Supplementary Material on "Uniform Bahadur Representation for Nonparametric Censored Quantile Regression: A Redistribution-of-Mass Approach"

Efang Kong
University of Electronic Science and Technology, China
and *University of Kent at Canterbury, UK*

Yingcun Xia
National University of Singapore, Singapore
and *University of Electronic Science and Technology, China*

This section contains detailed proofs of all the theoretical results from the main text, together with additional lemmas and propositions. Equations which first appear in this section would be labelled as (A??) while those which have already appeared in the main text are referenced by the Arabic number already assigned to them. Also in this section, symbols $\sup_{\mathbf{x}}$, $\sup_{\mathbf{x},\tau}$ or $\sup_{\mathbf{x},\tau,\boldsymbol{\beta}}$ stand for sup lim taken with respect to $\mathbf{x} \in \mathcal{D}, \tau \in [\tau^*, 1-\tau^*]$, and $\boldsymbol{\beta} \in R^{n(A)}$. Similarly, the phrase *uniformly in* $\mathbf{x}$, τ means *uniformly in* $\mathbf{x} \in \mathcal{D}, \tau \in [\tau^*, 1-\tau^*]$.

Proof of Lemma 1.With $h_n \to 0$, $nh_n^p/\log n \to \infty$ as $n \to \infty$, and the sequence of weights $B_{nk}(\mathbf{x})$ chosen to be the Gasser-Muller's type weights, Theorem 2.2 of Gonzalez-Manteigaa and Cadarso-Suarez (1994) states that

$$\sup_{\mathbf{x},\tau} |\hat{F}_{KM}(t|\mathbf{x}) - F_0(t|\mathbf{x})| = O\Big(h_n^2 + \Big(\frac{\log n}{nh_n^p}\Big)^{1/2}\Big),$$

$$\hat{F}_{KM}(t|\mathbf{x}) - F_0(t|\mathbf{x}) = \frac{\{1-F_0(t|\mathbf{x})\}}{nh_n^p}\sum_{k=1}^n \tilde{K}_{h_n}(\mathbf{X}_{k\mathbf{x}})\varphi(Y_k, d_k, t, \mathbf{x}) + O\Big(h_n^2 + \Big(\frac{\log n}{nh_n^p}\Big)^{3/4}\Big),$$

uniformly in $\mathbf{x} \in \mathcal{D}$ and t. Their line of arguments is still valid with other forms of weights, as long as $\sum_{k=1}^n B_{nk}(\mathbf{x}) = 1$. Specifically, for the generalized K-M estimator $\hat{F}_{KM}(.|\mathbf{X}_i)$ of (10) with local polynomial weights, we have, with probability one

$$\sup_{\mathbf{x},\tau} |\hat{F}_{KM}(t|\mathbf{x}) - F_0(t|\mathbf{x})| = O\Big(h_n^{s_3} + \Big(\frac{\log n}{nh_n^p}\Big)^{1/2}\Big),$$

$$\hat{F}_{KM}(t|\mathbf{x}) - F_0(t|\mathbf{x}) = \frac{\{1-F_0(t|\mathbf{x})\}}{nh_n^p}\sum_{k=1}^n \tilde{B}_{nk}(\mathbf{x})\varphi(Y_k, d_k, t, \mathbf{x}) + O\Big(h_n^{s_3} + \Big(\frac{\log n}{nh_n^p}\Big)^{3/4}\Big). \quad (A.1)$$

Note that with $nh_n^{p+4/3s_3}/\log n < \infty$ (Assumption [A7]), the term $h_n^{s_3}$ can be absorbed to form part of $O\Big((nh_n^p/\log n)^{-3/4}\Big)$.

Plug (A.1) into the definition of $\hat{F}_n^S(.|.)$, and we have again with probability one,

$$\begin{aligned}\hat{F}_n^S(t|\mathbf{x}) &= \int \Big\{F_0(s|\mathbf{x}) + \frac{1}{nh_n^p}\sum_{k=1}^n \tilde{B}_{nk}(\mathbf{x})\{1-F_0(s|\mathbf{x})\}\varphi(Y_k, d_k, s, \mathbf{x})\Big\}\bar{K}(\frac{s-t}{h_{1n}})\frac{1}{h_{1n}}ds \\ &= F_0(t|\mathbf{x}) + \frac{\{1-F_0(t|\mathbf{x})\}}{nh_n^p}\sum_{k=1}^n \tilde{B}_{nk}(\mathbf{x})\chi(Y_k, d_k, t, \mathbf{x}) + O\Big(h_{1n}^2 + \Big(\frac{\log n}{nh_n^p}\Big)^{3/4}\Big), \quad \text{(A.2)}\end{aligned}$$

where

$$\chi(Y_k, d_k, t, \mathbf{x}) = \int \{1-F_0(s|\mathbf{x})\}\varphi(Y_k, d_k, s, \mathbf{x})\frac{1}{h_{1n}}\bar{K}\Big(\frac{s-t}{h_{1n}}\Big)ds.$$

Since $E[\chi(Y_k, d_k, t, \mathbf{x})] = 0$, $E|\varphi(Y_k, d_k, t + h_{1n}s, \mathbf{x}) - \varphi(Y_k, d_k, t, \mathbf{x})| = O(h_{1n})$, and $h_{1n} = O((nh_n^p/\log n)^{-1/2})$ (Assumption [A8]), we have, with probability one,

$$\begin{aligned}&\frac{\{1-F_0(t|\mathbf{x})\}}{nh_n^p}\sum_{k=1}^n \tilde{B}_{nk}(\mathbf{x})\chi(Y_k, d_k, t, \mathbf{x}) - \frac{\{1-F_0(t|\mathbf{x})\}}{nh_n^p}\sum_{k=1}^n \tilde{B}_{nk}(\mathbf{x})\varphi(Y_k, d_k, t, \mathbf{x}) \\ &= O(\{\frac{h_{1n}\log n}{nh_n^p}\}^{1/2}) = O(\{\frac{\log n}{nh_n^p}\}^{3/4}). \quad \text{(A.3)}\end{aligned}$$

With $h_{1n}^2 = O(h_n^{s_3})$ (Assumption [A7]), Lemma 1 thus follows from (A.2) and (A.3). □

To keep the exposition simple, we illustrate by focusing on the case where $K(.)$ is the uniform density on $[-1,1]^p$. The arguments could be easily adapted for the case of a general symmetric probability density in R^p with a compact support. Rewrite (9) as

$$\sum_{i\in S_n(\mathbf{x})}\Big[w_{in}(\tau)\rho_\tau(Y_i - \boldsymbol{\beta}^\top\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})) + (1-w_{in}(\tau))\rho_\tau(Y^{+\infty} - \boldsymbol{\beta}^\top\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}))\Big], \quad \text{(A.4)}$$

where the index set $S_n(\mathbf{x}) = \{i : 1 \le i \le n,\ |\mathbf{X}_{i\mathbf{x}}| \le \delta_n\}$ with cardinality $N_n(\mathbf{x}) = \sharp(S_n(\mathbf{x}))$. Denote the minimizer of (A.4) as $\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x})$; and to facilitate the discussion on its properties, we implicitly assume the following simple facts; see also Chaudhuri (1991) and Kong , Linton and Xia (2013).

[FACT1] For any positive integer m, let $\mathbf{x}$ be a vector in R^m and $p(\mathbf{x})$ be an arbitrary nonzero polynomial in $\mathbf{x}$. Then, the Lebesgue measure of the set $\{\mathbf{x}|p(\mathbf{x}) = 0\}$ is 0.

[FACT2] Let $\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}$ be independent random vectors in R^m with the property that $Prob(\mathbf{X}^{(i)} \in \mathbf{H}) = 0$ for all $i = 1, \cdots, n$, and any given linear subspace $\mathbf{H}$ of R^m such that $\dim(\mathbf{H}) \le m-1$. Then the collection $\{\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(m)}\}$ is almost surely linearly independent.

[FACT3] For any $\mathbf{x} \in \mathcal{D}$, denote by $\omega_{\delta_n}(\mathbf{t}, \mathbf{x})$, the conditional density of $\delta_n^{-1}(\mathbf{X} - \mathbf{x})$, given that $|\mathbf{X} - \mathbf{x}| \le \delta_n$. Then $\omega_{\delta_n}(\mathbf{t}, \mathbf{x})$ converges to the uniform density on $[-1,1]^p$ uniformly in $\mathbf{t}$ as well as in $\mathbf{x} \in \mathcal{D}$.

For any given $\mathbf{x} \in \mathcal{D}$, let $[DX_n(\mathbf{x})]$ be the $N_n(\mathbf{x}) \times n(A)$ matrix with rows $\{\boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}}), i \in S_n(\mathbf{x})\}$, and $VY_n(\mathbf{x})$ be the $N_n(\mathbf{x}) \times 1$ vector $\{Y_i, i \in S_n(\mathbf{x})\}$. For any subset $\mathbf{h} \subset S_n(\mathbf{x})$ such that $\sharp(\mathbf{h}) = n(A)$, denote by $[DX_n(\mathbf{x}, \mathbf{h})]$, the $n(A) \times n(A)$ matrix with rows $\{\boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}}),\ i \in \mathbf{h}\}$, and by $VY_n(\mathbf{x}, \mathbf{h})$, the $n(A) \times 1$ vector $\{Y_i, i \in \mathbf{h}\}$. Define

$$\mathrm{H}_n(\mathbf{x}) = \{\mathbf{h} : \mathbf{h} \subset S_n(\mathbf{x}),\ \sharp(h) = n(A), [DX_n(\mathbf{x}, \mathbf{h})] \text{ has full rank}\}.$$

The following two propositions describe two critical facts about the minima of (A.4).

Proposition 1. *For any given $\mathbf{x} \in \mathcal{D}$ such that $[DX_n(\mathbf{x})]$ has rank $n(A)$, then there is a subset $\mathbf{h} \in \mathrm{H}_n(\mathbf{x})$, such that (A.4) has at least one minimum of the form*

$$\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}) = [DX_n(\mathbf{x}, \mathbf{h})]^{-1} VY_n(\mathbf{x}, \mathbf{h}).$$

Proof. Similar to the proof for Theorem 3.1 in Koenker and Bassett (1978), this roughly follows from the following linear programming formulation of the minimization of (A.4):

$$\min_{u_i^+,\ u_i^-} \sum_{i \in S_n(\mathbf{x})} w_{in}(\tau)[\tau u_i^+ + (1-\tau)u_i^-] + \sum_{i \in S_n(\mathbf{x})} \tau(1 - w_{in}(\tau))(u_i^+ - u_i^-), \tag{A.5}$$

subject to

$$\begin{aligned} &Y_i - \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) = u_i^+ - u_i^-; \\ &\boldsymbol{\beta} \in R^{n(A)};\ u_i^+,\ u_i^- \geq 0,\ i = 1, \cdots, N_n(\mathbf{x}). \end{aligned}$$

Note that we have implicitly used the following fact

$$\rho_\tau\{Y^{+\infty} - \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\} = \tau(Y^{+\infty} - Y_i + u_i^+ - u_i^-).$$

An equivalence to problem (A.5) is

$$\min_{u_i^+,\ u_i^-} \sum_{i \in S_n(\mathbf{x})} [\tau u_i^+ + (w_{in}(\tau) - \tau)u_i^-], \tag{A.6}$$

subject to

$$\begin{aligned} &VY_n(\mathbf{x}) - [DX_n(\mathbf{x})]\boldsymbol{\beta} = \mathbf{u}^+ - \mathbf{u}^-; \\ &(\boldsymbol{\beta}, \mathbf{u}^+, \mathbf{u}^-) \in R^{n(A)} \times R_+^{2N_n(\mathbf{x})}; \\ &\mathbf{u}^+ = (u_i^+, i \in S_n(\mathbf{x})); \quad \mathbf{u}^- = (u_i^-, i \in S_n(\mathbf{x})). \end{aligned}$$

According to Theorem 7.7.4 in Gill, Murray and Wright (1991), problem (A.6) has a vertex solution, i.e. there exists some subset $\mathbf{h} \in \mathrm{H}_n(\mathbf{x})$, such that $u_i^+ = u_i^- = 0,\ i \in \mathbf{h}$, which is equivalent to the statement in Proposition 1. $\square$

Let $\mathbf{1}_{n(A)}$ stand for the $n(A) \times 1$ vector of ones, and for any $\mathbf{x} \in \mathcal{D}$ and $\mathbf{h} \in \mathrm{H}_n(\mathbf{x})$, define $\underline{\mathbf{w}}_n(\mathbf{h}|\tau) = (w_{in}(\tau),\ i \in \mathbf{h})$, and

$$L_n(\mathbf{x}, \mathbf{h}|\tau) = \Big[DX_n(\mathbf{x}, \mathbf{h})\Big]^{-1} \sum_{i \in \bar{\mathbf{h}}} \Big[w_{in}(\tau) I\{Y_i \leq \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x})\} - \tau\Big] \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}),$$

where $\bar{\mathbf{h}} = S_n(\mathbf{x}) \backslash \mathbf{h}$ denotes the complement of $\mathbf{h}$ in $S_n(\mathbf{x})$.

Proposition 2. *If, for some* $\mathbf{h} \in \mathrm{H}_n(\mathbf{x})$, $\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}) = [DX_n(\mathbf{x}, \mathbf{h})]^{-1} VY_n(\mathbf{x}, \mathbf{h})$ *is a minimum (not necessarily a unique one) of (A.4), then we must have*

$$L_n(\mathbf{x}, \mathbf{h}|\tau) \in [\tau \mathbf{1}_{n(A)} - \underline{\mathbf{w}}_n(\mathbf{h}|\tau), \tau \mathbf{1}_{n(A)}], \tag{A.7}$$

where $[\tau \mathbf{1}_{n(A)} - \underline{\mathbf{w}}_n(\mathbf{h}|\tau), \tau \mathbf{1}_{n(A)}]$ *is the* $n(A)-$*fold Cartesian product of the closed intervals* $\{[\tau - w_{in}(\tau), \tau],\ i \in \mathbf{h}\}$. *Further,* $\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x})$ *is a unique minima if and only if* $L_n(\mathbf{x}, \mathbf{h}|\tau) \in (\tau \mathbf{1}_{n(A)} - \underline{\mathbf{w}}_n(\mathbf{h}|\tau), \tau \mathbf{1}_{n(A)})$.

Proof. Rewrite optimization problem (A.6) as minimizing with respect to $\boldsymbol{\beta}$ the following quantity

$$\psi(\boldsymbol{\beta}) = \sum_{i \in S_n(\mathbf{x})} [\tau - \frac{w_{in}(\tau)}{2} + \frac{w_{in}(\tau)}{2} \mathrm{sign}(Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\boldsymbol{\beta})][Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\boldsymbol{\beta}].$$

Its directional derivative in direction $\underline{a}$, a unit vector in $R^{n(A)}$,

$$\psi'(\boldsymbol{\beta}; \underline{a}) = \sum_{i \in S_n(\mathbf{x})} \Big\{\frac{w_{in}(\tau)}{2} - \frac{w_{in}(\tau)}{2} \mathrm{sign}^*[Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\boldsymbol{\beta}; -\boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\underline{a}] - \tau\Big\} \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\underline{a}, \tag{A.8}$$

where

$$\mathrm{sign}^*(u; z) = \begin{cases} \mathrm{sign}(u) & \text{if } u \neq 0, \\ \mathrm{sign}(z) & \text{if } u = 0. \end{cases}$$

This can be derived as follows. For small enough $\delta > 0$,

$$\psi(\boldsymbol{\beta} + \delta\underline{a}) = \sum_{i \in S_n(\mathbf{x})} \{\tau - \frac{w_{in}(\tau)}{2} + \frac{w_{in}(\tau)}{2} \mathrm{sign}[Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})(\boldsymbol{\beta} + \delta\underline{a})]\}\{Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})(\boldsymbol{\beta} + \delta\underline{a})\}.$$

Therefore,

$$\begin{aligned} & \psi(\boldsymbol{\beta} + \delta\underline{a}) - \psi(\boldsymbol{\beta}) \\ = & \sum_{i \in S_n(\mathbf{x})} \frac{w_{in}(\tau)}{2} \Big[\mathrm{sign}\{Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})(\boldsymbol{\beta} + \delta\underline{a})\} - \mathrm{sign}\{Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\boldsymbol{\beta}\}\Big] \{Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\boldsymbol{\beta}\} \\ & -\delta\underline{a}^\top \sum_{i \in S_n(\mathbf{x})} [\tau - \frac{w_{in}(\tau)}{2} + \frac{w_{in}(\tau)}{2} \mathrm{sign}\{Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})(\boldsymbol{\beta} + \delta\underline{a})\}] \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}). \end{aligned}$$

Consider each individual term and as $\delta \to 0$, the first term vanishes, and the second term when divided by δ converges to $\psi'(\boldsymbol{\beta};\underline{a})$ defined in (A.8).

Now since $\psi(\boldsymbol{\beta})$ is convex in $\boldsymbol{\beta}$, it attains a (local) minimum at $\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x})$ if and only if $\psi'(\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x});\underline{a}) > 0$ for all $\underline{a} \neq \underline{0}$. Note that at $\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}) = [DX_n(\mathbf{x},\mathbf{h})]^{-1}VY_n(\mathbf{x},\mathbf{h})$,

$$\begin{aligned}\psi'(\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x});\underline{a}) &= \sum_{i\in\mathbf{h}}[\frac{w_{in}(\tau)}{2} + \frac{w_{in}(\tau)}{2}\text{sign}(\boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\underline{a}) - \tau]\boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\underline{a} \\ &+ \sum_{i\in\bar{\mathbf{h}}}\Big\{\frac{w_{in}(\tau)}{2} - \frac{w_{in}(\tau)}{2}\text{sign}^*[Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}); -\boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\underline{a}] - \tau\Big\}\boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\underline{a}.\end{aligned}$$

Letting $\underline{\nu} = (\nu_i; i\in\mathbf{h}) = [DX_n(\mathbf{x},\mathbf{h})]\underline{a}$, we have that $\psi'(\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x});\underline{a}) > 0$ for all $\underline{a}\neq\underline{0}$, if and only if

$$\begin{aligned}0 < \ &\sum_{i\in\mathbf{h}}\Big[(\frac{w_{in}(\tau)}{2} - \tau)\nu_i + \frac{w_{in}(\tau)}{2}|\nu_i|\Big] + \sum_{i\in\bar{\mathbf{h}}}\Big\{\frac{w_{in}(\tau)}{2} - \frac{w_{in}(\tau)}{2}\text{sign}^*\Big[Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}); \\ &-\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})[DX_n(\mathbf{x},\mathbf{h})]^{-1}\underline{\nu}\Big] - \tau\Big\} \times \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})^\top[DX_n(\mathbf{x},\mathbf{h})]^{-1}\underline{\nu},\end{aligned}$$

for all $\underline{\nu}\neq 0$. This is equivalent to

$$\begin{aligned}\tau\mathbf{1}_{n(A)} - \underline{\mathbf{w}}_n(\mathbf{h}) < &\sum_{i\in\bar{\mathbf{h}}}\Big\{\frac{w_{in}(\tau)}{2} - \frac{w_{in}(\tau)}{2}\text{sign}^*\Big[Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}); -\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})[DX_n(\mathbf{x},\mathbf{h})]^{-1}\underline{\nu}\Big] \\ &-\tau\Big\} \times [DX_n(\mathbf{x},\mathbf{h})]^{-1}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) < \tau\mathbf{1}_{n(A)},\end{aligned} \tag{A.9}$$

for all $\underline{\nu}\neq 0$. As once given $\mathbf{h}$ thus $\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x})$, observations $\{(Y_i,\mathbf{X}_i) : i\in\bar{\mathbf{h}}\}$ behave like independent random vectors, consequently according to [FACT2], the event that for some $i\in\bar{\mathbf{h}}$,

$$Y_i - \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})^\top\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}) = Y_i - \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})[DX_n(\mathbf{x},\mathbf{h})]^{-1}VY_n(\mathbf{x},\mathbf{h}) = \underline{\mathbf{0}}$$

has probability zero, since it means at least $n(A)+1$ independent random vectors lie on a certain hyperplane of dimension $n(A)$. Therefore, (A.9) reduces to

$$\begin{aligned}\tau\mathbf{1}_{n(A)} - \underline{\mathbf{w}}_n(\mathbf{h}) < &\sum_{i\in\bar{\mathbf{h}}}\Big\{\frac{w_{in}(\tau)}{2} - \frac{w_{in}(\tau)}{2}\text{sign}\Big[Y_i - \boldsymbol{\mu}_n^\top(\mathbf{X}_{i\mathbf{x}})\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x})\Big] \\ &-\tau\Big\} \times [DX_n(\mathbf{x},\mathbf{h})]^{-1}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) < \tau\mathbf{1}_{n(A)},\end{aligned}$$

as required. □

Below are some classical results in kernel smoothing which will be repeatedly referred to throughout this section: with probability one,

$$\sup_{\mathbf{x}}|(n\delta_n^p)^{-1}N_n(\mathbf{x}) - f_{\mathbf{X}}(\mathbf{x})| = o(1), \tag{A.10}$$

$$\sup_{\mathbf{x},\tau}|\Sigma_{n\tau}(\mathbf{x}) - [1-G(Q_\tau(\mathbf{x})|\mathbf{x})]f_0(Q_\tau(\mathbf{x})|\mathbf{x}\Sigma(A)| = O(\delta_n + (nh_n^p/\log n)^{-1/2}), \tag{A.11}$$

$$\sup_{\mathbf{x}}|\tilde{\Sigma}_n(\mathbf{x}) - f_{\mathbf{X}}(\mathbf{x})\Sigma(\tilde{A})| = O\Big((nh_n^p/\log n)^{-1/2} + h_n\Big). \tag{A.12}$$

Also for notational simplicity, we shall write $F_0(.|\mathbf{X}_i)$, $f_0(.|\mathbf{X}_i)$, $G(.|\mathbf{X}_i)$, $g(.|\mathbf{X}_i)$ as $F_{i0}(.)$, $f_i(.)$, $G_i(.)$ and $g_i(.)$, respectively. For any given $\mathbf{x} \in R^p$, $\boldsymbol{\beta} \in R^{n(A)}$ and $F(.|\mathbf{x})$ a conditional distribution function defined with any $\mathbf{x} \in \mathcal{D}$, write

$$
\begin{aligned}
&Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F|\tau) = \Big[w_i(\tau|F)I\{Y_i \leq \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\} - \tau\Big]\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}),\ i \in S_n(\mathbf{x}),\\
&\tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F|\tau) = E_i[Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F|\tau)],
\end{aligned}
$$

where $E_i(.)$ denotes expectation taken with respect to the joint distribution of $\xi_i = (Y_i, \mathbf{X}_i, d_i)$.

Regarding the convergence rate of $\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x})$ that is uniform in $\mathbf{x} \in \mathcal{D}$ and $\tau \in [\tau^*, 1-\tau^*]$, we have

Lemma 3. *Under conditions in Theorem 1, we have*

$$
\sup_{\mathbf{x},\tau} |\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}) - \boldsymbol{\beta}_{n\tau}(\mathbf{x})| = O_p(\tau_n). \tag{A.13}
$$

Proof. First of all, according to Proposition 2 there exists some finite constant ϕ_1 such that

$$
\sup_{\mathbf{x},\tau} |\sum_{i \in S_n(\mathbf{x})} Z_{ni}(\mathbf{x}, \hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}), \hat{F}_n(.|.)|\tau)| \leq \phi_1. \tag{A.14}
$$

Secondly, according to the first assertion of Proposition 7, if $\tilde{\tau}_n^{1-\alpha}/\{\delta_n^{2p\alpha}\log n\} < \infty$, then

$$
\sup_{\mathbf{x},\tau,\beta} |\sum_{i \in S_n(\mathbf{x})} [Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, \hat{F}_n(.|.)|\tau) - Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)]| = O_p((n\delta_n^p \log n)^{1/2}) \tag{A.15}
$$

On the other hand, based on Proposition 4 below, there exists positive constants ϵ_1^*, c_1^* and M_2^*, such that

$$
\Big|\tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)\Big]\Big| \geq \min(\epsilon_1^*, c_1^*|\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})|) \tag{A.16}
$$

uniformly in $\mathbf{x}, \tau$ and $\boldsymbol{\beta}$ whenever

$$
|\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})| \geq M_3^*[n\delta_n^p/\log n]^{-1/2},
$$

where $M_3^* > M_2^*$.

Combining (A.14), (A.15), (A.16) and the facts that $\sharp(\mathbf{h}) = n(A)$, $\min_{\mathbf{x},\tau}(n\delta_n^p)^{-1}\sharp(\bar{\mathbf{h}}) > 0$ almost surely, we can conclude that for large enough $K_1 (> M_2^*)$,

$$
\begin{aligned}
&\limsup_n Prob\{\sup_{\mathbf{x},\tau} |\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}) - \boldsymbol{\beta}_{n\tau}(\mathbf{x})| \geq K_1[n\delta_n^p/\log n]^{-1/2}\}\\
&\leq \lim_n Prob\{\sup_{\mathbf{x},\boldsymbol{\beta}} |\sum_i I(|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n)[Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau) - \tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)]| \geq \frac{K_1}{2}(n\delta_n^p \log n)^{1/2}\}.
\end{aligned}
$$

Finally, we invoke Theorem II.37 of Pollard (1984) to show that for large enough K_1, the probability on the right hand side of the above is of $o(1)$ as $n \to \infty$. To this aim, first note that the following classes of functions $(i) : \{\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) : \mathbf{x} \in \mathcal{D}\}$, $(ii) : \{I(Y_i \leq \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})) : \boldsymbol{\beta} \in R^{n(A)},\ \mathbf{x} \in \mathcal{D}\}$, $(iii) : \{I(|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n) : \mathbf{x} \in \mathcal{D}\}$ and $(iv)\{w_i(\tau|F_0) : \tau \in [\tau^*, 1-\tau^*]\}$ are all Euclidean for a constant envelope (Lemma (2.13) of Pakes and Pollard, 1989; Lemma 18 and Lemma 22 of Nolan and Pollard, 1987). The closer properties of Euclidean classes further dictates that $\{I(|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n) Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau) :\ \boldsymbol{\beta} \in R^{n(A)},\ \mathbf{x} \in \mathcal{D}, \tau \in [\tau^*, 1-\tau^*]\}$ is also Euclidean, thus the conditions required by Theorem II.37 of Pollard (1984) are met. The proof is thus complete by noting that $E[I(|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n) Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)] = O(\delta_n^p)$ uniformly in $\boldsymbol{\beta} \in R^{n(A)},\ \mathbf{x} \in \mathcal{D}$ and $\tau \in [\tau^*, 1-\tau^*]$. □

As the random vector $Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F|\tau)$ plays a central role in the proof of Theorem 1, let us take a look at its expectation. Observing that

$$\begin{aligned} & w_i(\tau|F) I\{Y_i \leq \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\} - \tau \\ = \; & I\{C_i > \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}), T_i \leq \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\} + I\{C_i \leq \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}), T_i \leq C_i\} \\ & + I\{C_i \leq \boldsymbol{\beta}^\top \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}), T_i > C_i\} \Big[1 - \frac{1-\tau}{1-F_i(C_i)} I\{F_i(C_i) < \tau\}\Big] - \tau, \end{aligned} \tag{A.17}$$

$$\begin{aligned} E\{I(C_i > t, T_i < t)|\mathbf{X}_i = \mathbf{x}\} &= [1 - G(t|\mathbf{x})] F_0(t|\mathbf{x}), \\ E\{I(C_i > t, T_i < C_i)|\mathbf{X}_i = \mathbf{x}\} &= \int_{-\infty}^{t} F_0(u|\mathbf{x}) g(u|\mathbf{x}) du, \end{aligned}$$

we have

$$\tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F|\tau) = \int_{[-1,1]^p} \boldsymbol{\mu}(\mathbf{t}) \{R_1(\boldsymbol{\beta}^\top \boldsymbol{\mu}(\mathbf{t})|\mathbf{x} + \delta_n \mathbf{t}) + R_F(\boldsymbol{\beta}^\top \boldsymbol{\mu}(\mathbf{t})|\mathbf{x} + \delta_n \mathbf{t}, \tau) - \tau\} \omega_{\delta_n}(\mathbf{t}, \mathbf{x}) d\mathbf{t},$$

where

$$\begin{aligned} R_1(s|\mathbf{x}) &= [1 - G(s|\mathbf{x})] F_0(s|\mathbf{x}) + \int_{-\infty}^{s} F_0(u|\mathbf{x}) g(u|\mathbf{x}) du, \\ R_F(s|\mathbf{x}, , \tau) &= \int_{-\infty}^{s} \{1 - F_0(u|\mathbf{x})\} \Big[1 - \frac{1-\tau}{1-F(u|\mathbf{x})} I\{F(u|\mathbf{x}) < \tau\}\Big] g(u|\mathbf{x}) du. \end{aligned}$$

In the case where $F(.|.) = F_0(.|.)$, we have

$$R_1(s|\mathbf{x}) + R_{F_0}(s|\mathbf{x}, \tau) = [1 - G(s|\mathbf{x})] F_0(s|\mathbf{x}) + G(s|\mathbf{x}) - (1-\tau) \int_{-\infty}^{\min(s, Q_\tau(\mathbf{x}))} g(u|\mathbf{x}) du.$$

Proof of Theorem 1. This consists of the following steps.

Step 1: For any given $\mathbf{t} \in [-1,1]^p$, and $\boldsymbol{\beta}$ close enough to $\boldsymbol{\beta}_{n\tau}(\mathbf{x})$, let $s_1 = \boldsymbol{\mu}(\mathbf{t})^\top \boldsymbol{\beta}_{n\tau}(\mathbf{x})$ and $s_2 = \boldsymbol{\mu}(\mathbf{t})^\top \boldsymbol{\beta}$. Bear in mind the observation that for C_i lying above $Q_\tau(\mathbf{X}_i)$, the quantile fit will not be affected if the entire weight is shifted to $Y^{+\infty}$; see also Wang and Wang (2009). Therefore, assumptions [A2]-[A4] and [FACT3] imply that

$$\begin{aligned} & R_1(s_2|\mathbf{x}) + R_{F_0}(s_2|\mathbf{x},\tau) - R_1(s_1|\mathbf{x}) - R_{F_0}(s_1|\mathbf{x},\tau) \\ = \;& [1 - G(Q_\tau(\mathbf{x})|\mathbf{x})] f_0(Q_\tau(\mathbf{x})|\mathbf{x}) \boldsymbol{\mu}(\mathbf{t})^\top \{\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})\} + O(\tau_n^2 + |\delta_n|^{2s_2}), \end{aligned} \tag{A.18}$$

and thus for any $\boldsymbol{\beta}$ close enough to $\boldsymbol{\beta}_{n\tau}(\mathbf{x})$,

$$\begin{aligned} & \tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau) - \tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}_{n\tau}(\mathbf{x}), F_0|\tau) \\ = \;& \Sigma_n(\mathbf{x})\{\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})\} + O(|\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})|^2 + |\delta_n|^{2s_2}), \end{aligned} \tag{A.19}$$

uniformly in $\mathbf{t}$, $\mathbf{x}$ and τ.

Step 2: Define the following $n(A)-$dimensional random vector

$$\begin{aligned} \chi_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau) = & \sum_i I(|\mathbf{X}_{i\mathbf{x}}| \le \delta_n)[Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau) - \tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)] \\ & - \sum_i I(|\mathbf{X}_{i\mathbf{x}}| \le \delta_n)[Z_{ni}(\mathbf{x}, \boldsymbol{\beta}_{n\tau}(\mathbf{x}), F_0|\tau) - \tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}_{n\tau}(\mathbf{x}), F_0|\tau)]. \end{aligned}$$

As argued in the last paragraph in the proof of Lemma 3, $\{I(|\mathbf{X}_{i\mathbf{x}}| \le \delta_n) Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau) : \boldsymbol{\beta} \in R^{n(A)},\ \mathbf{x} \in \mathcal{D}, \tau \in (0,1)\}$ is Euclidean, hence the conditions required by Theorem II.37 of Pollard (1984) are met. As $E[I(|\mathbf{X}_{i\mathbf{x}}| \le \delta_n)\{Z_{ni}(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau) - Z_{ni}(\mathbf{x}, \boldsymbol{\beta}_{n\tau}(\mathbf{x}), F_0|\tau)\}]^2 = O(\delta_n^p \tau_n)$ uniformly in $\mathbf{x}, \tau$ and $\boldsymbol{\beta} \in R^{n(A)}$, whenever $|\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})| \le K_1 \tau_n$, there exists some finite $K_1 > 0$, such that with probability one,

$$\sup_{\mathbf{x}, \tau, |\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})| \le K_1 \tau_n} |\chi_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)| = O((n\delta_n^p \tau_n \log n)^{1/2}). \tag{A.20}$$

Step 3: Combining (A.19) and (A.20), we have

$$\begin{aligned} & \frac{1}{N_n(\mathbf{x})} \sum_{i \in S_n(\mathbf{x})} \Big[w_{i0} I\{Y_i \le Q_{n\tau}(\mathbf{X}_i, \mathbf{x})\} - \tau \Big] \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) \\ = \;& \frac{1}{N_n(\mathbf{x})} \chi_n(\mathbf{x}) + \tilde{H}_n(\mathbf{x}, \hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}), F_0|\tau) - \tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}_{n\tau}(\mathbf{x}), F_0|\tau) \\ & + \frac{1}{N_n(\mathbf{x})} \sum_{i \in S_n(\mathbf{x})} [w_{i0}(\tau) I\{Y_i \le \hat{Q}_{n\tau}(\mathbf{X}_i, \mathbf{x})\} - \tau] \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) \\ = \;& \Sigma_n(\mathbf{x})[\hat{\boldsymbol{\beta}}_{n\tau}(\mathbf{x}) - \boldsymbol{\beta}_{n\tau}(\mathbf{x})] + \frac{1}{N_n(\mathbf{x})} \sum_{i \in S_n(\mathbf{x})} \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})[w_{in}(\tau) I\{Y_i \le \hat{Q}_{n\tau}(\mathbf{X}_i, \mathbf{x})\} - \tau] \\ & + \frac{1}{N_n(\mathbf{x})} \sum_{i \in S_n(\mathbf{x})} \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})(w_{i0}(\tau) - w_{in}(\tau)) \Big[I\{Y_i \le \hat{Q}_{n\tau}(\mathbf{X}_i, \mathbf{x})\} - I\{Y_i \le Q_{n\tau}(\mathbf{X}_i, \mathbf{x})\} \Big] \\ & + \frac{1}{N_n(\mathbf{x})} \sum_{i \in S_n(\mathbf{x})} \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})(w_{i0}(\tau) - w_{in}(\tau)) I\{Y_i \le Q_{n\tau}(\mathbf{X}_i, \mathbf{x})\} + O(\tau_n^{3/4} + \delta_n^{2s_2}). \end{aligned}$$

The assertion in Theorem 1 thus follows from Propositions 7, 2 and 5. □

Proof of Corollary 1. We only need to show that the difference between the covariance matrices of these two estimators is positive definite. For illustration purposes, we focus on the covariance matrices of their respective 'staple' terms, not in the least because as discussed in Section 4, the 'correction' terms will become relatively negligible if the K-M estimator converges fast enough.

To this aim, we start with quantifying the variances of the following two random variables.

$$(A)\ \frac{d_i}{1-G(Y_i|\mathbf{X}_i)}[\tau - I\{Y_i \le Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\}] \qquad (B)\ w_{i0}(\tau)I\{Y_i \le Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\} - \tau.$$

Given $\mathbf{X}_i$, the second moments of term (A) is in the first order equal to

$$E\{[\tau - I\{T_i \le Q_\tau(\mathbf{X}_i)\}]^2/(1-G(T_i|\mathbf{X}_i))\}.$$

As for that of term (B), first note that

$$\begin{aligned} B^2 &= (1-\tau)^2 I\{C_i > T_i, T_i < Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\} + \tau^2 I\{C_i > T_i, T_i > Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\} \\ &\quad + \tau^2 I\{C_i < T_i, C_i > Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\} \\ &\quad + (1-\tau)^2 \frac{F_i^2(C_i|\mathbf{X}_i)}{[1-F_i(C_i|\mathbf{X}_i)]^2} I\{C_i < T_i, C_i < Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\} \\ &= B_1 + B_2 + B_3 + B_4 \end{aligned}$$

As $B_2+B_3 = \tau^2 I\{C_i > Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\}I\{T_i > Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\}$, we have $E(B_2+B_3) = \tau^2(1-\tau)[1-G(Q_\tau(\mathbf{X}_i)|\mathbf{X}_i)] + o(1)$. For the remaining two terms,

$$B_1 = (1-\tau)^2 I\{T_i < Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\} - (1-\tau)^2 I\{C_i < T_i, T_i < Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\}$$

and consequently, $E(B_1+B_4) = (1-\tau)^2\tau + (1-\tau)^2 E\Big\{I\{C_i \le Q_\tau(\mathbf{X}_i)\}\frac{F_0(C_i|\mathbf{X}_i)}{1-F_0(C_i|\mathbf{X}_i)}\Big\} + o(1)$. Therefore,

$$EB^2 = \tau(1-\tau)[1-G(Q_\tau(\mathbf{X}_i)|\mathbf{X}_i)] + (1-\tau)^2 E\Big[I\{C_i \le Q_\tau(\mathbf{X}_i)\}\frac{F_0(C_i|\mathbf{X}_i)}{1-F_0(C_i|\mathbf{X}_i)}\Big] + o(1).$$

Consequently, up to the first order we have

$$\begin{aligned} \mathrm{Cov}(\boldsymbol{\beta}_{n\tau}(\mathbf{x})) &= \frac{\Sigma^{-1}(A)]\tilde{\Sigma}(A)\Sigma^{-1}(A)}{\delta_n^p f_0^2(Q_\tau(\mathbf{x})|\mathbf{x})}\{1-G(Q_\tau(\mathbf{x})|\mathbf{x})\}^{-2} \\ &\quad \times\{\tau(1-\tau)[1-G(Q_\tau(\mathbf{X}_i)|\mathbf{X}_i)] + (1-\tau)^2 E\Big[I\{C_i \le Q_\tau(\mathbf{X}_i)\}\frac{F_0(C_i|\mathbf{X}_i)}{1-F_0(C_i|\mathbf{X}_i)}\Big]\} \\ &\le \frac{\Sigma^{-1}(A)]\tilde{\Sigma}(A)\Sigma^{-1}(A)}{\delta_n^p f_0^2(Q_\tau(\mathbf{x})|\mathbf{x})}\frac{\tau(1-\tau)}{\{1-G(Q_\tau(\mathbf{x})|\mathbf{x})\}^2}, \\ \mathrm{Cov}(\hat{\mathbf{c}}_{n\tau}(\mathbf{x})) &= \frac{\Sigma^{-1}(A)]\tilde{\Sigma}(A)\Sigma^{-1}(A)}{\delta_n^p f_0^2(Q_\tau(\mathbf{x})|\mathbf{x})} E\Big\{\frac{[\tau - I\{T_i \le Q_\tau(\mathbf{X}_i)\}]^2}{1-G(T_i|\mathbf{X}_i)}\Big\} \end{aligned}$$

This finishes the proof. $\square$

Proposition 4. *If $\delta_n^{s_2} = O(\tau_n)$, then there exist positive constants ϵ_1, c_1 and M_1, such that for all $\mathbf{x}$ and τ,*

$$|\tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)| \geq min(\epsilon_1, c_1|\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})|) \tag{A.21}$$

whenever $|\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})| \geq M_1(n\delta_n^p/\log n)^{-1/2}$.

Proof. This is split into several steps.

Steps 1: We show that there exist constants $M_2 > 0$ and $\epsilon_1 > 0$, such that for all $\mathbf{x}$ and τ,$|\tilde{H}_n(\mathbf{x}, \boldsymbol{\beta}, F_0|\tau)| \geq \epsilon_1$, whenever $|\boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})| \geq M_2$.

If this is false, we can construct three sequences of vectors $\{\boldsymbol{\beta}_{n^*}\}$ in $R^{n(A)}$, $\{\mathbf{x}_{n^*}\}$ in R^p, $\{\tau_{n^*}\}$ such that with $\Delta_{n^*} \equiv \boldsymbol{\beta}_{n^*} - \boldsymbol{\beta}_{n0}(\mathbf{x}_{n^*})$ and $n^* \to \infty$, we have $|\Delta_{n^*}| \to \infty$ and $\tilde{H}_n(\mathbf{x}_{n^*}, \boldsymbol{\beta}_{n^*}, F_0|\tau_{n*}) \to 0$. Without loss of generality, we can assume $\Delta_{n^*}/\|\Delta_{n^*}\| \to$ some $\Delta^* \in R^{n(A)}$, such that $\|\Delta^*\| = 1$, $\mathbf{x}_{n^*} \to$ some $\mathbf{x}^0 \in \mathcal{D}$ and $\tau_{n*} \to$ some $\tau^0 \in [\tau^*, 1-\tau^*]$.

As $\boldsymbol{\mu}(\mathbf{t})^\top \boldsymbol{\beta}_{n\tau}(\mathbf{x})$ is uniformly bounded over $\mathbf{t} \in [-1,1]^p$, $\mathbf{x} \in \mathcal{D}$ and $\tau \in [\tau^*, 1-\tau^*]$, we have for any given $\mathbf{t} \in [-1,1]^p$, $\boldsymbol{\mu}(\mathbf{t})^\top \boldsymbol{\beta}_{n^*}$ must tend to either $+\infty$ or $-\infty$, depending on whether $\boldsymbol{\mu}(\mathbf{t})^\top \Delta^*$ is positive or negative, respectively. Specifically, for those $\mathbf{t} \in [-1,1]^p$ such that $\boldsymbol{\mu}(\mathbf{t})^\top \Delta^* < 0$,

$$R_1(\boldsymbol{\mu}(\mathbf{t})^\top \boldsymbol{\beta}_{n^*}|\mathbf{x}_{n^*} + \delta_{n^*}\mathbf{t}) + R_{F_0}(\boldsymbol{\mu}(\mathbf{t})^\top \boldsymbol{\beta}_{n^*}|\mathbf{x}_{n^*} + \delta_{n^*}\mathbf{t}, \tau_{n*}) \to 0, \tag{A.22}$$

as $n^* \to \infty$; while for those $\mathbf{t} \in [-1,1]^p$, such that $\boldsymbol{\mu}(\mathbf{t})^\top \Delta^* > 0$, we have as $n^* \to \infty$,

$$\begin{aligned} R_1(\mathbf{t}^\top \boldsymbol{\beta}_{n^*}|\mathbf{x}_{n^*} + \delta_{n^*}\mathbf{t}) + R_F(\mathbf{t}^\top \boldsymbol{\beta}_{n^*}|\mathbf{x}_{n^*} + \delta_{n^*}\mathbf{t}, \tau_{n*}) - \tau_{n*} \\ \to (1-\tau_0)\{1 - G(Q_{\tau^0}(\mathbf{x}^0)|\mathbf{x}^0)\}. \end{aligned} \tag{A.23}$$

Since the region $[-1,1]^p \cap \{\mathbf{t} : \boldsymbol{\mu}(\mathbf{t})^\top \Delta^* = 0\}$ must have Legesque measure zero, if $|\tilde{H}_n(\boldsymbol{\beta}_{n^*}, F, \mathbf{x}_{n^*})| \to 0$, as $n^* \to \infty$, (A.22), (A.23) and a straightforward application of the dominated convergence theorem then yield

$$\tau^0 \int_{[-1,1]^p \cap \{\mathbf{t}^\top \Delta^* < 0\}} \boldsymbol{\mu}(\mathbf{t}) d\mathbf{t} = (1-\tau^0)\{1 - G(Q_\tau(\mathbf{x}^0)|\mathbf{x}^0)\} \int_{[-1,1]^p \cap \{\mathbf{t}^\top \Delta^* > 0\}} \boldsymbol{\mu}(\mathbf{t}) d\mathbf{t}.$$

Multiplying either side by Δ^*, we get

$$\tau^0 \int_{[-1,1]^p \cap \{\boldsymbol{\mu}(\mathbf{t})^\top \Delta^* < 0\}} \boldsymbol{\mu}(\mathbf{t})^\top \Delta^* d\mathbf{t} = (1-\tau^0)\{1 - G(Q_{\tau^0}(\mathbf{x}^0)|\mathbf{x}^0)\} \int_{[-1,1]^p \cap \{\boldsymbol{\mu}(\mathbf{t})^\top \Delta^* > 0\}} \boldsymbol{\mu}(\mathbf{t})^\top \Delta^* d\mathbf{t}.$$

As $\tau^0 \in [\tau^*, 1-\tau^*]$ and whence $G(Q_{\tau^0}(\mathbf{x}^0)|\mathbf{x}^0) < 1$, the above equality implies that the two regions $[-1,1]^p \cap \{\boldsymbol{\mu}(\mathbf{t})^\top \Delta^* < 0\}$ and $[-1,1]^p \cap \{\boldsymbol{\mu}(\mathbf{t})^\top \Delta^* > 0\}$ must both have Legesque measure zero. This contradicts [FACT1].

Step 2: Note that for any $s \in R$ and $\mathbf{x} \in R^p$,

$$R_1(s|\mathbf{x}) + R_{F_0}(s|\mathbf{x},\tau) - \tau = \{1 - G(s|\mathbf{x})\}\{F_0(s|\mathbf{x}) - \tau\} + (1-\tau)\{G(s|\mathbf{x}) - \int_{-\infty}^{\min(s, Q_\tau(\mathbf{x}))} g(u|\mathbf{x})du\}. \quad \text{(A.24)}$$

For any given $\boldsymbol{\beta} \in R^{n(A)}$, write $\Delta \equiv \boldsymbol{\beta} - \boldsymbol{\beta}_{n\tau}(\mathbf{x})$, and thus for any given $\mathbf{t} \in [-1,1]^p$, $\boldsymbol{\beta}^\top \boldsymbol{\mu}(\mathbf{t}) = \Delta^\top \boldsymbol{\mu}(\mathbf{t}) + Q_\tau(\mathbf{x} + \mathbf{t}\delta_n) - r(\mathbf{t}\delta_n, \mathbf{x})$, and

$$F_0(\boldsymbol{\beta}^\top \boldsymbol{\mu}(\mathbf{t})|\mathbf{x} + \delta_n \mathbf{t}) - \tau = h(\Delta, \mathbf{x}, \mathbf{t}, \tau)\Delta^\top \boldsymbol{\mu}(\mathbf{t}) + O(\delta_n^{s_2}) \quad \text{(A.25)}$$

uniformly in $\mathbf{t}, \mathbf{x}, \tau$ and Δ, where

$$h(\Delta, \mathbf{x}, \mathbf{t}, \tau) = \frac{F_0(\Delta^\top \boldsymbol{\mu}(\mathbf{t}) + Q_\tau(\mathbf{x} + \mathbf{t}\delta_n) - r(\mathbf{t}\delta_n, \mathbf{x})|\mathbf{x} + \delta_n \mathbf{t}) - F_0(Q_\tau(\mathbf{x} + \mathbf{t}\delta_n) - r(\mathbf{t}\delta_n, \mathbf{x})|\mathbf{x} + \delta_n \mathbf{t})}{\Delta^\top \boldsymbol{\mu}(\mathbf{t})}$$

if $\Delta^\top \boldsymbol{\mu}(\mathbf{t}) \neq 0$, and is defined arbitrarily otherwise, since for nonzero Δ, the set $[-1,1]^p \cap \{\mathbf{t} : \boldsymbol{\mu}(\mathbf{t})^\top \Delta = 0\}$ has Lebesque measure zero. In view of [A3], there exists $b > 0$, such that the density $f_0(.|\mathbf{x})$ is continuous and bounded away from zero in the interval $[Q_\tau(\mathbf{x}) - b, Q_\tau(\mathbf{x}) + b]$ uniformly in $\mathbf{x}$ and $\tau \in [\tau^*, 1-\tau^*]$. Suppose s_2 in [A2] and the bandwidth δ_n are such that $|r(\mathbf{t}\delta_n, \mathbf{x}| = O(\delta_n^{s_2}) < b$ for all $\mathbf{x}$ and $\mathbf{t} \in [-1,1]^p$. Therefore, there exists some $M_3 > 0$, such that

$$\inf_{\substack{|\Delta| \leq M_2, \mathbf{t} \in [-1,1]^p \\ \mathbf{x}, \tau \in [\tau^*, 1-\tau^*]}} h(\Delta, \mathbf{x}, \mathbf{t}) > M_3. \quad \text{(A.26)}$$

Similar to (A.25), it could be shown that

$$\begin{aligned} & G(\boldsymbol{\beta}^\top \boldsymbol{\mu}(\mathbf{t})|\mathbf{x} + \mathbf{t}\delta_n) - Pr\{C_i < \min[Q_\tau(\mathbf{x} + \mathbf{t}\delta_n), \boldsymbol{\beta}^\top \boldsymbol{\mu}(\mathbf{t})]|\mathbf{x} + \delta_n \mathbf{t}\} \\ = \; & I\{\Delta^\top \boldsymbol{\mu}(\mathbf{t}) > r(\mathbf{t}\delta_n, \mathbf{x})\}\tilde{h}(\Delta, \mathbf{x}, \mathbf{t}, \tau)\Delta^\top \boldsymbol{\mu}(\mathbf{t}) + O(\delta_n^{s_4}), \end{aligned}$$

again uniformly in $\mathbf{t}, \mathbf{x}, \tau \in [\tau^*, 1-\tau^*]$ and Δ, where

$$\tilde{h}(\Delta, \mathbf{x}, \mathbf{t}) = \frac{G(\Delta^\top \boldsymbol{\mu}(\mathbf{t}) + Q_\tau(\mathbf{x} + \mathbf{t}\delta_n) - r(\mathbf{t}\delta_n, \mathbf{x})|\mathbf{x} + \delta_n \mathbf{t}) - G(Q_\tau(\mathbf{x} + \mathbf{t}\delta_n) - r(\mathbf{t}\delta_n, \mathbf{x})|\mathbf{x} + \delta_n \mathbf{t})}{\Delta^\top \boldsymbol{\mu}(\mathbf{t})},$$

which is always nonnegative.

Step 3: Write $\hat{Q}_{n\tau}(\mathbf{x}+\mathbf{t}\delta_n,\mathbf{x}) = \hat{\beta}_{n\tau}^{\top}(\mathbf{x})\boldsymbol{\mu}(\mathbf{t})$. Combining the results in *Step 2*, we have

$$\begin{aligned}
&\int_{[-1,1]^p}\boldsymbol{\mu}(\mathbf{t})\Big[\{1-G(s|\mathbf{x}+\delta_n\mathbf{t})\}\{F_0(s|\mathbf{x}+\delta_n\mathbf{t})-\tau\} \\
&\qquad\qquad +(1-\tau)\{G(s|\mathbf{x}+\delta_n\mathbf{t})-Pr\{C_i<\min(Q_\tau(\mathbf{x}+\delta_n\mathbf{t}),s)|\mathbf{x}\}\}\Big]d\mathbf{t} \\
\geq\ & \Big[\int_{[-1,1]^p}\{1-G(s|\mathbf{x}+\delta_n\mathbf{t})\}h(\Delta,\mathbf{x},\mathbf{t},\tau)\boldsymbol{\mu}(\mathbf{t})\boldsymbol{\mu}(\mathbf{t})^{\top}d\mathbf{t}\Big]\Delta + O(\delta_n^{s_4}) \\
\geq\ & M_3(1-b_2)\Big[\int I\{|\boldsymbol{\mu}(\mathbf{t})|\leq b_1/2\}\boldsymbol{\mu}(\mathbf{t})\boldsymbol{\mu}(\mathbf{t})^{\top}d\mathbf{t}\Big]\Delta, \qquad\qquad (A.27)
\end{aligned}$$

where the last inequality(A.27) follows from [A4] and (A.26). Denote by $\lambda_1(>0)$, the smallest eigenvalue of the positive definite matrix

$$\int I\{|\boldsymbol{\mu}(\mathbf{t})|\leq b_1/2\}\boldsymbol{\mu}(\mathbf{t})\boldsymbol{\mu}(\mathbf{t})^{\top}d\mathbf{t}.$$

It can be seen that

$$|\tilde{H}_n(\mathbf{x},\boldsymbol{\beta},F_0|\tau)| \geq \lambda_1 M_3(1-b_2)|\boldsymbol{\beta}-\boldsymbol{\beta}_{n\tau}(\mathbf{x})|;$$

this together with the conclusion in *Step 1* leads to (A.21). □

Proposition 5. *Suppose conditions in Theorem 1 hold. We have*

$$\begin{aligned}
&\frac{1}{N_n(\mathbf{x})}\sum_{i\in S_n(\mathbf{x})}\Big[\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})(w_{i0}(\tau)-w_{in}(\tau))I\{Y_i\leq Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\} \\
&\qquad -E_i\{\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})(w_{i0}(\tau)-w_{in}(\tau))I\{Y_i\leq Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\}\}\Big] = O_p\Big\{\tau_n\Big(\frac{\tilde{\tau}_n^{1-\alpha}}{\delta_n^{p\alpha}\log n}\Big)^{1/2}\Big\}
\end{aligned}$$

uniformly in $\mathbf{x},\tau$*, where* $E_i(.)$ *stands for expectation taken with respect to the joint distribution of* $(\mathbf{X}_i,Y_i)$ *with the other argument held fixed.*

The proof of Proposition 5 is based on the concept of stochastic equicontinuity in empirical process. To start with, we introduce the definition of $\epsilon-$covering number. Let $\mathbf{S}$ be a sample space and $\mathcal{S}$, the sigma field. Let P be a probability measure on $\mathcal{S}$ and $\mathcal{F}$ be a class of measurable functions from $\mathbf{S}$ to R^d. Denote by $(\mathcal{F},\|.\|_\infty)$ and $(\mathcal{F},\|.\|_{L_2(P)})$, the subset of metric spaces equipped with different norms, the first with the supremum norm and the second, the $L_2(P)$ norm. The $\epsilon-$covering number $N(\epsilon,\|.\|_{L_2(P)},\mathcal{F})$ is defined to be the minimum number of balls of radius ϵ with respect to the $L_2(P)$ norm needed to cover the set $\mathcal{F}$. $N(\epsilon,\|.\|_\infty,\mathcal{F})$ is similarly defined.

Proof. A useful identity is that, for given $\mathbf{x}$, τ, and a generic conditional distribution function $F(.|.)$,

$$\begin{aligned}
&(w_{i0}(\tau)-w_{in}(\tau))I\{Y_i\leq Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) \\
&=(1-\tau)[Z_{n1}(\xi_i|\mathbf{x},F_0,\tau)-Z_{n1}(\xi_i|\mathbf{x},F,\tau)+Z_{n2}(\xi_i|\mathbf{x},F,\tau)],\ i\in S_n(\mathbf{x}), \qquad (A.28)
\end{aligned}$$

where $w_{in}(\tau) \equiv w_i(\tau|F(.|.))$,

$$Z_{n1}(\xi_i|\mathbf{x}, F, \tau) = \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) I\{|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n\} I\{C_i \leq Q_{n\tau}(\mathbf{X}_i, \mathbf{x}), T_i > C_i\} \frac{I\{F_i(C_i) \leq \tau\}}{1 - F_{i0}(C_i)},$$

$$\begin{aligned} Z_{n2}(\xi_i|\mathbf{x}, F, \tau) &= I\{|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n\} I\{C_i \leq Q_{n\tau}(\mathbf{X}_i, \mathbf{x}), T_i > C_i, F_i(C_i) \leq \tau\} \\ &\quad \times \frac{F_{i0}(C_i) - F_i(C_i)}{(1 - F_{i0}(C_i))(1 - F_i(C_i))} \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}). \end{aligned}$$

We start with proving

$$\begin{aligned} \sup_{F(.|.) \in \mathcal{F}_n} \Big| \frac{1}{n\delta_n^p} \sum_{i=1}^n \Big[& Z_{n1}(\xi_i|\mathbf{x}, F, \tau) - Z_{n1}(\xi_i|\mathbf{x}, F_0, \tau) \\ & - E\{Z_{n1}(\xi_i|\mathbf{x}, F, \tau) - Z_{n1}(\xi_i|\mathbf{x}, F_0, \tau)\}\Big]\Big| = o_p\{(n\delta_n^p/\epsilon)^{-1/2}\}. \quad \text{(A.29)} \end{aligned}$$

Let $\mathcal{H} = \{Z_{n1}(\xi|\mathbf{x}, F, \tau) : \mathbf{x} \in \mathcal{D},\ F(.|.) \in \mathcal{C}_M^{s_4}((0,1) \otimes \mathcal{D})\}$. The proof follows the same lines as those for the Equicontinuity Lemma (Pollard, 1984, pp. 150), provided that the $\epsilon-$covering number $N(\epsilon, \|.\|_{L_2(P_n)}, \mathcal{H})$ satisfies condition (16) therein, where $P_n(.)$ denotes the empirical measure that puts mass n^{-1} at each of $Z_{n1}(\xi_i|\mathbf{x}, F, \tau),\ i = 1, \cdots, n$.

Next, note the following simple but useful fact: for any real values $|a| < 1,\ |b| < 1$, and $x,\ y \in \{0, 1\}$, it holds that $|ax - by|^2 \leq |x - y| + |a - b|^2$. An application of this inequality yields

$$\begin{aligned} |Z_{n1}(\xi|\mathbf{x}, F, \tau) - Z_{n1}(\xi|\mathbf{x}', F', \tau)|^2 \leq |\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) I\{|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n\} I\{C_i \leq Q_{n\tau}(\mathbf{X}_i, \mathbf{x})\} - \\ \boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}'}) I\{|\mathbf{X}_{i\mathbf{x}'}| \leq \delta_n\} I\{C_i \leq Q_{n\tau}(\mathbf{X}_i, \mathbf{x}')\}| + |I\{F_i(C_i) \leq \tau\} - I\{F_i'(C_i) \leq \tau\}|. \end{aligned}$$

This in turn implies that the upper bound for $\epsilon-$covering number $N(\epsilon, \|.\|_{L_2(P_n)}, \mathcal{H})$ is, the product of $N(\epsilon, \|.\|_{L_2(P_n)}, \mathcal{F}_1)$, the $\frac{\epsilon}{2}-$covering number of

$$\mathcal{F}_1 = \{\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}}) I\{|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n\} I\{C_i \leq Q_{n\tau}(\mathbf{X}_i, \mathbf{x})\} : \mathbf{x} \in \mathcal{D}, \tau \in [\tau^*, 1 - \tau^*]\} \quad \text{(A.30)}$$

and $\epsilon^2/4-$covering number $N(\epsilon^2/4, \|.\|_\infty, \mathcal{C}_M^{s_4}((0,1) \otimes \mathcal{D}))$. On one hand, by Lemmas 2.6.15, 2.6.18 and 2.6.20 of van der Vaart and Wellner (1989) and Example 38 of Pollard (1984) (pp.35), $\mathcal{F}_1$ of (A.30) is VC-subgraph class of functions with a constant envelope and thus according to Theorem 2.6.7 of van der Vaart et al. (1989), there exist some universal constants $K_1 > 0,\ W > 0$ such that

$$\sup_P N(\epsilon, \mathcal{F}_1, \|.\|_{L_2(P)}) \leq K_1 \epsilon^{-W} \text{ for any } \epsilon > 0. \quad \text{(A.31)}$$

On the other hand, Theorem 2.7.1 of van der Vaart et al. (1989) states that there exists some constant K_2 depending only on $p,\ s_4$ and the Legesque measure of the set $(0,1) \otimes \mathcal{D}$, such that

$$\log N(\epsilon^2/4, \mathcal{C}_M^{s_4}((0,1) \otimes \mathcal{D}), \|.\|_\infty) \leq K_2 \epsilon^{-2(p+1)/s_4} \quad \text{for any } \epsilon > 0.$$

This together with (A.31) implies that there exists some constant K_3 such that

$$\sup_P \log N(\epsilon, \|.\|_{L_2(P)}, \mathcal{H}) \leq K_3 \epsilon^{-2(p+1)/s_4}. \tag{A.32}$$

If $s_4 > p+1$, (A.32) means that the requirement on the covering integral-condition (16) of the Equi-continuity Lemma (Pollard, 1984, pp. 150) is satisfied and (A.29) consequently holds.

The handling of $Z_{n2}(\xi_i|\mathbf{x}, F, \tau)$ is much simpler and leads to results similar to (A.29). $\square$

Proposition 6. *Suppose conditions in Theorem 1 hold. We have*

$$\begin{aligned}
&E_i[(w_{i0}(\tau) - w_{in}(\tau))I\{Y_i \leq Q_{n\tau}(\mathbf{X}_i, \mathbf{x})\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})] \\
&= \frac{1}{n}\sum_j E_i\Big[\tilde{B}_{h_n}(\mathbf{X}_j; \mathbf{X}_i)I\{|\mathbf{X}_{i\mathbf{x}}| \leq \delta_n\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\Phi(\mathbf{X}_i, Y_j, d_j|\tau)\Big] \\
&\quad + O\Big(\delta_n^p\|F(.|.) - F_0(.|.)\|^2 + \delta_n^p(nh_n^p/\log n)^{-3/4} + \delta_n^p h_n^{s_4}\Big).
\end{aligned} \tag{A.33}$$

where the term $O(.)$ is uniform in $\mathbf{x}$, τ and $F(.|.) \in \mathcal{F}_n$.

Proof. This implies dealing with the terms introduced in (A.28). First note that through arguments used to derive (A.19), for $F(.|.)$ such that $\|F(.|.) - F_0(.|.)\|$ can be arbitrarily small, we have

$$\begin{aligned}
& E_i[Z_{n1}(\xi_i|\mathbf{x}, F, \tau) - Z_{n1}(\xi_i|\mathbf{x}, F_0, \tau)] \\
= \;& E_i\Big[I\{i \in S_n(\mathbf{x})\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\Big\{G(F^{-1}(\tau|\mathbf{X}_i)|\mathbf{X}_i) - G(Q_\tau(\mathbf{X}_i)|\mathbf{X}_i)\Big\}\Big] \\
= \;& E_i\Big[I\{i \in S_n(\mathbf{x})\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\{\frac{g}{f_0}\}(Q_\tau(\mathbf{X}_i)|\mathbf{X}_i)\{\tau - F(Q_\tau(\mathbf{X}_i)|\mathbf{X}_i)\}\Big] \\
& + O(\delta_n^p\|F(.|.) - F_0(.|.)\|^2),
\end{aligned}$$

where the second equality follows from the following observations

$$G(F^{-1}(\tau|\mathbf{x})|\mathbf{x}) - G(Q_\tau(\mathbf{x})|\mathbf{x}) = g(Q_\tau(\mathbf{x})|\mathbf{x})\{F^{-1}(\tau|\mathbf{x}) - F_0^{-1}(\tau|\mathbf{x})\}\{1 + o(1)\},$$

$$F^{-1}(\tau|\mathbf{x}) - F_0^{-1}(\tau|\mathbf{x}) = \frac{\tau - F(Q_\tau(\mathbf{x})|\mathbf{x})}{f_0(Q_\tau(\mathbf{x})|\mathbf{x})} + O(\|F(.|.) - F_0(.|.)\|^2),$$

where again the terms $O(.)$ and $o(.)$ are both unform in $\mathbf{x}$, τ and $F(.|.) \in \mathcal{F}_n$. Note that second equality above was proved by van der Vaart (1998) and was also used in Wang et al. (2009) for linear CQR. With $F(.|.) = \hat{F}_n^S$, invoke Lemma 1 with $t = Q_\tau(\mathbf{X}_i)$, $\mathbf{x} = \mathbf{X}_i$, plug it into (A.34) and we have

$$\begin{aligned}
& E_i[Z_{n1}(\xi_i|\mathbf{x}, F, \tau) - Z_{n1}(\xi_i|\mathbf{x}, F_0, \tau)] \\
&= \frac{1}{nh_n^p}\sum_{j=1}^n E_i\Big[I\{i \in S_n(\mathbf{x})\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\tilde{B}_{nj}(\mathbf{X}_j; \mathbf{X}_i)\{\frac{g}{f_0}\}(Q_\tau(\mathbf{X}_i)|\mathbf{X}_i)\varphi(Y_j, d_j, Q_\tau(\mathbf{X}_i), \mathbf{X}_i)\Big] \\
&\qquad + O\{\delta_n^p\|F(.|.) - F_0(.|.)\|^2 + \delta_n^p\Big(\frac{\log n}{nh_n^p}\Big)^{3/4} + \delta_n^p h_n^{s_4}\},
\end{aligned}$$

once again unform in $\mathbf{x}$, τ and $F(.|.) \in \mathcal{F}_n$. Similar results can be proved for the expectation of $Z_{n2}(\xi_i|\mathbf{x}, F, \tau)$. Note that

$$\frac{1}{n}\sum_j E_i\Big[\tilde{B}_{h_n}(\mathbf{X}_j;\mathbf{X}_i)I\{|\mathbf{X}_{i\mathbf{x}}| \le \delta_n\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})I\{C_i \le Q_{n\tau}(\mathbf{X}_i,\mathbf{x}), F_i(C_i) \le \tau\}\frac{F_{i0}(C_i) - F_i(C_i)}{(1 - F_i(C_i))}\Big]$$
$$= \frac{1}{n}\sum_j E_i\Big[\tilde{B}_{h_n}(\mathbf{X}_j;\mathbf{X}_i)I\{|\mathbf{X}_{i\mathbf{x}}| \le \delta_n\}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})I\{C_i \le Q_{\tau}(\mathbf{X}_i)\}\frac{F_{i0}(C_i) - F_i(C_i)}{(1 - F_{i0}(C_i))}\Big]$$
$$+ O(\delta_n^{s_2}\|F(.|.) - F_0(.|.)\| + \|F(.|.) - F_0(.|.)\|^2),$$

and consequently (A.33) holds. Depending on the ratio δ_n/h_n, the leading term in (A.33) above admits the following asymptotic expressions:

- $\delta_n = o(h_n)$,

$$\boldsymbol{\gamma}(A)f_{\mathbf{X}}(\mathbf{x})\frac{\delta_n^p}{n}\sum_j \tilde{B}_{h_n}(\mathbf{X}_j;\mathbf{x})\Phi(\mathbf{x}, Y_j, d_j) + O(\delta_n^{p+1}\tilde{\tau}_n); \tag{A.34}$$

- if $\delta_n/h_n \to \infty$,

$$\frac{1}{n}\sum_{j\in S_n(\mathbf{x})} f_{\mathbf{X}}(\mathbf{X}_j)\boldsymbol{\mu}_n(\mathbf{X}_{j\mathbf{x}})\Phi(\mathbf{X}_j, Y_j, d_j) + O(\delta_n^p h_n\tau_n/\delta_n); \tag{A.35}$$

- $\delta_n \asymp h_n$,

$$f_{\mathbf{X}}(\mathbf{x})\frac{\delta_n^p}{n}\sum_j B_{h_n}(\mathbf{X}_j;\mathbf{x})\Phi(\mathbf{x}, Y_j, d_j) + O(\delta_n^{p+1}\tilde{\tau}_n), \tag{A.36}$$

where $B_{h_n}(\mathbf{X}_j;\mathbf{x}) = \int_{[-1,1]^p}\tilde{B}_{h_n}(\mathbf{X}_j;\mathbf{x} + \delta_n\mathbf{t})\mu(\mathbf{t})d\mathbf{t}$. □

Proposition 7. *Suppose conditions in Theorem 1 hold. We have*

$$\sup_{\tau,\mathbf{x},\boldsymbol{\beta}}\frac{1}{N_n(\mathbf{x})}\Big|\sum_{i\in S_n(\mathbf{x})}[Z_{ni}(\mathbf{x},\boldsymbol{\beta},\hat{F}_n|\tau) - Z_{ni}(\mathbf{x},\boldsymbol{\beta},F_0|\tau)]\Big| = O_p\Big\{\tau_n + \Big(\frac{\tilde{\tau}_n^{(1-\alpha)}}{\delta_n^{p\alpha}\log n}\Big)^{1/2}\tau_n\Big\};$$
$$\sup_{\tau,\mathbf{x},\boldsymbol{\beta}}\frac{1}{N_n(\mathbf{x})}\Big|\sum_{i\in S_n(\mathbf{x})}\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})(w_{i0}(\tau) - w_{in}(\tau))\Big[I\{Y_i \le \boldsymbol{\beta}^\top\boldsymbol{\mu}_n(\mathbf{X}_{i\mathbf{x}})\} - I\{Y_i \le Q_{n\tau}(\mathbf{X}_i,\mathbf{x})\}\Big]\Big|$$
$$= O_p\{\tau_n\tilde{\tau}_n + \Big(\frac{\tilde{\tau}_n^{(1-\alpha)}}{\delta_n^{p\alpha}\log n}\Big)^{1/2}\tau_n^{(3-\alpha)/2}\}.$$

Proof. The arguments are exactly the same as those used to prove Propositions 5 and 6. □

REFERENCES

Chaudhuri, P. (1991) Nonparametric estimates of regression quantiles and their local Bahadur representation. *Annals of Statistics* **19**, 760-777.

Gill, P.E., Murray, W. & Wright, M.H. (1991) *Numerical Linear Algebra and Optimization.* Vol. 1. Addison-Wesley.

Gonzalez-Manteigaa, W. & Cadarso-Suarez, C. (1994) Asymptotic properties of a generalized Kaplan-Meier estimator with some applications. *Journal of Nonparametric Statistics* 4, 65-78.

Koenker, R. & Bassett, G. (1978) Regression quantiles. *Econometrica* 46, 33-50.

Kong, E., Linton, O. & Xia, Y. (2013) Global Bahadur representation for nonparametric censored regression quantiles and its applications. *Econometric Theory* 29, 941-968.

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

Pakes, A. & Pollard, D. (1989) Simulation and the asymptotics of optimization estimators. *Econometrica* 57, 1027-1057.

Pollard, D. (1984) *Convergence of Stochastic Processes.* Springer-Verlag.

van der Vaart, A.W. & Wellner, J. (1996) *Weak Convergence and Empirical Pprocesses: with Applications to Statistics.* Springer-Verlag.

van der Vaart, A.W. (1998) *Asymptotic Statistics.* Cambridge University Press.

Wang, H. & Wang, L. (2009) Locally weighted censored quantile regression. *Journal of the American Statistical Association* 104, 1117-1128.