SFB 823

Discussion Paper
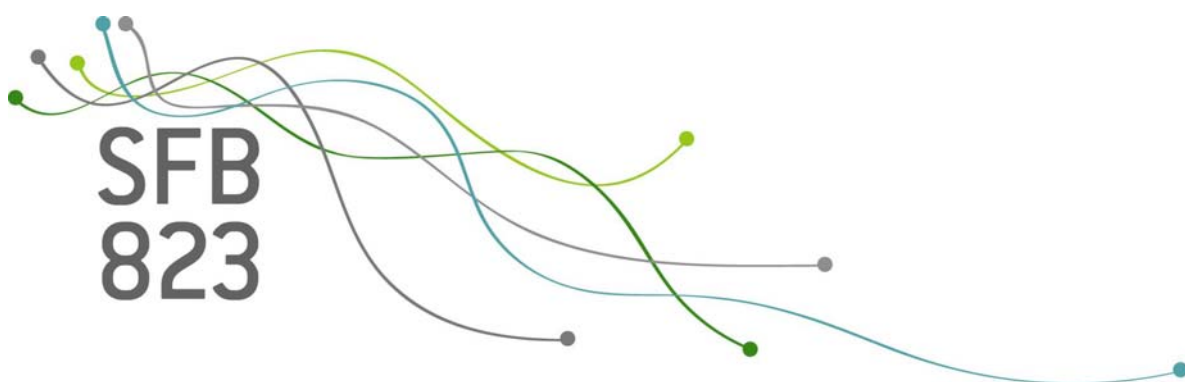
Local constant and local bilinear multiple-output quantile regression

Marc Hallin, Zudi Lu, Davy Paindaveine, Miroslav Siman

SFB 823

LOCAL CONSTANT AND LOCAL BILINEAR MULTIPLE-OUTPUT QUANTILE REGRESSION

By Marc Hallin[*,†,¶,‖,**], Zudi Lu[‡,††], Davy Paindaveine[§,¶,‖]
and Miroslav Šiman[‡‡]

Université Libre de Bruxelles[‖], Princeton University[**], University of
Adelaide[††], and Institute of Information Theory and Automation of the
Academy of Science of the Czech Republic[‡‡]

A new quantile regression concept, based on a directional version of Koenker and Bassett's traditional single-output one, has been introduced in [Hallin, Paindaveine and Šiman, *Annals of Statistics* 2010, 635-703] for multiple-output regression problems. The polyhedral contours provided by the empirical counterpart of that concept, however, cannot adapt to nonlinear and/or heteroskedastic dependencies. This paper therefore introduces local constant and local linear versions of those contours, which both allow to asymptotically recover the conditional halfspace depth contours of the response. In the multiple-output context considered, the local linear construction actually is of a bilinear nature. Bahadur representation and asymptotic normality results are established. Illustrations are provided both on simulated and real data.

1. Introduction. A multiple-output extension of Koenker and Bassett's celebrated concept of regression quantiles was recently proposed in Hallin, Paindaveine, and Šiman [18] (hereafter HPŠ). That extension provides regions that are enjoying, at population level, a double interpretation in terms of quantile and halfspace depth regions. In the empirical case, those contours are polyhedral, and computable via parametric linear programming techniques.

Denote by $(\mathbf{X}_i', \mathbf{Y}_i')' = (X_{i1}, \ldots, X_{ip}, Y_{i1}, \ldots, Y_{im})'$, $i = 1, \ldots, n$, an observed n-tuple of independent copies of $(\mathbf{X}', \mathbf{Y}')'$, where $\mathbf{Y} := (Y_1, \ldots, Y_m)'$ is an m-dimensional response and $\mathbf{X} := (1, \mathbf{W}')'$ a p-dimensional random vector of covariates. For any $\tau \in (0, 1)$ and any direction $\mathbf{u}$ in the unit sphere $\mathcal{S}^{m-1}$ of the m-dimensional space of the response $\mathbf{Y}$, the HPŠ concept produces a hyperplane $\pi_{\tau\mathbf{u}}$ ($\pi_{\tau\mathbf{u}}^{(n)}$ in the empirical case) which is defined as the classical Koenker and Bassett regression quantile hyperplane of order τ once $(\mathbf{0}_{p-1}', \mathbf{u}')'$ has been chosen as the "vertical direction" in the computation of the relevant L_1 deviations. More specifically, decompose $\mathbf{y} \in \mathbb{R}^m$ into $(\mathbf{u}'\mathbf{y})\mathbf{u} + \mathbf{\Gamma}_{\mathbf{u}}(\mathbf{\Gamma}_{\mathbf{u}}'\mathbf{y})$, where $\mathbf{\Gamma}_{\mathbf{u}}$ is such that $(\mathbf{u}, \mathbf{\Gamma}_{\mathbf{u}})$ is an $m \times m$ orthogonal matrix; then the *directional quantile hyperplanes* $\pi_{\tau\mathbf{u}}$ and $\pi_{\tau\mathbf{u}}^{(n)}$ are the hyperplanes with equations

$$(1.1) \quad \mathbf{u}'\mathbf{y} - \mathbf{c}_{\boldsymbol{\tau}}'\mathbf{\Gamma}_{\mathbf{u}}'\mathbf{y} - \mathbf{a}_{\boldsymbol{\tau}}'(1, \mathbf{w}')' = 0 \quad \text{and} \quad \mathbf{u}'\mathbf{y} - \mathbf{c}_{\boldsymbol{\tau}}^{(n)'}\mathbf{\Gamma}_{\mathbf{u}}'\mathbf{y} - \mathbf{a}_{\boldsymbol{\tau}}^{(n)'}(1, \mathbf{w}')' = 0$$

($\mathbf{w} \in \mathbb{R}^{p-1}$) minimizing, with respect to $\mathbf{c} \in \mathbb{R}^{m-1}$ and $\mathbf{a} \in \mathbb{R}^p$,

$$(1.2) \quad \mathrm{E}[\rho_\tau(\mathbf{u}'\mathbf{Y} - \mathbf{c}'\mathbf{\Gamma}_{\mathbf{u}}'\mathbf{Y} - \mathbf{a}'\mathbf{X})] \quad \text{and} \quad \sum_{i=1}^{n} \rho_\tau(\mathbf{u}'\mathbf{Y}_i - \mathbf{c}'\mathbf{\Gamma}_{\mathbf{u}}'\mathbf{Y}_i - \mathbf{a}'\mathbf{X}_i),$$

respectively, where

$$\zeta \mapsto \rho_\tau(\zeta) := \zeta(\tau - \mathrm{I}[\zeta < 0]) = \max\{(\tau-1)\zeta, \tau\zeta\} = (|\zeta| + (2\tau-1)\zeta)/2, \quad \zeta \in \mathbb{R}$$

as usual denotes the well-known τ-quantile *check function*.

HPŠ show that $\pi_{\tau\mathbf{u}}$ and $\pi_{\tau\mathbf{u}}^{(n)}$ equivalently can be defined, in a more symmetric way, as the hyperplanes with equations

$$(1.3) \qquad \mathbf{b}_{\boldsymbol{\tau}}'\mathbf{y} - \mathbf{a}_{\boldsymbol{\tau}}'(1, \mathbf{w}')' = 0 \quad \text{and} \quad \mathbf{b}_{\boldsymbol{\tau}}^{(n)'}\mathbf{y} - \mathbf{a}_{\boldsymbol{\tau}}^{(n)'}(1, \mathbf{w}')' = 0,$$

minimizing, with respect to $\mathbf{b} \in \mathbb{R}^m$ satisfying $\mathbf{b}'\mathbf{u} = 1$ and $\mathbf{a} \in \mathbb{R}^p$, the L_1 criteria

$$(1.4) \qquad \mathrm{E}[\rho_\tau(\mathbf{b}'\mathbf{Y} - \mathbf{a}'\mathbf{X})] \quad \text{and} \quad \sum_{i=1}^{n} \rho_\tau(\mathbf{b}'\mathbf{Y}_i - \mathbf{a}'\mathbf{X}_i),$$

respectively.

For $p = 1$, the multiple-output regression model reduces to a multivariate location one: $\mathbf{a}_{\boldsymbol{\tau}}$ and $\mathbf{a}_{\boldsymbol{\tau}}^{(n)}$ reduce to scalars, $a_{\boldsymbol{\tau}}$ and $a_{\boldsymbol{\tau}}^{(n)}$, while the equations describing $\pi_{\tau\mathbf{u}}$ and $\pi_{\tau\mathbf{u}}^{(n)}$ take the simpler forms

$$(1.5) \qquad \mathbf{u}'\mathbf{y} - \mathbf{c}_{\boldsymbol{\tau}}'\boldsymbol{\Gamma}_{\mathbf{u}}'\mathbf{y} - a_{\boldsymbol{\tau}} = 0 \quad \text{and} \quad \mathbf{u}'\mathbf{y} - \mathbf{c}_{\boldsymbol{\tau}}^{(n)'}\boldsymbol{\Gamma}_{\mathbf{u}}'\mathbf{y} - a_{\boldsymbol{\tau}}^{(n)} = 0,$$

respectively. Those *location* quantile hyperplanes $\pi_{\tau\mathbf{u}}$ and $\pi_{\tau\mathbf{u}}^{(n)}$ are studied in detail in HPŠ, where it is shown that their fixed-τ collections characterize regions and contours that actually coincide with the Tukey halfspace depth ones. Consistency, asymptotic normality and Bahadur-type representation results for the $\pi_{\tau\mathbf{u}}^{(n)}$'s are also provided there, together with a linear programming method for their computation.

Those results establish a strong and quite fruitful link between two seemingly unrelated statistical worlds—on one hand the typically one-dimensional concept of quantiles, deeply rooted into the strong ordering features of the real line and L_1 optimality, with linear programming algorithms, and traditional central-limit asymptotics; the intrinsically multivariate concept of depth on the other hand, with geometric characterizations, computationally intensive combinatorial algorithms, and nonstandard asymptotics. From their relation to depth, quantile hyperplanes and regions inherit a variety of geometric properties—connectedness, nestedness, convexity, affine-equivariance... while, through its relation to quantiles, depth accedes to L_1

optimality, feasible linear programming algorithms, and tractable asymptotics.

The situation is less satisfactory in the general regression case ($p \geq 2$). The above definitions still produce regions and contours indexed by τ and, in the empirical case, efficient linear programming methods are still available; see [28]. Those regions and contours still admit an interpretation in terms of (joint) directional quantiles. However, that interpretation is only remotely related to the regression problem under study. If indeed $\mathbf{Y} = (Y_1, \ldots, Y_m)'$ is a response and $\mathbf{X} = (1, \mathbf{W}')'$ a vector of covariates, the objective is an analysis of the influence of the covariate(s) $\mathbf{W}$ on the response $\mathbf{Y}$, that is, a study of the distribution of $\mathbf{Y}$ conditional on $\mathbf{W}$. The contours of interest, thus, are the collection of the population *conditional quantile/depth contours* of $\mathbf{Y}$, indexed by the values $\mathbf{w} \in \mathbb{R}^{p-1}$ of $\mathbf{W}$—that is, for each $\mathbf{w}$, the collection of conditional (on $\mathbf{W} = \mathbf{w}$) location ($p = 1$) quantile/depth contours. Equations (1.2) or (1.4) being of a *global* (with respect to $\mathbf{W}$ or $\mathbf{X}$) nature, the resulting hyperplanes and contours, unfortunately, in general carry very little information about conditional distributions, and rather produce some averaged (over the covariate space) quantile/depth contours.

Of course, this problem is not specific to the multiple-output context. In the traditional single-output setting, it has motivated weighted, local polynomial and nearest-neighbor versions of quantile regression, among others. We refer to [39–41] for conceptual insight and practical information, to [3, 7, 16, 17, 24, 42] for some recent asymptotic results, and to [1, 4, 13, 14, 20–22, 34] for some less recent ones.

Our objective in this paper is to extend those local estimation ideas to the HPŠ concept of multiple-output regression quantiles. Since local constant and local linear methods have been shown to perform extremely well in the single-output single-regressor case ([40]), we will concentrate on local

constant and local bilinear approaches—in the multiple-output context, indeed, it turns out that the adequate extension of locally linear procedures are of a *bilinear* nature. Just as in the single-output case, the local methods we propose in this paper will not require any a priori knowledge of the trend and will still allow to characterize asymptotically the whole conditional distribution of $\mathbf{Y}$ given $\mathbf{W} = \mathbf{w}$ for any $\mathbf{w} \in \mathbb{R}^{p-1}$.

A major application of this local approach to multiple-output quantile regression is the analysis of multivariate growth charts. Growth charts (reference curves, percentile curves) have been used for a long time by practitioners in order to assess the impact of regressors on the quantiles of some given single variable of interest. Many methods have been developed (see, e.g., [2, 5, 36, 38], and the references therein), including single-response quantile regression (see [12, 37]). Only a few attempts have been made, mainly in the bivariate case ([11, 30]), to adapt that daily practice instrument to a multiple-response context. The only method available for that case is, to the best of our knowledge, the recent proposal by [35], that defines a new concept of dynamic quantile regression contours. Our local methodology, which is based on entirely different principles, appears as a natural alternative. See [26] for a real-data example of bivariate growth charts based on the methods we are describing here.

The rest of this paper is organized as follows. Section 2 defines the (population) conditional regression quantile regions and contours we would like to estimate in the sequel. This estimation will make use of (empirical) *weighted* multiple-output regression quantiles, which we introduce in Section 3. Section 4 explains how these weighted quantiles lead to local constant (Section 4.2) and local bilinear (Section 4.3) quantiles. Section 5 provides asymptotic results (Bahadur representation and asymptotic normality) both for the local constant and local bilinear cases. In Section 6, simulated and real

data are used to demonstrate the usefulness of the proposed local quantile regions. Finally, the Appendix collects proofs of asymptotic results.

2. Conditional multiple-output quantile regression. As in the Introduction, consider the regression setup involving the m-variate response $\mathbf{Y}$ and the p-variate covariate $\mathbf{X} = (1, \mathbf{W}')'$, with the objective of analyzing the distribution of $\mathbf{Y}$ conditional on $\mathbf{W}$, that is, of fully investigating the dependence of $\mathbf{Y}$ on $\mathbf{W}$—in strong contrast with traditional regression, where investigation is limited to the mean of $\mathbf{Y}$ conditional on $\mathbf{W}$. The relevant quantile hyperplanes, depth regions and contours of interest are the location quantile/depth regions and contours associated (in the sense of HPŠ) with the m-dimensional distributions of $\mathbf{Y}$ conditional on $\mathbf{W}$—more precisely, with the distributions $\mathrm{P}^{\mathbf{Y}|\mathbf{W}=\mathbf{w}_0}$ of $\mathbf{Y}$ conditional on $\mathbf{W} = \mathbf{w}_0$ ($\mathbf{w}_0 \in \mathbb{R}^{p-1}$). We now carefully define these objects, that we will call $\mathbf{w}_0$-*conditional $\boldsymbol{\tau}$-quantile hyperplanes, regions and contours.*

Let $\tau \in (0, 1)$ and $\mathbf{u} \in \mathcal{S}^{m-1} := \{\mathbf{u} \in \mathbb{R}^m : \|\mathbf{u}\| = 1\}$ (the unit sphere in $\mathbb{R}^m$), and write $\boldsymbol{\tau} := \tau \mathbf{u}$. Denoting by $\mathbf{w}_0$ some fixed point of $\mathbb{R}^{p-1}$ at which the marginal density $f^{\mathbf{W}}$ of $\mathbf{W}$ does not vanish (in order for the distribution of $\mathbf{Y}$ conditional on $\mathbf{W} = \mathbf{w}_0$ to make sense), define the *extended* and *restricted* $\mathbf{w}_0$-*conditional $\boldsymbol{\tau}$-quantile hyperplanes of* $\mathbf{Y}$ as the $(m+p-2)$-dimensional and $(m-1)$-dimensional hyperplanes

$$(2.1) \qquad \boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0} := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{p-1} \times \mathbb{R}^m \,|\, \mathbf{b}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{y} - a_{\boldsymbol{\tau};\mathbf{w}_0} = 0\}$$

and

$$(2.2) \qquad \pi_{\boldsymbol{\tau};\mathbf{w}_0} := \{(\mathbf{w}'_0, \mathbf{y}')' \in \mathbb{R}^{p-1} \times \mathbb{R}^m \,|\, \mathbf{b}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{y} - a_{\boldsymbol{\tau};\mathbf{w}_0} = 0\},$$

respectively, where $a_{\boldsymbol{\tau};\mathbf{w}_0}$ and $\mathbf{b}_{\boldsymbol{\tau};\mathbf{w}_0}$ minimize

$$(2.3) \qquad \Psi_{\boldsymbol{\tau};\mathbf{w}_0}(a, \mathbf{b}) := \mathrm{E}[\rho_\tau(\mathbf{b}'\mathbf{Y} - a) \,|\, \mathbf{W} = \mathbf{w}_0] \quad \text{subject to } \mathbf{b}'\mathbf{u} = 1,$$

with the check function ρ_τ defined in Page 2. Comparing (2.3) with (1.4) immediately shows that $\pi_{\boldsymbol{\tau};\mathbf{w}_0}$ is the (location) $(m-1)$-dimensional $\boldsymbol{\tau}$-quantile hyperplane of $\mathbf{Y}$ associated with the distribution of $\mathbf{Y}$ conditional on $\mathbf{W} = \mathbf{w}_0$. Of course, $\pi_{\boldsymbol{\tau};\mathbf{w}_0}$ is also the intersection of $\boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0}$ with the m-dimensional hyperplane $C_{\mathbf{w}_0} := \{(\mathbf{w}_0', \mathbf{y}')' \,|\, \mathbf{y} \in \mathbb{R}^m\}$. This, and the fact that $\boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0}$ is "parallel to the space of covariates" (in the sense that if $(\mathbf{w}_0', \mathbf{y}_0')' \in \boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0}$, then $(\mathbf{w}', \mathbf{y}_0')' \in \boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0}$ for all $\mathbf{w}$), fully characterizes $\boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0}$.

Associated with $\boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0}$ are the extended *upper and lower* $\mathbf{w}_0$-*conditional $\boldsymbol{\tau}$-quantile halfspaces*

$$\boldsymbol{H}_{\boldsymbol{\tau};\mathbf{w}_0}^+ := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{p-1} \times \mathbb{R}^m \,|\, \mathbf{b}_{\boldsymbol{\tau};\mathbf{w}_0}'\mathbf{y} - a_{\boldsymbol{\tau};\mathbf{w}_0} \geq 0\}$$

and

$$\boldsymbol{H}_{\boldsymbol{\tau};\mathbf{w}_0}^- := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{p-1} \times \mathbb{R}^m \,|\, \mathbf{b}_{\boldsymbol{\tau};\mathbf{w}_0}'\mathbf{y} - a_{\boldsymbol{\tau};\mathbf{w}_0} < 0\},$$

together with the extended (cylindrical) $\mathbf{w}_0$-*conditional quantile/depth regions*

$$(2.4) \qquad \boldsymbol{R}_{\mathbf{w}_0}(\tau) := \bigcap_{\mathbf{u} \in \mathcal{S}^{m-1}} \{\boldsymbol{H}_{\tau\mathbf{u};\mathbf{w}_0}^+\}$$

and their boundaries $\partial \boldsymbol{R}_{\mathbf{w}_0}(\tau)$, the extended $\mathbf{w}_0$-*conditional quantile/depth contours*. The intersections of those extended regions $\boldsymbol{R}_{\mathbf{w}_0}(\tau)$ (resp., contours $\partial \boldsymbol{R}_{\mathbf{w}_0}(\tau)$) with $C_{\mathbf{w}_0}$ are the restricted $\mathbf{w}_0$-conditional quantile/depth regions $R_{\mathbf{w}_0}(\tau)$ (resp., contours $\partial R_{\mathbf{w}_0}(\tau)$), that is, the location HPŠ regions (resp., contours) for $\mathbf{Y}$, conditional on $\mathbf{W} = \mathbf{w}_0$. It follows from HPŠ that those regions are compact, convex, and nested. As a consequence, the regions $\boldsymbol{R}_{\mathbf{w}_0}(\tau)$ are closed, convex, and nested.

Finally, define the *nonparametric τ-quantile/depth regions* as

$$\boldsymbol{R}(\tau) := \bigcup_{\mathbf{w}_0 \in \mathbb{R}^{p-1}} R_{\mathbf{w}_0}(\tau) = \bigcup_{\mathbf{w}_0 \in \mathbb{R}^{p-1}} \left(\boldsymbol{R}_{\mathbf{w}_0}(\tau) \cap C_{\mathbf{w}_0}\right)$$

and write $\partial \boldsymbol{R}(\tau)$ for their boundaries. The regions $\boldsymbol{R}(\tau)$ are still closed and nested but they adapt to the general dependence of $\mathbf{Y}$ on $\mathbf{W}$: in particu-

lar, $\partial\boldsymbol{R}(\tau)$, for any τ, goes through *all* corresponding $\partial R_{\mathbf{w}_0}(\tau)$'s, $\mathbf{w}_0 \in \mathbb{R}^{p-1}$. Consequently, the regions $\boldsymbol{R}(\tau)$ in general are no longer convex.

The fixed-$\mathbf{w}_0$ collection (over $\tau \in (0, 1/2)$) of the $\mathbf{w}_0$-conditional location quantile/depth contours $\partial R_{\mathbf{w}_0}(\tau)$ (which, by construction, are the intersections of $\partial\boldsymbol{R}(\tau)$ with the "vertical hyperplanes" $C_{\mathbf{w}_0}$) will be called $\mathbf{w}_0$-*quantile/depth cut* or simply $\mathbf{w}_0$-*cut*. Such cuts are of crucial interest, since they entirely characterize the distribution of $\mathbf{Y}$ conditional on $\mathbf{W} = \mathbf{w}_0$, hence provide a full description of the dependence of the response $\mathbf{Y}$ on the regressors $\mathbf{W}$. Note that the nonparametric contours $\partial\boldsymbol{R}(\tau)$, via the location depth interpretation, for fixed $\mathbf{w}_0$, of the $\partial R_{\mathbf{w}_0}(\tau)$'s, inherit a most interesting interpretation as "regression depth contours". Clearly, this concept of regression depth, that defines regression depth of any point $(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{m+p-1}$, is not of the same nature as the regression depth concept proposed in [31], that defines the depth of any regression "fit" (i.e., of any regression hyperplane).

3. Weighted multiple-output empirical quantile regression. Under the assumption of absolute continuity of the distribution of $\mathbf{W}$, the number of observations, in a sample of size n, belonging to $C_{\mathbf{w}_0}$ clearly is (a.s.) zero, which implies that no empirical version of the conditional regression hyperplanes (2.1) or (2.2) can be constructed. If nonparametric τ-quantile regions or contours, or simply some selected cuts, are to be estimated, local smoothing techniques have to be considered. Those local techniques will typically be based on weighted versions, involving adequate sequences $\boldsymbol{\omega}_{\mathbf{w}_0}^{(n)} = (\omega_{\mathbf{w}_0,i}^{(n)}, i = 1, \ldots, n)$ of weights, of the original concept of empirical quantile regression hyperplanes developed in HPŠ. In this section, we provide general definitions and basic results for such weighted concepts, under fixed sample size n and fixed weights ω_i. In Section 4, we will then consider sequences of kernel-based weights to be used in the local approach.

Consider a sample of size n, with observations $(\mathbf{X}_i', \mathbf{Y}_i')' = ((1, \mathbf{W}_i'), \mathbf{Y}_i')'$, $i = 1, \ldots, n$, along with n nonnegative weights ω_i satisfying (without any loss of generality) $\sum_{i=1}^n \omega_i = n$ ($\omega_i \equiv 1$ then yields the unweighted case). The definitions of HPŠ extend, *mutatis mutandis*, quite straightforwardly, into the following weighted versions. The coefficients $\mathbf{a}_{\boldsymbol{\tau};\omega}^{(n)} \in \mathbb{R}^p$ and $\mathbf{b}_{\boldsymbol{\tau};\omega}^{(n)} \in \mathbb{R}^m$ of the *weighted empirical $\boldsymbol{\tau}$-quantile hyperplane*

$$(3.1) \qquad \boldsymbol{\pi}_{\boldsymbol{\tau};\omega}^{(n)} := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{p-1} \times \mathbb{R}^m \,|\, \mathbf{b}_{\boldsymbol{\tau};\omega}^{(n)\prime} \mathbf{y} - \mathbf{a}_{\boldsymbol{\tau};\omega}^{(n)\prime}(1, \mathbf{w}')' = 0\}$$

(an $(m + p - 2)$-dimensional hyperplane) are defined as the minimizers, under $\mathbf{b}'\mathbf{u} = 1$, of

$$(3.2) \qquad \Psi_{\tau;\omega}^{(n)}(\mathbf{a}, \mathbf{b}) := \frac{1}{n} \sum_{i=1}^n \omega_i \rho_\tau(\mathbf{b}'\mathbf{Y}_i - \mathbf{a}'\mathbf{X}_i) \quad \text{subject to } \mathbf{b}'\mathbf{u} = 1.$$

As usual in the empirical case, the minimizer may not be unique, but the minimizers always form a convex set. When substituted for the $\boldsymbol{\pi}_{\boldsymbol{\tau};\mathbf{w}_0}$'s in the definitions of upper and lower conditional $\boldsymbol{\tau}$-quantile halfspaces, those $\boldsymbol{\pi}_{\boldsymbol{\tau};\omega}^{(n)}$'s also characterize upper and lower weighted $\boldsymbol{\tau}$-quantile halfspaces $\boldsymbol{H}_{\boldsymbol{\tau};\omega}^{(n)+}$ and $\boldsymbol{H}_{\boldsymbol{\tau};\omega}^{(n)-}$, with *weighted τ-quantile regions and contours*

$$\boldsymbol{R}_\omega^{(n)}(\tau) := \bigcap_{\mathbf{u} \in \mathcal{S}^{m-1}} \{\boldsymbol{H}_{\tau\mathbf{u};\omega}^{(n)+}\} \quad \text{and} \quad \partial\boldsymbol{R}_\omega^{(n)}(\tau),$$

respectively.

Note that the objective function in (3.2) rewrites as

$$\Psi_{\tau;\omega}^{(n)}(\mathbf{a}, \mathbf{b}) = \frac{1}{n} \sum_{i=1}^n \rho_\tau(\mathbf{b}'\mathbf{Y}_{i;\omega} - \mathbf{a}'\mathbf{X}_{i;\omega}),$$

with $\mathbf{X}_{i;\omega} := \omega_i \mathbf{X}_i$ and $\mathbf{Y}_{i;\omega} := \omega_i \mathbf{Y}_i$. As an important consequence, the weighted quantile hyperplanes, contours and regions can be computed in the same way as their non-weighted counterparts because the corresponding algorithm in [28] allows to have $(\mathbf{X}_i)_1 \neq 1$. Due to the quantile crossing phenomenon, however, and contrary to the population regions and contours

defined in the previous section, the $\boldsymbol{R}_{\omega}^{(n)}(\tau)$'s need not be nested for $p \geq 2$; if nestedness is required, one may rather consider the regions

$$\boldsymbol{R}_{\omega\cap}^{(n)}(\tau) := \bigcap_{0 < t \leq \tau} \{\boldsymbol{R}_{\omega}^{(n)}(t)\}.$$

The necessary sample subgradient conditions for $(\mathbf{a}_{\boldsymbol{\tau};\omega}^{(n)\prime}, \mathbf{b}_{\boldsymbol{\tau};\omega}^{(n)\prime})'$ can be derived as for the unweighted case. They state in particular that

(3.3)
$$\frac{1}{n}\sum_{i=1}^{n} \omega_i \, \mathrm{I}[\mathbf{b}_{\boldsymbol{\tau};\omega}^{(n)\prime}\mathbf{Y}_i - \mathbf{a}_{\boldsymbol{\tau};\omega}^{(n)\prime}\mathbf{X}_i < 0] \leq \tau \leq \frac{1}{n}\sum_{i=1}^{n} \omega_i \, \mathrm{I}[\mathbf{b}_{\boldsymbol{\tau};\omega}^{(n)\prime}\mathbf{Y}_i - \mathbf{a}_{\boldsymbol{\tau};\omega}^{(n)\prime}\mathbf{X}_i \leq 0],$$

which controls the probability contents of $\boldsymbol{H}_{\boldsymbol{\tau};\omega}^{(n)-}$ with respect to the distribution putting probability mass ω_i/n on $(\mathbf{W}_i', \mathbf{Y}_i')'$, $i = 1, \ldots, n$. The width of the interval in (3.3) depends only on the weights ω_i associated with those data points $(\mathbf{W}_i', \mathbf{Y}_i')'$ that belong to $\boldsymbol{\pi}_{\boldsymbol{\tau};\omega}^{(n)}$. Another consequence worth mentioning is that there always exists a $\boldsymbol{\pi}_{\tau\mathbf{u};\omega}^{(n)}$ hyperplane containing at least $(m + p - 1)$ data points of the form $(\omega_i \mathbf{W}_i, \omega_i \mathbf{Y}_i)$. With probability one, thus, the intersection defining the regions $\boldsymbol{R}_{\omega}^{(n)}(\tau)$ is finite.

Note that, unlike the extended conditional quantile hyperplanes (2.1), the weighted empirical quantile hyperplanes (3.1) involve an unrestricted coefficient $\mathbf{a} \in \mathbb{R}^p$. As a consequence, $\boldsymbol{\pi}_{\boldsymbol{\tau};\omega}^{(n)}$ is not necessarily parallel to the space of covariates (as defined in Page 7). That degree of freedom will be exploited in the local linear approach described in Section 4.3 (in an augmented regressor space, though, which makes it bilinear rather than linear). If we impose the constraint $\mathbf{a} = (a_1, 0, \ldots, 0)'$ in (3.1), we obtain hyperplanes of the form

(3.4) $\boldsymbol{\pi}_{\boldsymbol{\tau};\omega}^{(n)} := \{(\mathbf{w}', \mathbf{y}')' \in \mathbb{R}^{p-1} \times \mathbb{R}^m \,|\, \mathbf{b}_{\boldsymbol{\tau};\omega}^{(n)\prime}\mathbf{y} - (\mathbf{a}_{\boldsymbol{\tau};\omega}^{(n)})_1 = 0\}.$

The corresponding minimization problem yields hyperplanes that are parallel to the space of covariates, hence cylindrical weighted regions and contours

that are "horizontal". For the sake of simplicity, we avoid introducing a specific notation for them; such cylindrical contours will be considered in the local constant approach described in Section 4.2.

Finally, it should be pointed out that ($\mathbf{y}$ and/or $\mathbf{w}$)-affine-invariant weights $\omega_i := \omega(\mathbf{W}_i, \mathbf{Y}_i)$ imply good ($\mathbf{y}$ and/or $\mathbf{w}$)-affine-equivariance properties of the corresponding weighted quantile hyperplanes, halfspaces, regions, and contours considered.

4. Local multiple-output quantile regression.

4.1. *From weighted to local quantile regression.* The weighted quantiles of Section 3 have an interest on their own. They can be used for handling multiple identical observations (allowing, for instance, for bootstrap procedures), or for downweighting observations that are suspected to be outliers or leverage points. Above all, weighted regression quantiles allow for a nonparametric approach to regression quantiles that will take care of the drawbacks of the unweighted approach of HPŠ (see the example considered in the Introduction). In particular, adequate sequences of weights will allow to estimate the conditional contours described in Section 2, thus extending to the multiple-output case the *local constant* and *local linear* approaches to regression quantiles proposed, for instance, by [39, 40] in the single-output context.

The basic idea is very standard: in order to estimate $\mathbf{w}_0$-conditional quantile hyperplanes, regions or contours, we will consider weighted quantile hyperplanes, regions or contours, with sequences of weights $\omega_i^{(n)} := \omega_{\mathbf{w}_0}^{(n)}(\mathbf{W}_i)$ based on weight functions of the form

$$(4.1) \qquad \mathbf{w} \mapsto \omega_{\mathbf{w}_0}^{(n)}(\mathbf{w}) := \det(\mathbf{H}_0^{(n)})^{-1} K\big((\mathbf{H}_0^{(n)})^{-1}(\mathbf{w} - \mathbf{w}_0)\big),$$

where $\mathbf{H}_0^{(n)}$ is a sequence of symmetric positive definite $(p-1) \times (p-1)$ bandwidth matrices and K is a nonnegative kernel (density) function over $\mathbb{R}^{p-1}$.

The literature proposes a variety of possible kernels, and there is no compelling reason for not considering the most usual ones, such as

(i) the rectangular (uniform) kernel $K_1(\mathbf{w}) = 2^{-(p-1)} \, \mathrm{I}[\mathbf{w} \in [-1,1]^{p-1}]$,

(ii) the Epanechnikov kernel $K_2(\mathbf{w}) = \dfrac{(p^2-1)\Gamma(\frac{p-1}{2})}{4\pi^{(p-1)/2}}(1-\mathbf{w}'\mathbf{w})\,\mathrm{I}[\mathbf{w}'\mathbf{w} \leq 1]$, or

(iii) the (spherical) Gaussian kernel $K_3(\mathbf{w}) = (2\pi)^{-(p-1)/2}\exp(-\mathbf{w}'\mathbf{w}/2)$.

As for the bandwidth matrices, we will restrict to the simple scalar case, that is, to $\mathbf{H}_0^{(n)} = h_n\mathbf{I}_{p-1}$ and write $K_h(\mathbf{w}-\mathbf{w}_0)$ for the weight $\omega_{\mathbf{w}_0}^{(n)}(\mathbf{w})$.

Since we typically intend to compute by means of parametric programming, for any fixed $\tau \in (0,1)$, the directional quantile hyperplanes for all $\mathbf{u} \in \mathcal{S}^{m-1}$, we should use the same weights for all of them. This is why we only consider $\mathbf{u}$-independent (actually, even $\boldsymbol{\tau}$-independent) bandwidths. However, exact computation of all quantiles (for each fixed τ) is possible in the local constant case, but not in the local bilinear one. In the latter case, quantile contours will be approximated by sampling the unit sphere, which of course allows us, if we wish, to have $\mathbf{u}$-dependent bandwidths.

The weights considered above cover both kernel and nearest-neighbor quantile regression but exclude more sophisticated techniques such as double-kernel-, supersmoother- or LOWESS-based modifications. On the other hand, the choice of weights has no impact on computational issues, and special kernels (and bandwidths) can be selected for extreme $\mathbf{w}_0$'s to take care of boundary effects, for instance.

4.2. *Local constant quantile contours.* If we only care about $\mathbf{w}_0$-conditional contours, that is, $\mathbf{w}_0$-cuts, for a selected number of $\mathbf{w}_0$ values, the above weighting scheme can be applied in the computation of weighted cylindrical regions generated by the hyperplanes in (3.4) (that are parallel to the space of covariates); more precisely, these cylindrical regions, with edges that are

parallel to the space of covariates, are obtained by computing the intersection (over all directions $\mathbf{u}$, for fixed τ) of the upper quantile halfspaces associated with the quantile hyperplanes in (3.4); see Figure 1(a).

The intersection with the $\mathbf{w} = \mathbf{w}_0$ hyperplane of these cylindrical regions yields a local constant estimate, $\partial \hat{R}_{\mathbf{w}_0}^{(n)\mathrm{const}}(\tau)$ say, of the corresponding population $\mathbf{w}_0$-cut $\partial R_{\mathbf{w}_0}(\tau)$; see Section 5 for asymptotic results. Of course, the resulting local constant τ-quantile/depth contours, namely

$$\partial \hat{\boldsymbol{R}}^{(n)\mathrm{const}}(\tau) := \bigcup_{\mathbf{w}_0 \in \mathbb{R}^{p-1}} \partial \hat{R}_{\mathbf{w}_0}^{(n)\mathrm{const}}(\tau),$$

are not (globally) cylindrical, but rather adapt to the underlying possibly nonlinear and/or heteroskedastic dependence structures.
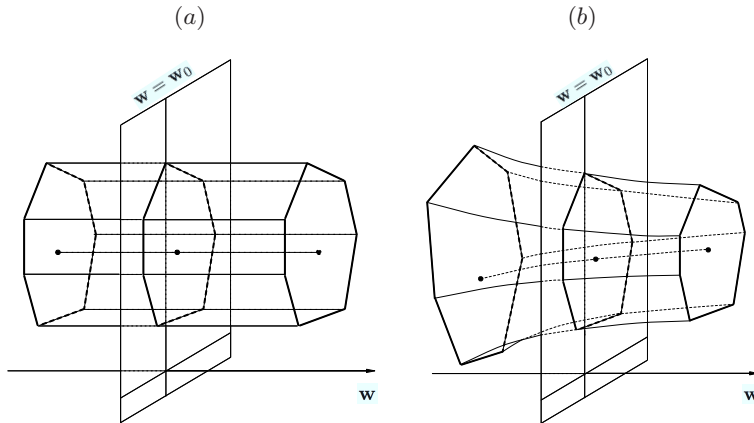

FIG 1. *Construction of the local constant (a) and local bilinear (b) τ-quantile regions from the intersection (over all possible directions $\mathbf{u}$) of the upper quantile halfspaces associated with the constrained-to-be-parallel-to-the-space-of-covariates (a) or unconstrained (b) $(\tau\mathbf{u})$-quantile hyperplanes.*

This approach, which constitutes a generalization of the local constant approach adopted elsewhere for single-output regression, has many advantages. The main one is parsimony: each quantile hyperplane involved in the construction of the weighted contours only entails m parameters, that is,

considerably less than the local bilinear approach described in the next section. On the other hand, the local constant approach does not provide any information on, nor does take any advantage of, the behavior of $\mathbf{w}$-cuts for $\mathbf{w}$ values in the neighborhood of $\mathbf{w}_0$, and its boundary performances are likely to be poor. These two reasons, in traditional contexts, have motivated the development of local linear and local polynomial methods; see [8] for a classical reference. Local linear methods were successfully used in single-output quantile regression ([39–42]). Considering them in the present context, thus, is a quite natural idea.

4.3. *Local bilinear quantile contours.* Assume that the distribution of $(\mathbf{W}', \mathbf{Y}')'$ is smooth enough that the coefficients of $\mathbf{w}$-conditional quantile hyperplanes are differentiable with respect to $\mathbf{w}$. Getting back to the first characterization (1.1)-(1.2) of quantile hyperplanes, the (restricted) $\mathbf{w}_0$-conditional $\boldsymbol{\tau}$-quantile hyperplane of $\mathbf{Y}$ defined in (2.2)-(2.3) has equation (in $\mathbf{y}$—of course, in $\mathbf{w}$, we just have $\mathbf{w} = \mathbf{w}_0$)

$$(4.2) \qquad \mathbf{u}'\mathbf{y} - \left(a_{\boldsymbol{\tau};\mathbf{w}_0}, \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\right) \begin{pmatrix} 1 \\ \boldsymbol{\Gamma}'_{\mathbf{u}}\mathbf{y} \end{pmatrix} = 0.$$

The same hyperplane equation, relative to a point $\mathbf{w}$ in the neighborhood of $\mathbf{w}_0$, takes the form

$$(4.3) \quad \mathbf{u}'\mathbf{y} - \left(a_{\boldsymbol{\tau};\mathbf{w}_0}, \ \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\right) \begin{pmatrix} 1 \\ \boldsymbol{\Gamma}'_{\mathbf{u}}\mathbf{y} \end{pmatrix}$$
$$- (\mathbf{w} - \mathbf{w}_0)' \left(\dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}_0}, \ \dot{\mathbf{c}}'_{\boldsymbol{\tau};\mathbf{w}_0}\right) \begin{pmatrix} 1 \\ \boldsymbol{\Gamma}'_{\mathbf{u}}\mathbf{y} \end{pmatrix} + o(\|\mathbf{w} - \mathbf{w}_0\|) = 0,$$

where $\dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}_0}$ stands for the gradient of $\mathbf{w} \mapsto a_{\boldsymbol{\tau};\mathbf{w}}$ and $\dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w}}$ for the Jacobian matrix of $\mathbf{w} \mapsto \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}}$, respectively, both taken at $\mathbf{w} = \mathbf{w}_0$. In order to express this equation into the equivalent quantile formulation in (1.3)-(1.4), note that we have $\mathbf{b}_{\boldsymbol{\tau};\mathbf{w}_0} = \mathbf{u} - \boldsymbol{\Gamma}_{\mathbf{u}}\mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}$, which entails $\dot{\mathbf{b}}_{\boldsymbol{\tau};\mathbf{w}_0} = -\boldsymbol{\Gamma}_{\mathbf{u}}\dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w}_0}$,

where $\dot{\mathbf{b}}_{\boldsymbol{\tau};\mathbf{w}_0}$ is the Jacobian matrix of $\mathbf{w} \mapsto \mathbf{b}_{\boldsymbol{\tau};\mathbf{w}}$ at $\mathbf{w} = \mathbf{w}_0$. Neglecting the $o(\|\mathbf{w} - \mathbf{w}_0\|)$ term, (4.3) then rewrites, after some algebra, as

$$(4.4) \quad \left(\mathbf{b}'_{\boldsymbol{\tau};\mathbf{w}_0} - \mathbf{w}'_0 \dot{\mathbf{b}}'_{\boldsymbol{\tau};\mathbf{w}_0} \right) \mathbf{y}$$

$$- \left(a_{\boldsymbol{\tau};\mathbf{w}_0} - \mathbf{w}'_0 \dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}_0}, \ \dot{\mathbf{a}}'_{\boldsymbol{\tau};\mathbf{w}_0}, \ -(\text{vec } \dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w}_0})' \right) \begin{pmatrix} 1 \\ \mathbf{w} \\ \mathbf{w} \otimes (\boldsymbol{\Gamma}'_{\mathbf{u}} \mathbf{y}) \end{pmatrix} = 0.$$

Letting $\bar{\mathbf{x}} := (1, \bar{\mathbf{w}}')' := (1, \mathbf{w}', (\mathbf{w} \otimes \boldsymbol{\Gamma}'_{\mathbf{u}} \mathbf{y})')'$, the latter equation is of the form

$$\boldsymbol{\beta}'_{\boldsymbol{\tau}} \mathbf{y} - \boldsymbol{\alpha}'_{\boldsymbol{\tau}} (1, \bar{\mathbf{w}}')' = 0,$$

with $\boldsymbol{\beta}'_{\boldsymbol{\tau}} \mathbf{u} = (\mathbf{b}'_{\boldsymbol{\tau};\mathbf{w}_0} - \mathbf{w}'_0 \dot{\mathbf{b}}'_{\boldsymbol{\tau};\mathbf{w}_0}) \mathbf{u} = \mathbf{b}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{u} = 1$, since $\dot{\mathbf{b}}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{u} = -\dot{\mathbf{c}}'_{\boldsymbol{\tau};\mathbf{w}_0} \boldsymbol{\Gamma}'_{\mathbf{u}} \mathbf{u} = 0$. Comparing with (1.3), this suggests a local linear approach based on weighted quantile hyperplanes (in the mp-dimensional regressor-response space associated with the augmented regressor $\bar{\mathbf{x}}$, that is, the $(\bar{\mathbf{w}}', \mathbf{y}')'$-space), yielding weighted empirical quantile hyperplanes with equations

$$(4.5) \qquad\qquad \boldsymbol{\beta}^{(n)\prime}_{\boldsymbol{\tau};\omega} \mathbf{y} - \boldsymbol{\alpha}^{(n)\prime}_{\boldsymbol{\tau};\omega} (1, \bar{\mathbf{w}}')' = 0,$$

based on the same sequences of weights $\omega_i^{(n)} := \omega_{\mathbf{w}_0}^{(n)}(\mathbf{W}_i)$, $i = 1, \dots, n$, as in Section 4.1. Interpretation of the results, however, is easier from (4.3) than from (4.4). The left-hand side of (4.3) indeed splits naturally into two parts of independent interest: (i) the first one, made of the first two terms, yields the equation of the $\mathbf{w}_0$-conditional $\boldsymbol{\tau}$-quantile hyperplane of $\mathbf{Y}$, hence provides the required information for constructing the empirical $\mathbf{w}_0$-cuts, whereas (ii) the second part (the third term) provides the linear (linear with respect to $(\mathbf{w} - \mathbf{w}_0)$; actually, bilinear in $(\mathbf{w} - \mathbf{w}_0)$ and $\boldsymbol{\Gamma}'_{\mathbf{u}} \mathbf{y})$ correction required for a small perturbation $(\mathbf{w} - \mathbf{w}_0)$ of the value of the conditioning variable. Therefore, the important quantities to be recovered from $\boldsymbol{\alpha}^{(n)}_{\boldsymbol{\tau};\omega}$ and $\boldsymbol{\beta}^{(n)}_{\boldsymbol{\tau};\omega}$ are estimations of these two parts, which are easily obtained by

(i) letting $\mathbf{w} = \mathbf{w}_0$ in (4.5), which yields the equation

$$\boldsymbol{\beta}_{\boldsymbol{\tau};\omega}^{(n)\prime}\mathbf{y} - \boldsymbol{\alpha}_{\boldsymbol{\tau};\omega}^{(n)\prime}(1, \mathbf{w}_0', (\mathbf{w}_0 \otimes \boldsymbol{\Gamma}_{\mathbf{u}}'\mathbf{y})')' = 0$$

of an empirical hyperplane providing an estimate of the two first terms in (4.3), namely, the $\mathbf{w}_0$-conditional $\boldsymbol{\tau}$-quantile hyperplane;

(ii) subtracting the latter equation from (4.5), which provides the bilinear correction term.

The bilinear nature of the local approximation in (ii) is easily explained by the fact that, in general, unless the $\mathbf{w}_0$-conditional and $\mathbf{w}$-conditional $\boldsymbol{\tau}$-quantile hyperplanes are parallel to each other, no higher-dimensional hyperplane can run through both (for instance, two mutually skew non-intersecting straight lines in $\mathbb{R}^3$ do not span a plane). Omitting the additional $\mathbf{W} \otimes (\boldsymbol{\Gamma}_{\mathbf{u}}'\mathbf{Y})$ regressors (in (i) above) may result in inconsistent estimators of the $\mathbf{w}_0$-conditional $\boldsymbol{\tau}$-quantile hyperplanes. The resulting regions in $\mathbb{R}^{m+p}$, are not polyhedral anymore, but delimited by ruled quadrics (hyperbolic paraboloids for $m = 2$ and $p = 1$), the intersections of which with the $\mathbf{w} = \mathbf{w}_0$ hyperplane yield polyhedral estimated $\mathbf{w}_0$-cuts; see Figure 1(b).

The local bilinear approach is more informative than the local constant one, and should be more reliable at boundary points; however, due to the presence of the regressors $\mathbf{W}$ and $\mathbf{W} \otimes (\boldsymbol{\Gamma}_{\mathbf{u}}'\mathbf{Y})$ in (4.5), it may suffer from a substantial increase of the covariate space dimension, hence of the number of free parameters (mp free parameters instead of m for the local constant method).

5. Asymptotics. Throughout this section, we fix $\mathbf{w}_0 \in \mathbb{R}^{p-1}$ and $\boldsymbol{\tau} = \tau\mathbf{u} \in (0,1) \times \mathcal{S}^{m-1}$, and we write, for notational simplicity, $Y_{\mathbf{u}} := \mathbf{u}'\mathbf{Y}$ and $\mathbf{Y}_{\mathbf{u}}^{\perp} := \boldsymbol{\Gamma}_{\mathbf{u}}'\mathbf{Y}$. Asymptotic results require some regularity assumptions on the density of the observations, the kernel, and the bandwidth.

ASSUMPTION (A1)(i) The n-tuple $(\mathbf{W}_i', \mathbf{Y}_i')'$, $i = 1, \ldots, n$ is an i.i.d. sample

from $(\mathbf{W}', \mathbf{Y}')'$. (ii) The density $\mathbf{w} \mapsto f^{\mathbf{W}}(\mathbf{w})$ of $\mathbf{W}$ is continuous and strictly positive at $\mathbf{w}_0$. (iii) For any $\mathbf{t} \in \mathbb{R}^{m-1}$, the density $s \mapsto f^{Y_{\mathbf{u}} | \mathbf{Y}_{\mathbf{u}}^{\perp} = \mathbf{t}, \mathbf{W} = \mathbf{w}}(s)$ of $Y_{\mathbf{u}}$ conditional on $\mathbf{Y}_{\mathbf{u}}^{\perp} = \mathbf{t}$ and $\mathbf{W} = \mathbf{w}$ is continuous with respect to s in a neighborhood of $a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{t}$, uniformly in $\mathbf{w}$ over a neighborhood of $\mathbf{w}_0$, and continuous with respect to $\mathbf{w}$ in a neigborhood of $\mathbf{w}_0$ for all s in a neighborhood of $a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{t}$. (iv) The density $f^{\mathbf{Y}_{\mathbf{u}}^{\perp} | \mathbf{W} = \mathbf{w}}(\mathbf{t})$ of $\mathbf{Y}_{\mathbf{u}}^{\perp}$ conditional on $\mathbf{W} = \mathbf{w}$ is continuous with respect to $\mathbf{w}$ over a neighborhood of $\mathbf{w}_0$, except perhaps for a set of $\mathbf{t}$ of $(f^{\mathbf{Y}_{\mathbf{u}}^{\perp}}\text{-})$measure zero. (v) The $m \times m$ matrix

(5.1)

$$\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} := \int_{\mathbb{R}^{m-1}} \begin{pmatrix} 1 & \mathbf{t}' \\ \mathbf{t} & \mathbf{t}\mathbf{t}' \end{pmatrix} f^{Y_{\mathbf{u}} | \mathbf{Y}_{\mathbf{u}}^{\perp} = \mathbf{t}, \mathbf{W} = \mathbf{w}_0}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{t}) \, f^{\mathbf{Y}_{\mathbf{u}}^{\perp} | \mathbf{W} = \mathbf{w}_0}(\mathbf{t}) \, \mathrm{d}\mathbf{t}$$

is finite and positive definite.

ASSUMPTION (A2) (i) The kernel function K is a bounded density over $\mathbb{R}^{p-1}$ that has a compact support (S_K, say). (ii) $\int_{\mathbb{R}^{p-1}} \mathbf{w} K(\mathbf{w}) \, \mathrm{d}\mathbf{w} = \mathbf{0}$ and $\boldsymbol{\mu}_2^K := \int_{\mathbb{R}^{p-1}} \mathbf{w}\mathbf{w}' K(\mathbf{w}) \, \mathrm{d}\mathbf{w}$ is finite and positive definite

ASSUMPTION (A3) The bandwidth h_n is such that $\lim_{n\to\infty} h_n = 0$ and $\lim_{n\to\infty} n h_n^{p-1} = \infty$.

The conditions we are imposing in Assumption (A1) are quite mild. For example, (A1)(ii) is the same as Condition (A)(iii) in [9] and (A1)(i) in [17]; (A1)(iii)-(v) are similar to Condition (A)(i, iv) in [9] and Condition (A1)(ii) in [17], where the existence and positive-definiteness ensure the invertibility of $\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0}$ in Theorem 5.1.

Assumptions (A2) and (A3) on the kernel function and the bandwidth also are quite standard in the nonparametric literature. For example, any compactly supported symmetric density function with second-order moments satisfies Assumption (A2). The compact support of K in Assumption (A2)

is only a technical assumption to simplify the proof of theorems. In practice, Gaussian kernels can be considered; indeed, at the cost of more involved proof, the compact support assumption in Theorems 5.1-5.2 can be replaced with the assumption that both $C_0^K := \int_{\mathbb{R}^{p-1}} K^2(\mathbf{w}) \, d\mathbf{w}$ and $C_2^K := \int_{\mathbb{R}^{p-1}} \mathbf{w}\mathbf{w}' K^2(\mathbf{w}) \, d\mathbf{w}$ are finite. As for Assumption (A3), it is the usual one in the i.i.d. setting.

Let $\boldsymbol{\mathcal{X}}_{\mathbf{u}}^c := (1, \mathbf{Y}_{\mathbf{u}}^{\perp\prime})'$ and $\boldsymbol{\mathcal{X}}_{\mathbf{u}}^\ell := (1, \mathbf{Y}_{\mathbf{u}}^{\perp\prime})' \otimes (1, (\mathbf{W} - \mathbf{w}_0)')'$, where the superscript c and ℓ stand for the local constant and local bilinear cases, respectively. For $(\mathbf{W}, \mathbf{Y}) = (\mathbf{W}_i, \mathbf{Y}_i)$, we use the notation $Y_{i\mathbf{u}}$, $\mathbf{Y}_{i\mathbf{u}}^\perp$, $\boldsymbol{\mathcal{X}}_{i\mathbf{u}}^c$, $\boldsymbol{\mathcal{X}}_{i\mathbf{u}}^\ell$, etc. in an obvious way.

Referring to (4.2) for the notation, the parameter of interest for the local constant case is $\boldsymbol{\theta}^c = \boldsymbol{\theta}_{\boldsymbol{\tau};\mathbf{w}_0}^c := (a_{\boldsymbol{\tau};\mathbf{w}_0}, \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}')'$, whereas, in the local bilinear case (see (4.3)), we rather have to estimate

$$(5.2) \qquad \boldsymbol{\theta}^\ell = \boldsymbol{\theta}_{\boldsymbol{\tau};\mathbf{w}_0}^\ell := \text{vec} \begin{pmatrix} a_{\boldsymbol{\tau};\mathbf{w}_0} & \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}' \\ \dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}_0} & \dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w}_0}' \end{pmatrix}.$$

The local constant and local bilinear methods described in the previous sections provide estimators of the form $\hat{\boldsymbol{\theta}}^{c(n)} := (\hat{a}, \hat{\mathbf{c}}')'$ and

$$(5.3) \qquad \hat{\boldsymbol{\theta}}^{\ell(n)} := \text{vec} \begin{pmatrix} \hat{a} & \hat{\mathbf{c}}' \\ \hat{\dot{\mathbf{a}}} & \hat{\dot{\mathbf{c}}}' \end{pmatrix}$$

(we should actually discriminate between $(\hat{a}, \hat{\mathbf{c}}') = (\hat{a}^c, \hat{\mathbf{c}}^{c\prime})$ and $(\hat{a}, \hat{\mathbf{c}}') = (\hat{a}^\ell, \hat{\mathbf{c}}^{\ell\prime})$, but will not do so in order to avoid making the notation too heavy); those estimators are defined as the corresponding minimizer $\boldsymbol{\theta}^r$ of

$$(5.4) \qquad \sum_{i=1}^n K_h(\mathbf{W}_i - \mathbf{w}_0)\rho_\tau(Y_{i\mathbf{u}} - \boldsymbol{\theta}^{r\prime}\boldsymbol{\mathcal{X}}_{i\mathbf{u}}^r), \qquad r = c, \ell.$$

The following result provides Bahadur representations for $\hat{\boldsymbol{\theta}}^{c(n)}$ and $\hat{\boldsymbol{\theta}}^{\ell(n)}$.

THEOREM 5.1.   (Bahadur representations) *Let Assumptions (A1), (A2)(i),* *and (A3) hold, assume that* $\mathbf{w} \mapsto (a_{\boldsymbol{\tau};\mathbf{w}}, \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}})'$ *is continuously differentiable at* $\mathbf{w}_0$, *and write* $\psi_\tau(y) := \tau - I[y < 0]$. *Then, as* $n \to \infty$,

$$\sqrt{nh_n^{p-1}}\, \mathbf{M}_h^r\big(\hat{\boldsymbol{\theta}}^{r(n)} - \boldsymbol{\theta}^r\big)$$

$$(5.5) \quad = \quad \frac{\boldsymbol{\eta}^r_{\boldsymbol{\tau};\mathbf{w}_0}}{\sqrt{nh_n^{p-1}}} \sum_{i=1}^{n} K\left(\frac{\mathbf{W}_i - \mathbf{w}_0}{h_n}\right) \psi_\tau(Z_{i\mathbf{u}}^r(\boldsymbol{\theta}))(\mathbf{M}_h^r)^{-1}\boldsymbol{\mathcal{X}}_{i\mathbf{u}}^r + o_{\mathrm{P}}(1),$$

where $Z_{i\mathbf{u}}^r(\boldsymbol{\vartheta}) := Y_{i\mathbf{u}} - \boldsymbol{\vartheta}' \boldsymbol{\mathcal{X}}_{i\mathbf{u}}^r$ $(r = c, \ell)$, $\mathbf{M}_h^c := \mathbf{I}_m$, $\mathbf{M}_h^\ell := \mathbf{I}_m \otimes \mathrm{diag}(1, h_n \mathbf{I}_{p-1})$,

(5.6)

$$\boldsymbol{\eta}^c_{\boldsymbol{\tau};\mathbf{w}_0} := \frac{1}{f^{\mathbf{W}}(\mathbf{w}_0)}\, \mathbf{G}^{-1}_{\boldsymbol{\tau};\mathbf{w}_0}, \;\; and\;\; \boldsymbol{\eta}^\ell_{\boldsymbol{\tau};\mathbf{w}_0} := \frac{1}{f^{\mathbf{W}}(\mathbf{w}_0)}\, \mathbf{G}^{-1}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathrm{diag}\big(1, (\boldsymbol{\mu}_2^K)^{-1}\big),$$

with $\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0}$ *defined in (5.1) (the result for the local constant case does not require (A2)(ii)).*

This result, along with Assumption (A4) below, entails the asymptotic normality of $\hat{\boldsymbol{\theta}}^{r(n)}$, $r = c, \ell$. That assumption deals with the existence, at $\mathbf{w} = \mathbf{w}_0$, of the second derivatives of $\mathbf{w} \mapsto (a_{\boldsymbol{\tau};\mathbf{w}}, \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}})'$. With $\mathbf{c}_{\boldsymbol{\tau};\mathbf{w}} =: (c_{\boldsymbol{\tau};\mathbf{w},1}, \ldots, c_{\boldsymbol{\tau};\mathbf{w},m-1})'$, denote by $\dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}}$ and $\dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},j}$ the $(p-1) \times 1$ vectors of first derivatives and by $\ddot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}}$ and $\ddot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},j}$ the $(p-1) \times (p-1)$ matrices of second derivatives (when they exist) of $\mathbf{w} \mapsto a_{\boldsymbol{\tau};\mathbf{w}}$ and $\mathbf{w} \mapsto c_{\boldsymbol{\tau};\mathbf{w},j}$, respectively (recall that $\dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}}$ and $\dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w}} = (\dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},1}, \ldots, \dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},m-1})'$ were already defined in page 14). Finally, write $\ddot{\mathbf{c}}'_{\boldsymbol{\tau};\mathbf{w}}$ for the $(p-1) \times (m-1)(p-1)$ matrix $(\ddot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},1}, \ldots, \ddot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},m-1})$.

ASSUMPTION (A4) (i) The function $\mathbf{w} \mapsto (a_{\boldsymbol{\tau};\mathbf{w}}, \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}})'$ is twice continuously differentiable at $\mathbf{w} = \mathbf{w}_0$, that is, $\ddot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}}$ and $\ddot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w}}$ exist in a neighborhood of $\mathbf{w}_0$ and are continuous with respect to $\mathbf{w}$ at $\mathbf{w}_0$. (ii) The function $\mathbf{w} \mapsto f^{\mathbf{W}}(\mathbf{w})$ is continuously differentiable at $\mathbf{w} = \mathbf{w}_0$, that is, the $(p-1) \times 1$ vector of first derivatives of $f^{\mathbf{W}}$, $\dot{f}^{\mathbf{W}}(\mathbf{w})$, exists in a neighborhood of $\mathbf{w}_0$ and is continuous with respect to $\mathbf{w}$ at $\mathbf{w}_0$.

The following matrices are involved in the asymptotic bias and variance expressions of the asymptotic normality result in Theorem 5.2 below. Define

$$(5.7)\ \ \mathbf{\Sigma}_{\mathbf{w}}^c := \tau(1-\tau)f^{\mathbf{W}}(\mathbf{w})\,C_0^K \boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}}^c \left[\int_{\mathbb{R}^{m-1}} f^{\mathbf{Y}_{\mathbf{u}}^\perp|\mathbf{W}=\mathbf{w}}(\mathbf{t})\begin{pmatrix} 1 & \mathbf{t}' \\ \mathbf{t} & \mathbf{t}\mathbf{t}' \end{pmatrix}d\mathbf{t}\right]\boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}}^c,$$

$$\mathbf{\Sigma}_{\mathbf{w}}^\ell := \tau(1-\tau)f^{\mathbf{W}}(\mathbf{w})\,\boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}}^\ell$$

$$(5.8)\qquad\qquad \times\left[\int_{\mathbb{R}^{m-1}} f^{\mathbf{Y}_{\mathbf{u}}^\perp|\mathbf{W}=\mathbf{w}}(\mathbf{t})\begin{pmatrix} 1 & \mathbf{t}' \\ \mathbf{t} & \mathbf{t}\mathbf{t}' \end{pmatrix}d\mathbf{t}\otimes\mathrm{diag}\,(C_0^K, C_2^K)\right]\boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}}^\ell,$$

and, for $r = c,\ell$,

$$(5.9)\quad \mathbf{B}_{\mathbf{w}}^r := f^{\mathbf{W}}(\mathbf{w})\,\boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}}^r$$

$$\times\int_{\mathbb{R}^{m-1}} f^{Y_{\mathbf{u}}|\mathbf{Y}_{\mathbf{u}}^\perp=\mathbf{t},\mathbf{W}=\mathbf{w}}(a_{\boldsymbol{\tau};\mathbf{w}}+\mathbf{c}_{\boldsymbol{\tau};\mathbf{w}}'\mathbf{t})f^{\mathbf{Y}_{\mathbf{u}}^\perp|\mathbf{W}=\mathbf{w}}(\mathbf{t})\begin{pmatrix} 1 \\ \mathbf{t} \end{pmatrix}\otimes\left[\mathbf{B}_{\mathbf{w};0}^r\begin{pmatrix} 1 \\ \mathbf{t} \end{pmatrix}\right]d\mathbf{t},$$

where (putting $\ddot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},0} := \ddot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}}$) $\mathbf{B}_{\mathbf{w};0}^c$ is the $1\times m$ matrix with jth entry

$$B_{\mathbf{w};0,j}^c := \mathrm{tr}\left[\left(\ddot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},j-1} + 2\frac{\dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},j-1}(\dot{f}^{\mathbf{W}}(\mathbf{w}))'}{f^{\mathbf{W}}(\mathbf{w})}\right)\boldsymbol{\mu}_2^K\right],\quad j=1,\ldots,m,$$

and $\mathbf{B}_{\mathbf{w};0}^\ell$ denotes the $p\times m$ matrix with (i,j)th entry

$$B_{\mathbf{w};0,ij}^\ell := \mathrm{tr}\left[\ddot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w},j-1}\int_{\mathbb{R}^{p-1}} w_{i-1}\mathbf{w}\mathbf{w}'K(\mathbf{w})\,d\mathbf{w}\right],\quad i=1,\ldots,p\ \ j=1,\ldots,m;$$

here, we wrote $\mathbf{w}=(w_1,w_2,\ldots,w_{p-1})'$, $w_0=1$. We then have

THEOREM 5.2.   (Asymptotic normality) *Let Assumptions (A1)-(A4) hold. Then, for $r=c,\ell$,*

$$(5.10)\qquad \sqrt{nh_n^{p-1}}\,\mathbf{M}_h^r\left(\hat{\boldsymbol{\theta}}^{r(n)} - \boldsymbol{\theta}^r - \frac{h^2}{2}\mathbf{B}_{\mathbf{w}_0}^r\right)\overset{\mathcal{L}}{\to}\mathcal{N}(\mathbf{0},\mathbf{\Sigma}_{\mathbf{w}_0}^r),$$

as $n\to\infty$, where $\overset{\mathcal{L}}{\to}$ denotes convergence in distribution (the result for the local bilinear case does not require (A4)(ii)).

Remark 1: The local bilinear fitting has the expression of bias that is independent of $\dot{f}^{\mathbf{W}}$. In contrast, the local constant fitting has a large bias

at the regions where the derivative of $f^{\mathbf{W}}$ is large, that is, it cannot adapt to highly-skewed designs (see [8, 10]). Another important advantage of local bilinear fitting over the local constant approach is its much better boundary behavior. This advantage often has been emphasized in the usual regression settings when the regressors take values on a compact subset of $\mathbb{R}^{p-1}$. For example, considering a univariate random regressor W ($p = 2$) with bounded support ([0, 1], say), it can be proved, using an argument similar to the one developed in the corresponding proof in [8], that asymptotic normality (with the same rate) still holds at boundary points of the form ch_n, where $c \in \mathbb{R}_0^+$, with asymptotic bias and variances of the same form as in the local bilinear ($r = \ell$) versions of (5.9) and (5.8), with $p = 2$, $\mathbf{w}_0$ replaced by $w_0 = 0^+$, and $\int_{\mathbb{R}^{p-1}}$ by $\int_{-c}^{\infty}$; see, for example, page 666 of [17].

Remark 2: In practice, we may be concerned with the estimation of the quantile regression functions at different $\boldsymbol{\tau}$'s simultaneously. Restricting to the estimation of $(\boldsymbol{\theta}'_{\boldsymbol{\tau}1;\mathbf{w}_0}, \boldsymbol{\theta}'_{\boldsymbol{\tau}2;\mathbf{w}_0})'$, with $\boldsymbol{\tau}_k \in (0,1) \times \mathcal{S}^{m-1}$, $k = 1, 2$, it can be shown by proceeding as in the proof of Theorem 5.2 that $(\hat{\boldsymbol{\theta}}'_{\boldsymbol{\tau}1;\mathbf{w}_0}, \hat{\boldsymbol{\theta}}'_{\boldsymbol{\tau}2;\mathbf{w}_0})'$ is asymptotically normal with a block-diagonal asymptotic covariance matrix, that is, $\hat{\boldsymbol{\theta}}_{\boldsymbol{\tau}1;\mathbf{w}_0}$ and $\hat{\boldsymbol{\theta}}_{\boldsymbol{\tau}2;\mathbf{w}_0}$ are asymptotically independent for $\boldsymbol{\tau}_1 \neq \boldsymbol{\tau}_2$.

6. Simulated and real data illustrations. This section illustrates the use of the proposed local quantile regions on simulated data (Section 6.1) and on real data (Section 6.2).

6.1. *Simulated data.* We first generated 999 points from the model $(Y_1, Y_2) = (W, W^2) + (1 + \frac{3}{2}(\sin(\frac{\pi}{2}W))^2)\boldsymbol{\varepsilon}$, where $W \sim U([-2, 2])$ is independent of the bivariate standard normal vector $\boldsymbol{\varepsilon}$, and plotted the $\tau = .2$ and $\tau = .4$ HPŠ regression quantile contours obtained by using the covariate vector $\mathbf{X} = (1, W)'$ (Figure 2(a)) or $\mathbf{X} = (1, W, W^2)'$ (Figure 2(b)). More precisely, these figures provide cuts of the HPŠ contours by hyperplanes orthogonal to the

w-axis at fixed w-values $-1.89, -1.83, -1.77, \ldots, 1.89$.

Clearly, the results are very poor : Figure 2(a) does not reveal the trend nor the heteroskedasticity pattern in the data. Although it is obtained by fitting the true regression function, Figure 2(b) does much better with the trend, but still fails to model the heteroskedasticity correctly. Instead of providing genuine conditional quantile/depth contours, the HPŠ methodology, as announced in the Introduction, produces some averaged (over the w-space) contours.

In contrast, the cuts obtained from the proposed local constant and local bilinear methods—that do not use any knowledge on the true regression function—exhibit a very good agreement with the population contours (see Figures 2(c)-(e) to which we refer for details); both trend and heteroskedascticity components are now appropriately recovered. Note that, compared to the local constant approach, the local bilinear one does better, as expected, close to the boundary of the regressor space (in particular, the local constant approach seems to miss the decay of the conditional scale when w converges to -2). Similar comments can be made for smaller sample sizes; see Figure 3, that is based on 499 data points.

The second example involves a homoskedastic setup and a heteroskedastic one. More specifically, we generated $n = 999$ points from the homoskedastic model $(Y_1, Y_2) = (W, W^2) + \boldsymbol{\varepsilon}$ and from the heteroskedastic model $(Y_1, Y_2) = (W, W^2) + (1 + W^2)\boldsymbol{\varepsilon}$, where $W \sim U([-2, 2])$ and $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, 1/4)^2$ are mutually independent. As above, cuts of the local constant and local bilinear ($\tau = 0.2$ and $\tau = 0.4$) quantile regions, associated with the values $w \in \{-1.89, -1.83, -1.77, \ldots, 1.89\}$, are provided in Figure 4. Parallel to the previous example, these cuts remarkably approximate their population counterparts. In particular, the inner regions mimic the trend faithfully even for quite extreme regressor values. Again, the local bilinear method seems

to provide a better boundary behavior than its local constant counterpart; the latter indeed seems to underestimate the conditional scale for extreme value of W.

6.2. *A real data example.* In order to illustrate the data-analytic power of the proposed method, we consider the "body girth measurement" dataset from [19], that was already investigated in HPŠ. The dataset consists of joint measurements of nine skeletal and twelve body girth dimensions, along with weight, height, and age, in a group of 247 young men and 260 young women. As in HPŠ, we discard the male observations, we restrict to the calf maximum girth (Y_1) and the thigh maximum girth (Y_2) for the response, and we use a single random regressor W (weight, height, age, or BMI). Figures 5 and 6 provide cuts—for the same w- and τ-values as in HPŠ—obtained from the proposed local constant and local bilinear approaches, respectively.

These cuts confirm most of the global analysis conducted in HPŠ and moreover reveal some interesting new features. For instance,

(a) for the dependence on weight, the local bilinear approach confirms the positive trend in location, the increase in dispersion, and the evolution of "principal directions" (as weight increases, the first "principal direction" rotates from horizontal to vertical), and it further indicates that high weights give rise to simultaneously large extreme values in Y_1 and Y_2. The differences, for low and high values of the covariate (weight), between the contours resulting from the local bilinear and local constant approaches illustrate the sensitivity of the latter to boundary effect.

(b) for the dependence on age, the local regression quantile regions, parallel to their global HPŠ counterparts, do indicate that the location and the first principal direction (along the main bisector) are constant over age. Still as in HPŠ, the local approaches confirm that the shapes
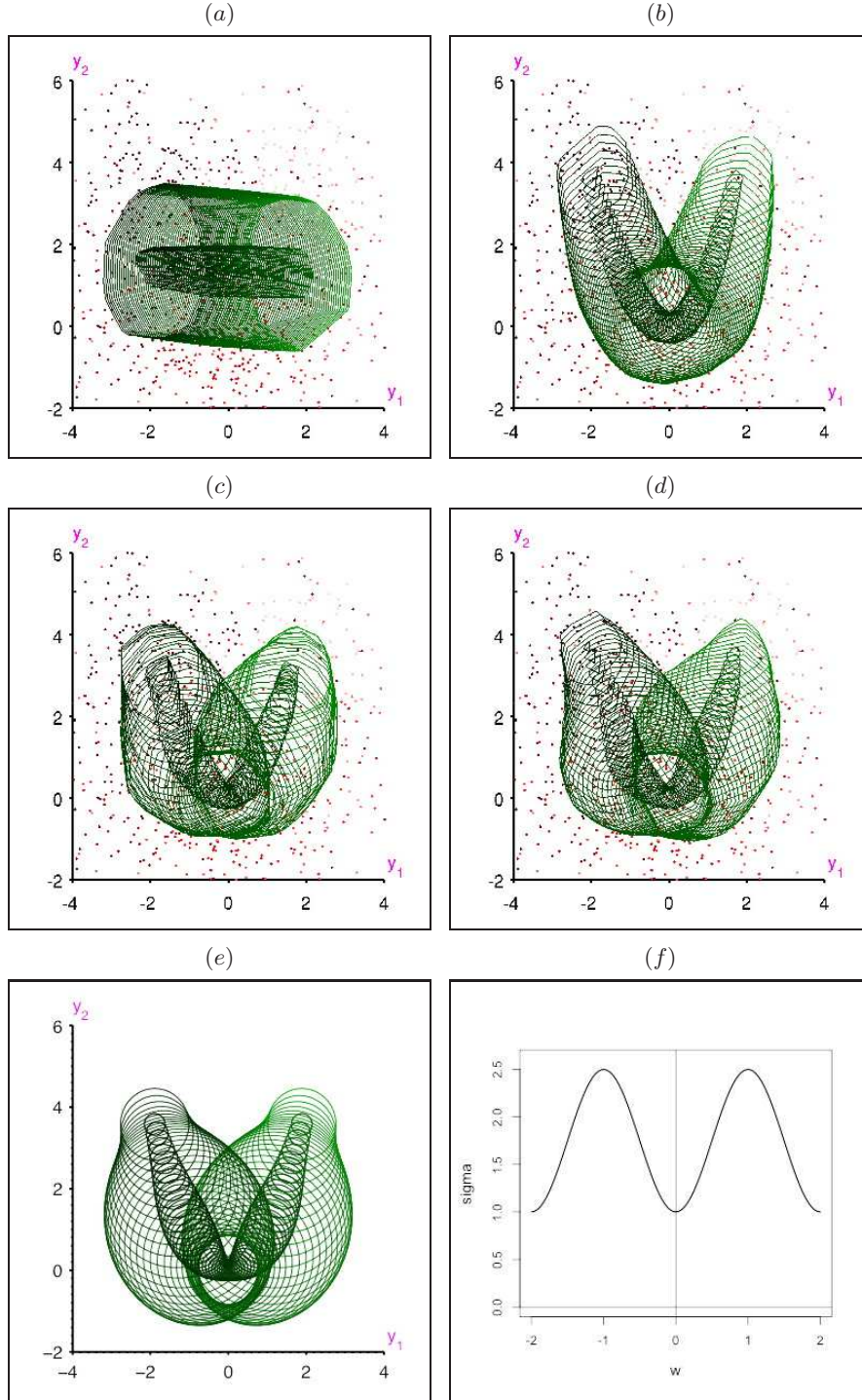
FIG 2. For 999 points following the model $(Y_1, Y_2) = (W, W^2) + (1 + \frac{3}{2}(\sin(\frac{\pi}{2}W))^2)\boldsymbol{\varepsilon}$, where $W \sim U([-2, 2])$ and $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, 1)^2$ are independent, the plots above show cuts, by hyperplanes orthogonal to the w-axis at fixed w-values $-1.89, -1.83, -1.77, \ldots, 1.89$, of (a) the HPŠ regression quantile regions with the single random regressor W, (b) the HPŠ regression quantile regions with the random regressors W and W^2, and (c-d) the proposed local constant and local bilinear regression quantile regions (in each case, $\tau = .2$ and $\tau = .4$ are considered). For the sake of comparison, the corresponding population (conditional) halfspace depth regions are provided in (e). The conditional scale function $w \mapsto 1 + \frac{3}{2}(\sin(\frac{\pi}{2}w))^2$ is plotted in (f). Local methods use a Gaussian kernel and bandwidth value $H = .37$, and 360 equispaced directions $\mathbf{u} \in \mathcal{S}^1$ were used to obtain results in (d).
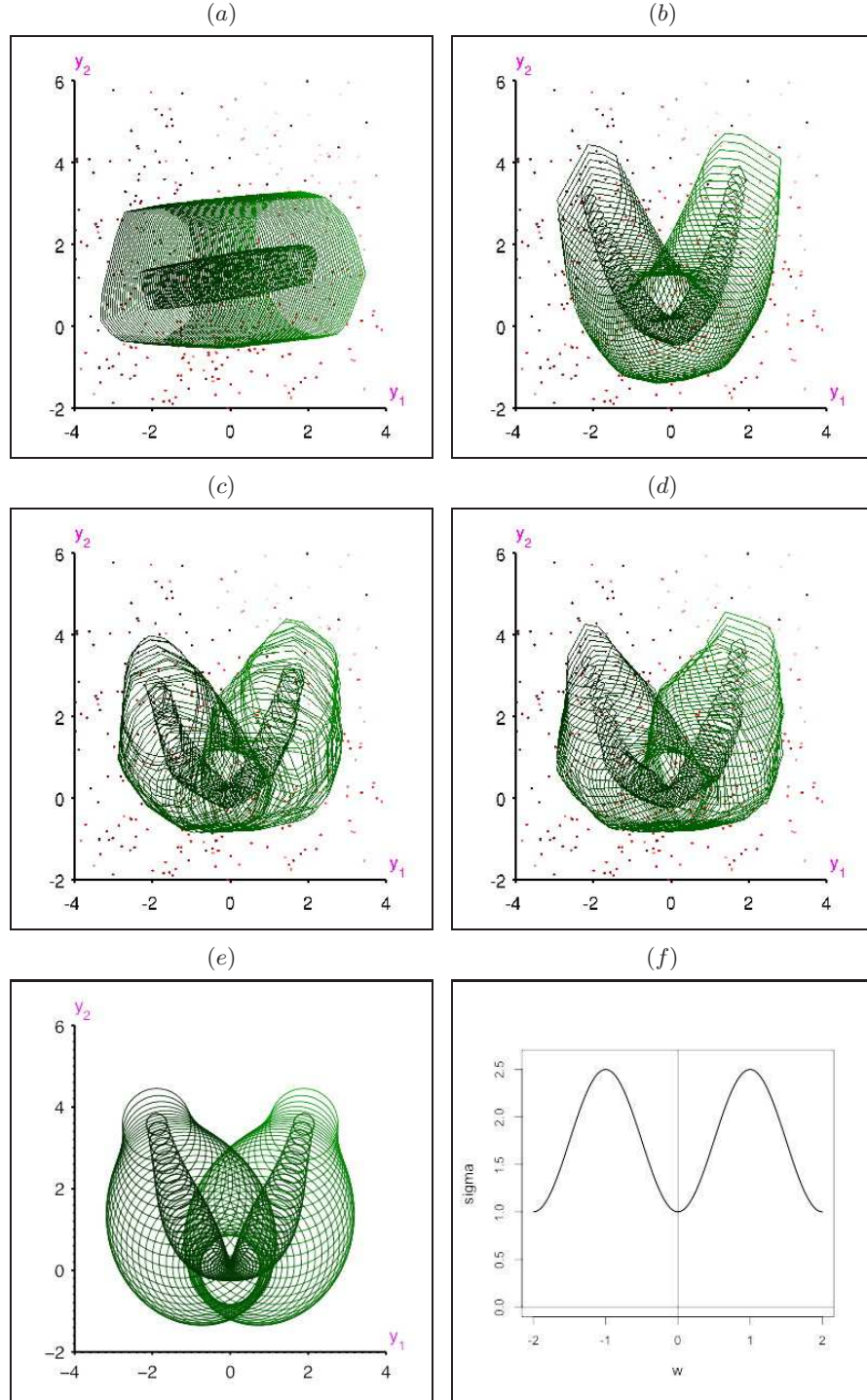
FIG 3. *For 499 points following the model* $(Y_1, Y_2) = (W, W^2) + (1 + \frac{3}{2}(\sin(\frac{\pi}{2}W))^2)\boldsymbol{\varepsilon}$, *where* $W \sim U([-2, 2])$ *and* $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, 1)^2$ *are independent, the plots above show cuts, by hyperplanes orthogonal to the w-axis at fixed w-values* $-1.89, -1.83, -1.77, \ldots, 1.89$, *of (a) the HPŠ regression quantile regions with the single random regressor* W, *(b) the HPŠ regression quantile regions with the random regressors* W *and* W^2, *and (c-d) the proposed local constant and local bilinear regression quantile regions (in each case,* $\tau = .2$ *and* $\tau = .4$ *are considered). For the sake of comparison, the corresponding population (conditional) halfspace depth regions are provided in (e). The conditional scale function* $w \mapsto 1 + \frac{3}{2}(\sin(\frac{\pi}{2}w))^2$ *is plotted in (f). Local methods use a Gaussian kernel and bandwidth value* $H = .41$, *and 360 equispaced directions* $\mathbf{u} \in \mathcal{S}^1$ *were used to obtain results in (d).*
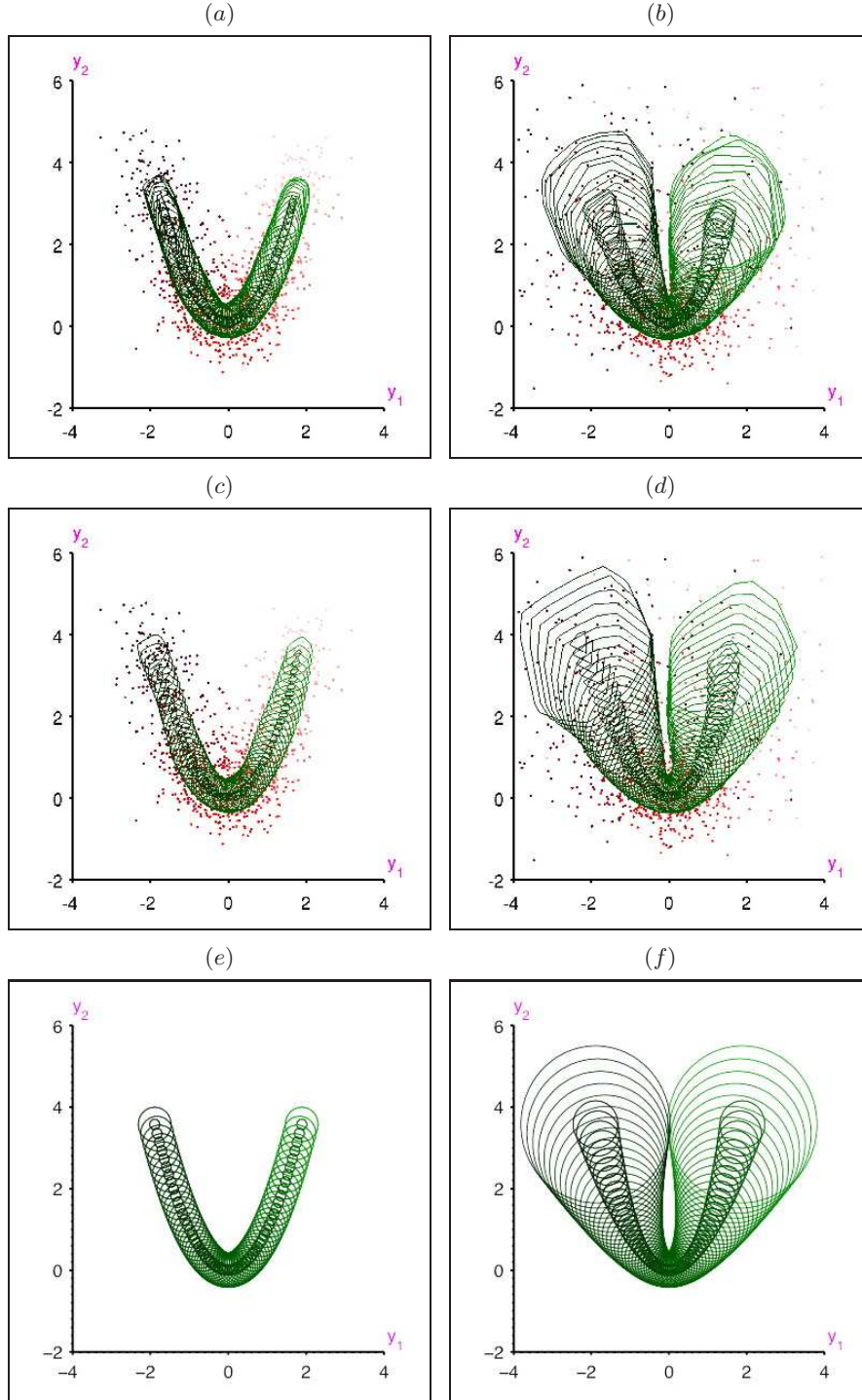
Fig 4. *Local multiple-output quantile regression with Gaussian kernel and ad-hoc bandwidth $H = .37$: cuts through $\mathbf{w} \in \{-1.89, -1.83, -1.77, \ldots, 1.89\}$ for $\tau = 0.2$ and $\tau = 0.4$ corresponding to $n = 999$ random points drawn from a homoskedastic model $(Y_1, Y_2) = (W, W^2) + \boldsymbol{\varepsilon}$ ((a), (c)) or a heteroskedastic model $(Y_1, Y_2) = (W, W^2) + (1 + W^2)\boldsymbol{\varepsilon}$ ((b), (d)), where $W \sim U([-2, 2])$ and $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, 1/4)^2$ are independent. Cuts are obtained either from the local constant method ((a), (b)) or the local bilinear one ((c), (d)). Color scaling of the points (resp., of the cuts) mimics their regressor values whose higher values are indicated by lighter red (resp., lighter green). For the sake of comparison, the population (conditional) halfspace depth regions are provided in (e) and (f).*
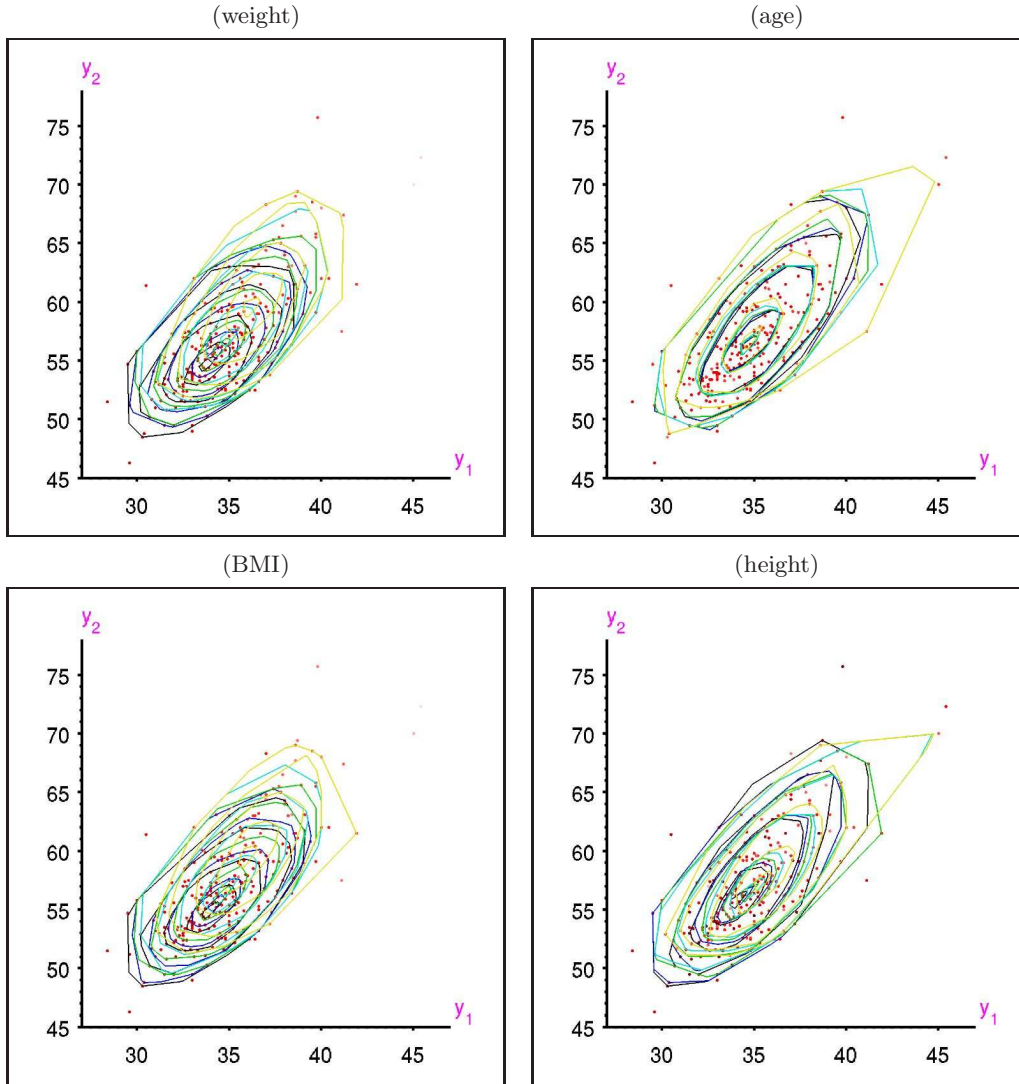
FIG 5. *Four empirical (local constant) regression quantile plots from the body girth measurements dataset (women subsample; see [19]). Throughout, the bivariate response $(Y_1, Y_2)'$ involves calf maximum girth (Y_1) and thigh maximum girth (Y_2), while a single random regressor is used: weight, age, BMI, or height. The plots are providing, for $\tau = .01, .03, .10, .25,$ and $.40$, the cuts of the local constant regression τ-quantile contours, at the empirical p-quantiles of the regressors, for p=.10 (black), .30 (blue), .50 (green), .70 (cyan) and .90 (yellow). Data points are shown in red (the lighter the red color, the higher the regressor value). The results are based on a Gaussian kernel and bandwidth $H = 1.3 \times \sigma_w n^{-1/5}$, where σ_w stands for the empirical standard deviation of the regressor (the corresponding cuts obtained from linear regression are provided in Figure 7 of HPŠ).*

(weight)   (age)

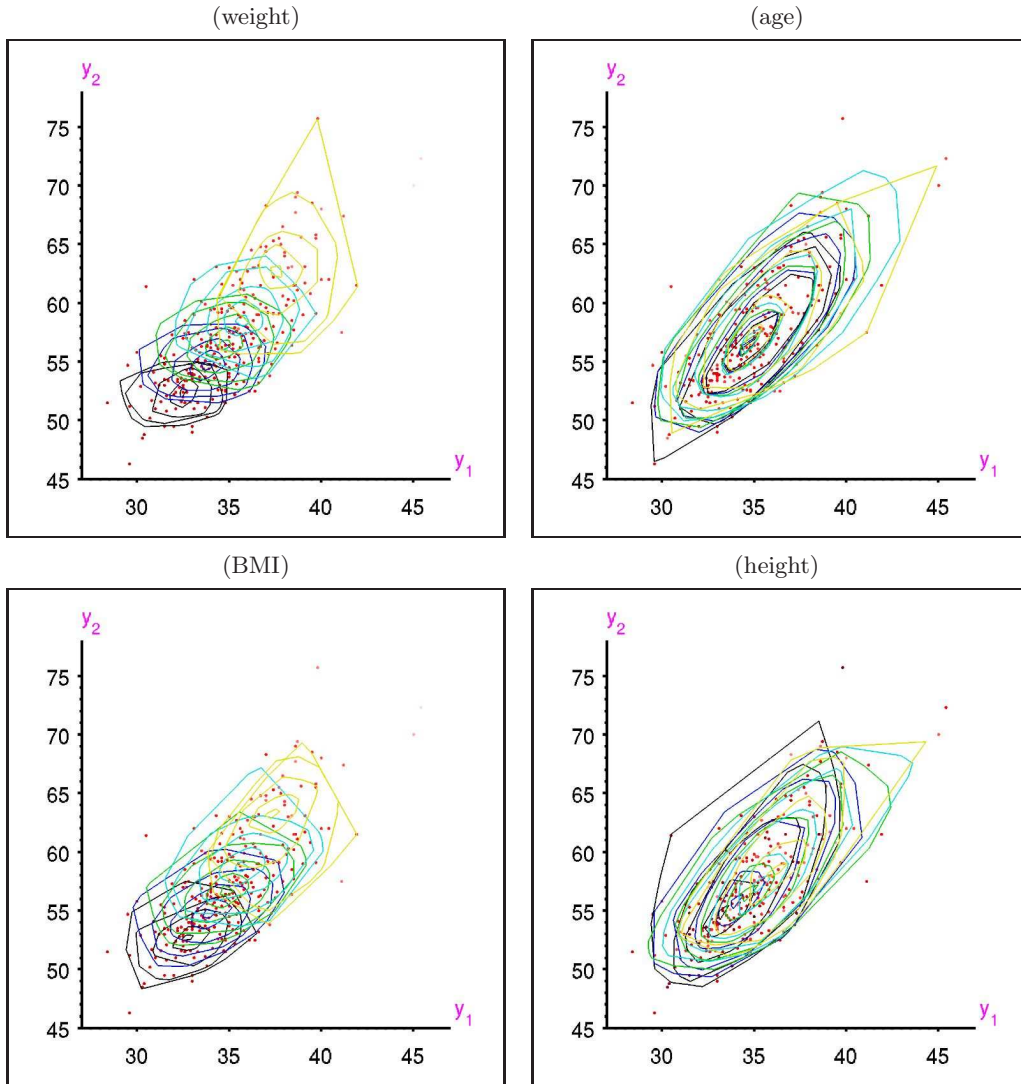(BMI)   (height)

FIG 6. *Same quantities as in Figure 5, here obtained from the local bilinear approach, with the same kernel and bandwidth as in Figure 5 (the computation was based on 360 equispaced directions $\mathbf{u} \in \mathcal{S}^1$).*

of outer contours vary quite significantly with age, indicating an increasing (with age) simultaneous variability of both calf and thigh girth largest values. Now, compared to HPŠ, the local bilinear approach further shows that *young* women present a large simultaneous variability of both calf and thigh girth *smallest* values;

(c) for the dependence on height, the local methods confirm the regression effect specific to inner contours. The local bilinear approach further shows that there is also a regression effect for outer contours, that, as height increases, get more widespread in the direction **u** corresponding to simultaneously large values of both responses).

This short application demonstrates how the local quantile regression analysis proposed here complements and refines the findings obtained from the global approach introduced in HPŠ by revealing the possible non-linear, heteroskedastic, skewness ... features of the distributions of **Y** conditional on **W** = **w**. We refer to [26] for a further application, in the context of bivariate growth charts.

APPENDIX A: PROOFS OF ASYMPTOTIC RESULTS

In this appendix, we prove Theorems 5.1 and 5.2. We will actually only prove the results in the local bilinear case (the proofs for the local constant case are entirely similar). The proofs rely on several lemmas, that require introducing some further notation.

Referring to (5.2)-(5.3), we let

$$\boldsymbol{\theta}^\ell = \mathrm{vec}\left(\begin{array}{cc} a_{\boldsymbol{\tau};\mathbf{w}_0} & \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \\ \dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}_0} & \dot{\mathbf{c}}'_{\boldsymbol{\tau};\mathbf{w}_0} \end{array}\right) =: \mathrm{vec}\left(\begin{array}{c} \boldsymbol{\varpi}'_{\mathbf{w}_0} \\ \dot{\boldsymbol{\varpi}}'_{\mathbf{w}_0} \end{array}\right)$$

and

$$\hat{\boldsymbol{\theta}}^{\ell(n)} = \mathrm{vec}\left(\begin{array}{cc} \hat{a} & \hat{\mathbf{c}}' \\ \hat{\dot{\mathbf{a}}} & \hat{\dot{\mathbf{c}}}' \end{array}\right) =: \mathrm{vec}\left(\begin{array}{c} \widehat{\boldsymbol{\varpi}}'_{\mathbf{w}_0} \\ \widehat{\dot{\boldsymbol{\varpi}}}'_{\mathbf{w}_0} \end{array}\right).$$

Denote by $\boldsymbol{\varpi}_1 = (a_1, \mathbf{c}_1')'$ and $\widetilde{\boldsymbol{\varpi}}_1 = (\tilde{a}_1, \tilde{\mathbf{c}}_1')'$ two arbitrary vectors of $\mathbb{R}^m$, by $\boldsymbol{\varpi}_2 = (\mathbf{a}_2, \mathbf{c}_2')'$ and $\widetilde{\boldsymbol{\varpi}}_2 = (\tilde{\mathbf{a}}_2, \tilde{\mathbf{c}}_2')'$ two arbitrary $m \times (p-1)$ matrices. Letting $H_n := \sqrt{nh_n^{p-1}}$, write then

$$\boldsymbol{\varphi} := H_n \, \mathbf{M}_h^\ell \, \text{vec}\left(\begin{array}{c} (\boldsymbol{\varpi}_1 - \boldsymbol{\varpi}_{\mathbf{w}_0})' \\ (\boldsymbol{\varpi}_2 - \dot{\boldsymbol{\varpi}}_{\mathbf{w}_0})' \end{array} \right), \qquad \tilde{\boldsymbol{\varphi}} := H_n \, \mathbf{M}_h^\ell \, \text{vec}\left(\begin{array}{c} (\widetilde{\boldsymbol{\varpi}}_1 - \boldsymbol{\varpi}_{\mathbf{w}_0})' \\ (\widetilde{\boldsymbol{\varpi}}_2 - \dot{\boldsymbol{\varpi}}_{\mathbf{w}_0})' \end{array} \right),$$

and

(A.1)
$$\boldsymbol{\varphi}^{(n)} := H_n \, \mathbf{M}_h^\ell \, \text{vec}\left(\begin{array}{c} (\widehat{\boldsymbol{\varpi}}_{\mathbf{w}_0} - \boldsymbol{\varpi}_{\mathbf{w}_0})' \\ (\widehat{\dot{\boldsymbol{\varpi}}}_{\mathbf{w}_0} - \dot{\boldsymbol{\varpi}}_{\mathbf{w}_0})' \end{array} \right),$$

and note that

$$\boldsymbol{\varphi}^{(n)} = \sqrt{nh_n^{p-1}} \, \mathbf{M}_h^\ell \left(\hat{\boldsymbol{\theta}}^{\ell(n)} - \boldsymbol{\theta}^\ell \right).$$

Put $\mathbf{W}_{hi} := (\mathbf{W}_i - \mathbf{w}_0)/h_n$, $K_{hi} := K(\mathbf{W}_{hi})$ and

$$\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell := (\mathbf{M}_h^\ell)^{-1} \boldsymbol{\mathcal{X}}_{i\mathbf{u}}^\ell = (1, \mathbf{Y}_{i\mathbf{u}}^{\perp\prime})' \otimes (1, \mathbf{W}_{hi}')'.$$

Let $Z_{i\mathbf{u}}^\ell = Z_{i\mathbf{u}}^\ell(\boldsymbol{\theta}^\ell) := Y_{i\mathbf{u}} - \boldsymbol{\theta}^{\ell\prime} \boldsymbol{\mathcal{X}}_{i\mathbf{u}}^\ell$ as in Theorem 5.1, and define

$$T_{ni} := h_n \dot{\mathbf{a}}_{\boldsymbol{\tau};\mathbf{w}_0}' \mathbf{W}_{hi} + h_n (\text{vec} \, \dot{\mathbf{c}}_{\boldsymbol{\tau};\mathbf{w}_0})' (\mathbf{Y}_{i\mathbf{u}}^\perp \otimes \mathbf{W}_{hi}),$$

$$Z_{ni}^*(\boldsymbol{\varphi}) := Z_{i\mathbf{u}}^\ell - H_n^{-1} \boldsymbol{\varphi}' \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell, \quad U_{ni} = U_{ni}(\boldsymbol{\varphi}) := T_{ni} + H_n^{-1} \boldsymbol{\varphi}' \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell$$

(note that these latter two quantities depend on the choice of $\boldsymbol{\varpi}_1$ and $\boldsymbol{\varpi}_2$). The following properties will be useful in the sequel:

(A.2)
$$Z_{i\mathbf{u}}^\ell = Y_{i\mathbf{u}} - (a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}' \mathbf{Y}_{i\mathbf{u}}^\perp) - T_{ni},$$

(A.3)
$$Z_{ni}^*(\boldsymbol{\varphi}) = Y_{i\mathbf{u}} - (a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}' \mathbf{Y}_{i\mathbf{u}}^\perp) - U_{ni}(\boldsymbol{\varphi}) = Y_{i\mathbf{u}} - \left(\text{vec}(\boldsymbol{\varpi}_1, \boldsymbol{\varpi}_2)' \right)' \boldsymbol{\mathcal{X}}_{i\mathbf{u}}^\ell.$$

Let C be a generic constant whose value may vary from line to line. Since K is a bounded density with a bounded support, we have, whenever $K_{hi} > 0$,

(A.4)
$$\|\mathbf{W}_{hi}\| \leq C \quad \text{and} \quad \|\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell\| \leq C(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|),$$

and, when moreover $\|\boldsymbol{\varphi}\| \leq M$,

$$(A.5) \quad |T_{ni}| \leq C h_n (1 + \|\mathbf{Y}_{i\mathbf{u}}^{\perp}\|) \text{ and } |U_{ni}| \leq C(h_n + H_n^{-1})(1 + \|\mathbf{Y}_{i\mathbf{u}}^{\perp}\|).$$

It follows from the definition of $\hat{\boldsymbol{\theta}}^{\ell(n)}$ as the argmin of (5.4) that

$$(A.6) \quad \boldsymbol{\varphi}^{(n)} = \operatorname{argmin}_{\boldsymbol{\varphi} \in \mathbb{R}^{mp}} \sum_{i=1}^{n} K_{hi} \rho_{\tau}(Z_{ni}^{*}(\boldsymbol{\varphi})).$$

Recalling that $\psi_{\tau}(y) := \tau - I[y < 0]$, define

$$(A.7) \quad \mathbf{V}_n(\boldsymbol{\varphi}) := H_n^{-1} \sum_{i=1}^{n} K_{hi} \psi_{\tau}(Z_{ni}^{*}(\boldsymbol{\varphi})) \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell}.$$

In order to prove Theorem 5.1, we need the following lemma.

LEMMA A.1. *Let* $\mathbf{V}_n(\cdot) : \mathbb{R}^{mp} \to \mathbb{R}^{mp}$ *be a sequence of functions that satisfies the following two properties:*

(i) for all $\lambda \geq 1$ and all $\boldsymbol{\psi} \in \mathbb{R}^{mp}$, $-\boldsymbol{\psi}' \mathbf{V}_n(\lambda \boldsymbol{\psi}) \geq -\boldsymbol{\psi}' \mathbf{V}_n(\boldsymbol{\psi})$ a.s.;

(ii) there exist a $p \times p$ positive definite matrix $\mathbf{D}$ and a sequence of mp-dimensional random vectors $\mathbf{A}_n$ satisfying $\|\mathbf{A}_n\| = O_{\mathrm{P}}(1)$ such that, for all $M > 0$, $\sup_{\|\boldsymbol{\psi}\| \leq M} \|\mathbf{V}_n(\boldsymbol{\psi}) + (\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D})\boldsymbol{\psi} - \mathbf{A}_n\| = o_{\mathrm{P}}(1)$, where $\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0}$ is given in (5.1).

Then, if $\boldsymbol{\psi}_n$ is such that $\|\mathbf{V}_n(\boldsymbol{\psi}_n)\| = o_{\mathrm{P}}(1)$, it holds that $\|\boldsymbol{\psi}_n\| = O_{\mathrm{P}}(1)$ and

$$(A.8) \quad \boldsymbol{\psi}_n = (\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D})^{-1} \mathbf{A}_n + o_{\mathrm{P}}(1).$$

Proof. The proof follows along the same lines as in page 809 of [23]; details are left to the reader. □

The proof of Theorem 5.1 then consists in checking that the assumptions of Lemma A.1 hold for $\mathbf{V}_n$ defined in (A.7). To do this, we will make use of the next lemma.

LEMMA A.2.   *Under Assumptions (A1)-(A3),*

$$
\begin{aligned}
\mathrm{E}\Big[K_{hi}|\psi_\tau(Z_{ni}^*(\boldsymbol{\varphi})) - \psi_\tau(Z_{ni}^*(\tilde{\boldsymbol{\varphi}}))|\Big] &\leq C\mathrm{E}\Big[K_{hi}I[|Z_{ni}^*(\tilde{\boldsymbol{\varphi}})| < CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|]\Big] \\
&\leq Ch_n^{p-1}H_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}\Big[K_{hi}^2|\psi_\tau(Z_{ni}^*(\boldsymbol{\varphi})) - \psi_\tau(Z_{ni}^*(\tilde{\boldsymbol{\varphi}}))|^2\Big] &\leq C\mathrm{E}\Big[K_{hi}^2I[|Z_{ni}^*(\tilde{\boldsymbol{\varphi}})| < CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|]\Big] \\
&\leq Ch_n^{p-1}H_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|.
\end{aligned}
$$

for any $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$ such that $\max(\|\boldsymbol{\varphi}\|, \|\tilde{\boldsymbol{\varphi}}\|) \leq M$, and n large enough.

Proof. The claim, in this lemma, is similar to that of Lemma A.3 in [17], which essentially follows from the same argument as in the time series case (cf. [25]). However, the details of the proof here are quite different.

It follows from (A.4) that

$$
\begin{aligned}
K_{hi}|\psi_\tau(Z_{ni}^*(\boldsymbol{\varphi})) - \psi_\tau(Z_{ni}^*(\tilde{\boldsymbol{\varphi}}))| &= K_{hi}\,|I[Z_{ni}^*(\boldsymbol{\varphi}) < 0] - I[Z_{ni}^*(\tilde{\boldsymbol{\varphi}}) < 0]| \\
&= K_{hi}\,|I[Z_{ni}^*(\tilde{\boldsymbol{\varphi}}) < H_n^{-1}(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}})'\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell] - I[Z_{ni}^*(\tilde{\boldsymbol{\varphi}}) < 0]| \\
&\leq K_{hi}\,I[|Z_{ni}^*(\tilde{\boldsymbol{\varphi}})| < CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|)].
\end{aligned}
$$

Hence, from (A.3) and the mean value theorem, we obtain

$$
\begin{aligned}
&\mathrm{E}\Big[K_{hi}|\psi_\tau(Z_{ni}^*(\boldsymbol{\varphi})) - \psi_\tau(Z_{ni}^*(\tilde{\boldsymbol{\varphi}}))|\Big] \\
&\leq \mathrm{E}\Big[K_{hi}I[|Z_{ni}^*(\tilde{\boldsymbol{\varphi}})| < CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|)]\Big] \\
&= \mathrm{E}\Big[K_{hi}\mathrm{P}[\,|Z_{ni}^*(\tilde{\boldsymbol{\varphi}})| < CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|)\,|\,\mathbf{Y}_{i\mathbf{u}}^\perp, \mathbf{W}_i]\Big] \\
&= \mathrm{E}\Big[K_{hi}F^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}\big(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}'\mathbf{Y}_{i\mathbf{u}}^\perp + U_{ni}(\tilde{\boldsymbol{\varphi}}) + CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|)\big)\Big] \\
&\quad - \mathrm{E}\Big[K_{hi}F^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}\big(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}'\mathbf{Y}_{i\mathbf{u}}^\perp + U_{ni}(\tilde{\boldsymbol{\varphi}}) - CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|)\big)\Big] \\
&\leq \mathrm{E}\Big[K_{hi}(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|)f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}\big(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}'\mathbf{Y}_{i\mathbf{u}}^\perp + U_{ni}(\tilde{\boldsymbol{\varphi}}) + \lambda CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|(1 + \|\mathbf{Y}_{i\mathbf{u}}^\perp\|)\big)\Big] \\
&\quad \times 2CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|,
\end{aligned}
$$

for some $\lambda \in (-1, 1)$. Assumptions (A1)-(A3), together with (A.5), therefore yield that, for $\boldsymbol{\varphi}$, $\tilde{\boldsymbol{\varphi}} \in \{\boldsymbol{\varphi} : \|\boldsymbol{\varphi}\| \leq M\}$ and n large enough,

$$\mathrm{E}\Big[K_{hi}\,|\psi_\tau(Z_{ni}^*(\boldsymbol{\varphi})) - \psi_\tau(Z_{ni}^*(\tilde{\boldsymbol{\varphi}}))|\Big]$$

$$\leq\ CH_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|\mathrm{E}\Big[K_{hi}\int_{\mathbb{R}^{m-1}}(1 + \|\mathbf{t}\|)f^{Y_{\mathbf{u}}|(\mathbf{Y}_{\mathbf{u}}^\perp = \mathbf{t}, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{t})f^{\mathbf{Y}_{\mathbf{u}}^\perp|\mathbf{W}}(\mathbf{t})\,\mathrm{d}\mathbf{t}\Big]$$

$$=\ Ch_n^{p-1}H_n^{-1}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}\|f^{\mathbf{W}}(\mathbf{w}_0)$$

$$\times \int_{\mathbb{R}^{m-1}}(1 + \|\mathbf{t}\|)f^{Y_{\mathbf{u}}|(\mathbf{Y}_{\mathbf{u}}^\perp = \mathbf{t}, \mathbf{W} = \mathbf{w}_0)}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{t})f^{\mathbf{Y}_{\mathbf{u}}^\perp|\mathbf{W} = \mathbf{w}_0}(\mathbf{t})\,\mathrm{d}\mathbf{t},$$

which proves the first inequality of Lemma A.2. The second one can be proved similarly. $\qquad\square$

LEMMA A.3.   *Under Assumptions (A1)-(A3), we have that, as $n \to \infty$,*

$$(\mathrm{A}.9) \qquad \sup_{\|\boldsymbol{\varphi}\| \leq M}\|\mathbf{V}_n(\boldsymbol{\varphi}) - \mathbf{V}_n(\mathbf{0}) - \mathrm{E}[\mathbf{V}_n(\boldsymbol{\varphi}) - \mathbf{V}_n(\mathbf{0})]\| = o_\mathrm{P}(1).$$

Proof. The proof of this lemma is quite similar, in view of Lemma A.2, to that of Lemma A.4 in [17]. Details are therefore omitted. $\qquad\square$

LEMMA A.4.   *Under Assumptions (A1)-(A3), we have that, as $n \to \infty$,*

$$(\mathrm{A}.10) \qquad \sup_{\|\boldsymbol{\varphi}\| \leq M}\|\mathrm{E}[\mathbf{V}_n(\boldsymbol{\varphi}) - \mathbf{V}_n(\mathbf{0})] + (\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D})\boldsymbol{\varphi}\| = o(1),$$

where $\mathbf{D} = f^{\mathbf{W}}(\mathbf{w}_0)\,\mathrm{diag}(1, \boldsymbol{\mu}_2^K)$.

Proof. Note that

$$(\mathrm{A}.11) \qquad \mathbf{V}_n(\boldsymbol{\varphi}) - \mathbf{V}_n(\mathbf{0}) = H_n^{-1}\sum_{i=1}^n K_{hi}[\psi_\tau(Z_{ni}^*(\boldsymbol{\varphi})) - \psi_\tau(Z_{i\mathbf{u}}^\ell)]\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell.$$

It follows from (A.2)-(A.3) that

$$\mathrm{E}[\mathbf{V}_n(\boldsymbol{\varphi}) - \mathbf{V}_n(\mathbf{0})] = nH_n^{-1}\mathrm{E}\Big[K_{hi}(I[Z_{i\mathbf{u}}^\ell < 0] - I[Z_{ni}^*(\boldsymbol{\varphi}) < 0])\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell\Big]$$

$$= H_nh_n^{-(p-1)}\mathrm{E}\Big[K_{hi}(F^{Y_{\mathbf{u}}|(\mathbf{Y}_{\mathbf{u}}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{Y}_{i\mathbf{u}}^\perp + T_{ni})$$

$$-F^{Y_{\mathbf{u}}|(\mathbf{Y}_{\mathbf{u}}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0}\mathbf{Y}_{i\mathbf{u}}^\perp + U_{ni}))\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell\Big].$$

Then, similar to the proof of Lemma A.2, by the mean value theorem, since $U_{ni} - T_{ni} = H_n^{-1} \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime} \boldsymbol{\varphi}$, there exists $\xi \in (0,1)$ such that

$$\sup_{\|\boldsymbol{\varphi}\| \leq M} \|\mathrm{E}[\mathbf{V}_n(\boldsymbol{\varphi}) - \mathbf{V}_n(\mathbf{0})] + (\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D})\boldsymbol{\varphi}\|$$

$$= \sup_{\|\boldsymbol{\varphi}\| \leq M} \|(\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D})\boldsymbol{\varphi}$$
$$- h_n^{-(p-1)} \mathrm{E}\big[K_{hi} f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{Y}_{i\mathbf{u}}^\perp + T_{ni} + \xi H_n^{-1} \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime} \boldsymbol{\varphi}) \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime} \boldsymbol{\varphi}\big]\|$$

$$= \sup_{\|\boldsymbol{\varphi}\| \leq M} \|\{(\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D}) - h_n^{-(p-1)} \mathrm{E}\big[K_{hi} f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{Y}_{i\mathbf{u}}^\perp) \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime}\big]\} \boldsymbol{\varphi}$$
$$- h_n^{-(p-1)} \mathrm{E}\big[K_{hi}(f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{Y}_{i\mathbf{u}}^\perp + T_{ni} + \xi H_n^{-1} \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime} \boldsymbol{\varphi}$$
$$- f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{Y}_{i\mathbf{u}}^\perp)) \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime} \boldsymbol{\varphi}\big]\|$$

$$\leq C \|(\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D}) - h_n^{-(p-1)} \mathrm{E}\big[K_{hi} f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{Y}_{i\mathbf{u}}^\perp) \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime}\big]\|$$
$$+ C \sup_{\|\boldsymbol{\varphi}\| \leq M} h_n^{-(p-1)} \mathrm{E}\big[K_{hi}|f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{Y}_{i\mathbf{u}}^\perp + T_{ni} + \xi H_n^{-1} \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime} \boldsymbol{\varphi})$$
$$- f^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp, \mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}_0} \mathbf{Y}_{i\mathbf{u}}^\perp)| \|\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell \boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^{\ell\prime}\|\big]$$

$$= o(1),$$

where we used Assumptions (A1) and (A2), together with (A.5). $\qquad \square$

LEMMA A.5.   *Let Assumptions (A2) and (A3) hold. Then the random vector $\boldsymbol{\varphi}^{(n)}$ defined in (A.1) satisfies $\|\mathbf{V}_n(\boldsymbol{\varphi}^{(n)})\| = o_\mathrm{P}(1)$.*

Proof. The proof follows from a similar argument to that of Lemma A.2 in Page 836 of [33]. $\qquad \square$

LEMMA A.6.   *Under Assumptions (A1)-(A3), for any $\mathbf{d} \in \mathbb{R}^{mp}$,*

$$\lim_{n\to\infty} \mathrm{E}\Big[\{\mathbf{d}'(\mathbf{V}_n(\mathbf{0}) - \mathrm{E}[\mathbf{V}_n(\mathbf{0})])\}^2\Big]$$
$$= \tau(1-\tau) f^\mathbf{W}(\mathbf{w}_0) \int_{\mathbb{R}^{p-1}} \int_{\mathbb{R}^{m-1}} ([(1, \mathbf{t}') \otimes (1, \mathbf{w}')]\mathbf{d})^2 f^{Y_\mathbf{u}^\perp|\mathbf{W}=\mathbf{w}_0}(\mathbf{t}) K^2(\mathbf{w}) \, \mathrm{d}\mathbf{t} \, \mathrm{d}\mathbf{w}.$$

Proof. Set $\tilde{v}_i = K_{hi}\psi_\tau(Z_{i\mathbf{u}}^\ell)\mathbf{d}'\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell = K_{hi}\psi_\tau(Z_{i\mathbf{u}}^\ell)[(1,\mathbf{Y}_{i\mathbf{u}}^{\perp\prime}) \otimes (1,\mathbf{W}_{hi}')]\mathbf{d}$. A simple calculation yields

$$(A.12) \quad \mathrm{E}\Big[\{\mathbf{d}'(\mathbf{V}_n(\mathbf{0}) - \mathrm{E}[\mathbf{V}_n(\mathbf{0})])\}^2\Big] = H_n^{-2}n\mathrm{Var}[\tilde{v}_1] = h_n^{-(p-1)}\mathrm{Var}[\tilde{v}_1].$$

Note that, for $k = 1, 2$,

$$\lim_{n\to\infty} h_n^{-(p-1)}\mathrm{E}\Big[K_{h1}^k I[Z_{1\mathbf{u}}^\ell < 0](\mathbf{d}'\boldsymbol{\mathcal{X}}_{h1\mathbf{u}}^\ell)^k\Big]$$
$$= \lim_{n\to\infty} h_n^{-(p-1)}\mathrm{E}\Big[K_{h1}^k F^{\mathbf{Y}_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp,\mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}'\mathbf{Y}_{1\mathbf{u}}^\perp + T_{n1})(\mathbf{d}'\boldsymbol{\mathcal{X}}_{h1\mathbf{u}}^\ell)^k\Big]$$
$$= \tau f^{\mathbf{W}}(\mathbf{w}_0)\int_{\mathbb{R}^{p-1}}\int_{\mathbb{R}^{m-1}} K^k(\mathbf{w})\left([(1,\mathbf{t}') \otimes (1,\mathbf{w}')]\mathbf{d}\right)^k f^{\mathbf{Y}_\mathbf{u}^\perp|\mathbf{W}=\mathbf{w}_0}(\mathbf{t})\,\mathrm{d}\mathbf{t}\,\mathrm{d}\mathbf{w},$$

which leads to

$$\lim_{n\to\infty} h_n^{-(p-1)}\mathrm{E}[\tilde{v}_1]$$
$$= \lim_{n\to\infty} h_n^{-(p-1)}\mathrm{E}\Big[K_{h1}(\tau - I[Z_{1\mathbf{u}}^\ell < 0])(\mathbf{d}'\boldsymbol{\mathcal{X}}_{h1\mathbf{u}}^\ell)\Big]$$
$$= (\tau - \tau)f^{\mathbf{W}}(\mathbf{w}_0)\int_{\mathbb{R}^{p-1}}\int_{\mathbb{R}^{m-1}} K(\mathbf{w})\left([(1,\mathbf{t}') \otimes (1,\mathbf{w}')]\mathbf{d}\right)f^{\mathbf{Y}_\mathbf{u}^\perp|\mathbf{W}=\mathbf{w}_0}(\mathbf{t})\,\mathrm{d}\mathbf{t}\mathrm{d}\mathbf{w} = 0$$

and

$$\lim_{n\to\infty} h_n^{-(p-1)}\mathrm{E}[\tilde{v}_1^2]$$
$$= \lim_{n\to\infty} h_n^{-(p-1)}\mathrm{E}\Big[K_{h1}^2(\tau^2 - 2\tau I[Z_{1\mathbf{u}}^\ell < 0] + I[Z_{1\mathbf{u}}^\ell < 0])(\mathbf{d}'\boldsymbol{\mathcal{X}}_{h1\mathbf{u}}^\ell)^2\Big]$$
$$= \tau(1-\tau)f^{\mathbf{W}}(\mathbf{w}_0)\int_{\mathbb{R}^{p-1}}\int_{\mathbb{R}^{m-1}} K^2(\mathbf{w})\left([(1,\mathbf{t}') \otimes (1,\mathbf{w}')]\mathbf{d}\right)^2 f^{\mathbf{Y}_\mathbf{u}^\perp|\mathbf{W}=\mathbf{w}_0}(\mathbf{t})\,\mathrm{d}\mathbf{t}\,\mathrm{d}\mathbf{w}.$$

Therefore,

$$\lim_{n\to\infty} h_n^{-(p-1)}\mathrm{Var}[\tilde{v}_1]$$
$$= \left(\lim_{n\to\infty} h_n^{-(p-1)}\mathrm{E}[\tilde{v}_1^2]\right) - \left(h_n^{-(p-1)}(\mathrm{E}[\tilde{v}_1])^2\right)$$
$$= \tau(1-\tau)f^{\mathbf{W}}(\mathbf{w}_0)\int_{\mathbb{R}^{p-1}}\int_{\mathbb{R}^{m-1}} K^2(\mathbf{w})\left([(1,\mathbf{t}') \otimes (1,\mathbf{w}')]\mathbf{d}\right)^2 f^{\mathbf{Y}_\mathbf{u}^\perp|\mathbf{W}=\mathbf{w}_0}(\mathbf{t})\,\mathrm{d}\mathbf{t}\,\mathrm{d}\mathbf{w},$$

which, together with (A.12), establishes the result. $\square$

Proof of Theorem 5.1. The proof consists in checking that the conditions of Lemma A.1 are satisfied. Lemmas A.3 and A.4 entail condition (ii) of Lemma A.1, with $\mathbf{D} = f^{\mathbf{W}}(\mathbf{w}_0)\,\mathrm{diag}(1,\boldsymbol{\mu}_2^K)$ (which yields to $(\mathbf{G}_{\boldsymbol{\tau};\mathbf{w}_0} \otimes \mathbf{D})^{-1} = \boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}_0}^\ell$) and

$$\mathbf{A}_n = \mathbf{V}_n(\mathbf{0}) = H_n^{-1}\sum_{i=1}^{n}K_{hi}\psi_\tau(Z_{i\mathbf{u}}^\ell)\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell,$$

which, by Lemma A.6, is $O_{\mathrm{P}}(1)$. As for Condition (ii) of Lemma A.1, the fact that

$$\lambda \mapsto -\boldsymbol{\varphi}'\mathbf{V}_n(\lambda\boldsymbol{\varphi}) = H_n^{-1}\sum_{i=1}^{n}K_{hi}\psi_\tau(Z_{i\mathbf{u}}^\ell - \lambda H_n^{-1}\boldsymbol{\varphi}'\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell)(-\boldsymbol{\varphi}'\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell)$$

is non-decreasing directly follows from the fact $y \mapsto \psi_\tau(y)$ is non-decreasing. Since, moreover, $\|\mathbf{V}_n(\boldsymbol{\varphi}^{(n)})\|$ is $o_{\mathrm{P}}(1)$ (this follows from Lemma A.5 and Assumptions (A2)-(A3)), Lemma A.1 applies and establishes the result. $\square$

Proof of Theorem 5.2. On the basis of the Bahadur representation of Theorem 5.1, the asymptotic normality of $\hat{\boldsymbol{\theta}}^{\ell(n)}$ follows exactly as in the corresponding proofs for usual nonparametric regression in the i.i.d. case (see, e.g., [8]), yielding the asymptotic normality with the bias (i.e., the expectation) of the first term on the right-hand side of (5.5) as

$$\mathrm{E}\!\left[\frac{\boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}_0}^\ell}{\sqrt{nh_n^{p-1}}}\sum_{i=1}^{n}K_{h1}\psi_\tau(Z_{i\mathbf{u}}^\ell)\boldsymbol{\mathcal{X}}_{hi\mathbf{u}}^\ell\right]$$

$$= \frac{\boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}_0}^\ell}{\sqrt{nh_n^{p-1}}}\,n\mathrm{E}\!\left[K_{h1}\psi_\tau(Z_{1\mathbf{u}}^\ell)\boldsymbol{\mathcal{X}}_{h1\mathbf{u}}^\ell\right]$$

$$= \boldsymbol{\eta}_{\boldsymbol{\tau};\mathbf{w}_0}^\ell\sqrt{nh_n^{p-1}}\,h_n^{-(p-1)}\mathrm{E}\!\left[K_{h1}\big(F^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp,\mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{W}} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{W}}'\mathbf{Y}_{1\mathbf{u}}^\perp)\right.$$

$$\left.-F^{Y_\mathbf{u}|(\mathbf{Y}_\mathbf{u}^\perp,\mathbf{W})}(a_{\boldsymbol{\tau};\mathbf{w}_0} + \mathbf{c}_{\boldsymbol{\tau};\mathbf{w}_0}'\mathbf{Y}_{1\mathbf{u}}^\perp + T_{n1})\big)\,\boldsymbol{\mathcal{X}}_{h1\mathbf{u}}^\ell\right]$$

$$= \sqrt{nh_n^{p-1}}\left(\frac{h_n^2}{2}\,\mathbf{B}_{\mathbf{w}_0}^\ell + o(h_n^2)\right),$$

where the last equality is derived from a first-order Taylor expansion of $y \mapsto F^{Y_{\mathbf{u}}|(\mathbf{Y}_{\mathbf{u}}^{\perp}, \mathbf{X})}(y)$ and a second-order Taylor expansion of $\mathbf{w} \mapsto (a_{\boldsymbol{\tau};\mathbf{w}}, \mathbf{c}'_{\boldsymbol{\tau};\mathbf{w}})'$ at $\mathbf{w} = \mathbf{w}_0$ (these expansions exist in view of Assumptions (A1) and (A4)). The $o(h_n^2)$ term is taken care of by Assumption (A3). The asymptotic variance of the theorem readily follows from Lemma A.6. Details are omitted. $\square$

REFERENCES

[1] BHATTACHARYA, P. K., and GANGOPADHYAY, A. K. (1990). Kernel and nearest-neighbor estimation of a conditional quantile. *Ann. Statist.* **18** 1400–1415.

[2] BONELLIE, S. R., and RAAB, G. M. (1996). A comparison of different approaches for fitting centile curves to birthweight data. *Stat. Med.* **15** 2657–2667.

[3] CAI, Z., and XU, X. (2009). Nonparametric quantile estimations for dynamic smooth coefficient models. *J. Amer. Statist. Assoc.* **104** 371–383.

[4] CHAUDHURI, P. (1991). Nonparametric estimates of regression quantiles and their local Bahadur representation. *Ann. Statist.* **19** 760–777.

[5] COLE, T. J. (1988). Fitting smoothed centile curves to reference data. *J. R. Statist. Soc. A* **151** 385–418.

[6] DOŠLÁ, Š. (2009). Conditions for bimodality and multimodality of a mixture of two unimodal densities. *Kybernetika* **45** 279-292.

[7] EL GHOUCH, A., and VAN KEILEGOM, I. (2009). Local linear quantile regression with dependent censored data. *Statist. Sinica* **19** 1621–1640.

[8] FAN, J., and GIJBELS, I. (1996). *Local Polynomial Modelling and its Applications.* Chapman & Hall, London.

[9] FAN, J., HU, T.C., and TRUONG, Y.K. (1994). Robust nonparametric function estimation. *Scand. J. Statist.* **21** 433–446.

[10] FAN, J., and YAO, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods.* New York: Springer-Verlag.

[11] FATTI, L. P., SENAOANA, E. M., and THOMPSON, M. L. (1998). Bayesian updating in reference centile charts. *J. R. Statist. Soc. A* **161** 103–115.

[12] GANNOUN, A., GIRARD, S., GUINOT, C., and SARACCO, J. (2002). Reference curves based on non-parametric quantile regression. *Stat. Med.* **21** 3119–3135.

[13] GANNOUN, A., SARACCO, J., and YU, K. (2003). Nonparametric prediction by conditional median and quantiles. *J. Statist. Plann. Inference* **117** 207–223.

[14] GILONI, A., and SIMONOFF, J. (2005). The conditional breakdown properties of least absolute value local polynomial estimators. *J. Nonparametr. Statist.* **17** 15–30.

[15] HABERMAN, S. J. (1989). Concavity and estimation. *Ann. Statist.* **17** 1631–1661.

[16] HAERDLE, W. K., and SONG, S. (2010). Confidence bands in quantile regression. *Econometric Theory* **26** 1180–1200.

[17] HALLIN, M., LU, Z., and YU, K. (2009). Local linear spatial quantile regression. *Bernoulli* **15** 659–686.

[18] HALLIN, M., PAINDAVEINE, D., and ŠIMAN, M. (2010). Multivariate quantiles and multiple-output regression quantiles : From L1 optimization to halfspace depth (with discussion). *Ann. Statist.* **38** 635–669.

[19] HEINZ, G., PETERSON, L. J., JOHNSON, R. W., and KERK, C. J. (2003). Exploring relationships in body dimensions. *J. Statist. Education* **11**, online at www.amstat.org/publications/jse/v11n2/datasets.heinz.html.

[20] HONDA, T. (2000). Nonparametric estimation of a conditional quantile for α-mixing processes. *Ann. Inst. Statist. Math.* **52** 459–470.

[21] HONDA, T. (2004). Quantile regression in varying coefficient models. *J. Statist. Plann. Inference* **121** 113–125.

[22] IOANNIDES, D. A. (2004). Fixed design regression quantiles for time series. *Statist. Probab. Lett.* **68** 235-245.

[23] KOENKER, R. and ZHAO, Q. (1996). Conditional quantile estimation and inference for ARCH models. *Econometric Theory* **12** 793–813.

[24] KONG, E., LINTON, O., and XIA, Y. (2010). Uniform Bahadur representation for local polynomial estimates of m-regression and its application to the additive model. *Econometric Theory* **26** 1529–1564.

[25] LU, Z., HUI, Y.V., and ZHAO, Q. (1998). Local linear quantile regression under dependence: Bahadur representation and application. Discussion paper, Dept. Management Sciences, City Univ. Hong Kong.

[26] McKeague, I.W., López-Pintado, S., Hallin, M., and Šiman, M. (2011). Analyzing growth trajectories. *Journal of Developmental Origins of Health and Disease*, to appear.

[27] PAINDAVEINE, D. and ŠIMAN, M. (2011a). On directional multiple-output quantile regression. *J. Multivariate Anal.* **102** 193–212.

[28] PAINDAVEINE, D., and ŠIMAN, M. (2011b). Computing multiple-output regression quantile regions. *Comput. Statist. Data Anal.*, to appear.

[29] PAINDAVEINE, D., and ŠIMAN, M. (2011c). Computing multiple-output regression

quantile regions from projectional quantiles. *Comput. Statist.*, to appear.

[30] PETERSEN, J. H. (2003). Two bivariate geometrically defined reference regions with applications to male reproductive hormones and human growth. *Stat. Med.* **22** 2603–2618.

[31] ROUSSEEUW, P. J., and HUBERT, M. (1999). Regression depth. *J. Amer. Statist. Assoc.* **94** 388–402.

[32] ROUSSEEUW, P. J., and RUTS, I. (1999). The depth function of a population distribution. *Metrika* **49** 213–244.

[33] RUPPERT, D., and CARROLL, R. J. (1980). Trimmed least squares estimation in the linear model. *J. Amer. Statist. Assoc.* **75** 828–837.

[34] WANG, F. T., and SCOTT, D. W. (1994). The L1 method for robust nonparametric regression. *J. Amer. Statist. Assoc.* **89** 65–76.

[35] WEI, Y. (2008). An approach to multivariate covariate-dependent quantile contours with application to bivariate conditional growth charts. *J. Amer. Statist. Assoc.* **103** 397–409.

[36] WEI, Y., and HE, X. (2006). Conditional growth charts. *Ann. Statist.* **34** 2069–2097.

[37] WEI, Y., PERE, A., KOENKER, R., and HE, X. (2006). Quantile regression methods for reference growth charts. *Stat. Med.* **25** 1369–1382.

[38] WRIGHT, E. M., and ROYSTON, P. (1997). A comparison of statistical methods for age-related reference intervals. *J. R. Statist. Soc. A* **160** 47–69.

[39] YU, K., and JONES, M. C. (1997). A comparison of local constant and local linear regression quantile estimators. *Comput. Statist. Data Anal.* **25** 159–166.

[40] YU, K., and JONES, M. C. (1998). Local linear quantile regression. *J. Amer. Statist. Assoc.* **93** 228–237.

[41] YU, K., and LU, Z. (2004). Local linear additive quantile regression. *Scand. J. Statist.* **31** 333–346.

[42] ZHOU, Z., and WU, W. B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729.

Marc Hallin, Davy Paindaveine
Université libre de Bruxelles
E.C.A.R.E.S., CP114/04
50, Avenue F.D. Roosevelt
B-1050 Brussels, Belgium
E-mail: mhallin@ulb.ac.be
        dpaindav@ulb.ac.be
URL: http://homepages.ulb.ac.be/˜dpaindav

Zudi Lu
School of Mathematical Sciences
The University of Adelaide
North Terrace, Adelaide
SA 5005, Australia
E-mail: zudi.lu@adelaide.edu.au

Miroslav Šiman
Institute of Information Theory and Automation of the ASCR
Department of Stochastic Informatics
Pod Vodárenskou věží 4
CZ-182 08 Prague 8, Czech Republic