Quantile-Regression Inference With Adaptive Control of Size
SUPPLEMENTARY MATERIAL

Juan Carlos Escanciano*
Indiana University
and
Chuan Goh†
University of Guelph

July 17, 2018

Abstract

This supplement contains precise statements of the assumptions underlying the theoretical results in the main body of our paper. This supplement also contains proofs of those results, analyses of the established estimators of $\boldsymbol{G}_0(\alpha)$ proposed by Powell (1991) and Hendricks and Koenker (1992), details regarding a data-driven bandwidth that can be used to implement the proposed Wald-type testing procedure, results of further simulations and further discussion of the empirical example presented in the paper.

*Universidad Carlos III de Madrid, Calle Madrid 126 28903, Getafe (Madrid), Spain. Phone: +34 91 624 61 98. Fax: +34 91 624 93 29. E-mail: jescanci@eco.uc3m.es. Research funded by the Spanish Plan Nacional de I+D+I, reference number ECO2014-55858-P.

†Department of Economics and Finance, University of Guelph, 50 Stone Road East, Guelph, ON, Canada, N1G 2W1. Phone: +1 (519) 824 4120. Fax: +1 (519) 763 8497. E-mail: gohc@uoguelph.ca. Web Page: http://www.chuangoh.org/

A Assumptions

Precise statements of the regularity conditions underlying the theorems presented in the main body of our paper are collected here. Henceforth, $(\mathbf{X}_1^\top, Y_1), \ldots, (\mathbf{X}_n^\top, Y_n)$ is am iid sample of size n. Let $\|\cdot\|$ denote the Euclidean norm, let $\mathcal{X}$ denote the common support of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, and define

$$\mathcal{B} \equiv \left\{ \boldsymbol{x}^\top \boldsymbol{\beta}(\alpha) : \, \boldsymbol{x} \in \mathcal{X}, \, \alpha \in \mathcal{A} \right\}. \tag{1}$$

Assumption 1. *The following conditions hold: (1) Uniformly in* $\alpha \in \mathcal{A}$, *where* $\mathcal{A} = [a_1, a_2]$ *for* $0 < a_1 < a_2 < 1$, $\boldsymbol{G}_n(\alpha) \equiv n^{-1} \sum_{i=1}^n f_i \left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \boldsymbol{X}_i \boldsymbol{X}_i^\top = \boldsymbol{G}_0(\alpha) \left(1 + O_p \left(n^{-1/2} \right) \right)$ *as* $n \to \infty$*; (2)*

$$E\left[\|\boldsymbol{X}_i\|^4 \right] < \infty; \tag{2}$$

and (3) $\boldsymbol{G}_0(\alpha)$ *and* $\boldsymbol{H} = E\left[\boldsymbol{X}_i \boldsymbol{X}_i^\top \right]$ *are positive definite.*

Assumption 2. *The smoothing kernel* $K(\cdot)$ *satisfies the following conditions: (1)* $K(\cdot)$ *is nonnegative, symmetric and bounded with support* $[-1/2, 1/2]$, *with* $\|K\|_2 \equiv \sqrt{\int_{-1/2}^{1/2} K^2(w)dw} \in (0, \infty)$*; (2)* $\int_{-1/2}^{1/2} K(w)dw = 1$ *and* $\left| \int_{-1/2}^{1/2} w^k K(w)dw \right| < \infty$ *for* $k \leq 4$*; and (3)* $K(\cdot)$ *is three-times continuously differentiable on* $\mathbb{R}$, *where the derivatives* $K^{(k)}(w)$ *satisfy* $\int_{-1/2}^{1/2} \left| K^{(k)}(w) \right| dw < \infty$ *for* $k = 1, 2, 3$.

Assumption 3. *For* $m \to \infty$ *as* $n \to \infty$ *at a rate no slower than* $[n/(\log n)^{11/5}]^{5/4}$*, the bandwidth sequence* $\{h_m\}$ *satisfies (1)* $h_m \to 0$*; (2)* mh_m^5, $nh_m^4 \to \infty$*; and (3)* $mh_m / \left[(\log m)^2 \sqrt{\log h_m^{-1}} \right] \to \infty$.

Assumption 4. *For each each* $i = 1, \ldots, n$*: (1) The conditional moment* $\Pr\left[Y_i \leq \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \middle| \boldsymbol{X}_i \right] = \alpha$ *holds almost surely for* $\alpha \in [a_1, a_2]$*. (2) The conditional distribution function* F_i *is absolutely continuous, with corresponding density* f_i *such that* $f_i(\cdot)$ *is uniformly continuous on the closure of* $\mathcal{B}$*. In addition, there exists an interval* $[-b, b]$ *with* $b \in (0, \infty)$ *such that* $\mathcal{B} \subset [-b, b]$*, where* $\mathcal{B}$ *is as given above in (1). (3) The densities* $f_i(y)$ *are five-times differentiable for all* $y \in \mathcal{B}$ *with* $\max_{1 \leq i \leq n} \sup_{y \in \mathcal{B}} \left| f_i^{(k)}(y) \right| < \infty$ *a.s. for each* $k = 0, 1, \ldots, 5$*. (4) There exist constants* $0 < l_1 \leq l_2 < \infty$ *such that* $0 < l_1 \leq f_i(y) \leq l_2 < \infty$ *for all* $y \in \mathcal{B}$.

Assumption 5. *The first component of the design vector* $\boldsymbol{X}_i$ *is an intercept, i.e.,* $\boldsymbol{X}_i = [\ 1 \quad \tilde{\boldsymbol{X}}_i^\top\ ]^\top$ *for some* $(d-1)$*-variate* $\tilde{\boldsymbol{X}}_i$*. Let* $\boldsymbol{\beta}^{(1)}(\alpha) \equiv (d/d\alpha) \boldsymbol{\beta}(\alpha)$*, i.e., the gradient vector, and let* $\tilde{\phi}_i(t)$ *denote the conditional characteristic function given* $\tilde{\boldsymbol{X}}_i$ *of the random variable*

$$\tilde{\boldsymbol{X}}_i \cdot \left(1 \left\{ Y_i \leq \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(\alpha) + n^{-1/2} \boldsymbol{\beta}^{(1)}(\alpha) \right) \right\} - \alpha \right).$$

For any $\epsilon > 0$, *there exists* $\eta \in (0,1)$ *such that* $\inf_{\|t\|>\epsilon} \prod_{i=1}^{n} \tilde{\phi}_i(t) \leq \eta^n$ *uniformly in* $\alpha \in [\epsilon, 1-\epsilon]$.

Remark 1. *The requirement of part 1 of Assumption 1 that* $\boldsymbol{G}_n(\alpha)$ *converge uniformly in* $\alpha \in \mathcal{A}$ *to* $\boldsymbol{G}_0(\alpha)$ *is used only in the proof of Theorem 3 appearing below in Appendix C. Theorems 1 and 2, whose proofs appear below in Appendix B only involve a requirement of pointwise convergence, i.e., that* $\boldsymbol{G}_n(\alpha)$ *converge to* $\boldsymbol{G}_0(\alpha)$ *for any* $\alpha \in \mathcal{A}$.

Remark 2. *Although we do not make this explicit in the conditions of Assumption 3, our results do allow for stochastic bandwidths. This is certainly relevant to the discussion presented below in Appendix D in which a data-driven bandwidth is derived.*

Remark 3. *Part 1 of Assumption 4 requires the correct specification of the quantile regression model on* $[a_1, a_2]$. *This condition may restrict the choice of* a_1 *and* a_2 *in practice. Specification tests developed in e.g., Escanciano and Goh (2014) and related papers can be used to check this condition. On the other hand, unreported simulations do suggest that Wald tests incorporating the proposed estimator of* $\boldsymbol{G}_0(\alpha)$ *have a certain degree of robustness in finite samples to incorrect specification of the underlying quantile regression model on a given interval of quantiles* $[a_1, a_2]$. *An extension of the analysis presented in this paper to the case where our estimator of* $\boldsymbol{G}_0(\alpha)$ *is computed using a shrinking neighborhood* $[a_{1n}, a_{2n}]$ *of the quantile level* α, *on which the quantile-regression specification holds, may be desirable.*

Remark 4. *Assumption 5 is taken from Portnoy (2012) and can be shown to hold if the distribution of* $\tilde{\boldsymbol{X}}_i$ *is appropriately smooth and bounded. In addition, the condition of part 2 of Assumption 4 , which implies that* $\|\boldsymbol{X}_i\|$ *is uniformly bounded on its support, is required in the derivation of both the precise form of the quantity* $\boldsymbol{T}_{2nm}(\alpha)$ *appearing in the statement of Theorem 1 of our paper and of the error rate appearing in the statement of Portnoy (2012, Theorem 5). Relaxation of this condition will not affect the* $O_p\left(\sqrt{\log h_m^{-1}/(mh_m)}\right)$ *convergence rate of* $\boldsymbol{T}_{2nm}(\alpha)$ *stated in Theorem 1 of our paper, but will likely increase the power of* $\log n$ *that appears in the conclusion of Portnoy (2012, Theorem 5); see in this connection the discussion in Portnoy (2012, p. 1720).*

B Proofs of Theorems 1 and 2

B.1 A useful lemma

We begin by introducing a useful lemma on uniform-in-bandwidth rates of convergence of kernel-type estimators that is instrumental in proving our main results. Recall that $\{U_j\}_{j=1}^{m}$ denotes a random sample of size m, distributed as $U \sim Unif[a_1, a_2]$ for $[a_1, a_2] \subset (0,1)$. For a given

$\boldsymbol{x} \in \mathbb{R}^d$, let $F_{\boldsymbol{x}}(\cdot)$ and $f_{\boldsymbol{x}}(\cdot)$ denote the cdf and Lebesgue density of the random variable $\boldsymbol{x}^\top \boldsymbol{\beta}(U)$. Define

$$g_{m,h}(w, \boldsymbol{x}) \equiv \frac{1}{m} \sum_{j=1}^{m} \varphi\left(\frac{\boldsymbol{x}^\top \boldsymbol{\beta}(U_j) - w}{h}\right),$$

where $\varphi(\cdot)$ is either $(a_2 - a_1) K(\cdot)$, $(a_2 - a_1) K^{(1)}(\cdot)$ or $(a_2 - a_1) K^{(2)}(\cdot)$. Let $\mathcal{X} \subset \mathbb{R}^d$ denote the support of $\boldsymbol{x}$, and let wp. 1 stand for with probability one.

Lemma 1. *Under Assumptions 2–4, for* $c > 0$, *and* $0 < h_0 < 1$,, *the following holds wp. 1:*

$$\limsup_{m\to\infty} \sup_{c \log m/m \le h \le h_0} \sup_{\boldsymbol{x}\in\mathcal{X}} \sup_{w\in\mathbb{R}} \frac{\sqrt{m}\left|g_{m,h}(w, \boldsymbol{x}) - E\left[g_{m,h}(w, \boldsymbol{x})\right]\right|}{\sqrt{h|\log h|}} \equiv A(c) < \infty, \tag{3}$$

for $\varphi(\cdot)$ *equal to* $(a_2 - a_1) K(\cdot)$, $(a_2 - a_1) K^{(1)}(\cdot)$ *or* $(a_2 - a_1) K^{(2)}(\cdot)$.

Proof. We provide the proof for $\varphi(\cdot) = (a_2 - a_1) K(\cdot)$; the proof for $\varphi(\cdot) = (a_2 - a_1) K^{(1)}(\cdot)$ or $\varphi(\cdot) = (a_2 - a_1) K^{(2)}(\cdot)$ is the same. In particular, Lemma 1 follows from an application of the main result of Mason and Swanepoel (2011, p. 73) applied to the class of functions

$$\mathcal{G} = \left\{(u, h) \to K\left(\frac{\boldsymbol{x}^\top \boldsymbol{\beta}(u) - w}{h}\right) : \boldsymbol{x} \in \mathcal{X}, w \in \mathbb{R}\right\}.$$

We proceed to verify their conditions. Since K is bounded, Mason and Swanepoel (2011, Condition (G.i)) is trivially satisfied. To verify Mason and Swanepoel (2011, Condition (G.ii)) note that

$$\begin{aligned}
E\left[K^2\left(\frac{\boldsymbol{x}^\top \boldsymbol{\beta}(U) - w}{h}\right)\right] &= \frac{1}{a_2 - a_1} \int_{a_1}^{a_2} K^2\left(\frac{\boldsymbol{x}^\top \boldsymbol{\beta}(u) - w}{h}\right) du \\
&= \frac{1}{a_2 - a_1} \int_{\boldsymbol{x}^\top \boldsymbol{\beta}(a_1)}^{\boldsymbol{x}^\top \boldsymbol{\beta}(a_2)} K^2\left(\frac{q - w}{h}\right) f_{\boldsymbol{x}}(q) dq \\
&= \frac{h}{a_2 - a_1} \int_{(\boldsymbol{x}^\top \boldsymbol{\beta}(a_1) - w)/h}^{(\boldsymbol{x}^\top \boldsymbol{\beta}(a_2) - w)/h} K^2(t) f_{\boldsymbol{x}}(w + ht) dt \\
&\le \frac{h}{a_2 - a_1} \|K\|_2^2 l_2,
\end{aligned}$$

where $\|K\|_2 = \sqrt{\int_{-1/2}^{1/2} K^2(w) dw} < \infty$ by Assumption 2 and where l_2 is as given in Assumption 4 above. Hence, Mason and Swanepoel (2011, Condition (G.ii)) holds. The class

$$\mathcal{G}_0 = \left\{u \to K\left(\frac{\boldsymbol{x}^\top \boldsymbol{\beta}(u) - w}{h}\right) : \boldsymbol{x} \in \mathcal{X}, w \in \mathbb{R}, h \in (0, 1]\right\}$$

is a VC class, which is also pointwise measurable, see e.g. Nolan and Pollard (1987). It follows that Mason and Swanepoel (2011, Conditions (F.i) and (F.ii)) hold. This completes the proof. □

B.2 Proof of Theorem 1

Begin by noting that

$$\begin{aligned}
\hat{\boldsymbol{G}}_n(\alpha) &= \frac{1}{n}\sum_{i=1}^{n} \hat{f}_{ni}\left(\boldsymbol{X}_i^\top \hat{\boldsymbol{\beta}}_n(\alpha)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top \\
&= \frac{a_2-a_1}{nmh_m}\sum_{i=1}^{n}\sum_{j=1}^{m} K\left(\frac{1}{\sqrt{n}h_m}\boldsymbol{X}_i^\top\left[\sqrt{n}\left(\hat{\boldsymbol{\beta}}_n\left(U_j\right)-\boldsymbol{\beta}\left(U_j\right)\right)-\sqrt{n}\left(\hat{\boldsymbol{\beta}}_n(\alpha)-\boldsymbol{\beta}(\alpha)\right)\right.\right. \\
&\left.\left.+\sqrt{n}\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right]\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top \\
&= \frac{a_2-a_1}{nmh_m}\sum_{i=1}^{n}\sum_{j=1}^{m} K\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)+\frac{1}{\sqrt{n}h_m}\boldsymbol{X}_i^\top\boldsymbol{D}_{nj}(\alpha)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top,
\end{aligned}$$

where

$$\boldsymbol{D}_{nj}(\alpha) = \sqrt{n}\left[\left(\hat{\boldsymbol{\beta}}_n\left(U_j\right)-\boldsymbol{\beta}\left(U_j\right)\right)-\left(\hat{\boldsymbol{\beta}}_n(\alpha)-\boldsymbol{\beta}(\alpha)\right)\right].$$

By a Taylor expansion we accordingly have

$$\begin{aligned}
\hat{\boldsymbol{G}}_n(\alpha) = \frac{a_2-a_1}{nmh_m}\sum_{i=1}^{n}\sum_{j=1}^{m}&\left[K\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right)\right. \\
&+\frac{1}{\sqrt{n}h_m}\boldsymbol{X}_i^\top\boldsymbol{D}_{nj}(\alpha)K^{(1)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right) \\
&\left.+\frac{1}{2nh_m^2}\left(\boldsymbol{X}_i^\top\boldsymbol{D}_{nj}(\alpha)\right)^2 K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\bar{\boldsymbol{\beta}}\left(U_j\right)-\bar{\boldsymbol{\beta}}(\alpha)\right)\right)\right]\boldsymbol{X}_i\boldsymbol{X}_i^\top,
\end{aligned}$$

where

$$\left\|\left(\bar{\boldsymbol{\beta}}\left(U_j\right)-\bar{\boldsymbol{\beta}}(\alpha)\right)-\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right\| < \left\|\frac{1}{\sqrt{n}}\boldsymbol{D}_{nj}(\alpha)\right\|$$

for each $j = 1,\ldots,m$.

Then

$$\hat{\boldsymbol{G}}_n(\alpha) = \frac{1}{n}\sum_{i=1}^{n}\tilde{f}_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top$$

$$
+\frac{a_2-a_1}{nmh_m^2}\sum_{i=1}^{n}\boldsymbol{X}_i^\top\left[\sum_{j=1}^{m}\frac{1}{\sqrt{n}}\boldsymbol{D}_{nj}(\alpha)K^{(1)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right)\right]\boldsymbol{X}_i\boldsymbol{X}_i^\top
$$

$$
+\frac{a_2-a_1}{nmh_m^3}\sum_{i=1}^{n}\boldsymbol{X}_i^\top\left[\sum_{j=1}^{m}\frac{1}{n}\boldsymbol{D}_{nj}(\alpha)\boldsymbol{D}_{nj}(\alpha)^\top K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\bar{\boldsymbol{\beta}}\left(U_j\right)-\bar{\boldsymbol{\beta}}(\alpha)\right)\right)\right]\boldsymbol{X}_i\cdot\boldsymbol{X}_i\boldsymbol{X}_i^\top, \tag{4}
$$

where $\tilde{f}_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)$ denotes the empirically infeasible estimator of $f_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)$ defined in 9 of our paper.

In what follows, the three terms appearing in the representation of $\hat{\boldsymbol{G}}_n(\alpha)$ given in (4) are analyzed in sequence. For convenience, we suppress the dependence on n of the quantile grid size m. We show that the following holds as $n\to\infty$ for a fixed quantile $\alpha\in\mathcal{A}$:

1.
$$
\frac{1}{n}\sum_{i=1}^{n}\tilde{f}_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top=\frac{1}{n}\sum_{i=1}^{n}f_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top+T_{1nm}(\alpha)+\boldsymbol{T}_{2nm}(\alpha)
$$
$$
+o_p\left(h_m^2+\sqrt{\frac{-\log h_m}{mh_m}}\right), \tag{5}
$$

   where $\boldsymbol{T}_{1nm}(\alpha)$ and $\boldsymbol{T}_{2nm}(\alpha)$ are as given above in the statement of Theorem 1, and where $\boldsymbol{T}_{1nm}(\alpha)=O_p\left(h_m^2\right)$ and $\boldsymbol{T}_{2nm}(\alpha)=O_p\left(\sqrt{-\log h_m/(mh_m)}\right)$. It follows that the remainder term in (5) is of strictly smaller order than $\boldsymbol{T}_{1nm}(\alpha)+\boldsymbol{T}_{2nm}(\alpha)$.

2. We also show that
$$
\boldsymbol{T}_{2nm}(\alpha)\equiv\frac{a_2-a_1}{nmh_m^2}\sum_{i=1}^{n}\boldsymbol{X}_i^\top\left[\sum_{j=1}^{m}\frac{1}{\sqrt{n}}D_{nj}(\alpha)K^{(1)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right)\right]\boldsymbol{X}_i\boldsymbol{X}_i^\top
$$
$$
=O_p\left(\frac{1}{\sqrt{n}}\right); \tag{6}
$$

3. and finally that
$$
\frac{a_2-a_1}{nmh_m^3}\sum_{i=1}^{n}\boldsymbol{X}_i^\top\left[\sum_{j=1}^{m}\frac{1}{n}\boldsymbol{D}_{nj}(\alpha)\boldsymbol{D}_{nj}(\alpha)^\top K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\bar{\boldsymbol{\beta}}\left(U_j\right)-\bar{\boldsymbol{\beta}}(\alpha)\right)\right)\right]\boldsymbol{X}_i\cdot\boldsymbol{X}_i\boldsymbol{X}_i^\top
$$

$$= O_p\left(\frac{1}{n} + \frac{1}{n^{\frac{3}{2}} h_m^4}\right). \tag{7}$$

Combining (4) with (5)–(7) yields the desired conclusion; namely, that

$$\begin{aligned}
\hat{\boldsymbol{G}}_n(\alpha) &= \frac{1}{n}\sum_{i=1}^{n} \tilde{f}_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \boldsymbol{X}_i \boldsymbol{X}_i^\top + T_{1nm}(\alpha) + \boldsymbol{T}_{2nm}(\alpha) + \boldsymbol{T}_{2nm}(\alpha) + \boldsymbol{R}_{nm}(\alpha) \\
&= G_0(\alpha) + \boldsymbol{T}_{1nm}(\alpha) + \boldsymbol{T}_{2nm}(\alpha) + T_{3nm}(\alpha) + \boldsymbol{R}_{nm}(\alpha),
\end{aligned}$$

where $\boldsymbol{R}_{nm}(\alpha)$ denotes the sum of the remainder term in (5) and the expression in (7). In particular,

$$\boldsymbol{R}_{nm}(\alpha) = o_p\left(h_m^2 + \sqrt{\frac{-\log h_m}{m h_m}}\right) + O_p\left(\frac{1}{n} + \frac{1}{n^{\frac{3}{2}} h_m^4}\right) = o_p\left(\boldsymbol{T}_{1nm}(\alpha) + \boldsymbol{T}_{2nm}(\alpha) + \boldsymbol{T}_{2nm}(\alpha)\right).$$

Claim 1. The following holds:

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left[\tilde{f}_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) - f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right] \boldsymbol{X}_i \boldsymbol{X}_i^\top \\
&= \boldsymbol{T}_{1nm}(\alpha) + \boldsymbol{T}_{2nm}(\alpha) + o_p\left(h_m^2 + \sqrt{\frac{-\log h_m}{m h_m}}\right),
\end{aligned}$$

where

$$\boldsymbol{T}_{1nm}(\alpha) = O_p\left(h_m^2\right)$$

and

$$\boldsymbol{T}_{2nm}(\alpha) = O_p\left(\sqrt{\frac{-\log h_m}{m h_m}}\right).$$

We begin by establishing the rates of convergence of $\boldsymbol{T}_{1nm}(\alpha)$ and $\boldsymbol{T}_{2nm}(\alpha)$. In particular, we have

$$\begin{aligned}
\|\boldsymbol{T}_{1nm}(\alpha)\| &\le \frac{h_m^2}{2}\int_{-1/2}^{1/2} w^2 K(w) dw \cdot \max_{1\le i\le n}\left|f_i^{(2)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right| \cdot \frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{X}_i\|^2 \\
&= O_p\left(h_m^2\right),
\end{aligned} \tag{8}$$

where use has been made of the conditions of Assumptions 1, 2 and 4 that $E\left[\|\boldsymbol{X}_i\|^4\right] < \infty$, $\int_{-1/2}^{1/2} w^2 K(w)dw < \infty$ and $\max_{1\le i\le n}\sup_{y\in\mathcal{B}}\left|f_i^{(2)}(y)\right| < \infty$ a.s., where $\mathcal{B}$ is as given above in (1). In addition, we have

$$\begin{aligned}\|\boldsymbol{T}_{2nm}(\alpha)\| &\le \sqrt{\frac{-\log h_m}{mh_m}\cdot\|K\|_2^2\cdot\max_{1\le i\le n}\sup_{y\in\mathcal{X}_i}|f_i(y)|}\cdot\frac{1}{n}\sum_{i=1}^n\|\boldsymbol{X}_i\|^2\\ &= O_p\left(\sqrt{\frac{-\log h_m}{mh_m}}\right)\end{aligned} \tag{9}$$

via a similar argument.

Next, define the quantities

$$V_{\tilde{f}_i}(\alpha) \equiv \tilde{f}_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right) - E\left[\frac{a_2-a_1}{h_m}K\left(\frac{1}{h_m}\boldsymbol{X}_i^\top(\boldsymbol{\beta}(U)-\boldsymbol{\beta}(\alpha))\right)\middle|\boldsymbol{X}_i\right]$$

and

$$B_{\tilde{f}_i}(\alpha) \equiv E\left[\frac{a_2-a_1}{h_m}K\left(\frac{1}{h_m}\boldsymbol{X}_i^\top(\boldsymbol{\beta}(U)-\boldsymbol{\beta}(\alpha))\right)\middle|\boldsymbol{X}_i\right] - f_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)$$

for each $i\in\{1,\ldots,n\}$. We have

$$\frac{1}{n}\sum_{i=1}^n\tilde{f}_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top = \frac{1}{n}\sum_{i=1}^n f_i\left(\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top + \frac{1}{n}\sum_{i=1}^n V_{\tilde{f}_i}(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top + \frac{1}{n}\sum_{i=1}^n B_{\tilde{f}_i}(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top. \tag{10}$$

Recall that $\mathcal{X}\subset\mathbb{R}^d$ denotes the common support of $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3,\ldots$. In addition, let $f_{\boldsymbol{x}}(\cdot)$ denote the density of the conditional distribution of Y_i given $\boldsymbol{X}_i=\boldsymbol{x}$, and let

$$\tilde{f}_{\boldsymbol{x}}(\boldsymbol{x}^\top\boldsymbol{\beta}(\alpha)) \equiv (mh_m)^{-1}\sum_{j=1}^m K\left(h_m^{-1}\boldsymbol{x}^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right).$$

Recall the condition of Assumption 4 that for some finite $b>0$, $\mathcal{B}\subset[-b,b]$. The following holds almost surely:

$$\left\|\frac{1}{n}\sum_{i=1}^n V_{\tilde{f}_i}(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top - \boldsymbol{T}_{2nm}(\alpha)\right\|$$

$$= \left\| \sqrt{\frac{\log h_m^{-1}}{mh_m}} \cdot \frac{1}{n} \sum_{i=1}^{n} \sqrt{f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)} \boldsymbol{X}_i \boldsymbol{X}_i^\top \right.$$
$$\left. \cdot \left(\sqrt{\frac{mh_m}{\log h_m^{-1}}} \cdot \frac{\tilde{f}_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) - E\left[\tilde{f}_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \middle| \boldsymbol{X}_i\right]}{\sqrt{f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)}} - \|K\|_2 \right) \right\|$$
$$\leq \sup_{\boldsymbol{x}\in\mathcal{X}} \sup_{y\in[-b,b]} \left| \sqrt{\frac{mh_m}{\log h_m^{-1}}} \cdot \frac{\tilde{f}_{\boldsymbol{x}}(y) - E\left[\tilde{f}_{\boldsymbol{x}}(y)\right]}{\sqrt{f_{\boldsymbol{x}}(y)}} - \|K\|_2 \right|$$
$$\cdot \sqrt{\frac{\log h_m^{-1}}{mh_m}} \cdot \frac{1}{n} \sum_{i=1}^{n} \sqrt{f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)} \|\boldsymbol{X}_i\|^2 . \tag{11}$$

By Giné et al. (2004, p. 185) we have for any $z \in \mathbb{R}$ that

$$\lim_{n\to\infty} \Pr\left[(-\log h_m) \left(\frac{\sqrt{mh_m}}{\sqrt{-\log h_m}} \sup_{\boldsymbol{x}\in\mathcal{X}} \sup_{y\in[-b,b]} \left| \frac{\tilde{f}_{\boldsymbol{x}}(y) - E\left[\tilde{f}_{\boldsymbol{x}}(y)\right]}{\sqrt{f_{\boldsymbol{x}}(y)}} \right| - \|K\|_2 \right) \leq z \right] = \exp\left(-e^{-z}\right).$$

It follows that

$$\left| \sqrt{\frac{mh_m}{\log h_m^{-1}}} \cdot \sup_{\boldsymbol{x}\in\mathcal{X}} \sup_{y\in[-b,b]} \frac{\tilde{f}_{\boldsymbol{x}}(y) - E\left[\tilde{f}_{\boldsymbol{x}}(y)\right]}{\sqrt{f_{\boldsymbol{x}}(y)}} - \|K\|_2 \right| = o_p(1). \tag{12}$$

Combining (12) with (11), the condition of Assumption 4 that $f_i(\cdot)$ is uniformly bounded on $\mathcal{B}$ and the condition that $n^{-1}\sum_{i=1}^{n} \|\boldsymbol{X}_i\|^2 = O_p(1)$ (implied by (2) of Assumption 1) yields the result that

$$\left\| \frac{1}{n} \sum_{i=1}^{n} V_{\tilde{f}_i}(\alpha) \boldsymbol{X}_i \boldsymbol{X}_i^\top - \boldsymbol{T}_{2nm}(\alpha) \right\| = o_p\left(\sqrt{\frac{-\log h_m}{mh_m}} \right). \tag{13}$$

Next, consider that for each $i \in \{1, \ldots, n\}$ and sufficiently small h_m that

$$E\left[\frac{a_2 - a_1}{h_m} K\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(U) - \boldsymbol{\beta}(\alpha)\right) \right) \middle| \boldsymbol{X}_i \right]$$
$$= \int_{a_1}^{a_2} \frac{1}{h_m} K\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(u) - \boldsymbol{\beta}(\alpha)\right) \right) du$$
$$= \int_{\boldsymbol{X}_i^\top \boldsymbol{\beta}(a_1)}^{\boldsymbol{X}_i^\top \boldsymbol{\beta}(a_2)} \frac{1}{h_m} K\left(\frac{1}{h_m} \left(t_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \right) f_i\left(t_i\right) dt_i$$

$$
\begin{aligned}
&= \int_{-1/2}^{1/2} K(w_i)\, f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) + h_m w_i\right) dw_i \\
&= f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \int_{-1/2}^{1/2} K(w_i)\, dw_i + \frac{h_m^2}{2} f_i^{(2)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \int_{-1/2}^{1/2} w_i^2 K(w_i)\, dw_i \\
&+ \frac{h_m^3}{6} f_i^{(3)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \int_{-1/2}^{1/2} w_i^3 K(w_i)\, dw_i + o_p\left(h_m^3\right)
\end{aligned}
$$

as $n \to \infty$, where we have exploited the conditions of Assumptions 2 and 4 that $\left|\int_{-1/2}^{1/2} w^4 K(w) dw\right| < \infty$ and $\max_{1 \le i \le n} \sup_{y \in \mathcal{B}} \left|f_i^{(4)}(y)\right| < \infty$ a.s.

It follows that

$$
\begin{aligned}
&\left\| \frac{1}{n} \sum_{i=1}^n B_{\tilde{f}_i}(\alpha) \boldsymbol{X}_i \boldsymbol{X}_i^\top - \boldsymbol{T}_{1nm}(\alpha) \right\| \\
&= \frac{h_m^3}{6} \int_{-1/2}^{1/2} w^3 K(w) dw \cdot \left\| \frac{1}{n} \sum_{i=1}^n f_i^{(3)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \boldsymbol{X}_i \boldsymbol{X}_i^\top \right\| + o_p\left(h_m^3\right) \\
&\le \frac{h_m^3}{6} \int_{-1/2}^{1/2} w^3 K(w) dw \cdot \max_{1 \le i \le n} \left| f_i^{(3)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) \right| \cdot \frac{1}{n} \sum_{i=1}^n \|\boldsymbol{X}_i\|^2 + o_p\left(h_m^3\right) \\
&= O_p\left(h_m^3\right) O_p(1) \\
&= o_p\left(h_m^2\right), \qquad (14)
\end{aligned}
$$

where we have additionally exploited the conditions of Assumptions 2 and 4 that $\int_{-1/2}^{1/2} w^3 K(w) dw < \infty$ and $\max_{1 \le i \le n} \sup_{y \in \mathcal{B}} \left|f_i^{(3)}(y)\right| < \infty$ a.s.

The desired conclusion follows from the combination of (8)–(10) and (13)–(14).

Claim 2. The following holds:

$$
\boldsymbol{T}_{2nm}(\alpha) = O_p\left(\frac{1}{\sqrt{n}}\right)
$$

By Lemma 1 applied with $\varphi(\cdot) = (a_2 - a_1) K^{(1)}(\cdot)$,

$$
(a_2 - a_1) \cdot \sqrt{\frac{m h_m^3}{\log h_m^{-1}}}
$$

$$
\cdot \left| \frac{1}{mh_m^2} \sum_{j=1}^{m} \left| K^{(1)} \left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha) \right) \right) \right| - E\left[\frac{1}{h_m^2} \left| K^{(1)} \left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(U) - \boldsymbol{\beta}(\alpha) \right) \right) \right| \,\middle|\, \boldsymbol{X}_i \right] \right|
$$
$$
< \infty, \tag{15}
$$

a.s. as $n \to \infty$.

Let $\mathcal{R}_+$ and $\mathcal{R}_-$ denote the regions $\{t \in [\boldsymbol{X}_i^\top \boldsymbol{\beta}(a_1), \boldsymbol{X}_i^\top \boldsymbol{\beta}(a_2)] : K^{(1)}\left(\left(t - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)/h_m\right) > 0\}$ and $\{t \in [\boldsymbol{X}_i^\top \boldsymbol{\beta}(a_1), \boldsymbol{X}_i^\top \boldsymbol{\beta}(a_2)] : K^{(1)}\left(\left(t - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)/h_m\right) < 0\}$, respectively. Then using integration by parts and applying the assumption that $K(\cdot)$ has bounded support, we have for a sufficiently small h_m that

$$
\begin{aligned}
& E\left[\frac{a_2 - a_1}{h_m^2} \left| K^{(1)} \left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(U) - \boldsymbol{\beta}(\alpha) \right) \right) \right| \,\middle|\, \boldsymbol{X}_i \right] \\
&= \int_{a_1}^{a_2} \frac{1}{h_m^2} \left| K^{(1)} \left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(u) - \boldsymbol{\beta}(\alpha) \right) \right) \right| du \\
&= \int_{\boldsymbol{X}_i^\top \boldsymbol{\beta}(a_1)}^{\boldsymbol{X}_i^\top \boldsymbol{\beta}(a_2)} \frac{1}{h_m^2} \left| K^{(1)} \left(\frac{1}{h_m} \left(t_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right) \right| f_i\left(t_i\right) dt_i \\
&= \frac{-1}{h_m} \left(\int_{\mathcal{R}_+} K \left(\frac{1}{h_m} \left(t_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right) f_i^{(1)}\left(t_i\right) dt_i - \int_{\mathcal{R}_-} K \left(\frac{1}{h_m} \left(t_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right) f_i^{(1)}\left(t_i\right) dt_i \right) \\
&= - \int_{\{K^{(1)}>0\}} K\left(w_i\right) f_i^{(1)} \left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) + h_m w_i \right) dw_i + \int_{\{K^{(1)}<0\}} K\left(w_i\right) f_i^{(1)} \left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) + h_m w_i \right) dw_i \\
&= f_i^{(1)} \left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \left(\int_{\{K^{(1)}<0\}} K(w) dw - \int_{\{K^{(1)}>0\}} K(w) dw \right) \\
&+ h_m f_i^{(2)} \left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \left(\int_{\{K^{(1)}<0\}} wK(w) dw - \int_{\{K^{(1)}>0\}} wK(w) dw \right) + O_p\left(h_m^2\right) \\
&\le \left| f_i^{(1)} \left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right| \left| \int_{\{K^{(1)}<0\}} K(w) dw - \int_{\{K^{(1)}>0\}} K(w) dw \right| \\
&+ h_m \left| f_i^{(2)} \left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right| \cdot \left| \int_{\{K^{(1)}<0\}} wK(w) dw - \int_{\{K^{(1)}>0\}} wK(w) dw \right| + O_p\left(h_m^2\right),
\end{aligned} \tag{16}
$$

where the assumption that $\max_i \sup_{y \in \mathcal{B}} \left| f_i^{(3)}(y) \right| < \infty$ a.s. (Assumption 4) has also been used.

Combine (15) and (16) to deduce that

$$
\begin{aligned}
&\max_i \frac{a_2 - a_1}{m h_m^2} \sum_{j=1}^{m} \left| K^{(1)}\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha) \right) \right) \right| \\
&\leq \max_i \left| \frac{a_2 - a_1}{m h_m^2} \sum_{j=1}^{m} \left| K^{(1)}\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha) \right) \right) \right| - E\left[\frac{a_2 - a_1}{h_m^2} \left| K^{(1)}\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(U) - \boldsymbol{\beta}(\alpha) \right) \right) \right| \middle| \boldsymbol{X}_i \right] \right| \\
&+ \max_i E\left[\frac{a_2 - a_1}{h_m^2} \left| K^{(1)}\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(U) - \boldsymbol{\beta}(\alpha) \right) \right) \right| \middle| \boldsymbol{X}_i \right] \\
&= O_p\left(\sqrt{\frac{-\log h_m}{m h_m^3}} \right) + \max_i \left| f_i^{(1)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right| \left| \int_{\{K^{(1)}<0\}} K(w) dw - \int_{\{K^{(1)}>0\}} K(w) dw \right| + O_p\left(h_m\right) \\
&= \max_i \sup_{\alpha \in \mathcal{A}} \left| f_i^{(1)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right| \left| \int_{\{K^{(1)}<0\}} K(w) dw - \int_{\{K^{(1)}>0\}} K(w) dw \right| + O_p\left(\sqrt{\frac{-\log h_m}{m h_m^3}} + h_m \right),
\end{aligned} \tag{17}
$$

where the assumptions that $\max_i \sup_{y \in \mathcal{B}} f_i(y) < \infty$ a.s. (Assumption 4), $\int_{-1/2}^{1/2} \left(K^{(1)}(w) \right)^2 dw < \infty$ (Assumption 2) and $\max_i \sup_{y \in \mathcal{B}} \left| f_i^{(3)}(y) \right| < \infty$ (Assumption 4) have been used.

Therefore

$$
\begin{aligned}
&\max_i \left| \frac{a_2 - a_1}{m h_m^2} \sum_{j=1}^{m} \frac{1}{\sqrt{n}} \boldsymbol{D}_{nj}(\alpha) K^{(1)}\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha) \right) \right) \right| \\
&\leq \max_i \frac{a_2 - a_1}{m h_m^2} \sum_{j=1}^{m} \left| K^{(1)}\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha) \right) \right) \right| \cdot \left[\frac{1}{\sqrt{n}} \max_j \left\| \boldsymbol{D}_{nj}(\alpha) \right\| \right] \\
&\leq \left(\max_i \left| f_i^{(1)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha) \right) \right| \cdot \left| \int_{\{K^{(1)}<0\}} K(w) dw - \int_{\{K^{(1)}>0\}} K(w) dw \right| + O_p\left(\sqrt{\frac{-\log h_m}{m h_m^3}} + h_m \right) \right) \\
&\cdot O_p\left(\frac{1}{\sqrt{n}} \right) \\
&= O_p\left(\frac{1}{\sqrt{n}} \right),
\end{aligned} \tag{18}
$$

where we have used (17), the condition of Assumption 4 that $\max_i \sup_{y\in\mathcal{B}} \left|f_i^{(1)}(y)\right| < \infty$ and the result (e.g., Angrist et al., 2006, Theorem 3) that

$$\begin{aligned}\max_j \left\|\boldsymbol{D}_{nj}(\alpha)\right\| &\leq 2\sup_{\alpha\in\mathcal{A}}\left\|\sqrt{n}\left(\hat{\boldsymbol{\beta}}_n(\alpha)-\boldsymbol{\beta}(\alpha)\right)\right\| \\ &= O_p(1).\end{aligned}$$

Applying (18) we have

$$\begin{aligned}&\left\|\frac{a_2-a_1}{n^{\frac{3}{2}}mh_m^2}\sum_{i=1}^{n}\sum_{j=1}^{m}\boldsymbol{X}_i^\top\left(\frac{1}{\sqrt{n}}\boldsymbol{D}_{nj}(\alpha)\right)K^{(1)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top\right\| \\ &\leq \max_i\left|\frac{a_2-a_1}{mh_m^2}\sum_{j=1}^{m}\frac{1}{\sqrt{n}}\boldsymbol{D}_{nj}(\alpha)K^{(1)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right)\right|\cdot\frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{X}_i\|^3 \\ &= O_p\left(\frac{1}{\sqrt{n}}\right)O_p(1) \\ &= O_p\left(\frac{1}{\sqrt{n}}\right),\end{aligned}$$

where we have also applied condition (2) of Assumption 1. The desired conclusion follows.

Claim 3: The following holds:

$$\begin{aligned}&\frac{a_2-a_1}{nmh_m^3}\sum_{i=1}^{n}\boldsymbol{X}_i^\top\left[\sum_{j=1}^{m}\frac{1}{n}\boldsymbol{D}_{nj}(\alpha)\boldsymbol{D}_{nj}(\alpha)^\top K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\bar{\boldsymbol{\beta}}\left(U_j\right)-\bar{\boldsymbol{\beta}}(\alpha)\right)\right)\right]\boldsymbol{X}_i\cdot\boldsymbol{X}_i\boldsymbol{X}_i^\top \\ &= O_p\left(\frac{1}{n}+\frac{1}{n^{\frac{3}{2}}h_m^4}\right).\end{aligned}$$

We first show that

$$\frac{a_2-a_1}{nmh_m^3}\sum_{i=1}^{n}\boldsymbol{X}_i^\top\left[\sum_{j=1}^{m}\frac{1}{n}\boldsymbol{D}_{nj}(\alpha)\boldsymbol{D}_{nj}(\alpha)^\top K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top\left(\boldsymbol{\beta}\left(U_j\right)-\boldsymbol{\beta}(\alpha)\right)\right)\right]\boldsymbol{X}_i\cdot\boldsymbol{X}_i\boldsymbol{X}_i^\top$$

$$= O_p\left(\frac{1}{n}\right). \tag{19}$$

In this connection, Lemma 1 applied with $\varphi(\cdot) = (a_2 - a_1)\, K^{(2)}(\cdot)$ yields

$$(a_2 - a_1) \cdot \lim_{m\to\infty} \sqrt{\frac{mh_m^5}{\log h_m^{-1}}}$$
$$\cdot \left| \frac{1}{mh_m^3} \sum_{j=1}^{m} K^{(2)}\left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha)\right)\right) - E\left[\frac{1}{h_m^3} K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(U) - \boldsymbol{\beta}(\alpha)\right)\right) \middle| \boldsymbol{X}_i\right]\right| < \infty \tag{20}$$

almost surely, where we have exploited the assumption that $\int_{-1/2}^{1/2} \left(K^{(2)}(w)\right)^2 dw < \infty$. In addition, by a derivation similar to that leading to (16) above, we have for all $\alpha \in \mathcal{A}$ that

$$E\left[\frac{a_2 - a_1}{h_m^3} \left| K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top \left(\boldsymbol{\beta}(U) - \boldsymbol{\beta}(\alpha)\right)\right)\right| \middle| \boldsymbol{X}_i\right] < \infty \tag{21}$$

wp. 1. Combine (20) and (21) to deduce (19), to wit:

$$\left\| \frac{a_2 - a_1}{nmh_m^3} \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{1}{n} \boldsymbol{X}_i^\top \boldsymbol{D}_{nj}(\alpha) \boldsymbol{D}_{nj}(\alpha)^\top \boldsymbol{X}_i K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha)\right)\right) \boldsymbol{X}_i \boldsymbol{X}_i^\top \right\|$$
$$\leq \frac{a_2 - a_1}{n} \cdot \max_j \left\|\boldsymbol{D}_{nj}(\alpha)\right\|^2$$
$$\cdot \max_i \frac{1}{mh_m^3} \sum_{j=1}^{m} \left| K^{(2)}\left(\frac{1}{h_m}\boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha)\right)\right)\right| \cdot \frac{1}{n} \sum_{i=1}^{n} \|\boldsymbol{X}_i\|^4$$
$$= O_p\left(\frac{1}{n}\right) O_p(1)$$
$$= O_p\left(\frac{1}{n}\right),$$

where we have applied the assumption that $\max_i \sup_{y\in\mathcal{B}} \left| f_i^{(2)}(y)\right| < \infty$ (Assumption 4) and the condition (2) of Assumption 1.

Next, we show that

$$\frac{a_2 - a_1}{n^2 m h_m^3} \sum_{i=1}^{n} \sum_{j=1}^{m} \left\| \boldsymbol{D}_{nj}(\alpha) \right\|^2 \left| K^{(2)} \left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha) \right) \right) - K^{(2)} \left(\frac{1}{h_m} \boldsymbol{X}_i^\top \left(\bar{\boldsymbol{\beta}}\left(U_j\right) - \bar{\boldsymbol{\beta}}(\alpha) \right) \right) \right| \|\boldsymbol{X}_i\|^4$$

$$= O_p \left(\frac{1}{n^{\frac{3}{2}} h_m^4} \right). \tag{22}$$

In particular, (22) follows directly from the mean value theorem, $\left\|K^{(3)}\right\|_\infty < \infty$, the result

$$\max_j \left\| \left(\bar{\boldsymbol{\beta}}\left(U_j\right) - \bar{\boldsymbol{\beta}}(\alpha)\right) - \left(\boldsymbol{\beta}\left(U_j\right) - \boldsymbol{\beta}(\alpha)\right) \right\| = O_p\left(\frac{1}{\sqrt{n}}\right)$$

(e.g., Angrist et al., 2006, Theorem 3) and condition (2) of Assumption 1.

The desired conclusion follows from (19) and (22).

B.3 Proof of Theorem 2

Begin by considering the following expansion of $\hat{\boldsymbol{W}}_n$ about $\boldsymbol{W}_n(\boldsymbol{G}_0(\alpha))$, which is a consequence of Theorem 1 appearing in the main body of our paper:

$$\hat{\boldsymbol{W}}_n = \boldsymbol{W}_n(\boldsymbol{G}_0(\alpha)) + \boldsymbol{U}_{1nm}(\alpha) + \boldsymbol{U}_{2nm}(\alpha) + \boldsymbol{U}_{3nm}(\alpha), \tag{23}$$

where $\hat{\boldsymbol{W}}_n$ is as given in (11) of our paper, and where we have

$$\boldsymbol{W}_n(\boldsymbol{G}_0(\alpha)) = \boldsymbol{W}_0 + O_p\left(n^{-\frac{1}{2}}\right) \tag{24}$$

by Assumption 1, where $\boldsymbol{W}_0$ is as given above in (12) of our paper. In addition, we have by the binomial inverse theorem,

$$\begin{aligned}
\boldsymbol{U}_{1nm}(\alpha) = \boldsymbol{W}_0 \cdot \Big[\boldsymbol{R} \Big(& \boldsymbol{G}_0^{-1}(\alpha) \boldsymbol{T}_{1nm}(\alpha) \boldsymbol{G}_0^{-1}(\alpha) \boldsymbol{H} \boldsymbol{G}_0^{-1}(\alpha) \\
& + \boldsymbol{G}_0^{-1}(\alpha) \boldsymbol{H} \boldsymbol{G}_0^{-1}(\alpha) \boldsymbol{T}_{1nm}(\alpha) \boldsymbol{G}_0^{-1}(\alpha) \Big) \boldsymbol{R}^\top \Big] \cdot \boldsymbol{W}_0 \\
& + \text{smaller-order terms} \\
& \equiv \bar{\boldsymbol{U}}_{1nm}(\alpha) + \text{smaller-order terms};
\end{aligned} \tag{25}$$

$$\boldsymbol{U}_{2nm}(\alpha) = \boldsymbol{W}_0 \cdot \Big[\boldsymbol{R} \Big(\boldsymbol{G}_0^{-1}(\alpha) \boldsymbol{T}_{2nm}(\alpha) \boldsymbol{G}_0^{-1}(\alpha) \boldsymbol{H} \boldsymbol{G}_0^{-1}(\alpha)$$

$$
\begin{aligned}
&+\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{H}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{T}_{2nm}(\alpha)\boldsymbol{G}_0^{-1}(\alpha)\Big)\,\boldsymbol{R}^\top\Big]\cdot\boldsymbol{W}_0\\
&+\text{smaller-order terms}\\
&\equiv \bar{\boldsymbol{U}}_{2nm}(\alpha)+\text{smaller-order terms;} \qquad (26)
\end{aligned}
$$

and

$$
\begin{aligned}
\boldsymbol{U}_{3nm}(\alpha) = \boldsymbol{W}_0\cdot\Big[\boldsymbol{R}\Big(&\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{T}_{2nm}(\alpha)\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{H}\boldsymbol{G}_0^{-1}(\alpha)\\
&+\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{H}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{T}_{3nm}(\alpha)\boldsymbol{G}_0^{-1}(\alpha)\Big)\,\boldsymbol{R}^\top\Big]\cdot\boldsymbol{W}_0\\
&+\text{smaller-order terms}\\
&\equiv \bar{\boldsymbol{U}}_{3nm}(\alpha)+\text{smaller-order terms.} \qquad (27)
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\boldsymbol{U}_{1nm}(\alpha) &= O_p\left(\boldsymbol{T}_{1nm}(\alpha)\right) = O_p\left(h_m^2\right),\\
\boldsymbol{U}_{2nm}(\alpha) &= O_p\left(\boldsymbol{T}_{2nm}(\alpha)\right) = O_p\left(\sqrt{\frac{-\log h_m}{mh_m}}\right)
\end{aligned}
$$

and that

$$
\boldsymbol{U}_{3nm}(\alpha) = O_p\left(\boldsymbol{T}_{2nm}(\alpha)\right) = O_p\left(\frac{1}{\sqrt{n}}\right).
$$

Combining (23)–(27) we find that for each $\alpha\in\mathcal{A}$,

$$
\begin{aligned}
\hat{\boldsymbol{W}}_n &= \boldsymbol{W}_0+\bar{\boldsymbol{U}}_{1nm}(\alpha)+\bar{\boldsymbol{U}}_{2nm}(\alpha)+\bar{\boldsymbol{U}}_{3nm}(\alpha)\\
&\quad+\text{smaller-order terms.} \qquad (28)
\end{aligned}
$$

Next, consider that if $H_0:\ \boldsymbol{R}\boldsymbol{\beta}(\alpha)-\boldsymbol{r}=\boldsymbol{0}$ is true, then the first-order asymptotic approximation for $\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha)$ has the form

$$
\sqrt{n}\left(\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha)-\boldsymbol{r}\right) = \boldsymbol{R}(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n}), \qquad (29)
$$

where $\boldsymbol{Z}(\alpha)\sim N(\boldsymbol{0},\boldsymbol{V}(\alpha))$ where $\boldsymbol{V}(\alpha)$ is as given above in (2) of our paper, and $\boldsymbol{S}_{1n}=o_p(1)$ as $n\to\infty$.

We can now consider the empirical size of a nominal level-τ Wald test of $H_0:\ \boldsymbol{R}\boldsymbol{\beta}(\alpha)-\boldsymbol{r}=\boldsymbol{0}$ incorporating the proposed estimator $\hat{\boldsymbol{G}}_n(\alpha)$ of $\boldsymbol{G}_0(\alpha)$. In particular, the representation appearing in (13) of our paper follows from the representations in (28) and (29). Let $\boldsymbol{S}_{2nm}$ denote the $o_p\left(h_m^2 + \sqrt{\log h_m^{-1}/(mh_m)} + n^{-1/2}\right)$ remainder term in (28). Then

$$\begin{aligned}
&\pi_0(h_m) \\
&\equiv \Pr\left[\frac{n}{\alpha(1-\alpha)}\left(\hat{\boldsymbol{\beta}}_n^\top(\alpha)\boldsymbol{R}^\top - \boldsymbol{r}^\top\right)\hat{\boldsymbol{W}}_n\left(\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{r}\right) > \chi^2_{J,\tau}\middle| H_0\right] \\
&= \Pr\left[\frac{1}{\alpha(1-\alpha)}\left[(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n})^\top\boldsymbol{R}^\top\left(\boldsymbol{W}_0 + \bar{\boldsymbol{U}}_{1nm}(\alpha) + \bar{\boldsymbol{U}}_{2nm}(\alpha) + \bar{\boldsymbol{U}}_{3nm}(\alpha)\right)\right.\right. \\
&\cdot\left.\left.\boldsymbol{R}\left(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n}\right) + (\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n})^\top\boldsymbol{R}^\top\boldsymbol{S}_{2nm}\boldsymbol{R}\left(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n}\right)\right] > \chi^2_{J,\tau}\right] \\
&= \Pr\left[\frac{1}{\alpha(1-\alpha)}(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n})^\top\boldsymbol{R}^\top\left(\boldsymbol{W}_0 + \bar{\boldsymbol{U}}_{1nm}(\alpha) + \bar{\boldsymbol{U}}_{2nm}(\alpha) + \bar{\boldsymbol{U}}_{3nm}(\alpha)\right)\right. \\
&\cdot\left.\boldsymbol{R}\left(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n}\right) > \chi^2_{J,\tau} - \Xi_{nm}(0)\right] \\
&= \Pr\left[\frac{1}{\alpha(1-\alpha)}\left(\boldsymbol{Z}(\alpha)^\top\boldsymbol{R}^\top\boldsymbol{W}_0\boldsymbol{R}\boldsymbol{Z}(\alpha) + \boldsymbol{Z}(\alpha)^\top\boldsymbol{R}^\top\boldsymbol{W}_0\boldsymbol{R}\boldsymbol{S}_{1n} + \boldsymbol{S}_{1n}^\top\boldsymbol{R}^\top\boldsymbol{W}_0\boldsymbol{R}\boldsymbol{Z}(\alpha) + \boldsymbol{S}_{1n}^\top\boldsymbol{R}^\top\boldsymbol{W}_0\boldsymbol{R}\boldsymbol{S}_{1n}\right.\right. \\
&\left.\left.+h_m^2\Lambda_{1nm}(\alpha,0) + \sqrt{\frac{-\log h_m}{mh_m}}\Lambda_{2nm}(\alpha,0) + \frac{1}{\sqrt{n}}\Lambda_{3nm}(\alpha,0)\right) > \chi^2_{J,\tau} - \Xi_{nm}(0)\right] \\
&= \Pr\left[\frac{1}{\alpha(1-\alpha)}\boldsymbol{Z}(\alpha)^\top\boldsymbol{R}^\top\boldsymbol{W}_0^{-1}\boldsymbol{R}\boldsymbol{Z}(\alpha) > \chi^2_{J,\tau}\right. \\
&\left.-\frac{1}{\alpha(1-\alpha)}\left(h_m^2\Lambda_{1nm}(\alpha,0) + \sqrt{\frac{-\log h_m}{mh_m}}\Lambda_{2nm}(\alpha,0) + \frac{1}{\sqrt{n}}\Lambda_{3nm}(\alpha,0)\right) - \Theta_n(0) - \Xi_{nm}(0)\right],
\end{aligned} \tag{30}$$

where

$$\Xi_{nm}(0) = (\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n})^\top\boldsymbol{R}^\top\boldsymbol{S}_{2nm}\boldsymbol{R}(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n}). \tag{31}$$

In addition, the quantities $\Lambda_{1nm}(\alpha,0)$, $\Lambda_{2nm}(\alpha,0)$ and $\Lambda_{3nm}(\alpha,0)$ appearing in (30) are given by

$$\Lambda_{1nm}(\alpha,0) = \frac{1}{h_m^2}(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n})^\top\boldsymbol{R}^\top\bar{\boldsymbol{U}}_{1nm}(\alpha)\boldsymbol{R}(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n}), \tag{32}$$

$$\Lambda_{2nm}(\alpha,0) = \sqrt{\frac{mh_m}{-\log h_m}}(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n})^\top\boldsymbol{R}^\top\bar{\boldsymbol{U}}_{2nm}(\alpha)\boldsymbol{R}(\boldsymbol{Z}(\alpha)+\boldsymbol{S}_{1n}) \tag{33}$$

and

$$\Lambda_{3nm}(\alpha, 0) = \sqrt{n}(\boldsymbol{Z}(\alpha) + \boldsymbol{S}_{1n})^\top \boldsymbol{R}^\top \bar{\boldsymbol{U}}_{3nm}(\alpha)\boldsymbol{R}(\boldsymbol{Z}(\alpha) + \boldsymbol{S}_{1n}), \tag{34}$$

while

$$\Theta_n(0) = \frac{1}{\alpha(1-\alpha)} \left(2\boldsymbol{Z}(\alpha)^\top \boldsymbol{R}^\top \boldsymbol{W}_0 \boldsymbol{R}\boldsymbol{S}_{1n} + \boldsymbol{S}_{1n}^\top \boldsymbol{R}^\top \boldsymbol{W}_0 \boldsymbol{R}\boldsymbol{S}_{1n}\right). \tag{35}$$

This establishes the representation of the size function given in (13) of our paper.

The remainder of Theorem 2 follows straightforwardly from the expression for the empirical size function given by $\pi_0(h_m)$ in (30) and the observation that if $G_0(\alpha)$ does not need to be estimated, the infeasible level-τ Wald test of $H_0 : \boldsymbol{R}\boldsymbol{\beta}(\alpha) - \boldsymbol{r} = \boldsymbol{0}$ has an empirical size function given by

$$\begin{aligned} &\Pr\left[\frac{n}{\alpha(1-\alpha)}\left(\hat{\boldsymbol{\beta}}_n(\alpha)^\top \boldsymbol{R}^\top - r^\top\right)\boldsymbol{W}_n(\boldsymbol{G}_0)\left(\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{r}\right) > \chi^2_{J,\tau}\middle| H_0\right] \\ &= \Pr\left[\frac{1}{\alpha(1-\alpha)}\boldsymbol{Z}(\alpha)^\top \boldsymbol{R}^\top \boldsymbol{W}_0 \boldsymbol{R}\boldsymbol{Z}(\alpha) > \chi^2_{J,\tau} - \Theta_n(0)\right]. \end{aligned}$$

C Analysis of Wald-type Tests Implemented Using the Estimators of Powell (1991) and Hendricks and Koenker (1992)

We show in this appendix that the well-known estimators of $\boldsymbol{G}_0(\alpha)$ proposed by Powell (1991) and Hendricks and Koenker (1992) for the express purpose of quantile-regression inference cannot actually be used to generate Wald-type tests that control size adaptively in large samples. For $i = 1, \ldots, n$ and α a fixed quantile in $\mathcal{A}$, let

$$\hat{\mathbf{G}}_n^P(\alpha) \equiv \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{Y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_n(\alpha)}{h_n}\right)\boldsymbol{X}_i\mathbf{X}_i^\top \tag{36}$$

denote the kernel-based estimator of $\mathbf{G}_0(\alpha)$ proposed by Powell (1991), where in this case $K(\cdot)$ is taken to denote a smoothing kernel satisfying the conditions of Assumption 2 above, while the bandwidth h_n is assumed to satisfy the constraints $h_n \to 0$ and $nh_n^3 \to \infty$ as $n \to \infty$.

In addition, let

$$\hat{f}_{ni}^{HK}(\alpha) \equiv \frac{2h_n}{\mathbf{X}_i^\top\left(\hat{\boldsymbol{\beta}}_n(\alpha + h_n) - \hat{\boldsymbol{\beta}}_n(\alpha - h_n)\right)} \tag{37}$$

denote the difference-quotient estimator of $f_i\left(\boldsymbol{X}_i^\top\beta(\alpha)\right)$ suggested by Hendricks and Koenker (1992). Let

$$\hat{\boldsymbol{G}}_n^{HK}(\alpha) \equiv \frac{1}{n}\sum_{i=1}^{n}\hat{f}_{ni}^{HK}(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top \tag{38}$$

denote the corresponding estimator of $\mathbf{G}_0(\alpha)$.

We establish that optimal implementations of $\hat{\mathbf{G}}_n^P(\alpha)$ or $\hat{\mathbf{G}}_n^{HK}(\alpha)$ from the point of view of maximizing the rate of decay of the empirical size distortion of a Wald-type test of H_0 : $\mathbf{R}\beta(\alpha) - \mathbf{r} = \mathbf{0}$ are still sub-optimal in the sense that the resulting tests will exhibit size distortions that decay at rates that are strictly slower than the $O_p\left(n^{-1/2}(\log n)^{3/2}\right)$ adaptive rate. Similarly, a Wald-type confidence interval for a given linear combination of $\beta(\alpha)$ constructed using $\hat{\mathbf{G}}_n^P(\alpha)$ or $\hat{\mathbf{G}}_n^{HK}(\alpha)$ will not exhibit a level error that decays at the $O_p\left(n^{-1}(\log n)^3\right)$ adaptive rate.

Theorem 3. *Suppose the validity of Assumptions 1, 4 and 5 as given in Appendix A above. Let $\{h_n\}$ denote a bandwidth sequence in which $h_n \longrightarrow 0$ and $nh_n^3 \longrightarrow \infty$ as $n \longrightarrow \infty$. We have the following as $n \longrightarrow \infty$:*

1. *Suppose $K(\cdot)$ is a smoothing kernel satisfying the conditions of Assumption 2. Then the magnitude of the empirical size distortion of a Wald-type test of H_0 : $\mathbf{R}\beta(\alpha) - \mathbf{r} = \mathbf{0}$ in which the estimator $\hat{\mathbf{G}}_n^P(\alpha)$ is embedded can be no smaller than $n^{-2/5}$-order. This rate of convergence is attained when $h_n \propto n^{-1/5}$. In addition, the level error of a Wald-type confidence interval for a linear combination of the elements of $\beta(\alpha)$ that incorporates $\hat{\mathbf{G}}_n^P(\alpha)$ can be no smaller than $n^{-4/5}$-order, a magnitude attained when $h_n \propto n^{-1/5}$.*

2. *The magnitude of the empirical size distortion of a test of H_0 : $\mathbf{R}\beta(\alpha) - \mathbf{r} = \mathbf{0}$ based on (2) as given in the main body of the paper and in which the estimator $\hat{\mathbf{G}}_n^{HK}(\alpha)$ is embedded can be no smaller than $n^{-2/7}$-order. This rate of convergence is attained when $h_n \propto n^{-1/7}$. In addition, the level error of a Wald-type confidence interval for a linear combination of the elements of $\beta(\alpha)$ that incorporates $\hat{\mathbf{G}}_n^{HK}(\alpha)$ can be no smaller than $n^{-4/7}$-order, a magnitude attained when $h_n \propto n^{-1/7}$.*

Proof. The proof appears below in Appendix C.1. □

C.1 Proof of Theorem 3

We consider the two conclusions of Theorem 3 in sequence.

First conclusion of Theorem 3:

By Taylor expansion, the following holds for each $\alpha \in \mathcal{A}$:

$$\begin{aligned}&\hat{\boldsymbol{G}}_n^P(\alpha)\\ &= \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{1}{h_n}\left(Y_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top - \frac{1}{n^{\frac{3}{2}}h_n^2}\sum_{i=1}^n K^{(1)}\left(\frac{1}{h_n}\bar{U}_i(\alpha)\right)\boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top,\end{aligned} \tag{39}$$

where $\left|\bar{U}_i(\alpha) - \left(Y_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right| < \left|n^{-1/2}\boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\right|$ for each $i = 1, \ldots, n$ and where

$$\boldsymbol{B}_n(\alpha) = \sqrt{n}\left(\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{\beta}(\alpha)\right).$$

Consider the first term in (39). Standard calculations show that for each $\alpha \in \mathcal{A}$,

$$\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{1}{h_n}\left(Y_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right)\boldsymbol{X}_i\boldsymbol{X}_i^\top = \boldsymbol{G}_0(\alpha) + O_p\left(h_n^2 + \frac{1}{\sqrt{nh_n}}\right). \tag{40}$$

Consider the second term in (39). In particular, for $U_i(\alpha) \equiv Y_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)$ we have

$$\begin{aligned}&\frac{1}{n^{\frac{3}{2}}h_n^2}\sum_{i=1}^n K^{(1)}\left(\frac{1}{h_n}\bar{U}_i(\alpha)\right)\boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top\\ &= \frac{1}{n^{\frac{3}{2}}h_n^2}\sum_{i=1}^n K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top\\ &+ \frac{1}{n^{\frac{3}{2}}h_n^2}\sum_{i=1}^n \left(K^{(1)}\left(\frac{1}{h_n}\bar{U}_i(\alpha)\right) - K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\right)\boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top.\end{aligned} \tag{41}$$

Standard calculations show that

$$\frac{1}{nh_n^2}\sum_{i=1}^n \left|K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\right| - E\left[\frac{1}{h_n^2}\left|K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\right|\middle| \boldsymbol{X}_i\right] = O_p\left(\frac{1}{\sqrt{nh_n^3}}\right), \tag{42}$$

while

$$E\left[\frac{1}{h_n^2}\left|K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\right|\middle| \boldsymbol{X}_i\right]$$

$$\leq \left| f_i^{(1)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right| \cdot \left| \int_{\{K^{(1)}<0\}} K(w)dw - \int_{\{K^{(1)}>0\}} K(w)dw \right|$$

$$+ h_n \left| f_i^{(2)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right| \cdot \left| \int_{\{K^{(1)}<0\}} wK(w)dw - \int_{\{K^{(1)}>0\}} wK(w)dw \right| + O_p\left(h_n^2\right) \tag{43}$$

as $n \to \infty$. Combining (42) with (43) we find that

$$\frac{1}{nh_n^2}\sum_{i=1}^n \left| K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\right|$$

$$\leq \max_i \left| f_i^{(1)}\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right)\right| \cdot \left| \int_{\{K^{(1)}<0\}} K(w)dw - \int_{\{K^{(1)}>0\}} K(w)dw \right|$$

$$+ O_p\left(\frac{1}{\sqrt{nh_n^3}} + h_n\right). \tag{44}$$

The asymptotic normality result given in (2) of our paper implies that $\boldsymbol{B}_n(\alpha) = O_p(1)$ for every $\alpha \in \mathcal{A}$. Combine this result with (44) above to deduce that the following holds for each $\alpha \in \mathcal{A}$:

$$\left\| \frac{1}{n^{\frac{3}{2}}h_n^2}\sum_{i=1}^n K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right) \boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top \right\|$$

$$\leq \frac{1}{\sqrt{n}} \max_i \|\boldsymbol{X}_i\|^3 \cdot \|\boldsymbol{B}_n(\alpha)\| \cdot \frac{1}{nh_n^2}\sum_{i=1}^n \left| K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\right|$$

$$= O_p\left(\frac{1}{\sqrt{n}}\right) \cdot O_p(1) \cdot O_p\left(1 + \frac{1}{\sqrt{nh_n^3}} + h_n\right)$$

$$= O_p\left(\frac{1}{\sqrt{n}}\right). \tag{45}$$

Next, note that

$$\frac{1}{n^{\frac{3}{2}}h_n^2}\sum_{i=1}^n \left(K^{(1)}\left(\frac{1}{h_n}\bar{U}_i(\alpha)\right) - K^{(1)}\left(\frac{1}{h_n}U_i(\alpha)\right)\right) \boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\boldsymbol{X}_i\boldsymbol{X}_i^\top .$$

$$= O_p\left(\frac{1}{n^{\frac{3}{2}}h_n^3}\right)$$

$$= o_p\left(\frac{1}{\sqrt{n}}\right), \tag{46}$$

a result that follows from an application of the mean value theorem, the assumption that $\left\|K^{(2)}\right\|_\infty < \infty$ and $\|\boldsymbol{X}_i\| < \infty$, (2) of our paper, the assumption that $nh_n^3 \to \infty$ and the result that

$$\begin{aligned}\max_i \left|\bar{U}_i(\alpha) - U_i(\alpha)\right| &< \max_i \left|n^{-\frac{1}{2}} \boldsymbol{X}_i^\top \boldsymbol{B}_n(\alpha)\right| \\ &\leq \frac{1}{n^{\frac{1}{2}}} \max_i \|\boldsymbol{X}_i\| \cdot O_p(1) \\ &= O_p\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

Combining (39), (40), (41), (45) and (46) we find that

$$\begin{aligned}\hat{\boldsymbol{G}}_n^P(\alpha) &= \boldsymbol{G}_0(\alpha) + O_p\left(h_n^2 + \frac{1}{\sqrt{nh_n}} + \frac{1}{\sqrt{n}}\right) \\ &= \boldsymbol{G}_0(\alpha) + O_p\left(h_n^2 + \frac{1}{\sqrt{nh_n}}\right) \qquad (47)\end{aligned}$$

The expansion of $\hat{\boldsymbol{G}}_n^P(\alpha)$ given in (47) can then be combined with the binomial inverse theorem to deduce an expansion of $\left[\boldsymbol{R}\left(\hat{\boldsymbol{G}}_n^P(\alpha)\right)^{-1} \boldsymbol{H}_n \left(\hat{\boldsymbol{G}}_n^P(\alpha)\right)^{-1} \boldsymbol{R}^\top\right]^{-1}$ about $\left(\boldsymbol{R}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{H}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{R}^\top\right)^{-1}$ of the form

$$\begin{aligned}&\left[\boldsymbol{R}\left(\hat{\boldsymbol{G}}_n^P(\alpha)\right)^{-1} \boldsymbol{H}_n \left(\hat{\boldsymbol{G}}_n^P(\alpha)\right)^{-1} \boldsymbol{R}^\top\right]^{-1} \\ &= \left(\boldsymbol{R}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{H}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{R}^\top\right)^{-1} + \boldsymbol{U}_{1n}^P(\alpha) + \boldsymbol{U}_{2n}^P(\alpha) + \boldsymbol{R}_{n2}^P(\alpha), \qquad (48)\end{aligned}$$

where for each $\alpha \in \mathcal{A}$, we have $\boldsymbol{U}_{1n}^P(\alpha) = O_p\left(h_n^2\right)$, $\boldsymbol{U}_{2n}^P(\alpha) = O_p\left((nh_n)^{-1/2}\right)$ and $\boldsymbol{R}_{n2}^P(\alpha) = o_p\left(h_n^2 + (nh_n)^{-1/2}\right)$ as $n \to \infty$.

The remainder of the proof of the first part of Theorem 3 follows arguments similar to those used in the proof of Theorem 2.

Second conclusion of Theorem 3:

Let $\alpha \in \mathcal{A} \equiv [a_1, a_2] \subset (0, 1)$ be a fixed quantile. Define

$$\boldsymbol{D}_n(\alpha) \equiv \sqrt{2nh_n}\left[\left(\hat{\boldsymbol{\beta}}_n(\alpha + h_n) - \boldsymbol{\beta}(\alpha + h_n)\right) - \left(\hat{\boldsymbol{\beta}}_n(\alpha - h_n) - \boldsymbol{\beta}(\alpha - h_n)\right)\right].$$

In particular, $\boldsymbol{D}_n(\alpha)$ denotes an instance of an appropriately normalized regression-quantile spacing local to $\left(\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{\beta}(\alpha)\right)$ whose asymptotic behavior is analyzed in Portnoy (2012). We make use of the following result derived from Portnoy (2012, Theorem 1, Ingredients 1–7 of the proof of Theorem 2):

Proposition 1 ((Portnoy, 2012))**.** *Suppose Assumptions 1, 4 and 5 hold. Then there exists a constant W such that for $\|\boldsymbol{D}_n(\alpha)\| \leq W\sqrt{\log n}$, the density of $\boldsymbol{D}_n(\alpha)$ at t satisfies*

$$f_{\boldsymbol{D}_n(\alpha)}(t) = \phi_{\boldsymbol{D}(\alpha)}(t)\left(1 + O\left(\sqrt{\frac{(\log n)^3}{2nh_n}}\right)\right),$$

where $\phi_{\boldsymbol{D}(\alpha)}(\cdot)$ denotes the density of a mean-zero Gaussian random vector $\boldsymbol{D}(\alpha)$ with covariance structure given in Portnoy (2012, eq. (7.3)).

In particular, Proposition 1 implies the existence of a mean-zero Gaussian random variable $\boldsymbol{D}(\alpha)$ such that

$$\|\boldsymbol{D}_n(\alpha) - \boldsymbol{D}(\alpha)\| = O_p\left(\sqrt{\frac{(\log n)^3}{nh_n}}\right). \tag{49}$$

Now consider the conditional density estimator $\hat{f}_{ni}^{HK}(\alpha)$ given in (37) above. For each $\alpha \in \mathcal{A}$ we have

$$\begin{aligned}
\hat{f}_{ni}^{HK}(\alpha) &= \frac{2h_n}{\boldsymbol{X}_i^\top \left[\frac{1}{\sqrt{2nh_n}}\boldsymbol{D}_n(\alpha) + \left(\boldsymbol{\beta}\left(\alpha + h_n\right) - \boldsymbol{\beta}\left(\alpha - h_n\right)\right)\right]} \\
&= \frac{1}{\boldsymbol{X}_i^\top \left[\frac{1}{2^{\frac{3}{2}}\sqrt{nh_n^3}}\boldsymbol{D}_n(\alpha) + \frac{1}{2h_n}\left(\boldsymbol{\beta}\left(\alpha + h_n\right) - \boldsymbol{\beta}\left(\alpha - h_n\right)\right)\right]} \\
&= \frac{1}{\frac{1}{2^{\frac{3}{2}}\sqrt{nh_n^3}} \cdot \boldsymbol{X}_i^\top \boldsymbol{D}_n(\alpha) + \frac{1}{f_i(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha))} + O_p\left(h_n^2\right)} \\
&= f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) + O_p\left(h_n^2 + \frac{(\log n)^{\frac{3}{2}}}{nh_n^2} + \frac{1}{\sqrt{nh_n^3}}\right) \\
&= f_i\left(\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)\right) + O_p\left(h_n^2 + \frac{1}{\sqrt{nh_n^3}}\right),
\end{aligned}$$

where (49) and Taylor expansions of $\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha + h_n)$ and $\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha - h_n)$ about $\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)$ have been applied. Arguments similar to those used in the proof of Theorem 1 can then be used to show that for each $\alpha \in \mathcal{A}$,

$$\hat{\boldsymbol{G}}_n^{HK}(\alpha) = \boldsymbol{G}_0(\alpha) + \boldsymbol{T}_{1n}^{HK}(\alpha) + \boldsymbol{T}_{2n}^{HK}(\alpha) + \boldsymbol{R}_{n1}^{HK}(\alpha), \tag{50}$$

where $\boldsymbol{T}_{1n}^{HK}(\alpha) = O_p\left(h_n^2\right)$, $\boldsymbol{T}_{2n}^{HK}(\alpha) = O_p\left(\left(nh_n^3\right)^{-1/2}\right)$ and $\boldsymbol{R}_{n1}^{HK}(\alpha) = o_p\left(h_n^2 + \left(nh_n^3\right)^{-1/2}\right)$. The expansion of $\hat{\boldsymbol{G}}_n^{HK}(\alpha)$ given in (50) can then be combined with the binomial inverse theorem to deduce an expansion of $\left[\boldsymbol{R}\left(\hat{\boldsymbol{G}}_n^{HK}(\alpha)\right)^{-1} \boldsymbol{H}_n \left(\hat{\boldsymbol{G}}_n^{HK}(\alpha)\right)^{-1} \boldsymbol{R}^\top\right]^{-1}$ about $\boldsymbol{W}_0$ of the form

$$\begin{aligned}
&\left[\boldsymbol{R}\left(\hat{\boldsymbol{G}}_n^{HK}(\alpha)\right)^{-1} \boldsymbol{H}_n \left(\hat{\boldsymbol{G}}_n^{HK}(\alpha)\right)^{-1} \boldsymbol{R}^\top\right]^{-1} \\
&= \left(\boldsymbol{R}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{H}\boldsymbol{G}_0^{-1}(\alpha)\boldsymbol{R}^\top\right)^{-1} + \boldsymbol{U}_{1n}^{HK}(\alpha) + \boldsymbol{U}_{2n}^{HK}(\alpha) + \boldsymbol{R}_{n2}^{HK}(\alpha),
\end{aligned} \tag{51}$$

where for each $\alpha \in \mathcal{A}$, we have $\boldsymbol{U}_{1n}^{HK}(\alpha) = O_p\left(h_n^2\right)$, $\boldsymbol{U}_{2n}^{HK}(\alpha) = O_p\left(\left(nh_n^3\right)^{-1/2}\right)$ and $\boldsymbol{R}_{n2}^{HK}(\alpha) = o_p\left(h_n^2 + \left(nh_n^3\right)^{-1/2}\right)$ as $n \to \infty$.

The remainder of the proof of the second conclusion of Theorem 3 follows arguments similar to those used in the proof of Theorem 2.

D A Data-Driven and Rate-Optimal Bandwidth

We present details regarding the derivation and estimation of a particular rate-optimal bandwidth usable in the implementation of our proposed estimate of $\boldsymbol{G}_0(\alpha)$. This bandwidth differs from the fixed bandwidth used in the simulations presented in Section 4 of the paper in that it is data-driven, i.e., it involves a leading constant that must be estimated. This bandwidth is nevertheless rate-optimal in that it decays at a rate such that any Wald-type test in which the corresponding estimate of $\boldsymbol{G}_0(\alpha)$ is embedded exhibits adaptive size control as $n \to \infty$. We present simulation evidence on the finite-sample performance of the resulting data-driven bandwidth in Appendix E.2 below. We also make use of our proposed data-driven bandwidth in the empirical analysis presented in Section 5 of our paper.

Assuming that the vector $\boldsymbol{\beta}(\alpha)$ of α-quantile regression coefficients is d-dimensional, let $\boldsymbol{R}$ denote a fully specified $(J \times d)$-matrix of rank J. In addition, let $\boldsymbol{r} \in \mathbb{R}^J$ be fully specified. Suppose that one wishes to test the hypothesis $H_0 : \boldsymbol{R}\boldsymbol{\beta}(\alpha) = \boldsymbol{r}$. Consider the size function given in (13) of our paper. Differentiating the size function with respect to h_m we find that the magnitude of the empirical size distortion is minimized by the solution to

$$h_m^5 = \frac{1}{16m}[\left(\log h_m^{-1}\right)^{\frac{1}{2}} - \left(\log h_m^{-1}\right)^{-\frac{1}{2}} h_m^2]^2 \left(\frac{\Lambda_{2nm}(\alpha,0)}{\Lambda_{1nm}(\alpha,0)}\right)^2, \tag{52}$$

where $\Lambda_{1nm}(\alpha, 0)$ and $\Lambda_{2nm}(\alpha, 0)$ are as given above in (32) and (33), respectively.

Note that

$$\frac{1}{16m}[\left(\log h_m^{-1}\right)^{\frac{1}{2}} - \left(\log h_m^{-1}\right)^{-\frac{1}{2}} h_m^2]^2 \left(\frac{\Lambda_{2nm}(\alpha,0)}{\Lambda_{1nm}(\alpha,0)}\right)^2 \approx \frac{1}{m}\log h_m^{-1}\left(\frac{\Lambda_{2nm}(\alpha,0)}{4\Lambda_{1nm}(\alpha,0)}\right)^2$$

when m is large. It follows that for large m an approximate solution to (52) is given implicitly by the relation

$$\frac{h_m^5}{\log h_m^{-1}} = \left(\frac{\Lambda_{2nm}(\alpha,0)}{4\Lambda_{1nm}(\alpha,0)}\right)^2 \frac{1}{m},$$

which implies that for large m, the optimal value of h_m has the form

$$h_m^* = \kappa\left[\left(\frac{\Lambda_{2nm}(\alpha,0)}{\Lambda_{1nm}(\alpha,0)}\right)^2 \cdot \frac{\log m}{m}\right]^{\frac{1}{5}}, \tag{53}$$

where $\kappa > 0$ is a proportionality constant. Experimentation with simulations involving various settings of κ suggest that the choice $\kappa = 1$ works well in practice.

We present in this connection a plug-in estimate of the optimal bandwidth h_m^* given in (53) with $\kappa = 1$. In particular, we show how one might estimate the unknown quantities in the leading constant appearing in the expression for h_m^*. Let $\hat{\boldsymbol{\beta}}_n(\alpha)$ denote the regression α-quantile based on a random sample of observations given by $(\boldsymbol{X}_1^\top, Y_1), \ldots, (\boldsymbol{X}_n^\top, Y_n)$. Let $\boldsymbol{X}$ denote the $(n \times d)$ matrix of regressors whose ith row is given by $\boldsymbol{X}_i^\top$ $(i = 1, \ldots, n)$. We propose to estimate the optimal bandwidth given in (53) under the setting $\kappa = 1$ by

$$\hat{h}_m^* \equiv \left[\left(\frac{\hat{\Lambda}_{2nm}(\alpha,0)}{\hat{\Lambda}_{1nm}(\alpha,0)}\right)^2 \cdot \frac{\log m}{m}\right]^{\frac{1}{5}}, \tag{54}$$

where

$$\hat{\Lambda}_{1nm}(\alpha,0) = (\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{r})^\top \hat{\boldsymbol{U}}_{1nm}(\alpha)(\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{r});$$

$$\hat{\Lambda}_{2nm}(\alpha,0) = (\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{r})^\top \hat{\boldsymbol{U}}_{2nm}(\alpha)(\boldsymbol{R}\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{r}),$$

and where the quantities $\hat{\boldsymbol{U}}_{1nm}(\alpha)$ and $\hat{\boldsymbol{U}}_{2nm}(\alpha)$ are given by

$$\begin{aligned}\hat{\boldsymbol{U}}_{1nm}(\alpha) = (\boldsymbol{R}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{R}^\top)^{-1}\cdot\Big(&\boldsymbol{R}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\hat{\boldsymbol{T}}_{1nm}(\alpha)\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\\ &+\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\hat{\boldsymbol{T}}_{1nm}(\alpha)\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{R}^\top\Big)(\boldsymbol{R}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{R}^\top)^{-1}\end{aligned} \tag{55}$$

and

$$\begin{aligned}\hat{\boldsymbol{U}}_{2nm}(\alpha) = (\boldsymbol{R}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{R}^\top)^{-1}\Big(&\boldsymbol{R}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\hat{\boldsymbol{T}}_{2nm}(\alpha)\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\\ &+\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\hat{\boldsymbol{T}}_{2nm}(\alpha)\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{R}^\top\Big)(\boldsymbol{R}\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{X}^\top\boldsymbol{X})\tilde{\boldsymbol{G}}_n(\alpha)^{-1}\boldsymbol{R}^\top)^{-1},\end{aligned} \tag{56}$$

respectively, where $\tilde{\boldsymbol{G}}_n(\alpha)$ denotes our proposed estimate of the matrix $\boldsymbol{G}_0(\alpha)$ given by $\hat{\boldsymbol{G}}_n(\alpha)$ in (6) of the main body of our paper, but implemented with the preliminary bandwidth

$$\tilde{h}_{m1} \equiv m^{-\frac{1}{5}}. \tag{57}$$

The quantities $\hat{\boldsymbol{T}}_{1nm}(\alpha)$ and $\hat{\boldsymbol{T}}_{2nm}(\alpha)$ appearing in (55) and (56), respectively, are given by

$$\hat{\boldsymbol{T}}_{1nm}(\alpha) = \frac{1}{n}\boldsymbol{X}^\top\tilde{\boldsymbol{f}}''_{nm}(\alpha)\boldsymbol{X}$$

and

$$\hat{\boldsymbol{T}}_{2nm}(\alpha) = \frac{1}{n}\boldsymbol{X}^\top\tilde{\boldsymbol{S}}_{nm}(\alpha)\boldsymbol{X}, \tag{58}$$

where $\tilde{\boldsymbol{f}}''_{nm}(\alpha)$ denotes the diagonal $(n\times n)$-matrix whose ith diagonal element is a kernel estimate of the second derivative of the conditional response density at the point $\boldsymbol{X}_i^\top\boldsymbol{\beta}(\alpha)$ given by

$$\tilde{\boldsymbol{f}}^2_{ni}(\alpha) \equiv \frac{1}{m\tilde{h}^3_{m2}}\sum_{j=1}^m K^{(2)}\left(\frac{1}{\tilde{h}_{m2}}\boldsymbol{X}_i^\top(\hat{\boldsymbol{\beta}}_n(U_j) - \hat{\boldsymbol{\beta}}_n(\alpha))\right),$$

where $U_1,\ldots,U_m$ denote a random sample of $Unif[a_1,a_2]$-variates generated by Monte Carlo, and where $K(\cdot)$ denotes the standard Gaussian kernel, $K^{(2)}(\cdot)$ denotes its second derivative and

$\tilde{h}_{m2} = m^{-1/9}$. Similarly, the quantity $\tilde{\boldsymbol{S}}_{nm}(\alpha)$ appearing in (58) is the diagonal $(n \times n)$-matrix whose ith diagonal element is the square root of a kernel estimate of the conditional response density at the point $\boldsymbol{X}_i^\top \boldsymbol{\beta}(\alpha)$ given by

$$\boldsymbol{S}_{ni}(\alpha) \equiv \sqrt{\frac{1}{m\tilde{h}_{m1}} \sum_{j=1}^{m} K\left(\frac{1}{\tilde{h}_{m1}} \boldsymbol{X}_i^\top (\hat{\boldsymbol{\beta}}_n(U_j) - \hat{\boldsymbol{\beta}}_n(\alpha))\right)},$$

where again $U_1, \ldots, U_m$ is a simulated random sample of $Unif[a_1, a_2]$-variates and $K(\cdot)$ is the standard Gaussian kernel. The bandwidth $\tilde{h}_{m1}$ is as given above in (57).

E Further Numerical Evidence

This appendix contains additional simulation evidence in the context of the same family of data-generating processes described in Section 4 of the main text. These additional simulations also consider the size and power performance of tests of the same hypothesis of no treatment-effect heterogeneity described in the paper. Samples of sizes $n = 100$ and $n = 300$, generated by 1000 Monte Carlo replications, continue to be examined. The simulations reported in Appendices E.1, E.2 and E.4, like those reported in Section 4 of the paper, involve $N(0, 1)$-errors. Appendix E.3 presents results for the same family of data-generating processes, but with t_3-errors. Appendix E.4 presents results for "F-tests" of the joint hypothesis of QTE-homogeneity in two different covariates , while Appendices E.1, E.2 and E.3 contain results for the same "t"-test of QTE-homogeneity in a single covariate considered in Section 4 of the main text.

E.1 Sensitivity analysis: Results induced by our method with a fixed bandwidth and varying choices of k and c, $N(0, 1)$-errors

This appendix analyzes the sensitivity of Wald-type tests implemented according to our procedure to variation in the the pseudo-sample size m or to the smoothing parameter h_m, which are assumed to take the forms given in (14) and (15) in the main text, respectively. Wald-type tests implemented according to our procedure are seen to have size or power performance that is not much affected across quantiles or models by variation in the pseudo-sample size provided that the pseudo-sample is sufficiently large. In particular, variation in the tuning parameter k when $k \geq 5$ is seen not to exert strong effects on the size or power performance of Wald-type tests implemented with our method. The size performance of these same tests, on the other hand, is seen to be somewhat more sensitive to variation in the bandwidth leading constant, in particular, to variation in the parameter c. We also note the relative insensitivity of the size-corrected power of these tests to variation in c.

In what follows, Tables 1–6 display empirical sizes and size-corrected powers for Wald-type tests implemented according to our procedure with the bandwidth constant c appearing in (15) of the main text fixed at $c = 1.5$. Tables 7–12 display the same quantities, but for tests implemented according to the proposed procedure in which the pseudo-sample size constant k appearing in (14) is fixed at $k = 5$.

Table 1: Empirical rejection percentages (size and size-corrected powers), Model 1. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
k/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1	7.2	15.7	31.3	46.5	4.7	18.9	38.3	57.6	5.9	18.9	40.2	57.0
3	5.6	19.4	36.6	58.6	3.7	23.4	42.4	65.7	3.9	19.6	37.9	59.7
5	6.0	17.0	37.8	61.9	4.2	23.5	42.8	66.7	4.6	19.4	38.4	61.7
7	6.1	18.7	38.7	60.9	4.7	19.0	46.7	75.4	5.4	19.9	45.8	71.9
9	6.5	19.4	47.8	77.7	5.5	20.4	43.8	70.0	6.0	19.5	44.0	69.9
11	6.5	20.4	45.6	67.6	3.9	24.0	54.2	81.0	7.2	16.5	38.2	66.6
13	6.0	20.8	40.8	69.2	6.1	19.8	47.9	75.9	5.5	20.6	45.7	73.2
15	6.4	20.1	40.5	66.2	5.1	21.7	43.0	67.9	6.2	22.9	47.7	75.4
$n = 300$												
1	4.8	31.6	69.5	86.4	3.8	32.4	70.1	88.7	4.9	29.1	65.7	82.1
3	6.0	26.6	66.0	90.6	5.1	33.0	74.6	95.1	4.7	29.5	72.4	94.2
5	5.0	30.2	74.7	95.6	5.2	37.3	85.1	99.3	5.2	28.2	73.4	95.2
7	3.8	38.5	81.8	98.4	4.3	39.7	87.6	99.3	5.0	38.9	83.3	98.6
9	5.1	32.4	80.8	98.3	3.7	41.8	89.6	99.9	6.3	30.7	76.9	97.1
11	5.8	32.1	78.7	98.3	5.5	34.9	82.2	98.8	6.6	32.1	81.1	98.7
13	4.6	35.4	81.0	98.7	4.8	37.4	86.1	99.4	6.0	32.0	80.2	98.6
15	6.7	28.8	77.6	97.9	5.2	40.2	88.2	99.8	7.0	27.2	73.5	97.9

Table 2: Empirical rejection percentages (size and size-corrected powers), Model 2. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
k/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1	6.4	15.8	23.6	33.4	3.9	22.2	39.4	53.5	5.5	20.2	36.9	48.9
3	5.2	17.9	28.2	42.7	4.2	20.5	43.6	62.7	6.2	21.5	47.1	68.3
5	5.0	16.0	31.2	50.7	5.5	16.8	33.4	57.0	5.2	26.1	54.8	75.0
7	6.6	14.8	29.1	48.1	3.6	24.4	49.0	69.2	5.6	22.7	53.4	72.5
9	7.9	10.2	26.3	46.1	5.4	22.4	49.9	71.9	5.7	26.8	52.5	75.5
11	6.9	15.1	29.7	47.7	5.8	24.5	50.8	75.0	6.1	27.3	57.7	80.1
13	6.3	19.2	31.8	48.8	4.6	25.1	50.1	72.0	6.4	25.6	52.0	73.6
15	6.5	15.9	33.0	52.5	5.7	20.2	48.5	70.9	5.0	28.3	57.2	80.2
$n = 300$												
1	5.8	18.1	46.3	71.3	3.9	33.0	70.6	85.1	5.2	38.5	75.2	84.8
3	5.4	21.9	49.4	76.7	3.5	36.0	78.1	96.2	5.0	38.3	77.7	94.1
5	4.1	22.2	57.9	87.1	3.4	39.1	86.2	98.5	5.2	40.9	83.0	96.1
7	5.4	21.7	54.2	83.4	3.8	34.1	78.1	95.5	6.2	43.2	86.5	97.8
9	5.6	21.0	56.8	86.6	3.8	40.9	88.0	98.7	4.2	41.8	86.4	98.1
11	5.5	20.2	52.1	83.3	4.0	45.4	87.8	97.7	5.7	42.6	87.1	98.2
13	4.8	22.8	58.7	86.9	4.7	37.1	82.2	97.9	7.7	44.0	89.5	98.5
15	6.9	19.4	60.2	88.4	4.8	36.3	84.3	97.3	5.7	39.7	87.2	98.4

Table 3: Empirical rejection percentages (size and size-corrected powers), Model 3. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
k/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1	5.8	15.7	25.5	36.2	4.2	17.8	30.4	43.9	5.7	20.8	38.8	53.8
3	5.6	14.0	24.9	43.5	3.8	18.9	38.6	60.6	6.5	22.6	47.1	69.8
5	6.2	12.6	28.8	50.8	4.9	18.8	43.9	67.7	4.9	26.4	53.3	74.4
7	6.6	14.1	29.0	49.5	4.2	22.9	49.2	73.6	6.5	22.4	47.9	70.6
9	5.9	13.8	23.7	39.7	5.5	15.8	39.0	60.0	7.2	24.6	52.9	73.1
11	6.9	12.9	26.0	42.9	4.5	19.0	44.9	67.2	6.8	24.6	52.2	75.8
13	6.9	15.8	27.8	45.3	5.0	19.0	42.0	64.0	6.0	26.4	55.5	76.4
15	6.8	13.6	28.9	49.1	4.4	22.4	44.2	66.9	6.9	22.7	53.2	73.5
$n = 300$												
1	5.2	16.9	39.3	63.6	3.6	31.5	67.4	82.7	7.2	37.5	71.2	86.0
3	5.3	16.8	54.2	84.4	3.5	35.1	78.1	95.7	5.3	39.6	80.6	95.0
5	5.9	17.1	49.8	81.7	4.9	31.4	83.4	97.3	5.8	48.9	90.1	98.8
7	4.9	18.7	56.2	87.4	5.1	33.0	81.4	96.6	5.1	44.4	86.8	98.9
9	7.7	13.0	46.1	78.3	4.0	33.9	77.5	96.2	4.9	44.0	88.1	98.1
11	6.5	15.2	45.5	81.6	2.9	41.0	86.4	99.0	5.6	43.8	88.4	98.8
13	6.4	14.5	51.1	82.5	4.9	30.1	78.6	97.1	5.3	51.2	93.2	99.4
15	6.1	15.2	49.4	82.6	5.1	33.4	83.7	98.4	5.9	51.3	91.6	99.6

Table 4: Empirical rejection percentages (size and size-corrected powers), Model 4. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
k/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1	5.9	17.1	27.7	40.6	5.0	18.5	38.2	52.2	4.9	20.8	40.4	54.7
3	4.6	17.7	33.9	49.6	4.2	23.7	50.5	71.2	4.8	21.3	45.4	67.6
5	7.4	15.8	33.9	52.8	4.4	22.1	49.5	71.7	6.4	19.4	46.2	72.1
7	6.5	14.5	27.1	48.1	3.9	22.4	49.8	74.9	5.3	27.9	57.7	80.4
9	7.0	15.8	27.5	47.7	3.9	19.9	44.0	68.4	7.2	20.2	50.0	75.3
11	7.4	15.5	29.3	46.7	5.0	23.1	50.8	74.4	6.8	22.4	50.1	74.3
13	6.5	14.8	34.3	56.1	3.9	23.2	51.8	73.6	6.1	19.6	46.0	70.0
15	7.6	16.4	33.9	54.1	5.7	19.8	41.9	68.2	6.4	21.6	44.9	70.9
$n = 300$												
1	5.4	23.7	51.2	71.8	3.3	40.8	79.2	91.4	6.9	35.9	71.6	85.7
3	5.3	22.3	60.1	86.2	3.5	40.0	82.8	97.0	4.4	43.9	87.0	97.2
5	6.7	18.3	50.2	79.6	4.6	37.4	86.9	98.5	4.2	42.7	86.9	98.5
7	5.9	24.2	65.0	90.7	3.9	42.2	87.3	99.4	6.2	42.3	87.3	98.9
9	4.9	29.1	67.2	92.2	5.0	37.3	83.6	98.2	6.4	48.8	93.3	99.4
11	5.5	20.5	60.4	88.6	5.2	36.8	80.9	97.6	6.5	47.4	92.9	99.7
13	5.4	24.9	67.4	90.9	4.0	44.1	90.7	99.0	5.6	39.5	87.0	98.9
15	4.8	29.9	67.0	92.0	5.0	41.7	89.3	99.1	6.1	44.2	92.8	99.3

Table 5: Empirical rejection percentages (size and size-corrected powers), Model 5. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
k/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1	5.7	14.3	28.5	45.6	5.1	18.2	34.7	49.6	5.9	18.5	35.7	50.6
3	5.1	18.4	34.6	58.9	3.9	19.6	44.2	69.8	5.5	20.7	44.4	66.7
5	5.1	17.2	35.1	56.9	5.6	18.9	41.4	63.8	6.1	21.2	50.7	75.9
7	5.8	17.0	39.0	63.0	4.4	22.8	51.0	78.6	4.7	24.9	52.7	76.2
9	5.8	18.1	33.7	54.7	5.6	19.8	45.9	70.3	4.6	25.7	54.2	76.5
11	6.1	17.9	41.6	64.8	5.0	20.6	48.9	75.9	6.3	21.1	47.1	73.4
13	5.9	20.1	38.2	62.4	5.7	21.6	47.8	72.7	5.3	23.1	52.1	80.0
15	7.0	13.8	31.5	55.2	5.6	17.4	42.9	72.4	6.0	23.6	49.7	76.3
$n = 300$												
1	4.3	24.1	57.4	81.3	3.7	35.4	74.7	89.6	5.7	30.4	64.2	81.4
3	4.8	29.3	71.3	93.4	4.5	34.4	81.1	97.6	6.0	36.5	78.8	96.5
5	5.5	24.7	61.6	88.2	5.3	31.2	79.8	97.6	5.6	35.0	81.8	98.4
7	4.3	32.5	73.8	94.9	5.2	37.3	83.5	99.1	5.0	37.3	84.1	98.9
9	5.0	27.2	67.0	92.5	4.1	40.5	84.7	99.2	6.1	38.0	83.7	98.3
11	5.9	26.7	70.2	94.4	4.6	37.5	87.6	99.6	5.5	35.1	82.5	98.5
13	5.3	25.8	71.0	94.8	3.3	42.1	90.6	99.2	5.9	45.5	90.9	99.7
15	5.7	29.1	73.1	95.0	4.0	41.3	85.7	98.7	5.7	42.5	90.2	99.5

Table 6: Empirical rejection percentages (size and size-corrected powers), Model 6. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
k/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1	7.7	21.0	42.3	61.9	3.7	23.0	43.3	61.5	6.3	13.7	28.6	45.1
3	4.9	25.9	54.5	78.9	4.1	22.2	46.2	71.3	6.3	13.9	28.0	49.5
5	6.3	20.3	43.3	68.1	5.6	18.2	38.4	67.3	6.2	15.3	34.5	62.8
7	7.3	25.4	51.3	73.7	4.3	21.7	45.5	72.1	5.6	16.0	39.7	65.5
9	5.7	25.9	53.7	77.5	5.3	20.0	47.3	73.2	8.3	11.0	27.4	53.7
11	6.0	24.0	47.5	71.4	4.9	22.0	48.0	77.1	6.2	14.0	30.2	53.3
13	7.7	26.8	55.7	82.4	5.4	22.1	42.3	65.3	7.8	15.6	34.1	61.2
15	5.8	26.3	51.0	77.0	4.6	22.3	50.6	75.7	8.0	12.9	33.0	64.3
$n = 300$												
1	5.3	41.3	75.7	88.6	3.7	35.3	74.0	90.4	5.3	22.0	53.7	79.0
3	4.4	42.1	80.5	95.9	4.4	36.1	82.3	97.7	5.5	19.4	55.6	85.6
5	4.2	43.2	82.0	97.7	3.5	38.8	85.4	98.8	5.9	19.0	66.8	94.4
7	5.3	47.8	88.8	98.9	3.2	43.0	88.7	99.6	6.7	22.2	69.9	96.2
9	7.1	43.4	89.0	99.6	4.5	39.7	88.0	99.7	5.5	18.2	64.6	93.4
11	6.4	40.1	83.4	98.6	4.3	34.8	87.9	99.3	6.3	19.7	65.6	95.6
13	5.5	50.5	93.4	99.9	5.2	36.8	88.0	99.1	5.8	23.0	68.9	96.4
15	4.3	48.5	87.8	99.1	4.5	36.8	86.8	99.7	6.1	23.8	77.1	98.3

Table 7: Empirical rejection percentages (size and size-corrected powers), Model 1. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $k = 5$

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
c/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1.00	7.80	20.60	33.60	50.20	7.30	17.80	45.50	71.50	8.50	17.50	40.20	64.40
1.25	7.20	17.80	40.10	63.80	6.00	20.80	43.90	70.70	6.30	18.70	38.10	62.00
1.50	6.30	20.70	44.40	72.20	4.10	20.80	49.90	78.70	5.10	19.60	46.60	73.90
1.75	6.00	18.30	46.80	75.80	4.30	21.30	51.70	81.60	5.20	19.30	39.90	64.60
2.00	3.10	23.10	47.30	73.00	2.00	30.60	63.10	87.90	3.90	18.40	38.10	65.30
$n = 300$												
1.00	6.20	33.00	72.00	93.20	6.00	33.20	76.20	96.90	8.20	30.40	77.20	94.70
1.25	5.00	37.80	81.30	98.00	6.00	30.30	75.20	96.60	6.50	33.00	78.50	96.40
1.50	5.00	29.30	70.80	94.10	4.70	38.20	84.90	98.90	5.70	27.20	68.30	93.60
1.75	4.30	36.70	83.90	98.70	3.40	40.60	88.70	99.40	4.60	34.30	80.90	98.00
2.00	3.50	36.10	83.30	99.20	2.70	39.30	84.90	98.80	4.30	35.50	85.50	99.50

Table 8: Empirical rejection percentages (size and size-corrected powers), Model 2. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $k = 5$

$n = 100$			$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$	
c/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1.00	9.50	16.30	29.50	44.70	8.00	21.90	49.20	72.00	7.50	23.40	44.40	64.10
1.25	7.90	14.80	27.70	44.50	4.20	20.90	45.70	67.40	7.30	21.20	43.10	64.60
1.50	7.20	13.40	25.50	44.20	4.90	20.90	44.20	66.20	6.10	23.80	48.10	66.70
1.75	4.90	14.40	30.00	49.00	3.00	21.30	47.00	69.60	4.40	22.00	42.20	60.00
2.00	3.90	14.90	28.90	44.40	2.90	23.40	48.50	70.00	4.30	22.20	46.90	68.80
$n = 300$												
1.00	6.00	24.80	56.40	82.20	6.00	36.20	79.20	95.20	6.90	40.40	78.10	93.60
1.25	5.60	23.10	56.90	83.00	5.20	33.70	78.90	95.40	5.50	43.20	85.60	96.30
1.50	5.80	18.30	52.70	80.50	4.10	44.10	85.70	98.00	6.50	47.20	89.60	98.50
1.75	3.40	22.70	59.00	86.30	3.40	37.20	84.80	97.50	4.80	50.40	91.90	98.60
2.00	4.40	24.80	63.10	90.10	2.90	44.90	87.60	98.70	3.60	52.80	93.80	99.10

Table 9: Empirical rejection percentages (size and size-corrected powers), Model 3. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $k = 5$

$n = 100$			$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$	
c/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1.00	8.60	12.20	22.50	39.70	6.80	18.70	36.10	56.70	8.60	22.70	43.80	62.70
1.25	6.50	13.60	24.50	41.70	5.60	19.20	45.30	69.00	6.40	27.10	54.40	76.10
1.50	4.60	14.70	29.60	49.50	4.10	20.50	45.40	67.20	5.00	27.90	58.00	78.10
1.75	5.40	15.40	27.50	47.10	3.40	20.40	49.20	75.30	4.30	32.10	66.30	82.70
2.00	3.90	15.20	32.40	52.20	2.80	18.20	46.70	71.10	4.40	27.00	58.30	80.80
$n = 300$												
1.00	8.90	14.20	42.00	76.60	4.90	35.60	79.90	95.80	7.40	45.90	86.60	97.10
1.25	6.50	13.70	41.10	73.90	5.80	25.40	68.90	92.60	5.90	42.40	84.90	96.60
1.50	5.70	16.60	49.90	82.90	3.50	38.50	83.50	97.70	5.20	48.80	91.50	98.40
1.75	3.70	19.70	52.40	85.80	3.40	34.90	80.00	98.10	4.90	44.60	86.80	98.60
2.00	4.70	17.30	50.40	82.90	3.70	36.40	81.80	97.20	3.60	50.90	93.20	99.30

Table 10: Empirical rejection percentages (size and size-corrected powers), Model 4. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $k = 5$

$n = 100$			$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$	
c/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1.00	8.00	15.90	27.00	41.70	7.60	18.20	40.60	61.20	8.20	25.20	57.40	78.80
1.25	6.20	20.00	33.30	51.80	5.90	21.20	42.40	63.80	7.80	21.90	46.50	66.50
1.50	6.10	14.00	28.10	46.70	5.20	21.40	47.30	71.10	5.50	23.10	48.50	70.30
1.75	4.80	14.50	30.00	51.10	3.60	20.40	42.30	66.30	4.90	22.10	47.50	67.30
2.00	3.40	21.40	42.30	64.70	2.70	25.80	53.40	75.70	3.90	20.10	39.90	65.70
$n = 300$												
1.00	7.40	22.90	54.30	79.40	6.50	36.40	80.50	96.30	7.20	41.10	87.20	97.80
1.25	6.40	23.10	58.90	84.80	4.30	39.80	82.50	96.80	5.80	40.10	85.70	97.30
1.50	4.70	22.90	57.00	85.10	5.00	37.40	84.50	98.70	4.70	46.20	89.00	98.80
1.75	4.20	26.20	60.50	87.60	3.60	34.50	81.20	97.50	5.30	53.00	92.80	99.70
2.00	3.40	23.30	63.10	90.50	3.20	43.30	87.20	99.10	4.80	39.50	87.20	98.80

Table 11: Empirical rejection percentages (size and size-corrected powers), Model 5. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $k = 5$

$n = 100$			$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$	
c/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1.00	9.30	16.30	31.70	49.70	7.60	21.30	47.90	73.60	7.70	19.80	40.90	63.10
1.25	6.10	18.30	37.50	59.60	5.30	21.80	43.80	71.50	7.70	19.70	37.00	57.50
1.50	5.70	18.00	35.30	56.80	4.00	25.40	50.30	74.30	6.00	19.80	39.70	64.10
1.75	4.40	19.10	37.50	60.60	3.70	25.50	55.20	80.70	5.80	21.00	52.00	76.90
2.00	4.00	16.50	37.00	57.80	2.30	22.30	46.10	73.10	3.50	22.40	45.90	69.50
$n = 300$												
1.00	6.80	28.00	65.80	90.60	4.90	39.20	83.70	98.60	6.20	39.70	81.70	97.50
1.25	6.70	23.70	58.70	84.40	6.30	31.60	78.90	97.00	5.60	35.80	78.80	96.50
1.50	5.50	24.20	62.90	90.60	4.50	34.60	80.70	98.00	4.20	43.80	88.20	98.90
1.75	5.70	27.10	70.80	93.70	2.90	38.50	88.50	99.40	5.10	41.10	88.40	98.90
2.00	4.50	27.00	69.90	94.20	3.10	35.80	86.10	99.30	4.00	36.10	80.80	98.30

Table 12: Empirical rejection percentages (size and size-corrected powers), Model 6. 1000 Monte Carlo replications; procedure “weg” implemented with fixed bandwidth, $k = 5$

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
c/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
1.00	8.70	20.80	41.60	64.20	6.90	20.40	46.40	73.60	9.50	14.90	29.50	50.60
1.25	5.90	23.20	41.50	58.20	6.00	21.40	42.90	68.30	7.50	12.00	32.80	55.60
1.50	5.60	23.70	53.80	77.30	5.40	19.00	47.00	73.80	6.10	14.70	31.80	58.10
1.75	5.20	23.60	43.50	66.70	3.10	20.50	45.70	72.70	4.40	15.40	34.80	57.90
2.00	4.00	25.80	50.50	75.30	2.60	24.20	52.40	80.20	3.90	16.60	38.30	66.20
$n = 300$												
1.00	7.10	45.50	86.90	97.90	5.20	34.70	77.60	96.60	6.30	19.50	61.10	87.60
1.25	5.40	47.80	88.40	98.30	4.40	41.40	88.90	99.20	5.80	23.20	71.50	95.80
1.50	4.60	47.90	88.60	98.70	4.10	40.10	87.60	98.90	6.30	18.90	58.20	89.50
1.75	5.30	43.90	88.40	99.10	2.50	45.70	88.70	99.60	4.40	21.00	68.50	94.50
2.00	4.00	47.50	88.40	98.70	2.80	45.70	93.00	99.70	4.10	20.60	65.40	94.90

E.2 Data-driven bandwidth: Results induced by our method with $k = 5$ and the data-driven bandwidth derived in Appendix D

This appendix presents the performance of Wald-type tests implemented according to our procedure but in which the bandwidth takes the form given by h_m^* above in (54). The rate-optimal data-driven bandwidth given above in (54) is seen to induce Wald-type tests with good size and power performance across quantiles and data-generating processes. These results are for the same “t-test” of QTE-homogeneity in a single covariate considered in Section 4 of the main text, and involve the same series of six data-generating processes with $N(0, 1)$-errors considered in the main text. Tables 13–18 below repeat the relevant entries in Tables 1–6 in the main text for ease of reference.

Table 13: Empirical rejection probabilities (size and size-corrected powers), Model 1. 1000 Monte Carlo replications; procedure "weg" implemented with data-driven bandwidth, $k = 5$; other procedures implemented using summary.rq.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	7.2	14.2	36.8	59.6	4.7	21.3	45.9	72.2	4.8	19.8	45.3	71.1
wiid	9.1	10	22.2	39.5	7.3	15.5	45.7	75.9	8.2	12.5	31.1	56.3
wnid	8.1	8.3	18.7	37.9	6.8	17.5	51	80	7.4	12.2	33.1	59.9
wker	1.3	13.2	31.5	53.7	0.3	17	51.2	80.8	1.9	17.7	41.8	69.5
riid	7.9	8.6	21.4	39.4	8.6	17.7	46.5	76.9	7.5	15.3	35.5	61.5
rnid	5.9	7.4	19	37.7	6.5	17.5	46.7	76.5	5.1	15.2	34.7	61.3
bxy	3.1	9.6	23.6	44.7	2.9	16.7	49.8	80	3.2	14.8	37	65.7
bpwy	1.2	9.7	23.7	44.3	2.4	17.1	49.4	80.4	1.6	17.5	41.1	69.6
bmcmb	3.3	8.8	23.2	43.3	3.7	16	48.9	79.2	3.4	16.6	39.7	66.7
bwxy	4.1	9.3	22.9	44.5	3	16	48.4	79.9	4.4	13.7	36	64.6
bwild	6.9	10.9	24	46.2	7.2	14.1	42.7	76.1	6.2	16.2	37	65.4
$n = 300$												
weg	6.1	26.9	71.4	93.9	4.6	36.4	85.1	98.8	5.9	27.3	68.8	93.3
wiid	7.9	25.4	74.2	98.1	3.7	33.6	84.3	98.5	6	30.5	84.5	99.6
wnid	8.2	26.2	76.1	98.6	3.9	34.9	86.4	98.6	5.9	32.5	84.7	99.3
wker	3	28.4	79.5	99.3	1.3	34.5	85.9	98.7	2	34.3	87	99.7
riid	7.7	27	75.8	97.6	5	31.4	80.5	98.1	5.6	31.7	81.6	98.8
rnid	6.6	26.5	74.7	97.6	4.7	31.4	80.4	98	4.7	31	82.3	98.6
bxy	4.4	29.4	79.2	98.3	2.5	34.1	84.4	98.4	3	32.7	85.5	99.4
bpwy	3.4	28.9	78.8	98.7	2.2	34.4	84.9	98.4	2.3	34.5	85.9	99.3
bmcmb	5.9	26.9	77.9	98.4	3.7	33.7	82.4	98.3	3.8	32.5	84.6	99.2
bwxy	4.9	29.2	79.1	98.8	2.7	32	82.4	98.4	3.1	31.5	83.9	99.2
bwild	7.1	29	79.1	98.7	4.8	32.3	82	98.3	4.9	31.9	85.7	99.6

Table 14: Empirical rejection probabilities (size and size-corrected powers), Model 2. 1000 Monte Carlo replications; procedure "weg" implemented with data-driven bandwidth, $k = 5$; other procedures implemented using summary.rq.

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	7	12.4	21.3	39.2	6.2	17.4	38	58.7	6.1	20.5	44.1	65.4
wiid	8.4	10.2	20.5	38.9	8.9	12.7	34.5	62.7	9	15.7	39.9	63.1
wnid	7.4	7.3	21.8	40.2	9.1	12.6	37.2	65.3	8.6	14.5	42.1	64.4
wker	1.5	8	21.9	39.9	1.1	12.5	36.5	63.2	1.7	11.3	37.7	61
riid	7.7	7.9	20.2	36.7	8.7	11.1	31.7	55.1	8.2	14.7	37.6	60.1
rnid	5.8	7.5	20.2	36	7.2	11.4	31.1	54.1	6	14	35.9	57.1
bxy	3.4	7.9	20.3	37.6	3.4	12.6	36.2	60.3	4.1	14.6	39.3	62.1
bpwy	1.8	7.1	20.8	40.2	2.9	12.8	37.1	62.7	2.5	12.7	40	62.6
bmcmb	3.4	8	20.5	36.7	4.1	12.7	36.2	60.1	4.6	15.3	39.2	61
bwxy	4.5	8.3	20.6	37.9	4.2	13.2	37.1	60.2	5.2	13.5	38.7	61.5
bwild	7.4	7.3	18.7	35.4	8.4	12.9	35.1	57	7.3	14.1	38.5	59.2
$n = 300$												
weg	5.7	15	45.9	77.2	5	32.6	77.1	96.1	5.6	35.4	81.2	95.7
wiid	5.5	20.7	58.8	88.4	5	32.2	81.3	98.5	8	34	83.4	98.3
wnid	5.9	19.6	60.1	88.6	4.8	35.6	84.6	98.5	8.4	36	86.5	98.9
wker	2.3	18.4	57	86	1	35.9	82.2	97.9	2.3	36	85.7	98.7
riid	6	17.9	55	83.9	5.4	31.8	77.7	96.5	7.5	35.6	82	97.3
rnid	4.6	17.3	53.3	83.1	5.1	30.7	76.9	96.2	6.8	33.8	80.9	96.8
bxy	2.6	20.7	58.8	84.2	3.7	32.7	79.9	96.9	3.7	38.4	84.6	98
bpwy	2.4	18.1	55.3	83.7	3	32.7	79.1	97	3	38.5	84.7	98.4
bmcmb	4.3	18.3	53.1	82.7	4.4	31.2	78.4	97	5	37.9	84.2	97.4
bwxy	2.6	17.8	53.8	81.8	3.6	31.5	78.6	96.7	4	36.1	82.8	97.3
bwild	5.1	19.1	55.8	84.1	5	30.7	78.9	96.4	6.1	36.1	84.6	98.5

Table 15: Empirical rejection probabilities (size and size-corrected powers), Model 3. 1000 Monte Carlo replications; procedure “weg” implemented with data-driven bandwidth, $k = 5$; other procedures implemented using summary.rq.

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
`weg`	7.4	10.9	21.5	39.8	6.2	17	36.9	57.5	6.3	23.3	53.1	76.8
`wiid`	9.7	6.1	14.5	26.4	7.5	11	28.5	53.3	7.7	16.1	44.9	71.8
`wnid`	7.9	8.4	19	36.8	6.7	11	31.9	56.8	7.2	18.2	47.4	72.1
`wker`	1.4	8.1	19.7	39.2	0.7	12.5	33.4	58	1.4	18.9	52.6	78.1
`riid`	7.5	6.5	15.7	32.6	7.3	9.4	26.7	47.4	8	16.9	43.9	68.2
`rnid`	5.3	6.7	16.6	32.2	6.5	9.3	27.8	45.6	5.5	17.3	45.4	68.3
`bxy`	2.4	8.3	19.1	37.9	2.8	12.3	32.3	55.7	3	19.3	49.2	75.2
`bpwy`	1.2	8.1	20.3	38.2	2.4	11.6	31.8	54.2	1.5	18.7	50.2	75.7
`bmcmb`	2.6	7.5	18.5	34.5	3.6	11.6	31.8	54.7	3.1	18.1	47.1	73
`bwxy`	3.1	8.5	20.2	37.6	3.5	10.7	30.9	54.2	3.9	18.9	49.5	74.3
`bwild`	6.3	7.7	18.5	35.7	7.6	10	27.7	50.2	7	17.1	47.2	73.6
$n = 300$												
`weg`	5.2	17.4	52.4	82.3	4.8	30.3	79.4	97.1	4.9	49.5	90.5	98.4
`wiid`	6.6	12.5	46.4	81.3	6.9	24.4	74	96.3	6.9	41.1	91.1	99.5
`wnid`	6.8	14.7	52.7	84.1	5.8	28.7	78.2	97.3	7.7	41.4	92	99.7
`wker`	3.3	15.4	52.7	84.5	1.6	28.2	76.7	96.2	3.2	40	90.4	99.7
`riid`	5.8	15.6	49.7	82.2	6.4	26	72.1	95	7.3	38.3	87.3	98.9
`rnid`	5	15	48.1	80.5	6	25.4	70.4	94.4	6.4	37.9	86.5	99
`bxy`	3.7	16.1	50.3	83.3	3.5	27.3	74.7	95.6	3.8	41.1	89.9	99.6
`bpwy`	3.1	15.6	52	83.7	3	28.2	75.4	95.9	2.8	38.5	89.8	99.2
`bmcmb`	4.7	14.8	49.7	81	4.7	28.7	76.5	96	5.1	40.8	90.5	99.4
`bwxy`	3.7	14.9	51	82.8	3.7	28.5	75.7	96	4.2	39.9	90	99.6
`bwild`	6.3	13.9	48.7	81.9	5.9	25.3	73.3	95.7	6.8	37.8	88.9	99.5

Table 16: Empirical rejection probabilities (size and size-corrected powers), Model 4. 1000 Monte Carlo replications; procedure "weg" implemented with data-driven bandwidth, $k = 5$; other procedures implemented using summary.rq.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	4.9	13.3	24.4	41.8	4.8	15.8	38.6	60.2	5	22.5	46.2	66.9
wiid	9.8	6	16	30.4	7.5	14.2	41.1	68.6	9.8	13	36.8	66.9
wnid	8.5	6.3	15.6	32.2	7.8	14	42.8	69	8.2	15.9	43.8	71.8
wker	1.4	11.7	24.5	43.2	1.1	12.7	43	66.9	1.7	16.2	45.1	73.2
riid	7.7	7.4	17.8	31.6	7.4	15.3	40.3	63.1	7.9	14.4	41	67.2
rnid	5.4	8.2	18.8	34.5	6.3	13.8	39.5	62.1	5.5	15.9	41.7	68.4
bxy	3.2	9.1	19.6	37.8	3.6	14.6	42.5	65.3	3.1	17.6	46.3	72.6
bpwy	1.5	8.5	20.7	38	2.7	13.8	40.6	64.2	1.1	17.1	47.3	74.9
bmcmb	4.4	6.7	17.2	33.3	4.1	14	41.2	64.2	3.2	17.3	45.9	71.4
bwxy	4.4	8.9	20.2	37.7	3.9	15	42.9	66.2	4.3	17.7	47	72.5
bwild	7.4	9.2	20.7	37.3	6.7	13.6	40	64.3	7.8	15.3	41.5	68.7
$n = 300$												
weg	4.6	26.6	64.8	90.2	5.4	30	73.5	94.2	5.7	42.3	89.6	98.4
wiid	6.5	14.5	48.1	81.5	6.9	28.7	79.6	97.7	5.9	40.1	88.3	99.2
wnid	7.3	17.6	53.3	84.3	7.2	28.2	79.6	97.6	5.9	42.9	90.3	99.4
wker	3.5	23.3	59.9	87.1	2.1	28.9	78.6	97.6	2.3	41.1	88.9	98.8
riid	7.2	17.4	49.2	81.6	8	26.7	76.5	96	5.6	40.4	86.6	98.2
rnid	6.1	17.5	50.8	81.9	6.8	25.6	76.1	95.5	4.8	41	86	98.1
bxy	4.5	18.6	52.2	82.5	3.8	27.8	77.6	96.4	3.3	40.8	87.6	98.3
bpwy	4	18	55	84.5	4.4	28.9	77.2	96.5	2.3	42.2	87.9	98.5
bmcmb	5.6	17	50.8	81.5	5.7	28.8	78.2	96.5	4.6	41.1	87.4	98.1
bwxy	4.5	18.2	52.5	82.5	4.8	28.1	76.7	96.2	3.3	43.5	88.6	98.4
bwild	6.6	17.9	53.1	82.9	6.5	25.9	75.7	96	5	41.5	88.3	98.7

Table 17: Empirical rejection probabilities (size and size-corrected powers), Model 5. 1000 Monte Carlo replications; procedure "weg" implemented with data-driven bandwidth, $k = 5$; other procedures implemented using summary.rq.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	5.6	13.7	32.4	59.2	5.4	15.3	37.7	63.3	6.3	17.4	41	64.4
wiid	8.5	10.9	22.7	42.1	7.2	11.5	27.7	55.6	9.6	11.3	28.9	54.6
wnid	8.2	10.1	25.5	46.1	6.7	11	32.5	59.5	8.2	10.7	31	57.4
wker	1.1	13.1	30.4	54.3	0.7	12.3	34	60.5	1.5	12.8	35.9	65.2
riid	7.3	11.1	25.7	45.9	8.1	9.4	27.2	51.8	8.4	11	29.5	58.8
rnid	5.3	11.2	26.1	45.8	7	10.7	27.3	51.9	6.2	11.7	28.9	56.8
bxy	2.7	11.4	27.1	49.4	2.5	11.9	32.7	58.7	3.3	12.4	33.3	62.7
bpwy	1.2	12.1	28.8	50.7	2.6	12.5	33.8	60.3	2	12.1	34	64
bmcmb	2.9	10.8	27.6	47.5	3.7	11.3	31.8	59.5	3.5	11.7	32.8	59.4
bwxy	4.2	11.3	27.7	48.8	3.6	11.3	32.4	58.6	4.4	11.8	32.7	61.8
bwild	6.8	12.4	26.9	47.2	7	9.7	28.6	53.9	7.4	10.5	31.8	61.2
$n = 300$												
weg	6.3	22	62.8	89.9	5.6	29	78.1	97	5.2	36.6	84.7	98.1
wiid	7.2	24	66.6	94.6	6.6	25.9	73.1	96.5	6.6	33.6	83.5	99.3
wnid	6.9	24.5	68.4	95.6	6.3	29	76.7	97.5	7	37.4	86.6	99.3
wker	2.7	26.5	72.1	96.6	1.7	30.4	77.3	97.8	2.6	38.3	87.9	99.4
riid	6.4	20.3	63.8	91.7	5.9	25.9	72.7	95.5	6.9	33.3	82.5	98.5
rnid	5.4	22.2	66.5	92.9	5.5	26.7	73.3	95.5	5.7	34.1	83.8	98.6
bxy	3.6	24.7	70.3	95.5	3.8	29.4	75.4	97.4	4	34.3	84.9	99
bpwy	3.5	23	68.2	95	3.6	28.6	75.9	97.1	2.7	37.7	85.8	99.2
bmcmb	4.9	24	68.6	95.3	5	28.1	75.5	96.9	4.6	36.2	85.4	99.3
bwxy	4	24.4	69.7	95.7	4.1	29.2	75.7	97.1	4.1	35	85.2	99.1
bwild	6.4	23.1	69.1	95.7	6	28.1	74.7	97.1	5.7	35.3	85.1	99.1

Table 18: Empirical rejection probabilities (size and size-corrected powers), Model 6. 1000 Monte Carlo replications; procedure "weg" implemented with data-driven bandwidth, $k = 5$; other procedures implemented using summary.rq.

n = 100		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	5.4	20.1	42.6	64.1	5.3	19	41	67.9	5.7	14.1	32.3	56.9
wiid	9.7	12.2	31.5	57.4	7.2	13.6	44.1	80.4	10	6.7	18.7	40
wnid	7.3	16.3	39	66.6	5.9	16.3	52.4	86.7	8.1	7.8	21.8	43.7
wker	1.3	20.2	47.4	75.8	0.8	16.3	53.5	89.2	2.2	8.9	26.2	52.1
riid	8.4	15	36.8	62.3	7.5	14.7	46.5	80.5	7.8	5.7	19.8	41.2
rnid	6.7	13.1	35	60.1	5.5	15	46.8	82.8	5.6	6.2	20.8	43.7
bxy	2.7	17.6	41.4	70.9	2.4	16.7	52.5	86.9	3.1	8.7	25.1	50.4
bpwy	1.5	17.7	42.7	71.8	1.9	16.9	51.1	87.2	1.7	8	22.5	48.7
bmcmb	3.1	15.8	40.3	69.1	3.4	15.8	51.8	85.7	3.6	8.6	23.7	50.7
bwxy	3.9	17.7	41.7	71.5	2.9	17.5	52.7	87.3	4.2	8	23.2	49.4
bwild	6.9	16.2	40.2	70	6.7	14	46	83.3	7.3	7.3	21.8	46.2
n = 300												
weg	5	43.4	86.1	98	6.2	28.5	76	97.4	6	18	60.2	91.5
wiid	6.4	39.7	87.4	99.2	8.3	25.9	76.7	97.9	6.5	13.9	54.3	91.3
wnid	6.4	42.8	89.3	99.7	8.1	26.7	78.4	98.9	6.5	16.6	60.6	94.2
wker	3.2	43.6	91	99.7	2.4	31.5	83.4	99	2.6	16	60.5	94.5
riid	6.9	39.3	86.5	99	7.5	25.2	71.8	97	6.4	15.1	56.7	92
rnid	6.1	39.7	86.1	99	6.9	27.8	76.3	97.6	5.3	15.2	55.7	92
bxy	3.2	43.8	89.3	99.4	4.4	30.3	80.1	98.4	3.1	16.1	59.3	93.5
bpwy	3.2	42.5	88.1	99.4	4.2	29.4	80.9	98.5	3.1	16.3	58.1	93.2
bmcmb	5.2	40.8	88.3	99.4	6.3	28.3	78.6	98.5	4.7	15.2	58.3	93.3
bwxy	4.6	39.8	87.7	99.5	5	28.8	79.4	98.4	3.5	16.5	58.5	93.7
bwild	5.7	39.7	87.7	99.5	6.7	29.2	80.4	98.7	6.1	14.2	56.8	93.1

E.3 Results for models with Student-t errors

This appendix repeats the simulations presented in Section 4 of the main text, but in which the $N(0,1)$-errors specified are replaced with t_3-errors. The corresponding simulation results are displayed in Tables 19–24.

We see that the empirical size accuracy and size-adjusted power of Wald-type tests induced by the proposed estimate of $\boldsymbol{G}_0(\alpha)$ are quite competitive with the other methods considered.

Table 19: Empirical rejection percentages (size and size-corrected powers), Model 1 with t_3-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	6.1	17.7	37.4	62.1	4.6	20.9	44.4	70.9	8.5	14.9	28.9	52.6
wiid	8.6	12.9	27.6	51.5	7.9	10.7	32.2	58.6	10.4	9.1	21.7	41.4
wnid	5.3	14.2	35	60.9	6.6	12.1	38.3	67.4	7.1	10.1	27.1	51.8
wker	1.2	12	30.8	57.8	0.5	15.8	43	73.3	2.4	13.6	34	59
riid	6.9	13.4	33.7	58.1	5.9	14.4	37.4	65.5	8.3	10.1	26.5	49.2
rnid	4.3	13	33.1	57.6	4.6	13.9	38	65.9	6	8.7	24.4	47.7
bxy	1.9	14.5	36.8	62	2.3	15	41.3	70.9	2.8	11.1	29	55.4
bpwy	0.9	13.1	34.7	59.7	1.4	15.6	41.4	72.2	1.9	14.5	34.3	60.5
bmcmb	2.9	15	34	59.5	2.9	16.1	41.9	72.4	3.6	10.4	27.7	54.1
bwxy	2.8	14.4	37.5	62.9	2.4	16.4	40.7	71.9	3.9	12.5	31.3	57.9
bwild	6.3	13.1	32	57.7	4.7	14.1	38.2	67.3	7.5	9.9	25.7	51.2
$n = 300$												
weg	6.9	21.8	61.9	89.1	4.5	33.8	80.9	98.3	6.6	29.9	69.2	93.9
wiid	8.5	16.3	57.1	90.3	7	21.9	74.5	97.5	7.8	22.7	70.4	95.1
wnid	7.3	20.7	67.3	94	6.1	28	79.3	98.2	7.3	24.2	72.6	95.9
wker	2.7	20.4	66.7	93.2	1.2	33.5	82.9	98.8	3.6	26.9	73.7	96.6
riid	6.9	18.3	62.3	91	6.1	27.8	78.5	98	5.9	28.3	74.7	95.6
rnid	5.6	19.7	63.7	91.2	5.7	27.9	78	97.9	5.1	26.6	74.3	94.9
bxy	3.2	22.2	68.4	93.2	3.2	29.6	80.1	98.3	3.2	30.6	76.7	97
bpwy	2.9	22.7	68.1	93.9	3.2	29.3	80.7	98.3	2.4	28.5	75.1	97
bmcmb	5.5	21.4	65.6	92.9	5	28.6	79.7	98.2	4.8	29.5	75.9	96.3
bwxy	3.9	22.8	69.2	93.9	3.6	29.4	81.1	98.3	3.7	30.1	76.6	97.3
bwild	7.3	18.9	63.7	92.2	5.8	26.4	77.1	98	6.4	26.3	73.2	96.1

Table 20: Empirical rejection percentages (size and size-corrected powers), Model 2 with t_3-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq.

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	7.1	12	24.3	40.7	4	19.5	39.5	57.2	7.5	20.6	42.7	60.2
wiid	9.8	7.9	15.9	31.3	6	10.6	25	45.4	8.8	14.6	34.7	56.6
wnid	6.3	9.1	20.8	36	6.5	10.3	28.3	47.9	6.5	15.5	37.8	57.9
wker	1.9	7.4	16.5	32	0.7	11.1	31	49.6	2	17.3	40.3	61.1
riid	7.4	8.8	17.2	30.1	6.4	8.7	23.7	39.8	7.1	15.4	34.2	54.3
rnid	5.2	8.5	17.1	29.1	5.1	9.4	23.7	40.2	4.9	15.1	34.2	53.8
bxy	2.2	8.2	18	32.2	2	10.5	28.7	45.8	2.4	16.7	39.7	59.6
bpwy	0.9	9.3	19.3	33.6	1.6	11.1	30.3	47.9	1.5	19.1	43.3	62.5
bmcmb	2.2	8.2	18.2	31.7	2.8	11.3	29.2	45.7	2.6	16	36.2	55.6
bwxy	3.3	8.2	18.8	31.6	2.2	10.4	29.6	47.6	2.9	18.3	40.3	59.2
bwild	7.3	7.8	16	30.5	5.2	10.3	25.2	42.2	6.9	14.8	35.4	56.2
$n = 300$												
weg	6.2	14.9	40.5	71.4	4.1	36	76.4	95.8	6.4	39.4	81.5	95.8
wiid	6.9	12.5	38.4	70.1	5.9	27.1	72.8	95.2	6.8	34.6	80.4	96.1
wnid	6	15.1	44.2	74	5	30.8	77.1	96.3	5.4	40.7	86.3	97.3
wker	3.3	12.2	37.1	67.6	0.9	35.8	78.1	96.2	3.4	38.7	82.5	96.3
riid	8.7	10.8	31	59.9	5.8	27.1	68	91.5	7.5	31.4	74	92.1
rnid	7.7	10.9	30.7	58.6	5.1	27.5	67.2	91.1	6.1	33.3	75	93
bxy	3.8	13.9	36.7	64.9	2.7	30	72.5	93.9	3.8	36.7	80.3	95
bpwy	3.5	12.4	37.2	66.7	2.8	30	73	93.3	3.9	35.7	78.9	95.4
bmcmb	5.7	13.1	35.8	62.8	3.9	30.2	72.3	93.1	4.7	39.1	80.5	95
bwxy	4.1	13.7	36.6	63.8	3.1	30.1	72.5	93.6	4.3	35.6	79	94.7
bwild	6.8	13.2	34.1	63.9	4.8	29	71.5	93.3	6.2	34.3	78.4	94.6

Table 21: Empirical rejection percentages (size and size-corrected powers), Model 3 with t_3-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	7.1	10.9	18.8	33.4	5.6	18.7	39.3	61.5	7.8	19.4	41.4	62.8
wiid	10.5	4.8	10.2	20.7	6.6	13.8	34.2	57.5	10.5	10.7	28	48.2
wnid	6.2	6.3	13.5	29.1	6	13.9	35.1	59.7	5.6	19.1	41.3	66.2
wker	1.5	7.4	18	35.3	0.5	14.7	35.9	60.6	1.3	17.9	40.9	65.6
riid	6.3	7.3	16.2	30.6	7	13.6	32	54.6	6.7	15.6	35	58.1
rnid	4.4	6.5	15.5	30.2	4.8	14.7	34	56.4	4.6	17.7	36.7	60.1
bxy	2.6	6.2	16.5	32.2	2.4	14.7	34.1	60	2.3	20.3	42.9	66.9
bpwy	1.2	6.7	15.6	31.8	1.6	14.2	33.7	59	0.8	20.7	43.8	68.5
bmcmb	3.4	5.3	13.5	27.1	3.2	12.1	30.6	54.6	2.6	17.5	38.1	63.2
bwxy	3.1	6.9	17.2	33.7	3.3	14.4	34.5	59.8	2.5	20.6	43.5	67.5
bwild	5.9	6.3	15.2	30.7	5.8	12.7	31	55.6	6.9	16.6	37.1	60.2
$n = 300$												
weg	5.3	15.5	41	71.4	3.3	27.8	75.3	95.3	6.9	33.7	79.7	96.1
wiid	8.5	9.5	31.5	64.7	6.1	20.3	66.8	94.9	8.1	28.1	78.3	97.6
wnid	7	9.9	35.8	71	5.2	23.3	73.8	96.7	8.1	31.3	82.2	98.1
wker	3.2	11.4	36.5	71.5	1.2	27.3	77.4	97.1	3.7	32.5	83.5	98.1
riid	6.1	12.3	35.5	67.6	6.3	20.4	66.4	94.2	6.7	27.6	75.4	96.6
rnid	5.3	13.3	35.3	67.6	5.8	21.1	68.1	93.3	5.9	29.8	77.4	96.5
bxy	2.9	13	37.5	71.8	3.7	21.9	70.2	94.6	3.9	31.1	81.2	97.6
bpwy	3.1	11.8	37.2	72.2	2.7	23	70.3	94.7	3.2	32.9	82.8	98.1
bmcmb	5.4	11.4	34.8	68.7	4.4	23.4	71	95.1	5.8	32.1	82	97.8
bwxy	3.8	12.5	36.3	71.7	3.4	23.6	70.7	94.8	4.3	32	81.6	97.7
bwild	7	11.1	34.7	70.2	5	22.8	70.8	95.4	6.8	27.9	79.4	97.7

Table 22: Empirical rejection percentages (size and size-corrected powers), Model 4 with t_3-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq.

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
`weg`	6.9	13.4	21.3	37.5	6.5	17.2	38.2	60.5	6.3	21	45.2	66.4
`wiid`	10	7.3	14.5	26	8	8.9	27.7	52.5	9.4	12.1	31.6	57.9
`wnid`	6.8	8.4	16.3	31.9	7.2	12.8	36.3	60.3	5.6	14.8	40.8	65.9
`wker`	1.7	9.1	18	33.5	0.8	12.8	37.1	60.6	1.7	13.4	40.4	66.7
`riid`	7.4	7.5	17.2	31.2	8.5	11.1	30.8	53.4	5.8	14	36.3	60.5
`rnid`	4.3	7.7	17.3	30.2	7.1	10.7	30.4	51.3	4.7	14.4	34.9	60
`bxy`	1.6	9.1	19.4	34.8	3.7	11.4	35.3	57.6	2.5	15	41.1	67.1
`bpwy`	0.8	9	18.8	35.4	2.8	12.2	36.6	58.9	1.4	17.2	43.7	68.3
`bmcmb`	2.3	7.5	15.6	29.9	4.2	10.9	33.3	55.7	2.8	13.4	36.7	62.6
`bwxy`	2.8	9.2	19.4	34.7	4.1	11.4	35.3	56.8	3	16.5	42.9	66.8
`bwild`	7.6	8.9	17.8	31.4	6.7	11.1	32.5	55.3	6.5	12.9	37.2	62.5
$n = 300$												
`weg`	7.1	14.8	47.4	78.4	4.4	36.4	83	97.5	7.1	32.2	77.7	95.8
`wiid`	7.3	11.3	42	78	6.7	29.6	80.7	97.6	7.7	26.9	77.3	96.8
`wnid`	7.7	12.6	44.6	80.6	6.3	32	81.4	98.5	7.3	31.6	83.4	98.5
`wker`	3.1	12.9	44.3	77.9	0.9	34.5	84	98.5	4.1	30.1	82.3	98.4
`riid`	6.8	13.6	45.6	77.5	7.3	27.3	75.8	95.7	7	30.1	79.2	97.3
`rnid`	5.6	14.4	46.4	77.5	6.7	26.5	74	94.4	5.5	32	80.4	97.3
`bxy`	3.7	13.3	45.9	77.6	3.3	32.2	80	97.2	4.3	33	83.9	98.2
`bpwy`	3.2	14.5	47.3	79.4	3	32.4	79.8	97.2	3.4	33.2	83.4	98.3
`bmcmb`	5.1	13.5	46.7	78.3	4.7	32.2	79.4	97.9	5.8	29.6	81.1	97.8
`bwxy`	4	14.5	46.4	78.1	3.4	32.2	79.5	96.9	4.6	31.4	82	98.2
`bwild`	7.4	11	40.9	74.4	6.2	29.6	78	96.2	7.2	28.6	80.2	97.6

Table 23: Empirical rejection percentages (size and size-corrected powers), Model 5 with t_3-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq.

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	8.7	11.4	24.6	42.5	4	22.9	45.1	67.3	6.8	19.1	44.9	72.4
wiid	10.1	7.7	15.3	30.7	6.3	10.9	29.6	55.2	8.6	11.7	33.2	58.7
wnid	7	7.8	19.2	38.9	6	13.3	33.3	60	6.1	14.1	41.2	69.7
wker	2.1	9.2	22	43.7	0.5	17.8	42.7	70.3	0.8	14.7	43	74.4
riid	7.5	8	17.6	35.1	5.7	15.5	35.3	59.6	6.5	14.4	37.8	65.7
rnid	5	8.9	19.5	37.9	4.2	14.6	35	60.2	4.4	15.2	38.2	66.3
bxy	2.7	9.8	22.5	43.3	2	16.6	38.6	65.2	1.4	16	44	75
bpwy	1.4	9.3	21.8	42.6	1.5	17	39.6	66.8	0.7	18.6	46	75.9
bmcmb	3.2	8.8	20.6	38.5	2.5	16	38.2	65	2.1	16.5	43.8	72.2
bwxy	3.6	9.4	21.7	41.7	2.4	16.5	40	65.7	2.2	17.9	45	74.7
bwild	7.2	9.2	19.5	37.4	5	14.9	35	60.8	5.9	15.6	40.3	70.1
$n = 300$												
weg	6.6	23.7	59.2	88.7	3.7	32.9	81.7	97.5	5.2	33.5	78.6	95.7
wiid	6.8	20.6	57.4	89.4	5.8	28.6	80	98.5	6.6	27.1	75.3	96.7
wnid	5.9	24.1	63.5	93.4	5.5	30.9	81.1	98.7	6.4	29.3	79.9	97.8
wker	3.3	20.1	59.7	91.6	1.1	31.3	81.9	98.3	2.9	29.5	81.1	97.8
riid	7.1	20.7	58.6	88.8	7.2	26.4	76.8	96.6	6.6	26.2	73.8	96.4
rnid	5.6	20.9	57	88.3	6.6	25.2	75.7	96.5	5.6	26.9	74.6	96.7
bxy	3.4	22.5	60.4	91.4	3.7	29.2	79.7	97.8	3.8	29.6	78.3	97.3
bpwy	2.6	22.2	59.2	91.1	3.3	29.7	79.5	97.9	3.1	29.3	79	97.5
bmcmb	4.9	21.7	60.7	91.4	4.9	30.5	80.2	98	5.2	26.1	77.7	96.8
bwxy	3.6	22.7	61.7	91.8	3.7	28.8	80	97.7	4.3	27.4	77.1	96.8
bwild	7.1	20.2	57	90.2	6.1	26.8	77.2	97.6	6	28.2	79	97.3

Table 24: Empirical rejection percentages (size and size-corrected powers), Model 6 with t_3-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	6	19.9	41.4	64.5	4.1	23.1	47.2	73.3	7.5	13.1	29.5	54.8
wiid	9.7	9.9	24.6	46.3	6.3	15.2	36.6	66.4	10	6.9	19	39.3
wnid	6.5	15.6	35.6	61.3	5.4	15.6	41.9	73.2	5.5	9.6	27.1	54.5
wker	1.9	12.3	32.6	59	0.1	20	47	76.6	1.6	9.3	27.1	56.4
riid	6.4	16.2	35.2	58.9	6.1	16.7	40.7	69.6	6.4	9.5	25.2	49.5
rnid	4.1	17	36.7	60.8	5	17.3	39.8	68.9	4.6	9.6	25	49.8
bxy	1.7	18.1	40	64.5	1.2	19.6	46.6	77.4	2.1	10.6	31.1	58.7
bpwy	0.8	17.1	40.1	64.3	0.8	18.6	44.4	75.5	1.1	10.5	29	57
bmcmb	2.4	14.9	34.9	61.3	1.5	19.2	45.1	75.3	2.5	9	27	55
bwxy	2.5	17	38.7	64.5	1.4	20	45.8	76.3	2.9	10.9	30	58.6
bwild	6.5	13.4	33.8	59	5	16.4	40.7	70.8	5.4	8.8	25.1	53.2
$n = 300$												
weg	6.4	35.4	74.8	94.9	3.8	42.1	86.4	98.3	5.3	19.8	57.7	87.9
wiid	7.5	30.2	72.4	96.7	6	31	80.2	98.7	5.8	14.6	55.2	88.7
wnid	6.3	37.8	81.1	98.7	7.2	29.9	81.8	98.8	4.9	18.4	58.8	92
wker	3.5	33.1	77.8	98.5	0.9	38.2	87.2	99.1	2	18	60.2	91.7
riid	6.6	31.8	74.9	96.8	6.1	32.6	81.2	98.5	5.6	16.7	55.7	88.5
rnid	5.2	33.9	76.3	97.4	5.2	33.2	81.4	98.5	4.6	17.1	57	89.3
bxy	3.4	35.9	78.8	98.3	2.8	34.9	84.8	98.9	2.4	19	61.5	91.6
bpwy	2.8	35.7	79.1	98.2	2.4	34.4	84	98.6	2	18.9	60.2	90.9
bmcmb	4.8	34.7	78	98	4.9	34.1	83.5	99	3.7	16.1	57.3	90.3
bwxy	3.9	37.3	80	98.1	2.9	36.1	85.1	99	2.3	18.1	59.8	90.9
bwild	6.8	34.2	77.1	98	5.4	32.5	83	98.9	5.1	16.3	57.5	90.9

E.4 Results for tests of a joint hypothesis

This appendix considers a joint hypothesis of significance for a bivariate subvector of the vector of coefficients in a linear quantile regression. Specifically, we consider the family of data-generating processes given by $Y = 1 + \sum_{j=1}^{4} X_j + D + \delta_a(U)DX_1 + \gamma(U)X_5 + F^{-1}(U)$, where $\{X_j\}_{j=1}^{4}$, D, U and $\{\delta_a(\cdot) : a \in \mathbb{R}\}$ are as described in Section 4 of the main text, and where $P[\gamma(U) \equiv 0] = 1$, $X_5 \sim N(0, 1)$ and X_5 is independent of $[\ X_1 \quad X_2 \quad X_3 \quad X_4 \quad D \quad U\ ]^\top$. That is, X_5 is an irrelevant regressor. In what follows we consider, for quantiles $\alpha \in \{.25, .50, .75\}$, tests of the null hypothesis $H_0 : \delta_a(\alpha) = \gamma(\alpha) = 0$. We examine the empirical power of these tests against alternatives in which $\delta_a(\alpha) \neq 0$ with $a \in \{.50, 1.00, 1.50\}$ and $\gamma(\alpha) = 0$.

The corresponding simulation results are displayed in Tables 25–30 for samples of sizes $n = 100$ and $n = 300$. These tables present the results of "F-test" implementations of our proposed procedure with pseudo-sample size m given by expression (14) in the main text with

$k = 5$. The corresponding bandwidth h_m is as given by (15) in the main text with constant $c = 1.5$. We also present the results of "F-tests" implemented using most of the other testing methods considered in Section 4 of the main text. Each of these other testing methods, with the exception of `riid`, were implemented by direct computation of the corresponding test statistic using the corresponding estimated asymptotic covariance generated by the `summary.rq` feature of the `quantreg` package. We also examined implementations of `wiid`, `wnid` and `wker` using `anova.rq`, but found that these implementations generated tests having empirical performances that were virtually identical to those of their counterparts implemented using `summary.rq`.

We note that `riid` can only be applied to tests of joint hypotheses using `anova.rq`. We also note that at present there exists no possibility of applying the `rnid` method to tests of joint hypotheses within `quantreg`.

We see that the empirical sizes and size-corrected powers of Wald-type tests induced by the proposed estimate of $\boldsymbol{G}_0(\alpha)$ are competitive with the alternative methods available. These results, along with those reported above in this appendix, supply further evidence of the potential of our method to generate tests with good size and power performance.

Table 25: Empirical rejection percentages (size and size-corrected powers), "F-test". Model 1 with $N(0,1)$-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq unless otherwise indicated.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	6.7	19.7	37.9	61.8	4.5	22.3	43.3	65.8	6.7	19.2	37.1	59.4
wiid	11.9	9.5	22.6	43	8.1	10.3	24.8	51.9	13	7.4	16.8	35.9
wnid	11.9	9.2	24.1	47.4	8.1	12.1	31.2	59.1	13	6.5	18.6	38.2
wker	1.3	9.1	28.8	56.3	0.5	12.8	34.9	64.8	1.2	10.2	28.2	54.3
riid (anova.rq)	3.7	12	31.5	57.7	3	11.9	34	63.5	4.8	8	22.7	49.2
bxy	3.3	10	26	54.1	2.1	13.2	32.9	62.2	3	10.8	26.2	51.3
bpwy	1.5	9.1	26.3	53.4	1.5	12.9	33.1	62.3	1.3	9.7	25.8	51.4
bmcmb	3.8	7.5	23.7	48.4	2.8	12.5	30.8	59.9	3.2	10.1	26	50.2
bwxy	4.7	9.3	26.3	53.2	2.8	13.7	33.1	62	4.3	9.3	25.7	51.2
bwild	8.7	6.9	24.3	47.7	8	11.5	27.9	55.5	9.6	9.7	24.8	48.7
$n = 300$												
weg	5.7	26.5	68.4	94.1	4.7	29	72.7	95.5	5.4	30.6	68	92
wiid	8.7	17	57.7	92.7	6.9	17.2	63.2	95.4	8.8	19.6	59.2	91.2
wnid	8.7	19	64.4	95.4	6.9	20.3	68.7	96.8	8.8	21.6	65.2	94.7
wker	3.7	19.4	67.6	96.3	1.5	20.2	70.9	97.7	3.1	24.5	68.1	95.5
riid (anova.rq)	4.7	21.3	69.8	96.5	4.4	25	73.3	97.8	5.4	22.5	67.4	95.6
bxy	4	20.9	67.3	96.7	4.4	19.1	66.1	96.1	4.3	24.2	66.7	94.3
bpwy	3.7	20.3	67.5	96	4.3	18.8	64.4	95.8	3.5	23.3	66.5	94.2
bmcmb	5.8	20.8	67.7	96	6.5	20	65.3	96	5.9	23.4	63.4	94.4
bwxy	4.8	19.5	66.2	96.1	4.4	19.8	65.8	95.8	4.4	24.5	65.8	94.6
bwild	8.1	21.4	68.1	96.2	7.2	19.5	65.6	95.2	7.6	22	63.4	93.4

Table 26: Empirical rejection percentages (size and size-corrected powers), "F-test". Model 2 with $N(0, 1)$-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq unless otherwise indicated.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	8.1	18.3	27.4	42.7	5	20.7	41.2	63	7.5	21.8	42.5	63.2
wiid	11.1	8	14.2	26.2	7.9	9.4	26	46.7	11.7	10.4	26.4	50.6
wnid	11.1	7.6	12.9	26.4	7.9	10.8	29	50.1	11.7	11.5	30	51.6
wker	1.3	11.1	20.5	35.6	0.6	8.7	25.8	48.3	1.5	13	35.6	62.4
riid (anova.rq)	4.1	7.9	15.9	29.6	4.2	11.7	27.1	48.7	5.2	10.7	30.9	54.4
bxy	3.6	7.6	15.6	29.7	2	9	23.8	45.2	3.1	12.7	33	59.6
bpwy	1.5	8.1	15.8	31.2	1.6	8.7	23.9	44.7	1.5	11.8	32	58
bmcmb	3.7	9.5	14.9	29.5	3.5	10.1	25.6	45	3.5	13.1	32.1	56
bwxy	5.1	8.4	15.3	30.2	3.3	9.3	26.1	46.9	4.9	12.1	32.6	57.8
bwild	8.8	8.4	14.2	30	7.9	9.5	22.1	42.6	8.3	12.1	30.8	55.1
$n = 300$												
weg	6.2	21	55.4	81.5	3.4	34.9	75.7	94	5.4	33.8	71.6	92.1
wiid	8.4	14.8	46.5	81.4	7.3	21.9	67.4	93.2	8.1	24.3	62.2	91.4
wnid	8.4	18.7	56	84.3	7.3	24.2	73.3	95	8.1	25	69.6	94.3
wker	2.7	12.8	46.3	79.3	1.1	26.6	71.7	94.1	2.8	30.7	72.9	94.6
riid (anova.rq)	6.3	14.1	44.2	75.9	4.5	25.8	70.8	94.3	5.3	23.5	69.5	93.3
bxy	4.4	14.3	45.3	73.1	3.6	24.4	67.8	91.6	4	28.1	70	92.2
bpwy	3.7	14.2	46.8	75.9	2.9	25.7	67.1	91.6	3.8	27.3	68.8	92.3
bmcmb	6.3	12.9	43.8	73.8	5.1	23.7	66.3	90.6	5.7	26.5	65	90.9
bwxy	5.8	12.2	41.5	71.1	3.5	23.3	66.1	91.8	4.6	26.7	67.8	91.1
bwild	8.2	13.2	44.3	74.6	5.9	24.6	66.1	91.5	6.8	27.9	69	92.3

Table 27: Empirical rejection percentages (size and size-corrected powers), "F-test". Model 3 with $N(0,1)$-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq unless otherwise indicated.

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	7	16.9	25.8	37.5	4.3	25.5	38.7	57.8	7.2	24.8	47.1	65
wiid	10.2	6.3	11.8	22.7	8	7.7	17.5	32.9	11.3	10.9	24.1	47.3
wnid	10.2	6.3	12	23.5	8	9	20	36.6	11.3	10.9	27.6	47
wker	1.3	6.6	15	29.1	0.4	11.7	24.8	42.6	1.7	11.9	37.4	61.7
riid (anova.rq)	4	6.7	15.1	29.7	4.3	8.6	19.8	38.5	4.8	13.5	31.4	55.1
bxy	2.9	6.8	16.3	28.2	2.7	11.1	21.7	37.8	3.3	12.8	34.9	57.6
bpwy	1.3	7.4	14.5	28.5	2.2	9.5	19.6	36	1.9	12.9	36.5	59.9
bmcmb	3.8	5.6	14.6	23.6	3.2	11	20.1	36.1	3.8	11.4	31.2	52
bwxy	4.8	6.2	13.8	26.7	3.4	11.2	21.8	40	5.5	11.5	32.9	55.5
bwild	7.7	6.9	15.3	27.4	8.5	10	18.4	33.6	8.3	12.9	35.2	56
$n = 300$												
weg	6.3	17	41.3	71.3	4.3	29.7	72.8	95.3	5.7	42.6	85.3	98
wiid	8.4	12.1	35.4	69.2	6.9	22.3	66.4	95.7	6.4	37.5	84	98.7
wnid	8.4	14.4	42	74	6.9	24.4	70.8	96.6	6.4	42	88.7	99.5
wker	2.7	8.4	32.5	68.6	1.3	24.5	71.2	96.9	2.7	45.4	91	99.6
riid (anova.rq)	4.7	10.3	36.8	72.3	5.2	22.9	71.4	96.6	4.4	38.2	87.1	98.7
bxy	4.6	9.5	31.9	64.7	3.7	23.6	67.6	94.3	4	38.4	86.4	98.6
bpwy	3.7	10	34.1	67.4	3.7	23.9	69.5	94.6	2.5	41.9	87.9	98.7
bmcmb	6	10.3	31.2	64.9	5.8	22.7	67.2	93.7	5.3	39.8	86.6	98.4
bwxy	4.9	10.2	34.1	66.3	3.9	24.8	69.6	94.8	4.2	38.7	85.6	98.7
bwild	8.1	9.6	31.2	66.9	7	20.9	64	91.8	7.3	38.7	86.5	98.7

Table 28: Empirical rejection percentages (size and size-corrected powers), "F-test". Model 4 with $N(0,1)$-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq unless otherwise indicated.

$n = 100$		$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$		
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	7.7	19.8	28.9	40.5	4.3	22.4	43.3	66.4	7.8	22.9	40.8	60
wiid	12.9	6	10.7	18.5	7.2	13.7	30.6	56.3	12.1	8.1	18.5	36.5
wnid	12.9	5.8	10.4	22.7	7.2	14.7	33.4	57.5	12.1	8.8	21	41.2
wker	1.7	9.2	18.2	34.2	0.4	9.6	28	51	1.7	9.3	27.4	50.9
riid (anova.rq)	4.9	5.9	14.9	26.7	3.9	12	28.5	53.1	6.1	7.7	22.1	43.7
bxy	3.8	7.7	14.6	25.7	2.2	11.7	27.4	49.4	4.7	8.4	24.3	45.9
bpwy	1.6	7.2	14.6	27.4	2.1	11.1	26.5	50.5	2	10.4	27.9	50.8
bmcmb	4.2	7.4	13.9	26.7	3.4	10.8	27.5	49.6	4.3	10.9	26.6	44.7
bwxy	4.9	8.8	16.3	30.2	3	11.6	29.6	52.5	5.3	11	27.6	48.9
bwild	8.6	8.5	15.2	26.9	6.5	11.7	28.2	47.7	10.4	7.4	21.2	40
$n = 300$												
weg	6.8	19.3	49.3	76.4	4.9	32.1	73.7	93.5	5.5	38.6	81.2	97.8
wiid	8.4	12.7	39.4	73.4	6.3	23.6	67.4	93.4	8.7	24.9	76.4	97.9
wnid	8.4	16.5	47.4	79.9	6.3	26.2	71.5	95.9	8.7	31.8	82.8	99
wker	3.3	13.4	42.6	76.1	1.2	24.8	69.6	95.5	2.2	30.2	85.2	99.1
riid (anova.rq)	5.7	12.1	45.2	79.8	4	26.4	71.4	95	5	30.6	83.4	98.9
bxy	4.4	14.3	42.1	73	3.8	23.9	65	91.8	4.8	26.4	79	98
bpwy	4	14.6	43	74.3	3.6	23.5	64.7	91.7	3.8	28.3	80.5	98.4
bmcmb	7.2	13.8	40	70	5.7	24	66.7	92.5	6.5	26.5	77.7	97.8
bwxy	5.3	14.1	43	73	4.3	24.1	65.6	91.6	5.3	26.1	79	98.4
bwild	9.3	14	40.5	71.5	7.9	22.5	63	91.2	7.3	24.5	76.3	97.9

Table 29: Empirical rejection percentages (size and size-corrected powers), “F-test”. Model 5 with $N(0, 1)$-errors. 1000 Monte Carlo replications; procedure “weg” implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq unless otherwise indicated.

$n = 100$	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	7.2	17.9	29.7	45.9	4.7	21.7	41	64.6	8.2	19.5	41.6	64.6
wiid	12.3	6.1	11.8	23.9	8.7	8.2	21.3	46.2	11.1	9.2	22.6	48
wnid	12.3	5.7	15.2	27.9	8.7	8.9	25.6	51.2	11.1	9.1	26.9	54.2
wker	1.1	7	17	36.3	0.3	9.7	29.2	56.3	1.6	10.6	32.6	63.2
riid (anova.rq)	5.1	6.5	16.9	34.7	4.8	8.9	28.6	54.8	3.9	12.1	33.9	62.3
bxy	3.7	5.6	15.3	31.8	3.2	8.9	27.1	51.4	3.6	10.8	30.8	58.6
bpwy	1.5	6.7	15.4	35.5	2.3	8.9	27	52.9	1.8	12.4	33.4	61.7
bmcmb	3.9	6.4	14.9	30.9	3.8	8	22.7	48.5	3.5	10.3	31.6	58.3
bwxy	5.3	6.2	16.1	34	4.3	7.7	26.1	52.8	4.7	11.9	32.6	60.2
bwild	10.1	7	15.6	32	8.6	10	25.4	49.9	10	9.2	26.6	52.5
$n = 300$												
weg	5	26.2	64	89.9	4.2	34.7	78.4	97	5.5	35.9	79.4	97.4
wiid	8.2	14.7	49.4	85.6	7.7	22.1	72.1	97.8	9.4	25.7	74.1	97.6
wnid	8.2	18.6	58.2	89.4	7.7	23.1	75.5	98.7	9.4	28.7	79	98.4
wker	2.9	18.2	61.4	91.4	0.9	23.6	75.7	98.2	2.9	33.1	83	99.1
riid (anova.rq)	4.1	21.6	60.3	92.8	4.3	27.6	77	97.7	5.2	29.2	81.1	98.6
bxy	3.9	16.7	55.6	87.9	4.3	23.3	71.6	96.6	4.3	29	79.2	98.4
bpwy	3	16.7	57.4	88.2	4.2	22.7	70.8	97	3.3	29.9	79.4	98.3
bmcmb	5.5	17.1	56.5	87.8	6.5	23.7	71.2	96.7	6.2	29.2	78.1	97.8
bwxy	4.7	16.8	57.5	87.5	4.8	24.2	72.6	97.3	5	29.6	79.2	98
bwild	7.3	17	56.5	88	7.8	20.8	68.2	96.2	8.3	29.9	78.5	97.9

Table 30: Empirical rejection percentages (size and size-corrected powers), "F-test". Model 6 with $N(0,1)$-errors. 1000 Monte Carlo replications; procedure "weg" implemented with fixed bandwidth, $c = 1.5$ and $k = 5$; other procedures implemented using summary.rq unless otherwise indicated.

$n = 100$			$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$	
Method/a	0	0.50	1.00	1.50	0	0.50	1.00	1.50	0	0.50	1.00	1.50
weg	9.1	18.5	38.5	61.5	4.5	21.4	44.9	72.2	7.8	14.9	29.6	50
wiid	14	7.8	20.8	42.4	10.5	9	24.1	51.4	12.2	6.8	15.1	35
wnid	14	9.4	22.6	44.6	10.5	10	27.4	57.7	12.2	5.8	13.8	29.5
wker	1.3	11.4	30.3	58.8	0.2	12.6	36.9	68.6	1.6	7.3	20.1	44.6
riid (anova.rq)	5.5	10.7	29.7	55.3	4.3	11.8	35.7	67.9	4.1	7.2	21	43.9
bxy	3.8	11	30.5	56.7	2.4	11.9	34.5	66	3.4	6.1	19.2	41.8
bpwy	2	11.2	28	53.7	1.8	12.9	38.8	68.3	1.5	6.6	17.3	38.7
bmcmb	4.2	10.1	26.7	51.4	3	12.7	35.1	66.5	3.6	6.9	17.2	37.6
bwxy	6	10	26.2	53.1	2.9	11.5	36.5	66.5	4.8	6.4	18.3	41.1
bwild	10.5	11.2	29.4	55.9	8.4	10.6	31.3	60.8	8.9	5.6	15.7	38
$n = 300$												
weg	6.3	35.6	75.2	95.7	3.2	36.6	80.7	98.6	5.5	18.3	49.3	83.1
wiid	9.2	24.1	67.3	94.7	6.3	24.8	73.9	97.2	7.8	11.1	40.9	80.7
wnid	9.2	29.1	73.3	97.4	6.3	29.8	78.8	98.5	7.8	12.8	45.3	86.2
wker	2.9	36.6	81.8	98.9	0.8	30.8	80.4	98.9	3	13.4	48.9	89.3
riid (anova.rq)	5.7	28.4	77.5	97.8	4.6	27.6	77.1	99	4.7	11.2	47.4	88.4
bxy	4.1	31.4	76.3	97.2	3.6	28.7	76.8	98	3.9	11.1	44.4	85.2
bpwy	3.5	31.8	76.4	97.4	3.3	27.9	76.3	97.6	3	12.7	48	86.3
bmcmb	5.8	30.8	75.5	97.4	4.8	30.5	77.7	98.5	5	12.2	46.9	87.1
bwxy	4.5	31.1	75.4	97.3	3.3	29	77.5	98.1	4.7	11.9	44.9	85.5
bwild	7.6	29.4	74.2	97.3	6.6	25.9	73.1	97.3	7.8	11.3	43.5	84.3

F Additional Material on the Empirical Example

We present in this appendix further details regarding the empirical application considered in Section 5 of our paper. Recall that we are concerned with estimating the distinct effects of treatment for experimental subjects at each quantile in a grid of 300 evenly spaced points in [.20, .80]. This is done in the context of the quantile-regression model

$$F^{-1}_{\log T|\boldsymbol{X}}(\alpha) = \boldsymbol{X}^\top \boldsymbol{\beta}(\alpha), \tag{59}$$

where $\alpha \in [.20, .80]$, where T denotes the duration of unemployment in weeks and where the regressors contained in $\boldsymbol{X}$ include a constant term, an indicator for assignment to treatment and various demographic or socioeconomic control variables listed in Koenker and Xiao (2002, p. 1603).

We depart in this appendix from the general question of treatment-effect heterogeneity considered in Section 5 of the main text by focusing on the questions of whether the effects of treatment by quantile differ significantly according to the age of the participants and also according to whether participants have some expectation to be recalled to a previously held job, although not to the extent of having a definite date of recall within 60 days of filing their applications for unemployment insurance (UI) benefits (Corson et al., 1992, p. 9). We note in this connection that of the 6384 participants in this experiment, 3460 (54%) were under the age of 35, while 753 (12%) indicated to the experimenters some expectation of being recalled to previous employment. Participants in the latter category were assumed by the experimenters to be similar to claimants with no stated expectation of returning to a previous job in terms of their assumed response to a promised bonus payment upon securing new employment within the qualifying period. On the other hand, UI claimants who indicated both an expectation of recall and a definite recall date were disqualified from participation in the experiment as their stated confidence in returning to work was assumed to make their hypothetical behavioral response to treatment systematically different from UI claimants expressing less confidence in returning to full-time employment.

Figure 1 displays, in the context of the model given above in (59), estimated differences in treatment effects between workers younger than 35 and those aged 35 and older at the time of the experiment. These estimated differences in treatment effects are plotted for each quantile in a grid of 300 points in [.20, .80]. The shaded area in Figure 1 indicates the union of 90% confidence intervals for the estimated difference in treatment effects at each quantile. These confidence intervals are computed using our proposed method with data-driven bandwidth given above in (54) and where the pseudo-sample size m is given by (14) in the main text with $k = 5$. These confidence intervals imply that workers younger than 35 tend to exit unemployment as a result of the treatment significantly more quickly than workers 35 and older for nearly all quantiles in the interval [.50, .80].

Estimated differences in treatment effects between workers with some expectation of being recalled to previous employment and those with no such expectation are displayed in Figure 2 for each quantile in a grid of 300 points in [.20, .80]. The shaded area in Figure 2, like that in Figure 1, denotes the union of 90% confidence intervals for the estimated difference in treatment effects, pointwise by quantile. These confidence intervals are computed in the same way as was done when generating the shaded area appearing in Figure 1. The confidence intervals in Figure 2 imply that the treatment has the effect of actually increasing unemployment durations for workers expecting a recall to a previous job for nearly all quantiles in the interval [.43, .74]. All this suggests that the cash bonus may not be as relevant as originally hoped for those claimants who indicated some degree of confidence in the temporary nature of their current spell of unemployment. In other words, this result suggests that the inclusion of these claimants in the experiment is as potentially problematic as the hypothetical inclusion of those excluded claimants who indicated both an expectation of recall and a definite recall date.

In summary, we have used our proposed method of inference to show that the treatment tends to cause participants having some expectation of being recalled to a previous job to exit unemployment more slowly than those not expecting to be recalled. This result further illustrates the utility, in terms of understanding behavioral responses to changes in unemployment insurance rules, of accounting for heterogeneity in treatment effects via the introduction of simple interaction terms in quantile-regression models.

Figure 1: Pennsylvania reemployment bonus experiment: 6384 observations. Differences in estimated treatment effects by quantile for workers younger than 35 and workers aged 35 and older, α-quantile regressions, $\alpha \in [.20, .80]$. The shaded area denotes the union of pointwise 90% confidence intervals, computed according to our proposal with data-driven bandwidth and $k = 5$, for each of 300 quantiles in [.20, .80]. Dotted vertical lines denote the .25-, .35-, .50-, .65- and .75-quantiles.

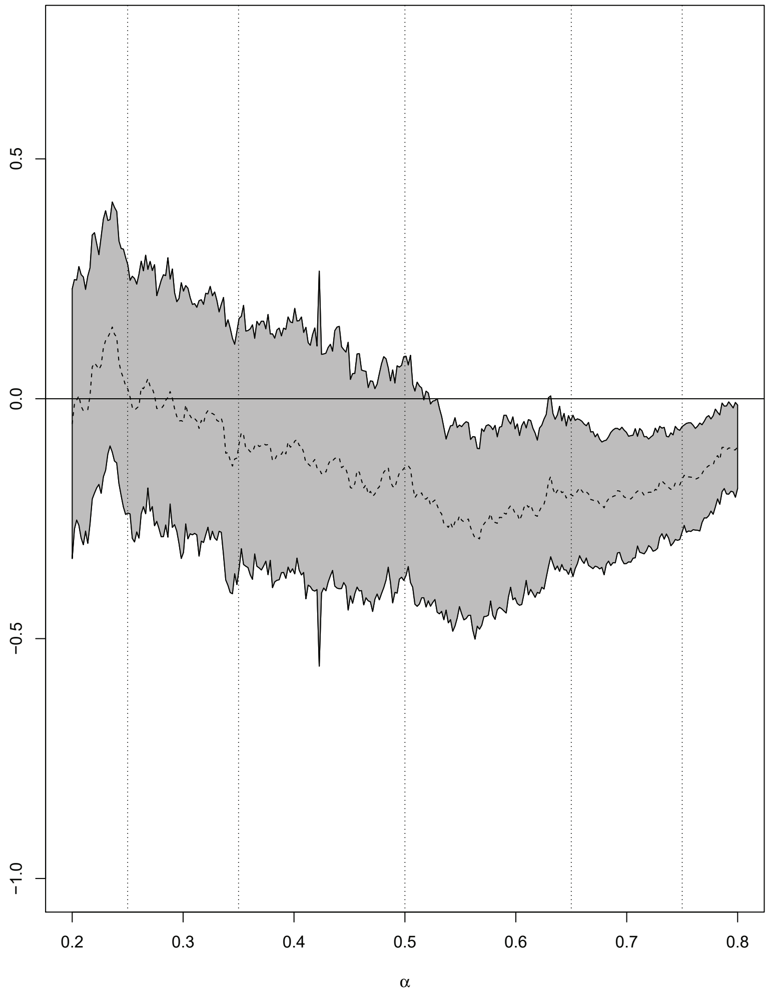
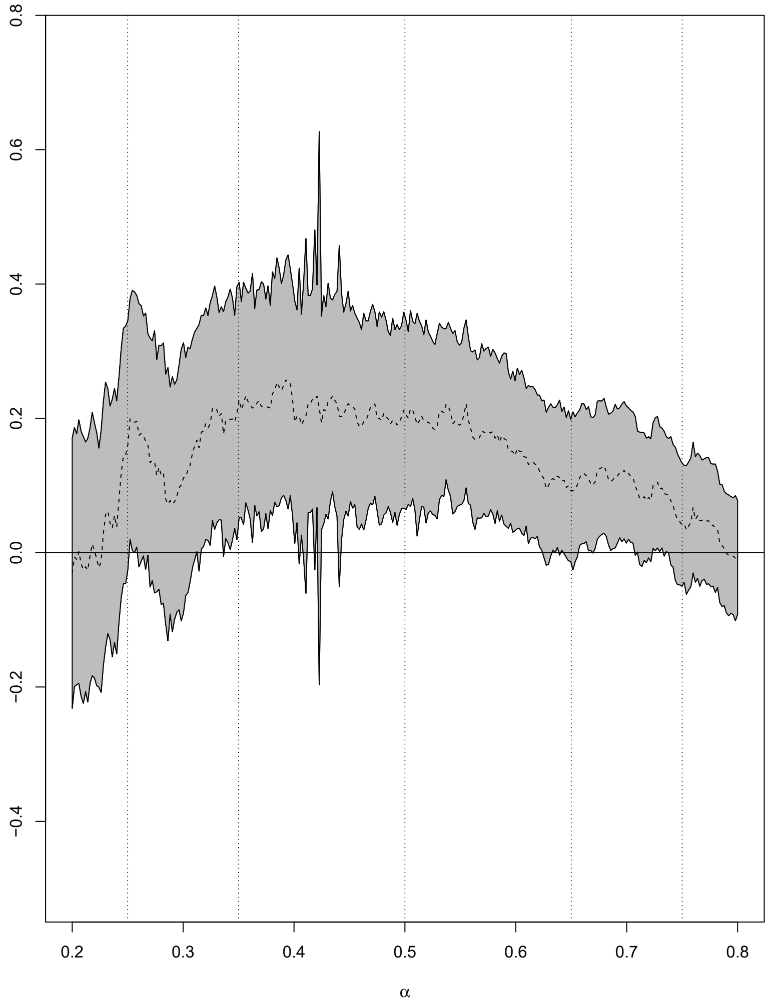


Figure 2: Pennsylvania reemployment bonus experiment: 6384 observations. Differences in estimated treatment effects by quantile for workers expecting and not expecting to be recalled to a previous job, α-quantile regressions, $\alpha \in [.20, .80]$. The shaded area denotes the union of pointwise 90% confidence intervals, computed according to our proposal with data-driven bandwidth and $k = 5$, for each of 300 quantiles in [.20, .80]. Dotted vertical lines denote the .25-, .35-, .50-, .65- and .75-quantiles.



References

Angrist, J., V. Chernozhukov, and I. Fernández-Val (2006). Quantile regression under misspecification, with an application to the u.s. wage structure. *Econometrica 74*, 539–563.

Corson, W., P. Decker, S. Dunstan, and S. Kerachsky (1992). Pennsylvania Reemployment Bonus Demonstration: Final Report. Unemployment Insurance Occasional Paper 92-1, United States Department of Labor.

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

Giné, E., V. Koltchinskii, and L. Sakhanenko (2004). Kernel density estimators: convergence in distribution for weighted sup-norms. *Probability Theory and Related Fields 130*, 167–198.

Hendricks, W. and R. Koenker (1992). Hierarchical spline models for conditional quantiles and the demand for electricity. *Journal of the American Statistical Association 87*, 58–68.

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

Mason, D. M. and J. W. H. Swanepoel (2011). A general result on the uniform in bandwidth consistency of kernel-type function estimators. *Test 20*, 72–94.

Nolan, D. and D. Pollard (1987). U-processes: Rates of convergence. *Annals of Statistics 15*, 780–799.

Portnoy, S. (2012). Nearly root-n approximation for regression quantile processes. *Annals of Statistics 40*, 1714–1736.

Powell, J. L. (1991). Estimation of monotonic regression models under quantile restrictions. In W. A. Barnett, J. L. Powell, and G. E. Tauchen (Eds.), *Nonparametric and semiparametric methods in econometrics and statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, Chapter 14, pp. 357–384. Cambridge, U.K.: Cambridge University Press.