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  
United Kingdom

October 2016 - (First version: November 2013)
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, 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 signal to the bidder covariate. The identification of this new model is established and some estimation methods are suggested.

\textit{JEL:} C14, L70

\textit{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, comments and encouragements from Klenio Barbosa, Christian Bontemps, Francesco Bravo, Estelle Cantillon, Valentina Corradi, Xiaohong Chen, Laurent Davezies, Dakshina De Silva, Pierre Dubois, Yanqin Fan, Jean-Pierre Florens, Eric Gautier, Matthew Gentry, Christian Gourieroux, Phil Haile, Xavier d’Haultfoeuille, Javier Hidalgo, Timothy Hubbard, Fabrizio Iacone, Asen Ivanov, Ingrid van Keleigom, Tatiana Komarova, Efang Kong, Gauthier Lanot, Stepana Lazarova, Hongjun Li, Qi Li, Oliver Linton, Nianqing Liu, Simon Lee, Aniol Llorente-Saguer, Jingfeng Lu, Adam McCloskey, Thierry Magnac, Vadim Marmer, Nour Meddahi, Serena Ng, Lars Neisheim, Andriy Norets, Tatsushi Oka, Taisuke Otsu, Harry Paarsch, Davy Paindavene, Elena Pastorino, Aureo de Paula, Isabelle Perrigne, Helena Perrone, Martin Pesendorfer, Eric Renault, Christoph Roth, Camille Sabbah, Marcia Schafgans, Suzanne Schennach, Juwon Seo, Art Shneyerov, Kevin Song, Sorawoot Srisuma, Denis Tkachenko, Thomas Verdebout, Michael Visser, Quang Vuong, Martin Weidner, Daniel Wilhem, Pai Xu, Hisayuki Yoshimoto, Andriy Zapechelnyuk, the participants of AMES2014, LAMES2014, 2014 Non and Semiparametric methods mini conference in York, 2014 Canadian Econometrics Study Group in Vancouver, 2014 MOITA Workshop, AMES2016, the 6th Shanghai Econometrics Workshop, the 2nd Dongbei Econometrics Workshop, the Séminaire Malinvaud, Queen Mary Economic Reading Group, New York University Econometrics Lunch, Kent, Columbia, Brown, UBC, Yale, FEA-USP, FEARP, Glasgow, Umeå, HKU, NUS, ULB-ECARES, UCL-CEMMAP, CUEB, TSE, LSE and Puc-Rio 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 replace 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

When exposing the identification strategy of Guerre, Perrigne and Vuong (2000, GPV hereafter), Milgrom (2001, Theorem 4.7) assumes a uniform signal and considers a valuation function which can be interpreted as the private value quantile function. Indeed, under risk neutrality and independent private value (IPV), the expected profit is the difference of the bidder private value, i.e. the private value quantile function taken at the signal, to the bid time the probability of winning. Under symmetric Bayesian Nash equilibrium bidding, 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 bid can be set equal to the bid quantile function, which is equal to the optimal strategy, taken at an arbitrary signal. Using the true signal is optimal and the associated first order condition (FOC hereafter) gives 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 GPV. 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 cannot capture interactions between the signal and the auctioned good covariate found in a timber data auction application. 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, as reviewed in Chen (2007), 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 characteristics observed by the econometrician. The considered parameters are the joint signal distribution and the valuation function of a specific bidder. As in the private value case, the signals can be easily recovered assuming strictly increasing strategies. The joint signal distribution can therefore 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. Such a restriction holds for instance in some simple auction with resale or in a revisited version of the Wilson (1998) model. Identification proceeds from the best response condition of bidder 1, which first identifies the expectation of bidder 1 valuation given that his bid is pivotal and his signal, as in HHS or Li, Perrigne and Vuong (2000, LPV hereafter). Differentiating with respect to the quantile level and with respect to other bidders characteristics gives an (integro) differential system, which has a unique solution under the considered functional restrictions. This line of proof seems to be new, although related to identification arguments used for competing risk models. See in

The identification arguments are also constructive and allows estimation the valuation function of bidder 1. ASQR procedures are potentially useful as it involves estimation of bid quantile functions and their derivatives at extreme quantiles. The proposed estimation procedure can be simple to implement for estimating linear valuation function or a revisited version of the Wilson (1998) model.

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 applied to timber auctions in Gimenes (2016). In the absence of covariate, Enache and Florens (2015a,b) have developed a Tikhonov regularization quantile framework for third-price and first-price auction models.

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 than 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.

**Signals.** Signals are a key element of interdependent value models and the choice of a signal distribution normalization is an important econometric issue. Athey and Haile (2007) have considered a normalization such that the expectation of the value given each individual
signal is equal to the signal. In the IPV case, this gives signals identical with private values as considered in GPV. As argued above for symmetric IPV, it may be more convenient for econometric purpose to use uniform private value signals. LPV achieved identification in a common value model through an assumption on the joint distribution of the value and the signals, which does not seem to impose a clearcut normalization on the signal distribution. For a common value model, Février (2008) has developed an original signal approach, which uses the bids as signals. In the IPV setting of GPV, this leads to take the inverse bidding strategy as a valuation function, and allows for a simple one step procedure that can compete with our quantile approach when there is no covariate.\footnote{Using the asymptotically unbasied inverse strategy estimator of Aryal et al. (2016) would for instance solve boundary bias issues which affects the GPV two step procedure. However estimating the inverse bidding strategy amounts to estimate the ratio of the conditional pdf and cdf, for which nonparametric specification not affected by the curse of dimensionality are difficult to justify. A quantile approach is probably more appropriate with that respect.}

Milgrom (2001) has used a uniform signal to present the identification strategy of GPV under symmetric IPV. As far as we are aware of, signals with marginal uniform distribution have not been considered in the interdependent case except in Somaini (2015) “for the sole purpose of simplifying notation”. Following Hubbard, Li and Paarsch (2012), Somaini (2015) adopts a copula approach which does not make use of the signal marginal distribution. His copula specification extends the lognormal specification of Wilson (1998) and Hong and Shum (2001) by allowing for a Gaussian factor structure which can be relevant for many applications. Somaini (2015) estimation strategy however use a uniform signal standardization to estimate the signals.
Identification and inference. Paarsch (1992) has considered common value specification, to be tested against IPV, see also HHS. Li, Perrigne and Vuong (2000) have used functional restriction to nonparametrically identify and estimate a common value specification which extends Wilson (1998). Hong and Shum (2002) have proposed a parametric nonlinear quantile regression approach for the common value Wilson model. Février (2008) considers alternative nonparametric specification where the support of the signal distribution depends upon the common value. Identification is established, as consistency of a nonparametric estimation procedure. He (2015) obtains constructive identification results for some policy parameters when the common value is an average of the signals plus a noise.

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 the conditional expectation of the value given all the signals, called the valuation function hereafter. 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 4.3.1.

2This 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 Econometric and statistical aspects

A quantile alternative to bid flexible specification. As mentioned earlier, GPV cannot be implemented with many covariates. In the IPV framework, some authors have proposed to use a “flexible” parametric distribution for the bids and to compute the associated private value distribution from the equilibrium FOC. See Jofre-Bonet and Pesendorfer (2003) or Athey et al. (2011) among others. As it is likely that the complexity of the bid flexible specification is going to increase with the sample size, this can be interpreted as a non-parametric sieve procedure, which sounds perfectly valid at first sight but which asymptotic properties remain to be clarified. This may be difficult due to the two step nature of such procedures. Our sieve quantile procedure actually parallels this approach but is much more simpler to study since it is one step. It also helps to focus on the interactions between the private signals and the observed covariates. Our results also suggest to use a quantile approach to test for the validity of a parametric specification.

Boundary issues. An interesting 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 esti-
mated, an important feature for applications based on such estimation as Cassola, Hortacşu and Kastl (2013). Ayral, Gabrielli and Vuong (2016) have recently developed a local polynomial version of the GPV private value estimator which is not affected by boundary issues but cannot cope with high dimensional covariates.

**Bandwidth choice.** 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 in the presence of bidder covariate within a quantile framework. 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. Bids are sealed so that a bidder does not know others’ bid when forming his own bid. The object is sold to the highest bidder who pays his bid \( B_i \) to the seller. Under the symmetric IPV paradigm, 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, but the joint distribution of the \( V_i \) is common knowledge. The private values are independently and identically drawn from a distribution given \( (x, I) \) with cdf \( F (\cdot | x, I) \), or equivalently with conditional quantile function

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

provided \( F (\cdot | x, I) \) is strictly increasing.

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

\[
A_i = F (V_i | x, I)
\]
has a uniform distribution over \([0, 1]\) and is independent of \(x\) and \(I\). It also follows from the IPV paradigm that the private value ranks \(A_i = 1, \ldots, I\) are independent. In other words, the dependence between the private value \(V_i\) and the auction covariates \(x\) and \(I\) is fully captured by the non separable model

\[
V_i = V(A_i|x, I), \quad A_i \sim \mathcal{U}_{[0,1]} \perp (x, I). \tag{2.1}
\]

Following Milgrom and Weber (1982) or Milgrom (2001), \(V(\cdot|x, I)\) can also be viewed as a valuation function, the private value rank \(A_i\) being the associated signal.

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

\[
B_i = \sigma(V_i|x, I) = s(A_i|x, i) \quad \text{where } s(\cdot|x, i) = \sigma(F(\cdot|x, I); x, I)
\]

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

\[ B_i = \sigma (V_i; x, I) = s (A_i; x, I) \text{ for all } i = 1, \ldots, I, \]  

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

i. [Signal identification] The conditional private value ranks \( A_i = F (V_i|x, I) \) are identical 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] The bids are given by

\[ B_i = B (A_i|x, I), \text{ for } i = 1, \ldots, I. \]  

Hence the signal bid function is identified with \( s (\cdot; x, I) = B (\cdot|x, I) \) while \( \sigma (\cdot; x, I) = B [F (\cdot|x, I)|x, I] \).

iii. [Probability of winning] Suppose bidder \( i \) bid is \( s (a; x, I) \) while his signal \( A_i \) is equal to \( \alpha \), while the other bidders bid \( B_j = s (A_j; x, I) \). Then the probability that bidder \( i \) wins the auction given \( A_i = \alpha \) and \( (x, I) \) 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 recover the signals from the bids, the covariate and the number of bidder. Lemma 1-(ii) directly follows inverting $G(B_i|x, I) = A_i$. The equations (2.3) together with (2.1) shows how to simulate $(V_i, B_i)$ from the private value and bid quantile functions and draws of the uniform signal $A_i$. This can be used to recover auction design parameters as the expected revenue

$$\mathbb{E} \left[ \max_{1 \leq i \leq I} B_i|x, I \right]$$

or the expected mark up

$$\mathbb{E} [V_i - B_i|x, I], \quad i = 1, \ldots, I$$

from an estimation of the private value and bid quantiles, without any need of estimating pdf. Lemma 1-(ii) also shows that the bid quantile function also identifies the signal bid function $s(\cdot; \cdot, \cdot)$, even when the bid function $\sigma(\cdot; \cdot, \cdot)$ 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 econometric 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 is strictly

\(^3\)As for instance in the nonparametric risk aversion setup considered in Campo, Guerre, Perrigne and Vuong (2011).
increasing and differentiable. By Lemma 1-(ii), the Bayesian Nash equilibrium bid function is identical to the conditional quantile function $B(\cdot|x,I)$. Suppose now that a bidder bids $B(a|x,I)$ while his signal is $\alpha$. Then Lemma 1-(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
\[
(V(\alpha|x,I) - B(a|x,I)) (I - 1) a^{I-2} - B^{(1)}(a|x,I) a^{I-1}
\]
\[
= (I - 1) a^{I-2} \left( V(\alpha|x,I) - B(a|x,I) - \frac{a B^{(1)}(a|x,I)}{I - 1} \right)
\]
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
\[
\alpha = \arg \max_{a \in [0,1]} (V(\alpha|x,I) - B(a|x,I)) a^{I-1}
\]
which gives a first order condition which implies for all $\alpha$ in $[0,1]$
\[
V(\alpha|x,I) = B(\alpha|x,I) + \frac{\alpha B^{(1)}(\alpha|x,I)}{I - 1}.
\] (2.4)
Hence the private value quantile function is easily identified from the bid quantile function and its derivative. When deriving the identification equation of GPV in a valuation function setup, Milgrom (2001, Theorem 4.7) derives a version of (2.4) with $1/g [B(\alpha|x,I) |x,I]$ instead of $B^{(1)}(\alpha|x,I)$. Guerre and Sabbah (2012), among others, have used (2.4) to estimate
the private value quantile function. The next section elaborates on (2.4) to propose parsimonious nonparametric specifications which are not subject to the curse of dimensionality.

### 2.1 Stability of quantile linear specifications

The identification equation (2.4) is a differentiable equation which can be solved using the initial condition \( B(0|x, I) = V(0|x, I) \) of Maskin and Riley (1984). 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. Then,

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

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

which is continuously differentiable over \([0, 1]\).

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

\[
V(\alpha|x, I) = B(\alpha|x, I) + \frac{\alpha B^{(1)}(\alpha|x, I)}{I - 1}.
\]

Equation (2.5) in Proposition 2-(i) shows that the bid quantile function is a linear functional of the private value quantile function. This linear transformation has an inverse which is given by (2.6). An important econometric consequence of Proposition 2-(i) is that a
linear specification for the private value quantile function is mapped into a similar linear specification for the bid quantile. An example of such private value specification is the homogenization technique of HHS, which postulates a regression model for the private value and generates a similar regression model for the bid, as also noted in Rezende (2008). The next sections extend this stability result to the quantile regression of Koenker and Bassett (1978) and to additive interactive quantile specification inspired by Horowitz and Lee (2005) and Andrews and Whang (1990).

2.1.1 Bid homogenization and quantile regression

Consider a $d$ dimensional covariate $x$. HHS and Rezende (2008) assume

$$V_i = \gamma_0 + x'\gamma_1 + v_i$$

where $v_i$ is a centred regression error term independent of $x$. This regression model implies that the private value quantile function writes

$$V(\alpha|x) = \gamma_0 + x'\gamma_1 + v(\alpha) \quad (2.7)$$

where $v(\alpha)$ is the quantile function of $v_i$. Since

$$\frac{I - 1}{\alpha^{I-1}} \int_0^\alpha a^{I-2}\gamma_1 da = \gamma_1$$
it follows that the associated bid quantile function is, by (2.5),

\[
B (\alpha|x, I) = \frac{I - 1}{\alpha I - 1} \int_0^\alpha a^{I-2} (\gamma_0 + x' \gamma_1 + v(a)) \, da = \gamma_0 + x' \gamma_1 + b(\alpha|I),
\]

\[
b(\alpha|I) = \frac{I - 1}{\alpha I - 1} \int_0^\alpha a^{I-2} v(a) \, da.
\]

This gives the bid regression model

\[
B_i = \beta_0 (I) + x' \gamma_1 + b_i, \quad \beta_0 (I) = \gamma_0 + \mathbb{E} [b(A_i|I)]
\]

where the regression error term \( b_i = b(A_i|I) - \mathbb{E} [b(A_i|I)] \) is centered. Following HHS, the coefficient \( \gamma_1 \) can be estimated regressing the bids on \([1, x']\) and the distribution of \( v_i \) can be estimated applying the two step method of GPV to the homogenized bids, which are the residuals \( B_i - x' \gamma_1 \).

However this approach requests independence between the regression error term \( v_i \) and the covariate \( x \). As seen from (2.7), this forbids interactions between the covariate and the signal. In particular, high values of the covariate \( x \) affect bidders of the top and the bottom of the distribution in the same way, an assumption which may be too restrictive in practice as found by Gimenes (2016) in a standard ascending auction timber dataset. This can be relaxed allowing the slope coefficients to vary with the quantile level \( \alpha \), as in the private value quantile regression model

\[
V (\alpha|x, I) = \gamma_0 (\alpha|I) + x' \gamma_1 (\alpha|I) = [1, x']' \gamma (\alpha|I). \tag*{(2.8)}
\]
Proposition 2-(i) implies that the conditional bid quantile function satisfies,

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

showing \( B(\alpha|x, I) \) belongs to the quantile regression specification. It follows from (2.6) that

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

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

The quantile regression specification is stable, i.e. that a quantile regression specification for the private value is equivalent to a quantile regression specification for the bid. Hence testing the correct specification of a bid quantile regression model is equivalent to test the correct specification of a private value quantile specification.\(^4\) This can be done using the quantile specification tests reviewed in Koenker (2005).

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

\(^4\)Note however that this ignores smoothness issue. For instance, a bid quantile regression with slope coefficients which are not twice continuously differentiable over \((0, 1]\) cannot generate private value one with continuously differentiable coefficients, as considered here.
function

\[ V(\alpha|x, I) = \sum_{j=1}^{d} V_j(\alpha; x_j, I) \] (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 following Andrews and Whang (1990) to allow for more covariate interactions. This leads to the additive interactive quantile specification with \( d_M \) interactions

\[ V(\alpha|x, I) = \sum_{D=1}^{d_M} \sum_{1 \leq j_1 < \cdots < j_D \leq d} V_{j_1 \cdots j_D}(\alpha; x_{j_1}, \ldots, x_{j_D}, I) \] (2.12)

where each functions \( V_{j_1 \cdots j_D}(\alpha; x_{j_1}, \ldots, x_{j_D}, I) \) can now depend upon \( D \) entries of \( x \) with \( D \leq d_M \leq d \). Setting \( d_M \) 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 \( d_M \) variables, so that \( d_M \) 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 \( d_M \) interaction will generate a bid quantile specification with the same number
of interactions: if (2.12) holds, then the bid quantile function satisfies

\[ B(\alpha|x, I) = \sum_{D=1}^{d_M} \sum_{1 \leq j_1 < \cdots < j_D \leq d} B_{j_1 \cdots j_D}(\alpha; x_{j_1}, \ldots, x_{j_D}, I) \]

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

\[ V_{j_1 \cdots j_D}(\alpha; x_{j_1}, \ldots, x_{j_D}, I) = B_{j_1 \cdots j_D}(\alpha; x_{j_1}, \ldots, x_{j_D}, I) + \frac{\alpha}{I-1} B_{j_1 \cdots j_D}^{(1)}(\alpha; x_{j_1}, \ldots, x_{j_D}, I) \]

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

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

where the sieve \( \{P_k(x), 1 \leq k \leq K\} \) 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 \( \alpha \). The choice of the sieve \( \{P_k(x), 1 \leq k \leq K\} \) should depend upon the order of interactions \( d_M \) in the model and the truncated \( \{P_k(x), 1 \leq k \leq K\} \) can be viewed as the covariate of a quantile regression specification. The model \( \mathcal{M} \) is therefore a sieve extension of the quantile regression, a *sieve quantile regression*. It follows from (2.5) in Proposition 2-(i) that, provided
the limit in (2.13) holds uniformly with respect to \( \alpha \),

\[
B (\alpha | x, I) = \lim_{K \to \infty} \sum_{k=1}^{K} \beta_k (\alpha | I) P_k (x), \quad \beta_k (\alpha | I) = \frac{I-1}{\alpha I-1} \int_{0}^{\alpha} t^{I-2}\gamma_k (t|I) dt, \tag{2.14}
\]

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

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 \( \{P_k (x), 1 \leq k \leq K\} \) 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 $d_M = 1$, a choice of sieve $\{P_k(x), 1 \leq k \leq K\}$ 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 ($d_M = 2$), 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 $d_M$ of interactions, the sieve $\{P_k(x), 1 \leq k \leq K\}$ is obtained by selecting similarly the product functions

$$p_{q_1}(x_{j_1}) \times \cdots \times p_{q_{d_M}}(x_{j_{d_M}}) = h^{-d_M/2} \prod_{k=1}^{d_M} p \left( \frac{x_j - q_k h}{h} \right) \quad (2.16)$$

for all possible $d_M$ indexes $1 \leq j_1 < \cdots < j_{d_M} \leq d$ and $K = K = O(1/h^{d_M})$.

### 2.2 Sieve expansion convergence rates

This section is 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

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

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

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 $K^{-(s+1)/d_M}$ in (2.17) replaces the order $h^{s+1}$ for the product sieve (2.16) with $d_M$ interactions since $K$ is of order $h^{-d_M}$. The approximation condition (2.17) strengthens the
more usual condition

$$\sup_{x \in \mathcal{X}} \left| V(\alpha | x, I) - \sum_{k=1}^{K} \gamma_k(\alpha | I) P_k(x) \right| = O \left( K^{-\frac{s+1}{4M}} \right)$$

which holds for higher order spline or wavelet provided that the partial derivatives of $V(\alpha | x, I)$ of order $s + 1$ with respect to $x$ are bounded, see Chen (2007, p.5573). Assuming in addition that these partial derivatives are continuous with respect to $\alpha$ 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.

- Higher order spline satisfying (2.17) are obtained from (2.16) using equispaced knots cardinal B-spline with an order larger than $s + 2$ as a function $p(\cdot)$, that is

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

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).

- Higher order wavelet example 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 $\alpha$ 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).

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 (2.5) and sieve coefficients $\{\gamma_k(\alpha|I), 1 \leq k \leq K\}$ of $V(\alpha|x, I)$ as in Property $S$

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

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

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

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

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

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

\[
\sup_{(\alpha, x) \in [0,1] \times \mathcal{X}} \left| \alpha B^{(1)} (\alpha|x, I) - \sum_{k=1}^{K} \alpha \beta_k^{(1)} (\alpha|x, I) P_k (x) \right| = o \left( K^{-\frac{s+1}{2s}} \right),
\]

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

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 $B^{(1)} (\alpha|x, I)$ with respect to $\alpha$. 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. As described above, the sieve component allows for flexible interactions between the signal $\alpha$ and the covariate $x$. The local polynomial part is new and will be used to estimate $B^{(1)} (\alpha|x, I)$, see Section 3.1. The main theoretical results in Section 3.2 focus on estimation of the conditional bid and private values quantile functions. Section 3.2.4 deals with estimation of the private values and of the optimal bidding strategy as a function of the private value. The choice of smoothing parameter and other implementation aspects is discussed later on with the simulation experiments, in Section 5.2.
3.1 Definition of the estimators

3.1.1 A local polynomial approach

The no covariate case. 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 \( V(\alpha|x, I) = V(\alpha|I) \) and \( B(\alpha|x, I) = B(\alpha|I) \). Let \( \rho_\alpha(u) \) be the check function,

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

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

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

Estimating the derivative \( B^{(1)}(\alpha|I) \) can be done by introducing local variation of the quantile level in the vicinity of \( \alpha \). Let \( K(\cdot) \geq 0 \) 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

\[
\{B(\alpha|I), \alpha \in [\alpha - h, \alpha + h] \cap [0, 1]\}
\]

\[
= \arg \min_{q(a)} \int_0^1 \mathbb{E}[\mathbb{I}(I_\ell = I) \rho_\alpha(B_{i\ell} - q(a))] \frac{1}{h} K \left( \frac{a - \alpha}{h} \right) da,
\]

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

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

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

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

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

$$
\hat{R}(b; \alpha, I) = \frac{1}{LI} \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_{0}^{1} \rho_\alpha \left(B_\ell t - \pi(a - \alpha)b\right) \frac{1}{h} K \left(\frac{a - \alpha}{h}\right) da
$$

$$
= \frac{1}{LI} \sum_{\ell=1}^{L} \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_{0}^{1} \rho_\alpha \left(B_\ell t - \pi(ht)b\right) K(t) dt.
$$

The augmented quantile estimator is $\hat{b}(\alpha|I) = \arg\min_b \hat{R}(b; \alpha, I)$, $\hat{\beta}_0(\alpha|I)$ and $\hat{\beta}_1(\alpha|I)$ being estimators of $B(\alpha|I)$ and its first derivative $B^{(1)}(\alpha|I)$, respectively.\footnote{When the private value distribution does not depend upon $I$, the bid quantile functions $B(\cdot|I)$ are such that the derivatives

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

do not depend upon $I$ as they are equal to $V^{(j)}(\alpha) = V^{(j)}(\alpha)$, $j = 0, \ldots, s + 1$. These constraints can be used to estimate $V(\alpha)$ using the parameters $\gamma = (\gamma_0, \ldots, \gamma_s)$, $\delta = (\delta_2, \ldots, \delta_I)$ where $\gamma_j$ is for $V^{(j)}(\alpha)$ and $\delta_I$ for the derivatives $B^{(s+1)}(\alpha|I)$, $I = 2, \ldots, I$ and $b_I(\gamma, \delta) = [b_0, I, \ldots, b_{s,I}, \delta_I]$ 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

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

}
private value quantile is
\[ \hat{V} (\alpha|I) = \hat{\beta}_0 (\alpha|I) + \frac{\alpha \hat{\beta}_1 (\alpha|I)}{I - 1}. \]

**Augmented quantile regression.** 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

\[ b = [\beta'_0, \beta'_1, \ldots, \beta'_{s+1}] \]

where all the \( \beta_j \) 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

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

The estimator of \( V (\alpha) \) is \( \hat{\gamma}_0 \) where \( \left( \hat{\gamma}, \hat{\delta} \right) = \arg \min_{\gamma, \delta} \sum_{l=2}^{T} \hat{R} (b_l (\gamma, \delta); \alpha, I). \)

6 Although not considered here, the augmented quantile estimation procedure can be used to estimate the p.d.f. \( f (v|I) \) of the private value using \( f (v|I) = 1/V^{(1)} [F (v|I)|I]. \) An estimator for \( F (\cdot|I) \) is \( \hat{V}^{-1} (\cdot|I). \) Set \( \hat{V}^{(1)} (\alpha|I) = \hat{\beta}_1 (\alpha|I) + \alpha \hat{\beta}_2 (\alpha|I)/(I - 1) \) and \( \hat{f} (v|I) = 1/\hat{V}^{(1)} [\hat{F} (v|I)|I]. \) This p.d.f. estimator can account for covariates by using the AQR and ASQR procedures introduced below.
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

\[
\hat{R}(b; \alpha, I) = \frac{1}{LI} \sum_{\ell=1}^{L} \mathbb{I}(I_{\ell} = I) \sum_{i=1}^{I_{\ell}} \int_{0}^{1} \rho_{a} \left( B_{i\ell} - P(x_{\ell}, a - \alpha) b \right) \frac{1}{h} K \left( \frac{a - \alpha}{h} \right) da
\]

which accounts for the covariate \( x_{\ell} \). The estimation of \( b(\alpha|I) \) is \( \hat{b}(\alpha|I) = \arg\min_{b} \hat{R}(b; \alpha, I) \) and the AQR private value quantile regression estimator is

\[
\hat{V}(\alpha|x, I) = [1, x']' \hat{\gamma}(\alpha|I) \quad \text{with} \quad \hat{\gamma}(\alpha|I) = \hat{\beta}_{0}(\alpha|I) + \frac{\hat{\alpha}\hat{\beta}_{1}(\alpha|I)}{I - 1}.
\]

The bid quantile function can be estimated using \( \hat{B}(\alpha|x, I) = [1, x']' \hat{\beta}_{0}(\alpha|I) \).

**Augmented sieve quantile regression.** The second extension is the augmented sieve quantile regression (ASQR) procedure which considers an interactive specification with \( d_{M} \) 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 \( K \), see Sections 2.1.2 and 2.2. Let \( K_{L} \) be a truncation parameter which diverges and will be taken of order \( h^{-d_{M}} \) 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 $\hat{R}(b; \alpha, I)$ as in the AQR case and

$$\hat{b}(\alpha|I) = \arg \min_b \hat{R}(b; \alpha, I).$$

The ASQR private value quantile estimator is, for

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

The ASQR bid quantile estimator is

$$\hat{B}(\alpha|x, I) = P(x)' \hat{\beta}_0(\alpha|I).$$

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).

### 3.1.2 Extreme quantile estimation and uniqueness issues

Standard quantile regression estimation is difficult to implement in the tails of the distribution. To see this, recall that the standard quantile regression estimators minimize the objective function

$$
\hat{R}_{QR}(b, \alpha, I) = \frac{1}{LI} \sum_{\ell=1}^{L} I(I_{\ell} = I) \sum_{i=1}^{I_{\ell}} \rho_{\alpha}(B_{i\ell} - [1, x_{\ell}]^\prime b) \text{ with } \rho_{\alpha}(q) = q(\alpha - I(q \leq 0)).
$$

For the extreme quantile $\alpha = 1$, $\rho_{\alpha}(q) = qI(q \geq 0)$ and

$$
\hat{R}_{QR}(b, 1, I) = \frac{1}{LI} \sum_{\ell=1}^{L} I(I_{\ell} = I) \sum_{i=1}^{I_{\ell}} (B_{i\ell} - [1, x_{\ell}]^\prime b) I(B_{i\ell} \geq [1, x_{\ell}]^\prime b)
$$

achieves its minimum value 0 for all $b$ satisfying

$$
B_{i\ell} \leq [1, x_{\ell}]^\prime b \text{ for all } i = 1, \ldots, I_{\ell} \text{ and } \ell \text{ with } I_{\ell} = I.
$$

If the entries of the $x_{\ell}$ are positive, this holds when any entry of $b$ is positive enough, showing that standard quantile regression cannot even identify the direction of the slope $\beta(1)$.\(^7\)

The augmented procedures proposed here are better behaved for extreme quantiles because the objective function $\hat{R}(\cdot ; \alpha, I)$ averages the check function $\rho_{\alpha}(\cdot)$ for quantile levels $\alpha$ in $[\alpha - ht, \alpha + ht] \cap [0, 1]$. For instance, if $\alpha = 1$ and $h \leq 1$, $\hat{R}(b; 1, I)$ averages

\(^7\)See also Figure 2 in Bassett and Koenker (1982) for a graphical illustration of this issue.
\( \rho_{1,+ht} (B_{i \ell} - P(x_{i \ell}, ht)'b) \) over \( t \) in \([-1, 0]\) so that \( \hat{\mathcal{K}}(b;1,I) \) will be large if \( b \) is too large.

Figure 1 below shows indeed that \( \hat{\mathcal{K}}(b;1,I) \) has no flat part when \( b \) grows, contrasting with the standard quantile regression objective functions.

![A path of the objective function \( \hat{\mathcal{K}}(b;1,I) \) (solid line) of the augmented quantile regression estimator and of the objective function \( \hat{\mathcal{K}}_{QR}(b;1,I) \) of the standard quantile regression estimator (dotted line) when \( b \) diverges in a given direction.](image)

Therefore the AQR and ASQR estimators are easier to define for the extreme quantile levels \( \alpha = 0 \) and \( \alpha = 1 \) than the standard quantile regression estimator. This is especially relevant for estimating auction models as the winner is expected to belong to the upper tail as soon as the number of bidders is large enough. In fact, it follows from the theoretical study of the objective function \( \hat{\mathcal{K}}(\cdot;\cdot,I) \) that the AQR and ASQR estimators are uniquely...
defined for all quantile levels with a probability tending to 1.\(^8\)

### 3.2 Main estimation results

As clear from Section 3.1, the AQR and ASQR estimation procedures can be nested in a general framework, which will be used to state the theoretical results. In this framework, the two procedures are identified by the sieve dimension \(d_M\), \(d_M = 0\) corresponding to the augmented regression case while \(d_M \geq 1\) is the order of interactions in the ASQR case.

#### 3.2.1 Main assumptions

The notations \(a \vee b\) and \(a \wedge b\) are used instead of \(\max(a, b)\) and \(\min(a, b)\). 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 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.

**Assumption A**

(i) The auction variables \(\{I_\ell, x_\ell, V_{i\ell}, B_{i\ell}, i = 1, \ldots, I_\ell\}\) are iid. The pdf \(f(x|I)\) of the covariates \(x_\ell\) 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 \(I_\ell\) 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

---

\(^8\) See the discussion following Theorem B.8 in Appendix B for a formal argument.
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) (2.5) 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 $d_M \in (0,d]$, the sieve satisfies the approximation property $S$.

Assumption A recalls the quantile implications of Bayesian Nash equilibrium bidding under symmetric IPV, see Assumption A-(iii). In Assumption A-(i), the existence of a conditional pdf for the covariate $x_t$ is only used for the infinite dimensional quantile regression specification. For a standard quantile regression specification, it is sufficient to assume that the matrix $\mathbb{E}[\Pi(I_t = I)X_tX_t']$ 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). For Assumption A-(ii), recall that

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

(3.3)
where \( f(v|x, I) \) is the conditional private value pdf. Hence Assumption A-(ii) amounts to assume that \( f(v|x, I) \) is bounded away from 0 and infinity on its support \([V(0|x, I), V(1|x, I)]\) as assumed for instance in Riley and Samuelson (1981), Maskin and Riley (1984) or GPV. The condition \( 0 < f(v|x, I) < \infty \) is also used for asymptotic normality of quantile regression estimator, see Koenker (2005). As discussed prior Proposition 3, Assumption S imposes that \( V(\cdot, I) \) is \( s+1 \) continuously differentiable. As a consequence \( B(\alpha|x, I) \) is \( s+2 \) continuously differentiable with respect to \( \alpha \) in \((0, 1]\). As a consequence, the AQR and ASQR procedures of Section 3.1 use a local polynomial expansion of order \( s+1 \). The next set of assumptions deals with sieve, kernel and bandwidth choices.

**Assumption R** In the AQR case the matrices \( \mathbb{E}[\mathbb{I}(I_t = I)X_tX'_t], I \in \mathcal{I} \), 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 \( K_{1L} \) with \( \log K_{1L} = O (\log L) \), it holds that

\[
\| P(x) - P(x') \| \leq K_{1L} \| x - x' \|^\eta \text{ for all } x, x' \text{ of } \mathcal{X}.
\]

**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

\[
\lim_{L \to \infty} \frac{\log L}{L h^{2(d_M+1)}} = 0.
\]

For the ASQR estimator, \( K_L \propto h^{-d_M} \).

Assumption R first imposes well conditioned matrices \( \mathbb{E} [\| (I_\ell = I) X_\ell X_\ell' \|] \) for the AQR case and \( \int_{\mathcal{X}} P(x) P'(x) \, dx \) for the ASQR case. The rest of Assumption R recalls some properties of the localized sieve (2.16) which are such that

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

which, when \( K_L \) is proportional to \( h^{-1/d_M} \) 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 \( p(\cdot) \) in the localized product sieve (2.16) has a compact support. Assumption R-(iii) holds when the bandwidth \( h \) of the sieve (2.16) decreases with a polynomial rate and provided \( p(\cdot) \) is Hölder with exponent \( \eta \). 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 \( \lim_{L \to \infty} \log L / (Lh^2) = 0 \) which is slightly more restrictive than the condition \( \lim_{L \to \infty} \log L / (Lh) = 0 \) used in nonparametric estimation.

3.2.2 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 \( \text{Id}_K \otimes s_1 \hat{\beta} (\alpha | I) = \hat{\beta}_1 (\alpha | I) \) is the estimator of sieve coefficient derivative \( \beta^{(1)} (\alpha) \). Let \( \Pi^1 (\alpha) \) be the second column of the inverse of \( \int \pi (t) \pi (t)' K (t) \, dt \), i.e.,

\[
\Pi^1 (\alpha) = \left( \int \pi (t) \pi (t)' K (t) \, dt \right)^{-1} s'_1
\]

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

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

\[
\Sigma (\alpha | I) = \frac{\alpha^2 v^2 (\alpha)}{(I - 1)^2} \mathbb{E}^{-1} \left[ \frac{P (x_\ell) P (x_\ell)' \mathbb{1} (I_\ell = I)}{B^{(1)} (\alpha | x_\ell, I_\ell)} \right]
\times \mathbb{E} \left[ P (x_\ell) P (x_\ell)' \mathbb{1} (I_\ell = I) \right] \mathbb{E}^{-1} \left[ \frac{P (x_\ell) P (x_\ell)' \mathbb{1} (I_\ell = I)}{B^{(1)} (\alpha | x_\ell, I_\ell)} \right],
\]

\[
\Sigma_{IL} = \int_X \int_0^1 P (x)' \Sigma (\alpha | I) P (x) \, d\alpha dx.
\]
That $v^2(\alpha)$, and then $\Sigma_{IL}$, 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

\[
\text{Bias}(\alpha|I) = \frac{1}{I-1} s_1 \left( \int \pi(t) \pi(t)^\prime K(t) \, dt \right)^{-1} \int \frac{t^{s+2} \pi(t)}{(s+2)!} K(t) \, dt
\]

\[\times \mathbb{E}^{-1} \left[ \frac{P(x_\ell) P(x_\ell)^\prime (I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)} \right] \mathbb{E} \left[ \frac{(I_\ell = I) P(x_\ell) \alpha B^{(s+2)}(\alpha|x_\ell, I_\ell)}{B^{(1)}(\alpha|x_\ell, I_\ell)} \right],\]

\[
\text{Bias}_{IL}^2 = \int_{\mathcal{X}} \int_0^1 (P(x)^\prime \text{Bias}(\alpha|I))^2 \, d\alpha dx.
\]

**Theorem 4** Suppose that the private value conditional quantile function $V(\cdot|\cdot)$ is a quantile regression (2.8), for which $d_M = 0$, or a sieve quantile regression (2.13) with $d_M$ interactions.

Then under Assumptions A, H, S and R-(i,ii) with $s \geq d_M/2$, there exists an approximation $\tilde{V}(\alpha|x, I)$ of $\hat{V}(\alpha|x, I)$ such that

\[
\mathbb{E} \left[ \int_{\mathcal{X}} \int_0^1 \left( \tilde{V}(\alpha|x, I) - V(\alpha|x, I) \right)^2 \, d\alpha dx \right] = h^{2(s+1)} \text{Bias}_{IL}^2 + \frac{\Sigma_{IL}}{L h^{d_M+1}}
\]

\[+ o \left( h^{2(s+1)} \frac{1}{L h^{d_M+1}} \right)\]

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

\[
\int_{\mathcal{X}} \int_0^1 \left( \tilde{V}(\alpha|x, I) - V(\alpha|x, I) \right)^2 \, d\alpha dx = o_F \left( \frac{1}{L h^{d_M+1}} \right). \quad (3.4)
\]

The quantile estimator $\hat{V}(\alpha|x, I)$ is nonlinear and defined in an implicit way, so that attempting a direct computation of its IMSE is difficult. Its approximation $\tilde{V}(\alpha|x, I)$ follows from a Bahadur linearization argument and is equal to $V(\alpha|x, I)$ plus a bias term, and the
score function of \( \hat{R}(b; \alpha, I) \) divided by a population Hessian, see equation (B.9) and Theorem B.9 in Appendix B. The rate in equation (3.4) is negligible with respect to the IMSE of \( \hat{V}(\alpha|x, I) \), showing that it is fair to replace \( \hat{V}(\alpha|x, I) \) by \( \tilde{V}(\alpha|x, I) \) to picture the IMSE of \( \hat{V}(\alpha|x, I) \).

Note that Theorem 4 holds over the full quantile level range \([0, 1]\). 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 \( h^{2(s+1)} \text{Bias}_{IL}^2 \). 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 \( \Sigma_{IL} / (LIh^{d_M+1}) \), which is similar to the asymptotic variance obtained for kernel estimation of a conditional pdf with \( d_M \) covariates. Indeed, the bid quantile derivative is homogeneous to a conditional pdf since

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

where \( g(\cdot|\cdot) \) is the bid conditional pdf. The bid quantile function is homogeneous to a cdf and converges with a faster rate. Note that the asymptotic variance term \( \Sigma_{IL} / (LIh^{d_M+1}) \) depends upon the number of interactions \( d_M \) and not the dimension of the covariate \( d \). Hence Theorem 4 illustrates the dimension reduction features of the AQR and ASQR 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

$$h_* = \left( \frac{(d_M + 1) \Sigma_{IL}}{2(s + 1) \text{Bias}_{IL}^2 L I} \right)^{\frac{1}{2s + d_M + 3}}. \tag{3.5}$$

As in kernel estimation, a pilot bandwidth can be computed using a simple private value quantile regression model to estimate $\Sigma_{IL}$ and $\text{Bias}_{IL}^2$ in a parametric way as implemented in Section 5.2.2 below. The corresponding IMSE rate is

$$L^{\frac{s+1}{2s + d_M + 3}}$$

which decreases with the number of interactions $d_M$, but does not depend upon the dimension $d$ of the covariate. In the AQR case with $d_M = 0$, the IMSE rate $L^{\frac{s+1}{2s + 3}}$ is as expected the optimal rate for estimating the marginal pdf 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 $L^{1/2}$.

Two assumptions limit the use of the optimal bandwidth (3.5). First, Theorem 4 assumes $s \geq d_M/2$ but this condition is only binding for a number of interactions $d_M$ 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 $2 (d_M + 1) / (2s + d_M + 3) < 1$ which boils down to the less stringent $s + 1 > d_M/2$. This contrasts with the smoothness condition $s > 0$ used in GPV and, for sieve regression estimator, in Belloni, Chernozhukov, Chetverikov and Kato (2015) and Chen and Christensen (2015). The stronger condition $s \geq d_M/2$ is due to the nonlinear nature of the sieve quantile regression estimator and is necessary to
approximate $\tilde{V}(\alpha|x, I)$ by $\tilde{V}(\alpha|x, I)$ as in (3.4). In a context where the covariate $d$ replaces $d_M$ but plays a similar role, Aryal et al. (2016) however use a condition $s + 1 > d$ to study a GMM version of GPV based on a local polynomial estimation of the private value which has no boundary bias.

### 3.2.3 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 $\tilde{V}(\alpha|x, I)$. Theorem 5 is also useful to better understand the pointwise properties of $\tilde{V}(\alpha|x, I)$, especially near the upper boundary $\alpha = 1$. Theorem 6 obtains its uniform convergence rates, a result allowing comparison with GPV. Theorem 6 also establishes the uniform convergence of $\hat{B}(\alpha|x, I)$, 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

$$\Pi^1_h(\alpha) = \left( \int_{-\frac{\alpha}{2}}^{\frac{1-\alpha}{2}} \pi(t) \pi(t)' K(t) \, dt \right)^{-1} s'_1,$$

$$v_h^2(\alpha) = \Pi^1_h(\alpha) \int_{-\frac{\alpha}{2}}^{\frac{1-\alpha}{2}} \int_{-\frac{\alpha}{2}}^{\frac{1-\alpha}{2}} \pi(t_1) \pi(t_2)' \min(t_1, t_2) K(t_1) K(t_2) \, dt_1 \, dt_2 \Pi^1_h(\alpha),$$
\[
\Sigma_h (\alpha | I) = \frac{\alpha^2 v_h^2 (\alpha)}{(I - 1)^2} \mathbb{E}^{-1} \left[ \frac{P(x_\ell) P(x_\ell)' \mathbb{I}(I_\ell = I)}{B^{(1)} (\alpha | x_\ell, I_\ell)} \right] \\
\times \mathbb{E} \left[ P(x_\ell) P(x_\ell)' \mathbb{I}(I_\ell = I) \right] \mathbb{E}^{-1} \left[ \frac{P(x_\ell) P(x_\ell)' \mathbb{I}(I_\ell = I)}{B^{(1)} (\alpha | x_\ell, I_\ell)} \right],
\]
(3.6)

\[
\text{Bias}_h (\alpha | I) = \frac{1}{I - 1}^{s_1} \left( \int_{-\frac{\alpha}{n}}^{\frac{1-\alpha}{n}} \pi (t) \pi (t)' K (t) dt \right)^{-1} \int_{-\frac{\alpha}{n}}^{\frac{1-\alpha}{n}} t^{s+2} \pi (t) K (t) dt \\
\times \mathbb{E}^{-1} \left[ \frac{P(x_\ell) P(x_\ell)' \mathbb{I}(I_\ell = I)}{B^{(1)} (\alpha | x_\ell, I_\ell)} \right] \mathbb{E} \left[ \frac{\mathbb{I}(I_\ell = I) P(x_\ell) \alpha B^{(s+2)} (\alpha | x_\ell, I_\ell)}{B^{(1)} (\alpha | x_\ell, I)} \right].
\]
(3.7)

**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 \( d_M \) interactions. Then under Assumptions A, H, S and R-(i,ii) with \( s \geq d_M/2 \) and

\[
\frac{\log^2 L}{L h^{2d_M + 1 + 1 + d_M}} = o (1),
\]

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

\[
\left( \frac{L h}{P(x)' \Sigma_h (\alpha | I) P(x)} \right)^{1/2} \left( \hat{V} (\alpha | x, I) - V (\alpha | x, I) - h^{s+1} P(x)' \text{Bias}_h (\alpha | I) + o (h^{s+1}) \right)
\]

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 $\hat{V}(0|x, I) = \hat{B}(0|x, I)$ has a smaller variance of order $1/(Lh^{d_M})$. For other quantile levels the private value conditional quantile estimator depends upon $\hat{B}^{(1)}(\alpha|x, I)$ so that the asymptotic variance of $\hat{V}(\alpha|x, I)$ has the larger order $1/(Lh^{d_M+1})$ 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 $\text{Bias}_h(\alpha|I)$ and $\Sigma_h(\alpha|I)$. Since the support of the kernel is $[-1, 1]$, it holds that

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

where $\text{Bias}(\alpha|I)$ 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 $\alpha$ 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 linearization remainder term combined with the order $h^{-d_M/2}$ of $P(x)P(x)'$ from Assumption R-(i), which now replaces the $\int_X P(x)P(x)'dx$ used for Theorem 4. Since the optimal bandwidths of Theorems 4 and 5 have the same order $L^{-1/(2s+d_M+3)}$, applying Theorem 5 with such bandwidths necessitates a larger smoothness index $s$. In addition to $s \geq d_M/2$, it must now hold that $s > d_M - 1$. 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 \( s \geq d_{\mathcal{M}}/4 \) 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 \( V (\cdot \mid \cdot) \) 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

\[
\frac{\log L}{L h^{2d_{\mathcal{M}}+1+(d_{\mathcal{M}}+1)}} = O (1),
\]

it holds

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

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

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

\[
\left( \frac{\log L}{L} \right)^{\frac{s+1}{2s+2d_{\mathcal{M}}+3}}
\]
which coincides with the GPV optimal minimax uniform consistency rate for the estimation of the private value conditional cdf in the presence of $d_M$ covariates. 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.

### 3.2.4 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 $A_{it}$ can be estimated by matching the estimated conditional bid quantile function with the observed bid,

$$
\hat{A}_{it} = \arg \min_{\alpha \in [0,1]} \left| B_{it} - \hat{B}(\alpha | x_t, I_t) \right|
$$

using an appropriate convention to break ties. Then (2.1) suggests the estimated private values,

$$
\hat{V}_{it} = \hat{V} \left( \hat{A}_{it} | x_t, I_t \right).
$$

The next Corollary gives the convergence rate of $\hat{V}_{it}$.

**Corollary 7** Under the conditions of Theorem 6

$$
\max_{\ell=1, \ldots, L} \max_{i=1, \ldots, I} \left| \hat{V}_{i\ell} - V_{i\ell} \right| = O_p \left( \left( \frac{\log L}{L h^{d_M+1}} \right)^{1/2} + h^{s+1} \right).
$$

---

9 GPV consider the pdf but the rate for cdf or quantile can be derived similarly.
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. This result can be useful in applications as in Cassola et al. (2013) which used estimated values, or two step estimation methods as in Aryal et al. (2016).

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

\[
\sigma(v|x, I) = B [F(v|x, I) | x, I]
\]

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

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

and a estimator of the optimal bidding strategy is

\[
\hat{\sigma}(v|x, I) = \hat{B} \left[ \hat{F}(v|x, I) | x, I \right].
\]

Arguing as in the proof of Corollary 7 shows that the estimated bidding 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 \( \hat{V}(0|x, I) \) and \( \hat{V}(1|x, I) \).
4 Extension to heterogenous interdependent value

4.1 Baseline framework

This section considers an extension where each bidder 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.\footnote{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 $A$.} 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 4.3.2. For the sake of brevity, the analysis will be restricted to univariate $z_i$ and the good covariate $x$ will be dropped out. The signal vector $A$ is assumed to be independent of $z = (z_1, \ldots, z_I)$ for the sake of brevity, but our results hold without this restriction. The individual signal $A_i$ all has a marginal uniform distribution but can be dependent. Due to the presence of the covariate $z$, this framework is asymmetric and bidding strategies should depend upon bidder’s identities,

$$B_i = s_i(A_i; z).$$ (4.1)

The observations consist on the bidder identities, bids $B_i$ and the covariate $z_i, i = 1, \ldots, I$.

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

\[ U_1 = U_1 (A; z) . \]

Our framework does not assume Bayesian Nash equilibrium bidding but retains some important equilibrium features. In particular, Lizerri and Persico (2002) have established that equilibrium bidding strategies must satisfy the so called \textit{terminal condition}\footnote{Lizerri 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.}

\[ s_1 (1; z) = \cdots = s_I (1; z) . \tag{4.2} \]

If (4.2) does not hold, all bidders \( i \) such that \( s_i (1; z) = \max_{1 \leq j \leq I} s_i (1; z) \) can jointly increase their profit at no cost by jointly decreasing their bids. Assume in a first step that the strategy functions \( s_i (\cdot; z) \) are continuous increasing for all \( i \). 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

\[ \alpha_i (z) = \max \left\{ \alpha \in [0, 1]; s_i (\alpha; z) \leq \max_{1 \leq j \neq i \leq I} s_j (0; z) \right\} , \tag{4.3} \]

which are well-defined under the terminal condition. There is always at least one bidder such that \( \alpha_i (z) = 0 \), so that the set of bidders with \( \alpha_i (z) = 0 \) 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 $D$ of dominated bidders $i$ with $\alpha_i(z) > 0$, or equivalently such that $s_i(0; z) < \max_{1 \leq j \leq I} s_j(0; z)$, is not empty. There is no aggressive bidding if $D$ is empty, that is $s_i(0; z) = \max_{1 \leq j \leq I} s_j(0; z)$ for all $i$, which is equivalent to

$$\alpha_i(z) = 0 \text{ for all } i = 1, \ldots, I.$$ 

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 $(0, \bar{z}]$ with $0 < \bar{z} \leq \infty$ and the support of $z$ is $\mathcal{Z} = (0, \bar{z}]^I$. Our framework relies on the following high-level assumption.

**Assumption DV** The signal vector $A$ is independent of the bidders characteristic $z_i$ 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 $U_i(A; z) = \alpha_i(z) < 1$, $s_i(\cdot; z)$ is continuous and increasing on $[\alpha_i(z), 1]$, with $s_i(\alpha; z) \leq s_i(\alpha_i(z); z)$ on $[0, \alpha_i(z)]$ and (4.2). 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 moreover strictly increasing on $[\alpha_i(z), 1]$.

(ii) For each $z$ in $\mathcal{Z}$ and for each $i = 1, \ldots, I$, $s_i(\cdot; z)$ is twice continuously differentiable on $[\alpha_i(z), 1]$ with $\frac{\partial}{\partial z} s_i(\cdot; z) > 0$ over $[\alpha_i(z), 1].$\footnote{Derivatives at $\alpha_i(z)$ are derivatives coming from the right. Our results also hold if the right first derivative vanishes at $\alpha_i(z) = 0$ with a right second derivative $\frac{\partial^2}{\partial z^2} s_i(\alpha_i(z); z) > 0.$}
(iii) For each $z$ in $Z$, the bidding strategy $s_1(\cdot; z)$ satisfies the best response condition

$$s_1(\alpha; z) \in \arg \max_b \mathbb{E} \left[ (U_1(A; z) - b) \mathbb{I} \left\{ b \geq \max_{2 \leq i \leq I} B_i \right\} | A_1 = \alpha, z \right]$$  \hspace{1cm} (4.4)

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

In addition to (4.2), Assumption DV retains the best response condition (4.4) among some key features of optimal Bayesian Nash Equilibrium bidding. Of course, other bidders can also bid strategically, in which case our framework allows for identification of their valuation functions. Such additional assumption would be necessary if the purpose is to simulate the outcomes of other auction mechanisms as it requires to recover all valuation functions.

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. 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 (4.5), the best response condition (4.4) does not provide any information about the bidding strategy for $\alpha < \underline{\alpha}_i(z)$. 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 vector distribution over its whole support $[0, 1]^I$. 

57
4.2 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 function, see the discussion following Lemma 9.

4.2.1 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 \( \alpha_i(z) \) strictly above 0. In this case, it may not be possible to recover signals \( A_i \) smaller than \( \alpha_i(z) \). It is therefore important to show that the \( \alpha_i(\cdot) \)'s are identified as follows from Lemma 8-(i).

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

i. **[Signal threshold identification]** For each \( z \) in \( Z \) and \( i = 1, \ldots, I \), the signal threshold \( \alpha_i(z) \) is equal to

\[
G_i \left[ \max_{1 \leq j \leq I} B_j(0 | z) \right]
\]
ii. **[Signal identification]** For each \( i = 1, \ldots, I \), the signals \( A_i \) satisfy

\[
A_i = G_i (B_i | z) \quad \text{provided} \quad B_i \geq B_i [a_i (z) | z]
\]

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

\[
\bigcup_{z \in \mathcal{Z}} [\alpha_1 (z), 1] \times \cdots \times [\alpha_I (z), 1].
\]

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

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

iv. **[Winning probability identification]** Suppose bidder 1 bid is \( s_1 (a; z) \) while his signal \( A_1 \) is equal to \( \alpha \). Then the probability \( \omega (a | \alpha, z) \) that bidder 1 wins the auction given \( A_1 = \alpha \) and \( z \) is identified provided \( \alpha \) is in \( [\alpha_1 (z), 1] \) 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 covariate, in which case the set of Lemma 8-(ii) is equal to $[0, 1]^{I}$ as $\alpha_i(z_0) = 0$ 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 heterogeneous 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 (3.8).\(^\text{13}\) 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 (4.4), as detailed in the next section to parallel Proposition 2.

\(^{13}\)Alternatively, the conditioning event $A_1 = \alpha$ can be equivalently written $B_1 = B_1 (\alpha | z)$ which suggests using kernel weights $K \left[ \left( B_{1 \ell} - \tilde{B}_1 (\alpha | z_\ell) \right) / h, (z_\ell - z) / h \right]$ instead of $K \left[ \left( A_{1 \ell} - \alpha \right) / h, (z_\ell - z) / h \right]$. 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.
4.2.2 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 $U_1(a|\alpha, z)$ of $U_1(A; z)$ over the event that bidder 1 wins, given $A_1 = \alpha \geq \alpha_1(z)$ and $z$. To see this, observe that the expected profit writes, setting $b = s_1(a; z)$,

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

Define

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

and observe that (4.4) gives

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

This gives the first order condition

$$
\frac{\partial \overline{U}_1(a|\alpha, z)}{\partial a} \bigg|_{a=\alpha} - s_1(\alpha; z) \frac{\partial \omega(a|\alpha, z)}{\partial a} \bigg|_{a=\alpha} - s_1^{(1)}(\alpha; z) \omega(\alpha|\alpha, z) = 0. \tag{4.5}
$$

Suppose first $\alpha < \alpha_1(z)$ and $s_1(\alpha; z) < s_1(\alpha_1(z); z)$, 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 $U_1(\cdot|\alpha, z)$ and $\omega(\cdot|\alpha, z)$

$$
\frac{\partial U_1(a|\alpha, z)}{\partial a}\bigg|_{a=\alpha} = \frac{\partial \omega(a|\alpha, z)}{\partial a}\bigg|_{a=\alpha} = \omega(\alpha|\alpha, z) = 0 \text{ for } \alpha < \overline{\alpha}_1(z)
$$

and (4.5) holds for any differentiable strategy satisfying $s_1(\alpha; z) < s_1(\overline{\alpha}_1(z); z)$. Hence best response strategies are not uniquely defined when $\alpha < \overline{\alpha}_1(z)$. For $\alpha \geq \overline{\alpha}_1(z)$ define\textsuperscript{14}

$$
\Omega(\alpha|z) = \frac{\omega(\alpha|\alpha, z)}{\partial \omega(a|\alpha, z)/\partial a}\bigg|_{a=\alpha}, \quad U_1(\alpha|z) = \frac{\partial U_1(a|\alpha, z)}{\partial a}\bigg|_{a=\alpha}.
$$

As shown in the proof of Lemma 9, these quantities are well defined for $\alpha \geq \overline{\alpha}_1(z)$. Since, by Lemma 8-(iii), $s_1(\alpha; z) = B_1(\alpha|z)$ when $\alpha \geq \overline{\alpha}_1(z)$, (4.6) in Lemma 9 shows that $U_1(\alpha|z)$ is identified for $\alpha$ in $[\overline{\alpha}_1(z), 1]$, 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 $\alpha$ in $[\overline{\alpha}_1(z), 1]$

$$
U_1(\alpha|z) = B_1(\alpha|z) + B_1^{(1)}(\alpha|z) \Omega(\alpha|z), \quad \text{(4.6)}
$$

$$
B_1(\alpha|z) = \exp\left(\int_\alpha^1 \frac{dt}{\Omega(t|z)}\right) \int_{\overline{\alpha}_1(z)}^\alpha \frac{U_1(a|z)}{\Omega(a|z)} \exp\left(-\int_a^1 \frac{dt}{\Omega(t|z)}\right) da, \quad \text{(4.7)}
$$

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

\textsuperscript{14}It can be shown that $U_1(\alpha|z)$ is the expectation of bidder 1 valuation given that his bid is pivotal and his signal, see HHS and Milgrom and Weber (1982).
The identity (4.6) extends (2.6), 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 non quantile versions of (4.6), respectively for testing common versus private values or identifying some common value models with symmetric homogeneous bidders. The equation (4.7) is obtained solving the differential equation (4.6) with the initial condition $B_1(\alpha_1(z)|z) = U_1(\alpha_1(z)|z)$ implied by (4.6) and $\Omega(\alpha_1(z)|z) = 0$. The identity (4.7) parallels (2.5) but is not as useful due to the presence of the covariate $z$. For instance, if $U_1(A;z) = \sum_{i=1}^I z_i\gamma_i(A_i)$, (4.7) 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.\(^{15}\)

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 $U_1(A;z)$ 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 (4.6). Likewise, the integral in (4.7) can be estimated and matched with the estimated bid quantile function $B_1(\alpha|z)$, see Enache and Florens (2015b) for such an attempt for the symmetric IPV case. In a parametric common value setup, Hong and Shum (2003) have implemented a nonlinear quantile regression which uses a version of (4.7) without bidder covariate to simulate the quantile function generated by their specification.

\(^{15}\)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 $(\gamma_1(\alpha|z), \gamma_2(\alpha|z))$.\)
4.3 The valuation function

The equation (4.6) 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 4.3.1. Section 4.3.2 introduces a nonparametric model of valuation functions which is shown to be identified.

4.3.1 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 $\alpha$, assuming $\alpha \geq \alpha_1(z)$. 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

$$U_1(a|\alpha, z) = \int U_1(\alpha, t; z) I[B_1(a|z) \geq B_2(t|z)] c(t|\alpha) \, dt$$

$$= \int_0^{G_{12}(a|z)} U_1(\alpha, t; z) c(t|\alpha) \, dt$$

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

$$\frac{\partial U_1(a|\alpha, z)}{\partial \alpha} = \frac{\partial}{\partial \alpha} \left[ \int_0^{G_{12}(a|z)} c(t|\alpha) \, dt \right] = g_{12}(a|z) U_1(\alpha, G_{12}(a|z); z) c(G_{12}(a|z)|\alpha)$$
where $g_{12}(a|z)$ is the derivative of $G_{12}(a|z)$ which is well-defined for $a \geq \alpha_1(z)$ 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 $[\alpha_1(z), 1]$ under Assumption DV-(ii). This gives that the identified $U_{1}(\alpha|z)$ is equal to

$$U_{1}(\alpha|z) = \left. \frac{\partial U_{1}(\alpha, z)}{\partial a} \right|_{a=\alpha} = U_{1}(\alpha, G_{12}(\alpha|z); z).$$

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 (4.8) 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 (4.8) 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 $\alpha$ 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 (4.2), which gives $B_1(1|z) = B_2(1|z)$. Hence this approach can only identify the valuation function near the diagonal $\alpha_1 = \alpha_2$ when $\alpha_1$ is close to 1. 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.

### 4.3.2 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

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

(4.9)

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

$$V_i(A_i; z_i) = \gamma_i(A_i) z_i, \; i = 1, \ldots, I$$

(4.10)

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 (4.9) and (4.10) but alternative conditions will be considered below.\(^{16}\) 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 \( \pi_i \gamma_i(A_i) z_i \) of the total valuation of the auctioned good, where the \( \pi_i \) may correct for the normalization \( \gamma_i(1) \) or indicate the preference of bidder 1. Hence the value of the auctioned good is\(^{17}\)

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

in which case \( \Phi(v_1, \ldots, v_I) = \sum_{i=1}^{I} \pi_i v_i \). Note that those \( \gamma_i(\cdot) \) associated with a vanishing \( \pi_i \) cannot be statistically identified. This specification can be useful for mineral rights auctions or more generally when bidders have specific expertise area, see He (2015) and the references therein for a similar specification without covariate.

- **Auction with resale.** Suppose each bidder is in contact with a final buyer to whom he can sell the good at a price \( \pi_i \gamma_i(A_i) z_i \) if he wins the auction. However the winner

\(^{16}\)Other normalizations 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 \( \bigcup_{z \in \mathcal{Z}} [\alpha_i(z), 1] \). Using \( \gamma_i(1/2) = 1 \) would request that \( \alpha_i(z) \geq 1/2 \) for some \( z \) and would reduce the identification set to \( \bigcup_{z \in \mathcal{Z}, \alpha_i(z) \geq 1/2} [\alpha_i(z), 1] \). This would also affect identification of the function \( \Phi(\cdot) \). However, when there is no aggressive bidding so that \( \alpha_i(z) = 0 \) for all \( i \) and \( z \), it is possible to consider the identification condition \( \gamma_i(0) = 1 \) as done below.

\(^{17}\)He (2015) considers the without covariate symmetric specification \( U_1(A) = \sum_{i=1}^{I} \gamma(A_i) / I \).
may be also in position to sell to other final buyers. Suppose bidder 1 can sell to final buyers in a subset $I$ of $\{1, \ldots, I\}$. If the prices $\pi_i \gamma_i(A_i) z_i$ are revealed from the bids or after the auction, bidder 1 may get

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

in which case $\Phi(v_1, \ldots, v_I) = \max_{i \in I} \{\pi_i v_i\}$. Other functions $\Phi(v_1, \ldots, v_I)$ may occur in more complex situations.

- **A revisited Wilson model.** Let $\Gamma(\cdot)$ be the quantile function of the standard normal and $A_0 = \Gamma^{-1}(\varepsilon_0)$ be a uniform signal unknown to the bidder, $\varepsilon_0$ having a $N(0, 1)$ distribution. The value of the good is $\gamma_0(A_0)$ and $\gamma_0(\cdot)$ is the parameter of interest. The bidders observe some public bidder covariate $\xi_i$ in $[0, 1]$ and a private signal

$$A_i = \Gamma^{-1}\left(\xi_i^{1/2} \varepsilon_0 + (1 - \xi_i)^{1/2} \varepsilon_i\right)$$

where the $\varepsilon_i$ are iid $N(0, 1)$, ensuring that $A_i$ has a uniform distribution. When $\xi_i = 1$, $A_i = A_0$ so that the bidder knows the value $\gamma_0(A_0)$ of the good, and the $\xi_i$ can be viewed as a bidder specific information variable. In this setup the valuation function
is, for \( z_i = \xi_i^{1/2} / (1 - \xi_i) \),

\[
U(A; z) = \mathbb{E}[\gamma_0 (A_0) | A_1, z_1, \ldots A_I, z_I] = \frac{\int \gamma_0 [\Gamma^{-1} (t)] \prod_{i=1}^{I} \exp \left( -\frac{(t - \xi_i^{1/2} \Gamma(A_i))^2}{2(1 - \xi_i)} \right) dt}{\int \prod_{i=1}^{I} \exp \left(-\frac{(t - \xi_i^{1/2} \Gamma(A_i))^2}{2(1 - \xi_i)} \right) dt}
\]

= \frac{\int \gamma_0 [\Gamma^{-1} (t)] \exp \left(t \sum_{i=1}^{I} z_i \Gamma (A_i) - \frac{t^2}{2} \sum_{i=1}^{I} \xi_i^{1/2} z_i \right) dt}{\int \exp \left(t \sum_{i=1}^{I} z_i \Gamma (A_i) - \frac{t^2}{2} \sum_{i=1}^{I} \xi_i^{1/2} z_i \right) dt}

= \Phi \left( \sum_{i=1}^{I} z_i \Gamma (A_i), \sum_{i=1}^{I} \xi_i^{1/2} z_i \right)

where the first integral expression recalls that

\[
\lim_{z_i \to -\infty} \Phi \left( \sum_{i=1}^{I} z_i \Gamma (A_i), \sum_{i=1}^{I} \xi_i^{1/2} z_i \right) = \gamma_0 (A_i), \quad i = 1, \ldots, I. \quad (4.11)
\]

Although the function \( \Phi (\cdot) \) depends upon the additional variable \( \sum_{i=1}^{I} \xi_i^{1/2} z_i \), restricting to those \( z_i \) such that \( \sum_{i=1}^{I} \xi_i^{1/2} z_i = C \) for some constant \( C > 0 \) gives a (family of) mixed signal models with known slope coefficients \( \gamma_i (\cdot) = \Gamma (\cdot), i = 1, \ldots, I. \) Identification is potentially simpler than for the two examples above, see footnote 18 below.

4.3.3 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 (4.8) writes,
for $\alpha \geq \alpha_1(z)$, \footnote{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. However, in the Wilson model example, $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ are known and the slope coefficient $\gamma_0(\cdot)$ of interest is identified from $U_1(\alpha|z)$ provided it is possible to allow for $z_1$ or $z_2$ going to minus infinity since it can be expected from (4.11) that \[ \gamma_0(\alpha) = \lim_{z_1 \to -\infty} U_1(\alpha|z) = \lim_{z_1 \to -\infty} U_1(G_{21}(\alpha|z)|z). \]}

\[ U_1(\alpha|z) = \Phi [\gamma_1(\alpha) z_1, \gamma_2[G_{12}(\alpha|z)] z_2] \quad \text{(4.12)} \]

where $G_{12}(\alpha|z) = B_2^{-1}[B_1(\alpha|z)|z]$ 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

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

In the private value case $\Phi(z_1, z_2) = \Phi(z_1)$, $\gamma_1(\cdot)$ is identified over $\bigcup_{z \in Z} [\alpha_1(z), 1]$, 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

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

to identify $\gamma_1(\cdot)$ over $[\alpha_1[(1, 0)], 1]$. This would however identify $\gamma_1(\cdot)$ over a smaller set than
\( \bigcup_{z \in Z} \left[ \alpha_1 (z), 1 \right] \). 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.\(^{19}\)

An alternative identification approach, which does not use such a specific \( z \), is to show that \( (\gamma_1 (\cdot), \gamma_2 [G_{12} (\cdot|z)]) \) solves a differential system which has a unique solution given the terminal condition

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

Differentiating (4.12) with respect to \( \alpha \) yields

\[
\Phi_{z_1} [\gamma_1 (\alpha) z_1, \gamma_2 [G_{12} (\alpha|z)] z_2] z_1 \frac{d\gamma_1 (\alpha)}{d\alpha} + \Phi_{z_2} [\gamma_1 (\alpha) z_1, \gamma_2 [G_{12} (\alpha|z)] z_2] z_2 \frac{\partial \{\gamma_2 [G_{12} (\alpha|z)]\}}{\partial \alpha} = \frac{\partial U_1 (\alpha|z)}{\partial \alpha}
\]

where \( \Phi_{z_i} (t_1, t_2) = \frac{\partial \Phi (t_1, t_2)}{\partial z_i} \). Before differentiating (4.12) with respect to \( z_2 \), observe

\[
\frac{\partial \{\gamma_2 [G_{12} (\alpha|z)]\}}{\partial z_2} = \gamma_2^{(1)} [G_{12} (\alpha|z)] \frac{\partial G_{12} (\alpha|z)}{\partial z_2} = \gamma_2^{(1)} [G_{12} (\alpha|z)] g_{12} (\alpha|z) \frac{\partial G_{12} (\alpha|z)}{\partial z_2} \frac{g_{12} (\alpha|z)}{g_{12} (\alpha|z)}
\]

\[
= \frac{\partial \{\gamma_2 [G_{12} (\alpha|z)]\}}{\partial \alpha} \frac{\partial G_{12} (\alpha|z)}{\partial z_2} \frac{g_{12} (\alpha|z)}{g_{12} (\alpha|z)}
\]

\[(4.13)\]

\(^{19}\)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.
Hence

\[ \Phi_{z_2} \left[ \gamma_1 (\alpha) \ z_1, \gamma_2 \ [G_{12} (\alpha|z)] \ z_2 \right] \gamma_2 \ [G_{12} (\alpha|z)] \]

\[ + \Phi_{z_2} \left[ \gamma_1 (\alpha) \ z_1, \gamma_2 \ [G_{12} (\alpha|z)] \ z_2 \right] \frac{\partial G_{12}(\alpha|z)}{\partial z_2} \frac{\partial \left\{ \gamma_2 \ [G_{12} (\alpha|z)] \right\}}{\partial \alpha} = \frac{\partial U_1 (\alpha|z)}{\partial z_2}. \]

Let

\[ D [\Phi, \gamma] (\alpha|z) = \begin{bmatrix} \Phi_{z_1} \left[ \gamma_1 (\alpha) \ z_1, \gamma_2 \ [G_{12} (\alpha|z)] \ z_2 \right] & 0 \\ 0 & \Phi_{z_2} \left[ \gamma_1 (\alpha) \ z_1, \gamma_2 \ [G_{12} (\alpha|z)] \ z_2 \right] \end{bmatrix}, \]

\[ G_2 (\alpha|z) = \begin{bmatrix} 1 & 1 \\ 0 & \frac{\partial G_{12}(\alpha|z)}{\partial z_2} \end{bmatrix}, \]

\[ \Psi [\Phi, \gamma] (\alpha|z) = \begin{bmatrix} \frac{\partial U_1 (\alpha|z)}{\partial \alpha} \\ g_{12} (\alpha|z) \left\{ \frac{\partial U_1 (\alpha|z)}{\partial z_2} - \Phi_{z_2} \left[ \gamma_1 (\alpha) \ z_1, \gamma_2 \ [G_{12} (\alpha|z)] \ z_2 \right] \gamma_2 \ [G_{12} (\alpha|z)] \right\} \end{bmatrix} \]

recalling that \( U_1 (\cdot|\cdot) \) and \( \Phi (\cdot) \) are identified at this stage. Combining the two differential equations derived from (4.12) shows that \( \gamma_1 (\cdot) \) and \( \gamma_2 \ [G_{12} (\cdot|z)] \) solves the differential system

\[ \begin{bmatrix} \frac{d}{d \alpha} \gamma_1 (\alpha) \\ \frac{d}{d \alpha} \left\{ \gamma_2 [G_{12} (\alpha|z)] \right\} \end{bmatrix} = \{ D [\Phi, \gamma] (\alpha|z) \ G_2 (\alpha|z) \}^{-1} \Psi [\Phi, \gamma] (\alpha|z) \]  \hspace{1cm} (4.14)

\( \alpha \) is in \( [\alpha_1 (z), 1] \), with the initial condition \( (\gamma_1 (1), \gamma_2 [G_{12} (1|z)]) = (1, 1) \). Standard rank and smoothness conditions on \( D [\Phi, \gamma] (\alpha|z) \ G_2 (\alpha|z) \) will ensure that its unique solution identifies \( \gamma_1 (\cdot) \) on \( \bigcup_{z \in \mathbb{Z}} \ [\alpha_1 (z), 1] \) and \( \gamma_2 (\cdot) \) on \( \bigcup_{z \in \mathbb{Z}} \ [\alpha_2 (z), 1] \) since \( \alpha_2 (z) = G_{12} (\alpha_1 (z)|z) \).
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

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

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 $\left. \frac{\partial \omega(\alpha|a,z)}{\partial a} \right|_{a=\alpha}$ is

$$
g_{12} (\alpha|z) \int_0^{G_{13}(\alpha|z)} \Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2 [G_{12} (\alpha|z)], z_3 \gamma_3 (t_3)] c [G_{12} (\alpha|z), t_3|\alpha] dt_3
\quad + g_{13} (\alpha|z) \int_0^{G_{12}(\alpha|z)} \Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 [G_{13} (\alpha|z)]] c [t_2, G_{13} (\alpha|z)|\alpha] dt_2.
\quad \text{(4.15)}
$$

The expression (4.15) is not as convenient as (4.8) to identify $\Phi (\cdot)$, due to the two integrals.\textsuperscript{20}

In the absence of aggressive bidding, that is if

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

\textsuperscript{20}Proceeding as in the two bidder case by recovering $\Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2 [G_{12} (\alpha|z)], z_3 \gamma_3 [G_{13} (\alpha|z)]]$ from (4.15) 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 $\alpha h' (\alpha) \neq 0$. 

73
it holds under standard continuity conditions that

$$\lim_{\alpha \to 0} U_1(\alpha | z) = \Phi [z_1 \gamma_1(0), z_2 \gamma_2(0), z_3 \gamma_3(0)] \quad (4.16)$$

as seen from (4.15), $G_{1i}(\alpha_1(z) | z) = 0$ for $i = 2, 3$ and recalling

$$\frac{\partial \omega(a|\alpha, z)}{\partial \alpha} \bigg|_{a=\alpha} = g_{12}(\alpha | z) \int_0^{G_{13}(\alpha | z)} c[G_{12}(\alpha | z), t_3|\alpha] \, dt_3$$

$$+ g_{13}(\alpha | z) \int_0^{G_{13}(\alpha | z)} c[t_2, G_{13}(\alpha | z)|\alpha] \, dt_2.$$

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 (4.15) with respect to $\alpha, z_2$ and $z_3$ to obtain an integro-differential system which can play the role of (4.14).

**The general case.** Let us now introduce some rank and smoothness conditions ensuring identification of the mixed signal specification (4.9-4.10). As above, $\Phi_{z_i}(z) = \frac{\partial}{\partial z_i} \Phi(z)$.

**Assumption MSM.** The valuation function of bidder 1 is given by (4.9), with a twice continuously differentiable $\Phi(\cdot)$ over the closure $\overline{Z}$ of $Z$, and (4.10) with continuously differentiable $\gamma_i(\cdot)$ taking value in $(0, 1]$. There is a subset $Z_0 = [z_0, \overline{z}_0]$ of $Z$ such 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 $Z$. For all $i$ of $\mathcal{I}$ $\Phi_{z_i}(z) \neq 0$ except for a finite number of $z$ of $\overline{Z}$ 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 $Z$
and all $\alpha > \alpha_1(z)$, $G_i [B_1 (\alpha | z) | z]$ is differentiable with respect to $z_j$ in $(0, \bar{z}]$. Moreover, for all $z$ of $\mathcal{Z}$,

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

except for a finite number of $\alpha > \alpha_1(z)$.

(iii) For all $i, j$ in $I$, all $z$ of $\mathcal{Z}$ and all $\alpha > \alpha_1(z)$, $g_{1i} (\alpha | z) = \frac{\partial}{\partial \alpha} \{G_i [B_1 (\alpha | z) | z]\}$ is continuously differentiable with respect to $\alpha$ and $z$.

It is important to note that the valuation functions of bidders $i = 2, \ldots, I$ is left unspec-iﬁed. 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 $\alpha$, see (ii). Note that this condition is in principle testable from the data. The subset $\mathcal{Z}_0$ of $\mathcal{Z}$ can be tailored to ensure that this condition holds.

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. 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.

**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, . . . , I. Then Φ(·) is identified over Z and, for all i in I, γ̂i(·) is identified over \( \bigcup_{z \in Z} [a_i(z), 1] \).

ii. Suppose \( I \geq 3 \), that \( a_i(z) = 0 \) for all \( z \) of \( Z \) and all i = 1, . . . , I, that the coefficients \( γ̂i(·) \) satisfy the identification restriction \( γ̂i(0) = 1, i = 1, . . . , I \), and that Assumption MSM-(iii) holds. Then Φ(·) is identified over Z and, for all i in I, γ̂i(·) 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 Φ(·) is achieved assuming nonaggressive 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.\(^{21}\)

The proofs of Theorem 10 (i) and (ii) are somehow similar, establishing first identification of Φ(·) and, in a second step, differentiating \( U_1(α|z) \) or \( \left. \frac{∂α(α|z)}{∂α} \right|_{α=α} U_1(α|z) \) with respect to \( α \) and \( z_i, i \geq 2 \) to find an (integro)differential system characterizing the coefficients \( γ̂i(·) \) in a unique way, as explained in the heuristic exposition above. This last step only involves one value of \( z \), suggesting that the \( γ̂i(·) \) are overidentified. Allowing for \( γ̂i(·) \) depending

\(^{21}\)Consider the two bidder case for the sake of brevity and \( U_1(α_1, α_2; z) = U_1(α_1, α_2; z_1) \) as in Somaini (2015). Then (4.8) shows that, for each \( z_1, U_1(α_1, α_2; z_1) \) is identified over

\[ \{ (α_1, α_2) ; \exists z_2 \in (0, τ) \text{ such that } α_1 ≥ a_1(z_1, z_2) \text{ and } α_2 = G_{12}(α_1|z_1, z_2) \} . \]

See Somaini (2015) for the general case. In particular \( α_2 \) must be equal to 1 if \( α_1 = 1 \) as \( G_{12}(1|z) = 1 \) for all \( z \) due to the terminal bidding condition. Since \( G_{12}(α_1|z) ≥ a_2(z) \) when \( α_1 ≥ a_1(z) \), the identification set of Somaini (2015) is strictly smaller than the identification set \([\min_{z \in Z} a_1(z), 1] \times [\min_{z \in Z} a_2(z), 1] \) of Theorem 10-(i).
upon \( z \) would still allow to identify \( \Phi(\cdot) \) under the condition of Theorem 10. Differentiating with respect to \( \alpha \) and \( z \) would give a partial (integro)differential system which may be more difficult to study.

4.4 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 (4.6) in Lemma 9

\[
\hat{U}_1(\alpha|z) = \hat{B}_1(\alpha|z) + \hat{B}_1^{(1)}(\alpha|z) \hat{\Omega}(\alpha|z)
\]

using unconstrained ASQR estimators to get consistency of \( \hat{U}_1(\alpha|z) \) at extreme quantiles. If \( I = 2 \) and under the identification condition of Theorem 10-(i), an estimator of \( \Phi(\cdot) \) is \( \hat{U}_1(1|z) \). Under the conditions of Theorem 10-(ii), estimators of \( \Phi(\cdot) \) are \( \hat{U}_1(0|z) \) or \( \hat{B}_1(0|z) \) which potentially converges with a faster rate. In the case of a parametric \( \Phi_\theta(\cdot) \) as in the additive valuation example, the parameter \( \theta \) can be estimated matching \( \Phi_\theta(\cdot) \) with the nonparametric \( \hat{\Phi}(\cdot) \), for instance minimizing \( \int_Z \left( \Phi_\theta(z) - \hat{\Phi}(z) \right)^2 dz \) with respect to \( \theta \) 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 \( \hat{U}_1(\alpha|z) \) with an estimation \( \hat{U}_1(\alpha|z, \Phi, \gamma) \) of \( U_1(\alpha|z) \) treating \( \hat{\Phi}(\cdot) \) and the \( \gamma_i(\cdot) \) as true values. For instance, if \( I = 2 \),
(4.12) suggests
\[ \hat{U}_1 \left( \alpha | z, \hat{\Phi}, \gamma \right) = \hat{\Phi} \left[ \gamma_1(\alpha) z_1, \gamma_2 \left( \hat{G}_{12}(\alpha | z) \right) \right] \]

where \( \hat{G}_{12}(\alpha | z) = \hat{G}_2 \left[ \hat{B}_1(\alpha | z) | z \right] \), \( \hat{G}_2(b | z) \) 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
\[ \int_{\Xi} \left( \hat{U}_1 \left( \alpha | z, \hat{\Phi}, \gamma \right) - \hat{U}_1(\alpha | z) \right)^2 d\alpha dz \]

over a sieve \( \Gamma_n \) satisfying the identification restriction used to construct \( \hat{\Phi}(\cdot) \), that is \( \gamma_1(1) = \gamma_2(1) = 1 \) if \( \hat{\Phi}(z) = \hat{U}_1(1 | z) \) or \( \gamma_1(0) = \gamma_2(0) = 1 \) if \( \hat{\Phi}(\cdot) \) is estimated using the lower extreme quantile. The case of a larger number of bidders complicates the choice of the estimator \( \hat{U}_1 \left( \alpha | z, \hat{\Phi}, \gamma \right) \) but this can be done using extensions of (4.15) 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 easy to implement when \( \hat{\Phi}(\cdot) \) 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

\[ V(\alpha|\mathbf{x}) = \gamma_0(\alpha) + x_1 + \gamma_2(\alpha) x_2, \]

\[ \gamma_2(\alpha) = 0.4 (1 - \exp(-6\alpha)), \quad \gamma_1(\alpha) = 1, \quad \gamma_0(\alpha) = -0.1 \log(1 - (1 - 1/e) \alpha). \]

The quantile function \( \gamma_0(\alpha) \) is the one of an exponential distribution with scale parameter 0.1, truncated over \([0, 0.1]\). The slope coefficient \( \gamma_2(\alpha) \) vanishes at 0, so that \( x_2 \) does not affect the value of bidders in the bottom of the distribution, and increases with \( \alpha \). By contrast, the slope coefficient of \( x_1 \) is equal to 1 independently of \( \alpha \). The covariates \( x_1 \) and \( x_2 \) are two independent uniform variables. The corresponding bid quantile regression model is, by (2.5) in Proposition 2,

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

The simulation experiment considers two numbers of bidders, \( I = 2 \) and \( I = 5 \). Since

\[ \hat{\gamma}_j(\alpha|I) = \hat{\beta}(\alpha|I) + \frac{\alpha \hat{\beta}^{(1)}(\alpha|I)}{I - 1} \]

a larger number of bidders may reduce the contribution of the term \( \hat{\beta}^{(1)}(\alpha|I) \), 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

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

with

\[
\begin{align*}
    b_0 (\alpha|I=2) &= \frac{(1 - \alpha) \log (1 - \alpha) + \alpha}{\alpha}, \\
    b_0 (\alpha|I=5) &= \frac{(12 - 12\alpha^4) \log (1 - \alpha) + 3\alpha^4 + 8\alpha^3 + 6\alpha^2 + 12\alpha}{12\alpha^4}, \\
    b_2 (\alpha|I=2) &= \frac{\exp (-\alpha) - 1 + \alpha}{\alpha}, \\
    b_2 (\alpha|I=5) &= \frac{(4\alpha^3 + 12\alpha^2 + 24\alpha + 24) \exp (-\alpha) + \alpha^4 - 24}{\alpha^4}.
\end{align*}
\]

In each simulation, the total number of bids is set to 100, which is 10 times 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 \( \alpha + h [-1, 1] \) with 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 \( \alpha_{k-1} \) (\( \alpha_{k+1} \) 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

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

where the disturbance error term \( e_{i\ell} \) has an exponential distribution with scaling parameter \( \lambda_1 \) truncated over \([0, \lambda_2]\), with a quantile function equal to

\[
Q_{\lambda} (\alpha) = -\frac{1}{\lambda_1} \log \left( 1 - (1 - \exp (-\lambda_1 \lambda_2)) \alpha \right).
\]
The parameters of this pseudo model can be estimated using the OLS slope coefficients $b_0$, $b_1$, $b_2$ and the OLS residuals $\varepsilon_{it}$. Estimators for $\beta_1$ and $\beta_2$ are $b_1$ and $b_2$ respectively. A natural estimator of $\lambda_2$ is the range

$$\hat{\lambda}_2 = \max_{1 \leq t \leq T, 1 \leq i \leq I} \varepsilon_{it} - \min_{1 \leq t \leq T, 1 \leq i \leq I} \varepsilon_{it}.$$  

$\beta_0$ and the disturbance terms $e_{it}$ are estimated using

$$\hat{\beta}_0 = b_0 + \min_{1 \leq t \leq T, 1 \leq i \leq I} \varepsilon_{it}, \quad \hat{e}_{it} = \varepsilon_{it} - \min_{1 \leq t \leq T, 1 \leq i \leq I} \varepsilon_{it}.$$  

The estimation of $\lambda_1$ is more difficult. The proposed estimator uses the formula

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

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

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

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

$$-\frac{1}{\hat{\lambda}_1} \log \left( 1 - \frac{1}{2} \left( 1 - \exp \left( -\hat{\lambda}_1 \lambda_2 \right) \right) \right)$$

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 the pseudo regression model (5.1) and using a modified Theorem 4 to propose an expansion for the modified IMSE

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

where $f(x|I)$ is the pdf of $x_\ell$ 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 $I$,

$$M_1 (\alpha) = \mathbb{E}_I \left[ \frac{X_\ell \alpha B^{(s+2)} (\alpha|x_\ell, I)}{B^{(1)} (\alpha|x_\ell, I)} \right], \quad M_2 (\alpha) = \mathbb{E}_I \left[ \frac{X_\ell X'_\ell}{B^{(1)} (\alpha|x_\ell, I)} \right],$$

$$M_2 = \mathbb{E}_I [X_\ell X'_\ell], \quad M_1 = \mathbb{E}_I [X_\ell],$$

$$\Pi_1 = \int \frac{t^{s+2} \pi (t)}{(s + 2)!} K(t) dt, \quad \Pi_2 = \int \pi(t) \pi(t)' K(t) dt,$$

$$\nu^2 = s_1 \Pi_2^{-1} \int \int \pi(t_1) \pi(t_2)' \min \left( t_1, t_2 \right) K(t_1) K(t_2) dt_1 dt_2 s_1'.$$
The expressions of the asymptotic bias and variance in (3.7) and (3.6) and Theorem 5 suggest that the modified IMSE leading term is

\[ \frac{v^2}{L_1 h} \int_0^1 \text{Tr} \left( M_2 (\alpha)^{-1} M_2 (\alpha)^{-1} M_2 \right) \frac{\alpha^2}{(I-1)^2} d\alpha 
+ h^{2(s+1)} \left( s_1^2 \Pi_2^{-1} \Pi_1 \right)^2 \int_0^1 \left( M_1 (\alpha) M_2 (\alpha)^{-1} M_2 (\alpha)^{-1} M_1 (\alpha)^\prime \right) d\alpha. \]

Minimizing this leading term gives the infeasible optimal bandwidth

\[ \left( \frac{v^2 \int_0^1 \text{Tr} \left( M_2 (\alpha)^{-1} M_2 (\alpha)^{-1} M_2 \right) \alpha^2 d\alpha}{2 (s+1) \left( s_1^2 \Pi_2^{-1} \Pi_1 \right)^2 \int_0^1 \left( M_1 (\alpha) M_2 (\alpha)^{-1} M_2 (\alpha)^{-1} M_1 (\alpha)^\prime \right) d\alpha} \right)^\frac{1}{2s+3}. \]

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

\[ M_1 = \mathbb{E}_I [X_1] \frac{\alpha Q_\lambda^{(s+2)} (\alpha)}{Q_\lambda^{(1)} (\alpha)} = M_1 \frac{\alpha Q_\lambda^{(s+2)} (\alpha)}{Q_\lambda^{(1)} (\alpha)}, \]

\[ M_2 (\alpha, I) = \mathbb{E}_I [X_1 X_2] \frac{M_2}{Q_\lambda^{(1)} (\alpha)} = \frac{M_2}{Q_\lambda^{(1)} (\alpha)} \]

with

\[ Q_\lambda^{(p)} (\alpha) = \frac{p!}{\lambda_1} \left( 1 - \exp (-\lambda_1 \lambda_2) \right)^p \frac{1}{\lambda_1} \left( 1 - (1 - \exp (-\lambda_1 \lambda_2) \alpha)^p \right)^p, \quad p = 1, s + 2. \]
Further elementary but cumbersome calculations yield that

\[ C_v = \int_0^1 \text{Tr} \left( M_2 (\alpha)^{-1} M_2 M_2 (\alpha)^{-1} M_2 \right) \alpha^2 d\alpha = \int_0^1 \left( \alpha Q_\lambda^{(1)}(\alpha) \right)^2 d\alpha \]

\[ = \frac{\exp (\lambda_1 \lambda_2) - \exp (-\lambda_1 \lambda_2) + 2\lambda_1 \lambda_2}{\lambda_1^2 (1 - \exp (-\lambda_1 \lambda_2))}, \]

\[ C_b = \int_0^1 \left( M_1 (\alpha)' M_2 (\alpha)^{-1} M_2 M_2 (\alpha)^{-1} M_1 (\alpha)' \right) d\alpha = M_1^2 M_2^2 M_2^2 M_1^2 \int_0^1 \left( \alpha Q_\lambda^{(s+2)}(\alpha) \right)^2 d\alpha \]

\[ = \frac{((s + 2)!^2 (1 - \exp (-\lambda_1 \lambda_2)))^{2s+1} \left( \exp ((2s + 3) \lambda_1 \lambda_2) - 1 \right)}{\lambda_1^2} \left( \frac{\exp ((2s + 3) \lambda_1 \lambda_2) - 1}{2s + 3} - \frac{\exp ((2s + 4) \lambda_1 \lambda_2) - 1}{s + 2} + \frac{\exp ((2s + 5) \lambda_1 \lambda_2) - 1}{2s + 5} \right) \]

which can be estimated plugging in \( \hat{\lambda}_1 \) and \( \hat{\lambda}_2 \). This together with \( \hat{M}_1 = \sum_{\ell=1}^L X_{\ell}/(LI) \) and \( \hat{M}_2 = \sum_{\ell=1}^L X_{\ell}X_{\ell}'/(LI) \) gives the pilot bandwidth

\[ \hat{h} = \left( \frac{\hat{C}_v}{2 (s + 1) \hat{C}_b \left( s_1 \Pi_2^{-1} \Pi_1 \right)^2 \hat{M}_1^2 \hat{M}_2^{-1} \hat{M}_2^{-1} \hat{M}_1 LI} \right)^{\frac{1}{2s+3}} \]

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 \( \hat{h} \) 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 \( \alpha = 0.5, 0.6 \), 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 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 slopes. In particular the estimation of private value quantile function for the average good looks very good.

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
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: 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.
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 $\hat{V}(\alpha|x)$ 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 $\hat{V}(\alpha|x,I)$ for a lower number of bidders $I = 2$ but still 100 bids. Since

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

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

Indeed Figure 4 suggests that $\hat{V}(\alpha|x,2)$ has a bigger variance than $\hat{V}(\alpha|x,5)$, 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 $\alpha$. This is expected due to the
Figure 4: AQR estimation of $V(\alpha|x)$ from 10,000 replications of a sample with $L = 50$, $I = 2$. 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.
multiplicative factor $\alpha$ in front of $\hat{B}^{(1)}(\alpha|x, I)$ in $\hat{V}(\alpha|x, I)$.

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 Koenker and Bassett (1978) quantile regression as a model for the private value quantile function. This gives a rich hierarchy of models allowing for various signal covariate interactions, that can be tested from the bids. The simple specification is quantile regression can be estimated with a fast nonparametric even with a high number of covariates. Other important features is the possibility of using simple bandwidth and consistency at the tails, allowing to recover the upper tail of the private value distribution which is important for winning 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, and the paper assumes that bidder specific covariate are observed. The considered
valuation function depends upon multiplicative functions tying up each bidder signal with his characteristic. This restriction, which covers simple case of auction with resale or a revisited Wilson model, 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. Because the procedure estimates quantile derivatives which involve density, the procedure can be used for the conditional private value density. 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


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 pdf 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 (2.1) 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 (2.2) gives

\[
G(B_i|x, I) = F(B_i|x, I) = F(V_i|x, I) = A_i,
\]

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 $U_{[0,1]} A_j$ given $(x,I)$,

$$
P\left(B_i > \max_{1 \leq j \neq i \leq I} B_j|I \right) = P\left(B(a|x,I) > \max_{1 \leq j \neq i \leq I} B(A_j|x,I)|x,I \right)
= P\left(a > \max_{1 \leq j \neq i \leq I} A_j \right) = a'^{-1}. \quad \square
$$

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

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

since $A_i$ has a uniform distribution over $[0,1]$ and $\alpha_i(z) \leq s^{-1}(b;z) \leq 1$. Hence $s_i(\alpha;z) = B_i(\alpha|z)$ over $[\alpha_i(z),1]$. (ii) easily follows. For (i), let $i_*$ be one of the most aggressive bidders, $i_* \in \arg \max_{i=1,...,I} s_i(0;z)$ so that $s_{i_*}(z) = 0$ and $s_{i_*}(\alpha;z) = B_{i_*}(\alpha|z)$ for all $\alpha$ in $[0,1]$ since $s_{i_*}(\cdot;z)$ is continuous strictly increasing by Assumption DV-(i). Then (4.3) gives

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

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

$$
\alpha_i(z) = G_i[B_{i*}(0|z)|z] = G_i\left[\max_{i \leq j \leq I} B_j(0|z)|z\right] = \max_{i \leq j \leq I} G_i[B_j(0|z)|z].
$$

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

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

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

**Proof of Proposition 2.** From Maskin and Riley (1984) and using the expression of $f(\cdot|x, I)$ from (3.3) to check that their pdf assumption holds,

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

for a strictly increasing and continuously differentiable $\sigma(\cdot; x, I)$. Hence

$$
B(\cdot|x, I) = \sigma(V(\cdot|x, I); x, I)
$$

and $B(\cdot|x, I)$ is continuoulsy differentiable over $[0,1]$. Proposition 2-(ii) is (2.4) and (i) follows by solving the differential equation (2.4) with the initial condition above as in the proof of Lemma 9. \hfill \qed
Proof of Proposition 3. By (2.5), \( B(\alpha|x, I) = (I - 1) \int_0^1 u^{I-2} V(\alpha u|x, I) \, du \), so that \( B^{(1)}(\alpha|x, I) = (I - 1) \int_0^1 u^{I-1} V'(\alpha u|x, I) \, du \) which implies the two first statements in (i) about lower and upper bounds for \( B^{(1)}(\alpha|x, I) \) and that \( B(\cdot|x, I) \) 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 (2.5). Observe now that for \( p = 1, \ldots, s+2 \)

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

with, for \( p = 1, \ldots, s+1 \)

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

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

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

Hence, when \( \alpha \) goes to 0

\[
\alpha B^{(s+2)}(\alpha|x, I) = -(I + s) B^{(s+1)}(0|x, I) + (I - 1) V^{(s+1)}(0|x, I) + o(1)
\]

\[
= -(I + s)(I - 1) \int_0^1 u^{I+s-1} V^{(s+1)}(0|x, I) \, du + (I - 1) V^{(s+1)}(0|x, I) + o(1)
\]

\[
= o(1)
\]

uniformly on \( x \).

For (ii), consider a sequence of \( \{\gamma_k(\alpha|I) : k \leq K\} \) approximating \( V(\alpha|x, I) \) and its deriv-
atives 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

\[
\sup_{(\alpha,x) \in [0,1] \times \mathcal{X}} \left| B^{(p)} (\alpha|x, I) - \sum_{k=1}^K \beta_k^{(p)} (\alpha|I) P_k (x) \right|
\]

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

\[
\leq \sup_{(\alpha,x) \in [0,1] \times \mathcal{X}} \left| V^{(p)} (\alpha|x, I) - \sum_{k=1}^K \gamma_k^{(p)} (\alpha|I) P_k (x) \right|
\]

which gives the sieve approximation result for \(B (\alpha|x, I)\) in (ii). Now, for \(\alpha B^{(1)} (\alpha|x, I)\), observe that

\[
\alpha B^{(1)} (\alpha|x, I) = (I - 1) \left[ V (\alpha|x, I) - B (\alpha|x, I) \right]
\]

and

\[
\alpha \beta_k^{(1)} (\alpha|I) = \alpha \times \left( -\frac{(I - 1)^2}{\alpha^4} \int_0^1 t^{I-2} \gamma_k (t|I) \, dt + \frac{I - 1}{\alpha} \gamma_k (\alpha|I) \right)
\]

\[
= (I - 1) \left[ \gamma_k (\alpha|I) - \beta_k (\alpha|I) \right].
\]
It follows

\[
\sup_{(\alpha,x)\in[0,1] \times \mathcal{X}} \left| \frac{\partial^p \left[ \alpha B^{(1)} (\alpha|x, I) \right]}{\partial \alpha^p} - \sum_{k=1}^{K} \frac{\partial^p \left[ \alpha \beta^{(1)}_k (\alpha|I) \right]}{\partial \alpha^p} P_k(x) \right| \\
\leq (I - 1) \sup_{(\alpha,x)\in[0,1] \times \mathcal{X}} \left| V^{(p)}(\alpha|x, I) - \sum_{k=1}^{K} \gamma^{(p)}_k (\alpha|I) P_k(x) \right| \\
+ (I - 1) \sup_{(\alpha,x)\in[0,1] \times \mathcal{X}} \left| B^{(p)}(\alpha|x, I) - \sum_{k=1}^{K} \beta^{(p)}_k (\alpha|I) P_k(x) \right| \\
\leq 2 (I - 1) \sup_{(\alpha,x)\in[0,1] \times \mathcal{X}} \left| V^{(p)}(\alpha|x, I) - \sum_{k=1}^{K} \gamma^{(p)}_k (\alpha|I) P_k(x) \right|
\]

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 \( \alpha \) decreases to \( \Omega_1 (z) \).

**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 (4.4), \( s_1 (1; z) = \cdots s_I (1; z) \) and

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

ii. It holds \( B_1 \left[ \Omega_1 (z) | z \right] = \cdots = B_I \left[ \Omega_I (z) | z \right] \text{ for each } z \text{ in } \mathcal{Z}. \)
iii. If Assumption DV-(ii) holds, $\Omega(\alpha|z)$ and $U_1(\alpha|z)$ are well defined for $\alpha$ in $(\bar{\alpha}_1(z), 1]$ with $\lim_{\alpha \to \bar{\alpha}_1(z)} \Omega(\alpha|z) = 0$, $\lim_{\alpha \to \bar{\alpha}_1(z)} U_1(\alpha|z) = U_1(\bar{\alpha}_1(z)|z)$ and

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

**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 $1 - \epsilon \geq \bar{\alpha}_I$ $s_I(1 - \epsilon) > s_{I-1}(1), \omega_I(1 - \epsilon|1) = 1$ with $\bar{U}_I(1 - \epsilon|1) < \bar{U}_I(1|1)$ 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 $\alpha_i = G_i(B_i(0))$ 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 $[\bar{\alpha}_1, 1]$ with $\frac{\partial}{\partial \alpha_i} G_{1i}(\alpha) > 0$ by Assumption DV-(ii) and

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

$$= \int_{G_{12}(a)}^{G_{12}(a)} \cdots \int_{G_{1I}(a)}^{G_{1I}(a)} c(\alpha_2, \ldots, \alpha_I|\alpha) \prod_{i=2}^{I} d\alpha_i$$
Hence $\Omega_1 (\alpha)$ is well defined for $\alpha$ in $(\alpha_1, 1]$ since $c(\cdot|\cdot)$ is positive. The limit for $\Omega (\alpha)$ when $\alpha$ decreases to $\alpha_1$ follows from the limit of $\Omega (\alpha) / (\alpha - \alpha_1)$. For the limit of $U_1 (\alpha)$, recall

$$
\mathcal{U}_1 (a|\alpha) = \int_0^{G_1(a)} \ldots \int_0^{G_1(a)} U_1 (\alpha, \alpha_2, \ldots, \alpha_I) c (\alpha_2, \ldots, \alpha_I|\alpha) \prod_{i=2}^{I} d\alpha_i \\
= \int_0^{G_1(a)} \ldots \int_0^{G_1(a)} U_1 (\alpha, \alpha_{-1}) c (\alpha_{-1}|\alpha) d\alpha_{-1}
$$

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

$$
\alpha_{-1,j} = (\alpha_2, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_I), \quad \alpha_{-1,2,j} = (\alpha_3, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_I), \\
d\alpha_{-1,j} = \prod_{i=2,i\neq j}^{I} d\alpha_i, \quad d\alpha_{-1,2,j} = \prod_{i=3,i\neq j}^{I} d\alpha_i,
$$

$$
U_1 (\alpha, G_jB_1 (a), \alpha_{-1,j}) = U_1 (\alpha, \alpha_2, \ldots, \alpha_{j-1}, G_jB_1 (a), \alpha_{j+1}, \ldots, \alpha_I), \\
c (G_jB_1 (a), \alpha_{-1,j}|\alpha) = c (\alpha_2, \ldots, \alpha_{j-1}, G_jB_1 (a), \alpha_{j+1}, \ldots, \alpha_I|\alpha), \\
c (G_2B_1 (t|z), G_jB_1 (a), \alpha_{-1,2,j}|\alpha) = c (G_2B_1 (t|z), \alpha_3, \ldots, \alpha_{j-1}, G_jB_1 (a), \alpha_{j+1}, \ldots, \alpha_I|\alpha).
$$

This gives for $a \geq \alpha_1$

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

$$
\frac{\partial \omega (a|\alpha)}{\partial a} = \sum_{j=2}^{I} \frac{\partial G_{1j} (a)}{\partial a} \int \prod_{2 \leq k \neq j \leq I} \mathbb{I} [\alpha_k \leq G_{1k} (a)] c (G_{1j} (a), \alpha_{-1,j}|\alpha) d\alpha_{-1,j}.
$$
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 \( B_j(0) = \min_{1 \leq i \leq B_i(0)} \). For some positive constants \( C \),

\[
\frac{\partial U_1(a|\alpha)}{\partial a} = C_U (a - \alpha_1)^{n_1-1} (1 + o(1)),
\]

\[
\omega(a|\alpha) = C_{\omega} (a - \alpha_1)^{n_1} (1 + o(1))
\]

\[
\frac{\partial \omega(a|\alpha)}{\partial a} = C'_{\omega} (a - \alpha_1)^{n_1-1} (1 + o(1))
\]

when \( a \) and \( \alpha \) decrease to \( \alpha_1 \). 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 \([\alpha_1, 1]\) with \( \alpha_1 < 1 \) follows from Lemma A.1. Then (4.5) gives (4.6)

\[
U_1(\alpha) = B_1(\alpha) + B_1^{(1)}(\alpha) \Omega(\alpha), \quad \alpha \in [\alpha_1, 1]
\]

with \( B_1(\alpha_1) = U_1(\alpha_1) \) since \( \Omega(\alpha_1) = 0 \) as established in Lemma A.1 and by Assumption DV-(ii) that ensures that \( B_1^{(1)}(\alpha_1) \) is finite. Hence Lemma 9-(i) holds. We now solve (4.6) to establish (4.7). 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 \( \Psi(\alpha) \sim C'(\alpha - \alpha_1)^{-1/C} \) when \( \alpha \) goes to \( \alpha_1 \) by Lemma A.1, where \( C > 0 \) is such that \( \Omega(\alpha) \sim C(\alpha - \alpha_1) \). Suppose that \( B_1(\alpha) = \)
\( \Psi (\alpha) b (\alpha) \) is a solution of (4.6). 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

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

and then

\[
B_1 (\alpha) = \Psi (\alpha) \times \left( C' + \int_\alpha^\infty \frac{U_1 (t)}{\Omega (t) \Psi (t)} dt \right).
\]

To show that \( C' = 0 \) observe that \( \lim_{\alpha \uparrow \alpha_1} \Psi (\alpha) = +\infty \), so that \( C' \neq 0 \) would give an infinite \( B_1 (\alpha_1) \). Now, by Lemma A.1-(iii),

\[
\lim_{\alpha \uparrow \alpha_1} \Psi (\alpha) \int_\alpha^\infty \frac{U_1 (t)}{\Omega_1 (t) \Psi (t)} dt = \lim_{\alpha \uparrow \alpha_1} C' (\alpha - \alpha_1)^{-1/C} \int_\alpha^\infty \frac{U_1 (\alpha_1)}{C'C (t - \alpha_1)^{1-1/C}} dt
\]

\[
= \lim_{\alpha \uparrow \alpha_1} \left\{ (\alpha - \alpha_1)^{-1/C} U_1 (\alpha_1) (t - \alpha_1)^{1/C} \right\}^\alpha_{\alpha_1} = U_1 (\alpha_1)
\]

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 (4.15) with
respect to $\alpha$, $z_2$ and $z_3$ as possible under Assumption MSM-(i,ii,iii). In what follows

$$
\Psi (\alpha | z) = \left. \frac{\partial U_1 (a | \alpha, z)}{\partial a} \right|_{a=\alpha}
$$

is the identified expression in (4.15). Recall also

$$
G_{ij} (\alpha | z) = G_j [B_i (\alpha | z) | z], \quad g_{ij} (\alpha | z) = \frac{\partial G_j B_i (\alpha | z)}{\partial \alpha},
$$

$$
c_{ij} (t_k | \alpha, z) = c (t_k | A_i = \alpha, A_j = G_{ij} (\alpha | z)).
$$

Recall also that $\Phi (\cdot)$ is identified over $Z = (0, \bar{z}]^3$ by (4.16), $\gamma_i (0) = 1$ for $i = 1, 2, 3$, and that $z_i \gamma_i [G_{1i} (\alpha | z)]$ belongs to $(0, \bar{z}]$ as the $\gamma_i (\cdot)$’s are valued in $(0, 1]$. Differentiating $\Psi (\alpha | z)$ with respect to $\alpha$ in $[0, 1]$ gives

$$
\left\{ g_{12} (\alpha | z) \int_0^{G_{13} (\alpha | z)} \Phi_{z_1} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (G_{12} (\alpha | z)), z_3 \gamma_3 (t_3)] c_{12} (t_3 | \alpha, z) \, dt_3 \\
+ g_{13} (\alpha | z) \int_0^{G_{12} (\alpha | z)} \Phi_{z_2} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (G_{13} (\alpha | z))] c_{13} (t_2 | \alpha, z) \, dt_2 \\
+ g_{12} (\alpha | z) \int_0^{G_{13} (\alpha | z)} \Phi_{z_3} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (G_{13} (\alpha | z))] c_{12} (t_3 | \alpha, z) \, dt_3
\right\} z_1 \gamma_1^{(1)} (\alpha)
$$

$$
\times z_2 \frac{\partial \{ \gamma_2 [G_{12} (\alpha | z)] \}}{\partial \alpha}
$$

$$
+ g_{13} (\alpha | z) \int_0^{G_{12} (\alpha | z)} \Phi_{z_3} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (G_{13} (\alpha | z))] c_{13} (t_2 | \alpha, z) \, dt_2
$$

$$
\times z_3 \frac{\partial \{ \gamma_3 [G_{13} (\alpha | z)] \}}{\partial \alpha}
$$

$$
= e_\alpha [\gamma] (\alpha | z)
$$
where

\[ e_{\alpha} [\gamma] (\alpha|z) = \frac{\partial \Psi (\alpha|z)}{\partial \alpha} \]

\[- \frac{\partial g_{12} (\alpha|z)}{\partial \alpha} \int_{0}^{G_{13}(\alpha|z)} \Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2, G_{12} (\alpha|z), z_3 \gamma_3 (t_3)] c_{12} (t_3|\alpha, z) dt_3 \]

\[- \frac{\partial g_{13} (\alpha|z)}{\partial \alpha} \int_{0}^{G_{12}(\alpha|z)} \Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (t_3)] c_{13} (t_2|\alpha, z) dt_2 \]

\[- 2g_{12} (\alpha|z) g_{13} (\alpha|z) \Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2, G_{12} (\alpha|z), z_3 [G_{13} (\alpha|z)] c [G_{12} (\alpha|z), G_{13} (\alpha|z)] |\alpha] \]

\[- g_{12} (\alpha|z) \int_{0}^{G_{13}(\alpha|z)} \Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2 [G_{12} (\alpha|z)], z_3 \gamma_3 (t_3)] \frac{\partial c_{12} (t_3|\alpha, z)}{\partial \alpha} dt_3 \]

\[- g_{13} (\alpha|z) \int_{0}^{G_{12}(\alpha|z)} \Phi [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (G_{13} (\alpha|z))] \frac{\partial c_{13} (t_2|\alpha, z)}{\partial \alpha} dt_2. \]

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

\[ \frac{\partial G_{12} (\alpha|z)}{\partial z_2} \int_{0}^{G_{13}(\alpha|z)} \Phi_{z_2} [z_1 \gamma_1 (\alpha), z_2 \gamma_2, G_{12} (\alpha|z), z_3 \gamma_3 (t_3)] c_{12} (t_3|\alpha, z) dt_3 \]

\[ \times \frac{\partial}{\partial \alpha} \{ \gamma_2 (G_{12} (\alpha|z)) \} \]

\[ + \frac{\partial G_{13} (\alpha|z)}{\partial z_2} \int_{0}^{G_{12}(\alpha|z)} \Phi_{z_3} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (G_{13} (\alpha|z))] c_{13} (t_2|\alpha, z) dt_2 \]

\[ \times \frac{\partial}{\partial \alpha} \{ \gamma_3 (G_{13} (\alpha|z)) \} \]

\[ = e_{z_2} [\gamma] (\alpha|z) , \]
where

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

\(e_{z_3} [\gamma]\) having a similar expression.

These three integro-differential equations can be stacked to obtain an expression similar to (4.14). Fix now a \(z\) in \(\mathbb{Z}\). Let \(e [\gamma] = [e_\alpha [\gamma] , e_{z_2} [\gamma] , e_{z_3} [\gamma]]'\). Consider the 3 \times 3 matrices
\(D [\gamma] (\alpha|z)\) and \(G (\alpha|z)\) with

\[
D_{\alpha \alpha} [\gamma] (\alpha|z) = z_1 g_{12} (\alpha|z) \int_0^{G_{13}(\alpha|z)} \Phi_{z_1} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (G_{12} (\alpha|z)), z_3 \gamma_3 (t_3)] c_{12} (t_3|\alpha, z) \, dt_3 \\
+ z_1 g_{13} (\alpha|z) \int_0^{G_{12}(\alpha|z)} \Phi_{z_1} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (G_{13} (\alpha|z))] c_{13} (t_2|\alpha, z) \, dt_2,
\]

\[
D_{z_2 z_2} [\gamma] (\alpha|z) = z_2 \int_0^{G_{13}(\alpha|z)} \Phi_{z_2} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (G_{12} (\alpha|z)), z_3 \gamma_3 (t_3)] c_{12} (t_3|\alpha, z) \, dt_3,
\]

\[
D_{z_3 z_3} [\gamma] (\alpha|z) = z_3 \int_0^{G_{12}(\alpha|z)} \Phi_{z_3} [z_1 \gamma_1 (\alpha), z_2 \gamma_2 (t_2), z_3 \gamma_3 (G_{13} (\alpha|z))] c_{13} (t_2|\alpha, z) \, dt_2,
\]

\(D [\gamma] (\alpha|z)\) being diagonal and

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

Then the three last equations above write

\[
G (\alpha|z) D [\gamma] (\alpha|z) \begin{bmatrix}
\gamma_1^{(1)} (\alpha) \\
\frac{\partial (\gamma_2 G_{12}(\alpha|z))}{\partial \alpha} \\
\frac{\partial (\gamma_3 G_{13}(\alpha|z))}{\partial \alpha}
\end{bmatrix} = e [\gamma] (\alpha|z) \quad (A.1.1)
\]
so that for all \( \alpha \) in \([0, 1]\)

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

assuming \( \Phi_{z_2} (\cdot) \) and \( \Phi_{z_3} (\cdot) \) do not vanish to ensure existence of the inverse matrix \( \{D [\gamma (\cdot|z)]\}^{-1} \) and ignoring that \( G (1|z) \) is not full rank as \( G_{ij} (1|z) = 1 \) for all \( z \). These issues are addressed using a regularized version of the operator \( E_\gamma [\cdot] \). 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 \( \sup_{x \in \mathbb{R}} |\tau^{(1)} (x)| < \infty \). 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 \( \sup_{x \in \mathbb{R}} |\tau_\epsilon^{(1)} (x)| \leq C \). 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 \( E_\gamma [\cdot] \) is

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

The role of the transformation \( T (\cdot) \) is to ensure that the entries of \( E_\gamma [\zeta] (\alpha|z) \) 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 $\epsilon$ 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 $\epsilon$. It follows

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

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

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 $\gamma^z(\cdot)$ with $\gamma^z(0) = [0, 0, 0]$ satisfying (A.1.2) so that

$$
\gamma^z(\alpha) - \gamma^z(\alpha) = \mathbf{E}_\gamma \gamma(\alpha|z) - \mathbf{E}_\gamma \gamma(\alpha|z) \text{ for all } \alpha \text{ in } [0, 1]. \quad \text{(A.1.3)}
$$

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

$$
\|\gamma^z(\alpha) - \gamma^z(\alpha)\| = \|\mathbf{E}_\gamma \gamma(\alpha|z) - \mathbf{E}_\gamma \gamma(\alpha|z)\| \leq C \alpha
$$

116
where the constant $C$ depends on $E_\gamma$, $\gamma^z(\cdot)$ and $\gamma^a_\alpha(\cdot)$ but does not depend upon $\alpha$. Substituting this bound in the fixed point condition (A.1.3) gives, after $k$ iterations and for the same constant $C$,

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

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

117
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 \( \|P(x)\| = (P(x)' P(x))^{1/2} \) is the standard Euclidean norm
and that, under Assumptions R-(i) and H-(ii),

\[
\max_{x \in \mathcal{X}} \|P(x)\| = O \left( K_L^{1/2} \right) = O \left( h^{-d_M/2} \right),
\]

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

with \( d_M = 0 \) 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 $E[P(x_\ell, ht) P(x_\ell, ht)']$ degenerates asymptotically. To avoid this, consider the change of parameters $b = Hb$ with $H = \text{Diag}(\pi(h)) \otimes \text{Id}_{K_L}$,

$$b = \begin{bmatrix}
\beta_{0,1}, \ldots, \beta_{0,K_L} \\
n_{\beta_{1,1}}, \ldots, h_{\beta_{1,K_L}}, \ldots, h_{s+1}^{*1}, \ldots, h_{s+1}^{*1,K_L}
\end{bmatrix}
$$

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

$$\hat{R}(b; \alpha, I) = \frac{1}{LIh} \sum_{\ell=1}^L \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_0^1 \rho_a \left( B_{it} - P \left( x_\ell, \frac{a - \alpha}{h} \right) ' b \right) K \left( \frac{a - \alpha}{h} \right) da$$

$$= \frac{1}{LI} \sum_{\ell=1}^L \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} \int_{\frac{a}{h}}^{1} \rho_{a+ht} \left( B_{it} - P(x_\ell, t)' b \right) K(t) dt,$$

$$\overline{R}(b; \alpha, I) = \mathbb{E} \left[ \hat{R}(b; \alpha, I) \right].$$

Note that $b \to \frac{1}{h} \int_0^1 \rho_{a+ht} (B_{it} - P(x_\ell, t)' b) K(t) dt$ is convex as an integral of convex functions. It follows that $\hat{R}(b; \alpha, I)$ and $\overline{R}(b; \alpha, I)$ have minimizers,

$$\hat{b}(\alpha|I) = \arg \min_b \hat{R}(b; \alpha, I) = H \hat{\beta}(\alpha|I),$$

$$\overline{b}(\alpha|I) = \arg \min_b \overline{R}(b; \alpha, I),$$

which uniqueness will be established in the next section. Set $\overline{b}(\alpha|I) = H^{-1} \overline{\beta}(\alpha|I)$ recalling

$$\overline{b}(\alpha|I) = \left[ \overline{\beta}_0(\alpha|I)', \ldots, \overline{\beta}_{s+1}(\alpha|I) \right]'$$

and define $\overline{B}(\alpha|x, I) = P(x)' \overline{\beta}_0(\alpha|I)$,

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

119
By Proposition 3 and its proof, there exists some $\beta^* (\cdot | I)$ grouping the entries in (2.14) such that

$$
\sup_{(\alpha, x) \in [0,1] \times X} | P(x) \beta^* (\alpha | I) - B(\alpha | x, I) | = o \left( \frac{K^\frac{4+1}{AM}}{h} \right) = o (h^{s+1}) .
$$

Let $b^* (\cdot | I)$ and $b^* (\cdot | I) = H b^* (\cdot | I)$ with

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

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

The next notations deal with the differentiability of the objective functions $\tilde{R}(\cdot; \alpha, I)$.

Since

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

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

$$
\tilde{R}^{(1)} (b; \alpha, I) = \frac{1}{L} \sum_{\ell=1}^{L} \sum_{i=1}^{L} \int_{-\frac{1}{h}}^{\frac{1}{h}} \{ I (B_{i\ell} \leq P(x_{\ell}, t)' b) - (\alpha + ht) \} P(x_{\ell}, t) K(t) \, dt
$$

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

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

120
is

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

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

\[
\frac{\partial}{\partial t} \left[ P(x, ht)' b^* (\alpha|I) \right] = h \left[ \pi^{(1)}(ht) \otimes P(x) \right]' b^* (\alpha|I)
= hP(x)' \beta^*_1 (\alpha|I) + O(h^2)
\]

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 \( P(x, t)' b^* (\alpha|I) \) is an increasing function of \( t \) in \( \mathcal{I}_{\alpha,h} \) for \( h \) small enough. Since \( \max_{(x,t) \in \mathcal{X} \times [-1,1]} \|P(x,t)\| = O(h^{-dM/2}), \ t \to P(x,t)' b \) is also strictly increasing provided \( b \) is close enough to \( b^* (\alpha|I) \).

In such case, it is convenient to redefine \( P(x,t)' b \) as follows

\[
\Psi(t|x, b) = \begin{cases} 
  P(x, I_{\alpha,h})' b & t > I_{\alpha,h} \\
  P(x,t)' b & t \in \mathcal{I}_{\alpha,h} \\
  P(x, L_{\alpha,h})' b & t < L_{\alpha,h}
\end{cases}
\]

\footnote{In principle \( \Psi(\cdot|\cdot) \) should be denoted \( \Psi_{\alpha,h}(\cdot|\cdot) \) to acknowledge that its definition depends upon \( \alpha \) 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.}
When $\Psi(\cdot| x, b)$ has an inverse, define

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

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

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

$$
\int_{\overline{T}_{\alpha,h}}^{\overline{T}_{\alpha,h}} \{ \mathbb{I} (B_{i\ell} \leq \Psi(t|x_\ell, b)) - (\alpha + ht) \} P(x_\ell, t)K(t) dt = \int_{\overline{T}_{\alpha,h}}^{\overline{T}_{\alpha,h}} \{ \mathbb{I} \left( \frac{\Phi(B_{i\ell}| x_\ell, b) - \alpha}{h} \leq t \right) - (\alpha + ht) \} P(x_\ell, t)K(t) dt
$$

$$
= \int_{\overline{T}_{\alpha,h}}^{\overline{T}_{\alpha,h}} \mathbb{I} (x_\ell, t) K(t) dt - \int_{\overline{T}_{\alpha,h}}^{\overline{T}_{\alpha,h}} (\alpha + ht) P(x_\ell, t)K(t) dt
$$

which is differentiable with respect to $b$, with for $B_{i\ell}$ in $\Psi(\overline{T}_{\alpha,h}| x, b)$

$$
\frac{\partial \Phi(B_{i\ell}| x_\ell, b)}{\partial b'} = -\frac{P(x, \Delta(B_{i\ell}| x_\ell, b))}{\Psi^{(1)}(\Delta(B_{i\ell}| x_\ell, b)| x_\ell, b)/h} \mathbb{I}[B_{i\ell} \in \Psi(\overline{T}_{\alpha,h}| x_\ell, b)].
$$

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

\[
\hat{R}^{(2)}(b; \alpha, I) = \frac{1}{L \ell h} \sum_{\ell=1}^{L} \sum_{i=1}^{I} \mathbb{I} [B_{i\ell} \in \Psi(I_{\alpha,h}, x_{\ell}, b), I_{\ell} = I] \nonumber
\]

\[
P(\ell, \Delta(B_{i\ell} | x_{\ell}, b)) P(\ell, \Delta(B_{i\ell} | x_{\ell}, b)') \Psi^{(1)} \Delta(B_{i\ell} | x_{\ell}, b) / h K(\Delta(B_{i\ell} | x_{\ell}, b)),
\]

\[
\overline{R}^{(2)}(b; \alpha, I) = \mathbb{E}\left[\hat{R}^{(2)}(b; \alpha, I)\right].
\]

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

\[
\begin{align*}
BT_{\alpha,h} &= \left\{ b : \min_{(t,x) \in I_{\alpha,h} \times X} \frac{\partial \Psi(t \cdot x, b)}{\partial t} > 0 \right\}, \\
\overline{BT}_{\alpha,h} &= \left\{ b : \min_{(t,x) \in I_{\alpha,h} \times X} \frac{\partial \Psi(t \cdot x, b)}{\partial t} > h / \tilde{f}, \max_{p=1,\ldots,s+1} \left( \frac{\max_{x \in X} |P(x)' b_p|}{h} \right) < \tilde{f} \right\},
\end{align*}
\]

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

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

i. \(b^* (\alpha | I)\) belongs to \(\overline{BT}_{\alpha,h} \subset BT_{\alpha,h}\) and for \(C\) small enough \(B(b^* (\alpha | I), C \ell h^{d_M/2+1})\) is a subset of \(\overline{BT}_{\alpha,h}\), for all \(\alpha\) in \([0, 1]\).
ii. For all $b$ in $BI_{\alpha,h}$ and all $u$ in $\Psi(I_{\alpha,h}|x,b)$

$$\frac{\partial \Phi(u|x,b)}{\partial b'} = -\frac{P(x,\Delta(u|x,b))}{\Psi(\Delta(u|x,b)|x,b)/h},$$

$$\frac{\partial \Phi(u|x,b)}{\partial u} = \frac{1}{\Psi(\Delta(u|x,b)|x,b)/h}.$$ 

iii. It holds that

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

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

and, recalling $b^*_1 (\alpha|I) = h\beta^*_1 (\alpha|I)$

$$\max_{(\alpha,x)\in [0,1] \times \mathcal{X}} \left| \Phi(x') \alpha \beta^*_1 (\alpha|I) - \alpha B^{(1)}(\alpha|x,I) \right| = o(h^{s+1}) ,$$

$$\max_{(\alpha,x)\in [0,1] \times \mathcal{X}} \max_{u \in \Psi[I_{\alpha,h}|x,b^*(\alpha|I)]} \left| \Phi(u|x,b^*(\alpha|I)) - G(u|x,I) \right| = o(h^{s+1}).$$
iv. There is a $C > 0$ such that for any $b_0$ and $b_1$ in $\mathcal{B}_{\alpha,h}$ and all $\alpha$ in $[0, 1]$

\[
\max_{(\alpha,x) \in [0,1] \times \mathcal{X}} \max_{t \in \mathcal{I}_{\alpha,h}} |\Psi(t|x, b_1) - \Psi(t|x, b_0)| ,
\]
\[
\max_{(\alpha,x) \in [0,1] \times \mathcal{X}} \max_{u \in \Psi[Z_{\alpha,h}|x,b_0] \cap \Psi[Z_{\alpha,h}|x,b_1]} |\Phi(u|x, b_1) - \Phi(u|x, b_0)| ,
\]
\[
\max_{(\alpha,x) \in [0,1] \times \mathcal{X}} \max_{u \in \Psi[Z_{\alpha,h}|x,b_0] \cap \Psi[Z_{\alpha,h}|x,b_1]} \left| \frac{\partial \Phi}{\partial u} (u|x, b_1) - \frac{\partial \Phi}{\partial u} (u|x, b_0) \right| ,
\]
\[
\max_{(\alpha,x) \in [0,1] \times \mathcal{X}} \max_{u \in \Psi[Z_{\alpha,h}|x,b_0] \cap \Psi[Z_{\alpha,h}|x,b_1]} |\Psi^{(1)}(\Delta(u|x, b_1)|x, b_1) - \Psi^{(1)}(\Delta(u|x, b_0)|x, b_0)| ,
\]

are all smaller or equal to $Ch^{-d,M/2} \|b_1 - b_0\|$.

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

\[
\Omega_h(\alpha) = \int_{\mathcal{L}_{\alpha,h}} \pi(t) \pi(t)^t K(t) dt = \left[ \int_{\mathcal{L}_{\alpha,h}} t^{p_1+p_2} K(t) dt, 0 \leq p_1, p_2 \leq s + 1 \right],
\]
\[
\Omega(0) = \int_{-1}^{0} \pi(t) \pi(t)^t K(t) dt, \quad \Omega(1) = \int_{0}^{1} \pi(t) \pi(t)^t K(t) dt,
\]
\[
\Omega_{1h}(\alpha) = \int_{\mathcal{L}_{\alpha,h}} t \pi(t) \pi(t)^t K(t) dt,
\]

While $\Omega_h(\alpha) \preceq \Omega$ for all $\alpha$ and $h$, it holds that for $h$ small enough $\Omega_h(\alpha) \succeq \Omega(0)$ for all $\alpha$ in $[0, 1/2]$ and $\Omega_h(\alpha) \succeq \Omega(1)$ for all $\alpha$ 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 $K_L^{-1/d,M} = O(h)$, $h$ small enough, all $I$ in $\mathcal{I}$, and any $C > 0$ small enough, (i) It
holds that $R^{(2)}(\cdot; \alpha, I)$ is continuously differentiable over $B(b^{*}(\alpha|I), Ch^{d_{\mathcal{M}}/2+1})$ with

$$
\max_{\alpha \in [0, 1]} \max_{b_1, b_0 \in B(b^{*}(\alpha|I), Ch^{d_{\mathcal{M}}/2+1})} \left\| R^{(2)}(b_1; \alpha, I) - R^{(2)}(b_0; \alpha, I) \right\| \leq O \left( h^{-d_{\mathcal{M}}/2} \right).
$$

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

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

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

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

if $h^s = o \left( h^{d_{\mathcal{M}}/2} \right),$

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

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

It then follows that the eigenvalues of $R^{(2)}(b; \alpha, I)$ stays bounded away from 0 and infinity uniformly in $\alpha$ 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 $\hat{R}(\cdot;\alpha,I)$ in a shrinking vicinity of $b^*(\alpha|I)$. In particular, Lemma B.3 implies that $\hat{R}(\cdot;\alpha,I)$ 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 $\log L = \frac{1}{L h^{d_\mathcal{M}+1}} = o(1)$.

Then, for any $C > 0$ small enough,

$$\max_{\alpha \in [0,1]} \max_{b \in \mathcal{B}(b^*(\alpha|I), Ch^{d_\mathcal{M}/2+1})} \left\| \hat{R}^{(2)}(b;\alpha,I) - \bar{R}^{(2)}(b;\alpha,I) \right\| = O_p \left( \frac{\log L}{L h^{d_\mathcal{M}+1}} \right)^{1/2}.$$

**Lemma B.4** Suppose Assumptions A, R-(i,ii) and S hold, and $\log L = \frac{1}{L h^{d_\mathcal{M}+1}} = o(1)$.

Then, for any $C > 0$,

$$\max_{\alpha \in [0,1]} \max_{b \in \mathcal{B}(b^*(\alpha|I), Ch^{d_\mathcal{M}/2+1})} \left\| \hat{R}^{(1)}(b;\alpha,I) - \bar{R}^{(1)}(b;\alpha,I) \right\| = O_p \left( \frac{\log L}{L h^{d_\mathcal{M}}} \right)^{1/2}.$$

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

$$\max_{\alpha \in [0,1]} \left\| \hat{R}^{(1)}(\bar{b}(\alpha|I);\alpha,I) \right\| = O_p \left( \frac{\log L}{L h^{d_\mathcal{M}}} \right)^{1/2}.$$

The next Lemma studies the leading term $\tilde{e}(\alpha|I)$ of $\hat{b}(\alpha|I) - \bar{b}(\alpha|I)$,

$$\tilde{e}(\alpha|I) = - \left[ \bar{R}^{(2)}(\bar{b}(\alpha|I);\alpha,I) \right]^{-1} \hat{R}^{(1)}(\bar{b}(\alpha|I);\alpha,I)$$
see Theorem B.9 below. Note that \( R(2) \) is not necessarily defined and invertible unless \( h^{s+1} = O(h^{d_{M}/2+1}) \) and \( \sup_{\alpha \in [0,1]} \| \tilde{b}(\alpha|I) - b^*(\alpha|I) \| = o(h^{s+1}) \) 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 \( 1/ (Lh^{d_{M}+1}) = o(1), \) \( s \geq d_{M}/2 \) and \( \sup_{\alpha \in [0,1]} \| \tilde{b}(\alpha|I) - b^*(\alpha|I) \| = o(h^{s+1}) \). Then (i) uniformly in \((\alpha, x)\) in \([0,1] \times \mathcal{X}\)

\[
\text{Var} \left[ P(x)' \hat{e}_0(\alpha|I) \right] = O \left( \frac{1}{Lh^{d_{M}}} \right)
\]

and \( \text{Var} \left[ P(x)' \hat{e}_1(\alpha|I) / h \right] = O \left( \frac{1}{Lh^{d_{M}+1}} \right) \) with \( \text{Var} \left[ \hat{e}_1(\alpha|I) / h \right] \) 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

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

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

### B.2 Asymptotic bias

The study of the bias \( \nabla (\alpha|x, I) - V(\alpha|x, I) \) 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.

**Lemma B.6** Let $\mathcal{F}(\cdot) : \mathbb{R}^D \to \mathbb{R}$ be a function. Suppose that there is a $x^* \in \mathbb{R}^D$ and some real numbers $\epsilon > 0$ and $C_0 > 0$ such that $\mathcal{F}(\cdot)$ is twice differentiable on $B(x^*, 2C_0\epsilon) = \{x \in \mathbb{R}^D; \|x - x^*\| < 2C_0\epsilon\}$. If, in addition,

i. $\|\mathcal{F}^{(1)}(x^*)\| \leq \epsilon$ and $\left\|\left[\mathcal{F}^{(2)}(x^*)\right]^{-1}\right\| \leq C_0$;

ii. There is a $C_1 > 0$ such that $\|\mathcal{F}^{(2)}(x) - \mathcal{F}^{(2)}(x')\| \leq C_1\|x - x'\|$ for all $x, x' \in B(x^*, 2C_0\epsilon)$;

iii. $C_0^2C_1\epsilon \leq 1/2$.

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

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 $\ell_1$ norm of the columns of the matrix

$$A_{\alpha, h} = E^{-1} \left[ I(I = I) \int_{I_{\alpha, h}}^T P(x_t, t) P(x_t, t)' K(t) dt \right]$$

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

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

\[
\text{bias}_h(\alpha|I) = \mathbb{E}^{-1} \left[ \frac{I_\ell = I \int_{L_{x,h}}^{T_{x,h}} P(x_\ell, t) P(x_\ell, t)' K(t) \, dt}{B(1)(\alpha|x_\ell, I_\ell)} \right] \\
\times \mathbb{E} \left[ \frac{I_\ell = I \, B^{(s+2)}(\alpha|x_\ell, I_\ell) \int_{L_{x,h}}^{T_{x,h}} t^{s+2} P(x_\ell, t) K(t) \, dt}{(s+2)!B(1)(\alpha|x_\ell, I_\ell)} \right],
\]

and

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

where the subvectors \(\text{bias}_{ph}(\alpha|I)\) are of dimension \(K_L\). While \(\text{bias}_h(\alpha|I)\) may not exist for \(\alpha = 0\), the function \(\text{Bias}_h(\alpha|I) = \alpha \text{bias}_h(\alpha|I)\) in (3.7) 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 \(h = O\left(K_L^{-1/d_M}\right)\) and \(s \geq d_M/2\). Then, for \(h\) small enough \(B(\alpha|I) = \arg\min_b R(b; \alpha, I)\) is unique for all \(\alpha\) in \([0, 1]\) and

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

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

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

\[
sup_{(\alpha, x, I) \in [0, 1] \times X \times I} \left| \alpha \left( B(\alpha|x, I) - B(\alpha|x, I) - h^{s+2} P(x)' \text{bias}_{0h}(\alpha|I) \right) \right| = o(h^{s+2})
\]
with \( \sup_{(\alpha, x, I) \in [0, 1] \times X \times \mathcal{I}} |P(x)' \alpha \text{bias}_{0h}(\alpha | I)| = O(1) \). Hence uniformly over \( x \) and \( \alpha \) in any compact subset of \( (0, 1] \),

\[
\bar{B}(\alpha | x, I) = B(\alpha | x, I) + h^{s+2} P(x)' \text{bias}_{0h}(\alpha | I) + o(h^{s+2}).
\]

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

\[
\sup_{\alpha \in [0, 1]} \sup_{b \in \mathcal{B}(\bar{b}(\alpha | I), Ch^{s+1})} \left| h^{-2(s+1)} \left\{ \hat{R}(b; \alpha, I) - \hat{R}(\bar{b}(\alpha | I); \alpha, I) - (b - \bar{b}(\alpha | I))' \hat{R}^{(1)}(\bar{b}(\alpha | I); \alpha, I) \right\} - \frac{h^{-2(s+1)}}{2} (b - \bar{b}(\alpha | I))' \hat{R}^{(2)}(\bar{b}(\alpha | I); \alpha, I) (b - \bar{b}(\alpha | I)) \right| = o_p(1).
\]

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

$$\max_{(\alpha, I) \in [0,1] \times \mathcal{I}} \| \tilde{R}^{(1)} (\mathbf{b}^* (\alpha|I) ; \alpha, I) \| = \epsilon_L \quad \text{with}$$

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

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

$$\max_{(\alpha, I) \in [0,1] \times \mathcal{I}} \| \tilde{R}^{(1)} (\mathbf{b}^* (\alpha|I) ; \alpha, I) \| = \epsilon_L$$

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

$$\max_{(\alpha, I) \in [0,1] \times \mathcal{I}} \| \tilde{R}^{(1)} (\mathbf{b}^* (\alpha|I) ; \alpha, I) \| = \epsilon_L$$

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

$$\max_{\alpha} | \theta' \tilde{R}^{(1)} (\mathbf{b}^* (\alpha|I) ; \alpha, I) |$$

$$= \mathbb{E} \left[ \mathbb{E} \left[ \mathbb{E} \left[ \right] \right] \right] \left[ \int_{L,h} G (P (x_t,t) \mathbf{b}^* (\alpha|I) |x_t, I_t) - G (B (\alpha + ht|x, I) |x_t, I_t) \right]$$

$$= \epsilon_L$$

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,

$$\sup_{(\alpha, I) \in [0,1] \times \mathcal{I}} \left\| \left[ \tilde{R}^{(2)} (\mathbf{b}^* (\alpha|I) ; \alpha, I) \right]^{-1} \right\| \leq C_0$$
Note that $s \geq d_M/2$ and $\epsilon_L = o(h^{s+1})$ gives that

$$\mathcal{B}(b^*(\alpha|I), 2C_0\epsilon_L) \subset \mathcal{B}(b^*(\alpha|I), Ch^{d_M/2+1})$$

for all $C_0, C > 0$ provided $L$ is large enough, for all $\alpha$ and all $I$. Condition (ii) in Lemma B.6 follows from Lemma B.2-(i) which ensures that for $C_1L = O(h^{d_M/2+1})$,

$$\|\overline{R}^{(2)}(b_1; \alpha, I) - \overline{R}^{(2)}(b_0; \alpha, I)\| \leq C_1L \|b_1 - b_0\|$$

for all $b_1, b_0$ in $\mathcal{B}(b^*(\alpha|I), 2C_0\epsilon_L)$ and all $\alpha, I$. For condition (iii) in Lemma B.6, $\epsilon_L = o(h^{s+1})$ and $s \geq d_M/2$ implies $C_0^2C_1L\epsilon_L = o(h^{s+d_M/2}) = o(1) < 1/2$ for $L$ large enough.

Hence Lemma B.6 ensures that, for $L$ large enough, all $\alpha$ and all $I$, there is a unique $\overline{b}(\alpha|I)$ in $\mathcal{B}(b^*(\alpha|I), 2C_0\epsilon_L)$ such that

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

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

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

133
Consider now $\alpha \overline{b} (\alpha|I) - \alpha b^* (\alpha|I)$. Define

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

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

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

$$= \int_0^1 g \left( B (\alpha + ht|x, I) + o (h^{s+1-d_M/2}) \right) |t, x, I| \, du$$

$$\geq (1 + o(1)) \max_{y \in [B(2h|x,I),B(1-2h|x,I)]} g (y|x, I) \geq C'' > 0$$

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

$$0 = \int \left( \int_{L_{\alpha,h}} \left\{ G \left[ \Psi \left( t|x, \overline{b} (\alpha|I) \right) \right] - (\alpha + ht) \right\} P (x, t) K(t) \, dt \right) f (x, I) \, dx$$

$$= \int \left( \int_{L_{\alpha,h}} \left\{ G \left[ \Psi \left( t|x, \overline{b} (\alpha|I) \right) \right] - G \left[ B (\alpha + ht|x, I) \right] \right\} P (x, t) K(t) \, dt \right) f (x, I) \, dx$$

$$= \int \left( \int_{L_{\alpha,h}} \overline{g} (\alpha|t, x, I) \left\{ \Psi \left( t|x, \overline{b} (\alpha|I) \right) - B (\alpha + ht|x, I) \right\} P (x, t) K(t) \, dt \right) f (x, I) \, dx$$

$$= \int \left( \int_{L_{\alpha,h}} \overline{g} (\alpha|t, x, I) \left\{ \Psi \left( t|x, \overline{b} (\alpha|I) \right) - \Psi \left( t|x, b^* (\alpha|I) \right) \right\} P (x, t) K(t) \, dt \right) f (x, I) \, dx$$

$$\quad + \int \left( \int_{L_{\alpha,h}} \overline{g} (\alpha|t, x, I) \left\{ \Psi \left( t|x, b^* (\alpha|I) \right) - B (\alpha + ht|x, I) \right\} P (x, t) K(t) \, dt \right) f (x, I) \, dx.$$
Since \( \{\Psi(t|x, b(\alpha|I)) - \Psi(t|x, b^*(\alpha|I))\} \) \( P(x,t) = P(x,t) P(x,t) P(x,t) K(t) dt \) \( f(x,I) \), by Assumption R-(i), and because \( g(\alpha|t,x,I), f(x,I) \) are bounded away from 0 and infinity

\[
\alpha (\mathcal{B}(\alpha|I) - b^*(\alpha|I)) = \left[ \int \left( \int_{L_{a,h}} h(\alpha|t,x,I) P(x,t) P(x,t) K(t) dt \right) f(x,I) dx \right]^{-1} \times \left( \int_{L_{a,h}} \bar{g}(\alpha|t,x,I) \left\{ \frac{(ht)^{s+2}}{(s+2)!} \alpha B^{(s+2)}(\alpha|I) + o(h^{s+2}) \right\} P(x,t) K(t) dt \right) f(x,I) dx
\]

uniformly in \( \alpha \) in \([0,1] \) by Lemma B.1-(iii). By Assumption R-(ii) which implies in particular

\[
\left\| \int (\int_{L_{a,h}} P(x,t) K(t) dt) f(x,I) \right\| = O(1), \text{ it follows}
\]

\[
\mathcal{B}(\alpha|I) - b^*(\alpha|I) = o(h^{s+1}) E^{-1} \left[ \frac{\mathbb{I}(I_{\ell} = I) \int_{L_{a,h}} P(x_t,t) P(x_t,t) K(t) dt}{B^{(1)}(\alpha|x_t,I_t)} \right] E \left[ \int_{L_{a,h}} P(x_t,t) K(t) dt \right],
\]

\[
\alpha (\mathcal{B}(\alpha|I) - b^*(\alpha|I)) = h^{s+2} \alpha \text{bias}_h(\alpha|I)
\]

\[
+ o(h^{s+2}) E^{-1} \left[ \frac{\mathbb{I}(I_{\ell} = I) \int_{L_{a,h}} P(x_t,t) P(x_t,t) K(t) dt}{B^{(1)}(\alpha|x_t,I_t)} \right] E \left[ \int_{L_{a,h}} P(x_t,t) K(t) dt \right],
\]

(B.5)

uniformly over \([0,1] \). Let

\[
A = A_{\alpha,h} = [A_1, \ldots A_{J_L}] = E^{-1} \left[ \frac{\mathbb{I}(I_{\ell} = I) \int_{L_{a,h}} P(x_t,t) P(x_t,t) K(t) dt}{B^{(1)}(\alpha|x_t,I_t)} \right]
\]

135
be a \( J_L \times J_L \) matrix with columns \( A_j, j = 1, \ldots, J_1, |A_j|_1 \) the associated \( \ell_1 \) norm and \( |A|_{1, \infty} = \max_{j \leq J_L} |A_j|_1 \), \( S \) a selection matrix which selects some columns of \( A \), \( a, b \) some conformable vectors and \( |a|_\infty \) the largest entry of \( a \).

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

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

\[
\sup_{(\alpha, x) \in [0, 1] \times \mathcal{X}} \left| P' (x) S \text{bias}_h (\alpha | I) \right| \leq C \left( \max_{x \in \mathcal{X}} \sum_{k=1}^{K_L} |P_k (x)| \right) \times \max_{1 \leq k \leq K_L} \int |P_k (x)| \, dx = O \left( 1 \right),
\]

\[
\sup_{(\alpha, x) \in [0, 1] \times \mathcal{X}} \left| P' (x) S \text{AE} \left[ \int_{\mathcal{I}_{\alpha, h}} |P (x, t)| K (t) \, dt \right] \right| \leq C \left( \max_{x \in \mathcal{X}} \| P (x) \| \right) \times \max_{1 \leq k \leq K_L} \int |P_k (x)| \, dx = O \left( 1 \right).
\]

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

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

136
\[
\sup_{(\alpha,x)\in [0,1] \times \mathcal{X}} \left| \alpha \left( \bar{B}^{(1)}(\alpha|x,I) - B^{(1)}(\alpha|x,I) \right) - h^{s+1} P'(x) \alpha S_1 \text{bias}_h(\alpha|I) \right|
\]

\[
= \sup_{(\alpha,x)\in [0,1] \times \mathcal{X}} \frac{1}{h} \left| \alpha P'(x) S_1 \left( \bar{b}(\alpha|I) - b^*(\alpha|I) - h^{s+2} P'(x) \text{bias}_h(\alpha|I) \right) \right|
\]

\[
+ \sup_{(\alpha,x)\in [0,1] \times \mathcal{X}} \frac{1}{h} \left| \alpha \left( P'(x) b_1^*(\alpha|I) - h B^{(1)}(\alpha|x,I) \right) \right|
\]

\[
= o(h^{s+1}).
\]

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

**B.3 Bahadur representation**

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

\[
\hat{e}(\alpha|I) = - \left( \bar{R}^{(2)}(\bar{b}(\alpha|I); \alpha, I) \right)^{-1} \hat{R}^{(1)}(\bar{b}(\alpha|I); \alpha, I), \tag{B.6}
\]

\[
\hat{d}(\alpha|I) = \hat{b}(\alpha|I) - \bar{b}(\alpha|I) - \hat{e}(\alpha|I). \tag{B.7}
\]

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

**Theorem B.9** Suppose Assumptions A, R-(i,ii) and S hold, \( s \geq d_M/2 \) and

\[
\frac{\log L}{L h^{2(d_M+1)}} = o(1).
\]

137
Then

\[
\max_{\alpha \in [0,1]} \left\| \frac{Lh^{d_M+(d_M^1)} / 2}{(h + \alpha (1 - \alpha))^{1/2} \log L} \left\{ \hat{\mathcal{B}}(\alpha|I) - \mathcal{B}(\alpha|I) + \left( \mathcal{R}^{(2)} (\mathcal{B}(\alpha|I); \alpha, I) \right)^{-1} \mathcal{R}^{(1)} (\mathcal{B}(\alpha|I); \alpha, I) \right\} \right\| = O_{\mathbb{P}} (1)
\]

with a diverging normalization term \( Lh^{d_M+(d_M^1)/2} / \log L \). Moreover, for \( \tilde{d}(\alpha|I) \) as in (B.7),

\[
\sup_{(\alpha, x) \in [0,1] \times \mathcal{X}} \left( Lh^{d_M+1} \right)^{1/2} \left\| P' (x) \hat{\mathcal{B}}_0 (\alpha|I) \right\| = O_{\mathbb{P}} \left( \frac{h^{1/2} \log L}{(Lh^{2d_M+(d_M^1)}/2)} \right),
\]

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

**Proof of Theorem B.9.** We first introduce some renormalizations. Let, for \( \hat{\mathcal{E}}(\alpha|I) \) as in (B.6),

\[
\varrho_{\alpha L} = \frac{(h + \alpha (1 - \alpha))^{1/2} \log L}{Lh^{d_M+(d_M^1)/2}},
\]

\[
\hat{\mathcal{R}} (d; \alpha, I) = \hat{\mathcal{R}} (\mathcal{B}(\alpha|I) + \hat{\mathcal{E}}(\alpha|I) + \varrho_{\alpha L} d; \alpha, I) - \hat{\mathcal{R}} (\mathcal{B}(\alpha|I) + \hat{\mathcal{E}}(\alpha|I); \alpha, I),
\]

which is such that \( \varrho_{\alpha L} = o (1) \) by \( \log L / (Lh^{2(d_M+1)}) = o (1) \)

\[
\frac{\hat{d}(\alpha|I)}{\varrho_{\alpha L}} = \arg \min_d \hat{\mathcal{R}} (d; \alpha, I).
\]
It follows that,

\[
\left\{ \sup_{\alpha \in [0,1]} \left\| \frac{d(\alpha|I)}{\ell_{\alpha L}} \right\| \geq t \right\} = \bigcup_{\alpha \in [0,1]} \left\{ \left\| \frac{d(\alpha|I)}{\ell_{\alpha L}} \right\| \geq t \right\}
\]

\[
\subset \bigcup_{\alpha \in [0,1]} \left\{ \inf_{\|d\| \geq t} \hat{R}(d; \alpha, I) \leq \inf_{\|d\| \leq t} \hat{R}(d; \alpha, I) \right\} \subset \bigcup_{\alpha \in [0,1]} \left\{ \inf_{\|d\| \geq t} \hat{R}(d; \alpha, I) \leq 0 \right\}
\]

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

\[
\hat{R}(d; \alpha, I) = \frac{\|d\|}{t} \left\{ t \frac{\|d\|}{\|d\|} \hat{R}\left( \frac{d}{\|d\|}; \alpha, I \right) + \left( 1 - \frac{t}{\|d\|} \right) \hat{R}(0; \alpha, I) \right\}
\]

\[
\geq \frac{\|d\|}{t} \hat{R}\left( \frac{d}{\|d\|}; \alpha, I \right)
\]

so that \( \inf_{\|d\| \geq t} \hat{R}(d; \alpha, I) \leq 0 \) implies \( \inf_{\|d\| = t} \hat{R}(d; \alpha, I) \leq 0 \) and then

\[
\left\{ \sup_{\alpha \in [0,1]} \left\| \frac{d(\alpha|I)}{\ell_{\alpha L}} \right\| \geq t \right\} \subset \left\{ \inf_{\alpha \in [0,1]} \inf_{\|d\| = t} \hat{R}(d; \alpha, I) \leq 0 \right\}.
\]

(B.8)
Thus it is sufficient to consider those $d$ with $\|d\| = t$. The expression of $\hat{R}(d; \alpha, I)$ gives, using two Taylor expansions with integral remainder,

$$
\hat{R}(d; \alpha, I) = \varrho_{\alpha L}d'' \hat{R}^{(1)}(\hat{b}(\alpha|I) + \hat{e}(\alpha|I); \alpha, I) \\
+ \varrho_{\alpha L}d'' \int_0^1 \hat{R}^{(2)}(\hat{b}(\alpha|I) + \hat{e}(\alpha|I) + u\varrho_{\alpha L}d; \alpha, I) (1 - u) \, du \, d' \\
= \varrho_{\alpha L}d'' \hat{R}^{(1)}(\hat{b}(\alpha|I); \alpha, I) \\
+ \varrho_{\alpha L}d'' \int_0^1 \hat{R}^{(2)}(\hat{b}(\alpha|I) + u\hat{e}(\alpha|I); \alpha, I) \, du \hat{e}(\alpha|I) \\
+ \varrho_{\alpha L}^2d'' \int_0^1 \hat{R}^{(2)}(\hat{b}(\alpha|I) + \hat{e}(\alpha|I) + u\varrho_{\alpha L}d; \alpha, I) (1 - u) \, du \, d' .
$$

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

$$
\hat{R}(d; \alpha, I) = \varrho_{\alpha L}d'' \left\{ \int_0^1 \hat{R}^{(2)}(\hat{b}(\alpha|I) + u\hat{e}(\alpha|I); \alpha, I) - \hat{R}^{(2)}(\hat{b}(\alpha|I); \alpha, I) \right\} \, du \hat{e}(\alpha|I) \\
+ \varrho_{\alpha L}^2d'' \left\int_0^1 \hat{R}^{(2)}(\hat{b}(\alpha|I) + \hat{e}(\alpha|I) + u\varrho_{\alpha L}d; \alpha, I) (1 - u) \, du \right\} d' .
$$

Lemma B.4 and (B.4) with $s \geq d_M/2$, $\log L / (L h^{2(d_M+1)}) = o(1)$, Lemma B.2-(ii) give

$$
\sup_{\alpha \in [0,1]} \left\| \frac{\hat{e}(\alpha|I)}{(h + \alpha (1 - \alpha))^{1/2}} \right\| = O_P \left( \left( \frac{\log L}{L h^{d_M}} \right)^{1/2} \right) = o_P \left( h^{d_M/2+1} \right) .
$$
Lemmas B.3 and B.2-(i) then imply for the first item in $\hat{R}(d; \alpha, I)$, uniformly in $\alpha$ and $d$

with $\|d\| = t$,

$$
\left| \varrho_{aL}d' \right| \left\{ \int_0^1 \left\{ \bar{R}^{(2)} (\bar{b} (\alpha | I) + u \bar{e} (\alpha | I) ; \alpha, I) - \bar{R}^{(2)} (\bar{b} (\alpha | I) ; \alpha, I) \right\} du \right\} \hat{e} (\alpha | I)

= \left| \varrho_{aL}d' \right| \int_0^1 \left\{ \bar{R}^{(2)} (\bar{b} (\alpha | I) + u \bar{e} (\alpha | I) ; \alpha, I) - \bar{R}^{(2)} (\bar{b} (\alpha | I) ; \alpha, I) \right\} du \left( \frac{\log L}{Lh^d + 1} \right) \frac{1}{2} \hat{e} (\alpha | I)

= \left| \varrho_{aL}d' \right| \left\{ \bar{R}^{(2)} (\bar{b} (\alpha | I) + u \bar{e} (\alpha | I) ; \alpha, I) - \bar{R}^{(2)} (\bar{b} (\alpha | I) ; \alpha, I) \right\} du \left( \frac{\log L}{Lh^d + 1} \right) \frac{1}{2} \hat{e} (\alpha | I)

= \left| \varrho_{aL}d' \right| \left\{ \bar{R}^{(2)} (\bar{b} (\alpha | I) + u \bar{e} (\alpha | I) ; \alpha, I) - \bar{R}^{(2)} (\bar{b} (\alpha | I) ; \alpha, I) \right\} du \left( \frac{\log L}{Lh^d + 1} \right) \frac{1}{2} \hat{e} (\alpha | I)

= t \varrho_{aL} \left\{ \bar{R}^{(2)} (\bar{b} (\alpha | I) + u \bar{e} (\alpha | I) ; \alpha, I) - \bar{R}^{(2)} (\bar{b} (\alpha | I) ; \alpha, I) \right\} du \left( \frac{\log L}{Lh^d + 1} \right) \frac{1}{2} \hat{e} (\alpha | I)

= t \varrho_{aL} \left( \frac{\log L}{Lh^d + 1} \right) \frac{1}{2} \hat{e} (\alpha | I)

= t \varrho_{aL} \left( \frac{\log L}{Lh^d + 1} \right) \frac{1}{2} \hat{e} (\alpha | I) = t \varrho_{aL} O_P (1).

Observe that the condition $\log L / (Lh^d + 1) = o (1)$ implies

$$
\log L \left( \frac{Lh^d + 1}{Lh^d + 1} \right) = o (1) \text{ and then } \varrho_{aL} = \left( \frac{(h + \alpha (1 - \alpha)) \log L}{Lh^d} \right)^{1/2}.
$$
Lemmas B.3 and B.2 then imply for the second item in $\widehat{R}(d; \alpha, I)$, uniformly in $\alpha$ and $d$ with $\|d\| = t$,

\[
\varrho_{\alpha L} d' \left[ \int_0^1 \widehat{R}^{(2)} (\tilde{\mathbf{b}}(\alpha|I) + \tilde{\mathbf{e}}(\alpha|I) + \varrho_{\alpha L} \mathbf{d}; \alpha, I) (1 - u) \, du \right] d'
\]

\[
= \varrho_{\alpha L} d' \left[ \int_0^1 \left\{ \widehat{R}^{(2)} (\tilde{\mathbf{b}}(\alpha|I) + \tilde{\mathbf{e}}(\alpha|I) + \varrho_{\alpha L} \mathbf{d}; \alpha, I) + O_P \left( \left( \frac{\log L}{Lh^{dM+1}} \right)^{1/2} \right) \right\} (1 - u) \, du \right] d'
\]

\[
= \varrho_{\alpha L} d' \left[ \int_0^1 \left\{ \widehat{R}^{(2)} (\tilde{\mathbf{b}}(\alpha|I)) + tO_P \left( \left( \frac{\log L}{Lh^{2dM}} \right)^{1/2} \right) + O_P \left( \left( \frac{\log L}{Lh^{dM+1}} \right)^{1/2} \right) \right\} (1 - u) \, du \right] d'
\]

\[
\geq C \varrho_{\alpha L}^2 t^2 (1 + tO_P(1)).
\]

Now (B.8) gives, with $O_P(1)$ and $o_P(1)$ which are uniform in $\alpha$,

\[
\mathbb{P} \left( \sup_{\alpha \in [0,1]} \left\| \frac{\widehat{d}(\alpha|I)}{\varrho_{\alpha L}} \right\| \geq t \right) \leq \mathbb{P} \left( \inf_{\alpha \in [0,1]} \left\{ C \varrho_{\alpha L}^2 t^2 (1 + tO_P(1)) + t\varrho_{\alpha L}^2 O_P(1) \right\} \leq 0 \right)
\]

\[
= \mathbb{P} (Ct (1 + tO_P(1)) + O_P(1) \leq 0)
\]

\[
\leq \mathbb{P} (t (1 + tO_P(1)) \leq |O_P(1)|)
\]

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 $\max_{\alpha \in [0,1]} \varrho_{\alpha L} = \log L/Lh^{dM+(dM+1)/2}$
so that, uniformly in $\alpha$ and $x$,

$$
\left| (Lh^{d_{M}+1})^{1/2} P(x)^t \tilde{d}_0 (\alpha|I) \right| = (Lh)^{1/2} h^{d_{M}/2} \max_{x \in \mathcal{X}} \|P(x)\|_2 \left\| \tilde{d} (\alpha|I) \right\| \\
= O_{\mathbb{P}} \left( (Lh)^{1/2} \omega_L \right) = O_{\mathbb{P}} \left( \frac{h^{1/2} \log L}{(Lh^{2d_{M}+(d_{M}+1)^{1/2}})_{\omega L}} \right),
$$

$$
\left| (Lh^{d_{M}+1})^{1/2} P(x)^t \frac{\tilde{d}_1 (\alpha|I)}{h} \right| = O_{\mathbb{P}} \left( \left( \frac{L}{h} \right)^{1/2} \omega_L \right) = O_{\mathbb{P}} \left( \frac{\log L}{(Lh^{2d_{M}+(d_{M}+1)^{1/2}})_{\omega L}} \right).
$$

This ends the proof of the Theorem. \hfill \Box

## 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 $s_0 = [1,0,\ldots,0]$, $S_0 = s_0 \otimes \text{Id}_{K_L}$, $S_1 = s_1 \otimes \text{Id}_{K_L}$ so that $\tilde{\beta}_j (\alpha|I) = S_j \tilde{\beta} (\alpha|I)$, $j = 0,1$ and

$$
\tilde{V} (\alpha|x,I) = P(x)^t \left[ S_0 + \frac{\alpha S_1}{h (I-1)} \right] \tilde{b} (\alpha|I),
$$

$$
\overline{V} (\alpha|x,I) = P(x)^t \left[ S_0 + \frac{\alpha S_1}{h (I-1)} \right] \overline{b} (\alpha|I).
$$

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

$$
\tilde{V} (\alpha|x,I) = \overline{V} (\alpha|x,I) + P(x)^t \left[ S_0 + \frac{\alpha S_1}{h (I-1)} \right] \hat{e} (\alpha|I) \tag{B.9}
$$
which is such, for \( \hat{d}(\alpha|I) \) as in (B.7),
\[
\hat{V}(\alpha|x, I) - \bar{V}(\alpha|x, I) = P(x)' \left[ S_0 + \frac{\alpha S_1}{h(I - 1)} \right] \hat{d}(\alpha|I).
\]

As the eigenvalues of \( \int_X P(x) P(x)' dx \) are bounded away from infinity under Assumption R-(i)
\[
\int_X \int_0^1 \left( \hat{V}(\alpha|x, I) - \bar{V}(\alpha|x, I) \right)^2 d\alpha dx = O \left( \frac{\sup_{\alpha \in [0,1]} \| \hat{d}(\alpha|I) \|^2}{h^2} \right)
\]
\[
= O_P \left( \left( \frac{\log L}{Lh^{d_M+1+(d_M \vee 1)/2}} \right)^2 \right)
\]
by Theorem B.9, which gives (3.4) since, by Assumption H,
\[
\frac{Lh^{d_M+1}}{\log L} \left( \frac{\log L}{Lh^{d_M+1+(d_M \vee 1)/2}} \right)^2 = \frac{\log L}{Lh^{d_M+1+(d_M \vee 1)}} = o \left( \frac{\log L}{Lh^{2(d_M+1)}} \right) = o(1).
\]
That \( \text{bias}_{IL}^2 = O(1) \) and \( \Sigma_{IL} = O(1) \) similarly follow from Assumption R-(i) and Proposition 3-(i).

It holds since \( \mathbb{E}[\hat{e}(\alpha|I)] = \bar{e}(\alpha|I); \alpha, I = 0 \) for all \( \alpha \) in \([0,1]\)
\[
\mathbb{E} \left[ \int_X \int_0^1 \left( \hat{V}(\alpha|x, I) - V(\alpha|x, I) \right)^2 d\alpha dx \right] = \int_X \int_0^1 (\bar{V}(\alpha|x, I) - V(\alpha|x, I))^2 d\alpha dx
\]
\[
+ \int_X \int_0^1 \mathbb{E} \left[ \left( P(x)' \left[ S_0 + \frac{\alpha S_1}{h(I - 1)} \right] \hat{e}(\alpha|I) \right)^2 \right] d\alpha dx.
\]
For the bias part, Theorem B.8 gives

\[
\int_{\mathcal{X}} \int_{0}^{1} (V(\alpha|x, I) - V(\alpha|x, I))^{2} \, d\alpha \, dx = \int_{\mathcal{X}} \int_{0}^{1} \left( \frac{h^{s+1}P(x)\alpha \text{bias}_{ih}(\alpha|I)}{I - 1} + o(h^{s+1}) \right)^{2} \, d\alpha \, dx
\]

\[
= h^{2(s+1)} \int_{\mathcal{X}} \int_{0}^{1} \left( \frac{P(x)\alpha \text{bias}_{ih}(\alpha|I)}{I - 1} \right)^{2} \, d\alpha \, dx + o(h^{2(s+1)}),
\]

Since \(\alpha \text{bias}_{ih}(\alpha|I) / (I - 1)\) differs from \(\text{bias}(\alpha|I)\) for \(\alpha\) in \([0, h]\) or \([1 - h, 1]\), it follows

\[
\int_{\mathcal{X}} \int_{0}^{1} (V(\alpha|x, I) - V(\alpha|x, I))^{2} \, d\alpha \, dx = h^{2(s+1)} \int_{\mathcal{X}} \int_{0}^{1} (P(x)\text{bias}(\alpha|I))^{2} \, d\alpha \, dx + o(h^{2(s+1)})
\]

\[
= h^{2(s+1)} \text{bias}_{IL}^{2} + o(h^{2(s+1)}).
\]

Arguing similarly with Lemma B.5-(i) yields

\[
\int_{\mathcal{X}} \int_{0}^{1} \mathbb{E} \left[ \left( P(x)' \left[ S_{0} + \frac{\alpha S_{1}}{h(I - 1)} \right] \bar{e}(\alpha|I) \right)^{2} \right] \, d\alpha \, dx
\]

\[
= \int_{\mathcal{X}} \int_{0}^{1} \mathbb{E} \left[ \left( \frac{P(x)' \alpha \bar{e}_{1}(\alpha|I)}{h(I - 1)} \right)^{2} \right] \, d\alpha \, dx + O \left( \frac{1}{Lh^{dM}} \right)
\]

\[
= \frac{\sigma_{IL}^{2}}{Lh^{dM+1}} + o \left( \frac{1}{Lh^{dM+1}} \right).
\]

Substituting in the bias-variance decomposition of the integrated mean squared error ends the proof of the Theorem. \(\Box\)

**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

$$\max_{x \in \mathcal{X}} \left( P(x) \right) = O \left( \max_{x \in \mathcal{X}} \| P(x) \|^2 \right) = O \left( h^{-d_M} \right).$$

By Theorem B.9, Lemma B.5, Assumptions R-(i), H, and using the same notations than in the proof of Theorem 4

$$\left( Lh^{d_M+1} \right)^{1/2} \left( \hat{V}(\alpha|x, I) - V(\alpha|x, I) - \frac{P'(x) \alpha S_1 \hat{e}(\alpha|I)}{h(I-1)} - (\hat{V}(\alpha|x, I) - V(\alpha|x, I)) \right)$$

$$= \left( Lh^{d_M+1} \right)^{1/2} \left\{ P'(x) \hat{e}_0(\alpha|I) + P'(x) \left[ S_0 + \frac{\alpha S_1}{h(I-1)} \right] \hat{d}(\alpha|I) \right\}$$

$$= \left( Lh^{d_M+1} \right)^{1/2} \left\{ O_P \left( \frac{1}{(Lh^{d_M})^{1/2}} \right) + O \left( \frac{\| P'(x) \hat{d}(\alpha|I) \|}{h} \right) \right\}$$

$$= O_P \left( h^{1/2} + \left( \frac{\log^2 L}{Lh^{2d_M-1+(d_M+1)}} \right)^{1/2} \right) = o_P(1).$$

Since $\hat{V}(\alpha|x, I) - V(\alpha|x, I) = h^{s+1} P(x) \hat{Bias}_h(\alpha|I) + o(h^{s+1})$, it remains to show that

$$\left( \frac{Lh}{P(x)' \Sigma_h(\alpha|I) P(x)} \right)^{1/2} \frac{\alpha P'(x) S_1 \hat{e}(\alpha|I)}{h(I-1)} \rightarrow N(0,1).$$

Write

$$\left( \frac{Lh}{P(x)' \Sigma_h(\alpha|I) P(x)} \right)^{1/2} \frac{\alpha P'(x) S_1 \hat{e}(\alpha|I)}{h(I-1)} = \sum_{\ell=1}^L r_\ell(\alpha|x, I)$$

146
with \( r_\ell (\alpha|x, I) = \mathbb{I}(I_\ell = I) \sum_{i=1}^{I_\ell} r_{i\ell} (\alpha|x, I) \) and

\[
    r_{i\ell} (\alpha|x, I) = \left( \frac{\alpha^2}{LIh(I-1)^2} \right)^{1/2} \frac{P(x)'}{P(x)' \Sigma h(\alpha|I) P(x)}^{1/2} S_1 \left[ \mathbb{R}^{(2)} (\mathbb{B}(\alpha|x, I); \alpha, I) \right]^{-1}
    \times \int_{\frac{1-\alpha}{2h}}^{1-\alpha} \{ \mathbb{I}(B_{i\ell} \leq P(x_\ell, t) \mathbb{B}(\alpha|x, I)) - (\alpha + ht) \} P(x_\ell, t) K(t) dt.
\]

Since \( \mathbb{E}[r_\ell (\alpha|x, I)] = 0 \) and \( \max_{1 \leq \ell \leq L} \text{Var}[r_\ell (\alpha|x, I)] - 1 \equiv o(1) \), it is sufficient to show that \( \max_{1 \leq \ell \leq L} \mathbb{E}[r^3_\ell (\alpha|x, I)] = o(1) \) 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),

\[
    |r_{i\ell} (\alpha|x, I)| \leq \frac{C}{(Lh)^{1/2}} \frac{\|P(x)\|}{\|P(x)\|} \times \max_{x \in \mathcal{X}} \|P(x)\| = O\left( \frac{1}{(Lh^{d+M+1})^{1/2}} \right).
\]

It follows that by Assumption H

\[
    \max_{1 \leq \ell \leq L} \mathbb{E}[r^3_\ell (\alpha|x, I)] \leq I \max_{1 \leq \ell \leq L, 1 \leq i \leq I_\ell} |r_{i\ell} (\alpha|x, I)| \max_{1 \leq \ell \leq L} \mathbb{E}[r^2_\ell (\alpha|x, I)]
    = O\left( \frac{1}{(Lh^{d+M+1})^{1/2}} \right) = o(1).
\]

This ends the proof of the Theorem. \( \square \)
Proof of Theorem 6.  By Theorems B.8 and B.9, Lemma B.5 and using the notations of the proof of Theorem 4

\[
\sup_{(\alpha, x) \in [0,1] \times X} \left| \hat{B}(\alpha|x, I) - B(\alpha|x, I) \right| \\
\leq \sup_{(\alpha, x) \in [0,1] \times X} \left| P(x)' S_0 \left[ \hat{b}(\alpha|I) - \bar{b}(\alpha|I) \right] \right| + \sup_{(\alpha, x) \in [0,1] \times X} \left| \bar{B}(\alpha|x, I) - B(\alpha|x, I) \right| \\
\leq \sup_{(\alpha, x) \in [0,1] \times X} \left| P(x)' \hat{e}_0(\alpha|I) \right| + \sup_{(\alpha, x) \in [0,1] \times X} \left| P(x)' \hat{d}_0(\alpha|I) \right| + o(h^{s+1}) \\
= O_P \left( \frac{\log L}{Lh^{d_M}} \right)^{1/2} \left\{ 1 + \left( \frac{\log L}{Lh^{2d_M+1+d_M/1}} \right)^{1/2} \right\} + o(h^{s+1}) \\
= O_P \left( \frac{\log L}{Lh^{d_M}} \right)^{1/2} + o(h^{s+1}) \\
\]

\[
\sup_{(\alpha, x) \in [0,1] \times X} \left| \hat{V}(\alpha|x, I) - V(\alpha|x, I) \right| \\
\leq \sup_{(\alpha, x) \in [0,1] \times X} \left| P(x)' \left( S_0 + \frac{\alpha}{h} S_1 \right) \left[ \hat{b}(\alpha|I) - \bar{b}(\alpha|I) \right] \right| + \sup_{(\alpha, x) \in [0,1] \times X} \left| \bar{V}(\alpha|x, I) - V(\alpha|x, I) \right| \\
\leq \sup_{(\alpha, x) \in [0,1] \times X} \left| P(x)' \hat{e}_0(\alpha|I) \right| + \sup_{(\alpha, x) \in [0,1] \times X} \left| P(x)' \frac{\hat{e}_1(\alpha|I)}{h} \right| + \sup_{(\alpha, x) \in [0,1] \times X} \left| P(x)' \left( \hat{d}_0 + \frac{\alpha \hat{d}_1(\alpha|I)}{h} \right) \right| + O(h^{s+1}) \\
= O_P \left( \frac{\log L}{Lh^{d_M+1}} \right)^{1/2} \left\{ 1 + \left( \frac{\log L}{Lh^{2d_M+1+d_M/1}} \right)^{1/2} \right\} + O(h^{s+1}) \\
= O_P \left( \frac{\log L}{Lh^{d_M+1}} \right)^{1/2} + O(h^{s+1}).
\]

This ends the proof of the Theorem.
Proof of Corollary 7. Let

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

Then, since \( \widehat{A}_{i\ell} = \arg\min_{\alpha \in [0,1]} B_{i\ell} - \widehat{B} (\alpha|x, I_\ell) \) and \( B_{i\ell} = B (A_{i\ell}|x, I_\ell) \) where \( B (\cdot|x, I_\ell) \) is increasing,

$$\left\{ \left| \widehat{A}_{i\ell} - A_{i\ell} \right| \geq t \right\} \subseteq \left\{ \min_{\alpha \in [0,1]: |A_{i\ell} - \alpha| \geq t} \left| B_{i\ell} - \widehat{B} (\alpha|x, I_\ell) \right| \leq \min_{\alpha \in [0,1]: |A_{i\ell} - \alpha| < t} \left| B_{i\ell} - \widehat{B} (\alpha|x, I_\ell) \right| \right\} \subseteq \left\{ \min_{\alpha \in [0,1]: |A_{i\ell} - \alpha| \geq t} \left| B (A_{i\ell}|x, I_\ell) - B (\alpha|x, I_\ell) \right| - \widehat{\theta}_B \leq \widehat{\theta}_B \right\} \subseteq \left\{ t \times \min_{(\alpha,x,I) \in [0,1] \times \mathcal{X} \times \mathcal{I}} B^{(1)} (\alpha|x, I) \leq 2\widehat{\theta}_B \right\},$$

and Proposition 3 implies that, for \( C > 2/\min_{(\alpha,x,I) \in [0,1] \times \mathcal{X} \times \mathcal{I}} B^{(1)} (\alpha|x, I), \)

$$\max_{\ell=1,\ldots,L} \max_{i=1,\ldots,I_{\ell}} \left| \widehat{A}_{i\ell} - A_{i\ell} \right| \leq C\widehat{\theta}_B.$$
Hence, since $\max_{(\alpha,x,I) \in [0,1] \times \mathcal{X} \times \mathcal{I}} V^{(1)}(\alpha|x,I) < \infty$,

\[
\max_{\ell=1,\ldots,L} \max_{i=1,\ldots,I} \left| \hat{V}_{i\ell} - V_{i\ell} \right| = \max_{\ell=1,\ldots,L} \max_{i=1,\ldots,I} \left| \hat{V} \left( \hat{A}_{i\ell}|x_{\ell},I_{\ell} \right) - V \left( A_{i\ell}|x_{\ell},I_{\ell} \right) \right|
\]
\[
\leq \max_{(\alpha,x,I) \in [0,1] \times \mathcal{X} \times \mathcal{I}} \left| \hat{V}(\alpha|x,I) - V(\alpha|x,I) \right| + \max_{\ell=1,\ldots,L} \max_{i=1,\ldots,I} \left| V \left( \hat{A}_{i\ell}|x_{\ell},I_{\ell} \right) - V \left( A_{i\ell}|x_{\ell},I_{\ell} \right) \right|
\]
\[
\leq \hat{\theta} + C\hat{\theta}.
\]

This gives the rate stated in the Corollary.

References


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),

\[
\begin{align*}
B (\alpha + ht|x, I) - P (x, t)' b^* (\alpha|I) \\
&= B (\alpha + ht|x, I) - \sum_{k=1}^{K_L} P_k (x) \beta_k (\alpha + ht|I) \\
&\quad + \sum_{k=1}^{K_L} P_k (x) \beta_k (\alpha + ht|I) - \sum_{k=1}^{K_L} P_k (x) \sum_{p=0}^{s+1} \frac{(ht)^p}{p!} \beta_k^{(p)} (\alpha|I) \\
&= B (\alpha + ht|x, I) - \sum_{k=1}^{K_L} P_k (x) \beta_k (\alpha + ht|I) \\
&\quad + \sum_{k=1}^{K_L} P_k (x) \left( \beta_k (\alpha + ht|I) - \sum_{p=0}^{s} \frac{(ht)^p}{p!} \beta_k^{(p)} (\alpha|I) \right) - \frac{(ht)^{s+1}}{(s+1)!} \sum_{k=1}^{K_L} P_k (x) \beta_k^{(s+1)} (\alpha|I) .
\end{align*}
\]

A Taylor expansion with integral remainder gives

\[
\beta_k (\alpha + ht|I) - \sum_{p=0}^{s} \frac{(ht)^p}{p!} \beta_k^{(p)} (\alpha|I) = \frac{(ht)^{s+1}}{s!} \int_0^1 \beta_k^{(s+1)} (\alpha + uht|I) (1 - u)^s du
\]

151
so that

\[
B(x + ht|x, I) - P(x, t) b^* (\alpha|I) = B(x + ht|x, I) - \sum_{k=1}^{K_L} P_k(x) \beta_k (\alpha + ht|I)
\]

\[
+ \frac{(ht)^{s+1}}{s!} \int_0^1 \left\{ \sum_{k=1}^{K_L} P_k(x) \beta_k^{(s+1)} (\alpha + uht|I) - B^{(s+1)} (\alpha + uht|I) \right\} (1 - u)^s \, du
\]

\[
+ \frac{(ht)^{s+1}}{s!} \int_0^1 \left\{ B^{(s+1)} (\alpha + uht|x, I) - B^{(s+1)} (\alpha|x, I) \right\} (1 - u)^s \, du
\]

\[
+ \frac{(ht)^{s+1}}{(s+1)!} \left\{ B^{(s+1)} (\alpha|x, I) - \sum_{k=1}^{K_L} P_k(x) \beta_k^{(s+1)} (\alpha|x, I) \right\}.
\]

Hence since \( B^{(s+1)} (\alpha|x, I) \) is continuous, by Property S and Proposition 3

\[
\max_{(\alpha,x) \in [0,1] \times x' \in x} \max_t |B(x + ht|x, I) - P(x, t) b^* (\alpha|I)| = o\left(h^{s+1}\right) + o \left( \frac{K_L}{\delta_{LM}} \right)
\]

\[
= o\left(h^{s+1}\right)
\]  \hspace{1cm} (C.1)

since \( K_L^{-1/\delta_{LM}} = O(h) \). Observe also that, uniformly in \( \alpha, x \) and \( t \) as above,

\[
\frac{\partial}{\partial t} [P(x, t) b^* (\alpha|I)] = \sum_{p=1}^{s+1} h^p \frac{t^{p-1}}{(p-1)!} \sum_{k=1}^{K_L} P_k(x) \beta_k^{(p)} (\alpha|I)
\]

\[
= h \left( B^{(1)} (\alpha|x, I) + o(1) \right) + h^2 \left( \sum_{p=2}^{s+1} h^{p-2} \frac{t^{p-1}}{(p-1)!} B^{(p)} (\alpha|x, I) + o(1) \right)
\]

\[
= h B^{(1)} (\alpha|x, I) + o(h)
\]
by Property S, which also gives,

\[
\max_{p=1,\ldots,s+1} \left( \max_{x \in X} \left| P(x)^t b_p^*(\alpha \mid I) \right| \right) = \max_{p=1,\ldots,s+1} \max_{(\alpha,x) \in [0,1] \times X} h^{p-1} \left| B(p)(\alpha \mid x, I) + o(1) \right|
\]

\[
= \max_{(\alpha,x) \in [0,1] \times X} B^{(1)}(\alpha \mid x, I) + o(1) \leq \bar{f}
\]

provided \( \bar{f} \) is large enough and \( h \) small enough, so that \( b^*(\alpha \mid I) \) is in \( B_{T,\alpha,h} \) since \( B^{(1)}(\cdot, \cdot, \cdot) \) is bounded away from 0 and infinity by Proposition 3. Suppose now that \( \|b - b^*(\alpha \mid I)\| \leq C h / K_1^{1/2} = C h^{d_M/2+1} \). Then

\[
\left| \frac{\partial}{\partial t} [P(x,t)^t b] \right| \geq \left| \frac{\partial}{\partial t} [P(x,t)^t b^*(\alpha \mid I)] \right| - \|b - b^*(\alpha \mid I)\| \|P(x)\|
\]

\[
\geq \left| \frac{\partial}{\partial t} [P(x,t)^t b^*(\alpha \mid I)] \right| - O(h),
\]

\[
|P(x)^t b_p| \leq |P(x)^t b_p^*(\alpha \mid I)| + \|b - b^*(\alpha \mid I)\| \|P(x)\|
\]

\[
\leq |P(x)^t b_p^*(\alpha \mid I)| - C h, \quad p = 1, \ldots, s + 1,
\]

and \( B \left( b^*(\alpha \mid I), C h^{d_M/2+1} \right) \subset B_{T,\alpha,h} \) when \( h \) is small enough provided \( C \) is small enough.

Hence (i) holds. (ii) follows from the Implicit Function Theorem and the definition of \( B_{T,\alpha,h} \).

The first equality of (iii) is (C.1). For the second, note that \( \alpha + h t \geq h > 0 \) when \( \alpha \geq 3h \)
for all $t$ in $\mathcal{I}_{\alpha,h}$. It holds

$$B(\alpha + ht|x, I) - P(x, t)^* b^*(\alpha|I)$$

$$= B(\alpha + ht|x, I) - \sum_{k=1}^{K_L} P_k(x) \beta_k(\alpha + ht|I)$$

$$+ \sum_{k=1}^{K_L} P_k(x) \left( \beta_k(\alpha + ht|I) - \sum_{p=0}^{s+1} \frac{(ht)^p}{p!} \beta_k^{(p)}(\alpha|I) \right)$$

with

$$\beta_k(\alpha + ht|I) - \sum_{p=0}^{s+1} \frac{(ht)^p}{p!} \beta_k^{(p)}(\alpha|I) = \frac{(ht)^{s+2}}{(s+1)!} \int_0^1 \beta_k^{(s+2)}(\alpha + uht|I)(1-u)^{s+1} du$$

recalling, as established in the proof of Proposition 3-(i) for $\alpha > 0$,

$$\beta_k^{(s+2)}(\alpha|I) = \frac{1}{\alpha} \left( (I - 1) \gamma_k^{(s+1)}(\alpha|I) - (I + s) \beta_k^{(s+1)}(\alpha|I) \right),$$

$$B^{(s+2)}(\alpha|x, I) = \frac{1}{\alpha} \left( (I - 1) V_k^{(s+1)}(\alpha|I) - (I + s) B^{(s+1)}(\alpha|x, I) \right). \quad (C.2)$$
Hence

\[
B(\alpha + ht|x, I) - P(x, t)' b^* (\alpha|I) - \frac{(ht)^{s+2}}{(s + 2)!} B^{(s+2)} (\alpha|I)
\]

\[
= B(\alpha + ht|x, I) - \sum_{k=1}^{K_L} P_k(x) \beta_k (\alpha + ht|I)
\]

\[
+ \frac{(ht)^{s+2}}{(s + 1)!} \int_0^1 \left\{ \sum_{k=1}^{K_L} P_k(x) \beta_k^{(s+2)} (\alpha + uht|I) - B^{(s+2)} (\alpha + uht|x, I) \right\} (1 - u)^{s+1} \, du
\]

\[
+ \frac{(ht)^{s+2}}{(s + 1)!} \int_0^1 \left\{ B^{(s+2)} (\alpha + uht|x, I) - B^{(s+2)} (\alpha|x, I) \right\} (1 - u)^{s+1} \, du,
\]

with, using the expressions \( \beta_k^{(s+2)} (\cdot|\cdot) \) and \( B^{(s+2)} (\cdot|\cdot) \) of the proof of Proposition 3

\[
\max_{(\alpha,x) \in [0,3h] \times X} \max_{t \in I_{\alpha,h}} \left| \alpha \left( B(\alpha + ht|x, I) - \sum_{k=1}^{K_L} P_k(x) \beta_k (\alpha + ht|I) \right) \right| = o \left( K_L^{-\frac{s+1}{s+2}} \right) = o \left( h^{s+2} \right),
\]

\[
\max_{(\alpha,x) \in [3h,1] \times X} \max_{t \in I_{\alpha,h}} \left| \alpha \int_0^1 \left\{ \sum_{k=1}^{K_L} P_k(x) \beta_k^{(s+2)} (\alpha + uht|I) - B^{(s+2)} (\alpha + uht|x, I) \right\} (1 - u)^{s+1} \, du \right|
\]

\[
\leq C \max_{(\alpha,x) \in [2h,1] \times X} \max_{t \in I_{\alpha,h}} \left\{ \frac{\alpha}{\alpha - h} \sum_{k=1}^{K_L} P_k(x) \beta_k^{(s+1)} (\alpha|I) - B(\alpha|x, I) \right\}
\]

\[
+ C \max_{(\alpha,x) \in [2h,1] \times X} \max_{t \in I_{\alpha,h}} \left\{ \frac{\alpha}{\alpha - h} \sum_{k=1}^{K_L} P_k(x) \gamma_k^{(s+1)} (\alpha|I) - V(\alpha|x, I) \right\} = o(1),
\]

\[
\max_{(\alpha,x) \in [3h,1] \times X} \max_{t \in I_{\alpha,h}} \left| \alpha \int_0^1 \left\{ B^{(s+2)} (\alpha + uht|x, I) - B^{(s+2)} (\alpha|x, I) \right\} (1 - u)^{s+1} \, du \right| = o(1).
\]

Substituting gives

\[
\max_{(\alpha,x) \in [3h,1] \times X} \max_{t \in I_{\alpha,h}} \left| \alpha \left( B(\alpha + ht|x, I) - P(x, t)' b^* (\alpha|I) - \frac{(ht)^{s+2}}{(s + 2)!} B^{(s+2)} (\alpha|x, I) \right) \right| = o \left( h^{s+2} \right)
\]
which implies the second statement in (iii) since by Proposition 3-(i) and (B.4)

\[
\max_{(\alpha, x) \in [0, 3h] \times X} \max_{t \in \mathcal{I}_{\alpha, h}} \left| \alpha \left( B (\alpha + ht|x, I) - P (x, t)' b^* (\alpha | I) \right) \right| = o \left( h^{s+2} \right),
\]

\[
\max_{(\alpha, x) \in [0, 3h] \times X} \max_{t \in \mathcal{I}_{\alpha, h}} \left| \alpha \frac{(ht)^{s+2}}{(s+2)!} B^{(s+2)} (\alpha | x, I) \right| = o \left( h^{s+2} \right).
\]

The third result in (iii) follows from Proposition 3-(iii). The fourth equality of (iii) follows from

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

\[
= \max_{(\alpha, x) \in [0, 1] \times X} \max_{u \in \Psi [\mathcal{I}_{\alpha, h}|x, b^* (\alpha | I)]} \left| \Psi \left[ \Delta (u|x, b^* (\alpha | I)) \right] \right| - B \left[ \alpha + h \Delta (u|x, b^* (\alpha | I)) \right] |x, I| \]

\[
= \max_{(\alpha, x) \in [0, 1] \times X} \max_{u \in \Psi [\mathcal{I}_{\alpha, h}|x, b^* (\alpha | I)]} \left| u - B \left[ \alpha + h \Delta (u|x, b^* (\alpha | I)) \right] \right| |x, I| \]

\[
= \max_{(\alpha, x) \in [0, 1] \times X} \max_{u \in \Psi [\mathcal{I}_{\alpha, h}|x, b^* (\alpha | I)]} \left| B \left[ \alpha + h \frac{G (u|x, I) - \alpha}{h} \right] \right| - B \left[ \alpha + h \Delta (u|x, b^* (\alpha | I)) \right] |x, I| \]

\[
\geq Ch \max_{(\alpha, x) \in [0, 1] \times X} \max_{u \in \Psi [\mathcal{I}_{\alpha, h}|x, b^* (\alpha | I)]} \left| \frac{G (u|x, I) - \alpha}{h} - \frac{\Phi (u|x, b^* (\alpha | I)) - \alpha}{h} \right|
\]

by Proposition 3-(i).

Consider now (iv). The first bound follows from the Cauchy-Schwarz inequality. This
bound implies for all $u$ in $\Psi [I_{\alpha,h}|x,b_1] \cap \Psi [I_{\alpha,h}|x,b_1]$

$$|\Psi [\Delta (u|x,b_1) |x,b_0] - \Psi [\Delta (u|x,b_0) |x,b_0]|$$

$$= |\Psi [\Delta (u|x,b_1) |x,b_0] - u|$$

$$= |\Psi [\Delta (u|x,b_1) |x,b_0] - \Psi [\Delta (u|x,b_1) |x,b_1]| \leq C h^{-d/2} \|b_1 - b_0\|.$$  

By definition of $B_{\alpha,h}$

$$|\Psi [\Delta (u|x,b_1) |x,b_0] - \Psi [\Delta (u|x,b_0) |x,b_0]|$$

$$\geq C h |\Delta (u|x,b_1) - \Delta (u|x,b_0)| = C |\Phi (u|x,b_1) - \Phi (u|x,b_0)|$$

and substituting shows that the second bound of (iv) holds. For the third bound in (iv), it holds uniformly in $\alpha, x, u, b_1$ and $b_0$

$$\left| \frac{\partial \Psi}{\partial t} [\Delta (u|x,b_1) |x,b_1] - \frac{\partial \Psi}{\partial t} [\Delta (u|x,b_0) |x,b_0] \right|$$

$$\leq \left| \frac{\partial \Psi}{\partial t} [\Delta (u|x,b_1) |x,b_1] - \frac{\partial \Psi}{\partial t} [\Delta (u|x,b_0) |x,b_1] \right|$$

$$+ \left| \frac{\partial \Psi}{\partial t} [\Delta (u|x,b_0) |x,b_1] - \frac{\partial \Psi}{\partial t} [\Delta (u|x,b_0) |x,b_0] \right|$$

$$\leq \max_{t \in I_{\alpha,h}} \frac{\partial^2 \Psi (t|x,b_1)}{\partial t^2} \frac{|\Phi (u|x,b_1) - \Phi (u|x,b_0)|}{h}$$

$$+ \max_{t \in I_{\alpha,h}} \left| \frac{\partial P (x,t)}{\partial t} (b_1 - b_0) \right|.$$
But, by definition of $\mathcal{B}_\alpha,h$

$$\max_{t \in \mathcal{T}_{\alpha,h}} \left| \frac{\partial^2 \Psi(t|x, b_1)}{\partial t^2} \right| \leq C h \max_{p=2, \ldots, s+1} \left| \frac{P(x) b_{p}}{h} \right| = O(h)$$

so that substituting and the bound for $\Phi(u|x, b_1) - \Phi(u|x, b_0)$ gives, uniformly in $\alpha, x, u, b_1$ and $b_0$

$$\left| \frac{\partial \Psi}{\partial t} \left[ \Delta(u|x, b_1)|x, b_1 \right] - \frac{\partial \Psi}{\partial t} \left[ \Delta(u|x, b_0)|x, b_0 \right] \right| \leq C h^{-d+1/2} \|b_1 - b_0\|,$$

which is the fourth inequality. The expression in (ii) of $\Phi(\cdot)$ and the definition of $\mathcal{B}_\alpha,h$ yield the third inequality.

**Proof of Lemma B.2.** It holds

$$\bar{R}^{(2)}(b; \alpha, I) = \mathbb{E} \left[ I \mathbb{I} [B_{i | \ell} \in \Psi(\mathcal{T}_{\alpha,h}|x_\ell, b), I_\ell = I] \right]$$

$$= \int \left[ \int \frac{P(x_\ell, \Delta(B_{i | \ell}|x_\ell, b)) P(x_\ell, \Delta(B_{i | \ell}|x_\ell, b))'}{\Psi(\Delta(B_{i | \ell}|x_\ell, b)|x_\ell, b)} K(\Delta(B_{i | \ell}|x_\ell, b)) P(x, \Delta(y|x, b)) P(x, \Delta(y|x, b))' \Psi(\Delta(y|x, b)|x_\ell, b) K(\Delta(y|x, b)) g(y, x, I)\,dy \right] \,dx.$$

Recall $\Delta[\Psi[t|x, b]|x, b] = t$ for all $t$ in $\mathcal{T}_{\alpha,h}$ and let

$$\mathcal{I}_{\alpha,h}(x, I; b) = \bar{I}_{\alpha,h} \land \Delta[B(1|x, I)|x, b], \quad \mathcal{L}_{\alpha,h}(x, I; b) = \bar{L}_{\alpha,h} \lor \Delta[B(0|x, I)|x, b].$$
The change of variable \( y = \Psi(t|x, b) \) yields that

\[
\bar{R}^{(2)}(b; \alpha, I) = \int \left[ \int_{\mathcal{L}_{\alpha,h}(x,I,b)} \mathcal{L}_{\alpha,h}(x,I,b) P(x,t) P(x,t)' K(t) g(\Psi(t|x,b), x, I) dt \right] dx.
\]

The Dominated Convergence Theorem and Proposition 3-(i)\(^{1}\), \( s \geq 1 \), yield that \( \bar{R}^{(2)}(\cdot; \alpha, I) \) is continuously differentiable over \( \mathcal{B}_{\mathcal{I}_{\alpha,h}} \) with, by the Liebniz integral rule,

\[
\begin{align*}
\bar{R}^{(3)}(b; \alpha, I)[d] &= \bar{R}^{(3)}_0(b; \alpha, I)[d] + \bar{R}^{(3)}_1(b; \alpha, I)[d] - \bar{R}^{(3)}_2(b; \alpha, I)[d], \\
\bar{R}^{(3)}_0(b; \alpha, I)[d] &= \int_X \left[ \int_{\mathcal{L}_{\alpha,h}(x,I,b)} P(x,t) P(x,t)' K(t) g^{(1)}(\Psi(t|x,b), x, I) [d' P(x,t)] dt \right] dx, \\
\bar{R}^{(3)}_1(b; \alpha, I)[d] &= \int_X P(x, \mathcal{T}_{\alpha,h}(x,I;b)) P(x, \mathcal{T}_{\alpha,h}(x,I;b))' K(\mathcal{T}_{\alpha,h}(x,I;b)) \\
&\quad \times g(\Psi(\mathcal{T}_{\alpha,h}(x,I;b), x, I)) \left[ d' \frac{\partial \mathcal{T}_{\alpha,h}(x,I;b)}{\partial b'} \right] dx, \\
\bar{R}^{(3)}_2(b; \alpha, I)[d] &= \int_X P(x, \mathcal{L}_{\alpha,h}(x,I;b)) P(x, \mathcal{L}_{\alpha,h}(x,I;b))' K(\mathcal{L}_{\alpha,h}(x,I;b)) \\
&\quad \times g(\Psi(\mathcal{L}_{\alpha,h}(x,I;b), x, I)) \left[ d' \frac{\partial \mathcal{L}_{\alpha,h}(x,I;b)}{\partial b'} \right] dx.
\end{align*}
\]

Proposition 3-(i) and Assumption R-(i) imply

\[
\left\| \bar{R}^{(3)}_0(b; \alpha, I)[d] \right\| \leq C \max_{x \in \mathcal{X}} \| P(x) \| \| d \| \leq Ch^{-d_M/2} \| d \|, 
\]

\(^{1}\)which implies that \( g(\cdot, I) \) is bounded away from 0 and infinity.
The operators $\mathcal{R}_i^{(3)}(b; \alpha, I)[d]$, $i = 1, 2$, can be studied in a similar way so that only $i = 1$ is considered. Observe

$$\frac{\partial T_{\alpha,h}(x, I; b)}{\partial b'} = \begin{cases} 
0 & \text{if } T_{\alpha,h} \leq \Delta [B(1|x, I)|x, b] \\
\frac{\partial \Delta [B(1|x, I)|x, b]}{\partial b'} = -\frac{P(x, \Delta (B(1|x, I)|x, b))}{\Psi^{(1)}(\Delta (B(1|x, I)|x, b))} & \text{if } T_{\alpha,h} > \Delta [B(1|x, I)|x, b]
\end{cases}$$

But, for $h$ small enough,

$$\Delta [B(1|x, I)|x, b] = \frac{\Phi [B(1|x, I)|x, b] - \alpha}{h} = \min \left\{ \alpha + hT_{\alpha,h}, \Phi [B(1|x, I)|x, b] \right\} - \alpha$$

$$\geq \min \left\{ \alpha + hT_{\alpha,h}, \Phi [B(1|x, I)|x, b^*(\alpha|I)] - Ch^{-d_{\mathcal{M}}/2} \|b - b^*(\alpha|I)\| \right\} - \alpha$$

$$\geq \min \left\{ \alpha + hT_{\alpha,h}, G [B(1|x, I)|x, I] - Ch^{s+1} - Ch \right\} - \alpha$$

$$\geq \min \left\{ \alpha + h \min \left( \frac{1-\alpha}{h}, 1 - Ch \right) , 1 - Ch \right\} - \alpha$$

uniformly in $\alpha$, $x$ and $b$ in $B(b^*(\alpha|I), Ch^{d_{\mathcal{M}}/2+1})$ by Lemma B.1. Hence, if $\alpha \leq 1 - C'h$ with $C' \geq 1$ large enough

$$\Delta [B(1|x, I)|x, b] \geq \min \left\{ \alpha + h, 1 - Ch \right\} - \alpha \geq 1 \geq T_{\alpha,h}$$

160
so that \( \frac{\partial \tau_{\alpha,b}(x,I;b)}{\partial b^r} = 0 \). Hence since \( B(b^*(\alpha|I), CH^{dM/2+1}) \subset BT_{\alpha,h} \) and by definition of \( BT_{\alpha,h} \)

\[
\| \overline{R}^{(3)}(b;\alpha,I) [d] \| \leq C \Pi [\alpha \geq 1 - C'h] \\
\times \left\| \int_{\mathcal{X}} P(x, T_{\alpha,h}(x,I;b)) P(x, T_{\alpha,h}(x,I;b)) \frac{d'P(x, \Delta(B(1|x,I)|x,b))}{\Psi(\Delta(B(1|x,I)|x,b)|x,b)} dx \right\| \\
\leq C h^{-1} \Pi [\alpha \geq 1 - C'h] \max_{x \in \mathcal{X}} \| P(x) \| \| d \| \leq C h^{-1} h^{-dM/2} \| d \| \Pi [\alpha \geq 1 - C'h] \\
\leq C \frac{h^{-dM/2}}{\alpha (1 - \alpha) + h} \| d \| .
\]

Substituting in the expression of \( \overline{R}^{(3)}(b;\alpha,I) [d] \) then gives uniformly in \( d \)

\[
\max_{\alpha \in [0,1]} \max_{b \in B(b^*(\alpha|I), CH^{dM/2+1})} (\alpha (1 - \alpha) + h) \left\| \overline{R}^{(3)}(b;\alpha,I) [d] \right\| \leq C h^{-dM/2} \| d \| .
\]

The Taylor inequality shows that (i) holds.

For (ii), the expression of \( \overline{R}^{(2)}(b;\alpha,I) \), Assumptions A and R-(i), Proposition 3-(i), which imply that the eigenvalues of \( \int P(x) P'(x) g[B(\alpha|x,I), x,I] dx \) stay bounded away 0 and infinity, Lemma B.1-(iii) and Proposition 3-(i) give that, uniformly in \( \alpha \) and \( x \)

\[
T_{\alpha,h}[x,I; b^*(\alpha|I)] = T_{\alpha,h} \wedge \frac{\Phi[B(1|x,I)|x,b^*(\alpha|I)] - \alpha}{h} \\
= T_{\alpha,h} \wedge \frac{1 + o(h^{s+1}) - \alpha}{h} = T_{\alpha,h} + o(h^s), \\
I_{\alpha,h}[x,I; b^*(\alpha|I)] = I_{\alpha,h} + o(h^s),
\]

\[161\]
\[
\mathcal{R}^{(2)}(b^*(\alpha|I);\alpha,I) = \int \left[ \int_{J_{\alpha,h}^{[x,I;b^*(\alpha|I)]}} \pi(t) \pi(t)' K(t) g(\Psi(t|x,b^*(\alpha|I))|x,I) \, dt \right] \\
\hspace{1cm} \otimes P(x) P(x)' f(x,I) \otimes dx \\
= \int \left[ \int_{J_{\alpha,h}^{o(h^*)}} \pi(t) \pi(t)' K(t) g\left[B(\alpha + ht|x,I) + o(h^{s+1})|x,I\right] dt \right] \\
\hspace{1cm} \otimes P(x) P(x)' f(x,I) \otimes dx \\
= \int \left[ \int_{J_{\alpha,h}^{o(h^*)}} \pi(t) \pi(t)' K(t) \left(\frac{1}{B^{(1)}(\alpha|x,I)} + o(h)\right) dt \right] \\
\hspace{1cm} \otimes P(x) P(x)' f(x,I) \otimes dx \\
= \int \Omega_h(\alpha) \otimes \frac{P(x) P(x)'}{B^{(1)}(\alpha|x,I)} f(x,I) dx + \\
\hspace{1cm} -\int \Omega_{1h}(\alpha) \otimes \frac{P(x) P(x)' B^{(2)}(\alpha|x,I)}{(B^{(1)}(\alpha|x,I))^2} f(x,I) dx + o(h)
\]

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 \( \int \mathcal{R}(x) P(x) P(x)' \, dx \) 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 $\alpha$ and $L$. It also gives
\[
\mathbb{E} \left[ \frac{\mathbb{E}^{1/2} \mathbb{E}^{1/2}}{I \kappa_1 (x_\ell) \mathbb{E}^{1/2} \mathbb{E}^{1/2}} \int_{L_{\alpha,h}}^{T_{\alpha,h}} P_{k_1} (x_\ell) \pi_{p_1} (t) P_{k_2} (x_\ell) \pi_{p_2} (t) K (t) \, dt \right] \leq \varrho < 1
\]
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 $\varrho$ 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 $\varrho \Id \otimes (T' + T)$, 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
\[
C \Id \otimes \left( \sum_{n=\infty}^{\infty} \varrho^n \left( T'^n + T^n \right) \right).
\]
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 $\alpha$ 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 $Z_\ell$ and, for a known real function $\xi (z, \theta)$ separable with respect to $\theta \in \Theta$, $Z_\ell (\theta) = \xi (Z_\ell, \theta)$ where $\theta$ is a parameter. Let $\underline{\xi} (\cdot) \leq \xi (\cdot) \leq \bar{\xi} (\cdot)$ be two functions. A *bracket* $[\xi, \bar{\xi}]$ is the set of all functions $\xi (\cdot)$ such that $\underline{\xi} (z) \leq \xi (z) \leq \bar{\xi} (z)$ for all $z$. The next proposition follows from Massart (2007, Theorem 6.8 and Corollary 6.9).

**Proposition C.1** Assume that $\sup_{\theta \in \Theta} |Z_\ell (\theta)| \leq M_\infty$, $\sup_{\theta \in \Theta} \text{Var} (Z_\ell (\theta)) \leq M_2^2$ for all $\ell$ and that for any $\epsilon > 0$ there exists brackets $[\underline{\xi}_j, \bar{\xi}_j] \subset [-b, b]$, $j = 1, \ldots, \exp (H (\epsilon))$, such that

$$
\mathbb{E} \left[ \left( \bar{\xi}_j (Z_i) - \underline{\xi}_j (Z_i) \right)^2 \right] \leq \frac{\epsilon^2}{2} \text{ and } \{ \xi (z, \theta), \theta \in \Theta \} \subset \bigcup_{j=1}^{\exp (H (\epsilon))} [\underline{\xi}_j, \bar{\xi}_j].
$$

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 $t \in [0, 10L^{1/2}M_2/M_\infty]$, 

$$
\mathbb{P} \left( \sup_{\theta \in \Theta} \left| \sum_{i=1}^{n} \{ Z_\ell (\theta) - \mathbb{E} [Z_\ell (\theta)] \} \right| \geq L^{1/2} \{ \mathcal{H}_L + t \} \right) \leq 2 \exp \left( -\frac{t^2}{25} \right).
$$

**Proof of Lemma B.3.** Note that $\tilde{R}^{(2)} (b; \alpha, I) - \tilde{R}^{(2)} (b; \alpha, I)$ is a $c (s + 2)$-band matrix, so that the order of its matrix norm is the same than the order of its largest entry. The generic
entry of $\hat{R}(2; b; \alpha, I) - R(2; b; \alpha, I)$ can be written as

$$\hat{r}(b; \alpha, I) = \frac{1}{Lh^{(d_{M}+1)/2}} \sum_{\ell=1}^{L} \xi_{\ell}(b; \alpha)$$

where the $\xi_{\ell}(b; \alpha)$ are centered iid with

$$\xi_{\ell}(b; \alpha) = \sum_{i=1}^{I_{\ell}} \{ \mathbb{I} [B_{i\ell} \in \Psi (I_{\alpha,h}|x_{\ell}, b), I_{\ell} = I] \xi_{i\ell}(b) \}
- \mathbb{E} \{ \mathbb{I} [B_{i\ell} \in \Psi (I_{\alpha,h}|x_{\ell}, b), I_{\ell} = I] \xi_{i\ell}(b) \}
\xi_{i\ell}(b) = \frac{h^{d_{M}/2}}{h^{1/2}} \frac{P_{k_{1}}(x_{\ell}) P_{k_{2}}(x_{\ell})}{\Psi^{(1)}(\Delta(B_{i\ell}|x_{\ell}, b)|x_{\ell}, b)/h} K_{p}(\Delta(B_{i\ell}|x_{\ell}, b)),
K_{p}(\Delta(B_{i\ell}|x_{\ell}, b)) = \frac{\Delta^{p_{1}+p_{2}}(B_{i\ell}|x_{\ell}, b)}{p_{1}!p_{2}!} K(\Delta(B_{i\ell}|x_{\ell}, b)).$$

The proof of the Lemma follows from Proposition C.1. Observe

$$|\xi_{\ell}(b; \alpha)| \leq C \frac{h^{d_{M}/2} \max_{x \in X} \|P(x)\|^2}{h^{1/2}} \leq M_{\infty} \text{ with } M_{\infty} \asymp h^{-(d_{M}+1)/2}.$$
for all $\alpha$ in $[0, 1]$ and all admissible $b$. For the variance, Lemma B.1-(iii,iv) gives

\[
|\Delta(B_{i\ell}|x_{\ell}, b)| = \frac{\left| \Phi(B_{i\ell}|x_{\ell}, b) - \alpha \right|}{h} \leq \left| \frac{G(B_{i\ell}|x_{\ell}, I_{\ell}) - \alpha}{h} \right| + \left| \frac{\Phi(B_{i\ell}|x_{\ell}, b^*(\alpha|I_{\ell})) - G(B_{i\ell}|x_{\ell}, b)}{h} \right| \\
+ \left| \frac{\Phi(B_{i\ell}|x_{\ell}, b) - \Phi(B_{i\ell}|x_{\ell}, b^*(\alpha|I_{\ell}))}{h} \right| \\
\leq \left| \frac{G(B_{i\ell}|x_{\ell}, I_{\ell}) - \alpha}{h} \right| + o(h^s) + O\left(\frac{h^{-dM/2} \times h^{dM/2+1}}{h}\right) \\
= \left| \frac{G(B_{i\ell}|x_{\ell}, I_{\ell}) - \alpha}{h} \right| + O(1)
\]

uniformly. It follows that, $U_{i\ell} = G(B_{i\ell}|x_{\ell}, I_{\ell})$ being a uniform random variable independent of $(x_{\ell}, I_{\ell})$

\[
\text{Var}\left(\xi_{\ell}(b; \alpha)\right) \leq CI_2^2 h^{dM} \max_{x \in \mathcal{X}} \|P(x)\|^2 \int_{\mathcal{X}} |P_k(x) P_{k_2}(x)| dx \int I_{[-\infty, C]} \left(\frac{u - \alpha}{h} \right) du \frac{du}{h} \\
\leq CI_2^2 h^{dM} \max_{x \in \mathcal{X}} \|P(x)\|^2 \left(\int_{\mathcal{X}} P_{k_1}^2(x) dx\right)^{1/2} \left(\int_{\mathcal{X}} P_{k_2}^2(x) dx\right)^{1/2} \\
\leq M_2^2 \text{ with } M_2 < \infty
\]

under Assumption R, uniformly in $b$ and $\alpha$.

Consider now the brackets covering. The key observation is that $\xi_{\ell}(b; \alpha)$ only depends on a finite dimension subvector of $b$, $b(k_1, k_2)$ 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 $b(k_1, k_2)$
is less than \(c(s+2)\) under Assumption R-(ii). Consequently the class to be bracketed is

\[
\mathcal{F} = \{ \xi \left( b^{(k_1,k_2)}; \alpha \right) ; \alpha \in [0,1], b^{(k_1,k_2)} \in \mathcal{B} \left( b^{(k_1,k_2)}* (\alpha|I) , Ch^{dM/2+1} \right) \}.
\]

Lemma B.1-(iii), \(1/(Lh^{dM+1}) = o(1)\), 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 \(H_L\) of Proposition C.1,

\[
H_L = O(\log L)^{1/2} + O \left( \frac{\log L}{Lh^{dM+1}} \right)^{1/2} = O(\log L)^{1/2}
\]

since \(1/(Lh^{dM+1})\) is bounded. Hence, by Proposition C.1 for \(t \leq 10L^{1/2}M_2/M_\infty\) diverges

\[
\mathbb{P} \left( \left( Lh^{dM+1} \right)^{1/2} \sup_{\alpha \in [0,1]} \sup_{b \in \mathcal{B} \left( b^{(\alpha,I)}, Ch^{dM/2+1} \right)} \left| \tilde{r} \left( b; \alpha, I \right) \right| \geq C \log^{1/2} L + t \right)
\leq 2 \exp \left( -\frac{t^2}{25} \right)
\]

uniformly over all the non zero entries \(\tilde{r} \left( b; \alpha, I \right)\) of the band matrix \(\hat{R}^{(2)} \left( b; \alpha, I \right) - \bar{R}^{(2)} \left( b; \alpha, I \right)\).
This gives, by the Bonferroni inequality

\[
P \left( \sup_{a \in [0,1]} \sup_{b \in \mathcal{B}(b^*(\alpha I), C h^{d,M/2+1})} \left\| \hat{R}^{(2)} (b; \alpha, I) - \bar{R}^{(2)} (b; \alpha, I) \right\| \geq \frac{C \log^{1/2} L + t}{(L h^{d,M+1})^{1/2}} \right) 
\leq C K_L \exp \left( - \frac{t^2}{25} \right)
\]

which implies the result of the lemma since \( t \leq 10 L^{1/2} M_2 / M_\infty = O \left( L h^{d,M+1} \right)^{1/2} \) can be set to \( t = \tau \log^{1/2} L \) for an arbitrary large \( \tau \) as \( \log L / (L h^{d,M+1}) = o(1) \). \(\square\)

**Proof of Lemma B.4.** The proof of Lemma B.4 is similar to the one of Lemma B.3. The generic entry of \( \hat{R}^{(1)} (b; \alpha, I) - \bar{R}^{(1)} (b; \alpha, I) \) writes

\[
\hat{\gamma} (b; \alpha, I) = \frac{1}{L} \sum_{\ell=1}^{L} \xi_\ell (b; \alpha)
\]

where the \( \xi_\ell (b; \alpha) \) are centered iid with, for \( K_p (t) = t^p K (t) / p! \),

\[
\xi_\ell (b; \alpha) = \sum_{i=1}^{I_\ell} \left( \mathbb{I} (I_\ell = I) \xi_{i\ell} (b; \alpha) - \mathbb{E} \left[ \mathbb{I} (I_\ell = I) \xi_{i\ell} (b; \alpha) \right] \right), \quad \xi_{i\ell} (b; \alpha) = P_k (x_\ell) \left\{ \int_{t_{n,h}}^{t_{n,h}+h} \{ B_{i\ell} \leq \Psi (t|x_\ell, b) \} K_p (t) dt \right\}. 
\]

This gives

\[
\left| \frac{\xi_\ell (b; \alpha)}{(h + \alpha (1 - \alpha))^{1/2}} \right| \leq C h^{-1/2} \max_{x \in \mathcal{X}} \| P (x) \| \leq M_\infty \text{ with } M_\infty \asymp h^{-(d,M+1)/2}. 
\]
For the computation of the variance, Lemma B.1-(iii,iv) and Proposition 3-(i) give uniformly in $\alpha$, $t$ in $\mathcal{I}_{\alpha,h}$ the admissible $b$ and $x_\ell$, and for the uniform $U_{i\ell} = G(B_{i\ell}|x_\ell, I_\ell)$,

$$
\mathbb{I} [B_{i\ell} \leq \Psi (t|x_\ell, b)] = \mathbb{I} [B_{i\ell} \leq \Psi (t|x_\ell, b^*(\alpha|I)) + O(h)]
$$

$$
= \mathbb{I} [B (U_{i\ell}|x_\ell, I_\ell) \leq B (\alpha + ht|x_\ell, I_\ell) + O(h)]
$$

$$
= \mathbb{I} [U_{i\ell} \leq G (B (\alpha + ht|x_\ell, I_\ell) + O(h) |x_\ell, I_\ell)]
$$

$$
= \mathbb{I} [U_{i\ell} \leq \alpha + ht + O(h)].
$$

It then follows, since $U_{i\ell}$ is independent of $(x_\ell, I_\ell)$

\[
\mathbb{E} \left[ \xi_{i\ell}^2 (b; \alpha) | I_\ell \right] \\
\leq \mathbb{E} \left[ P_\ell^2 (x_\ell) \int_{L_{\alpha,h}}^{L_{\alpha,h}} \int_{L_{\alpha,h}}^{L_{\alpha,h}} \mathbb{I} [U_{i\ell} \leq \alpha + h (t_1 \wedge t_2) + O(h)] K_\ell (t_1) K_\ell (t_2) dt_1 dt_2 | I_\ell \right] \\
- 2 \mathbb{E} \left[ P_\ell^2 (x_\ell) \int_{L_{\alpha,h}}^{L_{\alpha,h}} \int_{L_{\alpha,h}}^{L_{\alpha,h}} \mathbb{I} [U_{i\ell} \leq \alpha + ht_1 + O(h)] (\alpha + ht_2) K_\ell (t_1) K_\ell (t_2) dt_1 dt_2 | I_\ell \right] \\
+ \mathbb{E} \left[ P_\ell^2 (x_\ell) | I_\ell \right] \int_{L_{\alpha,h}}^{L_{\alpha,h}} \int_{L_{\alpha,h}}^{L_{\alpha,h}} (\alpha + ht_1) (\alpha + ht_2) K_\ell (t_1) K_\ell (t_2) dt_1 dt_2 \\
= \mathbb{E} \left[ P_\ell^2 (x_\ell) | I_\ell \right] \int_{L_{\alpha,h}}^{L_{\alpha,h}} \int_{L_{\alpha,h}}^{L_{\alpha,h}} \{ \alpha + O(h) - \alpha^2 \} K_\ell (t_1) K_\ell (t_2) dt_1 dt_2 \leq C (h + \alpha (1 - \alpha))
\]

uniformly in $\alpha$ and $b$. Hence, uniformly in $\alpha$ and $b$

$$
\text{Var} \left( \frac{\xi_{i\ell} (b; \alpha)}{(h + \alpha (1 - \alpha))^{1/2}} \right) \leq M_2^2 \text{ with } M_2 < \infty.
$$

169
The bracketing part of the proof is similar to the one of Lemma B.3 and gives

\[ H_L = O(\log L)^{1/2} + O\left(\frac{\log L}{L h^d M + 1}\right)^{1/2} = O(\log L)^{1/2}. \]

Arguing with Proposition C.1 then shows that the order of the largest entry in \( \hat{R}^{(1)}(b; \alpha, I) - R^{(1)}(b; \alpha, I) \) is \( O(\log L/L)^{1/2} \), which gives uniformly

\[ \left\| \hat{R}^{(1)}(b; \alpha, I) - R^{(1)}(b; \alpha, I) \right\| = K_L^{1/2} O(\log L/L)^{1/2} = O(\log L/L h^d M)^{1/2} \]

and the Lemma is proved.

\[ \square \]

**Proof of Lemma B.5.** For (i), define

\[
P = \mathbb{E} \left[ \mathbb{I}(I_\ell = I) P(x_\ell) P(x_\ell)' \right],
\]

\[
P_0 = \mathbb{E} \left[ \frac{\mathbb{I}(I_\ell = I) P(x_\ell) P(x_\ell)'}{B^{(1)}(\alpha|x_\ell, I_\ell)} \right],
\]

\[
P_1 = \mathbb{E} \left[ \frac{\mathbb{I}(I_\ell = I) B^{(2)}(\alpha|x_\ell, I_\ell) P(x_\ell) P(x_\ell)'}{(B^{(1)}(\alpha|x_\ell, I_\ell))^2} \right],
\]

and abbreviate \( \Omega_h(\alpha), \Omega_{ih}(\alpha) \) in \( \Omega, \Omega_1 \). It holds

\[
\text{Var}(\bar{e}(\alpha|I)) = \left[ \hat{R}^{(2)}(\bar{e}(\alpha|I); \alpha, I) \right]^{-1} \text{Var} \left[ \hat{R}^{(1)}(\bar{e}(\alpha|I); \alpha, I) \right] \left[ \hat{R}^{(2)}(\bar{e}(\alpha|I); \alpha, I) \right]^{-1}
\]
with by Lemma B.2

\[
\left( \tilde{R}^{(2)} (b(\alpha|I); \alpha, I) \right)^{-1} = \left[ \Omega \otimes P_0 - h \Omega_1 \otimes P_1 + o(h) \right]^{-1} \\
= \Omega^{-1} \otimes P_0^{-1} \left[ \text{Id} - h (\Omega^{-1} \Omega_1) \otimes (P_0^{-1} P_1) + o(h) \right]^{-1} \\
= \Omega^{-1} \otimes P_0^{-1} + h (\Omega^{-2} \Omega_1) \otimes (P_0^{-2} P_1) + o(h)
\]

uniformly in \( \alpha \) where the remainder term \( o(h) \) is with respect to the matrix norm. For \( \text{Var} \left[ \tilde{R}^{(1)} (b(\alpha|I); \alpha, I) \right] \), define

\[
\omega_0 = \int_{L_{\alpha, h}}^{T_{\alpha, h}} \pi(t)K(t)dt, \quad \omega_1 = \int_{L_{\alpha, h}}^{T_{\alpha, h}} \pi(t)K'(t)dt, \\
\Pi_m = \int_{L_{\alpha, h}}^{T_{\alpha, h}} \int_{L_{\alpha, h}}^{T_{\alpha, h}} \min(t_1, t_2) \pi(t_1) \pi(t_2)' K(t_1) K(t_2) dt.
\]

Now (B.4) in the proof of Theorem B.8 and Lemma B.1-(iii,iv) show that (LI) \( \text{Var} \left[ \tilde{R}^{(1)} (b(\alpha|I); \alpha, I) \right] \) admits the expansion, with uniform remainder terms,

\[
\mathbb{E} \left[ \int_{L_{\alpha, h}}^{T_{\alpha, h}} \int_{L_{\alpha, h}}^{T_{\alpha, h}} \left\{ G \left[ B (\alpha + ht_1 | x_\ell, I_\ell) \wedge B (\alpha + ht_2 | x_\ell, I_\ell) + o(h) | x_\ell, I_\ell \right] \\
- G \left[ B (\alpha + ht_1 | x_\ell, I_\ell) + o(h) | x_\ell, I_\ell \right] (\alpha + ht_2) - G \left[ B (\alpha + ht_2 | x_\ell, I_\ell) + o(h) | x_\ell, I_\ell \right] (\alpha + ht_1) \\
+ (\alpha + ht_1) (\alpha + ht_2) \right\} \pi(t_1) \pi(t_2)' K(t_1) K(t_2) dt_1 dt_2 \otimes I (I_\ell = I) P(x_\ell) P(x_\ell) \right]
\]

\[
= \int_{L_{\alpha, h}}^{T_{\alpha, h}} \int_{L_{\alpha, h}}^{T_{\alpha, h}} \left\{ \alpha + h (t_1 \land t_2) - \alpha^2 - h \alpha (t_1 + t_2) \right\} \pi(t_1) \pi(t_2)' K(t_1) K(t_2) dt_1 dt_2 + o(h)
\]

\[
= \alpha (1 - \alpha) \omega_0 \omega_0' \otimes P + h \left\{ \Pi_m - \alpha \left( \omega_0 \omega_1' + \omega_1 \omega_0' \right) \right\} \otimes P + o(h).
\]
Hence an elementary expansion gives, uniformly in \( \alpha \in [0, 1] \), \( \text{Var} (\bar{\mathbf{e}} (\alpha | I)) = \mathcal{V}_e / (L I) + o(h) \) with

\[
\mathcal{V}_e = \alpha (1 - \alpha) \Omega^{-1} \otimes P_0^{-1} \times \omega_0 \omega'_0 \otimes P \times \Omega^{-1} \otimes P_0^{-1}
- h \alpha (1 - \alpha) (\Omega^{-1} \otimes P_0^{-1})^2 \times \Omega_1 \otimes P_1 \times \omega_0 \omega'_0 \otimes P \times \Omega^{-1} \otimes P_0^{-1}
- h \alpha (1 - \alpha) \Omega^{-1} \otimes P_0^{-1} \times \omega_0 \omega'_0 \otimes P \times \Omega_1 \otimes P_1 \times (\Omega^{-1} \otimes P_0^{-1})^2
+ h \Omega^{-1} \otimes P_0^{-1} \times \{ \Pi_m \otimes P - \alpha (\omega_1 \omega'_0 + \omega_0 \omega'_1) \otimes P \} \times \Omega^{-1} \otimes P_0^{-1}
= \alpha (1 - \alpha) (\Omega^{-1} \times \omega_0 \omega'_0 \times \Omega^{-1}) \otimes (P_0^{-1} \times P \times P_0^{-1})
- h \alpha (1 - \alpha) (\Omega^{-2} \times \Omega_1 \times \omega_0 \omega'_0 \times \Omega^{-1}) \otimes (P_0^{-2} \times P_1 \times P \times P_0^{-1})
- h \alpha (1 - \alpha) (\Omega^{-1} \times \omega_0 \omega'_0 \times \Omega_1 \times \Omega^{-2}) \otimes (P_0^{-1} \times P \times P_1 \times P_0^{-2})
+ h (\Omega^{-1} \times \Pi_m \times \Omega^{-1} - \alpha \Omega^{-1} \times (\omega_1 \omega'_0 + \omega_0 \omega'_1) \times \Omega^{-1}) \otimes (P_0^{-1} \times P \times P_0^{-1}).
\]

Since the eigenvalues of \( P_0^{-1}, P, P_1, \Omega^{-1} \) and \( \Omega_1 \) are bounded away from infinity uniformly in \( \alpha \), it follows that \( \max_{\alpha \in [0, 1]} \| \text{Var} (\bar{\mathbf{e}}_0 (\alpha | I)) \| = O(1/L) \) and then

\[
\max_{(\alpha, x) \in [0, 1] \times \mathcal{X}} \text{Var} (P (x)^T \bar{\mathbf{e}}_0 (\alpha | I)) = O \left( \max_{x \in \mathcal{X}} \| P (x) \|^2 \right) = O \left( \frac{1}{L h^{d_M}} \right).
\]

For \( \text{Var} (\bar{\mathbf{e}}_1 (\alpha | I) / h) \), observe that \( \bar{\mathbf{e}}_1 (\alpha | I) = S_1 \bar{\mathbf{e}} (\alpha | I) \) with

\[
S_1 = s'_1 \otimes \text{Id}
\]
where \( \text{Id} \) 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

\[
s'_1 \Omega^{-1} \times (\omega_1 \omega_0' + \omega_0' \omega_1') \times \Omega^{-1} s_1 = s'_1 \Omega^{-1} \omega_1 s'_0 s_1 + s'_1 s_0 \omega_1' \Omega^{-1} s_1 = 0
\]

it follows

\[
S_1 V \epsilon S'_1 = h \left[ s'_1 \left( \Omega^{-1} \times \Pi_m \times \Omega^{-1} - \alpha \Omega^{-1} \times (\omega_1 \omega_0' + \omega_0' \omega_1') \times \Omega^{-1} \right) s_1 \right] \otimes (P_0^{-1} \times P \times P_0^{-1})
\]

\[
= h \left( s'_1 \Omega^{-1} \Pi_m \Omega^{-1} s_1 \right) \left( P_0^{-1} P P_0^{-1} \right)
\]

\[
= h v^2_h (\alpha) \mathbb{E}^{-1} \left[ \frac{\| (I_{\ell} = I) P(x_{\ell}) P(x_{\ell})' \|}{B^{(1)}(\alpha | x_{\ell}, I_{\ell})} \right]
\]

\[
\times \mathbb{E} \left[ \frac{\| (I_{\ell} = I) P(x_{\ell}) P(x_{\ell})' \|}{B^{(1)}(\alpha | x_{\ell}, I_{\ell})} \right] \mathbb{E}^{-1} \left[ \frac{\| (I_{\ell} = I) P(x_{\ell}) P(x_{\ell})' \|}{B^{(1)}(\alpha | x_{\ell}, I_{\ell})} \right]
\]

as \( v^2_h (\alpha) = s'_1 \Omega^{-1} \Pi_m \Omega^{-1} s_1 \). This gives the result for \( \text{Var} (\hat{\epsilon}_1 (\alpha I) / h) \) and \( \text{Var} (P(x)' \hat{\epsilon}_1 (\alpha I) / h) \).

For (ii), we just show that \( \max_{(\alpha, x) \in [0, 1] \times \mathcal{X}} \| P(x)' \hat{\epsilon}_1 (\alpha I) / h \| = O_P \left( (\log L / L h^{d_M+1})^{1/2} \right) \).

Since \( \max_{x \in [0, 1]} \| P(x) \| = O \left( h^{-d_M/2} \right) \) and

\[
\max_{(\alpha, x) \in [0, 1] \times \mathcal{X}} \left\| \frac{P(x)' \hat{\epsilon}_1 (\alpha I)}{h} \right\| \leq \left( \max_{(\alpha, x) \in [0, 1] \times \mathcal{X}} \left\| \frac{P(x)' \hat{\epsilon}_1 (\alpha I)}{h^{1/2} (1 + \| P(x) \|)} \right\| \right) \times h^{-1/2} \left( 1 + \max_{x \in [0, 1]} \| P(x) \| \right)
\]

it is sufficient to show

\[
\max_{(\alpha, x) \in [0, 1] \times \mathcal{X}} \left\| \frac{P(x)' \hat{\epsilon}_1 (\alpha I)}{h^{1/2} (1 + \| P(x) \|)} \right\| = O_P \left( \left( \frac{\log L}{L} \right)^{1/2} \right).
\]
Write
\[ \frac{P(x)' \tilde{e}_1(\alpha | I)}{h^{1/2} (1 + \|P(x)\|)} = \frac{1}{L} \sum_{\ell=1}^{L} \xi_{\ell}(\alpha, x) \]

with
\[ \xi_{\ell}(\alpha, x) = \sum_{i=1}^{I_{\ell}} (\mathbb{I}(I_\ell = I) \xi_{i\ell}(\alpha, x) - \mathbb{E}[\mathbb{I}(I_\ell = I) \xi_{i\ell}(\alpha, x)]) , \]
\[ \xi_{i\ell}(\alpha, x) = \frac{P(x)' S_1 \left[ \mathbb{R}^{(2)}(\bar{b}(\alpha | I) ; \alpha, I) \right]^{-1} P(x_\ell)}{h^{1/2} (1 + \|P(x)\|)} \]
\[ \times \left\{ \int_{L_{x, h}}^{1} \left( \mathbb{I}[B_{i\ell} \leq \Psi(t|x_\ell, \bar{b}(\alpha | I))] - (\alpha + ht) \right) K(t) dt \right\} . \]

This gives, for all \((\alpha, x) \in [0, 1]\)
\[ |\xi_{\ell}(\alpha, x)| \leq C h^{-1/2} \left( \frac{\max_{x \in \mathcal{X}} \|P(x)\|^2}{1 + \max_{x \in \mathcal{X}} \|P(x)\|} \right) \leq M_\infty \text{ with } M_\infty \asymp h^{-(d_M+1)/2} , \]
\[ \text{Var}(\xi_{\ell}(\alpha, x)) \leq C \left( \frac{\max_{x \in \mathcal{X}} \|P(x)\|^2}{(1 + \max_{x \in \mathcal{X}} \|P(x)\|)^2} \right) \leq M_2 \text{ with } M_2 \asymp 1. \]

The Implicit Function Theorem and the FOC \( \mathbb{R}^{(1)}(\bar{b}(\alpha | I) ; \alpha, I) = 0, \) Lemma B.2 with \( (B.4) \) and \( s \geq d_M/2 \) give that \( \alpha \mapsto \bar{b}(\alpha | I) \) is \( \|\cdot\|\)-Lipschitz with a Lipschitz constant of order \( L^C, \) as \( \alpha \mapsto \left[ \mathbb{R}^{(2)}(\bar{b}(\alpha | I) ; \alpha, I) \right]^{-1} \) and \( x \mapsto P(x) / (1 + \|P(x)\|). \) Lemma B.1-(iii), \( 1/(L h^{d_M+1}) = O(1), \) 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
\[ \exp(H_L(\epsilon)) \asymp \left( \frac{L^C}{\epsilon} \right)^C. \]

174
Arguing as in the proof of Lemma B.3 gives, for the item $H_L$ of Proposition C.1,

$$H_L = O((\log L)^{1/2}) + O\left(\frac{\log L}{Lh^{d_M+1}}\right)^{1/2} = O((\log L)^{1/2})$$

and then (C.3) holds.

References


