Online Appendix

A Comparison with Horowitz (1998)

It is possible to study Horowitz's (1998) smoothed QR estimator using the same tools we employ to document the asymptotic behavior of our convolution-type kernel QR estimator. Let $\tau \in (0,1)$ and Assumptions X, Q and K hold. Let now $\mathfrak{R}_h^{(j)}(b;\tau) := \mathbb{E}[\widehat{\mathfrak{R}}_h^{(j)}(b;\tau)]$ for $j = 0,1,2$ and $\mathfrak{b}_h(\tau) := \arg\min_b \mathfrak{R}_h(b;\tau)$. The latter corresponds to the unique solution of the first-order condition $\mathfrak{R}_h^{(1)}(\mathfrak{b}_h(\tau);\tau) = 0$ for h small enough. It turns out that $\widehat{\mathfrak{R}}_h^{(2)}(b;\tau) = \frac{1}{n}\sum_{i=1}^n X_i X_i' \,\kappa\big(e_i(b)/h\big)$, where $\kappa(t) := 2k(t) + tk^{(1)}(t)$. Integrating by parts shows that $\int t^j t k^{(1)}(t)\,\mathrm{d}t = -(j+1)\int t^j k(t)\,\mathrm{d}t$, so that $\kappa(\cdot)$ is a kernel function with the same order than $k(\cdot)$. Accordingly, Horowitz's (1998) smoothed objective function also satisfies Lemma 1.

Along the same lines as in the proof of Theorem 1,

$$\begin{aligned}\mathfrak{b}_h(\tau) - \beta_h(\tau) &= -\Big[\mathfrak{R}^{(2)}(\beta(\tau);\tau) + o(1)\Big]^{-1} \mathfrak{R}_h^{(1)}(\beta_h(\tau);\tau)\\ &= \big[D^{-1}(\tau) + o(1)\big]\,\mathbb{E}\left[X\,\frac{e\big(\beta_h(\tau)\big)}{h}\,k\bigg(-\frac{e\big(\beta_h(\tau)\big)}{h}\bigg)\right].\end{aligned}$$

Because $k(\cdot)$ is symmetric and of order $s+1$, Theorem 1 implies that

$$\begin{aligned}\mathbb{E}\left[X\,\frac{e\big(\beta_h(\tau)\big)}{h}\,k\bigg(-\frac{e\big(\beta_h(\tau)\big)}{h}\bigg)\right] &= h\,\mathbb{E}\Big[X\int zk(z) f\big(X'\beta_h(\tau) + hz \,|\, X\big)k(z)\,\mathrm{d}z\Big]\\ &= h^{s+1}\frac{\int z^{s+1}k(z)\mathrm{d}z}{s!}\,\mathbb{E}\Big[Xf^{(s)}\big(X'\beta(\tau)\,|\,X\big)\Big] + o(h^{s+1})\end{aligned}$$

and that $\mathfrak{b}_h(\tau) = \beta_h(\tau) + (s+1)h^{s+1}B(\tau) + o(h^{s+1}) = \beta(\tau) + sh^{s+1}B(\tau) + o(h^{s+1})$. This means that $\mathfrak{b}_h(\tau) - \beta(\tau) = -s\big(\beta_h(\tau) - \beta(\tau)\big) + o(h^{s+1})$, so that Horowitz's (1998) smoothing approach amplifies the bias by a factor $-s$ asymptotically.

We next consider the asymptotic covariance matrix of Horowitz's smoothed QR estimator. Consider $b_h(\tau) = \beta(\tau) + O(h^2)$ and let $\Delta_h(\tau) := \widehat{\mathfrak{R}}_h^{(1)}(b_h(\tau);\tau) - \widehat{R}_h^{(1)}(b_h(\tau)b;\tau)$. We first observe that $\mathbb{V}\big[\sqrt{n}\Delta_h(\tau)\big] = O(h)$, whereas using the fact $y = X'b_h(\tau) - hu$ yields under Assumption Q2 that

$$\begin{aligned}n\,\mathrm{Cov}\big(\widehat{R}_h^{(1)}(b_h(\tau);\tau), \Delta_h(\tau)\big) &= \mathbb{E}\left\{XX'\!\int\left[\tau - K\bigg(-\frac{e\big(b_h(\tau)\big)}{h}\bigg)\right]\frac{e\big(b_h(\tau)\big)}{h}\,k\bigg(-\frac{e\big(b_h(\tau)\big)}{h}\bigg)\,f(y|X)\,\mathrm{d}y\right\}\\ &= h\int[K(u) - \tau]\,uk(u)\,\mathrm{d}u\,\mathbb{E}\left[XX'f\big(X'b_h(\tau)|X\big)\mathrm{d}y\right] + O(h^2)\\ &= h\int_0^\infty[K(u) - K(-u)]\,uk(u)\,\mathrm{d}u\,\mathbb{E}\left[XX'f\big(X'\beta(\tau)|X\big)\,\mathrm{d}y\right] + O(h^2)\end{aligned}$$

for any symmetric kernel $k(\cdot)$. Because $\int_0^\infty [K(u) - K(-u)]\,uk(u)\,\mathrm{d}u > 0$ for second-order and *bona fide* higher-order kernels, there exists a symmetric positive M_τ such that

$$\mathbb{V}\left[\sqrt{n}\,\widehat{\mathfrak{R}}_h^{(1)}(\beta(\tau);\tau)\right] = \mathbb{V}\left[\sqrt{n}\,\widehat{R}_h^{(1)}(\beta(\tau);\tau)\right] + h\big[M_\tau + o(1)\big].$$

It then follows from Lemma 1 that $\mathfrak{R}_h^{(2)}\big(\mathfrak{b}_h(\tau);\tau\big) = D(\tau) + o(1)$, and hence

$$\mathbb{V}\Big[\mathfrak{R}_h^{(2)}\big(\mathfrak{b}_h(\tau);\tau\big)^{-1}\,\widehat{\mathfrak{R}}_h^{(1)}\big(\mathfrak{b}_h(\tau);\tau\big)\Big] = \mathbb{V}\Big[R_h^{(2)}\big(\beta_h(\tau);\tau\big)^{-1}\,\widehat{R}_h^{(1)}\big(\beta_h(\tau);\tau\big)\Big] + h\,D^{-1}(\tau)M_\tau D^{-1}(\tau) + o(h).$$

Horowitz's estimator has a Bahadur-Kiefer representation as in Theorem 2, ergo the above equality shows that the asymptotic covariance matrix of Horowitz's estimator is larger than ours at the second order.

B Technical proofs

Proof of Lemma 1 Under Assumption Q2, a Taylor expansion with integral remainder yields

$$f(v + hz \,|\, x) = \sum_{\ell=0}^{s} f^{(\ell)}(v \,|\, x)\,\frac{(hz)^\ell}{\ell!} + \frac{(hz)^s}{(s-1)!}\int_0^1 (1-w)^{s-1}\Big[f^{(s)}(v + whz \,|\, x) - f^{(s)}(v \,|\, x)\Big]\,\mathrm{d}w.$$

(i) Assumption K1 ensures that

$$\begin{aligned}
\mathbb{E}[k_h(v - Y) \,|\, x] - f(v \,|\, x) &= \int k_h(v-y)\, f(y \,|\, x)\,\mathrm{d}y - f(v \,|\, x) \\
&= \int k(z)\Big[f(v + hz \,|\, x) - f(v \,|\, x)\Big]\,\mathrm{d}z \\
&= \int_0^1 (1-w)^{s-1}\int \tfrac{(hz)^s}{(s-1)!}\, k(z)\Big[f^{(s)}(v + whz \,|\, x) - f^{(s)}(v \,|\, x)\Big]\,\mathrm{d}z\,\mathrm{d}w \qquad (23)
\end{aligned}$$

through a change of variables $y = v + hz$. Now, the check function is such that

$$\int \rho_\tau(v)\,\mathrm{d}G(v) = (1-\tau)\int_{-\infty}^0 G(v)\,\mathrm{d}v + \tau\int_0^\infty [1 - G(v)]\,\mathrm{d}v$$

for any arbitrary cdf G, and hence

$$R(b;\tau) = \int \bigg\{(1-\tau)\int_{-\infty}^0\int_{-\infty}^{t+x'b} f(v \,|\, x)\,\mathrm{d}v\,\mathrm{d}t + \tau\int_0^\infty\int_{t+x'b}^\infty f(v \,|\, x)\,\mathrm{d}v\,\mathrm{d}t\bigg\}\,\mathrm{d}F_X(x),$$

where $F_X(x)$ is the cdf of X. Similarly,

$$R_h(b;\tau) = \int \bigg\{(1-\tau)\int_{-\infty}^0\int_{-\infty}^{t+x'b} \mathbb{E}[k_h(v - Y) \,|\, x]\,\mathrm{d}v\,\mathrm{d}t + \tau\int_0^\infty\int_{t+x'b}^\infty \mathbb{E}[k_h(v - Y) \,|\, x]\,\mathrm{d}v\,\mathrm{d}t\bigg\}\,\mathrm{d}F_X(x).$$

It follows from (23) that

$$\begin{aligned}
L_1 &:= \bigg|\int_{-\infty}^0\int_{-\infty}^{t+x'b}\Big\{\mathbb{E}[k_h(v - Y) \,|\, x] - f(v \,|\, x)\Big\}\mathrm{d}v\,\mathrm{d}t\bigg| \\
&= \bigg|\int_0^1 (1-w)^{s-1}\int \tfrac{(hz)^s}{(s-1)!}\, k(z)\int_{-\infty}^0\int_{-\infty}^{t+x'b}\Big[f^{(s)}(v + whz \,|\, x) - f^{(s)}(v \,|\, x)\Big]\,\mathrm{d}v\,\mathrm{d}t\,\mathrm{d}z\,\mathrm{d}w\bigg| \\
&= \bigg|\int_0^1 (1-w)^{s-1}\int \tfrac{(hz)^s}{(s-1)!}\, k(z)\,\Big[f^{(s-2)}(x'b + whz \,|\, x) - f^{(s-2)}(x'b \,|\, x)\Big]\,\mathrm{d}z\,\mathrm{d}w\bigg|
\end{aligned}$$

given that $\int \big|z^{s+1}k(z)\big|\,\mathrm{d}z < \infty$ by Assumption K1 and that $f^{(s-2)}(\cdot \,|\, \cdot)$ is Lipschitz. Analogously, $\Big|\int_0^\infty\int_{t+x'b}^\infty \mathbb{E}[k_h(v - Y) \,|\, x] - f(v \,|\, x)\,\mathrm{d}v\,\mathrm{d}t\Big| \le C\,h^{s+1}$, establishing the result.

(ii) By the definitions of $R(b;\tau)$ and $R_h(b;\tau)$, it follows from the Lebesgue dominated convergence theorem that

$$R^{(1)}(b;\tau) = \mathbb{E}\Big[X\Big(F(X'b\,|\,X) - \tau\Big)\Big] = \int x\Big[\int_{-\infty}^{x'b} f(y\,|\,x)\,\mathrm{d}y - \tau\Big]\,\mathrm{d}F_X(x),$$

and that

$$R_h^{(1)}(b;\tau) = \mathbb{E}\Big\{X\left[K\left(\frac{X'b - Y}{h}\right) - \tau\right]\Big\} = \int x\Big\{\int_{-\infty}^{x'b} \mathbb{E}\big[k_h(v - Y)\,|\,x\big]\,\mathrm{d}v - \tau\Big\}\,\mathrm{d}F_X(x). \tag{24}$$

In view that $\int z^s k(z)\,\mathrm{d}z = 0$ and $\int \left|z^{s+1}k(z)\right|\mathrm{d}z < \infty$, integrating (23) yields

$$\begin{aligned}
L_2 :&= \left|\int_{-\infty}^{x'b} \mathbb{E}\big[k_h(v - Y)\,|\,x\big] - f(v\,|\,x)\,\mathrm{d}v\right| \\
&= \left|\int_0^1 (1-w)^{s-1}\int \tfrac{(hz)^s}{(s-1)!}\,k(z)\int_{-\infty}^{x'b}\Big[f^{(s)}(v + whz\,|\,x) - f^{(s)}(v\,|\,x)\Big]\,\mathrm{d}v\,\mathrm{d}z\,\mathrm{d}w\right| \\
&= \left|\int_0^1 (1-w)^{s-1}\int \tfrac{(hz)^s}{(s-1)!}\,k(z)\Big[f^{(s-1)}(x'b + whz\,|\,x) - f^{(s-1)}(x'b\,|\,x)\Big]\,\mathrm{d}z\,\mathrm{d}w\right| \\
&= \left|\int_0^1 w(1-w)^{s-1}\int \tfrac{(hz)^{s+1}}{(s-1)!}\,k(z)\int_0^1 f^{(s)}(x'b + twhz\,|\,x)\mathrm{d}t\,\mathrm{d}z\,\mathrm{d}w\right| \le C\,h^{s+1},
\end{aligned} \tag{25}$$

uniformly given that $f^{(s)}$ is bounded. The result then readily follows from Assumption X.

(iii) Differentiating $R^{(1)}(b;\tau)$ with respect to b results in

$$R^{(2)}(b;\tau) = \mathbb{E}\big[XX'f(X'b\,|\,X)\big] = \int xx'f(x'b\,|\,x)\,\mathrm{d}F_X(x)$$

and, likewise,

$$R_h^{(2)}(b;\tau) = \mathbb{E}\big[XX'k_h(X'b - Y)\big] = \int xx'\,\mathbb{E}\big[k_h(x'b - Y)\,|\,x\big]\,\mathrm{d}F_X(x).$$

Setting $v = x'b$ in (23) then yields

$$\begin{aligned}
\left\|R_h^{(2)}(b;\tau) - R^{(2)}(b;\tau)\right\| &\le C\,\left|\mathbb{E}\big[k_h(v - Y)\,|\,x\big] - f(v\,|\,x)\right| \\
&\le C\,h^s\int |z^sK(z)| \sup_{(x,y)\in\mathbb{R}^{d+1}}\ \sup_{t:\,|t|\le hz}\left|f^{(s)}(y + t\,|\,x) - f^{(s)}(y\,|\,x)\right|\mathrm{d}z = o(h^s),
\end{aligned}$$

under Assumptions X and Q2 by the Lebesgue dominated convergence theorem, as stated.

(iv) Recall that

$$R_h^{(2)}(b;\tau) = \mathbb{E}\big[XX'k_h(X'b - Y)\big] = \int k(z)\int xx'f(x'b + hz\,|\,x)\,\mathrm{d}F_X(x)\,\mathrm{d}z.$$

Under Assumption Q2, it ensues from $f(\cdot\,|\,\cdot)$ being Lipschitz that

$$\left\|R_h^{(2)}(b + \delta;\tau) - R_h^{(2)}(b;\tau)\right\| \le C\int |k(z)|\int \|xx'\|\,|x'\delta|\,\mathrm{d}F_X(x)\,\mathrm{d}z \le C\,\|\delta\|,$$

uniformly in (b, h, δ, τ), completing the proof. ■

Proof of Lemma 3 For $\eta > 0$,

$$\left\{\sup_{(\tau,h)} \left\|\widehat{\beta}_h(\tau) - \beta_h(\tau)\right\| \geq 2\eta\right\} = \bigcup_{(\tau,h)} \left\{\left\|\widehat{\beta}_h(\tau) - \beta_h(\tau)\right\| \geq 2\eta\right\}$$
$$\subset \bigcup_{(\tau,h)} \left\{\inf_{\{b:\|b-\beta_h(\tau)\|\geq 2\eta\}} \widehat{\mathcal{R}}_h(b;\tau) \leq \inf_{\{b:\|b-\beta_h(\tau)\|\leq 2\eta\}} \widehat{\mathcal{R}}_h(b;\tau)\right\}$$
$$\subset \bigcup_{(\tau,h)} \left\{\inf_{\{b:\|b-\beta_h(\tau)\|\geq 2\eta\}} \widehat{\mathcal{R}}_h(b;\tau) \leq \widehat{\mathcal{R}}_h\big(\beta_h(\tau);\tau\big)\right\}$$
$$= \bigcup_{(\tau,h)} \left\{\inf_{\{b:\|b-\beta_h(\tau)\|\geq 2\eta\}} \widehat{\mathcal{R}}_h(b;\tau) \leq 0\right\},$$

given that $\widehat{\mathcal{R}}_h\big(\beta_h(\tau);\tau\big) = 0$. Theorem 1 ensures that

$$\left\{b: \|b - \beta_h(\tau)\| \geq 2\eta\right\} \subset \left\{b: \|b - \beta(\tau)\| + \sup_{(\tau,h)} \|\beta_h(\tau) - \beta(\tau)\| \geq 2\eta\right\}$$
$$\subset \left\{b: \|b - \beta(\tau)\| + O(\bar{h}_n^{s+1}) \geq 2\eta\right\}$$
$$\subset \left\{b: \|b - \beta(\tau)\| \geq \eta\right\}$$

for all (τ, h) provided that n is large enough. This means that

$$\left\{\sup_{(\tau,h)} \left\|\widehat{\beta}_h(\tau) - \beta_h(\tau)\right\| \geq 2\eta\right\} \subset \bigcup_{(\tau,h)} \left\{\inf_{\{b:\|b-\beta(\tau)\|\geq \eta\}} \widehat{\mathcal{R}}_h(b;\tau) \leq 0\right\}.$$

As $t \mapsto \rho_\tau(t)$ is 1-Lipschitz, it follows from

$$\widehat{R}_h(b;\tau) = \frac{1}{nh}\sum_{i=1}^{n} \int \rho_\tau(t)\, k\left(\frac{t - (Y_i - X_i'b)}{h}\right) \mathrm{d}t = \frac{1}{n}\sum_{i=1}^{n} \int \rho_\tau(Y_i - X_i'b + hz)\, k(z)\, \mathrm{d}z$$

that

$$\left|\widehat{R}_h(b;\tau) - \widehat{R}(b;\tau)\right| = \left|\frac{1}{n}\sum_{i=1}^{n} \int \Big[\rho_\tau(Y_i - X_i'b + hz) - \rho_\tau(Y_i - X_i'b)\Big]\, k(z)\, \mathrm{d}z\right| \leq h \int |z\, k(z)|\, \mathrm{d}z < \infty,$$

for all b, τ and h by Assumption K1. Theorem 1 and the Lipschitz property of $b \mapsto \widehat{R}(b;\tau)$ then ensures that $\widehat{\mathcal{R}}_h(b;\tau) \geq \widehat{\mathcal{R}}(b;\tau) - C\,h$ uniformly in b and τ, so that

$$\left\{\sup_{(\tau,h)} \left\|\widehat{\beta}_h(\tau) - \beta_h(\tau)\right\| \geq 2\eta\right\} \subset \bigcup_{(\tau,h)} \left\{\inf_{\{b:\|b-\beta(\tau)\|\geq \eta\}} \widehat{\mathcal{R}}(b;\tau) \leq C\,h\right\}.$$

The next step is a convexity argument. We first perform the change of variables $b = \beta(\tau) + \rho u$ with $\|u\| = 1$ and $\rho \geq \eta$. In view that $b \mapsto \widehat{\mathcal{R}}(b;\tau)$ is convex with $\widehat{\mathcal{R}}\big(\beta(\tau);\tau\big) = 0$,

$$\frac{\eta}{\rho}\,\widehat{\mathcal{R}}\big(\beta(\tau) + \rho u;\tau\big) = \frac{\eta}{\rho}\,\widehat{\mathcal{R}}\big(\beta(\tau) + \rho u;\tau\big) + \left(1 - \frac{\eta}{\rho}\right)\widehat{\mathcal{R}}\big(\beta(\tau);\tau\big) \geq \widehat{\mathcal{R}}\big(\beta(\tau) + \eta u;\tau\big).$$

It follows from the above inequality that

$$\left\{\inf_{\{b:\|b-\beta(\tau)\|\geq \eta\}} \widehat{\mathcal{R}}(b;\tau) \leq C\,h\right\} \subset \left\{\inf_{\{b:\|b-\beta(\tau)\|= \eta\}} \widehat{\mathcal{R}}(b;\tau) \leq C\,h\right\},$$

and hence

$$\bigcup_{(\tau,h)}\left\{\left\|\widehat{\beta}_h(\tau)-\beta_h(\tau)\right\|\ge 2\eta\right\}\subset\bigcup_{\tau}\left\{\inf_{\{b:\|b-\beta(\tau)\|=\eta\}}\widehat{\mathcal{R}}(b;\tau)\le C\,\bar{h}_n\right\}$$
$$\subset\left\{\inf_{\tau}\inf_{\{b:\|b-\beta(\tau)\|=\eta\}}\left[\widehat{\mathcal{R}}(b;\tau)-\mathcal{R}(b;\tau)\right]\le C\,\bar{h}_n-\inf_{\tau}\inf_{\{b:\|b-\beta(\tau)\|=\eta\}}\mathcal{R}(b;\tau)\right\}.$$

We next establish an upper bound for $C\,\bar{h}_n-\inf_{\tau\in[\underline{\tau},\bar{\tau}]}\inf_{\{b:\|b-\beta(\tau)\|=\eta\}}\mathcal{R}(b;\tau)$ using the fact that the eigenvalues of $\mathcal{R}^{(2)}(b;\tau)$ are bounded away from 0 uniformly in b, for $\|b-\beta(\tau)\|\le 1$ and $\tau\in[\underline{\tau},\bar{\tau}]$. Given that $R^{(1)}\big(\beta(\tau),\tau\big)=0$, a second-order Taylor expansion of $\mathcal{R}(b;\tau)=R(b;\tau)-R\big(\beta(\tau);\tau\big)$ gives way to

$$\mathcal{R}(b;\tau)=0+\big(b-\beta(\tau)\big)'\left[\int_0^1(1-t)\mathcal{R}^{(2)}\big(\beta(\tau)+t\big[b-\beta(\tau)\big];\tau\big)\,\mathrm{d}t\right]\big(b-\beta(\tau)\big)\ge C\,\eta^2$$

for all b such that $\|b-\beta(\tau)\|=\eta$. This means that, for any $\eta_2=\eta-\epsilon_2<\eta$ with conformable ϵ_2 and $\bar{h}_n$ small enough,

$$\bigcup_{(\tau,h)}\left\{\left\|\widehat{\beta}_h(\tau)-\beta_h(\tau)\right\|\ge 2\eta\right\}\subset\left\{\sup_{\tau\in[\underline{\tau},\bar{\tau}]}\sup_{\{b:\|b-\beta(\tau)\|=\eta\}}\left|\widehat{\mathcal{R}}(b;\tau)-\mathcal{R}(b;\tau)\right|\ge C\,\eta_2^2\right\}.$$

Now, let $Z_i=(Y_i,X_i')'$, $\theta=(\tau,b')'$ and $g_1(Z_i,\theta)=\rho_\tau(Y_i-X_i'b)-\rho_\tau\big(Y_i-X_i'\beta(\tau)\big)$, so that

$$\widehat{\mathcal{R}}(b;\tau)-\mathcal{R}(b;\tau)=\frac{1}{n}\sum_{i=1}^n\left\{g_1(Z_i,\theta)-\mathbb{E}\big[g_1(Z_i,\theta)\big]\right\}.$$

Under Assumption X, it follows from $\eta\le 1$ that, for all b such that $\|b-\beta(\tau)\|=\eta$ and $\tau\in[\underline{\tau},\bar{\tau}]$,

$$|g_1(Z_i,\theta)|\le\|X_i\|\,\|b-\beta(\tau)\|\le C,$$

implying that $\mathbb{V}\big(g_1(Z_i,\theta)\big)\le\sigma^2\le C$. Observe also that pairing Assumption X with the Lipschitz conditions on $\tau\mapsto\beta(\tau)$ in Assumption Q1 and on $\tau\mapsto\rho_\tau(u)$ entails, for all admissible z,

$$|g_1(z,\theta_1-g_1(z,\theta_2)|\le C\,\|\theta_1-\theta_2\|\,, \tag{26}$$

where $\|\theta\|^2=\|b\|^2+|\tau|^2$. Next, for $\delta>0$, let θ_j, with $j=1,\ldots,J(\delta)\le C\,\delta^{-(d+1)}$, be such that

$$\Theta=\Big\{\theta=(b,\tau):\,\tau\in[\underline{\tau},\bar{\tau}],\|b-\beta(\tau)\|=\eta_1\Big\}\subset\bigcup_{j=1}^{J(\delta)}\mathcal{B}(\theta_j,\delta),$$

where $\mathcal{B}(\theta_j,\delta)$ is the $\|\cdot\|$-ball with center θ_j and radius δ. Define $\underline{g}_{1j}(\cdot)$ and $\bar{g}_{1j}(\cdot)$ respectively as $\underline{g}_{1j}(z):=\inf_{\theta\in\mathcal{B}(\theta_j,\delta)}g_1(z,\theta)$ and $\bar{q}_{1j}(z)=\sup_{\theta\in\mathcal{B}(\theta_j,\delta)}g_1(z,\theta)$, so that $\big\{g_1(\cdot,\theta):\,\theta\in\mathcal{B}(\theta_j,\delta)\big\}\subset[\underline{g}_{1j},\bar{g}_{1j}]$. Let $\mathcal{G}_{1,\Theta}:=\big\{g_1(\cdot,\theta):\,\theta\in\Theta\big\}\subset\bigcup_{j=1}^{J(\delta)}[\underline{g}_{1j},\bar{g}_{1j}]$. It follows from (26) that $\big|\bar{g}_{1j}(z)-\underline{g}_{1j}(z)\big|\le C\delta\le C$ and $\mathbb{E}\Big[\big|\bar{g}_{1j}(Z_i)-\underline{g}_{1j}(Z_i)\big|^2\Big]\le C\delta^2$. By conditions (i) and (ii) in Lemma 2, it follows from (18) that setting $H(\delta)=-(d+1)\ln\delta+C$ leads to

$$\Pr\left(\sup_{\theta\in\Theta}\left|\widehat{\mathcal{R}}(b;\tau)-\mathcal{R}(b;\tau)\right|\ge C\,\frac{1+\sqrt{r}+r/\sqrt{n}}{\sqrt{n}}\right)\le\exp(-r).$$

This means that, for n large enough with respect to η_2^2,

$$\Pr\left(\sup_{\tau}\sup_{\left\{b\colon \|b-\beta(\tau)\|=\eta_1\right\}}\left|\widehat{\mathcal{R}}(b;\tau)-\mathcal{R}(b;\tau)\right|\geq C\eta_2^2\right)\leq C\exp\left(-n\,C\eta_2^4\right),$$

and hence

$$\Pr\left(\sup_{(\tau,h)}\left\|\widehat{\beta}_h(\tau)-\beta_h(\tau)\right\|\geq 2\eta\right)\leq C\exp\left(-n\,C\eta_2^4\right),$$

completing the proof. ■

Proof of Lemma 4 We start with the first deviation probability. As $R_h^{(1)}\left(\beta_h(\tau),\tau\right)=0$,

$$\sup_{(\tau,h)}\left\|\sqrt{n}\,\widehat{R}_h^{(1)}\left[\beta_h(\tau),\tau\right]\right\|\leq\sup_{(\tau,h)}\sup_{\left\{b\colon \|b-\beta_h(\tau)\|\leq\eta\right\}}\left\|\sqrt{n}\left(\widehat{R}_h^{(1)}(b,\tau)-R_h^{(1)}(b,\tau)\right)\right\|.$$

However,

$$\widehat{R}_h^{(1)}(b,\tau)=\frac{\partial}{\partial b}\left[\frac{1}{n}\sum_{i=1}^n\int\rho_\tau(Y_i-X_i'b+hz)\,k(z)\,\mathrm{d}z\right]=\frac{1}{n}\sum_{i=1}^n X_i\left[\int\mathbb{I}(Y_i-X_i'b+hz<0)\,k(z)\,\mathrm{d}z-\tau\right],$$

implying that $\widehat{R}_h^{(1)}(b,\tau)=\sum_{i=1}^n g_2(Z_i,\theta)/n$, with

$$g_2(Z_i,\theta)=X_i\left[\int\mathbb{I}(Y_i-X_i'b+hz<0)\,k(z)\,\mathrm{d}z-\tau\right],$$

for $Z_i=(Y_i,X_i')'$ and $\theta\in\Theta:=\left\{(b',h,\tau)\colon(\tau,h)\in[\underline{\tau},\bar{\tau}]\times[\underline{h}_n,\bar{h}_n],\ \|b-\beta_h(\tau)\|\leq\eta\right\}$. We bound each of the entries of $\widehat{R}_h^{(1)}(b,\tau)$, so that there is no loss of generality in assuming that X_i is univariate. Note that $|g_2(Z_i,\theta)|\leq C$, $\mathbb{V}\big(g_2(Z_i,\theta)\big)\leq\sigma^2\leq C$, and $|g_2(Z_i,\theta_2)-g_2(Z_i,\theta_1)|\leq C$ for all θ_1 and θ_2. Let $\|\theta\|^2=\|b\|^2+|h|^2+|\tau|^2$ and let $\mathcal{B}(\theta,\delta^2)$ denote the $\|\cdot\|$-ball with center θ and radius δ^2. Assumption X ensures that, for any θ_1 and θ_2 in $\mathcal{B}\left(\theta,\delta^2\right)$,

$$|g_2(Z_i,\theta_2)-g_2(Z_i,\theta_1)|\leq C\left[\int\mathbb{I}\left(Y_i-X_i'b+hz\in[-C\delta^2,C\delta^2]\right)|k(z)|\,\mathrm{d}z+\delta^2\right]. \tag{27}$$

Consider a covering of Θ with $J(\delta^2)\leq C\,\delta^{-2(d+1)}$ balls $\mathcal{B}(\theta_j,\delta^2)$. Letting $\underline{g}_{2j}(z):=\inf_{\theta\in\mathcal{B}(\theta_j,\delta)}g_2(z,\theta)$ and $\bar{g}_{2j}(z)=\sup_{\theta\in\mathcal{B}(\theta_j,\delta)}g_2(z,\theta)$ implies not only that $\left\{g_2(\cdot,\theta)\colon\theta\in\mathcal{B}(\theta_j,\delta)\right\}\subset[\underline{g}_{2j},\bar{g}_{2j}]$, but also that $\mathcal{G}_{2,\Theta}:=\left\{g_2(\cdot,\theta)\colon\theta\in\Theta\right\}\subset\bigcup_{j=1}^{J(\delta^2)}[\underline{g}_{2j},\bar{g}_{2j}]$. Equation (27) ensures that, uniformly in j and $\delta^2\leq\sigma^2$,

$$\mathbb{E}\left[\left|\bar{g}_{2j}(Z_i)-\underline{g}_{2j}(Z_i)\right|^2\right]\leq C\,\delta^4+C\,\mathbb{E}\left[\int\mathbb{I}\left(Y_i-X_i'b+hz\in[-C\,\delta^2,C\,\delta^2]\right)|k(z)|\,\mathrm{d}z\right]^2.$$

Applying the Cauchy-Schwarz inequality under Assumptions K and Q2 then gives way to

$$\begin{aligned}
L_4&:=\mathbb{E}\left[\int\mathbb{I}\left(Y_i-X_i'b-hz\in[-C\,\delta^2,C\,\delta^2]\right)k(z)\,\mathrm{d}z\right]^2\\
&\leq\mathbb{E}\left[\int\mathbb{I}\left(Y_i-X_i'b-hz\in[-C\,\delta^2,C\,\delta^2]\right)|k(z)|\,\mathrm{d}z\right]\times\int|k(z)|\,\mathrm{d}z\\
&\leq\int\mathbb{E}\left\{\Pr\left(Y_i-X_ib-hz\in[-C\,\delta^2,C\,\delta^2]\,|\,X_i\right)\right\}|k(z)|\,\mathrm{d}z\times\int|k(z)|\,\mathrm{d}z\\
&\leq C\,\delta^2,
\end{aligned}$$

implying that $\mathbb{E}\left[\left|\bar{g}_{2j}(Z_i) - \underline{g}_{2j}(Z_i)\right|^2\right] \leq C(\delta^4 + \delta^2) \leq C\,\delta^2$, uniformly in j and $\delta^2 \leq \sigma^2$. As a result, conditions (i) and (ii) in Lemma 2 hold for $\ln H(\delta) = -2(d+1)\ln\delta + C$, so that (18) gives

$$\Pr\left(\sup_{\theta\in\Theta}\left\|\sqrt{n}\left(\widehat{R}_h^{(1)}(b,\tau) - R_h^{(1)}(b,\tau)\right)\right\| \geq C\,(\sqrt{r} + 1 + r/\sqrt{n})\right) \leq 2\exp(-r).$$

Accordingly, the first bound holds for n large enough. As for the second bound, there is no loss of generality to assume that X_i is unidimensional. Note that $\sqrt{nh/\ln n}\,\widehat{R}_h^{(2)}(b,\tau) = \sum_{i=1}^n g_3(Z_i,\theta)/\sqrt{n}$, with

$$g_3(Z_i,\theta) := \sqrt{\frac{1}{h\ln n}}\,X_i^2\,k\left(\frac{X_i'b - Y_i}{h}\right).$$

Assumptions K and X ensure that, uniformly for $\theta\in\Theta$,

$$|g_3(Z_i,\theta)| \leq C\sqrt{\frac{1}{h\ln n}} \leq C\,\frac{O(\sqrt{n})}{\ln^2 n}.$$

It also follows from Assumption Q2 that, uniformly for $\theta\in\Theta$,

$$\begin{aligned}\mathbb{V}\big(g_3(Z_i,\theta)\big) &\leq \frac{C}{h\ln n}\int\int k\left(\frac{x'b - y}{h}\right) f(y\,|\,x)\,\mathrm{d}y\,\mathrm{d}F_X(x)\\ &= \frac{C}{\ln n}\times\int\int k(v)\,f(x'b + hv\,|\,x)\,\mathrm{d}v\,\mathrm{d}F_X(x) \leq \frac{C}{\ln n} = \sigma_n^2.\end{aligned}$$

Assumption K posits that, for any θ_1 and θ_2 in Θ, $|g_3(Z_i,\theta_1) - g_3(Z_i,\theta_2)| \leq C\,n^C\,\|\theta_1 - \theta_2\|$. Consider a covering of Θ with $J(\delta/n^C) \leq C\,(\delta/n^C)^{-(d+1)}$ balls $\mathcal{B}(\theta_j,\delta/n^C)$ and let $\underline{g}_{3j}(z) := \inf_{\theta\in\mathcal{B}(\theta_j,\delta)} g_3(z,\theta)$ and $\bar{g}_{3j}(z) := \sup_{\theta\in\mathcal{B}(\theta_j,\delta)} g_3(z,\theta)$. It then turns out that $\big\{g_3(z,\theta) : \theta\in\mathcal{B}(\theta_j,\delta)\big\} \subset [\underline{g}_{3j},\bar{g}_{3j}]$ and hence $\mathcal{G}_{3,\Theta} = \{g_3(\cdot,\theta) : \theta\in\Theta\} \subset \bigcup_{j=1}^{J(\delta/n^C)}[\underline{g}_{3j},\bar{g}_{3j}]$, with $\mathbb{E}\left[\left|\bar{g}_3(Z_i) - \underline{g}_3(Z_i)\right|^2\right] \leq C\,\delta^2$. Conditions (i) and (ii) in Lemma 2 thus hold for $\ln H(\delta) = -2(d+1)\,(\ln\delta - C\ln n) + C$, so that (18) results for any $u > 0$ in

$$\Pr\left(\sup_{\theta\in\Theta}\left\|\sqrt{\frac{nh}{\ln n}}\left(\widehat{R}_h^{(2)}(b,\tau) - R_h^{(2)}(b,\tau)\right)\right\| \geq C\left(1 + \frac{\sqrt{u}}{\sqrt{\ln n}} + \frac{u}{\ln n}\right)\right) \leq 2\exp(-u).$$

Setting $u = r\ln n$ then yields the exponential inequality.

Suppose now, without loss of generality, that $\mathcal{B}$ is convex. Recall that

$$\widehat{R}_h^{(2)}(b_1,\tau) - \widehat{R}_h^{(2)}(b_0,\tau) = \frac{1}{n}\sum_{i=1}^n X_iX_i'X_i(b_1 - b_0)\int_0^1 \frac{1}{h^2}\,k^{(1)}\left(\frac{Y_iX_i'\big[b_1 + t(b_1 - b_0)\big]}{h}\right)\mathrm{d}t$$

and that the variance of $h^{-2}\,k^{(1)}\big((Y_i - X_i'b)/h\big)$ is of order $h^{-3} = o(n/\ln n)$ under Assumption K. Applying now the same arguments as in the proof of the exponential inequality yields

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^n X_iX_i'X_i(b_1 - b_0)\frac{1}{h^2}\,k^{(1)}\left(\frac{Y_i - X_i'\big[b_1 + t(b_1 - b_0)\big]}{h}\right)\\ &= \mathbb{E}\left[XX'X(b_1 - b_0)\int_{-\infty}^{\infty}\frac{1}{h}\,k\left(\frac{y - X\big[b_1 + t(b_1 - b_0)\big]}{h}\right) f^{(1)}(y|X)\mathrm{d}y\right] + O_p\left(\sqrt{\frac{\ln n}{nh^3}}\right),\end{aligned}$$

uniformly in (τ,h,b_0,b_1) for $t\in[0,1]$. The proofs of the remaining results follow similarly. ■

Proof of Proposition 3 Let

$$\mathcal{E}_n^3(\epsilon) := \left\{ \sup_{(\tau,h)} \left\| \widehat{\beta}_h(\tau) - \beta_h(\tau) \right\| \geq \epsilon^{1/4} \right\},$$

which is such that $\Pr\left(\mathcal{E}_n^3(\epsilon)\right) \leq C \exp(-C\, n\, \epsilon)$ by Lemma 3. The bounds for $\Pr\left(\mathcal{E}_n^1(r)\right)$ and $\Pr\left(\mathcal{E}_n^2(r)\right)$ follow from Lemma 4. In particular, $\lim_{n\to\infty} \Pr\left(\mathcal{E}_n^2(r)\right) = 0$, whereas Lemma 1 ensures under Assumption X that $b \mapsto \widehat{R}_h(b;\tau)$ is strictly convex for b in a vicinity of $\beta_h(\tau)$, for all τ in $[\underline{\tau}, \overline{\tau}]$ with probability at least $1 - \Pr\left(\mathcal{E}_n^1(r)\right) - \Pr\left(\mathcal{E}_n^2(r)\right)$. But, by Lemma 3 and Theorem 1, all minimizers of $\widehat{R}_h(b;\tau)$ lie in such a vicinity with a probability tending to 1. This means that we can make $1 - \Pr\left(\mathcal{E}_n^1(r)\right) - \Pr\left(\mathcal{E}_n^2(r)\right)$ arbitrarily close to 1 by increasing r, and hence $\widehat{\beta}_h(\tau)$ is unique with a probability going to 1 as n increases. It also follows that, in case $\breve{\mathcal{E}}_n^1(r)$, $\breve{\mathcal{E}}_n^2(r)$ and $\breve{\mathcal{E}}_n^3(\epsilon)$ are all true and n is large enough, $\widehat{\beta}_h(\tau)$ satisfies the first-order condition $\widehat{R}_h^{(1)}\big(\widehat{\beta}_h(\tau);\tau\big) = 0$. Recall from the proof of Theorem 1 that $\widehat{R}_h^{(2)}(\cdot;\tau)$ has an inverse in the vicinity of $\beta_h(\tau)$ for n large enough on $\mathcal{E}^2(r)$. Applying the implicit function theorem then yields $\widehat{\beta}_h(\tau)$ continuous over the admissible (τ, h). Accordingly,

$$\begin{aligned}
-\widehat{R}_h^{(1)}\big(\beta_h(\tau);\tau\big) &= \widehat{R}_h^{(1)}\big(\widehat{\beta}_h(\tau);\tau\big) - \widehat{R}_h^{(1)}\big(\beta_h(\tau);\tau\big) \\
&= \big[\widehat{\beta}_h(\tau) - \beta_h(\tau)\big] \int_0^1 \widehat{R}_h^{(2)}\Big(\beta_h(\tau) + t\big[\widehat{\beta}_h(\tau) - \beta_h(\tau)\big];\tau\Big)\,\mathrm{d}t.
\end{aligned}$$

Now, if ϵ in $\mathcal{E}_n^3(\epsilon)$ is small enough, the eigenvalues of the above matrix are in $[1/C, C]$ for a large C provided that n is large enough, uniformly in τ and h. This means that

$$\widehat{\beta}_h(\tau) - \beta_h(\tau) = -\left[\int_0^1 \widehat{R}_h^{(2)}\Big(\beta_h(\tau) + u\big[\widehat{\beta}_h(\tau) - \beta_h(\tau)\big];\tau\Big)\,\mathrm{d}u\right]^{-1} \widehat{R}_h^{(1)}\big(\beta_h(\tau);\tau\big). \tag{28}$$

Lemma 1(iv) then implies that, for a generic constant C coming from Bernstein-type inequalities,

$$\begin{aligned}
P_2 &:= \left\| \sqrt{n}\big(\widehat{\beta}_h(\tau) - \beta_h(\tau)\big) + \big[R_h^{(2)}(\beta_h(\tau);\tau)\big]^{-1} \sqrt{n}\, \widehat{R}_h^{(1)}(\beta_h(\tau);\tau) \right\| \\
&\leq C \left\| \int_0^1 \Big[\widehat{R}_h^{(2)}\Big(\beta_h(\tau) + u\big[\widehat{\beta}_h(\tau) - \beta_h(\tau)\big];\tau\Big) - R_h^{(2)}\Big(\beta_h(\tau) + u\big[\widehat{\beta}_h(\tau) - \beta_h(\tau)\big];\tau\Big)\Big]\,\mathrm{d}u \right\| \left\| \sqrt{n}\, \widehat{R}_h^{(1)}(\beta_h(\tau);\tau) \right\| \\
&\quad + C \left\| \int_0^1 \Big[R_h^{(2)}\Big(\beta_h(\tau) + u\big[\widehat{\beta}_h(\tau) - \beta_h(\tau)\big];\tau\Big) - R_h^{(2)}\big(\beta_h(\tau);\tau\big)\Big]\,\mathrm{d}u \right\| \left\| \sqrt{n}\, \widehat{R}_h^{(1)}\big(\beta_h(\tau);\tau\big) \right\| \\
&\leq C \left\{ \sqrt{\frac{\ln n}{nh}}\, r^2 + \left\| \widehat{\beta}_h(\tau) - \beta_h(\tau) \right\| \left\| \sqrt{n}\, \widehat{R}_h^{(1)}\big(\beta_h(\tau);\tau\big) \right\| \right\} \\
&\leq C \left\{ \sqrt{\frac{\ln n}{nh}}\, r^2 + n^{-1/2} \left\| \sqrt{n}\, \widehat{R}_h^{(1)}\big(\beta_h(\tau);\tau\big) \right\|^2 \right\} \\
&\leq C \left(\sqrt{\frac{\ln n}{nh}} + \frac{1}{\sqrt{n}} \right) r^2
\end{aligned}$$

on $\breve{\mathcal{E}}_n^1(r)$ and $\breve{\mathcal{E}}_n^2(r)$, implying that $\breve{\mathcal{E}}_n(r)$ holds as long as C_0 of the Proposition is large enough. ■

Proof of Lemma 5 Let $h = h_n$ to simplify notation. We first note that $\mathbb{E}\big(\sqrt{n}\, \widehat{S}_h(\tau)\big) = 0$. In addition, for any $\alpha, \tau \in [\underline{\tau}, \overline{\tau}]$, it follows that

$$\mathbb{V}\big(\sqrt{n}\, \widehat{S}_h(\tau), \sqrt{n}\, \widehat{S}_h(\varsigma)\big) = \mathbb{E}\left\{ XX' \left[K\left(-\frac{e\big(\beta_h(\tau)\big)}{h} \right) - \tau \right] \left[K\left(-\frac{e\big(\beta_h(\varsigma)\big)}{h} \right) - \varsigma \right] \right\}$$

converges to $\mathbb{E}\big\{XX'\big(\mathbb{I}[X'\beta(\tau) \geq Y] - \tau\big)\big(\mathbb{I}[X'\beta(\varsigma) \geq Y] - \varsigma\big)\big\}$ as $n \to \infty$. A simple computation using iterated expectations then yields the limiting covariance structure in (22).

By the Cramér-Wold device, in order to obtain weak convergence for the d-dimensional process $\big\{\sqrt{n}\,\widehat{S}_h : \tau \in [\underline{\tau}, \bar{\tau}]\big\}$, it suffices to consider the convergence in distribution of the linear form $\big\{\sqrt{n}\,\lambda'\widehat{S}_h : \tau \in [\underline{\tau}, \bar{\tau}]\big\}$, where λ is an arbitrary (fixed) vector in $\mathbb{R}^d$. Assume without loss of generality that $\|X\| \leq 1$ and $\|\lambda\| \leq 1$, and let $Z = (Y, X) \in \mathbb{R} \times \mathbb{R}^d$ and, similarly, $Z_i = (Y_i, X_i)$. Define now $g_{n,\tau} : \mathbb{R} \times \mathrm{supp}X \to \mathbb{R}$ for $z = (y, x)$ as

$$g_{n,\tau}(z) := x_\lambda \left\{ K\left(\frac{x'\beta_h(\tau) - y}{h}\right) - \tau \right\}, \tag{29}$$

where $x_\lambda = \lambda' x$ and $X_\lambda = \lambda' X$, and consider the class of functions $\mathcal{G}_n = \{g_{n,\tau} : \tau \in [\underline{\tau}, \bar{\tau}]\}$. Letting $\mathbb{P}$ and $\mathbb{P}_n$ respectively denote the distribution of Z and the empirical distribution of the sample $(Z_1, \ldots, Z_n)$ yields

$$\sqrt{n}\,\lambda'\widehat{S}_h(\tau) = \sqrt{n}\big(\mathbb{P}_n g_{n,\tau} - \mathbb{P} g_{n,\tau}\big).$$

In other words, the process $\big\{\sqrt{n}\lambda'\widehat{S}_h : \tau \in [\underline{\tau}, \bar{\tau}]\big\}$ is an empirical process indexed by a (changing) class of functions $\mathcal{G}_n$. By Theorem 19.28 in van der Vaart (1998), it suffices to establish that

$$\sup_{|\tau - \varsigma| < \delta(n)} \mathbb{E}\big|g_{n,\tau}(Z) - g_{n,\varsigma}(Z)\big|^2 \to 0 \tag{30}$$

and that, for any $\delta(n) \downarrow 0$,

$$\int_0^{\delta(n)} \sqrt{\ln N_{[]}(\epsilon, \mathcal{G}_n, L^2(\mathbb{P}))}\, \mathrm{d}\epsilon \to 0 \tag{31}$$

with $N_{[]}(\epsilon, \mathcal{G}_n, L^2(\mathbb{P}))$ denoting the minimum number of ϵ-brackets in $L^2(\mathbb{P})$ required to cover $\mathcal{G}_n$. The remaining requirements of Theorem 19.28 indeed hold trivially in view that the index set $[\underline{\tau}, \bar{\tau}]$ is a compact—and so, totally bounded—metric space, and that the changing classes $\mathcal{G}_n$ admit envelope functions $G_n \equiv 1$ for all n that satisfy the Lindeberg condition $\mathbb{E}_{\mathbb{P}}\big(G_n^2\, \mathbb{I}[G_n > \sqrt{n}\epsilon]\big) \to 0$.

Let $\partial_\tau := \frac{\partial}{\partial \tau}$. By Lemma 1 and Theorem 1, applying twice the implicit function theorem yields

$$\partial_\tau \beta_h(\tau) = -D_h(\tau)^{-1}\, \partial_\tau R_h^{(1)}\big(\beta_h(\tau); \tau\big) = D_h(\tau)^{-1}\mathbb{E}(X) = [D(\tau) + o(1)]^{-1}\, \mathbb{E}(X) = \partial_\tau \beta(\tau) + o(1)$$

uniformly for $(\tau, h) \in [\underline{\tau}, \bar{\tau}] \times [\underline{h}, \bar{h}]$. This implies, by Assumption Q1, that $\sup \|\partial_\tau \beta_h(\tau)\| \leq C$ for n large enough, with supremum taken over $(\tau, h) \in [\underline{\tau}, \bar{\tau}] \times [\underline{h}_n, \bar{h}_n]$, and so $\|\beta_h(\tau) - \beta_h(\varsigma)\| \leq C|\tau - \varsigma|$. It also follows from the inverse function theorem and Assumption Q1 that $\tau \mapsto x'\beta_h(\tau)$ is strictly increasing in τ, for any $x \in \mathrm{supp}X$ and n large enough. In what follows, we assume that n is large enough, so that the above holds.

Now, let $\underline{\tau} \leq \tau_{\mathrm{L}} \leq \tau_{\mathrm{U}} \leq \bar{\tau}$ and consider two random elements (possibly degenerate) $\widehat{\tau}$ and $\widehat{\varsigma}$ in $[\tau_{\mathrm{L}}, \tau_{\mathrm{U}}]$. The mean value theorem and Assumption Q2 then ensure that

$$\Pr\Big(x'\beta_h(\widehat{\tau} \wedge \widehat{\varsigma}) - hu \leq Y \leq x'\beta_h(\widehat{\tau} \vee \widehat{\varsigma}) - hu \,|\, X = x\Big) \leq C\, |\tau_{\mathrm{U}} - \tau_{\mathrm{L}}|\,, \tag{32}$$

uniformly for $u \in \mathbb{R}$ and $x \in \mathrm{supp}X$, given that $\big[x'\beta_h(\widehat{\tau} \wedge \widehat{\varsigma}),\, x'\beta_h(\widehat{\tau} \vee \widehat{\varsigma})\big] \subset \big[x'\beta_h(\tau_{\mathrm{L}}),\, x'\beta_h(\tau_{\mathrm{U}})\big]$ and $|x'\beta_h(\tau_{\mathrm{U}}) - x'\beta_h(\tau_{\mathrm{L}})| \leq C\, |\tau_{\mathrm{U}} - \tau_{\mathrm{L}}|$. Define $\Upsilon_u = \{X'\beta_h(\widehat{\tau} \wedge \widehat{\varsigma}) - Y \leq hu \leq X'\beta_h(\widehat{\tau} \vee \widehat{\varsigma}) - Y\}$.

It follows from $\left|g_{n,\widehat{\tau}}(Z) - g_{n,\widehat{\varsigma}}(Z)\right| \leq \int \mathbb{I}(\Upsilon_u)\,|k(u)|\,\mathrm{d}u + |\widehat{\tau} - \widehat{\varsigma}|$ that

$$\begin{aligned}
\mathbb{E}\left|g_{n,\widehat{\tau}}(Z) - g_{n,\widehat{\varsigma}}(Z)\right|^2 &\leq \mathbb{E}|\widehat{\tau} - \widehat{\varsigma}|^2 + 2\mathbb{E}\Big[\,|\widehat{\tau} - \widehat{\varsigma}| \int \mathbb{I}(\Upsilon_u)\,|k(u)|\,\mathrm{d}u\Big] + \mathbb{E}\Big[\int \mathbb{I}(\Upsilon_u)\,|k(u)|\,\mathrm{d}u\Big]^2 \\
&\leq |\tau_{\mathrm{U}} - \tau_{\mathrm{L}}|^2 + 2\,|\tau_{\mathrm{U}} - \tau_{\mathrm{L}}| \int \Pr(\Upsilon_u)\,|k(u)|\,\mathrm{d}u + C \int \Pr(\Upsilon_u)\,|k(u)|\,\mathrm{d}u \qquad (33) \\
&\leq C\,|\tau_{\mathrm{U}} - \tau_{\mathrm{L}}|\,, \qquad (34)
\end{aligned}$$

given that the Cauchy-Schwarz inequality implies that $\int \mathbb{I}(\Upsilon_u)\,|k(u)|\,\mathrm{d}u \int |k(u)|\,\mathrm{d}u$ is an upper bound for $\left(\int \mathbb{I}(\Upsilon_u)\,|k(u)|^{1/2}\,|k(u)|^{1/2}\,\mathrm{d}u\right)^2$, $\int |k(u)|\,\mathrm{d}u < \infty$ by Assumption K1, and $\Pr(\Upsilon_u) = \mathbb{E}\big[\Pr(\Upsilon_u \,|\, X)\big] \leq C\,|\tau_{\mathrm{U}} - \tau_{\mathrm{L}}|$ by iterated expectations and (32). Taking $\widehat{\tau}$ and $\widehat{\varsigma}$ to be deterministic shows that (30) holds, for all $\delta(n) \downarrow 0$.

We now obtain a set of brackets whose bracketing number is of order $1/\epsilon$. For $\epsilon > 0$ small enough, we cover the interval $[\underline{\tau}, \bar{\tau}]$ with $J(\epsilon) \leq \lceil(\bar{\tau} - \underline{\tau})/\epsilon + 1\rceil \leq 2/\epsilon$ open intervals $B_i = (\tau_i - \epsilon, \tau_i + \epsilon)$, and let $\bar{g}_n^i(z) = \sup_{\tau \in B_i} g_{n,\tau}(z)$ and $\underline{g}_n^i(z) = \inf_{\tau \in B_i} g_{n,\tau}(z)$. It is straightforward to appreciate that the collection formed by the brackets $[\underline{g}_n^i, \bar{g}_n^i]$, with $i = 1, \ldots, J(\epsilon)$, covers $\mathcal{G}_n$ and that these suprema and infima are attained in the closure of B_i.[6] In particular, $\bar{g}_n^i(Z) = g_{n,\widehat{\tau}_i}(Z)$ and $\underline{g}_n^i(Z) = g_{n,\widehat{\varsigma}_i}(Z)$, where $\widehat{\tau}_i$ and $\widehat{\varsigma}_i$ are random elements in $[\tau_i - \epsilon, \tau_i + \epsilon]$. Resorting to (34) once more then gives

$$\mathbb{E}|\bar{g}_n^i(Z) - \underline{g}_n^i(Z)|^2 \leq C\epsilon,$$

and, as a result, $N_{[\,]}(\epsilon, \mathcal{G}_n, L^2(P)) \leq C/\epsilon$. This ensures that (31) holds for all $\delta(n) \downarrow 0$. ■

[6] For $i = 1$ and $i = J(\epsilon)$, the intervals are actually $[\underline{\tau}, \tau_1 + \epsilon]$ and $[\tau_{J(\epsilon)} - \epsilon, \bar{\tau}]$, respectively. For simplicity of exposition, we keep the notation as above.