}&\max_{p=1,\ldots,s+1}\left(\frac{\max_{x\in\mathcal{X}}\left|P(x)'\,\mathsf{b}_p^*(\alpha|I)\right|}{h}\right)=\max_{p=1,\ldots,s+1}\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}h^{p-1}\left|B^{(p)}(\alpha|x,I)+o(1)\right|\\ &\quad=\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}B^{(1)}(\alpha|x,I)+o(1)\leq\overline{f}\end{aligned}$$

provided $\overline{f}$ is large enough and h small enough, so that $\mathsf{b}^*(\alpha|I)$ is in $\underline{\mathcal{BI}}_{\alpha,h}$ since $B^{(1)}(\cdot|\cdot,\cdot)$ is bounded away from 0 and infinity by Proposition 3. Suppose now that $\|\mathsf{b}-\mathsf{b}^*(\alpha|I)\|\leq Ch/K_L^{1/2}=Ch^{d_{\mathcal{M}}/2+1}$. Then

$$\begin{aligned}\left|\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}\right]\right| &\geq \left|\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}^*(\alpha|I)\right]\right|-\|\mathsf{b}-\mathsf{b}^*(\alpha|I)\|\,\|P(x)\|\\ &\geq \left|\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}^*(\alpha|I)\right]\right|-O(h),\\ \left|P(x)'\,\mathsf{b}_p\right| &\leq \left|P(x)'\,\mathsf{b}_p^*(\alpha|I)\right|+\|\mathsf{b}-\mathsf{b}^*(\alpha|I)\|\,\|P(x)\|\\ &\leq \left|P(x)'\,\mathsf{b}_p^*(\alpha|I)\right|-Ch,\quad p=1,\ldots,s+1,\end{aligned}$$

and $\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)\subset\underline{\mathcal{BI}}_{\alpha,h}$ when h is small enough provided C is small enough. Hence (i) holds. (ii) follows from the Implicit Function Theorem and the definition of $\mathcal{BI}_{\alpha,h}$.

The first equality of (iii) is (C.1). For the second, note that $\alpha+ht\geq h>0$ when $\alpha\geq 3h$

for all t in $\mathcal{I}_{\alpha,h}$. It holds

$$\begin{aligned}
& B\left(\alpha+ht|x,I\right)-P\left(x,t\right)'\mathsf{b}^{*}\left(\alpha|I\right) \\
&\quad = B\left(\alpha+ht|x,I\right)-\sum_{k=1}^{K_L}P_k\left(x\right)\beta_k\left(\alpha+ht|I\right) \\
&\quad +\sum_{k=1}^{K_L}P_k\left(x\right)\left(\beta_k\left(\alpha+ht|I\right)-\sum_{p=0}^{s+1}\frac{(ht)^p}{p!}\beta_k^{(p)}\left(\alpha|I\right)\right)
\end{aligned}$$

with

$$\beta_k\left(\alpha+ht|I\right)-\sum_{p=0}^{s+1}\frac{(ht)^p}{p!}\beta_k^{(p)}\left(\alpha|I\right)=\frac{(ht)^{s+2}}{(s+1)!}\int_0^1\beta_k^{(s+2)}\left(\alpha+uht|I\right)(1-u)^{s+1}\,du$$

recalling, as established in the proof of Proposition 3-(i) for $\alpha>0$,

$$\begin{aligned}
\beta_k^{(s+2)}\left(\alpha|I\right) &= \frac{1}{\alpha}\Big((I-1)\,\gamma_k^{(s+1)}\left(\alpha|I\right)-(I+s)\,\beta_k^{(s+1)}\left(\alpha|I\right)\Big), \\
B^{(s+2)}\left(\alpha|x,I\right) &= \frac{1}{\alpha}\Big((I-1)\,V_k^{(s+1)}\left(\alpha|I\right)-(I+s)\,B^{(s+1)}\left(\alpha|x,I\right)\Big).
\end{aligned} \tag{C.2}$$

Hence

$$\begin{aligned}
& B\left(\alpha+ht|x,I\right)-P\left(x,t\right)'\mathsf{b}^{*}\left(\alpha|I\right)-\frac{(ht)^{s+2}}{(s+2)!}B^{(s+2)}\left(\alpha|I\right) \\
&= B\left(\alpha+ht|x,I\right)-\sum_{k=1}^{K_L}P_k\left(x\right)\beta_k\left(\alpha+ht|I\right) \\
&+\frac{(ht)^{s+2}}{(s+1)!}\int_0^1\left\{\sum_{k=1}^{K_L}P_k\left(x\right)\beta_k^{(s+2)}\left(\alpha+uht|I\right)-B^{(s+2)}\left(\alpha+uht|x,I\right)\right\}(1-u)^{s+1}\,du \\
&+\frac{(ht)^{s+2}}{(s+1)!}\int_0^1\left\{B^{(s+2)}\left(\alpha+uht|x,I\right)-B^{(s+2)}\left(\alpha|x,I\right)\right\}(1-u)^{s+1}\,du,
\end{aligned}$$

with, using the expressions $\beta_k^{(s+2)}\left(\cdot|\cdot\right)$ and $B^{(s+2)}\left(\cdot|\cdot\right)$ of the proof of Proposition 3

$$\max_{(\alpha,x)\in[0,3h]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\left(B\left(\alpha+ht|x,I\right)-\sum_{k=1}^{K_L}P_k\left(x\right)\beta_k\left(\alpha+ht|I\right)\right)\right|=ho\left(K_L^{-\frac{s+1}{d_{\mathcal{M}}}}\right)=o\left(h^{s+2}\right),$$

$$\max_{(\alpha,x)\in[3h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\int_0^1\left\{\sum_{k=1}^{K_L}P_k(x)\beta_k^{(s+2)}(\alpha+uht|I)-B^{(s+2)}(\alpha+uht|x,I)\right\}(1-u)^{s+1}du\right|$$
$$\leq C\max_{(\alpha,x)\in[2h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left\{\frac{\alpha}{\alpha-h}\left|\sum_{k=1}^{K_L}P_k(x)\beta_k^{(s+1)}(\alpha|I)-B(\alpha|x,I)\right|\right\}$$
$$+C\max_{(\alpha,x)\in[2h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left\{\frac{\alpha}{\alpha-h}\left|\sum_{k=1}^{K_L}P_k(x)\gamma_k^{(s+1)}(\alpha|I)-V(\alpha|x,I)\right|\right\}=o(1),$$
$$\max_{(\alpha,x)\in[3h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\int_0^1\left\{B^{(s+2)}(\alpha+uht|x,I)-B^{(s+2)}(\alpha|x,I)\right\}(1-u)^{s+1}du\right|=o(1).$$

Substituting gives

$$\max_{(\alpha,x)\in[3h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\left(B(\alpha+ht|x,I)-P(x,t)'\mathsf{b}^*(\alpha|I)-\frac{(ht)^{s+2}}{(s+2)!}B^{(s+2)}(\alpha|x,I)\right)\right|=o\left(h^{s+2}\right)$$

which implies the second statement in (iii) since by Proposition 3-(i) and (B.4)

$$\max_{(\alpha,x)\in[0,3h]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\left(B(\alpha+ht|x,I)-P(x,t)'\mathsf{b}^*(\alpha|I)\right)\right|=o\left(h^{s+2}\right),$$
$$\max_{(\alpha,x)\in[0,3h]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\frac{(ht)^{s+2}}{(s+2)!}B^{(s+2)}(\alpha|x,I)\right|=o\left(h^{s+2}\right).$$

The third result in (iii) follows from Proposition 3-(iii). The fourth equality of (iii) follows from

$$\begin{aligned}
o\left(h^{s+1}\right)&=\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\Psi(t|x,\mathsf{b}^*(\alpha|I))-B(\alpha+ht|x,I)\right|\\
&=\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|\Psi\left[\Delta(u|x,\mathsf{b}^*(\alpha|I))|x,\mathsf{b}^*(\alpha|I)\right]\right.\\
&\qquad\left.-B\left[\alpha+h\Delta(u|x,\mathsf{b}^*(\alpha|I))|x,I\right]\right|\\
&=\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|u-B\left[\alpha+h\Delta(u|x,\mathsf{b}^*(\alpha|I))|x,I\right]\right|\\
&=\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|B\left[\alpha+h\frac{G(u|x,I)-\alpha}{h}|x,I\right]\right.\\
&\qquad\left.-B\left[\alpha+h\Delta(u|x,\mathsf{b}^*(\alpha|I))|x,I\right]\right|\\
&\geq Ch\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|\frac{G(u|x,I)-\alpha}{h}-\frac{\Phi(u|x,\mathsf{b}^*(\alpha|I))-\alpha}{h}\right|
\end{aligned}$$

by Proposition 3-(i).

Consider now (iv). The first bound follows from the Cauchy-Schwarz inequality. This

bound implies for all u in $\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]$

$$\begin{aligned}
&\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-\Psi\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right|\\
&\quad=\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-u\right|\\
&\quad=\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]\right|\leq Ch^{-d_{\mathcal{M}}/2}\left\|\mathsf{b}_1-\mathsf{b}_0\right\|.
\end{aligned}$$

By definition of $\underline{\mathcal{BI}}_{\alpha,h}$

$$\begin{aligned}
&\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-\Psi\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right|\\
&\quad\geq Ch\left|\Delta\left(u|x,\mathsf{b}_1\right)-\Delta\left(u|x,\mathsf{b}_0\right)\right|=C\left|\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)\right|
\end{aligned}$$

and substituting shows that the second bound of (iv) holds. For the third bound in (iv), it holds uniformly in α, x, u, b_1 and b_0

$$\begin{aligned}
&\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right|\\
&\quad\leq\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_1\right]\right|\\
&\quad+\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right|\\
&\quad\leq\max_{t\in\mathcal{I}_{\alpha,h}}\left|\frac{\partial^2\Psi\left(t|x,\mathsf{b}_1\right)}{\partial t^2}\right|\frac{\left|\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)\right|}{h}\\
&\quad+\max_{t\in\mathcal{I}_{\alpha,h}}\left|\frac{\partial P\left(x,t\right)}{\partial t}\left(\mathsf{b}_1-\mathsf{b}_0\right)\right|.
\end{aligned}$$

But, by definition of $\underline{\mathcal{BI}}_{\alpha,h}$

$$\max_{t\in\mathcal{I}_{\alpha,h}}\left|\frac{\partial^2\Psi\left(t|x,\mathsf{b}_1\right)}{\partial t^2}\right|\leq Ch\max_{p=2,\ldots,s+1}\left|\frac{P\left(x\right)\mathsf{b}_{1p}}{h}\right|=O\left(h\right)$$

so that substituting and the bound for $\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)$ gives, uniformly in α, x, u, b_1 and b_0

$$\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right|\leq Ch^{-d_{\mathcal{M}}/2}\left\|\mathsf{b}_1-\mathsf{b}_0\right\|,$$

which is the fourth inequality. The expression in (ii) of $\Phi\left(\cdot\right)$ and the definition of $\underline{\mathcal{BI}}_{\alpha,h}$ yield the third inequality. □

Proof of Lemma B.2. It holds

$$\begin{aligned}
\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) &= \mathbb{E}\Big[\mathbb{I}\left[B_{i\ell}\in\Psi\left(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}\right),I_\ell=I\right] \\
&\qquad \frac{P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)'}{\Psi\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)}K\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)\Big] \\
&= \int\left[\int_{\Psi\left(\underline{I}_{\alpha,h}|x,\mathsf{b}\right)\vee B(0|x,I)}^{\Psi\left(\overline{I}_{\alpha,h}|x,\mathsf{b}\right)\wedge B(1|x,I)}\frac{P\left(x,\Delta\left(y|x,\mathsf{b}\right)\right)P\left(x,\Delta\left(y|x,\mathsf{b}\right)\right)'}{\Psi\left(\Delta\left(y|x,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)}K\left(\Delta\left(y|x,\mathsf{b}\right)\right)g\left(y,x,I\right)dy\right]dx.
\end{aligned}$$

Recall $\Delta\left[\Psi\left[t|x,\mathsf{b}\right]|x,\mathsf{b}\right]=t$ for all t in $\mathcal{I}_{\alpha,h}$ and let

$$\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)=\overline{I}_{\alpha,h}\wedge\Delta\left[B\left(1|x,I\right)|x,\mathsf{b}\right],\quad \underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)=\underline{I}_{\alpha,h}\vee\Delta\left[B\left(0|x,I\right)|x,\mathsf{b}\right].$$

The change of variable $y=\Psi\left(t|x,\mathsf{b}\right)$ yields that

$$\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)=\int\left[\int_{\underline{I}_{\alpha,h}(x,I;\mathsf{b})}^{\overline{I}_{\alpha,h}(x,I;\mathsf{b})}P\left(x,t\right)P\left(x,t\right)'K\left(t\right)g\left(\Psi\left(t|x,\mathsf{b}\right),x,I\right)dt\right]dx.$$

The Dominated Convergence Theorem and Proposition 3-(i)[1], $s\geq 1$, yield that $\overline{\mathsf{R}}^{(2)}(\cdot;\alpha,I)$ is continuously differentiable over $\underline{\mathcal{BI}}_{\alpha,h}$ with, by the Liebniz integral rule,

$$\begin{aligned}
\overline{\mathsf{R}}^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \overline{\mathsf{R}}_0^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]+\overline{\mathsf{R}}_1^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]-\overline{\mathsf{R}}_2^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}], \\
\overline{\mathsf{R}}_0^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \int_{\mathcal{X}}\left[\int_{\underline{I}_{\alpha,h}(x,I;\mathsf{b})}^{\overline{I}_{\alpha,h}(x,I;\mathsf{b})}P\left(x,t\right)P\left(x,t\right)'K\left(t\right)g^{(1)}\left(\Psi\left(t|x,\mathsf{b}\right),x,I\right)\left[\mathsf{d}'P\left(x,t\right)\right]dt\right]dx, \\
\overline{\mathsf{R}}_1^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \int_{\mathcal{X}}P\left(x,\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)P\left(x,\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)'K\left(\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right) \\
&\qquad\times g\left(\Psi\left(\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)|x,\mathsf{b}\right),x,I\right)\left[\mathsf{d}'\frac{\partial\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)}{\partial\mathsf{b}'}\right]dx, \\
\overline{\mathsf{R}}_2^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \int_{\mathcal{X}}P\left(x,\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)P\left(x,\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)'K\left(\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right) \\
&\qquad\times g\left(\Psi\left(\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)|x,\mathsf{b}\right),x,I\right)\left[\mathsf{d}'\frac{\partial\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)}{\partial\mathsf{b}'}\right]dx.
\end{aligned}$$

Proposition 3-(i) and Assumption R-(i) imply

$$\left\|\overline{\mathsf{R}}_0^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]\right\|\preceq C\max_{x\in\mathcal{X}}\left\|P\left(x\right)\right\|\left\|\mathsf{d}\right\|\leq Ch^{-d_{\mathcal{M}}/2}\left\|\mathsf{d}\right\|.$$

[1] which implies that $g\left(\cdot|\cdot,I\right)$ is bounded away from 0 and infinity.

The operators $\overline{\mathsf{R}}_i^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]$, $i=1,2$, can be studied in a similar way so that only $i=1$ is considered. Observe

$$\frac{\partial \overline{I}_{\alpha,h}(x,I;\mathsf{b})}{\partial \mathsf{b}'}=\left\{\begin{array}{ll} 0 & \text{if } \overline{I}_{\alpha,h}\leq \Delta\left[B\left(1|x,I\right)|x,\mathsf{b}\right] \\ \frac{\partial \Delta[B(1|x,I)|x,\mathsf{b}]}{\partial \mathsf{b}'}=-\frac{P(x,\Delta(B(1|x,I)|x,\mathsf{b}))}{\Psi^{(1)}(\Delta(B(1|x,I)|x,\mathsf{b})|x,\mathsf{b})} & \text{if } \overline{I}_{\alpha,h}> \Delta\left[B\left(1|x,I\right)|x,\mathsf{b}\right] \end{array}\right. .$$

But, for h small enough,

$$\begin{aligned}
\Delta\left[B\left(1|x,I\right)|x,\mathsf{b}\right] &= \frac{\Phi\left[B\left(1|x,I\right)|x,\mathsf{b}\right]-\alpha}{h}=\frac{\min\left\{\alpha+h\overline{I}_{\alpha,h},\Phi\left[B\left(1|x,I\right)|x,\mathsf{b}\right]\right\}-\alpha}{h} \\
&\geq \frac{\min\left\{\alpha+h\overline{I}_{\alpha,h},\Phi\left[B\left(1|x,I\right)|x,\mathsf{b}^*\left(\alpha|I\right)\right]-Ch^{-d_{\mathcal{M}}/2}\left\|\mathsf{b}-\mathsf{b}^*\left(\alpha|I\right)\right\|\right\}-\alpha}{h} \\
&\geq \frac{\min\left\{\alpha+h\overline{I}_{\alpha,h},G\left[B\left(1|x,I\right)|x,I\right]-Ch^{s+1}-Ch\right\}-\alpha}{h} \\
&\geq \frac{\min\left\{\alpha+h\min\left(\frac{1-\alpha}{h},1\right),1-Ch\right\}-\alpha}{h}
\end{aligned}$$

uniformly in α, x and b in $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),Ch^{d_{\mathcal{M}}/2+1}\right)$ by Lemma B.1. Hence, if $\alpha\leq 1-C'h$ with $C'\geq 1$ large enough

$$\Delta\left[B\left(1|x,I\right)|x,\mathsf{b}\right]\geq\frac{\min\left\{\alpha+h,1-Ch\right\}-\alpha}{h}\geq 1\geq \overline{I}_{\alpha,h}$$

so that $\frac{\partial \overline{I}_{\alpha,h}(x,I;\mathsf{b})}{\partial \mathsf{b}'}=0$. Hence since $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),Ch^{d_{\mathcal{M}}/2+1}\right)\subset\underline{\mathcal{BI}}_{\alpha,h}$ and by definition of $\underline{\mathcal{BI}}_{\alpha,h}$

$$\begin{aligned}
\left\|\overline{\mathsf{R}}_1^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]\right\| &\leq C\mathbb{I}\left[\alpha\geq 1-C'h\right] \\
&\quad\times\left\|\int_{\mathcal{X}}P\left(x,\overline{I}_{\alpha,h}(x,I;\mathsf{b})\right)P\left(x,\overline{I}_{\alpha,h}(x,I;\mathsf{b})\right)'\frac{\mathsf{d}'P\left(x,\Delta\left(B\left(1|x,I\right)|x,\mathsf{b}\right)\right)}{\Psi\left(\Delta\left(B\left(1|x,I\right)|x,\mathsf{b}\right)|x,\mathsf{b}\right)}dx\right\| \\
&\leq Ch^{-1}\mathbb{I}\left[\alpha\geq 1-C'h\right]\max_{x\in\mathcal{X}}\left\|P(x)\right\|\left\|\mathsf{d}\right\|\leq Ch^{-1}h^{-d_{\mathcal{M}}/2}\left\|\mathsf{d}\right\|\mathbb{I}\left[\alpha\geq 1-C'h\right] \\
&\leq C\frac{h^{-d_{\mathcal{M}}/2}}{\alpha(1-\alpha)+h}\left\|\mathsf{d}\right\|.
\end{aligned}$$

Substituting in the expression of $\overline{\mathsf{R}}^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]$ then gives uniformly in d

$$\max_{\alpha\in[0,1]}\max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)}\left(\alpha(1-\alpha)+h\right)\left\|\overline{\mathsf{R}}^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]\right\|\leq Ch^{-d_{\mathcal{M}}/2}\left\|\mathsf{d}\right\|.$$

The Taylor inequality shows that (i) holds.

For (ii), the expression of $\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)$, Assumptions A and R-(i), Proposition 3-(i), which imply that the eigenvalues of $\int P(x)P'(x)g\left[B\left(\alpha|x,I\right),x,I\right]dx$ stay bounded away 0 and

infinity, Lemma B.1-(iii) and Proposition 3-(i) give that, uniformly in α and x

$$\begin{aligned}
\overline{I}_{\alpha,h}\left[x,I;\mathsf{b}^*(\alpha|I)\right] &= \overline{I}_{\alpha,h}\wedge\frac{\Phi\left[B\left(1|x,I\right)|x,\mathsf{b}^*(\alpha|I)\right]-\alpha}{h}\\
&= \overline{I}_{\alpha,h}\wedge\frac{1+o\left(h^{s+1}\right)-\alpha}{h}=\overline{I}_{\alpha,h}+o\left(h^{s}\right),\\
\underline{I}_{\alpha,h}\left[x,I;\mathsf{b}^*(\alpha|I)\right] &= \underline{I}_{\alpha,h}+o\left(h^{s}\right),
\end{aligned}$$

$$\begin{aligned}
\overline{\mathsf{R}}^{(2)}\left(\mathsf{b}^*(\alpha|I);\alpha,I\right) &= \int\left[\int_{\underline{I}_{\alpha,h}[x,I;\mathsf{b}^*(\alpha|I)]}^{\overline{I}_{\alpha,h}[x,I;\mathsf{b}^*(\alpha|I)]}\pi(t)\,\pi(t)'\,K(t)\,g\left(\Psi\left(t|x,\mathsf{b}^*(\alpha|I)\right)|x,I\right)dt\right]\\
&\qquad P(x)\,P(x)'\,f(x,I)\otimes dx\\
&= \int\left[\int_{\underline{I}_{\alpha,h}+o(h^s)}^{\overline{I}_{\alpha,h}+o(h^s)}\pi(t)\,\pi(t)'\,K(t)\,g\left[B\left(\alpha+ht|x,I\right)+o\left(h^{s+1}\right)|x,I\right]dt\right]\\
&\qquad\otimes P(x)\,P(x)'\,f(x,I)\,dx\\
&= \int\left[\int_{\underline{I}_{\alpha,h}+o(h^s)}^{\overline{I}_{\alpha,h}+o(h^s)}\pi(t)\,\pi(t)'\,K(t)\left(\frac{1}{B^{(1)}\left(\alpha+ht|x,I\right)}+o\left(h^{s+1}\right)\right)dt\right]\\
&\qquad\otimes P(x)\,P(x)'\,f(x,I)\,dx\\
&= \int\left[\int_{\underline{I}_{\alpha,h}+o(h^s)}^{\overline{I}_{\alpha,h}+o(h^s)}\pi(t)\,\pi(t)'\,K(t)\left(\frac{1}{B^{(1)}\left(\alpha|x,I\right)}-ht\frac{B^{(2)}\left(\alpha|x,I\right)}{\left(B^{(1)}\left(\alpha|x,I\right)\right)^{2}}+o(h)\right)dt\right]\\
&\qquad\otimes P(x)\,P(x)'\,f(x,I)\,dx\\
&= \int\Omega_h(\alpha)\otimes\frac{P(x)\,P(x)'}{B^{(1)}\left(\alpha|x,I\right)}f(x,I)\,dx+\\
&\qquad-\int\Omega_{1h}(\alpha)\otimes\frac{P(x)\,P(x)'\,B^{(2)}\left(\alpha|x,I\right)}{\left(B^{(1)}\left(\alpha|x,I\right)\right)^{2}}f(x,I)\,dx+o(h)
\end{aligned}$$

where the last $o(h)$ term is with respect of the matrix norm. This together the fact that the eigenvalues of the matrices $\Omega_h(\alpha)$ and $\int_{\mathcal{X}} P(x)\,P(x)'\,dx$ are bounded away from 0 and infinity, the fact that $B^{(1)}(\alpha|x,I)$ is bounded away from 0 and infinity shows that (ii) holds.$\square$

Proof of Lemma B.7. Write $A_{\alpha,h}^{-1}=D_{\alpha,h}+B_{\alpha,h}$ where $D_{\alpha,h}$ is the diagonal of $A_{\alpha,h}^{-1}$ and $B_{\alpha,h}=A_{\alpha,h}^{-1}-D_{\alpha,h}$. Provided the series converges

$$A_{\alpha,h}=D_{\alpha,h}^{-1/2}\left\{\sum_{n=0}^{\infty}\left(D_{\alpha,h}^{-1/2}B_{\alpha,h}D_{\alpha,h}^{-1/2}\right)^{n}\right\}D_{\alpha,h}^{-1/2}.$$

Proposition 3-(i) and Assumption R-(i) ensure that the entries of $D_{\alpha,h}^{-1/2}$ are bounded in absolute value by $C < \infty$ for all α and L. It also gives

$$\left| \frac{\mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)} \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P_{k_1}(x_\ell)\, \pi_{p_1}(t)\, P_{k_2}(x_\ell)\, \pi_{p_2}(t)\, K(t)\, dt \right]}{\mathbb{E}^{1/2}\left[\frac{\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)} \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P_{k_1}^2(x_\ell)\, \pi_{p_1}^2(t)\, K(t)\, dt \right] \mathbb{E}^{1/2}\left[\frac{\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)} \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P_{k_2}^2(x_\ell)\, \pi_{p_2}^2(t)\, K(t)\, dt \right]} \right| \leq \varrho < 1$$

for all $1 \leq k_1, k_2 \leq K_L$ and $0 \leq p_1, p_2 \leq s+1$, that is all the entries of $D_{\alpha,h}^{-1/2} B_{\alpha,h} D_{\alpha,h}^{-1/2}$ are bounded by ϱ in absolute value. By Assumption R-(ii), the entries of $D_{\alpha,h}^{-1/2} B_{\alpha,h} D_{\alpha,h}^{-1/2}$ are bounded by the ones of $\varrho \operatorname{Id} \otimes (T' + T)$, where T is a lower $c/2$ band matrix with band entries equal to 1 and Id is the $(s+2) \times (s+2)$ identity matrix. Hence the absolute value of the entries of $A_{\alpha,h}$ are bounded by the entries of

$$C \operatorname{Id} \otimes \left(\sum_{n=\infty}^{\infty} \varrho^n \left(T^{n'} + T^n \right) \right).$$

Since T is a triangular $c-$band nilpotent matrix, it follows that $|A_{\alpha,h}(j_1, j_2)| \leq C \rho^{|j_2 - j_1|}$ with $0 < \varrho \leq \rho < 1$, for all α and L. It follows

$$\max_L \max_{\alpha \in [0,1]} \max_{1 \leq j_1 \leq (s+1)K_L} \sum_{j_2=1}^{(s+1)K_L} |A_{\alpha,h}(j_1, j_2)| \leq C \sum_n \rho^n < \infty$$

which ends the proof of the Lemma. □

C.2 Lemmas B.3, B.4 and B.5

The proofs of the lemmas grouped here make use of a deviation inequality from Massart (2007). Consider n independent random variables Z_ℓ and, for a known real function $\xi(z, \theta)$ separable with respect to $\theta \in \Theta$, $Z_\ell(\theta) = \xi(Z_\ell, \theta)$ where θ is a parameter. Let $\underline{\xi}(\cdot) \leq \overline{\xi}(\cdot)$ be two functions. A *bracket* $\left[\underline{\xi}, \overline{\xi}\right]$ is the set of all functions $\xi(\cdot)$ such that $\underline{\xi}(z) \leq \xi(z) \leq \overline{\xi}(z)$ for all z. The next proposition follows from Massart (2007, Theorem 6.8 and Corollary 6.9).

Proposition C.1 *Assume that* $\sup_{\theta \in \Theta} |Z_\ell(\theta)| \leq M_\infty$, $\sup_{\theta \in \Theta} \operatorname{Var}(Z_\ell(\theta)) \leq M_2^2$ *for all* ℓ *and that for any* $\epsilon > 0$ *there exists brackets* $\left[\underline{\xi}_j, \overline{\xi}_j\right] \subset [-b, b]$, $j = 1, \ldots, \exp(H(\epsilon))$, *such that*

$$\mathbb{E}\left[\left(\overline{\xi}_j(Z_i) - \underline{\xi}_j(Z_i) \right)^2 \right] \leq \frac{\epsilon^2}{2} \text{ and } \{\xi(z, \theta), \theta \in \Theta\} \subset \bigcup_{j=1}^{\exp(H(\epsilon))} \left[\underline{\xi}_j, \overline{\xi}_j\right].$$

Let

$$\mathcal{H}_L = 54 \int_0^{M_2/2} \sqrt{\min(L, H(\epsilon))} d\epsilon + \frac{2(M_\infty + M_2) H(M_2)}{L^{1/2}}.$$

Then, for any $t \in \left[0, 10L^{1/2}M_2/M_\infty\right]$,

$$\mathbb{P}\left(\sup_{\theta\in\Theta}\left|\sum_{i=1}^{n}\left\{Z_\ell\left(\theta\right)-\mathbb{E}\left[Z_\ell\left(\theta\right)\right]\right\}\right| \geq L^{1/2}\left\{\mathcal{H}_L+t\right\}\right) \leq 2\exp\left(-\frac{t^2}{25}\right).$$

Proof of Lemma B.3. Note that $\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)-\overline{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)$ is a $c\left(s+2\right)$-band matrix, so that the order of its matrix norm is the same than the order of its largest entry. The generic entry of $\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)-\overline{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)$ can be written as

$$\widehat{\mathsf{r}}\left(\mathsf{b};\alpha,I\right)=\frac{1}{Lh^{(d_{\mathcal{M}}+1)/2}}\sum_{\ell=1}^{L}\xi_\ell\left(\mathsf{b};\alpha\right)$$

where the $\xi_\ell\left(\mathsf{b};\alpha\right)$ are centered iid. with

$$\begin{aligned}
\xi_\ell\left(\mathsf{b};\alpha\right) &= \sum_{i=1}^{I_\ell}\left\{\mathbb{I}\left[B_{i\ell}\in\Psi\left(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}\right),I_\ell=I\right]\xi_{i\ell}\left(\mathsf{b}\right)\right. \\
&\qquad\left.-\mathbb{E}\left[\mathbb{I}\left[B_{i\ell}\in\Psi\left(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}\right),I_\ell=I\right]\xi_{i\ell}\left(\mathsf{b}\right)\right]\right\} \\
\xi_{i\ell}\left(\mathsf{b}\right) &= \frac{h^{d_{\mathcal{M}}/2}}{h^{1/2}}\frac{P_{k_1}\left(x_\ell\right)P_{k_2}\left(x_\ell\right)}{\Psi^{(1)}\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)/h}K_p\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right), \\
K_p\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right) &= \frac{\Delta^{p_1+p_2}\left(B_{i\ell}|x_\ell,\mathsf{b}\right)}{p_1!p_2!}K\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right).
\end{aligned}$$

The proof of the Lemma follows from Proposition C.1. Observe

$$\left|\xi_\ell\left(\mathsf{b};\alpha\right)\right| \leq C\frac{h^{d_{\mathcal{M}}/2}\max_{x\in\mathcal{X}}\left\|P\left(x\right)\right\|^2}{h^{1/2}} \leq M_\infty \text{ with } M_\infty \asymp h^{-(d_{\mathcal{M}}+1)/2}.$$

for all α in $[0,1]$ and all admissible b. For the variance, Lemma B.1-(iii,iv) gives

$$
\begin{aligned}
|\Delta(B_{i\ell}|x_\ell,\mathsf{b})| &= \left|\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b})-\alpha}{h}\right| \\
&\leq \left|\frac{G(B_{i\ell}|x_\ell,I_\ell)-\alpha}{h}\right| + \left|\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b}^*(\alpha|I_\ell))-G(B_{i\ell}|x_\ell,\mathsf{b})}{h}\right| \\
&\quad + \left|\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b})-\Phi(B_{i\ell}|x_\ell,\mathsf{b}^*(\alpha|I_\ell))}{h}\right| \\
&\leq \left|\frac{G(B_{i\ell}|x_\ell,I_\ell)-\alpha}{h}\right| + o(h^s) + O\left(\frac{h^{-d_{\mathcal{M}}/2}\times h^{d_{\mathcal{M}}/2+1}}{h}\right) \\
&= \left|\frac{G(B_{i\ell}|x_\ell,I_\ell)-\alpha}{h}\right| + O(1)
\end{aligned}
$$

uniformly. It follows that, $U_{i\ell}=G(B_{i\ell}|x_\ell,I_\ell)$ being a uniform random variable independent of (x_ℓ,I_ℓ)

$$
\begin{aligned}
\operatorname{Var}(\xi_\ell(\mathsf{b};\alpha)) &\leq CI^2h^{d_{\mathcal{M}}}\max_{x\in\mathcal{X}}\|P(x)\|^2\int_{\mathcal{X}}|P_{k_1}(x)P_{k_2}(x)|\,dx\int\mathbb{I}_{[-C,C]}\left(\frac{u-\alpha}{h}\right)\frac{du}{h} \\
&\leq CI^2h^{d_{\mathcal{M}}}\max_{x\in\mathcal{X}}\|P(x)\|^2\left(\int_{\mathcal{X}}P_{k_1}^2(x)\,dx\right)^{1/2}\left(\int_{\mathcal{X}}P_{k_2}^2(x)\,dx\right)^{1/2} \\
&\leq M_2^2 \text{ with } M_2<\infty
\end{aligned}
$$

under Assumption R, uniformly in b and α.

Consider now the brackets covering. The key observation is that $\xi_\ell(\mathsf{b};\alpha)$ only depends on a finite dimension subvector of b, $\mathsf{b}^{(k_1,k_2)}$ which groups the entries of b corresponding to those $P_k(\cdot)$ such that $P_k(\cdot)P_{k_1}(\cdot)\neq 0$ or $P_k(\cdot)P_{k_2}(\cdot)\neq 0$, so that the dimension of $\mathsf{b}^{(k_1,k_2)}$ is less than $c(s+2)$ under Assumption R-(ii). Consequently the class to be bracketed is

$$
\mathcal{F}=\left\{\xi_\ell\left(\mathsf{b}^{(k_1,k_2)};\alpha\right);\alpha\in[0,1],\mathsf{b}^{(k_1,k_2)}\in\mathcal{B}\left(\mathsf{b}^{(k_1,k_2)*}(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)\right\}.
$$

Lemma B.1-(iii), $1/\left(Lh^{d_{\mathcal{M}}+1}\right)=o(1)$, van de Geer (1999, p.20) and arguing as Guerre and Sabbah (2012, 2014) imply that $\mathcal{F}$ can be bracketed with a number of brackets

$$
\exp(H_L(\epsilon))\asymp\left(\frac{L^C}{\epsilon}\right)^C
$$

so that

$$
\int_0^{M_2/2}\sqrt{\min(L,H_L(\epsilon))}d\epsilon\leq\left(\frac{M_2}{2}\right)^{1/2}\left(\int_0^{M_2/2}H_L(\epsilon)\,d\epsilon\right)^{1/2}=O(\log L)^{1/2}
$$

and for the item $\mathcal{H}_L$ of Proposition C.1,

$$\mathcal{H}_L = O\left(\log L\right)^{1/2} + O\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2} = O\left(\log L\right)^{1/2}$$

since $1/\left(Lh^{d_{\mathcal{M}}+1}\right)$ is bounded. Hence, by Proposition C.1 for $t \leq 10L^{1/2}M_2/M_\infty$ diverges

$$\mathbb{P}\left(\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2} \sup_{\alpha\in[0,1]} \sup_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)} \left|\widehat{\mathsf{r}}\left(\mathsf{b};\alpha,I\right)\right| \geq C\log^{1/2}L + t\right)$$
$$\leq 2\exp\left(-\frac{t^2}{25}\right)$$

uniformly over all the non zero entries $\widehat{\mathsf{r}}\left(\mathsf{b};\alpha,I\right)$ of the band matrix $\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right) - \overline{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)$. This gives, by the Bonferroni inequality

$$\mathbb{P}\left(\sup_{\alpha\in[0,1]} \sup_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{d_{\mathcal{M}}/2+1}\right)} \left\|\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right) - \overline{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)\right\| \geq \frac{C\log^{1/2}L + t}{\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}}\right)$$
$$\leq CK_L\exp\left(-\frac{t^2}{25}\right)$$

which implies the result of the lemma since $t \leq 10L^{1/2}M_2/M_\infty = O\left(Lh^{d_{\mathcal{M}}+1}\right)^{1/2}$ can be set to $t = \tau\log^{1/2}L$ for an arbitrary large τ as $\log L/\left(Lh^{d_{\mathcal{M}}+1}\right) = o\left(1\right)$. □

Proof of Lemma B.4. The proof of Lemma B.4 is similar to the one of Lemma B.3. The generic entry of $\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b};\alpha,I\right) - \overline{\mathsf{R}}^{(1)}\left(\mathsf{b};\alpha,I\right)$ writes

$$\widehat{\mathsf{r}}\left(\mathsf{b};\alpha,I\right) = \frac{1}{L}\sum_{\ell=1}^{L}\xi_\ell\left(\mathsf{b};\alpha\right)$$

where the $\xi_\ell\left(\mathsf{b};\alpha\right)$ are centered iid with, for $K_p\left(t\right) = t^pK\left(t\right)/p!$,

$$\xi_\ell\left(\mathsf{b};\alpha\right) = \sum_{i=1}^{I_\ell}\left(\mathbb{I}\left(I_\ell = I\right)\xi_{i\ell}\left(\mathsf{b};\alpha\right) - \mathbb{E}\left[\mathbb{I}\left(I_\ell = I\right)\xi_{i\ell}\left(\mathsf{b};\alpha\right)\right]\right),$$
$$\xi_{i\ell}\left(\mathsf{b};\alpha\right) = P_k\left(x_\ell\right)\left\{\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\overline{\mathsf{I}}_{\alpha,\mathsf{h}}}\left\{\mathbb{I}\left[B_{i\ell} \leq \Psi\left(t|x_\ell,\mathsf{b}\right)\right] - \left(\alpha + ht\right)\right\}K_p\left(t\right)dt\right\}.$$

This gives

$$\left|\frac{\xi_\ell\left(\mathsf{b};\alpha\right)}{\left(h+\alpha\left(1-\alpha\right)\right)^{1/2}}\right| \leq Ch^{-1/2}\max_{x\in\mathcal{X}}\left\|P\left(x\right)\right\| \leq M_\infty \text{ with } M_\infty \asymp h^{-(d_{\mathcal{M}}+1)/2}.$$

For the computation of the variance, Lemma B.1-(iii,iv) and Proposition 3-(i) give uniformly in α, t in $\mathcal{I}_{\alpha,h}$ the admissible b and x_ℓ, and for the uniform $U_{i\ell} = G\left(B_{i\ell}|x_\ell, I_\ell\right)$,

$$\begin{aligned}
\mathbb{I}\left[B_{i\ell} \leq \Psi\left(t|x_\ell,\mathsf{b}\right)\right] &= \mathbb{I}\left[B_{i\ell} \leq \Psi\left(t|x_\ell,\mathsf{b}^*\left(\alpha|I\right)\right) + O\left(h\right)\right] \\
&= \mathbb{I}\left[B\left(U_{i\ell}|x_\ell, I_\ell\right) \leq B\left(\alpha+ht|x_\ell, I_\ell\right) + O\left(h\right)\right] \\
&= \mathbb{I}\left[U_{i\ell} \leq G\left[B\left(\alpha+ht|x_\ell, I_\ell\right) + O\left(h\right)|x_\ell, I_\ell\right]\right] \\
&= \mathbb{I}\left[U_{i\ell} \leq \alpha + ht + O\left(h\right)\right].
\end{aligned}$$

It then follows, since $U_{i\ell}$ is independent of (x_ℓ, I_ℓ)

$$\begin{aligned}
&\mathbb{E}\left[\xi_{i\ell}^2\left(\mathsf{b};\alpha\right)|I_\ell\right] \\
&\leq \mathbb{E}\left[P_k^2\left(x_\ell\right)\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\mathbb{I}\left[U_{i\ell} \leq \alpha + h\left(t_1\wedge t_2\right) + O\left(h\right)\right]K_p\left(t_1\right)K_p\left(t_2\right)dt_1dt_2|I_\ell\right] \\
&\quad - 2\mathbb{E}\left[P_k^2\left(x_\ell\right)\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\mathbb{I}\left[U_{i\ell} \leq \alpha + ht_1 + O\left(h\right)\right]\left(\alpha+ht_2\right)K_p\left(t_1\right)K_p\left(t_2\right)dt_1dt_2|I_\ell\right] \\
&\quad + \mathbb{E}\left[P_k^2\left(x_\ell\right)|I_\ell\right]\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\left(\alpha+ht_1\right)\left(\alpha+ht_2\right)K_p\left(t_1\right)K_p\left(t_2\right)dt_1dt_2 \\
&= \mathbb{E}\left[P_k^2\left(x_\ell\right)|I_\ell\right]\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\left\{\alpha + O\left(h\right) - \alpha^2\right\}K_p\left(t_1\right)K_p\left(t_2\right)dt_1dt_2 \leq C\left(h+\alpha\left(1-\alpha\right)\right)
\end{aligned}$$

uniformly in α and b. Hence, uniformly in α and b

$$\operatorname{Var}\left(\frac{\xi_\ell\left(\mathsf{b};\alpha\right)}{\left(h+\alpha\left(1-\alpha\right)\right)^{1/2}}\right) \leq M_2^2 \text{ with } M_2 < \infty.$$

The bracketing part of the proof is similar to the one of Lemma B.3 and gives

$$\mathcal{H}_L = O\left(\log L\right)^{1/2} + O\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2} = O\left(\log L\right)^{1/2}.$$

Arguing with Proposition C.1 then shows that the order of the largest entry in $\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b};\alpha,I\right) -$

$\overline{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)$ is $O_{\mathbb{P}}(\log L/L)^{1/2}$, which gives uniformly

$$\left\|\widehat{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)-\overline{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)\right\| = K_L^{-1/2}O_{\mathbb{P}}\left(\frac{\log L}{L}\right)^{1/2} = O_{\mathbb{P}}\left(\frac{\log L}{Lh^{d_{\mathcal{M}}}}\right)^{1/2}$$

and the Lemma is proved. □

Proof of Lemma B.5. For (i), define

$$\begin{aligned}
\mathbf{P} &= \mathbb{E}\left[\mathbb{I}(I_\ell=I)\,P(x_\ell)\,P(x_\ell)'\right],\\
\mathbf{P}_0 &= \mathbb{E}\left[\frac{\mathbb{I}(I_\ell=I)\,P(x_\ell)\,P(x_\ell)'}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right],\\
\mathbf{P}_1 &= \mathbb{E}\left[\frac{\mathbb{I}(I_\ell=I)\,B^{(2)}(\alpha|x_\ell,I_\ell)\,P(x_\ell)\,P(x_\ell)'}{\left(B^{(1)}(\alpha|x_\ell,I_\ell)\right)^2}\right],
\end{aligned}$$

and abbreviate $\Omega_h(\alpha)$, $\Omega_{1h}(\alpha)$ in Ω, Ω_1. It holds

$$\operatorname{Var}(\widehat{\mathsf{e}}(\alpha|I)) = \left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]^{-1}\operatorname{Var}\left[\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]^{-1}$$

with by Lemma B.2

$$\begin{aligned}
\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]^{-1} &= \left[\Omega\otimes\mathbf{P}_0 - h\Omega_1\otimes\mathbf{P}_1 + o(h)\right]^{-1}\\
&= \Omega^{-1}\otimes\mathbf{P}_0^{-1}\left[\operatorname{Id} - h\left(\Omega^{-1}\Omega_1\right)\otimes\left(\mathbf{P}_0^{-1}\mathbf{P}_1\right)+o(h)\right]^{-1}\\
&= \Omega^{-1}\otimes\mathbf{P}_0^{-1} + h\left(\Omega^{-2}\Omega_1\right)\otimes\left(\mathbf{P}_0^{-2}\mathbf{P}_1\right)+o(h)
\end{aligned}$$

uniformly in α where the remainder term $o(h)$ is with respect to the matrix norm. For $\operatorname{Var}\left[\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]$, define

$$\begin{aligned}
\omega_0 &= \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \pi(t)\,K(t)\,dt, \quad \omega_1 = \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \pi(t)\,K(t)\,dt,\\
\mathbf{\Pi}_m &= \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \min(t_1,t_2)\,\pi(t_1)\,\pi(t_2)'\,K(t_1)\,K(t_2)\,dt.
\end{aligned}$$

Now (B.4) in the proof of Theorem B.8 and Lemma B.1-(iii,iv) show that $(LI)\,\mathrm{Var}\left[\widehat{\mathsf{R}}^{(1)}\left(\bar{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)\right]$ admits the expansion, with uniform remainder terms,

$$
\begin{aligned}
&\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\left\{G\left[B\left(\alpha+ht_1|x_\ell,I_\ell\right)\wedge B\left(\alpha+ht_2|x_\ell,I_\ell\right)+o\left(h\right)|x_\ell,I_\ell\right]\right.\right.\\
&\qquad -G\left[B\left(\alpha+ht_1|x_\ell,I_\ell\right)+o\left(h\right)|x_\ell,I_\ell\right]\left(\alpha+ht_2\right)-G\left[B\left(\alpha+ht_2|x_\ell,I_\ell\right)+o\left(h\right)|x_\ell,I_\ell\right]\left(\alpha+ht_1\right)\\
&\qquad \left.\left.+\left(\alpha+ht_1\right)\left(\alpha+ht_2\right)\right\}\pi\left(t_1\right)\pi\left(t_2\right)'K\left(t_1\right)K\left(t_2\right)dt_1dt_2\otimes\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)P\left(x_\ell\right)'\right]\\
&\ =\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\left\{\alpha+h\left(t_1\wedge t_2\right)-\alpha^2-h\alpha\left(t_1+t_2\right)\right\}\pi\left(t_1\right)\pi\left(t_2\right)'K\left(t_1\right)K\left(t_2\right)dt_1dt_2+o\left(h\right)\\
&\ =\alpha\left(1-\alpha\right)\omega_0\omega_0'\otimes\mathbf{P}+h\left\{\mathbf{\Pi}_m-\alpha\left(\omega_0\omega_1'+\omega_1\omega_0'\right)\right\}\otimes\mathbf{P}+o\left(h\right).
\end{aligned}
$$

Hence an elementary expansion gives, uniformly in $\alpha\in[0,1]$, $\mathrm{Var}\left(\widehat{\mathsf{e}}\left(\alpha|I\right)\right)=\mathsf{V}_e/\left(LI\right)+o\left(h\right)$ with

$$
\begin{aligned}
\mathsf{V}_e&=\alpha\left(1-\alpha\right)\Omega^{-1}\otimes\mathbf{P}_0^{-1}\times\omega_0\omega_0'\otimes\mathbf{P}\times\Omega^{-1}\otimes\mathbf{P}_0^{-1}\\
&\quad-h\alpha\left(1-\alpha\right)\left(\Omega^{-1}\otimes\mathbf{P}_0^{-1}\right)^2\times\Omega_1\otimes\mathbf{P}_1\times\omega_0\omega_0'\otimes\mathbf{P}\times\Omega^{-1}\otimes\mathbf{P}_0^{-1}\\
&\quad-h\alpha\left(1-\alpha\right)\Omega^{-1}\otimes\mathbf{P}_0^{-1}\times\omega_0\omega_0'\otimes\mathbf{P}\times\Omega_1\otimes\mathbf{P}_1\times\left(\Omega^{-1}\otimes\mathbf{P}_0^{-1}\right)^2\\
&\quad+h\Omega^{-1}\otimes\mathbf{P}_0^{-1}\times\left\{\mathbf{\Pi}_m\otimes\mathbf{P}-\alpha\left(\omega_1\omega_0'+\omega_0\omega_1'\right)\otimes\mathbf{P}\right\}\times\Omega^{-1}\otimes\mathbf{P}_0^{-1}\\
&=\alpha\left(1-\alpha\right)\left(\Omega^{-1}\times\omega_0\omega_0'\times\Omega^{-1}\right)\otimes\left(\mathbf{P}_0^{-1}\times\mathbf{P}\times\mathbf{P}_0^{-1}\right)\\
&\quad-h\alpha\left(1-\alpha\right)\left(\Omega^{-2}\times\Omega_1\times\omega_0\omega_0'\times\Omega^{-1}\right)\otimes\left(\mathbf{P}_0^{-2}\times\mathbf{P}_1\times\mathbf{P}\times\mathbf{P}_0^{-1}\right)\\
&\quad-h\alpha\left(1-\alpha\right)\left(\Omega^{-1}\times\omega_0\omega_0'\times\Omega_1\times\Omega^{-2}\right)\otimes\left(\mathbf{P}_0^{-1}\times\mathbf{P}\times\mathbf{P}_1\times\mathbf{P}_0^{-2}\right)\\
&\quad+h\left(\Omega^{-1}\times\mathbf{\Pi}_m\times\Omega^{-1}-\alpha\Omega^{-1}\times\left(\omega_1\omega_0'+\omega_0\omega_1'\right)\times\Omega^{-1}\right)\otimes\left(\mathbf{P}_0^{-1}\times\mathbf{P}\times\mathbf{P}_0^{-1}\right).
\end{aligned}
$$

Since the eigenvalues of $\mathbf{P}_0^{-1}$, $\mathbf{P}$, $\mathbf{P}_1$, Ω^{-1} and Ω_1 are bounded away from infinity uniformly in α, it follows that $\max_{\alpha\in[0,1]}\left\|\mathrm{Var}\left(\widehat{\mathsf{e}}_0\left(\alpha|I\right)\right)\right\|=O\left(1/L\right)$ and then

$$
\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\mathrm{Var}\left(P\left(x\right)'\widehat{\mathsf{e}}_0\left(\alpha|I\right)\right)=O\left(\frac{\max_{x\in\mathcal{X}}\left\|P\left(x\right)\right\|^2}{L}\right)=O\left(\frac{1}{Lh^{d_{\mathcal{M}}}}\right).
$$

For $\mathrm{Var}\left(\widehat{\mathsf{e}}_1\left(\alpha|I\right)/h\right)$, observe that $\widehat{\mathsf{e}}_1\left(\alpha|I\right)=S_1\widehat{\mathsf{e}}\left(\alpha|I\right)$ with

$$
S_1=s_1'\otimes\mathrm{Id}
$$

where Id is the $K_L\times K_L$ identity matrix and $s_1'=[0,1,0,\ldots]$ the row selection vector of dimension $s+2$. Let $s_0'=[1,0,\ldots]$ the row selection vector of dimension $s+2$, so that $s_1's_0=0$. Since $\Omega^{-1}\omega_0=s_0$, $\Omega^{-1}\omega_1=s_1$ which also gives

$$
s_1'\Omega^{-1}\times\left(\omega_1\omega_0'+\omega_0\omega_1'\right)\times\Omega^{-1}s_1=s_1'\Omega^{-1}\omega_1s_0's_1+s_1's_0\omega_1'\Omega^{-1}s_1=0
$$

it follows

$$
\begin{aligned}
S_1 \mathsf{V}_e S_1' &= h\left[s_1'\left(\Omega^{-1}\times\mathbf{\Pi}_m\times\Omega^{-1}-\alpha\Omega^{-1}\times\left(\omega_1\omega_0'+\omega_0\omega_1'\right)\times\Omega^{-1}\right)s_1\right]\otimes\left(\mathbf{P}_0^{-1}\times\mathbf{P}\times\mathbf{P}_0^{-1}\right)\\
&= h\left(s_1'\Omega^{-1}\mathbf{\Pi}_m\Omega^{-1}s_1\right)\left(\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right)\\
&= hv_h^2\left(\alpha\right)\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)P\left(x_\ell\right)'}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]\\
&\qquad\times\mathbb{E}\left[\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)P\left(x_\ell\right)'\right]\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)P\left(x_\ell\right)'}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]
\end{aligned}
$$

as $v_h^2\left(\alpha\right)=s_1'\Omega^{-1}\mathbf{\Pi}_m\Omega^{-1}s_1$. This gives the result for $\mathrm{Var}\left(\widehat{\mathsf{e}}_1\left(\alpha|I\right)/h\right)$ and $\mathrm{Var}\left(P\left(x\right)'\widehat{\mathsf{e}}_1\left(\alpha|I\right)/h\right)$.

For (ii), we just show that $\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\widehat{\mathsf{e}}_1\left(\alpha|I\right)/h\right|=O_{\mathbb{P}}\left(\left(\log L/Lh^{d_{\mathcal{M}}+1}\right)^{1/2}\right)$. Since $\max_{x\in[0,1]}\left\|P\left(x\right)\right\|=O\left(h^{-d_{\mathcal{M}}/2}\right)$ and

$$
\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{P\left(x\right)'\widehat{\mathsf{e}}_1\left(\alpha|I\right)}{h}\right|\leq\left(\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{P\left(x\right)'\widehat{\mathsf{e}}_1\left(\alpha|I\right)}{h^{1/2}\left(1+\left\|P\left(x\right)\right\|\right)}\right|\right)\times h^{-1/2}\left(1+\max_{x\in[0,1]}\left\|P\left(x\right)\right\|\right)
$$

it is sufficient to show

$$
\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{P\left(x\right)'\widehat{\mathsf{e}}_1\left(\alpha|I\right)}{h^{1/2}\left(1+\left\|P\left(x\right)\right\|\right)}\right|=O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{1/2}\right). \tag{C.3}
$$

Write

$$
\frac{P\left(x\right)'\widehat{\mathsf{e}}_1\left(\alpha|I\right)}{h^{1/2}\left(1+\left\|P\left(x\right)\right\|\right)}=\frac{1}{L}\sum_{\ell=1}^{L}\xi_\ell\left(\alpha,x\right)
$$

with

$$
\begin{aligned}
\xi_\ell\left(\alpha,x\right) &= \sum_{i=1}^{I_\ell}\left(\mathbb{I}\left(I_\ell=I\right)\xi_{i\ell}\left(\alpha,x\right)-\mathbb{E}\left[\mathbb{I}\left(I_\ell=I\right)\xi_{i\ell}\left(\alpha,x\right)\right]\right),\\
\xi_{i\ell}\left(\alpha,x\right) &= \frac{P\left(x\right)'S_1\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)\right]^{-1}P\left(x_\ell\right)}{h^{1/2}\left(1+\left\|P\left(x\right)\right\|\right)}\\
&\qquad\times\left\{\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\overline{\mathsf{I}}_{\alpha,\mathsf{h}}}\left\{\mathbb{I}\left[B_{i\ell}\leq\Psi\left(t|x_\ell,\overline{\mathsf{b}}\left(\alpha|I\right)\right)\right]-\left(\alpha+ht\right)\right\}K\left(t\right)dt\right\}.
\end{aligned}
$$

This gives, for all $(\alpha, x) \in [0,1]$

$$|\xi_\ell(\alpha, x)| \leq Ch^{-1/2} \frac{(\max_{x\in\mathcal{X}} \|P(x)\|)^2}{1+\max_{x\in\mathcal{X}} \|P(x)\|} \leq M_\infty \text{ with } M_\infty \asymp h^{-(d_\mathcal{M}+1)/2},$$

$$\operatorname{Var}(\xi_\ell(\alpha, x)) \leq C \frac{(\max_{x\in\mathcal{X}} \|P(x)\|)^2}{(1+\max_{x\in\mathcal{X}} \|P(x)\|)^2} \leq M_2 \text{ with } M_2 \asymp 1.$$

The Implicit Function Theorem and the FOC $\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I); \alpha, I\right) = 0$, Lemma B.2 with (B.4) and $s \geq d_\mathcal{M}/2$ give that $\alpha \mapsto \overline{\mathsf{b}}(\alpha|I)$ is $\|\cdot\|$-Lipshitz with a Lipshitz constant of order L^C, as $\alpha \mapsto \left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I); \alpha, I\right)\right]^{-1}$ and $x \mapsto P(x)/(1+\|P(x)\|)$. Lemma B.1-(iii), $1/\left(Lh^{d_\mathcal{M}+1}\right) = O(1)$, van de Geer (1999, p.20) and arguing as Guerre and Sabbah (2012, 2014) imply that $\{\xi_\ell(\alpha, x); (\alpha, x) \in [0,1] \times \mathcal{X}\}$ can be bracketed with a number of brackets

$$\exp(H_L(\epsilon)) \asymp \left(\frac{L^C}{\epsilon}\right)^C.$$

Arguing as in the proof of Lemma B.3 gives, for the item $\mathcal{H}_L$ of Proposition C.1,

$$\mathcal{H}_L = O(\log L)^{1/2} + O\left(\frac{\log L}{Lh^{d_\mathcal{M}+1}}\right)^{1/2} = O(\log L)^{1/2}$$

and then (C.3) holds. □

References

[1] 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.

[2] Guerre, E. & C. Sabbah (2014). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. http://arxiv.org/pdf/1105.5038.pdf

[3] 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.

[4] van de Geer, S. (1999). *Empirical processes in M-estimation.* Cambridge University Press.