The independence process in conditional quantile location-scale models and an application to testing for monotonicity

Melanie Birke, Natalie Neumeyer and Stanislav Volgushev
University of Bayreuth, University of Hamburg and University of Toronto

Supplementary material.

The supplementary material contains all proofs. The proofs of the main results (Theorem 1, Corollary 1 and Theorem 3) and the bootstrap versions (Theorems 2 and 4) are contained in Sections S1 and S2, respectively. Technical details needed in the proofs of those results can be found in Section S3.3. Finally, Section S3.1 contains basic results on linearized versions and differentiability of the quantile estimator $\hat{q}_\tau$, scale estimator $\hat{s}$ and the corresponding bootstrap versions, while Section S3.2 contains additional technical details.

S1 Proof of weak convergence results

Proof of Theorem 1. For the numerator $\bar{F}_{X,\varepsilon,n}(t,y) = \hat{F}_{X,\varepsilon,n}(t,y)(\hat{F}_{X,n}(1-2h_n) - \hat{F}_{X,n}(2h_n))$ of the joint empirical distribution function defined in (3.2) we have

$$\bar{F}_{X,\varepsilon,n}(t,y) = \frac{1}{n}\sum_{i=1}^{n} I\Big\{\varepsilon_i \leq y\frac{\hat{s}(X_i)}{s(X_i)} + \frac{\hat{q}_\tau(X_i) - q_\tau(X_i)}{s(X_i)}\Big\} I\{2h_n < X_i \leq t\}.$$

Note that in Lemma 9 it is shown that without changing the asymptotic distribution of the process the residuals $\hat{\varepsilon}_i$ can be replaced by their versions obtained from linearized estimators $\hat{q}_{\tau,L}$, $\hat{s}_L$ instead of $\hat{q}_\tau$, $\hat{s}$ (see Section S3.1 for the definitions). Thus we have

$$\bar{F}_{X,\varepsilon,n}(t,y) = \frac{1}{n}\sum_{i=1}^{n} I\Big\{\varepsilon_i \leq y\frac{\hat{s}_L(X_i)}{s(X_i)} + \frac{\hat{q}_{\tau,L}(X_i) - q_\tau(X_i)}{s(X_i)}\Big\} I\{2h_n < X_i \leq t\} + o_P(\frac{1}{\sqrt{n}}).$$

From this we obtain the expansion

$$\bar{F}_{X,\varepsilon,n}(t,y) = \frac{1}{n}\sum_{i=1}^{n} I\{\varepsilon_i \leq y\} I\{2h_n < X_i \leq t\}$$

$$+\int_{2h_n}^{1-2h_n}\Big(F_\varepsilon\Big(y\frac{\hat{s}_L(x)}{s(x)}+\frac{\hat{q}_{\tau,L}(x)-q_\tau(x)}{s(x)}\Big)-F_\varepsilon(y)\Big)I\{x\le t\}f_X(x)\,dx \qquad \text{(S1.1)}$$
$$+\,o_P(\frac{1}{\sqrt{n}})$$

uniformly with respect to $t \in [2h_n, 1-2h_n]$ and $y \in \mathbb{R}$ by the following argumentation. Consider the empirical process

$$G_n(\varphi) \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\varphi(X_i,\varepsilon_i)-E[\varphi(X_i,\varepsilon_i)]\Big), \quad \varphi\in\mathcal{F},$$

indexed by the following class of functions,

$$\mathcal{F} \;=\; \Big\{(X,\varepsilon)\mapsto I\{\varepsilon\le yd_2(X)+d_1(X)\}I\{h<X\}I\{X\le t\}-I\{\varepsilon\le y\}I\{h<X\}I\{X\le t\}$$
$$\Big|\; y\in\mathbb{R}, h,t\in[0,1], d_1\in C_1^{1+\delta}([0,1]), d_2\in\tilde{C}_2^{1+\delta}([0,1])\Big\},$$

for some arbitrary $\delta \in (0,1)$, where the function class $C_c^{1+\delta}([0,1])$ is defined as the set of differentiable functions $g : [0,1] \to \mathbb{R}$ with derivatives g' such that

$$\max\Big\{\sup_{x\in[0,1]}|g(x)|,\sup_{x\in[0,1]}|g'(x)|\Big\}+\sup_{x,z\in[0,1]}\frac{|g'(x)-g'(z)|}{|x-z|^\delta}\;\le\; c$$

[see van der Vaart and Wellner (1996, p. 154)]. We further by slight abuse of notation define the subset $\tilde{C}_2^{1+\delta}([0,1])$ of $C_1^{1+\delta}([0,1])$ by the additional constraint $\inf_{x\in[0,1]} g(x) \ge 1/2$. Now $\mathcal{F}$ is a product of the uniformly bounded Donsker classes $\{(X,\varepsilon)\mapsto I\{h<X\}I\{X\le t\}|h,t\in[0,1]\}$ and $\{(X,\varepsilon)\mapsto I\{\varepsilon\le yd_2(X)+d_1(X)\}-I\{\varepsilon\le y\}|y\in\mathbb{R}, d_1\in C_1^{1+\delta}([0,1]), d_2\in\tilde{C}_2^{1+\delta}([0,1])\}$ [the Donsker property for the second class is shown in Lemma 1 by Akritas and Van Keilegom (2001)] and is therefore Donsker as well (Ex. 2.10.8, van der Vaart and Wellner (1996), p. 192). The remaining part of the proof for equality (S1.1) follows exactly the lines of the end of the proof of Lemma 1, Akritas and Van Keilegom (2001), p. 567, using the inequality

$$\mathrm{Var}\Big(I\{\varepsilon_1\le yd_2(X_1)+d_1(X_1)\}I\{h<X_1\}I\{X_1\le s\}-I\{\varepsilon_1\le y\}I\{h<X_1\}I\{X_1\le s\}\Big)$$
$$\le\; E\Big[\Big(I\{\varepsilon_1\le yd_2(X_1)+d_1(X_1)\}-I\{\varepsilon_1\le y\}\Big)^2\Big].$$

Here one also needs $\hat{s}_L/s \in \tilde{C}_2^{1+\delta}([0,1])$, $(\hat{q}_{\tau,L}-q_\tau)/s \in C_1^{1+\delta}([0,1])$ with probability converging to one, which follows from uniform consistency results in Lemma 1. For $\varphi = \varphi_{h,t,y,d_1,d_2}$ we obtain

$$\sup_{\substack{y\in\mathbb{R},\\ t\in[2h_n,1-2h_n]}}\Big|G_n\Big(\varphi_{2h_n,t,y,\frac{\hat{q}_{\tau,L}-q_\tau}{s},\frac{\hat{s}_L}{s}}\Big)\Big| = o_P(1)$$

and thus (S1.1).

Further, by a Taylor expansion we obtain from (S1.1) together with assumption (**A4**) that

$$\bar F_{X,\varepsilon,n}(t,y) \;=\; \frac{1}{n}\sum_{i=1}^n I\{\varepsilon_i\le y\}I\{2h_n<X_i\le t\} + yf_\varepsilon(y)\int_{2h_n}^{1-2h_n}\frac{\hat s_L(x)-s(x)}{s(x)}I\{x\le t\}f_X(x)\,dx$$
$$+\, f_\varepsilon(y)\int_{2h_n}^{1-2h_n}\frac{\hat q_{\tau,L}(x)-q_\tau(x)}{s(x)}I\{x\le t\}f_X(x)\,dx + o_P(\frac{1}{\sqrt n})$$

uniformly with respect to $t\in[2h_n,1-2h_n]$ and $y\in\mathbb{R}$. In Lemma 10 expansions of the integrals in this decomposition are derived and it follows that

$$\begin{aligned}&\bar F_{X,\varepsilon,n}(t,y) \qquad\qquad \text{(S1.2)}\\ =\;& \frac{1}{n}\sum_{i=1}^n I\{\varepsilon_i\le y\}I\{2h_n<X_i\le t\} - \phi(y)\frac{1}{n}\sum_{i=1}^n (I\{\varepsilon_i\le 0\}-\tau)I\{2h_n<X_i\le t\}\\ &-\psi(y)\frac{1}{n}\sum_{i=1}^n\Big(I\{|\varepsilon_i|\le 1\}-\frac12\Big)I\{2h_n<X_i\le t\} + o_P(\frac{1}{\sqrt n}),\end{aligned}$$

where ϕ and ψ are defined in the assertion of the theorem. Thus noting that $\hat F_{X,n}(1-2h_n)-\hat F_{X,n}(2h_n) = F_X(1-2h_n)-F_X(2h_n)+o_P(1) = 1+o_P(1)$, from the definition (3.3) we obtain by Slutsky's lemma that

$$\begin{aligned}S_n(t,y) \;=\;& \frac{1}{\sqrt n}\sum_{i=1}^n\Big(I\{\varepsilon_i\le y\}-F_\varepsilon(y)-\phi(y)(I\{\varepsilon_i\le 0\}-\tau)-\psi(y)\Big(I\{|\varepsilon_i|\le 1\}-\frac12\Big)\Big)\\ &\times\Big(I\{2h_n<X_i\le t\}-I\{2h_n<X_i\le 1-2h_n\}\frac{\hat F_{X,n}(t)-\hat F_{X,n}(2h_n)}{\hat F_{X,n}(1-2h_n)-\hat F_{X,n}(2h_n)}\Big)\\ &+o_P(1).\end{aligned}$$

uniformly with respect to $t\in[2h_n,1-2h_n]$ and $y\in\mathbb{R}$. Note that the dominating part of this process vanishes in the boundary points $t=2h_n$ and $t=1-2h_n$. Further, from $\hat F_{X,n}(t)=F_X(t)+O_P(n^{-1/2})$ uniformly in $t\in[0,1]$ and $F_X(2h_n)\to 0$, $F_X(1-2h_n)\to 1$ we have

$$S_n(t,y) \;=\; S_{n,1}(t,y)+o_P(1),$$

uniformly with respect to $t\in[0,1]$, $y\in\mathbb{R}$, where $S_{n,1}(t,y)=0$ for $t\in[0,2h_n)\cup(1-2h_n,1]$ and

$$S_{n,1}(t,y) \;=\; \frac{1}{\sqrt n}\sum_{i=1}^n g(\varepsilon_i,y)\Big(I\{2h_n<X_i\le t\}-I\{2h_n<X_i\le 1-2h_n\}F_X(t)\Big)$$

for $t\in[2h_n,1-2h_n]$ and $y\in\mathbb{R}$, where $g(\varepsilon_i,y)=I\{\varepsilon_i\le y\}-F_\varepsilon(y)-\phi(y)(I\{\varepsilon_i\le 0\}-\tau)-\psi(y)(I\{|\varepsilon_i|\le 1\}-\frac12)$ is centered and independent of X_i. The first assertion of the theorem now follows if we show that for

$$S_{n,2}(t,y) \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^{n} g(\varepsilon_i,y)\Big(I\{X_i\le t\}-F_X(t)\Big), \quad t\in[0,1], y\in\mathbb{R},$$

we have $\sup_{t\in[0,1],y\in\mathbb{R}}|S_{n,1}(t,y)-S_{n,2}(t,y)| = o_P(1)$, which is equivalent to

$$\sup_{t\in[2h_n,1-2h_n],y\in\mathbb{R}} |S_{n,1}(t,y)-S_{n,2}(t,y)| \;=\; o_P(1) \tag{S1.3}$$

together with

$$\sup_{t\in[0,2h_n)\cup(1-2h_n,1],y\in\mathbb{R}} |S_{n,2}(t,y)| \;=\; o_P(1). \tag{S1.4}$$

We will only show (S1.3); (S1.4) follows by similar arguments. Note that $S_{n,1}(t,y)-S_{n,2}(t,y)=G_n(h_n,t,y)$ for $t\in[2h_n,1-2h_n]$, $y\in\mathbb{R}$, where the process

$$G_n(h,t,y)=\frac{-1}{\sqrt{n}}\sum_{i=1}^{n} g(\varepsilon_i,y)(I\{X_i\le t\}-F_X(t))I\{X_i\in[0,2h)\cup(1-2h,1]\}$$

indexed in $h\in[0,\frac{1}{4}]$, $t\in[0,1]$, $y\in\mathbb{R}$, converges weakly to a centered Gaussian process G with asymptotic variance

$$\begin{aligned}\mathrm{Var}(G(h,t,y)) \;=\; & E[g^2(\varepsilon_1,y)]\Big((F_X(t\wedge 2h)+F_X(t)-F_X(t\wedge(1-2h)))(1-2F_X(t))\\ & + F_X^2(t)(F_X(2h)+1-F_X(1-2h))\Big).\end{aligned}$$

For $h=h_n\to 0$ this asymptotic variance vanishes uniformly with respect to y and t. From asymptotic equicontinuity of G_n (confer van der Vaart and Wellner, 1996, p. 89/90), using the asymptotic variance as semi-metric, with $G_n(0,t,y)\equiv 0$ it follows that $\sup_{t,y}|G_n(h_n,t,y)| = o_P(1)$ and thus (S1.3).
Hence, we have shown the first assertion of the theorem, i. e. $S_n = S_{n,2}+o_P(1)$ uniformly. Weak convergence of $S_{n,2}$ (and thus of S_n) to a centered Gaussian process with the asserted covariance structure follows by standard arguments. □

Proof of Corollary 1. The asymptotic distribution of K_n directly follows from Theorem 1 and the continuous mapping theorem. From those theorems also follows that

$$\tilde{C}_n = \int_{\mathbb{R}}\int_{[0,1]} S_n^2(t,y)\,F_X(dt)\,F_\varepsilon(dy)$$

converges in distribution to the desired limit. It therefore remains to show that $C_n-\tilde{C}_n = o_P(1)$. To this end denote

$$\tilde{C}_n^{(1)} = \int_{\mathbb{R}}\int_{[0,1]} S_n^2(F_X^{-1}(\hat{F}_{X,n}(t)),F_\varepsilon^{-1}(\hat{F}_{\varepsilon,n}(y)))\,\hat{F}_{X,n}(dt)\,\hat{F}_{\varepsilon,n}(dy)$$

and let ϱ_n be some sequence specified later with $\varrho_n \to \infty$ for $n \to \infty$. Then

$$|C_n - \tilde{C}_n^{(1)}| \leq \Big| \int_{[-\varrho_n,\varrho_n]} \int_{[0,1]} \Big(S_n^2(t,y) - S_n^2(F_X^{-1}(\hat{F}_{X,n}(t)), F_\varepsilon^{-1}(\hat{F}_{\varepsilon,n}(y))) \Big) \hat{F}_{X,n}(dt)\, \hat{F}_{\varepsilon,n}(dy) \Big| + 2 \sup_{t,y} |S_n^2(t,y)| \int_{\mathbb{R}\setminus[-\varrho_n,\varrho_n]} \hat{F}_{\varepsilon,n}(dy).$$

The second term on the right hand side is $O_P(1)(1 - \hat{F}_{\varepsilon,n}(\varrho_n) + \hat{F}_{\varepsilon,n}(-\varrho_n)) = o_P(1)$ due to the results from Theorem 1 and because $\varrho_n \to \infty$ and $\hat{F}_{\varepsilon,n}$ converges to F_ε uniformly in probability (this follows from the proof of Theorem 1). The first term on the right hand side can further be bounded by

$$2 \sup_{t,y} |S_n(t,y)| \sup_{\substack{t\in[0,1] \\ y\in[-\varrho_n,\varrho_n]}} \Big| S_n(t,y) - S_n(F_X^{-1}(\hat{F}_{X,n}(t)), F_\varepsilon^{-1}(\hat{F}_{\varepsilon,n}(y))) \Big|.$$

From Theorem 1 it follows that the process S_n is asymptotically stochastic equicontinuous such that we obtain the desired rate $o_P(1)$ from

$$\sup_{t\in[0,1]|} |t - F_X^{-1}(\hat{F}_{X,n}(t))| \leq \sup_{\xi\in[0,1]} \frac{1}{f_X(\xi)} \sup_{t\in[0,1]} |\hat{F}_{X,n}(t) - F_X(t)| = o_P(1)$$

by assumption (**A1**) and

$$\sup_{y\in[-\varrho_n,\varrho_n]} |y - F_\varepsilon^{-1}(\hat{F}_{\varepsilon,n}(y)))| \leq \sup_{y\in[-\varrho_n,\varrho_n]} \sup_{\substack{\zeta \text{ between} \\ F_\varepsilon(y) \text{ and } \hat{F}_{\varepsilon,n}(y)}} \frac{1}{f_\varepsilon(F_\varepsilon^{-1}(\zeta))} \sup_{y\in\mathbb{R}} |\hat{F}_{\varepsilon,n}(y) - F_\varepsilon(y)| = o_P(1).$$

The latter rate follows because $\sup_{y\in\mathbb{R}} |\hat{F}_{\varepsilon,n}(y) - F_\varepsilon(y)| = O_P(n^{-1/2})$ (which can be deduced by $\hat{F}_{\varepsilon,n}(\cdot) = \overline{F}_{X,\varepsilon,n}(1-2h_n,\cdot)/(\hat{F}_{X,n}(1-2h_n) - \hat{F}_{X,n}(2h_n))$ and (S1.2) in the proof of Theorem 1) if we choose a sequence ϱ_n such that $n^{1/2} \inf_{y\in[-2\varrho_n,2\varrho_n]} f_\varepsilon(y) \to \infty$ for $n \to \infty$. This is possible by assumption (**A4**).

We have shown $C_n - \tilde{C}_n^{(1)} = o_P(1)$ and it remains to show that $\tilde{C}_n - \tilde{C}_n^{(1)} = o_P(1)$. To this end, note that almost surely

$$\begin{aligned} \tilde{C}_n - \tilde{C}_n^{(1)} &= \int_{[0,1]} \int_{[0,1]} S_n^2(F_X^{-1}(s), F_\varepsilon^{-1}(z))\, ds\, dz - \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n S_n^2(F_X^{-1}(\tfrac{i}{n}), F_\varepsilon^{-1}(\tfrac{j}{n})) \\ &= \sum_{i=1}^n \sum_{j=1}^n \int_{[\frac{i-1}{n},\frac{i}{n})} \int_{[\frac{j-1}{n},\frac{j}{n})} \Big(S_n^2(F_X^{-1}(s), F_\varepsilon^{-1}(z)) - S_n^2(F_X^{-1}(\tfrac{i}{n}), F_\varepsilon^{-1}(\tfrac{j}{n})) \Big)\, ds\, dz. \end{aligned}$$

We decompose the second sum into $\sum_{j=1}^{j_n} \ldots + \sum_{j=J_n+1}^{n} \ldots + \sum_{j=j_n+1}^{J_n} \ldots$ for sequences of integers with $1 \leq j_n < J_n \leq n$ and $j_n/n \to 0$, $J_n/n \to 1$ for $n \to \infty$. We obtain

$$\begin{aligned} |\tilde{C}_n - \tilde{C}_n^{(1)}| &\leq 2 \frac{j_n + n - J_n}{n} \sup_{t,y} |S_n^2(t,y)| \\ &\quad + 2 \sup_{t,y} |S_n(t,y)| \sup_{\substack{|s-u|\leq\frac{1}{n} \\ s,u\in[0,1]}} \sup_{\substack{|z-v|\leq\frac{1}{n} \\ z,v\in\left[\frac{j_n}{n},\frac{J_n}{n}\right]}} |S_n(F_X^{-1}(s), F_\varepsilon^{-1}(z)) - S_n(F_X^{-1}(u), F_\varepsilon^{-1}(v))|. \end{aligned}$$

By asymptotic stochastic equicontinuity of S_n this converges to zero in probability if

$$\sup_{\substack{|s-u|\le\frac{1}{n}\\ s,u\in[0,1]}} |F_X^{-1}(s) - F_X^{-1}(u)| \to 0$$

which follows from assumption (**A1**) and the mean value theorem, and

$$\sup_{\substack{|z-v|\le\frac{1}{n}\\ z,v\in\left[\frac{j_n}{n},\frac{J_n}{n}\right]}} |F_\varepsilon^{-1}(z) - F_\varepsilon^{-1}(v)| \to 0$$

which can be guaranteed by assumption (**A4**) and the mean value theorem if j_n/n and J_n/n converge slowly enough. □

Proof of Theorem 3. The assertion follows from Theorem 1 if we show that uniformly with respect to $t \in [0,1]$ and $y \in \mathbb{R}$, $S_n(t,y) = S_{n,I}(t,y) + o_P(1)$. To this end, observe that as in the proof of Theorem 1 we can replace the estimators $\hat{q}_\tau$ and $\hat{s}$ by their linearized versions $\hat{q}_{\tau,L}$ and $\hat{s}_L$ in the definition of S_n without changing the asymptotic properties. Denote the corresponding version of the process by $S_{n,L}$. Similarly, in the definition of $S_{n,I}$ the estimators $\hat{q}_{\tau,I}$ and $\hat{s}$ can be replaced by $\hat{q}_{\tau,L,I} = \Gamma_n(\hat{q}_{\tau,L})$ and $\hat{s}_L$, where $\hat{q}_{\tau,L,I}$ denotes the increasing rearrangement of the linearized estimator $\hat{q}_{\tau,L}$. More precisely, denoting this version of the process by $S_{n,L,I}$, we will show that

$$\sup_{t\in[0,1],y\in\mathbb{R}} |S_{n,L,I}(t,y) - S_{n,I}(t,y)| = o_P(1). \tag{S1.5}$$

To see this, let $c = \inf_{x\in[0,1]} q'_\tau(x)$ and note that by our assumptions $c > 0$ and by Lemma 1 we have for the set $\Omega_n := \{\sup_{x\in[h_n,1-h_n]} |\hat{q}'_{\tau,L}(x) - q'_\tau(x)| \le \frac{c}{2}\}$ that $P(\Omega_n) \to 1$ for $n \to \infty$. Observe that by a straightforward modification of the proof of Theorem 3.1 (a) in Neumeyer (2007), we have on the set Ω_n

$$\sup_{x\in[h_n,1-h_n]} |\Gamma_n(\hat{q}_{\tau,L})(x) - \Gamma_n(\hat{q}_\tau)(x)| \le C \sup_{x\in[h_n,1-h_n]} |\hat{q}_{\tau,L}(x) - \hat{q}_\tau(x)|$$

for a universal constant C which is independent of n. Thus Lemma 2 together with $P(\Omega_n) \to 1$ implies that

$$\sup_{x\in[h_n,1-h_n]} |\Gamma_n(\hat{q}_{\tau,L})(x) - \Gamma_n(\hat{q}_\tau)(x)| = o_P(n^{-1/2}).$$

Additionally, observe that the estimator $\hat{q}_{\tau,L}$ is strictly increasing provided that the event Ω_n holds, which implies that $P(\hat{q}_{\tau,L} \equiv \Gamma_n(\hat{q}_{\tau,L})) \ge P(\Omega_n) \to 1$. Now similar arguments as those used in the proof of Lemma 9 show that, defining $F_{X,\varepsilon_{L,I},n}$ in the same manner as $\hat{F}_{X,\varepsilon_I,n}$ but with $\hat{\varepsilon}_{i,L,I} := (Y_i - \Gamma_n(\hat{q}_{\tau,L})(x))/\hat{s}(X_i)$ instead of $\varepsilon_{i,I}$, we have

$$\hat{F}_{X,\varepsilon_I,n}(t,y) = \hat{F}_{X,\varepsilon_{L,I},n}(t,y) + o_P(n^{-1/2})$$

uniformly on $x \in [2h_n, 1-2h_n], y \in \mathbb{R}$. Combining this with arguments which are similar to those in the proof of Theorem 1, this shows the validity of (S1.5). Next, note that on Ω_n the estimator $\hat{q}_{\tau,L}$ is strictly increasing. For every $\epsilon > 0$ it follows that

$$\begin{aligned}
& P\Big(\sup_{t\in[2h_n,1-2h_n],y\in\mathbb{R}} |S_{n,L,I}(t,y) - S_{n,L}(t,y)| > \epsilon\Big) \\
=\ & P\Big(\sup_{t\in[2h_n,1-2h_n],y\in\mathbb{R}} |S_{n,L,I}(t,y) - S_{n,L}(t,y)| > \epsilon\Big) + o(1) \\
\leq\ & P\Big(\sup_{t\in[2h_n,1-2h_n],y\in\mathbb{R}} |S_{n,L,I}(t,y) - S_{n,L}(t,y)| > \epsilon\ ,\ \sup_{x\in[h_n,1-h_n]} |\hat{q}'_{\tau,L}(x) - q'_\tau(x)| \leq \frac{c}{2}\Big) + o(1) \\
\overset{(*)}{\leq}\ & P\Big(\sup_{t\in[2h_n,1-2h_n],y\in\mathbb{R}} |S_{n,L,I}(t,y) - S_{n,L}(t,y)| > \epsilon\ ,\ \inf_{x\in[h_n,1-h_n]} \hat{q}'_{\tau,L}(x) > 0\Big) + o(1) \\
=\ & o(1).
\end{aligned}$$

Here the last equality is due to the following argumentation. If $\inf_{x\in[h_n,1-h_n]} \hat{q}'_{\tau,L}(x) > 0$, then $\hat{q}_{\tau,L}$ is strictly increasing, and for any increasing function the increasing rearrangement equals the original function function and we have $\hat{q}_{\tau,L,I} = \hat{q}_{\tau,L}$ (see Section 4). But then, $S_{n,L}(t,y) = S_{n,L,I}(t,y)$ for all $t \in [2h_n, 1-2h_n], y \in \mathbb{R}$ and the probability in $(*)$ is zero. Finally, similar arguments as those in the proof of Theorem 1 show that, uniformly with respect to $t \in [0, 2h_n) \cup (1-2h_n, 1], y \in \mathbb{R}$, we have $S_{n,L,I}(t,y) = S_{n,L}(t,y) + o_P(1)$. This completes the proof. □

S2 Validity of bootstrap

Preliminaries.

Let $\tilde{f}_\varepsilon$ denote the density corresponding to $\tilde{F}_\varepsilon$. Then under assumptions (**B1**) analogous to Lemma 2 in Neumeyer (2009) it can be shown that

$$\begin{aligned}
&\sup_{y\in\mathbb{R}} |\tilde{f}_\varepsilon(y) - f_\varepsilon(y)| = o_P\Big(\Big(\frac{h_n}{\log n}\Big)^{1/2}\Big), \quad \sup_{y\in\mathbb{R}} |y\tilde{f}_\varepsilon(y) - yf_\varepsilon(y)| = o_P(1) \qquad \text{(S2.1)} \\
&\sup_{y,z\in\mathbb{R}} \frac{|\tilde{f}_\varepsilon(y) - f(y) - \tilde{f}_\varepsilon(z) + f(z)|}{|y-z|^{\delta/2}} = o_P(1), \quad \sup_{y\in\mathbb{R}} |\tilde{F}_\varepsilon(y) - F(y)| = o_P(1)
\end{aligned}$$

(with δ from assumption (**B1**)). Further note that under assumption (**B2**), Proposition 4 in Neumeyer (2009) is valid (with υ from assumption (**B2**)) and it follows that (for some constants d and L) we have $\tilde{F}_\varepsilon \in \mathcal{D}$ with probability converging to one. Here the function class is defined as

$$\mathcal{D} = \Big\{F : \mathbb{R} \to [0,1] \ \Big|\ F \text{ increasing and continuously differentiable with derivative}$$
$$f \text{ such that } \sup_{x\in\mathbb{R}} |f(x)| + \sup_{x,x'} \frac{|f(x) - f(x')|}{|x-x'|^{\delta/2}} \leq L,$$

$$|1-F(x)| \le \frac{d}{x^{\upsilon}} \forall x>0 \text{ and } |F(x)| \le \frac{d}{|x|^{\upsilon}} \forall x<0 \Big\}. \qquad \text{(S2.2)}$$

From Lemma 4 in Neumeyer (2009) and the conditions on δ and υ in assumption (**B2**) it follows that

$$\log N(\epsilon, \mathcal{D}, ||\cdot||_\infty) = O(\epsilon^{-a}) \text{ for some } a<1. \qquad \text{(S2.3)}$$

Proof of Theorem 2.

In Lemma 9 it is shown that in the process $\hat{F}^*_{X,\varepsilon,n}$ the residuals $\hat{\varepsilon}^*_i$ can be replaced by linearized versions $\hat{\varepsilon}^*_{i,L}$ (see Section S3.1 for the definitions). Using this, the preliminaries above as well as Lemma 1 (instead of Lemma 3 in Neumeyer (2009)) we obtain analogously to the proofs of Lemma 1(i) and Theorem 2 in the reference that

$$\begin{aligned}
&\hat{F}^*_{X,\varepsilon,n}(t,y)\\
&= \frac{1}{n}\sum_{i=1}^n I\{\hat{\varepsilon}^*_{i,L} \le y\} I\{4h_n < X_i \le t\} + o_P(\frac{1}{\sqrt{n}})\\
&= \frac{1}{n}\sum_{i=1}^n I\{\varepsilon^*_i \le y\} I\{4h_n < X_i \le t\}\\
&\quad + \int \Big(\tilde{F}_\varepsilon\Big(y\frac{\hat{s}^*_L(x)}{\hat{s}_L(x)} + \frac{\hat{q}^*_{\tau,L}(x)-\hat{q}_{\tau,L}(x)}{\hat{s}_L(x)}\Big) - \tilde{F}_\varepsilon(y)\Big) I\{4h_n < x \le t\} f_X(x)\, dx\\
&\quad + o_P(\frac{1}{\sqrt{n}})
\end{aligned}$$

uniformly with respect to $t \in (4h_n, 1-4h_n]$, $y \in \mathbb{R}$. One can further apply a Taylor expansion for $\tilde{F}_\varepsilon$. Lemma 10 gives expansions for the remaining integrals and we obtain

$$\begin{aligned}
\hat{F}^*_{X,\varepsilon,n}(t,y) \;=\; & \frac{1}{n}\sum_{i=1}^n I\{4h_n < X_i \le t\}\Big(I\{\varepsilon^*_i \le y\} - \tilde{\psi}_n(y)\Big(I\{|\varepsilon^*_i| \le 1\} - \frac{1}{2}\Big)\\
&\qquad - \tilde{\phi}_n(y)\Big(I\{\varepsilon^*_i \le 0\} - \tau\Big)\Big)\\
& + o_P(\frac{1}{\sqrt{n}})
\end{aligned}$$

uniformly with respect to $t \in (4h_n, 1-4h_n]$, $y \in \mathbb{R}$, where

$$\tilde{\psi}_n(y) = \frac{y\tilde{f}_\varepsilon(y)}{f_{|\varepsilon|}(1)}, \quad \tilde{\phi}_n(y) = \frac{\tilde{f}_\varepsilon(y)}{f_\varepsilon(0)}\Big(1 - y\frac{f_\varepsilon(1)-f_\varepsilon(-1)}{f_{|\varepsilon|}(1)}\Big).$$

By the definition of the process S^*_n one now directly has

$$\begin{aligned}
&S^*_n(t,y)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big(I\{\varepsilon^*_i \le y\} - \tilde{\psi}_n(y)\Big(I\{|\varepsilon^*_i| \le 1\} - \frac{1}{2}\Big) - \tilde{\phi}_n(y)\Big(I\{\varepsilon^*_i \le 0\} - \tau\Big)\Big)
\end{aligned}$$

$$\times\Big(I\{4h_n < X_i \le t\} - I\{4h_n < X_i \le 1-4h_n\}\frac{\hat F_{X,n}(t)-\hat F_{X,n}(4h_n)}{\hat F_{X,n}(1-4h_n)-\hat F_{X,n}(4h_n)}\Big)$$
$$+ o_P(1)$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^n g_n(\varepsilon_i^*,y)\Big(I\{4h_n < X_i \le t\} - I\{4h_n < X_i \le 1-4h_n\}\frac{\hat F_{X,n}(t)-\hat F_{X,n}(4h_n)}{\hat F_{X,n}(1-4h_n)-\hat F_{X,n}(4h_n)}\Big)$$
$$+ o_P(1)$$

uniformly with respect to $t \in (4h_n, 1-4h_n]$, $y \in \mathbb{R}$, with

$$g_n(\varepsilon_i^*,y)$$
$$= I\{\varepsilon_i^* \le y\} - \tilde F_\varepsilon(y) - \tilde\phi_n(y)\Big(I\{\varepsilon_i^* \le 0\} - \tilde F_\varepsilon(0)\Big) - \tilde\psi_n(y)\Big(I\{|\varepsilon_i^*| \le 1\} - \tilde F_\varepsilon(1) + \tilde F_\varepsilon(-1)\Big).$$

Note that $E[g_n(\varepsilon_i^*,y) \mid \mathcal{Y}_n] = 0$ and the dominating part of the process S_n^* vanishes in the boundary points $t = 4h_n$ and $t = 1-4h_n$, for all $y \in \mathbb{R}$. Similarly to the corresponding arguments in the proof of Theorem 1 (but with more technical effort) it can be shown that this process is equivalent in terms of conditional weak convergence in $\ell^\infty([0,1]\times\mathbb{R})$ in probability to the process

$$S_{n,2}^*(t,y) \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^n g_n(\varepsilon_i^*,y)\Big(I\{X_i \le t\} - \hat F_{X,n}(t)\Big), \quad t\in[0,1], y\in\mathbb{R}.$$

Details are omitted for the sake of brevity.

To finish the proof we have to show that (conditional on $\mathcal{Y} = ((X_1,Y_1),(X_2,Y_2),\ldots)$) the process $S_{n,2}^*$ converges weakly to S in probability ($n\to\infty$). To this end we may show that for each subsequence $(n_k)_k$ there exists a further subsequence $(n_{k_\ell})_\ell$ such that (conditional on $\mathcal{Y}$) $S_{n_{k_\ell},2}^*$ converges weakly to S almost surely ($\ell\to\infty$), cf. Sweeting (1989), p. 463. To this end we choose a subsequence $(n_{k_\ell})_\ell$ such that along this subsequence the convergences in (S2.1) hold almost surely ($\ell\to\infty$). To simplify notation for the remainder of the proof we simply assume that the sequences in (S2.1) converge almost surely ($n\to\infty$) and show that then $S_{n,2}^*$ converges weakly to S almost surely ($n\to\infty$).

It is easy to see that the conditional covariances $\mathrm{Cov}(S_{n,2}^*(s,y), S_{n,2}^*(t,z) \mid \mathcal{Y})$ converge almost surely to $\mathrm{Cov}(S(s,y),S(t,z))$ as defined in Theorem 1. Thus it remains to show conditional tightness and conditional fidi convergence of $S_{n,2}^*$. To obtain the latter we use Cramér-Wold's device. Let $k\in\mathbb{N}$, $(y_1,t_1),\ldots,(y_k,t_k)\in\mathbb{R}\times[0,1]$, $a_1,\ldots,a_k\in\mathbb{R}$ and $Z_n = \sum_{j=1}^k a_j S_{n,2}^*(t_j,y_j) = n^{-1/2}\sum_{i=1}^n z_{n,i}$. Note that for some constant c, $|g_n(\varepsilon_i^*,y)(I\{X_i\le t\} - \hat F_{X,n}(t))| \le 1 + c(1+y)\tilde f_\varepsilon(y)$, which converges almost surely to $1+c(1+y)f_\varepsilon(y)$ due to (S2.1) and thus is almost surely bounded. From this the validity of the conditional Lindeberg condition easily follows, i. e.

$$L_n(\delta) \;=\; \frac{1}{n}\sum_{i=1}^n E[z_{n,i}^2 I\{|z_{n,i}| > n^{1/2}\delta\} \mid \mathcal{Y}] \to 0 \text{ almost surely, for all } \delta > 0.$$

Finally, to prove conditional tightness we use the decomposition $S^*_{n,2}(t,y) = \sum_{k=0}^3 U_n^{(k)}(t,y)$, where

$$\begin{aligned}
U_n^{(0)}(t,y) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big(I\{\varepsilon_i^* \le y\} - \tilde F_\varepsilon(y)\Big) I\{X_i \le t\} \\
U_n^{(1)}(t,y) &= -\tilde\phi_n(y) V_{n,1}(t) \\
&\quad \text{with } V_{n,1}(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big(I\{\varepsilon_i^* \le 0\} - \tilde F_\varepsilon(0)\Big)\Big(I\{X_i \le t\} - \hat F_{X,n}(t)\Big) \\
U_n^{(2)}(t,y) &= -\tilde\psi_n(y) V_{n,2}(t) \\
&\quad \text{with } V_{n,2}(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big(I\{|\varepsilon_i^*| \le 1\} - \tilde F_\varepsilon(1) + \tilde F_\varepsilon(-1)\Big)\Big(I\{X_i \le t\} - \hat F_{X,n}(t)\Big) \\
U_n^{(3)}(t,y) &= -\hat F_{X,n}(t) W_n(y) \text{ with } W_n(y) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big(I\{\varepsilon_i^* \le y\} - \tilde F_\varepsilon(y)\Big).
\end{aligned}$$

Note that conditional weak convergence of $V_{n,1}$ and $V_{n,2}$ to centered Gaussian processes, almost surely, can be shown analogously to the proof of bootstrap validity in Birke and Neumeyer (2013). Further conditional weak convergence of W_n is completely analogous to Theorem 4 by Neumeyer (2009). From uniform almost sure convergence of ϕ_n, ψ_n and $\hat F_{X,n}$ to bounded functions, conditional tightness of $U_n^{(k)}$ follows for $k = 1, 2, 3$.
It remains to consider $U_n^{(0)}$. Applying Corollary 1 from Shorack and Wellner (1986), p. 622, (set $a = n^{-1}$, $b = \delta = \frac{1}{2}$, $\lambda = \sqrt{n}$) and the Borel-Cantelli lemma one obtains the existence of $c \in (0, \infty)$ such that with probability one

$$|\hat F_{X,n}(t) - \hat F_{X,n}(s)| \le c|s-t|^{1/2} \quad \forall s,t \text{ with } n^{-1}\Delta_1^{-1} \le |s-t| \le \frac{1}{2}\Delta_2 \tag{S2.4}$$

for all but finitely many n, where $\Delta_1 = \inf_x f_X(x) > 0$, $\Delta_2 = \sup_x f_X(x) < \infty$.
We proceed by applying Theorem 2.11.9 by van der Vaart and Wellner (1996). Define $\mathcal{F} := [0,1] \times \mathbb{R}$ and for $f = (t,y)$ let

$$Z_{ni}(f) := \frac{1}{\sqrt{n}}\Big(I\{\varepsilon_i^* \le y\} - \tilde F_\varepsilon(y)\Big) I\{X_i \le t\}.$$

Let $\eta > 0$ and let $N_{[]}(\eta, \mathcal{F}, L_2^n)$ denote the minimal number of sets N_η in a partition of $\mathcal{F}$ in subsets $\mathcal{F}^n_{\eta j}$, $j = 1, \ldots, N_\eta$, such that for every $\mathcal{F}^n_{\eta j}$

$$\sum_{i=1}^n E\Big[\sup_{f,g \in \mathcal{F}^n_{\eta j}} |Z_{ni}(f) - Z_{ni}(g)|^2 \,\Big|\, \mathcal{Y}\Big] \le \eta^2. \tag{S2.5}$$

Here the subsets are allowed to depend on n. Note also that we consider the conditional probability measure $P(\cdot \mid \mathcal{Y})$, so the sequence $(X_1, Y_1), (X_2, Y_2), \ldots$ is given and the subsets are allowed to depend on it. We distinguish two cases.

1. Let $n \geq \Delta_1^{-1}\eta^{-4}$.
Partition $[0,1]$ into $L = O(\eta^{-4})$ intervals $[t_{\ell-1}, t_\ell]$, $\ell = 1, \dots, L$ of length $\eta^4 \leq t_\ell - t_{\ell-1} \leq 2\eta^4$ $(\forall \ell)$. Partition $\mathbb{R}$ into $K = O(\eta^{-2})$ intervals $[y_{k-1}, y_k]$, $k = 1, \dots, K$, with $\tilde{F}_\varepsilon(y_k) - \tilde{F}_\varepsilon(y_{k-1}) \leq \eta^2$ (using quantiles of the smooth distribution function $\tilde{F}_\varepsilon$). The $N_\eta = LK$ intervals $[t_{\ell-1}, t_\ell] \times [y_{k-1}, y_k]$ define the subsets $\mathcal{F}^n_{\eta j}$, $j = 1, \dots, N_\eta$.
Now fix one subset and let $f, g \in \mathcal{F}^n_{\eta j} = [t_{\ell-1}, t_\ell] \times [y_{k-1}, y_k]$. Then for monotonicity reasons $Z_{ni}(f)$ as well as $Z_{ni}(g)$ are elements of the bracket $[Z_{ni}^{k,\ell,l}, Z_{ni}^{k,\ell,u}]$, where

$$
\begin{aligned}
Z_{ni}^{k,\ell,l} &= \frac{1}{\sqrt{n}}\Big(I\{\varepsilon_i^* \leq y_{k-1}\}I\{X_i \leq t_{\ell-1}\} - \tilde{F}_\varepsilon(y_k)I\{X_i \leq t_\ell\}\Big)\\
Z_{ni}^{k,\ell,u} &= \frac{1}{\sqrt{n}}\Big(I\{\varepsilon_i^* \leq y_k\}I\{X_i \leq t_\ell\} - \tilde{F}_\varepsilon(y_{k-1})I\{X_i \leq t_{\ell-1}\}\Big).
\end{aligned}
$$

Thus the left hand side of (S2.5) can be bounded by

$$
\begin{aligned}
&\sum_{i=1}^n E\Big[(Z_{ni}^{k,\ell,u} - Z_{ni}^{k,\ell,l})^2 \,\Big|\mathcal{Y}\Big]\\
\leq\ & \frac{2}{n}\sum_{i=1}^n (I\{X_i \leq t_\ell\} - I\{X_i \leq t_{\ell-1}\})^2\\
&+ \frac{2}{n}\sum_{i=1}^n E\Big[\Big(I\{\varepsilon_i^* \leq y_k\} - \tilde{F}_\varepsilon(y_{k-1}) - I\{\varepsilon_i^* \leq y_{k-1}\} + \tilde{F}_\varepsilon(y_k)\Big)^2 \,\Big|\mathcal{Y}\Big]\\
\leq\ & \frac{2}{n}\sum_{i=1}^n (I\{X_i \leq t_\ell\} - I\{X_i \leq t_{\ell-1}\})^2\\
&+ \frac{4}{n}\sum_{i=1}^n E\Big[I\{\varepsilon_i^* \leq y_k\} - I\{\varepsilon_i^* \leq y_{k-1}\} + \tilde{F}_\varepsilon(y_k) - \tilde{F}_\varepsilon(y_{k-1}) \,\Big|\mathcal{Y}\Big]\\
\leq\ & 2(\hat{F}_{X,n}(t_\ell) - \hat{F}_{X,n}(t_{\ell-1})) + 8(\tilde{F}_\varepsilon(y_k) - \tilde{F}_\varepsilon(y_{k-1}))\\
\leq\ & 2(\hat{F}_{X,n}(t_\ell) - \hat{F}_{X,n}(t_{\ell-1})) + 8\eta^2 \ \leq\ C\eta^2, \qquad \text{(S2.6)}
\end{aligned}
$$

where we have used (S2.4) and $t_\ell - t_{\ell-1} \geq \eta^4 \geq n^{-1}\Delta_1^{-1}$, and the constant C does not depend on n and η.
2. Let $n < \Delta_1^{-1}\eta^{-4}$.
As before we partition $\mathbb{R}$ into $K = O(\eta^{-4})$ intervals $[y_{k-1}, y_k]$, $k = 1, \dots, K$, with $\tilde{F}_\varepsilon(y_k) - \tilde{F}_\varepsilon(y_{k-1}) \leq \eta^2$. We partition $[0,1]$ into $n+2 = O(\eta^{-4})$ intervals $I_\ell = [t_{\ell-1}, t_\ell)$, $\ell = 1, \dots, n+1$, and $I_{n+2} = \{1\}$, where $t_0 = 0$, $t_\ell = X_{(\ell)}$ for $\ell = 1, \dots, n$ and $t_{n+1} = 1$. Here $X_{(1)}, \dots, X_{(n)}$ denote the order statistics of $X_1, \dots, X_n$. Now we proceed as in case 1 but replacing $Z_{ni}^{k,\ell,u}, Z_{ni}^{k,\ell,l}$ with

$$
\begin{aligned}
\tilde{Z}_{ni}^{k,\ell,l} &= \frac{1}{\sqrt{n}}\Big(I\{\varepsilon_i^* \leq y_{k-1}\}I\{X_i \leq t_{\ell-1}\} - \tilde{F}_\varepsilon(y_k)I\{X_i < t_\ell\}\Big)\\
\tilde{Z}_{ni}^{k,\ell,u} &= \frac{1}{\sqrt{n}}\Big(I\{\varepsilon_i^* \leq y_k\}I\{X_i < t_\ell\} - \tilde{F}_\varepsilon(y_{k-1})I\{X_i \leq t_{\ell-1}\}\Big).
\end{aligned}
$$

By definition, $\tilde{Z}_{ni}^{k,\ell,l} \leq Z_{ni}(f) \leq \tilde{Z}_{ni}^{k,\ell,u}$ for $f = (t,y) \in [t_{\ell-1}, t_\ell) \times [y_{k-1}, y_k]$. Noting that $\hat{F}_{X,n}(t_\ell -) - \hat{F}_{X,n}(t_{\ell-1}) = 0$ for all $\ell = 1, \ldots, n+1$, we obtain by similar arguments as used to derive (S2.6)

$$\sum_{i=1}^{n} E\Big[(\tilde{Z}_{ni}^{k,\ell,u} - \tilde{Z}_{ni}^{k,\ell,l})^2 \,\Big|\mathcal{Y}\Big] \;\leq\; 2(\hat{F}_{X,n}(t_\ell -) - \hat{F}_{X,n}(t_{\ell-1})) + 8\eta^2 = 8\eta^2.$$

The partitionings in both cases depend on n, but the bracketing number $N_{[]}(\eta, \mathcal{F}, L_2^n)$ can be bounded by $O(\eta^{-8})$, independent of n, such that the condition

$$\int_0^{\delta_n} \sqrt{\log N_{[]}(\eta, \mathcal{F}, L_2^n)}\, d\eta \longrightarrow 0 \text{ for every } \delta_n \searrow 0$$

is fulfilled (this corresponds to the third condition in Theorem 2.11.9 by van der Vaart and Wellner (1996)). Further, because $|Z_{ni}(f)| \leq n^{-1/2}\ \forall f$ we have

$$\sum_{i=1}^{n} E\Big[\sup_{f\in\mathcal{F}} |Z_{ni}(f)| I\{\sup_{f\in\mathcal{F}} |Z_{ni}(f)| > \eta\} \,\Big|\mathcal{Y}\Big] \longrightarrow 0 \text{ for every } \eta > 0$$

(this corresponds to the first condition in Theorem 2.11.9 by van der Vaart and Wellner (1996)). Moreover, $(\mathcal{F}, \rho)$ is a totally bounded semimetric space with $\rho((s,y),(t,z)) = |t-s| + |F_\varepsilon(z) - F_\varepsilon(y)|$. Now for $\delta_n \searrow 0$ we obtain similarly to the calculation in case 1 above (for some constant c),

$$\begin{aligned}
&\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{n} E\Big[(Z_{ni}(f) - Z_{ni}(g))^2 \,\Big|\mathcal{Y}\Big] \\
&\leq c\Big(\sup_{|t-s|\leq\delta_n} |\hat{F}_{X,n}(t) - \hat{F}_{X,n}(s)| + \sup_{\substack{z,y:\\ |F_\varepsilon(z)-F_\varepsilon(y)|\leq\delta_n}} |\tilde{F}_\varepsilon(z) - \tilde{F}_\varepsilon(y)|\Big) \\
&= o(1) \text{ almost surely}
\end{aligned}$$

by uniform convergence of $\hat{F}_{X,n}$ to F_X and $\tilde{F}_\varepsilon$ to F_ε (this corresponds to the second condition in Theorem 2.11.9 by van der Vaart and Wellner (1996)) and uniform continuity of F_X. From Theorem 2.11.9 one obtains

$$\lim_{\delta\searrow 0} \lim_{n\to\infty} P\Big(\sup_{\rho((s,y),(t,z))<\delta} |\tilde{U}_n^{(0)}(s,y) - \tilde{U}_n^{(0)}(t,z)| > \eta \,\Big|\, \mathcal{Y}\Big) = 0 \text{ for all } \eta > 0$$

for almost all sequences $\mathcal{Y}$. This completes the proof. □

Proof of Theorem 4.

Theorem 4 follows from Theorem 2 in the same manner as Theorem 3 follows from Theorem 1. □

S3 Technical results

We begin by recalling some notation from the main body of the paper that will be used throughout the proofs.

One fact that we will use throughout is that the bootstrap residuals ε_i^* can be represented as $\varepsilon_i^* = \tilde{F}_\varepsilon^{-1}(U_i)$ where $U_1, ..., U_n$ denote a sample of i.i.d. $\mathcal{U}[0,1]$ random variables that are independent of the original sample and

$$\tilde{F}_\varepsilon(y) = \frac{\frac{1}{n}\sum_{i=1}^n \Phi\Big(\frac{y-\hat{\varepsilon}_i}{\alpha_n}\Big) I\{2h_n < X_i \le 1-2h_n\}}{\hat{F}_{X,n}(1-2h_n) - \hat{F}_{X,n}(2h_n)}$$

denotes the distribution function of ε_1^* conditional on the sample, see (3.4). Additionally, we will use the abbreviation

$$r_n := \Big(\frac{\log n}{nh_n}\Big)^{1/2}.$$

Next, we introduce some additional notation that will be used throughout. First, introduce the functional

$$Q_{G,\kappa,\tau,b_n}(F) := G^{-1}\Big(\frac{1}{b_n}\int_0^1 \int_{-\infty}^\tau \kappa\Big(\frac{F(G^{-1}(u)) - v}{b_n}\Big) dv du\Big)$$

which is defined for arbitrary functions F that are uniformly bounded. Some properties of this functional are collected in Lemma 6. Additionally, define the quantities

$$\begin{aligned}
\hat{F}_Y^*(y|x) &:= \sum_{i=1}^n W_i(x)\Omega\Big(\frac{y-Y_i^*}{d_n}\Big), \quad \hat{q}_\tau^*(x) := Q_{G,\kappa,\tau,b_n}(\hat{F}_Y^*(\cdot|x)),\\
\hat{F}_{|e|}^*(y|x) &:= \sum_{i=1}^n W_i(x)\Omega\Big(\frac{y-|Y_i^*-\hat{q}_\tau^*(X_i)|}{d_n}\Big), \quad \hat{s}^*(x) := Q_{G,\kappa,1/2,b_n}(\hat{F}_{|e|}^*(\cdot|x)).
\end{aligned}$$

where the weights W_i are the same as in equation (2.3). Observe that the estimators $\hat{q}_\tau, \hat{s}$ which we introduced in the main body of the paper admit the representations

$$\hat{q}_\tau(x) = Q_{G,\kappa,\tau,b_n}(\hat{F}_Y(\cdot|x)), \quad \hat{s}(x) := Q_{G,\kappa,1/2,b_n}(\hat{F}_{|e|}(\cdot|x)).$$

In Section S3.1, we will introduce linearized versions of the estimators $\hat{q}_\tau, \hat{q}_\tau^*, \hat{s}, \hat{s}^*$, those will be denoted by $\hat{q}_{\tau,L}, \hat{q}_{\tau,L}^*, \hat{s}_L, \hat{s}_L^*$. Key results there are Lemma 1 and Lemma 2 which state that the linearized versions are uniformly close to the original estimators and that the linearized versions have certain smoothness properties, respectively. The rest of the supplement is organized as follows. Section S3.1 contains results about the estimators $\hat{q}_\tau, \hat{q}_\tau^*, \hat{s}, \hat{s}^*$ and their linearizations. The proofs of those results require additional technical Lemmas, that we collect and prove in Section S3.2. Finally, some key results which are used in the main body of the paper and whose proofs rely on findings in Sections S3.1 and S3.2 can be found in Section S3.3.

S3.1 Properties of $\hat{q}_\tau$ and $\hat{s}$

We start this section by introducing some notation and giving an overview of the derived results. Our first key result is an asymptotic representation of the form

$$\begin{aligned}\hat{F}_Y(y|x) &= \hat{F}_{Y,L,S}(y|x) + o_P(1/\sqrt{n}), \qquad \hat{F}_{|e|}(y|x) = F_{|e|,L,S}(y|x) + o_P(1/\sqrt{n}),\\ \hat{F}^*_Y(y|x) &= \hat{F}^*_{Y,L,S}(y|x) + o_P(1/\sqrt{n}), \qquad \hat{F}^*_{|e|}(y|x) = F^*_{|e|,L,S}(y|x) + o_P(1/\sqrt{n}),\end{aligned}$$

holding uniformly over x, y where the expressions on the right-hand side of the above equations are defined as

$$\begin{aligned}\hat{F}_{Y,L,S}(y|x) &:= F_Y(y|x) + u_1^t \mathcal{M}(K)^{-1}\Big(T_{n,0,L,S}(x,y), \dots, T_{n,p,L,S}(x,y)\Big)^t,\\ \hat{F}_{|e|,L,S}(y|x) &:= F_{|e|}(y|x) + u_1^t \mathcal{M}(K)^{-1}\Big(T_{|e|,n,0,L,S}(x,y), \dots, T_{|e|,n,p,L,S}(x,y)\Big)^t,\\ \hat{F}^*_{Y,L,S}(y|x) &:= F_Y(y|x) + u_1^t \mathcal{M}(K)^{-1}\Big(T^*_{n,0,L,S}(x,y), \dots, T^*_{n,p,L,S}(x,y)\Big)^t,\\ \hat{F}^*_{|e|,L,S}(y|x) &:= F_{|e|}(y|x) + u_1^t \mathcal{M}(K)^{-1}\Big(T^*_{|e|,n,0,L,S}(x,y), \dots, T^*_{|e|,n,p,L,S}(x,y)\Big)^t,\end{aligned}$$

$u_1^t := (1, 0, ..., 0)$ denotes the first unit vector in $\mathbb{R}^{p+1}$, $\mathcal{M}(K)$ denotes a $(p+1) \times (p+1)$ matrix with entries

$$\mathcal{M}(K)_{ij} = \mu_{i+j-2}(K) := \int u^{i+j-2} K(u) du,$$

and

$$\begin{aligned}T_{n,k,L,S}(x,y) &:= \frac{1}{nh}\sum_{i=1}^n \frac{1}{f_X(X_i)} K_{h,k}(x - X_i)\Big(\Omega\Big(\frac{y - Y_i}{d_n}\Big) - F_Y(y|X_i)\Big),\\ T_{|e|,n,k,L,S}(x,y) &:= \frac{1}{nh}\sum_{i=1}^n \frac{1}{f_X(X_i)} K_{h,k}(x - X_i)\Big(\Omega\Big(\frac{y - |Y_i - \hat{q}_{\tau,L}(X_i)|}{d_n}\Big) - F_{|e|}(y|X_i)\Big),\\ T^*_{n,k,L,S}(x,y) &:= \frac{1}{nh}\sum_{i=1}^n \frac{1}{f_X(X_i)} K_{h,k}(x - X_i)\Big(\Omega\Big(\frac{y - Y_i^*}{d_n}\Big) - F_Y(y|X_i)\Big),\\ T^*_{|e|,n,k,L,S}(x,y) &:= \frac{1}{nh}\sum_{i=1}^n \frac{1}{f_X(X_i)} K_{h,k}(x - X_i)\Big(\Omega\Big(\frac{y - |Y_i^* - \hat{q}^*_{\tau,L}(X_i)|}{d_n}\Big) - F_{|e|}(y|X_i)\Big).\end{aligned}$$

This, and further properties as differentiability and convergence rates of $\hat{F}_{Y,L,S}(y|x)$, $\hat{F}_{|e|,L,S}$, $\hat{F}^*_{Y,L,S}$, $\hat{F}^*_{|e|,L,S}$ is the subject of Lemma 3.

The results in Lemma 6 and properties of the estimators $\hat{F}_Y, \hat{F}_{|e|}, \hat{F}^*_Y, \hat{F}^*_{|e|}$ yield representations of the form

$$\begin{aligned}\hat{q}_\tau(x) &= \hat{q}_{\tau,L}(x) + o_P(n^{-1/2}), \qquad \hat{s}(x) = \hat{s}_L(x) + o_P(n^{-1/2}),\\ \hat{q}^*_\tau(x) &= \hat{q}^*_{\tau,L}(x) + o_P(n^{-1/2}), \qquad \hat{s}^*(x) = \hat{s}^*_L(x) + o_P(n^{-1/2})\end{aligned}$$

uniformly in x [see Lemma 2] where

$$
\begin{aligned}
\hat{q}_{\tau,L}(x) &:= q_\tau(x) - \frac{1}{f_e(0|x)} \int_{-1}^{1} \Big(\hat{F}_{Y,L,S}(q_{\tau+vb_n}(x)|x) - F_Y(q_{\tau+vb_n}(x)|x) \Big) \kappa(v) dv \\
&= q_\tau(x) - \frac{u_1^t \mathcal{M}(K)^{-1}}{f_e(0|x)} \int_{-1}^{1} \kappa(v) \Big(T_{n,0,L,S}(x, q_{\tau+vb_n}(x)), \ldots, T_{n,p,L,S}(x, q_{\tau+vb_n}(x)) \Big)^t dv \\
\hat{s}_L(x) &:= s(x) - \frac{1}{f_{|\varepsilon|}(1|x)} \int_{-1}^{1} \Big(\hat{F}_{|e|,L,S}(s_{1/2+vb_n}(x)|x) - F_{|e|}(s_{1/2+vb_n}(x)|x) \Big) \kappa(v) dv \\
&= s(x) - \frac{u_1^t \mathcal{M}(K)^{-1}}{f_{|\varepsilon|}(1)} \int_{-1}^{1} \kappa(v) \Big(T_{|e|,n,0,L,S}(x, s_{1/2+vb_n}(x)), \ldots, T_{|e|,n,p,L,S}(x, s_{1/2+vb_n}(x)) \Big)^t dv \\
\hat{q}^*_{\tau,L}(x) &:= q_\tau(x) - \frac{1}{f_e(0|x)} \int_{-1}^{1} \Big(\hat{F}^*_{Y,L,S}(q_{t+vb_n}(x)|x) - F_Y(q_{t+vb_n}(x)|x) \Big) \kappa(v) dv \\
&= q_\tau(x) - \frac{u_1^t \mathcal{M}(K)^{-1}}{f_e(0|x)} \int_{-1}^{1} \kappa(v) \Big(T^*_{n,0,L,S}(x, q_{\tau+vb_n}(x)), \ldots, T^*_{n,p,L,S}(x, q_{\tau+vb_n}(x)) \Big)^t dv \\
\hat{s}^*_L(x) &:= s(x) - \frac{1}{f_{|\varepsilon|}(1)} \int_{-1}^{1} \Big(\hat{F}^*_{|e|,L,S}(s_{1/2+vb_n}(x)|x) - F_{|e|}(s_{1/2+vb_n}(x)|x) \Big) \kappa(v) dv \\
&= s(x) - \frac{u_1^t \mathcal{M}(K)^{-1}}{f_{|\varepsilon|}(1)} \int_{-1}^{1} \kappa(v) \Big(T^*_{|e|,n,0,L,S}(x, s_{1/2+vb_n}(x)), \ldots, T^*_{|e|,n,p,L,S}(x, s_{1/2+vb_n}(x)) \Big)^t dv
\end{aligned}
$$

where $s_\alpha(x) := F^{-1}_{|e|}(\alpha|x)$. Differentiability properties and convergence rates of derivatives of these estimators can obviously be derived from the corresponding properties of the underlying distribution function estimators, see Lemma 1.

Lemma 1 *Let (**K1**)-(**K6**), (**A1**)-(**A5**), (**BW**) hold. Then for any* $k \le 2$

$$
\begin{aligned}
\sup_{x\in[h_n,1-h_n]} |\hat{q}^{(k)}_{\tau,L}(x) - q^{(k)}_\tau(x)| &= O_P\Big(\frac{\log h_n^{-1}}{nh_n(h_n \wedge d_n)^{2k}}\Big)^{1/2} = o_P(1), \\
\sup_{x\in[2h_n,1-2h_n]} |\hat{s}^{(k)}_L(x) - s^{(k)}(x)| &= O_P\Big(\frac{\log h_n^{-1}}{nh_n(h_n \wedge d_n)^{2k}}\Big)^{1/2} = o_P(1),
\end{aligned}
$$

*and under (**B1**)-(**B2**) it follows that*

$$
\begin{aligned}
\sup_{x\in[3h_n,1-3h_n]} |(\hat{q}^*_{\tau,L})^{(k)}(x) - q^{(k)}_\tau(x)| &= O_P\Big(\frac{\log h_n^{-1}}{nh_n(h_n \wedge d_n)^{2k}}\Big)^{1/2} = o_P(1), \\
\sup_{x\in[4h_n,1-4h_n]} |(\hat{s}^*_L)^{(k)}(x) - s^{(k)}(x)| &= O_P\Big(\frac{\log h_n^{-1}}{nh_n(h_n \wedge d_n)^{2k}}\Big)^{1/2} = o_P(1).
\end{aligned}
$$

Proof of Lemma 1 Since all claims share the same structure, we will only establish that

$$
\sup_{x\in[h_n,1-h_n]} |\hat{q}^{(k)}_{\tau,L}(x) - q^{(k)}_\tau(x)| = O_P\Big(\frac{\log h_n^{-1}}{nh_n(h_n \wedge d_n)^{2k}}\Big)^{1/2} = o_P(1).
$$

Observe that by definition of $\hat q_{\tau,L}$ we have

$$\hat q_{\tau,L}^{(k)}(x) - q_\tau^{(k)}(x) = -\frac{\partial^k}{\partial x^k}\Big(\frac{1}{f_e(0|x)}\int_{-1}^1 \Big(\hat F_{Y,L,S}(q_{\tau+vb_n}(x)|x) - F_Y(q_{\tau+vb_n}(x)|x)\Big)\kappa(v)dv\Big).$$

Observing that $f_e(0|x) = f_\varepsilon(0)/s(x)$, it suffices to show that

$$\sup_{\substack{x\in[h_n,1-h_n]\\ v\in[-1,1]}}\sup_{m\le k}\Big|\frac{\partial^m}{\partial x^m}\Big(\hat F_{Y,L,S}(q_{\tau+vb_n}(x)|x) - F_Y(q_{\tau+vb_n}(x)|x)\Big)\Big| \;=\; O_P\Big(\frac{\log h_n^{-1}}{nh_n(h_n\wedge d_n)^{2k}}\Big)^{1/2}.$$

Now by Remark 2 in the main body of the paper, the function $x \mapsto q_{\tau+vb_n}(x)$ is 2 times continuously differentiable and its derivatives are bounded uniformly over $x \in (0,1), v \in [-1,1]$. Thus the above assertion follows from (i) of Lemma 3 combined with the chain rule for derivatives. □

Lemma 2 *Let (**K1**)-(**K6**), (**A1**)-(**A5**), (**BW**) hold. Then*

$$(i)\quad \sup_{x\in[h_n,1-h_n]}|\hat q_\tau(x) - \hat q_{\tau,L}(x)| = o_P(1/\sqrt{n}),$$
$$(ii)\quad \sup_{x\in[2h_n,1-2h_n]}|\hat s(x) - \hat s_L(x)| = o_P(1/\sqrt{n}),$$

*and if additionally (**B1**)-(**B2**) hold, we also have*

$$(iii)\quad \sup_{x\in[3h_n,1-3h_n]}|\hat q_\tau^*(x) - \hat q_{\tau,L}^*(x)| = o_P(1/\sqrt{n}),$$
$$(iv)\quad \sup_{x\in[4h_n,1-4h_n]}|\hat s^*(x) - \hat s_L^*(x)| = o_P(1/\sqrt{n}).$$

Proof Since all assertions share a similar structure, we will only prove (iii). We begin by stating and intermediate result which we will establish in the end.

$$\sup_{y\in\mathbb{R}}\sup_{x\in[3h_n,1-3h_n]}|\hat F_Y^*(y|x) - F_Y(y|x)| = o_P(1). \tag{S3.7}$$

Note that, in contrast to the statements in Lemma 3 part (iii), the range for y is $\mathbb{R}$ instead of a bounded set. Now let $\delta > 0, c_0 > 0$ be such that $\inf_{x\in[0,1]}\inf_{|y-q_\tau(x)|\le 2\delta} f_Y(y|x) \ge c_0$ and define

$$F_Y^*(y|x) := \hat F_Y^*(y|x)I\{|y-q_\tau(x)| \le 2\delta/c_0\} + F_Y(y|x)I\{|y-q_\tau(x)| > 2\delta/c_0\}.$$

By the results in Lemma 3 parts (iii), (iii)' we have

$$\sup_{y\in\mathbb{R}}\sup_{x\in[3h_n,1-3h_n]}|F_Y^*(y|x) - F_Y(y|x)| = O_P\Big(\frac{\log n}{nh_n}\Big)^{1/2}, \tag{S3.8}$$

and

$$\sup_{x\in[3h_n,1-3h_n]}\sup_{|y-q_\tau(x)|\le 2\delta/c_0}|F_Y^*(y|x)-\hat F_{Y,L,S}^*(y|x)| = o_P(n^{-1/2}). \tag{S3.9}$$

Moreover, as we shall prove later, we have

$$P\Big(Q_{G,\kappa,\tau,b_n}(\hat F_Y^*(\cdot|x)) = Q_{G,\kappa,\tau,b_n}(F_Y^*(\cdot|x))\ \forall x\in[3h_n,1-3h_n]\Big)\to 1. \tag{S3.10}$$

Now apply part (c) of Lemma 6 with $F=F_1=F_Y(\cdot|x), F_2=F_Y^*(\cdot|x)$. A careful inspection of the remainder terms in the statement of Lemma 6 part (c) shows that, uniformly in $x\in[3h_n,1-3h_n]$,

$$\begin{aligned}&Q_{G,\kappa,\tau,b_n}(F_Y^*(\cdot|x)) - Q_{G,\kappa,\tau,b_n}(F_Y(\cdot|x))\\ =&\ -\frac{1}{f_e(0|x)}\int_{-1}^1\kappa(v)\Big(F_Y^*(q_{\tau+vb_n}(x)|x)-F_Y(q_{\tau+vb_n}(x)|x)\Big)dv + o_P(n^{-1/2}).\end{aligned} \tag{S3.11}$$

An application of Lemma 6, part (a) with $F=F_Y(\cdot|x)$ shows that

$$Q_{G,\kappa,\tau,b_n}(F_Y(\cdot|x)) = q_\tau(x)+O(b_n^2) = q_\tau(x)+o(n^{-1/2})$$

uniformly in $x\in[0,1]$. Combining this with (S3.9), (S3.10) and (S3.11) and observing that $\hat q_\tau^*(x) = Q_{G,\kappa,\tau,b_n}(\hat F_Y^*(\cdot|x))$ we obtain, uniformly in $x\in[3h_n,1-3h_n]$,

$$\begin{aligned}&\hat q_\tau^*(x)-q_\tau(x)\\ =&\ -\frac{1}{f_e(0|x)}\int_{-1}^1\kappa(v)\Big(\hat F_{Y,L,S}^*(q_{\tau+vb_n}(x)|x)-F_Y(q_{\tau+vb_n}(x)|x)\Big)dv + o_P(n^{-1/2}).\end{aligned}$$

Note that, by the definition of $\hat q_{\tau,L}^*(x)$, the leading term in this representation is equal to $\hat q_{\tau,L}^*(x)-q_\tau(x)$. This implies statement (iii), and thus it remains to prove (S3.7) and (S3.10).

Proof of (S3.7) Define (with W_i the same as defined in (2.3))

$$\hat F_{Y,U}^*(y|x) := \sum_{i=1}^n W_i(x)I\{Y_i^*\le y\}.$$

Since

$$\hat F_Y^*(y|x) = (\hat F_{Y,U}^*(\cdot|x) * \frac{1}{d_n}\omega(\cdot/d_n))(y)$$

and by the smoothness of F_Y, it suffices to prove that

$$\sup_{y\in\mathbb{R}}\sup_{x\in[3h_n,1-3h_n]}|\hat F_{Y,U}^*(y|x)-F_Y(y|x)| = o_P(1). \tag{S3.12}$$

Now by the definition of Y_i^* we have

$$\hat F_{Y,U}^*(y|x) = \sum_{i=1}^n W_i(x)I\{\hat q_\tau(X_i)+\hat s(X_i)\tilde F_\varepsilon^{-1}(U_i)\le y\} = \sum_{i=1}^n W_i(x)I\Big\{U_i\le \tilde F_\varepsilon\Big(\frac{y-\hat q_\tau(X_i)}{\hat s(X_i)}\Big)\Big\}.$$

From (S2.1) in the main body of the paper we obtain after a Taylor expansion

$$\sup_{x\in[3h_n,1-3h_n]}\sup_{y\in\mathbb{R}}\Big|\tilde F_\varepsilon\Big(\frac{y-\hat q_\tau(x)}{\hat s(x)}\Big)-\tilde F_\varepsilon\Big(\frac{y-q_\tau(x)}{s(x)}\Big)\Big|=o_P(1).$$

Since the conclusion of Lemma 2 in Neumeyer (2009) remains valid in our setting [see the discussion in the beginning of Section S2], it follows that $\sup_{z\in\mathbb{R}}|\tilde F_\varepsilon(z)-F_\varepsilon(z)|=o_P(1)$ and thus

$$\sup_{x\in[3h_n,1-3h_n]}\sup_{y\in\mathbb{R}}\Big|\tilde F_\varepsilon\Big(\frac{y-\hat q_\tau(x)}{\hat s(x)}\Big)-F_\varepsilon\Big(\frac{y-q_\tau(x)}{s(x)}\Big)\Big|=o_P(1).$$

Thus there exists a deterministic sequence $\gamma_n\to 0$ such that $P(D_n)\to 1$ where we defined the event

$$D_n:=\Big\{\sup_{x\in[3h_n,1-3h_n]}\sup_{y\in\mathbb{R}}\Big|\tilde F_\varepsilon\Big(\frac{y-\hat q_\tau(x)}{\hat s(x)}\Big)-F_\varepsilon\Big(\frac{y-q_\tau(x)}{s(x)}\Big)\Big|\le\gamma_n\Big\}.$$

Additionally, define the event

$$\tilde D_n:=\{\sup_i\sup_{x\in[h_n,1-h_n]}|W_i(x)|\le C(nh_n)^{-1}I\{|x-X_i|\le h_n\}\}$$

and observe that $P(\tilde D_n)\to 1$ by the definition of $W_i(x)$ and Lemma 4. Thus on $D_n\cap\tilde D_n$ we have

$$\begin{aligned}&\sup_{y\in\mathbb{R}}\sup_{x\in[3h_n,1-3h_n]}\Big|\hat F^*_{Y,U}(y|x)-\sum_{i=1}^n W_i(x)I\Big\{U_i\le F_Y(y|X_i)\Big\}\Big|\\ &\le\frac{C}{nh_n}\sup_{y\in\mathbb{R}}\sup_{x\in[3h_n,1-3h_n]}\sum_{i=1}^n I\{|X_i-x|\le h_n\}I\Big\{\Big|U_i-F_Y(y|X_i)\Big|\le\gamma_n\Big\}=o_P(1)\end{aligned}\qquad\text{(S3.13)}$$

where the last equality follows by a combination of parts 1, 4-6 of Lemma 8 with Lemma 7. Similarly, applying Lemma 4, parts 1,2 4-6 of Lemma 8 with Lemma 7 shows that

$$\sum_{i=1}^n W_i(x)\Big(I\Big\{U_i\le F_Y(y|X_i)\Big\}-F_Y(y|X_i)\Big)=o_P(1)\qquad\text{(S3.14)}$$

uniformly in $x\in[3h_n,1-3h_n],y\in\mathbb{R}$. Finally, by similar arguments as used in the proof of (S3.22) one can show that

$$\sum_{i=1}^n W_i(x)F_Y(y|X_i)=F_Y(y|x)+o_P(1)\qquad\text{(S3.15)}$$

uniformly in $x\in[3h_n,1-3h_n],y\in\mathbb{R}$. Combining (S3.13)-(S3.15) yields (S3.12) and completes the proof of (S3.7).

Proof of (S3.10) Define the events

$$
\begin{aligned}
D_{n1} &:= \Big\{\hat{F}^*_Y(y|x) = F^*_Y(y|x)\ \forall (x,y) \in \{(x,y) : |F^*_Y(y|x) - \tau| \le \delta, x \in [3h_n, 1-3h_n]\}\Big\} \\
D_{n2} &:= \Big\{\sup_{x\in[3h_n,1-3h_n], y\in\mathbb{R}} |\hat{F}^*_Y(y|x) - F^*_Y(y|x)| \le \delta/2\Big\} \\
D_{n3} &:= \Big\{\sup_{x\in[3h_n,1-3h_n], y\in\mathbb{R}} |\hat{F}^*_Y(y|x) - F_Y(y|x)| \le \delta/2\Big\}.
\end{aligned}
$$

Observe that on $D_{n1}\cap D_{n2}\cap D_{n3}$ we have $F^*_Y(y|x) \le \tau-\delta \Rightarrow \hat{F}^*_Y(y|x) \le \tau - \delta/2$, $F^*_Y(y|x) \ge \tau+\delta \Rightarrow \hat{F}^*_Y(y|x) \ge \tau+\delta/2$ and $|F^*_Y(y|x)-\tau| \le \delta \Rightarrow F^*_Y(y|x) = \hat{F}^*_Y(y|x)$. Thus on $D_{n1}\cap D_{n2}\cap D_{n3}$ we obtain $Q_{G,\kappa,\tau,b_n}(\hat{F}^*_Y(\cdot|x)) = Q_{G,\kappa,\tau,b_n}(F^*_Y(\cdot|x))$ provided that $b_n \le \delta/2$. It remains to prove that $P(D_{n1}\cap D_{n2}\cap D_{n3}) \to 1$. The fact that $P(D_{n2}\cap D_{n3}) \to 1$ follows from (S3.7), (S3.8), so that it remains to prove $P(D_{n1}) \to 1$ which follows from

$$
P\Big(\{(x,y) : |F^*_Y(y|x) - \tau| \le \delta, x\in[3h_n,1-3h_n]\} \subset \{(x,y) : |y - q_\tau(x)| \le 2\delta/c_0, x\in[3h_n,1-3h_n]\}\Big) \to 1.
$$

This in turn is a consequence of the fact that on D_{n3} (note that $|F^*_Y(y|x) - F_Y(y|x)| \le |\hat{F}^*_Y(y|x) - F_Y(y|x)|$)

$$
|F^*_Y(y|x) - \tau| \le \delta \Rightarrow |F_Y(y|x) - \tau| \le 3\delta/2 \Rightarrow |y - q_\tau(x)| \le 3\delta/(2c_0)
$$

by the definition of δ, c_0. This completes the proof.

□

Lemma 3 *Assume that conditions* ***(K1)****-****(K6)****,* ***(A1)****-****(A5)*** *and* ***(BW)*** *hold. Denote by* $\tilde{T}_{n,0,L,S}, \tilde{T}_{|e|,n,0,L,S}, \tilde{T}^*_{n,0,L,S}, \tilde{T}^*_{|e|,n,0,L,S}$ *versions of* $T_{n,0,L,S}, T_{|e|,n,0,L,S}, T^*_{n,0,L,S}, T^*_{|e|,n,0,L,S}$ *where* $1/f_X(X_i)$ *is replaced by* $1/f_X(x)$.
Then for any bounded $\mathcal{Y}_1 \subset \mathbb{R}, \mathcal{Y}_2 \subset \mathbb{R}^+$ *such that* $\mathcal{Y}_2$ *is bounded away from zero we have*

$$
(i)' \quad \hat{F}_Y(y|x) \ = \ \hat{F}_{Y,L,S}(y|x) + o_P(1/\sqrt{n}), \quad T_{n,0,L,S} = \tilde{T}_{n,0,L,S} + o_P(1/\sqrt{n}),
$$

uniformly in $y \in \mathcal{Y}_1, x \in [h_n, 1-h_n]$ *and*

$$
(ii)' \quad \hat{F}_{|e}|(y|x) \ = \ \hat{F}_{|e|,L,S}(y|x) + o_P(1/\sqrt{n}), \quad T_{|e|,n,0,L,S} = \tilde{T}_{|e|,n,0,L,S} + o_P(1/\sqrt{n}),
$$

uniformly in $y \in \mathcal{Y}_2, x \in [2h_n, 1-2h_n]$*. If additionally* ***(B1)****-****(B2)*** *hold,*

$$
(iii)' \quad \hat{F}^*_Y(y|x) \ = \ \tilde{T}^*_{n,0,L,S} + o_P(1/\sqrt{n}), \quad T^*_{n,0,L,S} = \tilde{T}^*_{n,0,L,S} + o_P(1/\sqrt{n}),
$$

uniformly in $y \in \mathcal{Y}_1, x \in [3h_n, 1-3h_n]$ *and*

$$
(iv)' \quad \hat{F}^*_{|e|}(y|x) \ = \ \hat{F}^*_{|e|,L,S}(y|x) + o_P(1/\sqrt{n}), \quad T^*_{|e|,n,0,L,S} = \tilde{T}^*_{|e|,n,0,L,S} + o_P(1/\sqrt{n}).
$$

uniformly in $y \in \mathcal{Y}_2, x \in [4h_n, 1-4h_n]$.

Moreover, (i)-(iv) hold under the assumptions of $(i)' - (iv)'$*, respectively.*

$$(i)\quad \forall k+l \le 2 \qquad \sup_{y\in\mathcal{Y}_1, x\in[h_n,1-h_n]} |\partial_x^k\partial_y^l \hat{F}_{Y,L,S}(y|x) - \partial_x^k\partial_y^l F_Y(y|x)| = O_P\Big(\frac{\log n}{nh_n^{2k+1}d_n^{2l}}\Big)^{1/2},$$

$$(ii)\quad \forall k+l \le 2 \qquad \sup_{y\in\mathcal{Y}_2, x\in[2h_n,1-2h_n]} |\partial_x^k\partial_y^l \hat{F}_{|e|,L,S}(y|x) - \partial_x^k\partial_y^l F_{|e|}(y|x)| = O_P\Big(\frac{\log n}{nh_n^{2k+1}d_n^{2l}}\Big)^{1/2},$$

$$(iii)\quad \forall k+l \le 2 \qquad \sup_{y\in\mathcal{Y}_1, x\in[3h_n,1-3h_n]} |\partial_x^k\partial_y^l \hat{F}^*_{Y,L,S}(y|x) - \partial_x^k\partial_y^l F_Y(y|x)| = O_P\Big(\frac{\log n}{nh_n^{2k+1}d_n^{2l}}\Big)^{1/2},$$

$$(iv)\quad \forall k+l \le 2 \qquad \sup_{y\in\mathcal{Y}_2, x\in[4h_n,1-4h_n]} |\partial_x^k\partial_y^l \hat{F}^*_{|e|,L,S}(y|x) - \partial_x^k\partial_y^l F_{|e|}(y|x)| = O_P\Big(\frac{\log n}{nh_n^{2k+1}d_n^{2l}}\Big)^{1/2}.$$

Proof of Lemma 3

We will only provide the arguments for (iv) and (iv)' since all other assertions can be derived analogously. Since $\mathcal{Y}_2$ is bounded away from zero, and since $d_n \to 0$, the fact that $\omega = \Omega'$ is symmetric and has support $[-1, 1]$ implies that for n sufficiently large

$$\Omega\Big(\frac{y-|z|}{d_n}\Big) = \Omega\Big(\frac{y-z}{d_n}\Big) - \Omega\Big(\frac{-y-z}{d_n}\Big) \quad \forall y \in \mathcal{Y}_2,\ z \in \mathbb{R}.$$

Thus we find that for n sufficiently large

$$\begin{aligned}
\hat{F}^*_{|e|}(y|x) &= \hat{F}^*_e(y|x) - \hat{F}^*_e(-y|x),\\
\hat{F}^*_{|e|,L,S}(y|x) &= \hat{F}^*_{e,L,S}(y|x) - \hat{F}^*_{e,L,S}(-y|x),\\
T^*_{|e|,n,0,L,S}(x,y) &= T^*_{e,n,0,L,S}(x,y) - T^*_{e,n,0,L,S}(x,-y),\\
\tilde{T}^*_{|e|,n,0,L,S} &= \tilde{T}^*_{e,n,0,L,S}(x,y) - \tilde{T}^*_{e,n,0,L,S}(x,-y),
\end{aligned}$$

where

$$\begin{aligned}
\hat{F}^*_e(y|x) &:= \sum_i W_i(x)\Omega\Big(\frac{y-(Y_i^* - \hat{q}^*_\tau(X_i))}{d_n}\Big),\\
\hat{F}^*_{e,L,S}(y|x) &:= F_e(y|x) + u_1^t\mathcal{M}(K)^{-1}\Big(T^*_{e,n,0,L,S}(x,y),\dots,T^*_{e,n,p,L,S}(x,y)\Big)^t,\\
T^*_{e,n,0,L,S}(x,y) &:= \frac{1}{nh_n}\sum_{i=1}^n \frac{1}{f_X(X_i)}K_{h_n,k}(x-X_i)\Big(\Omega\Big(\frac{y-(Y_i^*-\hat{q}^*_{\tau,L}(X_i))}{d_n}\Big) - F_e(y|X_i)\Big),\\
\tilde{T}^*_{e,n,0,L,S} &:= \frac{1}{nh_n}\sum_{i=1}^n \frac{1}{f_X(x)}K_{h_n,k}(x-X_i)\Big(\Omega\Big(\frac{y-(Y_i^*-\hat{q}^*_{\tau,L}(X_i))}{d_n}\Big) - F_e(y|X_i)\Big).
\end{aligned}$$

It thus suffices to establish, uniformly in $y \in \mathcal{Y} := \mathcal{Y}_2 \cup (-\mathcal{Y}_2), x \in [4h_n, 1-4h_n]$,

$$\hat{F}^*_e(y|x) = F_e(y|x) + u_1^t\mathcal{M}(K)^{-1}\Big(T^*_{e,n,0,L,S}(x,y),\dots,T^*_{e,n,p,L,S}(x,y)\Big)^t + o_P(n^{-1/2}), \tag{S3.16}$$

$$T^*_{e,n,0,L,S} = \tilde{T}^*_{e,n,0,L,S} + o_P(n^{-1/2}), \tag{S3.17}$$

$$\sup_{y\in\mathcal{Y}_2, x\in[4h_n,1-4h_n]} |\partial_x^k\partial_y^l \hat{F}^*_{e,L,S}(y|x) - \partial_x^k\partial_y^l F_e(y|x)| = O_P\Big(\frac{\log n}{nh_n^{2k+1}d_n^{2l}}\Big)^{1/2}. \tag{S3.18}$$

Define the quantities

$$T^*_{e,n,k,L}(x,y) := \frac{1}{nh_n}\sum_{i=1}^n \frac{1}{f_X(X_i)}K_{h_n,k}(x-X_i)\Big(I\{Y_i^* \le y + \hat{q}^*_{\tau,L}(X_i)\} - F_e(y|X_i)\Big),$$

$$\tilde{T}^*_{e,n,k,L}(x,y) := \frac{1}{nh_n}\sum_{i=1}^n \frac{1}{f_X(x)}K_{h_n,k}(x-X_i)\Big(I\{Y_i^* \le y + \hat{q}^*_{\tau,L}(X_i)\} - F_e(y|X_i)\Big),$$

and note that,uniformly in $y \in \mathcal{Y}, x \in [4h_n, 1-4h_n]$,

$$(T^*_{e,n,k,L}(x,\cdot) * \frac{1}{d_n}\omega(\cdot/d_n))(y) = T^*_{e,n,k,L,S}(x,y) + o(1/\sqrt{n}), \tag{S3.19}$$

$$(\tilde{T}^*_{e,n,k,L}(x,\cdot) * \frac{1}{d_n}\omega(\cdot/d_n))(y) = \tilde{T}^*_{e,n,k,L,S}(x,y) + o(1/\sqrt{n}). \tag{S3.20}$$

Also, let

$$\begin{aligned}\hat{F}^*_{e,U}(y|x) &:= \sum_{i=1}^n W_i(x) I\{Y_i^* - \hat{q}^*_\tau(X_i) \le y\} \\ &= \frac{1}{nh_n}u_1^t(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\begin{pmatrix}\sum_i K_{h_n,0}(x-X_i)I\{Y_i^* - \hat{q}^*_\tau(X_i) \le y\} \\ \vdots \\ \sum_i h_n^p K_{h_n,p}(x-X_i)I\{Y_i^* - \hat{q}^*_\tau(X_i) \le y\}\end{pmatrix}, \\ \hat{F}^*_{e,L,U}(y|x) &:= F_e(y|x) + u_1^t\mathcal{M}(K)^{-1}\Big(T^*_{e,n,0,L}(x,y),\dots,T^*_{e,n,p,L}(x,y)\Big)^t\end{aligned}$$

where the weights $W_i(x)$ are the same as in equation (2.3). At the end of the proof, we will establish the following assertions uniformly in $y \in \mathcal{Y}, x \in [4h_n, 1-4h_n]$

$$T^*_{e,n,0,L}(x,y) = \tilde{T}^*_{e,n,0,L}(x,y) + o_P(n^{-1/2}). \tag{S3.21}$$

$$\hat{F}^*_{e,U}(y|x) = \hat{F}^*_{e,L,U}(y|x) + o_P(n^{-1/2}), \tag{S3.22}$$

$$\partial_x^m T^*_{e,n,k,L}(x,y) = O_P\Big(\frac{\log n}{nh_n^{2m+1}}\Big)^{1/2}, \quad m = 0,1,2. \tag{S3.23}$$

Now assertions (S3.16), (S3.17) follows from (S3.19), (S3.22) and (S3.21) since

$$(\hat{F}^*_{e,U}(\cdot|x) * \frac{1}{d_n}\omega(\cdot/d_n))(y) = \hat{F}^*_e(y|x),$$

$$(F_e(\cdot|x) * \frac{1}{d_n}\omega(\cdot/d_n))(y) = F_e(y|x) + O(d_n^{p_\omega}) = F_e(y|x) + o(1/\sqrt{n}),$$

uniformly in $x \in [4h_n, 1-4h_n], y \in \mathcal{Y}$.
On the other hand we have

$$\partial_x^m \hat{F}^*_{e,L,U}(y|x) := \partial_x^m F_e(y|x) + u_1^t\mathcal{M}(K)^{-1}\Big(\partial_x^m T^*_{e,n,0,L}(x,y),\dots,\partial_x^m T^*_{e,n,p,L}(x,y)\Big)^t,$$

and thus (S3.23) implies, uniformly in $y \in \mathcal{Y}, x \in [4h_n, 1-4h_n]$,

$$\partial_x^m \hat{F}^*_{e,L,U}(y|x) = \partial_x^m F_e(y|x) + O_P\Big(\frac{\log n}{nh_n^{2m+1}}\Big)^{1/2}.$$

This entails (S3.18) since

$$\begin{aligned}\partial_x^k \partial_y^l \Big(\hat{F}^*_{e,L,S}(y|x) - F_e(y|x)\Big) \;=\; & \frac{1}{d_n^l}\Big[\Big(\partial_x^k \hat{F}^*_{e,L,U}(\cdot|x) - \partial_x^k F_e(\cdot|x)\Big) * \Big(\frac{1}{d_n}\omega^{(l)}\Big(\frac{\cdot}{d_n}\Big)\Big)\Big](y) \\ & + \Big((\partial_x^k \partial_y^l F_e(\cdot|x)) * \Big(\frac{1}{d_n}\omega\Big(\frac{\cdot}{d_n}\Big)\Big)\Big)(y) - \partial_x^k \partial_y^l F_e(y|x).\end{aligned}$$

Now, since by assumption $\partial_x^k F_e(y|x)$ is r times continuously differentiable with respect to y, the second summand is of order $d_n^{r-l} = O\Big(\frac{\log n}{nh_n^{2k+1}d_n^{2l}}\Big)^{1/2}$. The first summand can be bounded by $\frac{1}{d_n^l}O_P\Big(\frac{\log n}{nh_n^{2k+1}}\Big)^{1/2}$.

The proof will thus be complete after we establish (S3.21)-(S3.23). In order to do so, observe that there exists a set D_n such that the probability of D_n tends to one and such that on D_n we have, for any sequence c_n such that $c_n/r_n \to \infty$ [this is a consequence of (S2.1) and the uniform rates of convergence for $\hat{s}_L, \hat{q}_{\tau,L}, \hat{q}^*_{\tau,L}$ which follow from parts (i)-(iii) of Lemma 2 and Lemma 1]

$$\begin{aligned}&\Big|\tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big) - F_\varepsilon\Big(\frac{y}{s(X_i)}\Big)\Big| \\ &\le \sup_{y \in (1+\mathcal{Y}/c_s)} |\tilde{F}_\varepsilon(y) - F_\varepsilon(y)| + 0.5c_n \sup_{y \in (1+\mathcal{Y}/c_s)} |yf_\varepsilon(y)| \le c_n\end{aligned}$$

where the last bound follows from (S3.27) in Lemma 5. In particular, on D_n we have

$$I\Big\{U_i \le F_\varepsilon\Big(\frac{y}{s(X_i)}\Big) - c_n\Big\} \;\le\; I\Big\{U_i \le \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)\Big\} \qquad \text{(S3.24)}$$

$$\le\; I\Big\{U_i \le F_\varepsilon\Big(\frac{y}{s(X_i)}\Big) + c_n\Big\}. \qquad \text{(S3.25)}$$

Proof of (S3.21)

Recall that $Y_i^* = \hat{q}_\tau(X_i) + \hat{s}(X_i)\varepsilon_i^*$ and $\varepsilon_i^* = \tilde{F}_\varepsilon^{-1}(U_i)$. Observe the identity

$$I\{Y_i^* \le y + \hat{q}^*_{\tau,L}(X_i)\} = I\Big\{U_i \le \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}(X_i)}{\hat{s}(X_i)}\Big)\Big\}.$$

Moreover, a Taylor expansion shows that, with probability tending to one,

$$\begin{aligned}&\Big|I\Big\{U_i \le \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}(X_i)}{\hat{s}(X_i)}\Big)\Big\} - I\Big\{U_i \le \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)\Big\}\Big| \\ &\le I\Big\{\Big|U_i - \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)\Big| \le C\gamma_n \sup_{y \in 2\mathcal{Y}/c_s} |y\tilde{f}_\varepsilon(y)|\Big\}\end{aligned}$$

where $\gamma_n = o(1/\sqrt{n})$, and thus arguments similar to those in the proof of Lemma 9 yield

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \frac{K_{h_n,k}(x-u)}{h_n}\Big(\frac{1}{f_X(u)}-\frac{1}{f_X(x)}\Big)\Big(I\Big\{U_i \le \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}(X_i)}+\frac{\hat{q}^*_{\tau,L}(X_i)-\hat{q}(X_i)}{\hat{s}(X_i)}\Big)\Big\}\Big)\\
=&\frac{1}{n}\sum_{i=1}^n \frac{K_{h_n,k}(x-u)}{h_n}\Big(\frac{1}{f_X(u)}-\frac{1}{f_X(x)}\Big)\Big(I\Big\{U_i \le \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}+\frac{\hat{q}^*_{\tau,L}(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)\Big\}\Big)\\
&+o_P(1/\sqrt{n}).
\end{aligned}$$

Next, observe that by part (i)-(iii) of Lemma 2, Lemma 1 and by (S2.1) there exists a set D_n whose probability tends to one such that on D_n we have for some $\delta > 0$

$$\hat{s}_L \in \tilde{C}_C^{1+\delta}([3h_n, 1-3h_n]), \quad \hat{q}^*_{\tau,L}, \hat{q}_{\tau,L} \in C_C^{1+\delta}([3h_n, 1-3h_n]),$$

$$\sup_{u\in[3h_n,1-3h_n],y\in\mathcal{Y}}\Big|\tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(u)}+\frac{\hat{q}^*_{\tau,L}(u)-\hat{q}_{\tau,L}(u)}{\hat{s}_L(u)}\Big)-F_e(y|u)\Big| \le r_n h_n^{-1/4},$$

and $\tilde{F}_\varepsilon \in \mathcal{D}$ defined in (S2.2). Additionally, (S2.3) and the arguments from Proposition 3 in Neumeyer (2009) show that for the class of functions

$$\begin{aligned}
\mathcal{G}_n := \Big\{&(u,v)\mapsto I\Big\{u \le F\Big(\frac{y}{a_1(v)}+\frac{a_2(v)}{a_1(v)}\Big)\Big\}\\
&\Big|\ y\in\mathcal{Y}, F\in\mathcal{D}, a_1\in\tilde{C}_C^{1+\delta}([3h_n,1-3h_n]), a_2\in C_C^{1+\delta}([3h_n,1-3h_n])\Big\}
\end{aligned}$$

we have, denoting by P the product measure of the uniform random variable U_1 and the covariate X_1, $\sup_n \log N_{[\,]}(\varepsilon, \mathcal{G}_n, L^2(P)) \le C\varepsilon^{-2\alpha}$ for some $\alpha < 1$. Next, define the class of functions

$$\begin{aligned}
\mathcal{F}_n := \Big\{&(u,v)\mapsto \frac{K_{h_n,k}(x-u)}{h_n}\Big(\frac{1}{f_X(u)}-\frac{1}{f_X(x)}\Big)\times\\
&\times\Big(I\Big\{v \le \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(u)}+\frac{\hat{q}^*_{\tau,L}(u)-\hat{q}_{\tau,L}(u)}{\hat{s}_L(u)}\Big)\Big\}-F_e(y|u)\Big)\Big|x\in[4h_n,1-4h_n], y\in\mathcal{Y}\Big\}.
\end{aligned}$$

In particular, observe that, due to the continuous differentiability of f_X and the compact support of K, the functions in $\mathcal{F}_n$ are bounded uniformly over n. Additionally, combining the bound on $\sup_n \log N_{[\,]}(\varepsilon, \mathcal{G}_n, L^2(P))$ with parts 1, 3 and 4 of Lemma 8, we find that on D_n

$$\sup_n \log N_{[\,]}(\varepsilon, \mathcal{F}_n, L^2(P)) \le \tilde{C}\varepsilon^{-2\tilde{\alpha}}$$

for some $\tilde{\alpha} < 1$ and finite $\tilde{C}$. Moreover, again on D_n, we find that for each $f \in \mathcal{F}_n$

$$\mathbb{E}f(X_i,U_i) = O(h_n^{3/4}r_n) = o(1/\sqrt{n}), \quad \mathbb{E}f^2(X_i,U_i) = O(h_n).$$

To see the second statement, observe that every $f \in \mathcal{F}_n$ satisfies

$$|f(X_i, U_i)| \leq 2\Big|\frac{K_{h_n,k}(x - X_i)}{h_n}\Big(\frac{1}{f_X(X_i)} - \frac{1}{f_X(x)}\Big)\Big|,$$

the assertion now follows from a Taylor expansion of f_X. For the bound on $\mathbb{E}f(X_i, U_i)$, observe that

$$|\mathbb{E}[f(X_i, U_i)]| \leq \int \Big|\frac{K_{h_n,k}(x - u)}{h_n}\Big(\frac{1}{f_X(u)} - \frac{1}{f_X(x)}\Big)\Big| r_n h_n^{-1/4} f_X(u) du,$$

the claimed bound now follows from a Taylor expansion of $1/f_X(u)$ around x. Thus by Lemma 7 $\sup_{f\in\mathcal{F}_n} |\sum_i f(X_i, U_i)| = o_P(1/\sqrt{n})$ and (S3.21) follows.

Proof of (S3.22) Define $\mathcal{H} := \mathrm{diag}(1, h_n, ..., h_n^p)$ and observe that by (S3.21) we have uniformly in $x \in [4h_n, 1 - 4h_n], y \in \mathcal{Y}$

$$\begin{aligned}
\hat{F}^*_{e,U}(y|x) - \hat{F}^*_{e,L,U}(y|x) &= \sum_i W_i(x)(I\{Y_i^* - \hat{q}^*_\tau(X_i) \leq y\} - I\{Y_i^* - \hat{q}^*_{\tau,L}(X_i) \leq y\}) \\
&\quad + \frac{u_1^t(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}}{nh_n}\mathcal{H}\begin{pmatrix} \sum_i K_{h_n,0}(x - X_i)F_e(y|X_i) \\ \vdots \\ \sum_i K_{h_n,p}(x - X_i)F_e(y|X_i)\end{pmatrix} - F_e(y|x) \\
&\quad + \Big(u_1^t(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathcal{H} - \frac{u_1^t\mathcal{M}(K)^{-1}}{f_X(x)}\Big)\begin{pmatrix} f_X(x)\tilde{T}^*_{e,n,0,L}(x,y) \\ \vdots \\ f_X(x)\tilde{T}^*_{e,n,p,L}(x,y)\end{pmatrix} \\
&\quad + u_1^t\mathcal{M}(K)^{-1}\begin{pmatrix} \tilde{T}^*_{e,n,0,L}(x,y) - T^*_{e,n,0,L}(x,y) \\ \vdots \\ \tilde{T}^*_{e,n,p,L}(x,y) - T^*_{e,n,p,L}(x,y)\end{pmatrix} \\
&=: R_{n,1}(x,y) + R_{n,2}(x,y) + R_{n,3}(x,y) + R_{n,4}(x,y).
\end{aligned}$$

Note that a Taylor expansion of $F_e(y|X_i)$ with respect to X_i around the point x combined with the fact that

$$\frac{1}{nh_n}u_1^t(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\begin{pmatrix} h_n^k\sum_i K_{h_n,k}(x - X_i) \\ \vdots \\ h_n^{p+k}\sum_i K_{h_n,p+k}(x - X_i)\end{pmatrix} = I\{k = 0\}$$

for $k = 0, ..., p$ yields the representation

$$\frac{u_1^t(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}}{nh_n}\begin{pmatrix} \sum_i K_{h_n,0}(x - X_i)F_e(y|X_i) \\ \vdots \\ \sum_i h_n^p K_{h_n,p}(x - X_i)F_e(y|X_i)\end{pmatrix} = F_e(y|x) + O_P(h_n^{p+1}) = F_e(y|x) + o_P(n^{-1/2})$$

uniformly in $x \in [4h_n, 1-4h_n], y \in \mathcal{Y}$, so that $R_{n,2}$ is small.

Next, consider $R_{n,3}$. By Lemma 4 and observing that $u_1^t \mathcal{H}^{-1} = u_1^t$ we find

$$\Big(\frac{u_1^t(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}}{nh_n}\mathcal{H} - \frac{u_1^t\mathcal{M}(K)^{-1}}{f_X(x)}\Big) = O_P(h_n),$$

and together with the fact that

$$\sup_{x[4h_n,1-4h_n],y\in\mathcal{Y}} \sup_{k=0,...,p} |T^*_{e,n,k,L}(x,y)| = O_P\Big(\frac{\log n}{nh_n}\Big)^{1/2}$$

which follows by similar arguments as the proof of (S3.23), this shows that $R_{n,3}$ is small.

The negligibility of $R_{n,4}$ follows from (S3.21).

Finally, consider $R_{n,1}$. Observe that, by similar arguments as in the proof of (S3.21), there exists a deterministic sequence $\xi_n = o(n^{-1/2})$ such that, with probability tending to one, we have for any $X_i \in [3h_n, 1-3h_n]$

$$\Big|I\{Y_i^* - \hat{q}_\tau^*(X_i) \le y\} - I\{Y_i^* - \hat{q}^*_{\tau,L}(X_i) \le y\}\Big| \le I\Big\{\Big|U_i - \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)\Big| \le \xi_n\Big\}.$$

Now arguments similar to those in the proof of Lemma 9 yield for every $k = 0, ..., p$

$$d_{n,k} := \frac{1}{n}\sum_{i=1}^n \frac{|K_{h_n,k}(x-u)|}{h_n}\frac{1}{f_X(x)} I\Big\{\Big|U_i - \tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)} + \frac{\hat{q}^*_{\tau,L}(X_i) - \hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)\Big| \le \xi_n\Big\} = o_P(n^{-1/2})$$

uniformly over $x \in [4h_n, 1-4h_n], y \in \mathcal{Y}$. Moreover, by Lemma 4 we have

$$|R_{n,1}(x,y)| \le (p+1)\Big(\max_{k=0,...,p}(u_1^t(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathcal{H})_k\Big)\Big(\max_{k=0,...,p}|d_{n,k}(x,y)|\Big)$$

This shows that $R_{n,1}$ is negligible and completes the proof of (S3.22).

Proof of (S3.23) Consider the decomposition

$$\partial_x^m T^*_{e,n,k,L}(x,y) = A^+_{n,k,m}(x,y) + A^-_{n,k,m}(x,y)$$

where

$$A^+_{n,k,m}(x,y) := \frac{1}{nh_n}\frac{1}{h_n^m}\sum_{i=1}^n \frac{K^{(m)}_{h_n,k}(x-X_i)}{f_X(X_i)} I\Big\{K^{(m)}_{h_n,k}(x-X_i) > 0\Big\}\Big(I\{Y_i^* \le y + \hat{q}^*_{\tau,L}(X_i)\} - F_e(y|X_i)\Big)$$

and $A^-_{n,k,m}$ is defined analogously. On the set D_n (defined in the beginning of this proof) we have

$$A^+_{n,k,m}(x,y) \;\le\; \frac{1}{nh_n^{m+1}}\sum_{i=1}^n \frac{K^{(m)}_{h_n,k}(x-X_i)}{f_X(X_i)} I\Big\{K^{(m)}_{h_n,k}(x-X_i) > 0\Big\} \times$$

$$\times\Big(I\Big\{U_i \le F_\varepsilon\Big(\frac{y}{s(X_i)}\Big) + c_n\Big\} - F_\varepsilon\Big(\frac{y}{s(X_i)}\Big)\Big)$$

$$=: \frac{1}{nh_n^{m+1}} \sum_{i=1}^n g_{x,y}^{(n,m,+)}(X_i, U_i, c_n).$$

The expectation of each summand $g_{x,y}^{(n,m,+)}(X_i, U_i, c_n)$ in the above sum is of the order $O(h_n c_n)$. Moreover, the class of functions

$$\Big\{(u,v) \mapsto g_{x,y}^{(n,m,+)}(u,v,c_n)\Big| x \in [4h_n, 1-4h_n], y \in \mathcal{Y}\Big\}$$

is with probability tending to one contained in a class that satisfies the assumptions of part 2 of Lemma 7 with $\delta_n = h_n$, this follows from a combination of assumption (**K2**) with parts 1,2,4,6 of Lemma 8 where part 6 is applied with the class of functions $\mathcal{G} := \{v \mapsto F_\varepsilon(y/s(v)) + z | y \in \mathcal{Y}, z \in [0,1]\}$. This yields the bound

$$\frac{1}{nh_n^{m+1}} \sum_{i=1}^n g_{x,y}^{(n,m,+)}(X_i, U_i, c_n) = o\Big(\frac{c_n h_n}{h_n^{m+1}}\Big) + O_P\Big(\frac{\log n}{nh_n^{2m+1}}\Big)^{1/2}$$

uniformly in $x \in [4h_n, 1-4h_n], y \in \mathcal{Y}$. Since c_n/r_n can tend to infinity arbitrarily slowly, the above result implies

$$\frac{1}{nh_n^{m+1}} \sum_{i=1}^n g_{x,y}^{(n,m,+)}(X_i, U_i, c_n) = O_P\Big(\frac{\log n}{nh_n^{2m+1}}\Big)^{1/2}.$$

Summarizing, we have obtained the bound $A^+_{n,k,m}(x,y) \le O_P\Big(\frac{\log n}{nh_n^{2m+1}}\Big)^{1/2}$, and a corresponding lower bound can be obtained by similar arguments. Analogous reasoning yields a bound for $A^-_{n,k,m}(x,y)$ and altogether this implies (S3.23).
Thus we have established (S3.21)-(S3.23) and the proof of the Lemma is complete. □

Lemma 4 *Under assumptions (**K1**) and (**A1**) if additionally* $(nh_n)^{-1} = o(h_n\sqrt{\log n})$ *we have the decomposition (holding uniformly in* $x \in [h_n, 1-h_n]$*)*

$$nh_n(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1} = \frac{1}{f_X(x)}\mathcal{H}^{-1}\mathcal{M}(K)^{-1}\mathcal{H}^{-1} + \mathcal{H}^{-1}1_{(p+1)\times(p+1)}O_P(h)\mathcal{H}^{-1}$$

where $\mathcal{H} = diag(1, h_n, ..., h_n^p)$*, and* $1_{(p+1)\times(p+1)}$ *is a matrix with 1 in every entry.*

Proof The elements of the matrix $\mathbf{X}^t\mathbf{W}\mathbf{X}$ are of the form

$$\frac{1}{nh_n}(\mathbf{X}^t\mathbf{W}\mathbf{X})_{k,l} = \frac{1}{nh_n}\sum_i K_{h_n,0}(x-X_i)(x-X_i)^m = \frac{h_n^m}{nh_n^d}\sum_i K_{h_n,m}(x-X_i)$$

where $m = k+l-2$. In particular, continuous differentiability of f_X together with an application of Lemma 7 and Lemma 8 implies that

$$\frac{1}{nh_n}\sum_i K_{h_n,k}(x-X_i) = \mu_k f_X(x) + O_P(\Big(\frac{\log n}{nh_n}\Big)^{1/2} + h_n)$$

uniformly in x. Thus we obtain a representation of the form

$$\frac{1}{nh_n}\mathbf{X}^t\mathbf{W}\mathbf{X} = \mathcal{H}\Big(\mathcal{M}(K)f_X(x) + 1_{N\times N}O_P(h_n)\Big)\mathcal{H}$$

where $M_0 = \mathcal{M}(K)$ is invertible and $\mathcal{H}$ is a diagonal matrix with entries $1, h_n, ..., h_n^p$. Thus for h_n sufficiently small an application of the Neumann series yields the assertion with probability tending to one. □

S3.2 Additional technical results

Lemma 5 *Let $n\alpha_n^4 = o(1)$ and assume that the conditions of (i), (i)', (ii), (ii)' of Lemma 3 hold. Then for any bounded $\mathcal{Y} \subset \mathbb{R}$ and any $\delta_n \to 0$ we have*

$$\sup_{a,b\in\mathcal{Y},|a-b|\le\delta_n} \Big|\tilde{F}_\varepsilon(a) - \tilde{F}_\varepsilon(b) - \Big(\bar{F}_\varepsilon(a) - \bar{F}_\varepsilon(b)\Big)\Big| = o_P(1/\sqrt{n}), \quad \text{(S3.26)}$$

$$\sup_{y\in\mathcal{Y}}\Big|\tilde{F}_\varepsilon(y) - F_\varepsilon(y)\Big| = O_P\Big(\Big(\frac{\log n}{nh_n}\Big)^{1/2}\Big), \quad \text{(S3.27)}$$

where

$$\bar{F}_\varepsilon(a) := \frac{\sum_k I_{[2h_n,1-2h_n]}(X_k)F_Y(\hat{q}_{\tau,L}(X_k) + a\hat{s}_L(X_k)|X_k)}{\sum_l I_{[2h_n,1-2h_n]}(X_l)}.$$

Proof of Lemma 5 Recalling the definition of $\tilde{F}_\varepsilon$, it is easy to see that $\tilde{F}_\varepsilon(y) = \frac{1}{\alpha_n}\Big(\hat{F}_\varepsilon(\cdot) * \phi(\cdot/\alpha_n)\Big)(y)$ where

$$\hat{F}_\varepsilon(y) := \frac{\sum_k I_{[2h_n,1-2h_n]}(X_k)I\{Y_k - \hat{q}(X_k) \le y\hat{s}(X_k)\}}{\sum_l I_{[2h_n,1-2h_n]}(X_l)}.$$

Standard calculations show that

$$\frac{1}{\alpha_n}\Big(\bar{F}_\varepsilon(\cdot) * \phi(\cdot/\alpha_n)\Big)(y) = \bar{F}_\varepsilon(y) + o_P(1/\sqrt{n})$$

uniformly in $y \in \mathcal{Y}$. Thus it suffices to establish that, for any bounded $\tilde{\mathcal{Y}}$

$$\sup_{a,b\in\tilde{\mathcal{Y}},\ |a-b|\le\delta_n} \Big|\hat{F}_\varepsilon(a) - \hat{F}_\varepsilon(b) - \Big(\bar{F}_\varepsilon(a) - \bar{F}_\varepsilon(b)\Big)\Big| = o_P(1/\sqrt{n}) \quad \text{(S3.28)}$$

$$\sup_{y\in\tilde{\mathcal{Y}}}\Big|\hat{F}_\varepsilon(y) - F_\varepsilon(y)\Big| = O_P\Big(\Big(\frac{\log n}{nh_n}\Big)^{1/2}\Big). \quad \text{(S3.29)}$$

To simplify the notation, write $\mathcal{Y}$ for $\tilde{\mathcal{Y}}$.

Proof of (S3.28) Since $\frac{1}{n}\sum_l I_{[2h_n,1-2h_n]}(X_l) = 1 + o_P(1)$, we only need to consider the enumerator. Since $\mathcal{Y}$ is bounded we have, with probability tending to one, uniformly in $y \in \mathcal{Y}$

$$\Big|I\{Y_k - \hat{q}_\tau(X_k) \le y\hat{s}(X_k)\} - I\{Y_k - \hat{q}_{\tau,L}(X_k) \le y\hat{s}_L(X_k)\}\Big| \\ \le I\{Y_k - \hat{q}_{\tau,L}(X_k) - y\hat{s}_L(X_k) \le \gamma_n\} - I\{Y_k - \hat{q}_{\tau,L}(X_k) - y\hat{s}_L(X_k) \le -\gamma_n\}$$

for some $\gamma_n = o(1/\sqrt{n})$. Moreover an application of parts 1 and 6 of Lemma 8 combined with Theorem 2.7.1 in van der Vaart, Wellner (1996) shows that the functions

$$(u,v) \mapsto I\{v - \hat{q}_{\tau,L}(u) - y\hat{s}_L(u) \le \gamma_n\} - I\{v - \hat{q}_{\tau,L}(u) - y\hat{s}_L(u) \le -\gamma_n\}$$

are, with probability tending to one, contained in a class of functions satisfying the assumptions of the first part of Lemma 7 with the additional property that each element has expectation of order $o(1/\sqrt{n})$. Combined with parts 1 and 4 of Lemma 8, this implies

$$\sup_{y\in\mathcal{Y}}\Big|\sum_k I_{[2h_n,1-2h_n]}(X_k)\Big(I\{Y_k - \hat{q}_\tau(X_k) \le y\hat{s}(X_k)\} - I\{Y_k - \hat{q}_{\tau,L}(X_k) \le y\hat{s}_L(X_k)\}\Big)\Big| = o_P(1/\sqrt{n}),$$

and thus it remains to consider

$$\sup_{a,b\in\mathcal{Y},|a-b|\le\delta_n} \frac{1}{n}\sum_k I_{[2h_n,1-2h_n]}(X_i)\Big(I\{Y_k \le \hat{q}_{\tau,L}(X_k) + a\hat{s}_L(X_k)\} - I\{Y_k \le \hat{q}_{\tau,L}(X_k) + b\hat{s}_L(X_k)\} \\ - F_Y(\hat{q}_{\tau,L}(X_k) + a\hat{s}_L(X_k)|X_k) + F_Y(\hat{q}_{\tau,L}(X_k) + b\hat{s}_L(X_k)|X_k)\Big)$$

By arguments similar to those given above, it is easily seen that this quantity is of order $o_P(1/\sqrt{n})$ if one notes that the smoothness assumptions on F_Y imply that with $\hat{q}_{\tau,L}, \hat{s}_L \in C_C^{1+\delta}$ with probability tending to one the same holds for the function $u \mapsto F_Y(\hat{q}_{\tau,L}(u) + y\hat{s}_L(u)|u)$ uniformly in $y \in \mathcal{Y}$. This completes the proof of (S3.28).

Proof of (S3.29) Write

$$\hat{F}_\varepsilon(y) - F_\varepsilon(y) = \frac{n^{-1}\sum_k I_{[2h_n,1-2h_n]}(X_k)\Big(I\{Y_k - \hat{q}_\tau(X_k) \le y\hat{s}(X_k)\} - F_\varepsilon(y)\Big)}{n^{-1}\sum_l I_{[2h_n,1-2h_n]}(X_l)}.$$

Since $n^{-1}\sum_l I_{[2h_n,1-2h_n]}(X_l) = 1 + o_P(1)$, it suffices to consider the enumerator. Observe that

$$I\{Y_k - \hat{q}_\tau(X_k) \le y\hat{s}(X_k)\} = I\Big\{\varepsilon_k \le y\frac{\hat{s}(X_k)}{s(X_k)} + \frac{\hat{q}_\tau(X_k) - q_\tau(X_k)}{s(X_k)}\Big\}$$

and thus, for any $c_n/r_n \to \infty$ we have with probability tending to one, uniforly over $y \in \mathcal{Y}$

$$\Big|I\{Y_k - \hat{q}_\tau(X_k) \le y\hat{s}(X_k)\} - I\{\varepsilon_k < y\}\Big| \le I\{|\varepsilon_k - y| \le c_n\}.$$

Thus standard

$$\begin{aligned}&\sup_{y\in\mathcal{Y}}\Big|n^{-1}\sum_k I_{[2h_n,1-2h_n]}(X_k)\Big(I\{Y_k-\hat q_\tau(X_k)\le y\hat s(X_k)\}-I\{\varepsilon_k<y\}\Big)\Big|\\ \le&\ \sup_{y\in\mathcal{Y}} n^{-1}\sum_k I_{[2h_n,1-2h_n]}(X_k)I\{|\varepsilon_k-y|\le c_n\}=O_P(c_n),\end{aligned}$$

where the last equality follows by standard empirical process arguments. This shows that, uniformly in $y\in\mathcal{Y}$,

$$\hat F_\varepsilon(y)-F_\varepsilon(y)=\frac{n^{-1}\sum_k I_{[2h_n,1-2h_n]}(X_k)\Big(I\{\varepsilon_k\le y\}-F_\varepsilon(y)\Big)}{n^{-1}\sum_l I_{[2h_n,1-2h_n]}(X_l)}+O_P(c_n)=O_P(c_n).$$

Since c_n was arbitrary, this completes the proof of (S3.29) and hence also of the Lemma. □

Lemma 6 *Assume that κ is a symmetric, uniformly bounded density with support $[-1,1]$ and let $b_n=o(1)$.*
(a) If the function $F:[0,1]\to\mathbb{R}$ is strictly increasing and F^{-1} is k times continuously differentiable in a neighborhood of the point τ, we have for b_n small enough

$$H_{id,\kappa,\tau,b_n}(F)=F^{-1}(\tau)+\sum_{i=1}^k\frac{b_n^i}{i!}(F^{-1})^{(i)}(\tau)\mu_{i+1}(\kappa)+R_n(\tau)$$

with $|R_n(\tau)|\le C_k(\kappa)b_n^k\sup_{|s-\tau|\le b_n}|(F^{-1})^{(k)}(\tau)-(F^{-1})^{(k)}(s)|$, $\mu_i(\kappa):=\int u^i\kappa(u)du$ and a constant C_k depending only on k and κ. In particular, if we assume that $F:\mathbb{R}\to[0,1]$ is strictly increasing and F^{-1} is two times continuously differentiable in a neighborhood of τ and $G:\mathbb{R}\to(0,1)$ is two times continuously differentiable in a neighborhood of $F^{-1}(\tau)$ with $G'(F^{-1}(\tau))>0$ we have

$$|F^{-1}(\tau)-Q_{G,\kappa,\tau,b_n}(F)|\le Cb_n^2\sup_{|s-G\circ F^{-1}(\tau)|\le R_{n,1}}|(G^{-1})'(s)|\sup_{|s-\tau|\le b_n}|(G\circ F^{-1})''(s)|=:R_{n,2}$$

for some constant C that depends only on κ where $R_{n,1}:=Cb_n^2\sup_{|s-\tau|\le b_n}|(G\circ F^{-1})''(s)|$.

(b) Assume that κ is additionally differentiable with Lipschitz-continuous derivative and that the functions G,G^{-1} have derivatives that are uniformly bounded on any compact subset of $\mathbb{R}$ [the bound is allowed to depend on the interval]. Then for any increasing function F with uniformly bounded first derivative we have $|H(F_1)-H(F_2)|\le R_{n,3}+R_{n,4}$ and

$$|Q_{G,\kappa,\tau,b_n}(F_1)-Q_{G,\kappa,\tau,b_n}(F_2)|\le\sup_{u\in\mathcal{U}(H(F_1),H(F_2))}|(G^{-1})'(u)|(R_{n,3}+R_{n,4})$$

where C is a constant that depends only on κ, $\mathcal{U}(a,b):=[a\wedge b,a\vee b]$, and

$$R_{n,3}:=\frac{Cc_n}{b_n}\|F_1-F_2\|_\infty\sup_{|v-\tau|\le c_n}|(G\circ F^{-1})'(v)|,\quad R_{n,4}:=R_{n,3}\frac{\|F_1-F\|_\infty+\|F_1-F_2\|_\infty}{b_n}$$

with $c_n := b_n + 2\|F_1 - F_2\|_\infty + \|F_1 - F\|_\infty$.

(c) If additionally to the assumptions made in (b), the function F_1 *is two times continuously differentiable in a neighborhood of* $F^{-1}(\tau)$ *with* $F_1'(F_1^{-1}(\tau)) > 0$ *and* G *is two times continuously differentiable in a neighborhood of* $F_1^{-1}(\tau)$ *with* $G'(F^{-1}(\tau)) > 0$*, we have*

$$Q_{G,\kappa,\tau,b_n}(F_1) - Q_{G,\kappa,\tau,b_n}(F_2) \;=\; -\frac{1}{F_1'(F_1^{-1}(\tau))}\int_{-1}^{1}\kappa(v)\Big(F_2(F_1^{-1}(\tau+vb_n)) - F_1(F_1^{-1}(\tau+vb_n))\Big)dv + R_n,$$

where

$$|R_n| \;\le\; R_{n,5} + R_{n,6} + \frac{Cb_n\sup_{|s-\tau|\le b_n}(G\circ F^{-1})''(s)\|F_1-F_2\|_\infty + R_{n,4}}{G'(F_1^{-1}(\tau))}$$

with a constant C *depending only on* κ *and*

$$R_{n,5} \;:=\; \frac{1}{2}\sup_{u\in\mathcal{U}(H(F_1),H(F_2))}|(G^{-1})''(u)|(H(F_1)-H(F_2))^2$$

$$R_{n,6} \;:=\; \sup_{u\in\mathcal{U}(H(F_1),G(F_1^{-1})(\tau))}|(G^{-1})''(u)|\cdot|H(F_1)-G(F_1^{-1})(\tau)|\cdot|H(F_1)-H(F_2)|.$$

Proof See Volgushev et al. (2013).

Lemma 7 (Basic Lemma)

1. *Assume that the classes of functions* $\mathcal{F}_n$ *consist of uniformly bounded functions (with the bound, say* D*, not depending on* n*) with* $N_{[]}(\mathcal{F}_n,\varepsilon,L^2(P)) \le C\exp(-c\varepsilon^{-a})$ *for every* $\varepsilon \le \delta_n$ *for some* $a < 2$ *and constants* C, c *not depending on* n*. Then we have*

$$\sqrt{n}\sup_{f\in\mathcal{F}_n,\|f\|_{P,2}\le\delta_n}\Big(\int f dP_n - \int f dP\Big) = o_P^*(1)$$

where the * *denotes outer probability, see van der Vaart and Wellner (1996) for a more detailed discussion.*

2. *If under the assumptions of part one we have* $N_{[]}(\mathcal{F}_n,\varepsilon,L^2(P)) \le C\varepsilon^{-a}$ *for every* $\varepsilon \le \delta_n$*, some* $a > 0$ *and* C *not depending on* n*, it holds that for any* $\delta_n \sim n^{-b}$ *with* $b < 1/2$

$$\sqrt{n}\sup_{f\in\mathcal{F}_n,\|f\|_{P,2}\le\delta_n}\Big(\int f dP_n - \int f dP\Big) = O_P^*\Big(\delta_n|\log\delta_n|\Big)$$

Proof See Volgushev et al. (2013).

Lemma 8

1. *Define* $\mathcal{F}+\mathcal{G} := \{f+g|f\in\mathcal{F}, g\in\mathcal{G}\}, \mathcal{F}\mathcal{G} := \{fg|f\in\mathcal{F}, g\in\mathcal{G}\}$*. Then*
$$N_{[]}(\mathcal{F}+\mathcal{G},\varepsilon,\rho) \le N_{[]}(\mathcal{F},\varepsilon/2,\rho)N_{[]}(\mathcal{G},\varepsilon/2,\rho)$$
If additionally the classes $\mathcal{F},\mathcal{G}$ *are uniformly bounded by the constant* C*, we have*
$$N_{[]}(\mathcal{F}\mathcal{G},\varepsilon,\|.\|) \le N_{[]}^2(\mathcal{F},\varepsilon/4C,\|.\|)N_{[]}^2(\mathcal{G},\varepsilon/4C,\|.\|)$$
for any seminorm $\|.\|$ *with the additional property that* $|f_1|\le|f_2|$ *implies* $\|f_1\|\le\|f_2\|$*.*

2. *Assume that the Kernel* K *has compact support* $[-1,1]$*, that* $K_{1,k}^{(m)}$ *is uniformly bounded and Lipschitz-continuous, and that* f_X *is uniformly bounded. Then the* $L^2(P_X)$ *bracketing numbers* $N_{[]}(\mathcal{F}_n,\varepsilon,L^2(P_X))$ *of the set*
$$\mathcal{F}_n := \Big\{u\mapsto K_{h_n,k}^{(m)}(x-u)\Big|x\in[h_n,1-h_n]\Big\}$$
are bounded by $C\varepsilon^{-3}$ *for some constant* C *independent of* n*.*

3. *Assume that the Kernel* K *has compact support* $[-1,1]$*, that* K *is uniformly bounded and Lipschitz continuous, and that* f_X *is uniformly bounded away from zero on* $[0,1]$ *and Lipschitz-continuous. Then for the set of function*
$$\mathcal{F}_n := \Big\{u\mapsto \frac{1}{h_n}\Big(\frac{1}{f_X(x)}-\frac{1}{f_X(u)}\Big)K_{h_n,k}(x-u)\Big|x\in[h_n,1-h_n]\Big\}$$
we have $N_{[]}(\mathcal{F}_n,\varepsilon,L^2(P))\le C\varepsilon^{-5}$ *for some constant* C *independent of* n*.*

4. *For any measure* P *on the unit interval with uniformly bounded density* f*, the class of functions*
$$\mathcal{F} := \Big\{u\mapsto I\{u\le s\}\Big|s\in[0,1]\Big\}\cup\Big\{u\mapsto I\{u<s\}\Big|s\in[0,1]\Big\}$$
can be covered by $C\varepsilon^{-(2)}$ *brackets of* $L^2(P)$ *length* ε*.*

5. *Consider the class of distribution functions* $\mathcal{F} := \Big\{u\mapsto F(y|u)\Big|y\in\mathbb{R}\Big\}$ *with densities* $f(y|u)$ *and assume that* $\sup_{u,y}|y|^\alpha(F(y|u)\wedge(1-F(y|u))\le D$ *for some* $\alpha>0$ *and additionally* $\sup_{u,y}f(y|u)\le D$*. Then we have* $N_{[]}(\mathcal{F},\varepsilon,\|\ \|_\infty)\le C\varepsilon^{-\frac{\alpha+1}{\alpha}}$ *for some constant* C *independent of* α*.*

6. *For any measure* P *on* $\mathbb{R}\times\mathbb{R}^k$ *with uniformly bounded conditional density* $f_{V|U}$ *the class of functions*
$$\mathcal{G} := \Big\{(u,v)\mapsto I\{v\le f(u)\}\Big|f\in\mathcal{F}\Big\}$$
satisfies $N_{[]}(\mathcal{G},\varepsilon,\|.\|_{P,2})\le N_{[]}(\mathcal{F},C\varepsilon^2,\|.\|_\infty)$ *for some constant* C *independent of* ε*.*

Proof

Part 1 The first assertion is obvious from the definition of bracketing numbers. For the second assertion, note that $\mathcal{F}\mathcal{G} = (\mathcal{F}+C)(\mathcal{G}+C) - C\mathcal{F} - C\mathcal{G} + C^2$. Moreover, all elements of the classes $\mathcal{F}+C, \mathcal{G}+C$ are by construction non-negative and thus it also is possible to cover them with brackets consisting of non-negative functions and amounts equal to the brackets of $\mathcal{F}, \mathcal{G}$, respectively. Finally, observe that if $0 \leq f_l \leq f \leq f_u$ and $0 \leq g_l \leq g \leq g_u$, we also have $f_l g_l \leq fg \leq f_u g_u$. Moreover $\|f_l g_l - f_u g_u\| \leq C\|f_u - f_l\| + C\|g_u - g_l\|$. Thus the class $(\mathcal{F}+C)(\mathcal{G}+C)$ can be covered by at most $\leq N_{[]}(\mathcal{F}, \varepsilon, \|.\|)N_{[]}(\mathcal{G}, \varepsilon, \|.\|)$ brackets of length $2C\varepsilon$. Finding brackets for the classes $C\mathcal{F}, C\mathcal{G}$ is trivial, and applying the first assertion of the Lemma completes the proof.

Part 2+3 Without loss of generality, assume that $h = h_n < 1$. Note that the class of functions $\mathcal{F}_n$ from part 2 can be represented as $\mathcal{F}_n = \{u \mapsto g_x(u) | x \in [h_n, 1-h_n]\}$ where the functions g_x satisfy $\sup_{x\in[h_n,1-h_n]} \|g_x\|_\infty \leq C$, $\sup_{u\in\mathbb{R}} |g_x(u) - g_y(u)| \leq \tilde{C}|x-y|h_n^{-1}$ for some constants $C, \tilde{C}$ independent of n, x, y. To see the latter inequality, observe that by assumption $u \mapsto K_{1,k}^{(m)}(u)$ is uniformly bounded and Lipschitz continuous. Additionally, the support of the functions g_x is contained in $[x-h_n, x+h_n]$.

Similarly, $\mathcal{F}_n$ from part 3 can be represented as $\mathcal{F}_n = \{u \mapsto g_x(u) | x \in [h_n, 1-h_n]\}$ where the functions g_x satisfy $\sup_{x\in[h_n,1-h_n]} \|g_x\|_\infty \leq C$, $\sup_{u\in\mathbb{R}} |g_x(u) - g_y(u)| \leq \tilde{C}|x-y|h_n^{-2}$ for some constants $C, \tilde{C}$ independent of n, x, y (and possibly different from those for part 2), and the support of the functions g_x is contained in $[x-h_n, x+h_n]$.

Thus it suffices to establish that for any class of functions $\mathcal{F}$ of the form $\mathcal{F} = \{u \mapsto g_x(u) | x \in [h, 1-h]\}$ with $0 \leq h \leq 1/2$ with elements g_x that have support contained in $[x-h, x+h]$ and satisfy $\sup_{x\in[h,1-h]} \|g_x\|_\infty \leq C$, $\sup_{u\in\mathbb{R}} |g_x(u) - g_y(u)| \leq \tilde{C}|x-y|h^{-L}$ for some constants $C, \tilde{C}$ independent of h, x, y we have we have $N_{[\,]}(\mathcal{F}, \varepsilon, L^2(P_X)) \leq c\varepsilon^{-(2L+1)}$ for some c that does not depend on h.

To prove this statement, consider two cases.

1 $\varepsilon > 4h^{1/2}$

Divide $[0,1]$ into $N := 2/\varepsilon^2$ subintervals of length $2\alpha := \varepsilon^2$ with centers $r\alpha$ for $r = 1, ..., N$ and call the intervals $I_1, ..., I_N$. Note that two adjunct intervals overlap by $\alpha > 2h$. This construction ensures that every set of the form $[x-h, x+h]$ with $x \in [h, 1-h]$ is completely contained in at least one of the intervals defined above. Then a collection of N brackets of L^2-length $D\varepsilon$ for some $D > 0$ independent of h is given by $(-CI\{u \in I_j\}, CI\{u \in I_j\})$.

2 $\varepsilon \leq 4h^{1/2}$

Consider the points $t_i := i/(N+1), i = 1, ..., N$ with $N := 4^{2L+2}\tilde{C}/\varepsilon^{2L+1}$. By construction, to every $x \in [h, 1-h]$ there exists $i(x)$ with $|t_{i(x)} - x| \leq \varepsilon^{2L+1}/(4^{2L+2}\tilde{C})$.

This implies

$$\sup_u |g_x(u) - g_{t_{i(x)}}(u)| \le \tilde{C}\varepsilon^{2L+1}h^{-L}/(4^{2L+2}\tilde{C}) < \varepsilon/2$$

Then N $\|\cdot\|_\infty-$brackets of length ε covering $\mathcal{F}$ are given by $(g_{t_i}(\cdot) - \varepsilon/2, g_{t_i}(\cdot) + \varepsilon/2)$, $i = 1, ..., N$. From those one can easily construct $L^2(P_X)$-brackets.

Part 4 Follows by standard arguments.
Part 5 For any $\varepsilon > 0$, set $y_\varepsilon := \varepsilon^{-1/\alpha}D^{1/\alpha}$ and define $t_i := -y_\varepsilon + i\varepsilon/D$ for $i = 1, ..., N$ with N such that $1 + y_\varepsilon \ge t_N \ge y_\varepsilon$. Note that $N \le C\varepsilon^{-\frac{\alpha+1}{\alpha}}$ for some fixed, finite constant C which can depend on D but not on ε. The collection of brackets $(f \equiv 0, f \equiv \varepsilon), (f \equiv 1 - \varepsilon, f \equiv 1), (F(y_{t_i}|.) - \varepsilon/2, F(y_{t_i}|.) + \varepsilon/2)$ with $i = 1, ..., N$ covers the class $\mathcal{F}$. To see that, let $f \in \mathcal{F}$. Then there exists $y \in \mathbb{R}$ such that $f(\cdot) = F(y|\cdot)$. If $y < -y_\varepsilon$ we have

$$0 \le F(y|u) \le \sup_u F(-y_\varepsilon|u) \le y_\varepsilon^{-\alpha} \sup_u y_\varepsilon^\alpha F(-y_\varepsilon|u) \le D(\varepsilon^{-1/\alpha}D^{1/\alpha})^{-\alpha} = \varepsilon.$$

Similarly, $y > y_\varepsilon$ implies $1 - \varepsilon \le F(y|u) \le 1$ Finally, if $-y_\varepsilon \le y \le y_\varepsilon$, there exists $i \in \{1, ..., N\}$ such that $|y - t_i| \le \varepsilon/(2D)$. In that case

$$F(t_i|u) - \varepsilon/2 \le |F(t_i|u) - F(y|u)| + F(y|u) - \varepsilon/2 \le F(y|u) \le F(t_i|u) + \varepsilon/2$$

since $|F(t_i|u) - F(y|u)| \le D|t_i - y| \le \varepsilon/2$ by the assumption $\sup_{u,y} f(y|u) \le D$.
Part 6 Follows from $|I\{v \le g_1(u)\} - I\{v \le g_2(u)\}| \le I\{|v - g_1(u)| \le 2\|g_1 - g_2\|_\infty\}$.

□

S3.3 Main results for proofs

Define $\hat{\varepsilon}_{i,L}$ as the estimated residuals based on linearized versions $\hat{q}_{\tau,L}, \hat{s}_L$ [see Section S3.1 for their definition], i.e. $\hat{\varepsilon}_{i,L} := (Y_i - \hat{q}_{\tau,L}(X_i))/\hat{s}_L(X_i)$, and $\hat{\varepsilon}^*_{i,L}$ as the corresponding quantities in the bootstrap setting, that is

$$\hat{\varepsilon}^*_{i,L} = \frac{\hat{s}_L(X_i)\varepsilon_i^* + \hat{q}_{\tau,L}(X_i) - \hat{q}^*_{\tau,L}(X_i)}{\hat{s}^*_L(X_i)}$$

The following Lemma demonstrates, that the sequential empirical process based on the residuals $\hat{\varepsilon}_i = (Y_i - \hat{q}_\tau(X_i))/\hat{s}(X_i)$ computed from the initial estimators $\hat{q}_\tau, \hat{s}$ and the sequential empirical process of residuals based on $\varepsilon_{i,L}$ have the same first order expansion.

Lemma 9 *Assume that* ***(K1)***-***(K6)***, ***(A1)***-***(A5)***, ***(BW)*** *hold. Then*

$$\sup_{t\in[2h_n,1-2h_n],y\in\mathbb{R}} \Big|\frac{1}{\sqrt{n}}\sum_i I\{2h_n \le X_i \le t\}(I\{\hat{\varepsilon}_i \le y\} - I\{\hat{\varepsilon}_{i,L} \le y\})\Big| = o_P(1).$$

If additionally ***(B1)***-***(B2)*** *hold we also have*

$$\sup_{t\in[4h_n,1-4h_n],y\in\mathbb{R}} \Big|\frac{1}{\sqrt{n}}\sum_i I\{4h_n \le X_i \le t\}(I\{\hat{\varepsilon}^*_i \le y\} - I\{\hat{\varepsilon}^*_{i,L} \le y\})\Big| = o_P(1).$$

Proof We only proof the second assertion since the first one follows by similar but easier arguments. Start by observing that under the assumptions of the Lemma there exists a set D_n whose probability tends to one such that on D_n we have

$$(i)\ \sup_{x\in[4h_n,1-4h_n]} \max\Big(|\hat q_\tau(x)-\hat q_{\tau,L}(x)|, |\hat q^*_\tau(x)-\hat q^*_{\tau,L}(x)|, |\hat s(x)-\hat s_L(x)|, |\hat s^*(x)-\hat s^*_L(x)|\Big) \le \gamma_n$$

$$(ii)\qquad \inf_{x\in[4h_n,1-4h_n]} \min(\hat s_L(x), \hat s^*_L(x)) \ge c > 0$$

$$(iii)\qquad \sup_{y\in\mathbb{R}} |y\tilde f_\varepsilon(y)| \le C$$

for some deterministic sequence $\gamma_n = o(1/\sqrt{n})$ and finite constants $C, c > 0$. Here (i) and (ii) follow from Lemma 2 and Lemma 1 together Assumption (**A2**), while (iii) is a consequence of (S2.1) in the main body of the paper.

A standard Taylor expansion shows that on D_n

$$\begin{aligned}\Big|I\{\hat\varepsilon^*_i \le y\} - I\{\hat\varepsilon^*_{i,L} \le y\}\Big| \;&\le\; I\Big\{\Big|U_i - \tilde F_\varepsilon\Big(y\frac{\hat s^*_L(X_i)}{\hat s_L(X_i)} + \frac{\hat q^*_{\tau,L}(X_i)-\hat q_{\tau,L}(X_i)}{\hat s_L(X_i)}\Big)\Big| \le C\gamma_n\Big\} \\ &=:\; g_{n,y,C\gamma_n}(U_i, X_i),\end{aligned}$$

this follows from the representations

$$\begin{aligned} I\{\hat\varepsilon^*_i \le y\} \;&=\; I\Big\{U_i \le \tilde F_\varepsilon\Big(y\frac{\hat s^*(X_i)}{\hat s(X_i)} + \frac{\hat q^*_\tau(X_i)-\hat q_\tau(X_i)}{\hat s(X_i)}\Big)\Big\}, \\ I\{\hat\varepsilon^*_{i,L} \le y\} \;&=\; I\Big\{U_i \le \tilde F_\varepsilon\Big(y\frac{\hat s^*_L(X_i)}{\hat s_L(X_i)} + \frac{\hat q^*_{\tau,L}(X_i)-\hat q_{\tau,L}(X_i)}{\hat s_L(X_i)}\Big)\Big\},\end{aligned}$$

a Taylor expansion of $\tilde F_\varepsilon$ and (i)-(iii). In the same manner as the proof of Proposition 3 in Neumeyer (2009) it follows from assumptions (**B1**) and (**B2**) that, with probability tending to one

$$\mathcal{G}_n := \Big\{(u,v)\mapsto I\Big\{u \le z + \tilde F_\varepsilon\Big(y\frac{\hat s^*_L(v)}{\hat s_L(v)} + \frac{\hat q^*_{\tau,L}(v)-\hat q_{\tau,L}(v)}{\hat s_L(v)}\Big)\Big\}\Big|\ y\in\mathbb{R}, z\in[-2,2]\Big\} \tag{S3.30}$$

is contained in the class

$$\begin{aligned}\tilde{\mathcal{G}}_n = \Big\{(u,v)\mapsto I\Big\{u \le z + F\Big(y\frac{a_3(v)}{a_1(v)} + \frac{a_2(v)}{a_1(v)}\Big)\Big\}\Big|\ &F\in\mathcal{D}, a_1, a_3 \in \tilde C_C^{1+\delta}([4h_n, 1-4h_n]), \\ &a_2 \in C_C^{1+\delta}([4h_n, 1-4h_n]), y\in\mathbb{R}, z\in[-2,2]\Big\},\end{aligned}$$

where $\mathcal{D}$ is defined in (S2.2). Now, denoting by P the product measure of the uniform random variable U_1 and the covariate X_1,

$$\log N_{[\ ]}(\varepsilon, \tilde{\mathcal{G}}, L^2(P)) \le C\varepsilon^{-2\alpha} \tag{S3.31}$$

for some $\alpha < 1$, this can be shown by similar arguments as in the proof of Proposition 3 in Neumeyer (2009). Next, since $I\{|U_1 - a| \le b\} = I\{U_1 \le a+b\} - I\{U_1 \le a-b\}$ a.s., we find that, with probability tending to one

$$\mathcal{F}_n := \Big\{(u,v) \mapsto I\{s \le v \le t\} g_{n,y,C\gamma_n}(v,u) \Big| s,t \in [4h_n, 1-4h_n], y \in \mathbb{R}\Big\}$$
$$\subseteq \Big\{(u,v) \mapsto I\{s \le v \le t\}(g_1(v,u) - g_2(v,u)) \Big| s,t \in [4h_n, 1-4h_n], g_1, g_2 \in \tilde{\mathcal{G}}_n\Big\} =: \mathcal{G}_{n,1}.$$

Combining parts (1) and (4) of Lemma 8 thus yields that $\log N_{[\,]}(\varepsilon, \mathcal{F}_n, L^2(P)) \le \tilde{C}\varepsilon^{-2\alpha}$ for some constant $\tilde{C}$. Moreover, standard arguments (employing Taylor expansions and the bounds in (S2.1) from the main body of the paper) show that $\sup_{g\in\mathcal{F}_n} \int g dP = o(1/\sqrt{n})$ and $\sup_{g\in\mathcal{F}_n} \int g^2 dP = o(1)$. Here, P denotes the probability distribution of (X_i, U_i) and $g^2 = g$ for all $g \in \mathcal{F}_n$. Finally observe that, with probability tending to one,

$$\sup_{t\in[4h_n,1-4h_n], y\in\mathbb{R}} \frac{1}{\sqrt{n}} \sum_i \Big(I\{h_n \le X_i \le t\} g_{n,y,C\gamma_n}(U_i, X_i) - \int_{h_n}^t \int g_{n,y,C\gamma_n}(v,u) f_X(u) dv du]\Big)$$
$$\le \sqrt{n} \sup_{g\in\mathcal{F}_n} \Big(\int g dP_n - \int g dP\Big),$$

and the right-hand side of the inequality is of order $o_P(1)$ by part one of Lemma 7 . Moreover, standard arguments yield

$$\int_{h_n}^t \int g_{n,y,C\gamma_n}(v,u) f_X(u) dv du = o_P(1/\sqrt{n}).$$

Summarizing, we have obtained the estimate

$$\sup_{t\in[4h_n,1-4h_n], y\in\mathbb{R}} \frac{1}{\sqrt{n}} \sum_i I\{4h_n \le X_i \le t\} g_{n,y,C\gamma_n}(U_i, X_i) = o_P(1).$$

and thus the proof is complete. □

Lemma 10 *Assume that the conditions **(K1)**-**(K6)**, **(A1)**-**(A5)**, **(BW)** hold. Then*

$$\int_{h_n}^t \frac{\hat{q}_{\tau,L}(x) - q_\tau(x)}{s(x)} f_X(x) f_\varepsilon(0) dx = -\frac{1}{n} \sum_{i=1}^n (I\{\varepsilon_i \le 0\} - \tau) I_{[h_n,t]}(X_i) + o_P(1/\sqrt{n})$$

uniformly in $t \in [h_n, 1-h_n]$ *and*

$$\int_{2h_n}^t \frac{\hat{s}_L(x) - s(x)}{\hat{s}(x)} f_X(x)\, dx$$
$$= -\frac{1}{n} \sum_{i=1}^n \frac{I_{[2h_n,t]}(X_i)}{f_{|\varepsilon|}(1)} \Big(I\{|\varepsilon_i| \le 1\} - \frac{1}{2} - \frac{(I\{\varepsilon_i \le 0\} - \tau)(f_\varepsilon(1) - f_\varepsilon(-1))}{f_\varepsilon(0)}\Big) + o_P\Big(\frac{1}{\sqrt{n}}\Big)$$

uniformly in $t \in [2h_n, 1-2h_n]$.
*If additionally **(B1)**-**(B2)** hold*

$$\int_{3h_n}^t \frac{\hat{q}_\tau^*(x) - \hat{q}_{\tau,L}(x)}{\hat{s}_L(x)} f_X(x) dx = -\frac{1}{n} \sum_{i=1}^n \frac{I\{\varepsilon_i^* \le 0\} - \tau}{f_\varepsilon(0)} I_{[3h_n,t]}(X_i) + o_P(1/\sqrt{n})$$

uniformly in $t \in [3h_n, 1-3h_n]$ *and*

$$\int_{4h_n}^{t} \frac{\hat{s}^*(x) - \hat{s}(x)}{\hat{s}(x)} f_X(x)\,dx$$
$$= -\frac{1}{n}\sum_{i=1}^{n} \frac{I_{[4h_n,t]}(X_i)}{f_{|\varepsilon|}(1)}\Big(I\{|\varepsilon_i^*| \le 1\} - \frac{1}{2} - \frac{(I\{\varepsilon_i^* \le 0\} - \tau)(f_\varepsilon(1) - f_\varepsilon(-1))}{f_\varepsilon(0)}\Big) + o_P(\frac{1}{\sqrt{n}})$$

uniformly in $t \in [4h_n, 1-4h_n]$.

Proof We will only prove the representation for $\int_{3h_n}^{t} \frac{\hat{q}^*(x) - \hat{q}_{\tau,L}(x)}{\hat{s}_L(x)} f_X(x)dx$ since all other results can be derived by analogous arguments.

Observe the decomposition $\hat{q}_\tau^*(x) - \hat{q}_{\tau,L}(x) = \hat{q}_\tau^*(x) - q_\tau(x) + q_\tau(x) - \hat{q}_{\tau,L}(x)$. By Lemma 1 and Lemma 2 we have

$$\hat{q}_\tau^*(x) - \hat{q}_{\tau,L}^*(x) = o_P(1/\sqrt{n}), \quad \hat{q}_{\tau,L}^*(x) - q_\tau(x) = O_P(r_n), \quad \hat{s}_L(x) - s(x) = O_P(r_n),$$

uniformly in $x \in [3h_n, 1-3h_n]$. It thus suffices to establish

$$\int_{3h_n}^{t} \frac{\hat{q}_{\tau,L}^*(x) - q_\tau(x)}{s(x)} f_X(x)dx = \int_{3h_n}^{t} \frac{\hat{q}_{\tau,L}(x) - q_\tau(x)}{s(x)} f_X(x)dx - \frac{1}{n}\sum_{i=1}^{n} \frac{I\{\varepsilon_i^* \le 0\} - \tau}{f_\varepsilon(0)} I_{[3h_n,t]}(X_i)$$
$$+ o_P(1/\sqrt{n})$$

uniformly in $t \in [3h_n, 1-3h_n]$. By definition of $\hat{q}_{\tau,L}^*$, by part (iii)' of Lemma 3, and since $f_e(0|x) = s(x)f_\varepsilon(0)$ we have

$$\frac{f_X(x)(\hat{q}_{\tau,L}^*(x) - q_\tau(x))}{s(x)}$$
$$= -\frac{f_X(x)u_1^t\mathcal{M}(K)^{-1}}{f_\varepsilon(0)} \int_{-1}^{1} \kappa(v)\Big(\tilde{T}_{n,0,L,S}^*(x, q_{\tau+vb_n}(x)), \ldots, \tilde{T}_{n,p,L,S}^*(x, q_{\tau+vb_n}(x))\Big)^t dv + o_P(1\sqrt{n})$$

where

$$\tilde{T}_{n,k,L,S}^*(x,y) = \frac{1}{nh_n}\frac{1}{f_X(x)}\sum_{i=1}^{n} K_{h_n,k}(x - X_i)\Big(\Omega\Big(\frac{Y_i^* - y}{d_n}\Big) - F_Y(y|X_i)\Big).$$

The remaining proof is based on the following intermediate results which we will establish later on. First of all, uniformly in $t \in [3h_n, 1-3h_n]$, we have

$$\int_{3h_n}^{t} \tilde{T}_{n,k,L,S}^*(x, q_{\tau+vb_n}(x)) f_X(x)dx \tag{S3.32}$$
$$= \frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\int_{-1}^{1} K_{1,k}(u)\Big(\Omega\Big(\frac{Y_i^* - q_{\tau+vb_n}(X_i + uh_n)}{d_n}\Big)$$
$$-F_Y(q_{\tau+vb_n}(X_i + uh_n)|X_i)\Big)du + o_P(1/\sqrt{n}).$$

Moreover we have uniformly in $u \in [-1,1], t \in [3h_n, 1-3h_n]$

$$\begin{aligned}
&\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Omega\Big(\frac{Y_i^* - q_{\tau+vb_n}(X_i+uh_n)}{d_n}\Big) \\
= &\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\Omega\Big(\frac{\varepsilon_i^*\hat{s}_L(X_i)}{d_n}\Big) + vb_n\gamma_n(X_i) + \sum_{j=1}^p \xi_j(X_i,v,n)(uh_n)^j\Big) \qquad \text{(S3.33)} \\
&+\frac{f_\varepsilon(0)}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big) + o_P(n^{-1/2}) \\
= &\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(vb_n\gamma_n(X_i) + \sum_{j=1}^p \xi_j(X_i,v,n)(uh_n)^j\Big) \qquad \text{(S3.34)} \\
&+\frac{1}{n}\sum_i I_{[3h_n,t]}(X_i)I\{\varepsilon_i^*\leq 0\} + f_\varepsilon(0)\int_{3h_n}^t \frac{q_\tau(x)-\hat{q}_{\tau,L}(x)}{s(x)}f_X(x)dx + o_P(n^{-1/2}),
\end{aligned}$$

where ξ_j, γ_n denote some functions that do not depend on u. Additionally, a Taylor expansion of $(u,v)\mapsto F_Y(q_{\tau+vb_n}(X_i+uh_n)|X_i)$ shows that

$$\begin{aligned}
&\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)F_Y(q_{\tau+vb_n}(X_i+uh_n)|X_i) \\
= &\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\tau + vb_n + \sum_{j=1}^p \zeta_j(X_i,v,n)(uh_n)^j\Big) + o_P(n^{-1/2}), \qquad \text{(S3.35)}
\end{aligned}$$

where the remainder holds uniformly in $u\in[-1,1], t\in[3h_n,1-3h_n]$ and the functions ζ_j are again independent of u. Plugging (S3.34) and (S3.35) into (S3.32) we find that

$$\int_{-1}^1 \kappa(v)\int_{3h_n}^t \tilde{T}^*_{n,k,L,S}(x,q_{\tau+vb_n}(x))dxdv = \sum_{j=0}^p \mu_{k+j}(K)w_j(t) + o_P(1/\sqrt{n})$$

where

$$\begin{aligned}
w_0(t) &:= \Big(f_\varepsilon(0)\int_{3h_n}^t \frac{q_\tau(u)-\hat{q}_{\tau,L}(u)}{s(u)}f_X(u)du + \frac{1}{n}\sum_{i=1}^n I_{[3h_n,t]}(X_i)(I\{\varepsilon_i^*\leq 0\}-\tau)\Big), \\
w_j(t) &:= \frac{h_n^j}{n}\sum_{i=1}^n I_{[3h_n,t-h_n]}(X_i)\int_{-1}^1 \kappa(v)(\xi_j(X_i,v,n)-\zeta_j(X_i,v,n))dv, \quad j=1,...,p.
\end{aligned}$$

Thus, uniformly in $t\in[3h_n,1-3h_n]$,

$$\begin{aligned}
&f_X(x)\int_{-1}^1 \kappa(v)\Big(\tilde{T}^*_{n,0,L,S}(x,q_{\tau+vb_n}(x)),\dots,\tilde{T}^*_{n,p,L,S}(x,q_{\tau+vb_n}(x))\Big)^t dv \\
= &\mathcal{M}(K)(w_0(t),...,w_p(t))^t + o_P(1/\sqrt{n}).
\end{aligned}$$

Hence the proof will be complete once we establish (S3.32)-(S3.34).

Proof of (S3.32)

Recalling that K has support $[-1,1]$, we obtain for any $t \in [3h_n, 1-3h_n]$ the decomposition

$$K_{h_n,k}(x-X_i)I_{[3h_n,t]}(x) = K_{h_n,k}(x-X_i)I_{[3h_n,t]}(x)\Big(I_{(t-h_n,t+h_n]}(X_i)+I_{[2h_n,3h_n)}(X_i)+I_{[3h_n,t-h_n]}(X_i)\Big).$$

We will now show that the contributions corresponding to the summands containing $I_{[2h_n,3h_n)}(X_i)$ and $I_{(t-h_n,t+h_n]}(X_i)$ are negligible. Since both expressions can be treated analogously, we only provide the arguments for $I_{(t-h_n,t+h_n]}(X_i)$. By similar arguments as in the proof of Lemma 3 it is easy to show that

$$\sup_{t,x\in[3h_n,1-3h_n],y\in\mathcal{Y}} \Big|\frac{1}{nh_n}\sum_{i=1}^n \frac{K_{h_n,k}(x-X_i)}{f_X(x)} I_{(t-h_n,t+h_n]}(X_i)\Big(\Omega\Big(\frac{Y_i^*-y}{d_n}\Big) - F_Y(y|X_i)\Big)\Big|$$
$$=: A_n(\mathcal{Y}) = O_P(r_n)$$

for any bounded $\mathcal{Y} \subset \mathbb{R}$. Observe that $K_{h_n,k}$ vanishes outside $[-h_n, h_n]$, and since

$$I\{|x-X_i| \le h_n\}I_{[3h_n,t]}(x)I_{(t-h_n,t+h_n]}(X_i) \le I_{[t-2h_n,t+2h_n]}(x)I_{[t-h_n,t+h_n]}(X_i)$$

we obtain, for a suitably chosen $\mathcal{Y}$,

$$\Big|\int_{3h_n}^t \frac{1}{nh_n}\sum_{i=1}^n \frac{K_{h_n,k}(x-X_i)}{f_X(x)} I_{[t-h_n,t+h_n]}(X_i)\Big(\Omega\Big(\frac{Y_i^*-q_{\tau+vb_n}(x)}{d_n}\Big) - F_Y(q_{\tau+vb_n}(x)|X_i)\Big)dx\Big|$$
$$\le \int_{t-2h_n}^{t+2h_n} A_n(\mathcal{Y})dx = O_P(h_n r_n) = o_P(1/\sqrt{n})$$

uniformly in $t \in [3h_n, 1-3h_n], v \in [-1,1]$. This completes the proof of (S3.32).

Proof of (S3.33) Throughout this part of the proof, let $\mathcal{Y} \subset \mathbb{R}$ denote a fixed, bounded set containing the interval $[-d_n, d_n]$ for sufficiently large n. The following statement will be proved later

$$\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(I\{Y_i^* \le q_{\tau+vb_n}(X_i+uh_n)+y\} - I\{\varepsilon_i^* \le y/\hat{s}_L(X_i)\}\Big) \quad \text{(S3.36)}$$
$$= \frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\bar{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n) - \hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)$$
$$-\bar{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big) + f_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}\Big)\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big) + o_P(1/\sqrt{n})$$

uniformly in $t \in [3h_n, 1-3h_n], u, v \in [-1,1], y \in \mathcal{Y}$ where $\bar{F}_\varepsilon$ is defined in Lemma 5. Now convolving both sides of (S3.36) [with respect to the argument y] with $\frac{1}{d_n}\omega(\cdot/d_n)$ and evaluating the result in 0 yields the identity

$$\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\Omega\Big(\frac{Y_i^*-q_{\tau+vb_n}(X_i+uh_n)}{d_n}\Big) - \Omega\Big(\frac{\hat{s}_L(X_i)\varepsilon_i^*}{d_n}\Big)\Big)$$

$$= \frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\bar{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)-\bar{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)$$
$$+f_\varepsilon(0)\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)+o_P(1/\sqrt{n}).$$

Observe that the smoothness properties of $\bar{F}_\varepsilon$ (defined in Lemma 5) yield the representation

$$\bar{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)-\bar{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}\Big)$$
$$= vb_n\gamma_n(X_i)+\sum_{j=1}^p \xi_j(X_i,v,n)(uh_n)^j+r_{n,1}$$

where the remainder terms $r_{n,1}$ is of order $O(b_n^2+h_n^{p+1})=o(1/\sqrt{n})$ uniformly in u,v and ξ_j,γ_n denote some functions that do not depend on u. Thus the proof of (S3.33) will be complete once we establish (S3.36). To this end, observe that

$$I\Big\{Y_i^*\le q_{\tau+vb_n}(X_i+uh_n)+y\Big\}=I\Big\{\varepsilon_i^*\le\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_\tau(X_i)+y}{\hat{s}(X_i)}\Big\}$$

and

$$\frac{1}{n}\sum_i I_{[3h_n,t-3h_n]}(X_i)\Big(I\Big\{\varepsilon_i^*\le\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}(X_i)+y}{\hat{s}(X_i)}\Big\}-I\Big\{\varepsilon_i^*\le\frac{y}{\hat{s}_L(X_i)}\Big\}\Big)$$
$$= \frac{1}{n}\sum_i I_{[3h_n,t-3h_n]}(X_i)\Big(I\Big\{\varepsilon_i^*\le\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big\}-I\Big\{\varepsilon_i^*\le\frac{y}{\hat{s}_L(X_i)}\Big\}\Big)+o_P(1/\sqrt{n})$$
$$= \frac{1}{n}\sum_i I_{[3h_n,t-3h_n]}(X_i)\Big(\tilde{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\tilde{F}_\varepsilon(y/\hat{s}_L(X_i))\Big)+o_P(1/\sqrt{n})$$

uniformly in t,v,u, which follows by arguments similar to those used in the proof of Lemma 9. Consider the decomposition

$$\tilde{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}\Big)$$
$$= \tilde{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\tilde{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)$$
$$+\tilde{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}\Big).$$

For the first term in this decomposition, an application of Lemma 5 yields

$$\frac{1}{n}\sum_i I_{[3h_n,t-3h_n]}(X_i)\Big[\tilde{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\tilde{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)\Big]$$
$$= \frac{1}{n}\sum_i I_{[3h_n,t-3h_n]}(X_i)\Big[\bar{F}_\varepsilon\Big(\frac{q_{\tau+vb_n}(X_i+uh_n)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\bar{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)\Big]$$
$$+o_P(1/\sqrt{n}),$$

where $\bar{F}_\varepsilon$ is defined in Lemma 5. Noting that

$$\tilde{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}\Big)=\tilde{f}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}\Big)\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}+o_P(1/\sqrt{n}),$$

and recalling that $\tilde{f}_\varepsilon$ converges to f_ε uniformly with rate $o_P((h_n/\log n)^{1/2})$ [see (S2.1)] combined with $r_n(h_n/\log n)^{1/2}=o(1)$ yields

$$\tilde{F}_\varepsilon\Big(\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)+y}{\hat{s}_L(X_i)}\Big)-\tilde{F}_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}\Big)=f_\varepsilon\Big(\frac{y}{\hat{s}_L(X_i)}\Big)\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}+o_P(1/\sqrt{n})$$

which completes the proof of (S3.36) and thus (S3.33) is also established.
Proof of (S3.34) It suffices to show that, uniformly in $t\in[3h_n,1-3h_n]$

$$\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\Omega\Big(\frac{\hat{s}_L(X_i)\varepsilon_i^*}{d_n}\Big)-I\{\varepsilon_i^*\le 0\}\Big)=o_P(\frac{1}{\sqrt{n}}),\tag{S3.37}$$

$$\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)(I\{\varepsilon_i^*\le 0\}-\tau)=\frac{1}{n}\sum_i I_{[3h_n,t]}(X_i)(I\{\varepsilon_i^*\le 0\}-\tau)+o_P(\frac{1}{\sqrt{n}}),\tag{S3.38}$$

$$\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}=\int_{3h_n}^t\frac{q_\tau(u)-\hat{q}_{\tau,L}(u)}{s(u)}f_X(u)du+o_P(\frac{1}{\sqrt{n}}).\tag{S3.39}$$

The statement in (S3.39) follows since, for $t\in[4h_n,1-3h_n]$,

$$\begin{aligned}\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{\hat{s}_L(X_i)}&=\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\frac{q_\tau(X_i)-\hat{q}_{\tau,L}(X_i)}{s(X_i)}+o_P(1/\sqrt{n})\\&=\int_{3h_n}^{t-h_n}\frac{q_\tau(u)-\hat{q}_{\tau,L}(u)}{s(u)}f_X(u)du+o_P(1/\sqrt{n})\\&=\int_{3h_n}^{t}\frac{q_\tau(u)-\hat{q}_{\tau,L}(u)}{s(u)}f_X(u)du+o_P(1/\sqrt{n}),\end{aligned}$$

where the first equality follows from the rates of convergence for $\hat{q}_{\tau,L}-q_\tau$, $\hat{s}_L-s$ [see Lemma 1 and Lemma 2], the second equality is a consequence of the fact that $\hat{q}_{\tau,L}\in C_C^\delta$ with probability tending to one [see Lemma 1] combined with Lemma 7. For $t<4h_n$, the left-hand side of (S3.38) is zero and the right-hand side of order $o_P(n^{-1/2})$ by Lemma 1 and Lemma 2.
For a proof of (S3.37), observe that

$$\Omega\Big(\frac{\hat{s}_L(X_i)\varepsilon_i^*}{d_n}\Big)-I\{\varepsilon_i^*\le 0\}=\frac{1}{d_n}\int_{-d_n}^{d_n}\Big(I\{\varepsilon_i^*\le a/\hat{s}_L(X_i)\}-I\{\varepsilon_i^*\le 0\}\Big)\omega\Big(\frac{a}{d_n}\Big)da.$$

Define the sequence of sets

$$S(\delta_n):=\{(t,y_n,z_n)|t\in[3h_n,1-3h_n],y_n,z_n\in\mathcal{Y},|y_n-z_n|\le\delta_n\}$$

for some $\delta_n = o(1)$. Observe that, with probability tending to one,

$$\begin{aligned}
&\sup_{(t,y_n,z_n)\in S(\delta_n)} \Big|\frac{1}{n}\sum_{i=1}^n I_{[3h_n,t-3h_n]}(X_i)\Big(I\{\varepsilon_i^* \le y_n\} - I\{\varepsilon_i^* \le z_n\} + \tilde F_\varepsilon(z_n) - \tilde F_\varepsilon(y_n)\Big)\Big| \\
&= \sup_{(t,y_n,z_n)\in S(\delta_n)} \Big|\frac{1}{n}\sum_{i=1}^n I_{[3h_n,t-3h_n]}(X_i)\Big(I\{U_i \le \tilde F_\varepsilon(y_n)\} - I\{U_i \le \tilde F_\varepsilon(z_n)\} + \tilde F_\varepsilon(z_n) - \tilde F_\varepsilon(y_n)\Big)\Big| \\
&\le \sup_{(t,y_n,z_n)\in S(C\delta_n)} \Big|\frac{1}{n}\sum_{i=1}^n I_{[3h_n,t-3h_n]}(X_i)\Big(I\{U_i \le y_n\} - I\{U_i \le z_n\} + z_n - y_n\Big)\Big| \\
&= o_P(1/\sqrt{n}).
\end{aligned}$$

Here, for the first inequality we made use of (B.1). This implies that, with probability tending to one, $\tilde F_\varepsilon$ has a uniformly bounded derivative which shows that, with probability tending to one, $|y_n - z_n| \le \delta_n$ implies $|\tilde F_\varepsilon(y_n) - \tilde F_\varepsilon(z_n)| \le C\delta_n$ for some finite constant C. The last bound above follows by standard empirical process arguments provided that $\delta_n = o(1)$. Thus

$$\begin{aligned}
&\frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\Big(\Omega\Big(\frac{\hat s_L(X_i)\varepsilon_i^*}{d_n}\Big) - I\{\varepsilon_i^* \le 0\}\Big) \\
&= \frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\frac{1}{d_n}\int_{-d_n}^{d_n}\Big(\tilde F_\varepsilon(a/\hat s_L(X_i)) - \tilde F_\varepsilon(0)\Big)\omega\Big(\frac{a}{d_n}\Big)da + o_P(n^{-1/2}) \\
&= \frac{1}{n}\sum_i I_{[3h_n,t-h_n]}(X_i)\frac{1}{d_n}\int_{-d_n}^{d_n}\Big(\bar F_\varepsilon(a/\hat s_L(X_i)) - \bar F_\varepsilon(0)\Big)\omega\Big(\frac{a}{d_n}\Big)da + o_P(n^{-1/2}) \\
&= o_P(n^{-1/2})
\end{aligned}$$

where the second to last line follows by Lemma 5 and the last line is a consequence of the smoothness properties of $\bar F_\varepsilon$.

Thus (S3.37) follows and it remains to establish (S3.38). To this end, observe that it suffices to establish

$$\sup_{t\in[3h_n,1-3h_n]}\Big|\frac{1}{n}\sum_{i=1}^n I_{[t-h_n,t]}(X_i)(I\{\varepsilon_i^* \le 0\} - \tau)\Big| = o_P(n^{-1/2}).$$

Now

$$\frac{1}{n}\sum_{i=1}^n I_{[t-h_n,t]}(X_i)(I\{\varepsilon_i^* \le 0\} - \tau) = \frac{1}{n}\sum_{i=1}^n I_{[t-h_n,t]}(X_i)(I\{U_i \le \tilde F_\varepsilon(0)\} - \tau),$$

and by (S3.27) in Lemma 5 we have $\tilde F_\varepsilon(0) - \tau = \tilde F_\varepsilon(0) - F_\varepsilon(0) = O_P(r_n)$. Thus we have with probability tending to one $|\tilde F_\varepsilon(0) - \tau| \le r_n h_n^{-1/4}$ and in particular

$$\begin{aligned}
&\sup_{t\in[3h_n,1-3h_n]}\Big|\frac{1}{n}\sum_{i=1}^n I_{[t-h_n,t]}(X_i)(I\{\varepsilon_i^* \le 0\} - \tau)\Big| \\
&\le \sup_{t\in[3h_n,1-3h_n]}\ \sup_{|y|\le r_n h_n^{-1/4}}\Big|\frac{1}{n}\sum_{i=1}^n I_{[t-h_n,t]}(X_i)(I\{U_i \le y\} - \tau)\Big| = o_P(n^{-1/2})
\end{aligned}$$

where the first inequality holds with probability tending to one and the equality follows by standard empirical process arguments. Thus (S3.37) follows. This completes the proof of Lemma 10.

□

References

M. Akritas and I. Van Keilegom (2001). *Nonparametric estimation of the residual distribution.* Scand. J. Statist. 28, 549–567.

N. Neumeyer (2007). *A note on uniform consistency of monotone function estimators.* Statist. Probab. Lett. 77, 693–703.

N. Neumeyer (2009). *Smooth residual bootstrap for empirical processes of nonparametric regression residuals.* Scand. J. Statist. 36, 204–228.

G. R. Shorack and J. A. Wellner (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

T. J. Sweeting (1989). *On Conditional Weak Convergence.* J. Theoret. Probab. 2, 461–474.

A. W. van der Vaart and J. A. Wellner (1996). *Weak convergence and empirical processes.* Springer, New York.

S. Volgushev, M. Birke, H. Dette and N. Neumeyer (2013). *Significance testing in quantile regression.* Electron. J. Stat. 7, 105–145.

Melanie Birke, Universität Bayreuth, Fakultät für Mathematik, Physik und Informatik, 95440 Bayreuth, Germany, e-mail: Melanie.Birke@uni-bayreuth.de

Natalie Neumeyer, Universität Hamburg, Fachbereich Mathematik, Bundesstraße 55, 20146 Hamburg, Germany, e-mail: neumeyer@math.uni-hamburg.de

Stanislav Volgushev, University of Toronto, Department of Statistical Sciences 100 St. George Street, Toronto, Ontario, Canada, e-mail: volgushe@utstat.toronto.edu