Statistica Sinica: Supplement

STATISTICAL INFERENCE IN QUANTILE REGRESSION FOR ZERO-INFLATED OUTCOMES

Wodan Ling[1], Bin Cheng[*,2], Ying Wei[2], Joshua Z. Willey[2], and Ying Kuen Cheung[2]

[1]*Fred Hutchinson Cancer Research Center and* [2]*Columbia University*

Supplementary Material

The Supplementary Material contains the proofs of Theorems 1 and 2 (Section S1) and additional figures and tables (Section S2).

S1 Technical proofs

S1.1 Proof of Theorem 1

Proof of Theorem 1 (i): Since $Q_Y(\tau|\boldsymbol{x})$ is piecewise defined, we prove the consistency under three scenarios, namely, $\tau < 1-\pi(\boldsymbol{\gamma},\boldsymbol{x})$, $\tau = 1-\pi(\boldsymbol{\gamma},\boldsymbol{x})$, and $\tau > 1-\pi(\boldsymbol{\gamma},\boldsymbol{x})$.

If $\tau < 1-\pi(\boldsymbol{\gamma},\boldsymbol{x})$, then by the consistency of $\hat{\gamma}_n$, $P(B_n \cup C_n) \to 0$. For any $\epsilon > 0$,

$$\begin{aligned}
P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) - 0 \mid > \epsilon \Big) &= P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) \mid > \epsilon \Big) \\
&\leq P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) \mid > \epsilon, A_n \Big) + P(B_n \cup C_n) \\
&= P(B_n \cup C_n) \to 0.
\end{aligned}$$

If $\tau = 1-\pi(\boldsymbol{\gamma},\boldsymbol{x})$, then

$$\begin{aligned}
P(A_n) &= P\Big(\sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n,\boldsymbol{x}) - \pi(\boldsymbol{\gamma},\boldsymbol{x})\} < 0 \Big) \to \frac{1}{2}, \\
P(B_n) &= P\Big(0 \leq \sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n,\boldsymbol{x}) - \pi(\boldsymbol{\gamma},\boldsymbol{x})\} \leq n^{\frac{1}{2}-\delta} \Big) \to \frac{1}{2},
\end{aligned}$$

and $P(C_n) \to 0$. Thus,

$$\begin{aligned}
& P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) - 0 \mid > \epsilon \Big) \\
=\ & P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) \mid > \epsilon, A_n \Big) + P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) \mid > \epsilon, B_n \Big) + P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) \mid > \epsilon, C_n \Big) \\
\leq\ & 0 + P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) \mid > \epsilon, B_n \Big) + P(C_n) \\
\leq\ & P\Big[\mid \boldsymbol{x}^\top \widehat{\boldsymbol{\beta}}_n \{n^{-\delta} \pi(\widehat{\boldsymbol{\gamma}}_n,\boldsymbol{x})^{-1}\} n^{\delta} \{\pi(\widehat{\boldsymbol{\gamma}}_n,\boldsymbol{x}) - \pi(\boldsymbol{\gamma},\boldsymbol{x})\} \mid > \epsilon \Big] + P(C_n) \to 0,
\end{aligned}$$

which is because

$$
\begin{aligned}
& \boldsymbol{x}^\top \widehat{\boldsymbol{\beta}}_n \{n^{-\delta} \pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x})^{-1}\} n^{\delta} \{\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x})\} \\
= \;& \boldsymbol{x}^\top \widehat{\boldsymbol{\beta}}_n \{n^{-\delta} \pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x})^{-1}\} n^{\delta - \frac{1}{2}} \cdot n^{\frac{1}{2}} \{\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x})\} \\
= \;& o_p(1) \cdot O_p(1) = o_p(1).
\end{aligned}
$$

If $\tau > 1 - \pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x})$, then $P(C_n) \to 1$ and $P(A_n \cup B_n) \to 0$. Since $\widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau; \boldsymbol{x}, \widehat{\boldsymbol{\gamma}}_n)$ is consistent for $\boldsymbol{\beta} \circ \Gamma(\tau; \boldsymbol{x}, \boldsymbol{\gamma})$, we conclude that

$$
\begin{aligned}
& P\Big(\mid \widehat{Q}_Y(\tau|\boldsymbol{x}) - Q_Y(\tau|\boldsymbol{x}) \mid > \epsilon \Big) \\
\leq \;& P(A_n \cup B_n) + P\Big\{ \mid \boldsymbol{x}^\top \widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau; \boldsymbol{x}, \widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{x}^\top \boldsymbol{\beta} \circ \Gamma(\tau; \boldsymbol{x}, \boldsymbol{\gamma}) \mid > \epsilon \Big\} \\
\to \;& 0.
\end{aligned}
$$

□

Proof of Theorem 1 (ii): We now establish the asymptotic distribution of $\widehat{Q}_Y(\tau|\boldsymbol{x})$ under the three scenarios.

If $\tau < 1 - \pi(\boldsymbol{\gamma}, \boldsymbol{x})$, then

$$
\begin{aligned}
\sqrt{n}\Big(\widehat{Q}_Y(\tau|\boldsymbol{x}) - 0\Big) &= \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x}) I(A_n) + \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x}) I(B_n \cup C_n) \\
&= \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x}) I(B_n \cup C_n),
\end{aligned}
$$

which converges to 0 in probability because for any $\epsilon > 0$,

$$
P\Big(\mid \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x}) I(B_n \cup C_n) \mid > \epsilon \Big) \leq P(B_n \cup C_n) \to 0.
$$

Therefore, $\sqrt{n}\Big(\widehat{Q}_Y(\tau|\boldsymbol{x}) - 0\Big) \rightarrow_p 0$.

If $\tau = 1 - \pi(\boldsymbol{\gamma}, \boldsymbol{x})$, we have

$$
\begin{aligned}
& \sqrt{n}\Big(\widehat{Q}_Y(\tau|\boldsymbol{x}) - 0\Big) \\
= \;& \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x})I(B_n) + \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x})I(C_n) \\
= \;& \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x})I(B_n) + o_p(1) \\
= \;& n^{\delta}\boldsymbol{x}^{\top}\widehat{\boldsymbol{\beta}}_n\{n^{-\delta}\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x})^{-1}\} \\
& \cdot\sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x})\}I\{0 \le \pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x}) \le n^{-\delta}\} + o_p(1) \\
= \;& n^{\delta}\widehat{Q}_Y\{n^{-\delta}\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x})^{-1}|\boldsymbol{x}, Y > 0\} \\
& \cdot\sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x})\}I\{0 \le \pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x}) \le n^{-\delta}\} + o_p(1).
\end{aligned}
$$

Let $Z_n = \sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x})\}$. By the asymptotic distribution of $\widehat{\gamma}_n$ and the delta method, we have $Z_n \rightarrow_d Z$, where Z follows a normal distribution with mean 0. Consider the function $h = z \cdot I(z > 0)$, whose set of discontinuity points is $D_h = \{0\}$. Noticing $P(Z \in D_h) = 0$, we can apply the almost surely continuous mapping theorem to function h to obtain

$$
\begin{aligned}
& \sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x})\}I\{0 \le \pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x}) - \pi(\boldsymbol{\gamma}, \boldsymbol{x}) \le n^{-\delta}\} \\
\rightarrow_d \;& Z \cdot I(Z > 0) \\
= \;& \pi(\boldsymbol{\gamma}, \boldsymbol{x})\{1 - \pi(\boldsymbol{\gamma}, \boldsymbol{x})\}\sqrt{\boldsymbol{x}^{\top}\boldsymbol{D}_{1,\gamma}^{-1}\boldsymbol{x}}\; Z_0 I\{Z_0 > 0\},
\end{aligned}
$$

where $Z_0 \sim N(0,1)$. We prove that

$$n^{\delta}\widehat{Q}_Y\{n^{-\delta}\pi(\widehat{\boldsymbol{\gamma}}_n,\boldsymbol{x})^{-1}|\boldsymbol{x},Y>0\} \rightarrow_p Q'_Y(0|\boldsymbol{x},Y>0)\pi(\boldsymbol{\gamma},\boldsymbol{x})^{-1}.$$

Let $\Pi_{\boldsymbol{\gamma},\boldsymbol{x}}$ be a compact neighborhood of $\pi(\boldsymbol{\gamma},\boldsymbol{x})$ that excluding 0 and let $\pi_* = \sup \Pi_{\boldsymbol{\gamma},\boldsymbol{x}} < \infty$. We prove first that

$$\sup_{\pi\in\Pi_{\boldsymbol{\gamma},\boldsymbol{x}}} n^{\delta}\Big\{\widehat{Q}_Y(n^{-\delta}\pi^{-1}|\boldsymbol{x},Y>0) - Q_Y(n^{-\delta}\pi^{-1}|\boldsymbol{x},Y>0)\Big\} \rightarrow_p 0.$$

Let $q_n(t)$ and $u_n(t)$ denote the empirical processes associated with $Y|(\boldsymbol{x},Y>0)$ and the uniform distribution on $(0,1)$, respectively. And let $B_n(t)$ be the standard Brownian bridge process. For any $\epsilon>0$,

$$\begin{aligned}
& P\Big\{\sup_{\pi\in\Pi_{\boldsymbol{\gamma},\boldsymbol{x}}} n^{\delta} \mid \widehat{Q}_Y(n^{-\delta}\pi^{-1}|\boldsymbol{x},Y>0) - Q_Y(n^{-\delta}\pi^{-1}|\boldsymbol{x},Y>0) \mid > \epsilon\Big\} \\
\leq\ & P\Big[\sup_{\pi\in\Pi_{\boldsymbol{\gamma},\boldsymbol{X}}} f_{Y|Y>0}\{F_Y^{-1}(n^{-\delta}\pi^{-1}|\boldsymbol{x},Y>0)\}\sqrt{n} \mid \widehat{Q}_Y(n^{-\delta}\pi^{-1}|\boldsymbol{x},Y>0) \\
& -Q_Y(n^{-\delta}\pi^{-1}|\boldsymbol{x},Y>0) \mid > \epsilon n^{\frac{1}{2}-\delta} f_{Y|Y>0}\{F_Y^{-1}(n^{-\delta}\pi_*^{-1}|\boldsymbol{x},Y>0)\}\Big] \\
\leq\ & P\Big[\sup_{0<t<1} f_{Y|Y>0}\{F_Y^{-1}(t|\boldsymbol{x},Y>0)\} \mid q_n(t) \mid > \epsilon n^{\frac{1}{2}-\delta} f_{Y|Y>0}\{F_Y^{-1}(n^{-\delta}\pi_*^{-1}|\boldsymbol{x},Y>0)\}\Big] \\
\leq\ & P\Big[\sup_{0<t<1} \mid f_{Y|Y>0}\{F_Y^{-1}(t|\boldsymbol{x},Y>0)\}\ q_n(t) - u_n(t) \mid \\
& \qquad > \frac{\epsilon}{3} n^{\frac{1}{2}-\delta} f_{Y|Y>0}\{F_Y^{-1}(n^{-\delta}\pi_*^{-1}|\boldsymbol{x},Y>0)\}\Big] \\
& + P\Big[\sup_{0<t<1} \mid u_n(t) - B_n(t) \mid > \ \frac{\epsilon}{3} n^{\frac{1}{2}-\delta} f_{Y|Y>0}\{F_Y^{-1}(n^{-\delta}\pi_*^{-1}|\boldsymbol{x},Y>0)\}\Big] \\
& + P\Big[\sup_{0<t<1} \mid B_n(t) \mid > \frac{\epsilon}{3} n^{\frac{1}{2}-\delta} f_{Y|Y>0}\{F_Y^{-1}(n^{-\delta}\pi_*^{-1}|\boldsymbol{x},Y>0)\}\Big] = \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}$$

By Theorem 1 on Page 640 of Shorack and Wellner (1986), under the assumption that $f_{Y|Y>0}(\cdot \mid \boldsymbol{x})$ is positive and continuous on $[0, \infty)$, we have

$$\sup_{0\le t\le 1} \mid f_{Y|Y>0}\{F_Y^{-1}(t|\boldsymbol{x}, Y>0)\}\ q_n(t) - u_n(t) \mid \to_p 0.$$

So I $\to 0$. By Theorem B in Csörgö and Révész (1978),

$$\sup_{0<t<1} \mid u_n(t) - B_n(t) \mid =_{a.s.} O(n^{-\frac{1}{2}} \log n),$$

which implies that II $\to 0$. Finally, recalling that for any $c > 0$

$$P\left(\sup_{0<t<1} \mid B_n(t) \mid > c\right) = 2\sum_{i=1}^{\infty}(-1)^{i+1}\exp(-2i^2c^2),$$

we conclude that

$$\text{III} = P\Big[\sup_{0<t<1} \mid B_n(t) \mid > \frac{\epsilon}{3} n^{\frac{1}{2}-\delta} f_{Y|Y>0}\{F_Y^{-1}(n^{-\delta}\pi_*^{-1}|\boldsymbol{x}, Y>0)\}\Big] \to 0.$$

Taylor's expansion of $Q_Y\{n^{-\delta}\pi(\boldsymbol{\gamma}, \boldsymbol{x})^{-1} \mid \boldsymbol{x}, Y>0\}$ at 0 yields

$$Q_Y\{n^{-\delta}\pi(\boldsymbol{\gamma}, \boldsymbol{x})^{-1}|\boldsymbol{x}, Y>0\} = Q_Y'(0|\boldsymbol{x}, Y>0)\ n^{-\delta}\pi(\boldsymbol{\gamma}, \boldsymbol{x})^{-1} + o\{n^{-2\delta}\pi(\boldsymbol{\gamma}, \boldsymbol{x})^{-2}\},$$

which, together with (S1.1) and the consistency of $\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x})$ implies that

$$n^{\delta}\widehat{Q}_Y\{n^{-\delta}\pi(\widehat{\boldsymbol{\gamma}}_n, \boldsymbol{x})^{-1}|\boldsymbol{x}, Y>0\} \to_p Q_Y'(0|\boldsymbol{x}, Y>0)\pi(\boldsymbol{\gamma}, \boldsymbol{x})^{-1}.$$

By the Slutsky's theorem

$$\sqrt{n}\Big(\widehat{Q}_Y(\tau|\boldsymbol{x}) - 0\Big) \to_d \{1-\pi(\boldsymbol{\gamma}, \boldsymbol{x})\}\sqrt{\boldsymbol{x}^\top \boldsymbol{D}_{1,\gamma}^{-1}\boldsymbol{x}}\ Q_Y'(0|\boldsymbol{x}, Y>0) Z_0 I\{Z_0>0\},$$

where $Z_0 \sim N(0,1)$. Finally, if $\tau > 1-\pi(\boldsymbol{\gamma},\boldsymbol{x})$,

$$\begin{aligned}
&\sqrt{n}\Big\{\widehat{Q}_Y(\tau|\boldsymbol{x}) - Q_Y(\tau|\boldsymbol{x})\Big\} \\
= \;& \sqrt{n}\{\widehat{Q}_Y(\tau|\boldsymbol{x})I(C_n) - Q_Y(\tau|\boldsymbol{x})\} + \sqrt{n}\widehat{Q}_Y(\tau|\boldsymbol{x})I(A_n \cup B_n) \\
= \;& \boldsymbol{x}^\top\sqrt{n}\{\widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau;\boldsymbol{x},\widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\} + o_p(1).
\end{aligned}$$

We now proceed to establish the limiting distribution of $\sqrt{n}\{\widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau;\boldsymbol{x},\widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\}$. Define

$$\boldsymbol{m}(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma}) = \begin{pmatrix} \boldsymbol{X}\{\tau - I(Y - \boldsymbol{X}^\top\boldsymbol{\beta} < 0)\}I(Y>0) \\ \boldsymbol{X}\{I(Y>0) - \pi(\boldsymbol{\gamma},\boldsymbol{X})\} \end{pmatrix}.$$

Define

$$\boldsymbol{M}(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma}) = \int \boldsymbol{m}(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma})dP, \quad \boldsymbol{M}_n(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma}) = \int \boldsymbol{m}(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma})d\mathbb{P}_n.$$

Let $\boldsymbol{\Delta}$ be a compact neighborhood of the true $\boldsymbol{\gamma}$, T be a compact subset of interval $(0,1)$ containing $\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})$, and $\boldsymbol{\Psi}$ is a compact neighborhood of the true $\boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})$. The class

$$\mathcal{F} = \Big\{\boldsymbol{m}(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma}), \tau\times\boldsymbol{\gamma}\times\boldsymbol{\beta} \in T\times\boldsymbol{\Delta}\times\boldsymbol{\Psi}\Big\}$$

is clearly a VC class with a squared integrable envelope function $\boldsymbol{X}\boldsymbol{X}^\top$. Thus, $\mathcal{F}$ is a Donsker class. Define

$$\mathbb{G}_n(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma}) = \sqrt{n}\left\{\boldsymbol{M}_n(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma}) - \boldsymbol{M}(\boldsymbol{\beta}(\tau),\boldsymbol{\gamma})\right\}.$$

The fact that $\mathcal{F}$ is a Donsker class and both $\widehat{\boldsymbol{\gamma}}_n$ and $\widehat{\boldsymbol{\beta}}_n(\tau)$ are consistent implies that

$$\mathbb{G}_n(\widehat{\boldsymbol{\beta}}_n(\tau), \widehat{\boldsymbol{\gamma}}_n) = \mathbb{G}_n(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) + o_p(1),$$

which is equivalent to

$$\begin{aligned} o_P(1) &= \sqrt{n}\boldsymbol{M}_n(\widehat{\boldsymbol{\beta}}_n(\tau), \widehat{\boldsymbol{\gamma}}_n) \\ &= \sqrt{n}\boldsymbol{M}_n(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) + \sqrt{n}\boldsymbol{M}(\widehat{\boldsymbol{\beta}}_n(\tau), \widehat{\boldsymbol{\gamma}}_n) \\ &= \sqrt{n}\boldsymbol{M}_n(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) + \sqrt{n}\nabla\boldsymbol{M}(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) \begin{pmatrix} \widehat{\boldsymbol{\beta}}_n(\tau) - \boldsymbol{\beta}(\tau) \\ \widehat{\boldsymbol{\gamma}}_n - \boldsymbol{\gamma} \end{pmatrix}, \end{aligned}$$

where the first equality is by the definition of $\widehat{\boldsymbol{\gamma}}_n$ and $\widehat{\boldsymbol{\beta}}_n(\tau)$.

Thus,

$$\sqrt{n} \begin{pmatrix} \widehat{\boldsymbol{\beta}}_n(\tau) - \boldsymbol{\beta}(\tau) \\ \widehat{\boldsymbol{\gamma}}_n - \boldsymbol{\gamma} \end{pmatrix} = -\nabla\boldsymbol{M}^{-1}(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma})\sqrt{n}\boldsymbol{M}_n(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) + o_p(1).$$

By the central limit theorem,

$$\sqrt{n}\boldsymbol{M}_n(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) \to_d N\left\{\mathbf{0}, \begin{pmatrix} \tau(1-\tau)\boldsymbol{D}_0 & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & \boldsymbol{D}_{1,\boldsymbol{\gamma}} \end{pmatrix}\right\},$$

and calculation yields

$$-\nabla\boldsymbol{M}(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) = \begin{pmatrix} \boldsymbol{D}_{1,\boldsymbol{\beta}(\tau)} & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & \boldsymbol{D}_{1,\boldsymbol{\gamma}} \end{pmatrix},$$

where

$$\boldsymbol{D}_{1,\boldsymbol{\beta}(\tau)} = \mathbb{E}_{\boldsymbol{X}}\left[\pi(\boldsymbol{\gamma}, \boldsymbol{X}) f_{Y|Y>0}\{\boldsymbol{X}^\top\boldsymbol{\beta}(\tau) \mid \boldsymbol{X}\}\boldsymbol{X}\boldsymbol{X}^\top\right], \quad \boldsymbol{D}_0 = \mathbb{E}_{\boldsymbol{X}}\left\{\pi(\boldsymbol{\gamma}, \boldsymbol{X})\boldsymbol{X}\boldsymbol{X}^\top\right\},$$

$$\boldsymbol{D}_{1,\boldsymbol{\gamma}} = \mathbb{E}_{\boldsymbol{X}}\Big[\pi(\boldsymbol{\gamma},\boldsymbol{X})\{1-\pi(\boldsymbol{\gamma},\boldsymbol{X})\}\boldsymbol{X}\boldsymbol{X}^\top\Big].$$

Therefore,

$$\sqrt{n}\begin{pmatrix}\widehat{\boldsymbol{\beta}}_n(\tau)-\boldsymbol{\beta}(\tau)\\ \\ \widehat{\boldsymbol{\gamma}}_n-\boldsymbol{\gamma}\end{pmatrix}\to_d N\left\{\mathbf{0},\begin{pmatrix}\tau(1-\tau)\boldsymbol{D}_{1,\boldsymbol{\beta}(\tau)}^{-1}\boldsymbol{D}_0\boldsymbol{D}_{1,\boldsymbol{\beta}(\tau)}^{-1} & \mathbf{0}_{p\times p}\\ \\ \mathbf{0}_{p\times p} & \boldsymbol{D}_{1,\boldsymbol{\gamma}}^{-1}\end{pmatrix}\right\}.$$

By the delta method,

$$\sqrt{n}\begin{pmatrix}\widehat{\boldsymbol{\beta}}_n\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})-\boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\\ \\ \Gamma(\tau;\boldsymbol{x},\widehat{\boldsymbol{\gamma}}_n)-\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\end{pmatrix}\to_d N\left\{\mathbf{0},\begin{pmatrix}\boldsymbol{V}_1 & \mathbf{0}_{p\times 1}\\ \\ \mathbf{0}_{p\times 1}^\top & V_2\end{pmatrix}\right\},$$

where

$$\boldsymbol{V}_1=\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\{1-\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\}\boldsymbol{D}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})}^{-1}\boldsymbol{D}_0\boldsymbol{D}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})}^{-1},$$

$$V_2=\{1-\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\}^2\{1-\pi(\boldsymbol{\gamma},\boldsymbol{x})\}^2\boldsymbol{x}^\top\boldsymbol{D}_{1,\boldsymbol{\gamma}}^{-1}\boldsymbol{x}.$$

Applying the delta method again, and then Slutsky's theorem,

$$\begin{aligned}&\sqrt{n}\Big(\widehat{Q}_Y(\tau|\boldsymbol{x})-Q_Y(\tau|\boldsymbol{x})\Big)\\ &=\boldsymbol{x}^\top\sqrt{n}\{\widehat{\boldsymbol{\beta}}_n\circ\Gamma(\tau;\boldsymbol{x},\widehat{\boldsymbol{\gamma}}_n)-\boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\}+o_p(1)\to_d N(0,\,\Sigma_1+\Sigma_2),\end{aligned}$$

where

$$\Sigma_1=\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\{1-\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\}\boldsymbol{x}^\top\boldsymbol{D}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})}^{-1}\boldsymbol{D}_0\boldsymbol{D}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})}^{-1}\boldsymbol{x},$$

$$\Sigma_2=\{1-\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\}^2\{1-\pi(\boldsymbol{\gamma},\boldsymbol{x})\}^2\boldsymbol{x}^\top\boldsymbol{D}_{1,\gamma}^{-1}\boldsymbol{x}\,\boldsymbol{x}^\top\dot{\boldsymbol{\beta}}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})\,\dot{\boldsymbol{\beta}}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})^\top\boldsymbol{x},$$

and

$$\dot{\boldsymbol{\beta}}\circ\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})=\frac{d\boldsymbol{\beta}(\tau)}{d\tau}\bigg|_{\tau=\Gamma(\tau;\boldsymbol{x},\boldsymbol{\gamma})}.$$

□

S1.2 Proof of Theorem 2

Proof of Theorem 2: Without loss of generality, we derive the following theorem for the treatment effect of a binary covariate X_j, which takes values 0 or 1. By Theorem 1,

$$\begin{aligned}
& \sqrt{n}\Bigg[\Big\{\widehat{Q}_Y(\tau \mid X_j = 1, \boldsymbol{X}^{(-j)}) - \widehat{Q}_Y(\tau \mid X_j = 0, \boldsymbol{X}^{(-j)})\Big\} \qquad \text{(S1.1)} \\
& \qquad - \Big\{Q_Y(\tau \mid X_j = 1, \boldsymbol{X}^{(-j)}) - Q_Y(\tau \mid X_j = 0, \boldsymbol{X}^{(-j)})\Big\}\Bigg] \\
= \; & (1, \boldsymbol{X}^{(-j)})\boldsymbol{\beta}'(0)\pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})^{-1} \\
& \qquad \Big[\sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n, 1, \boldsymbol{X}^{(-j)}) - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})\}\Big]_+ I\{\tau = 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})\} \\
- \; & (0, \boldsymbol{X}^{(-j)})\boldsymbol{\beta}'(0)\pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})^{-1} \\
& \qquad \Big[\sqrt{n}\{\pi(\widehat{\boldsymbol{\gamma}}_n, 0, \boldsymbol{X}^{(-j)}) - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{X}^{(-j)})\}\Big]_+ I\{\tau = 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{X}^{(-j)})\} \\
+ \; & (1, \boldsymbol{X}^{(-j)})\sqrt{n}\Big\{\widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau; 1, \boldsymbol{X}^{(-j)}, \widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta} \circ \Gamma(\tau; 1, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma})\Big\} \\
& \qquad I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})\} \\
- \; & (0, \boldsymbol{X}^{(-j)})\sqrt{n}\Big\{\widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau; 0, \boldsymbol{X}^{(-j)}, \widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta} \circ \Gamma(\tau; 0, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma})\Big\} \\
& \qquad I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{X}^{(-j)})\} + o_p(1),
\end{aligned}$$

where $f_+ = \max\{f, 0\}$.

Let $\boldsymbol{\theta} = (\boldsymbol{\gamma}, \boldsymbol{\beta})$ and define

$$\begin{aligned} h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)}) &= Q_Y(\tau \mid 1, \boldsymbol{X}^{(-j)}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})\} \\ &\qquad - Q_Y(\tau \mid 0, \boldsymbol{X}^{(-j)}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{X}^{(-j)})\} \\ &= (1, \boldsymbol{X}^{(-j)})\boldsymbol{\beta} \circ \Gamma(\tau; 1, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})\} \\ &\qquad - (0, \boldsymbol{X}^{(-j)})\boldsymbol{\beta} \circ \Gamma(\tau; 0, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{X}^{(-j)})\}. \end{aligned}$$

Denote $\boldsymbol{X}^{0,(-j)}$ as the covariates from the current data excluding variable X_j, which is used in estimating the average quantile effect, and denote $\boldsymbol{X}^{(-j)}$ as new covariates, which have the same distribution $P_{\boldsymbol{X}^{(-j)}}$ as $\boldsymbol{X}^{0,(-j)}$. We first consider the case when $P_{\boldsymbol{X}^{(-j)}}$ is absolutely continuous with respect to the Lebesgue measure.

Write

$$
\begin{aligned}
& \sqrt{n}\Big(\hat{\Delta}_\tau(X_j;1,0)-\Delta_\tau(X_j;1,0)\Big) \\
= \;& \sqrt{n}\left\{\int h(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{X}^{0,(-j)})d\,\mathbb{P}_{n\,\boldsymbol{X}^{0,(-j)}} - \int h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})dP_{\boldsymbol{X}^{(-j)}}\right\} \\
= \;& \sqrt{n}\left\{\int h(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{X}^{0,(-j)})d\,\mathbb{P}_{n\,\boldsymbol{X}^{0,(-j)}} - \int h(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{X}^{0,(-j)})dP_{\boldsymbol{X}^{0,(-j)}}\right\} \\
& \quad +\sqrt{n}\left\{\int h(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{X}^{(-j)}) - h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})dP_{\boldsymbol{X}^{(-j)}}\right\} \\
= \;& \sqrt{n}\int h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})d(\mathbb{P}_{n\,\boldsymbol{X}^{0,(-j)}} - P_{\boldsymbol{X}^{0,(-j)}}) \\
& \quad +\sqrt{n}\left\{\int h(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{X}^{(-j)}) - h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})dP_{\boldsymbol{X}^{(-j)}}\right\} + o_p(1) \\
= \;& \int\left[\sqrt{n}\int h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})d(\mathbb{P}_{n\,\boldsymbol{X}^{0,(-j)}} - P_{\boldsymbol{X}^{0,(-j)}}) + \sqrt{n}\{h(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{X}^{(-j)}) - h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})\}\right]dP_{\boldsymbol{X}^{(-j)}} \\
& \quad + o_p(1),
\end{aligned}
$$

where the second last step is by asymptotic tightness of the empirical process $\mathbb{G}_n h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})$ because the fact that $\boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})$ and $\boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})$ as smooth functions of $\boldsymbol{X}^{(-j)}$ with compact supports implies that $\{h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})\}$ as indexed by $\boldsymbol{X}^{(-j)}$ is a Donsker class.

The integrand in the previous display is written as

$$
\begin{aligned}
& \sqrt{n}\int h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)}) d(\mathbb{P}_{n\,\boldsymbol{X}^{0,(-j)}} - P_{\boldsymbol{X}^{0,(-j)}}) + \sqrt{n}\{h(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{X}^{(-j)}) - h(\boldsymbol{\theta}, \boldsymbol{X}^{(-j)})\} \\
= \;& \sqrt{n}\Bigg(\frac{1}{n}\sum_{i=1}^{n}\Big[(1, \boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau; 1, \boldsymbol{x}_i^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{x}_i^{0,(-j)})\} \\
- \;& (0, \boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau; 0, \boldsymbol{x}_i^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{x}_i^{0,(-j)})\}\Big] - \Delta_\tau(X_j; 1, 0)\Bigg) \\
+ \;& \sqrt{n}\Big[(1, \boldsymbol{X}^{(-j)})\{\widehat{\boldsymbol{\beta}}_n\circ\Gamma(\tau; 1, \boldsymbol{X}^{(-j)}, \widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta}\circ\Gamma(\tau; 1, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma})\} I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{(-j)})\} \\
- \;& (0, \boldsymbol{X}^{(-j)})\{\widehat{\boldsymbol{\beta}}_n\circ\Gamma(\tau; 0, \boldsymbol{X}^{(-j)}, \widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta}\circ\Gamma(\tau; 0, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma})\} I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{X}^{(-j)})\}\Big],
\end{aligned}
$$

whose limiting distribution is derived as follows.

By the central limit theorem,

$$
\sqrt{n}\begin{pmatrix}
\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i^0\{\tau - I(y_i - \boldsymbol{x}_i^{0\top}\boldsymbol{\beta}(\tau) < 0)\} I(y_i > 0) - \mathbf{0} \\
\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i^0\,\{I(y_i > 0 - \pi(\boldsymbol{\gamma}, \boldsymbol{x}_i^0\}) - \mathbf{0} \\
\frac{1}{n}\sum_{i=1}^{n}\Big[(1, \boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau; 1, \boldsymbol{x}_i^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{x}_i^{0,(-j)})\} \\
-(0, \boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau; 0, \boldsymbol{x}_i^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{x}_i^{0,(-j)})\}\Big] - \Delta_\tau(X_j; 1, 0)
\end{pmatrix}
\\
\rightarrow_d N\left\{\mathbf{0}, \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0}_{2p\times 1} \\ \mathbf{0}_{2p\times 1}^\top & \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}} \end{pmatrix}\right\},
$$

where

$$
\begin{aligned}
\sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}} = \mathrm{Var}\Big[&(1, \boldsymbol{X}^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau; 1, \boldsymbol{X}^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{X}^{0,(-j)})\} \\
&- (0, \boldsymbol{X}^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau; 0, \boldsymbol{X}^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{X}^{0,(-j)})\}\Big].
\end{aligned}
$$

Applying the delta method, we get

$$\sqrt{n}\begin{pmatrix} \widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma}) - \boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma}) \\ \widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma}) - \boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma}) \\ \widehat{\boldsymbol{\gamma}}_n - \boldsymbol{\gamma} \\ \frac{1}{n}\sum_{i=1}^n \Big[(1,\boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{x}_i^{0,(-j)},\boldsymbol{\gamma})I\{\tau > 1-\pi(\boldsymbol{\gamma},1,\boldsymbol{x}_i^{0,(-j)})\} \\ -(0,\boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{x}_i^{0,(-j)},\boldsymbol{\gamma})I\{\tau > 1-\pi(\boldsymbol{\gamma},0,\boldsymbol{x}_i^{0,(-j)})\}\Big] - \Delta_\tau(X_j;1,0) \end{pmatrix}$$

$$\to_d N\left\{\mathbf{0}, \begin{pmatrix} \boldsymbol{V}_{\boldsymbol{\beta}_1} & \boldsymbol{V}_{\boldsymbol{\beta}_1,\boldsymbol{\beta}_0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{V}_{\boldsymbol{\beta}_0,\boldsymbol{\beta}_1} & \boldsymbol{V}_{\boldsymbol{\beta}_0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{D}_{1,\boldsymbol{\gamma}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}} \end{pmatrix}\right\},$$

where

$$\boldsymbol{V}_{\boldsymbol{\beta}_1} = \Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\{1-\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\}\boldsymbol{D}^{-1}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})}\boldsymbol{D}_0\boldsymbol{D}^{-1}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})},$$

$$\boldsymbol{V}_{\boldsymbol{\beta}_0} = \Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\{1-\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\}\boldsymbol{D}^{-1}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})}\boldsymbol{D}_0\boldsymbol{D}^{-1}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})},$$

$$\boldsymbol{V}_{\boldsymbol{\beta}_1,\boldsymbol{\beta}_0} = \boldsymbol{V}^\top_{\boldsymbol{\beta}_0,\boldsymbol{\beta}_1} = \{\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\wedge\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})$$
$$-\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\}\boldsymbol{D}^{-1}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})}\boldsymbol{D}_0\boldsymbol{D}^{-1}_{1,\boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})}.$$

Define $\boldsymbol{V}_\gamma = \mathrm{Cov}\{\boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\widehat{\boldsymbol{\gamma}}_n),\boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\widehat{\boldsymbol{\gamma}}_n)\}$. The delta method yields

$$\sqrt{n}\begin{pmatrix} \widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau;1,\boldsymbol{X}^{(-j)},\widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma}) \\ \widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau;0,\boldsymbol{X}^{(-j)},\widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma}) \\ \frac{1}{n}\sum_{i=1}^n \Big[(1,\boldsymbol{x}_i^{0,(-j)})^\top\boldsymbol{\beta}\circ\Gamma(\tau;1,\boldsymbol{x}_i^{0,(-j)},\boldsymbol{\gamma})I\{\tau>1-\pi(\boldsymbol{\gamma},1,\boldsymbol{x}_i^{0,(-j)})\} \\ -(0,\boldsymbol{x}_i^{0,(-j)})^\top\boldsymbol{\beta}\circ\Gamma(\tau;0,\boldsymbol{x}_i^{0,(-j)},\boldsymbol{\gamma})I\{\tau>1-\pi(\boldsymbol{\gamma},0,\boldsymbol{x}_i^{0,(-j)})\}\Big] - \Delta_\tau(X_j;1,0) \end{pmatrix}$$

$$\to_d N\left\{\boldsymbol{0}, \begin{pmatrix} \boldsymbol{V}_1^{'} & \boldsymbol{V}_{1,0}^{'} & \boldsymbol{0}_{p\times1} \\ \boldsymbol{V}_{0,1}^{'} & \boldsymbol{V}_0^{'} & \boldsymbol{0}_{p\times1} \\ \boldsymbol{0}_{p\times1}^\top & \boldsymbol{0}_{p\times1}^\top & \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}} \end{pmatrix}\right\},$$

$$\boldsymbol{V}_1^{'} = \boldsymbol{V}_{\boldsymbol{\beta}_1} + \{1-\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\}^2\{1-\pi(\boldsymbol{\gamma},1,\boldsymbol{X}^{(-j)})\}^2(1,\boldsymbol{X}^{(-j)})\boldsymbol{D}_{1,\gamma}^{-1}(1,\boldsymbol{X}^{(-j)})^\top$$
$$(1,\boldsymbol{X}^{(-j)})\dot{\boldsymbol{\beta}}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\dot{\boldsymbol{\beta}}\circ\Gamma(\tau;1,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})^\top(1,\boldsymbol{X}^{(-j)})^\top,$$

$$\boldsymbol{V}_0{}' = \boldsymbol{V}_{\boldsymbol{\beta}_0} + \{1-\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\}^2\{1-\pi(\boldsymbol{\gamma},0,\boldsymbol{X}^{(-j)})\}^2(0,\boldsymbol{X}^{(-j)})\boldsymbol{D}_{1,\gamma}^{-1}(0,\boldsymbol{X}^{(-j)})^\top$$
$$(0,\boldsymbol{X}^{(-j)})\dot{\boldsymbol{\beta}}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})\dot{\boldsymbol{\beta}}\circ\Gamma(\tau;0,\boldsymbol{X}^{(-j)},\boldsymbol{\gamma})^\top(0,\boldsymbol{X}^{(-j)})^\top,$$

$$\boldsymbol{V}_{1,0}^{'} = \boldsymbol{V}_{0,1}^{'T} = \boldsymbol{V}_{\boldsymbol{\beta}_1,\boldsymbol{\beta}_0} + \boldsymbol{V}_\gamma.$$

Applying the delta method again,

$$\sqrt{n}\int h(\boldsymbol{\theta},\boldsymbol{X}^{(-j)})d(\mathbb{P}_{n\,\boldsymbol{X}^{0,(-j)}} - P_{\boldsymbol{X}^{0,(-j)}}) + \sqrt{n}\{h(\widehat{\boldsymbol{\theta}}_n,\boldsymbol{X}^{(-j)}) - h(\boldsymbol{\theta},\boldsymbol{X}^{(-j)})\} \to_d G(\boldsymbol{X}^{(-j)}),$$

where $G(\boldsymbol{X}^{(-j)})$ is a Gaussian distribution indexed by $\boldsymbol{X}^{(-j)}$ whose variance is

$$
\begin{aligned}
&(1,\boldsymbol{X}^{(-j)})\boldsymbol{V}_1'(1,\boldsymbol{X}^{(-j)})^\top I\{\tau>1-\pi(\gamma,1,\boldsymbol{X}^{(-j)})\}\\
&\qquad +(0,\boldsymbol{X}^{(-j)})\boldsymbol{V}_0'(0,\boldsymbol{X}^{(-j)})^\top I\{\tau>1-\pi(\gamma,0,\boldsymbol{X}^{(-j)})\}\\
&-2(1,\boldsymbol{X}^{(-j)})\boldsymbol{V}_{1,0}'(0,\boldsymbol{X}^{(-j)})^\top I\{\tau>1-\pi(\gamma,1,\boldsymbol{X}^{(-j)}),\tau>1-\pi(\gamma,0,\boldsymbol{X}^{(-j)})\}+\sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}}.
\end{aligned}
$$

By the Donsker property of $\mathbb{G}_n h(\boldsymbol{\gamma},\boldsymbol{X}^{(-j)})$, the above convergence is uniform as a random process indexed by $\boldsymbol{X}^{(-j)}$. Therefore,

$$
\sqrt{n}(\widehat{\Delta}_\tau(X_j;1,0)-\Delta_\tau(X_j;1,0))\to_d \int G(\boldsymbol{X}^{(-j)})\ dP_{\boldsymbol{X}^{(-j)}}=N(0,\sigma^2),
$$

where

$$
\sigma^2=\int\int \mathrm{Cov}\{G(\boldsymbol{X}^{(-j)}),G(\boldsymbol{X}^{*(-j)})\}\ dP_{\boldsymbol{X}^{(-j)}}dP_{\boldsymbol{X}^{*(-j)}}
$$

for independent and identically distributed random vectors $\boldsymbol{X}^{(-j)}$ and $\boldsymbol{X}^{*(-j)}$ from the distribution $P_{\boldsymbol{X}^{(-j)}}$.

To simplify the expression for σ^2, without loss of generality we assume $\pi(\gamma,1,\boldsymbol{X}^{(-j)})>\pi(\gamma,0,\boldsymbol{X}^{(-j)})$ a.s. $\mathbb{P}_{\boldsymbol{X}^{(-j)}}$, which means that the outcome Y is less likely to have zero-inflation when $X_j=1$ than when $X_j=0$. Define

$$
\begin{aligned}
\Omega_1 &= \left\{\boldsymbol{X}^{(-j)}:\tau\le 1-\pi(\gamma,1,\boldsymbol{X}^{(-j)})\right\},\\
\Omega_2 &= \left\{\boldsymbol{X}^{(-j)}:1-\pi(\gamma,1,\boldsymbol{X}^{(-j)})<\tau\le 1-\pi(\gamma,0,\boldsymbol{X}^{(-j)})\right\},\\
\Omega_3 &= \left\{\boldsymbol{X}^{(-j)}:\tau> 1-\pi(\gamma,0,\boldsymbol{X}^{(-j)})\right\}.
\end{aligned}
$$

Let

$$
\begin{aligned}
G_1(\boldsymbol{X}^{(-j)}) &= \sqrt{n}(1, \boldsymbol{X}^{(-j)})\{\widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau; 1, \boldsymbol{X}^{(-j)}, \widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta} \circ \Gamma(\tau; 1, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma})\}, \\
G_0(\boldsymbol{X}^{(-j)}) &= \sqrt{n}(0, \boldsymbol{X}^{(-j)})\{\widehat{\boldsymbol{\beta}}_n \circ \Gamma(\tau; 0, \boldsymbol{X}^{(-j)}, \widehat{\boldsymbol{\gamma}}_n) - \boldsymbol{\beta} \circ \Gamma(\tau; 0, \boldsymbol{X}^{(-j)}, \boldsymbol{\gamma})\}, \\
G_{\boldsymbol{X}^{0,(-j)}} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big[(1, \boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta} \circ \Gamma(\tau; 1, \boldsymbol{x}_i^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 1, \boldsymbol{x}_i^{0,(-j)})\} \\
&- (0, \boldsymbol{x}_i^{0,(-j)})\boldsymbol{\beta} \circ \Gamma(\tau; 0, \boldsymbol{x}_i^{0,(-j)}, \boldsymbol{\gamma}) I\{\tau > 1 - \pi(\boldsymbol{\gamma}, 0, \boldsymbol{x}_i^{0,(-j)})\}\Big].
\end{aligned}
$$

Then, the double integral in the expression of σ^2 can be decomposed into a sum of integrals over the 9 subregions $\Omega_i \times \Omega_j, i, j = 1, 2, 3$. Specifically the expressions of $\mathrm{Cov}\{G(\boldsymbol{X}^{(-j)}), G(\boldsymbol{X}^{*(-j)})\}$ are

$$
\begin{aligned}
\Omega_1 \times \Omega_1 &: \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}}, \\
\Omega_2 \times \Omega_2 &: \mathrm{Cov}\{G_1(\boldsymbol{X}^{(-j)}), G_1(\boldsymbol{X}^{*(-j)})\} + \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}}, \\
\Omega_3 \times \Omega_3 &: \mathrm{Cov}\{G_1(\boldsymbol{X}^{(-j)}), G_1(\boldsymbol{X}^{*(-j)})\} + \mathrm{Cov}\{G_0(\boldsymbol{X}^{(-j)}), G_0(\boldsymbol{X}^{*(-j)})\} \\
&\quad -2\,\mathrm{Cov}\{G_1(\boldsymbol{X}^{(-j)}), G_0(\boldsymbol{X}^{*(-j)})\} + \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}}, \\
(\Omega_1 \times \Omega_2) \cup (\Omega_2 \times \Omega_1) &: \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}}, \\
(\Omega_1 \times \Omega_3) \cup (\Omega_3 \times \Omega_1) &: \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}}, \\
(\Omega_2 \times \Omega_3) \cup (\Omega_3 \times \Omega_2) &: \mathrm{Cov}\{G_1(\boldsymbol{X}^{(-j)}), G_1(\boldsymbol{X}^{*(-j)})\} \\
&\quad - \mathrm{Cov}\{G_1(\boldsymbol{X}^{(-j)}), G_0(\boldsymbol{X}^{*(-j)})\} + \sigma^2_{\tau,\boldsymbol{\beta},\boldsymbol{\gamma}}.
\end{aligned}
$$

When $P_{\boldsymbol{X}^{(-j)}}$ is not absolutely continuous with respect to the Lebesgue measure, as long as $\Delta_\tau(X_j; 1, 0)$ is well defined, it can be shown that the sequence of random processes (S1.2) indexed by $\boldsymbol{X}^{(-j)}$ converges uniformly to a tight process $G(\boldsymbol{X}^{(-j)})$, although not necessarily Gaussian. Then, by the almost surely continuous mapping theorem, $\sqrt{n}\Big(\widehat{\Delta}_\tau(X_j;1,0)-\Delta_\tau(X_j;1,0)\Big)$ converges in distribution to $\int G^*(\boldsymbol{X}^{(-j)})\; dP_{\boldsymbol{X}^{(-j)}}$, a distribution not necessarily normal.

$\square$

Bibliography

Csörgö, M. and P. Révész (1978). Strong approximations of the quantile process. *The Annals of Statistics 6*, 882–894.

Shorack, G. and J. Wellner (1986). Empirical processes with applications to statistics. *John Wiley & Sons, Inc: New York*.

S2 Additional figures and tables

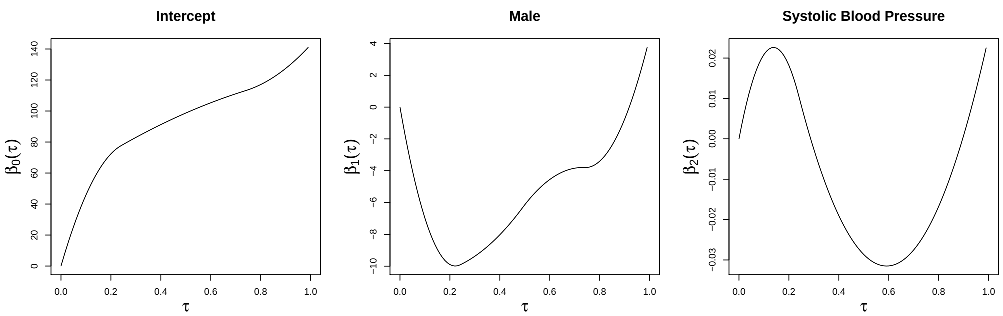


Figure S1: The true coefficient functions $\beta_0(\tau)$, $\beta_1(\tau)$ and $\beta_2(\tau)$ in simulation studies.

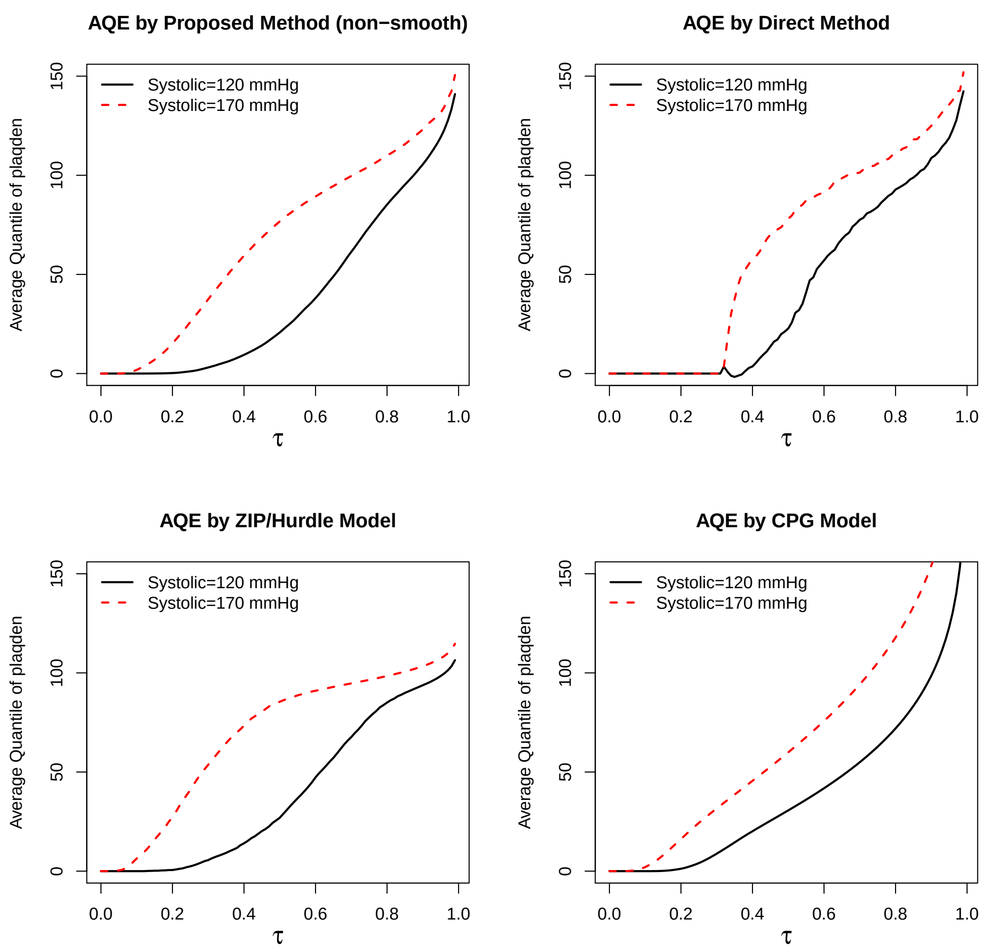
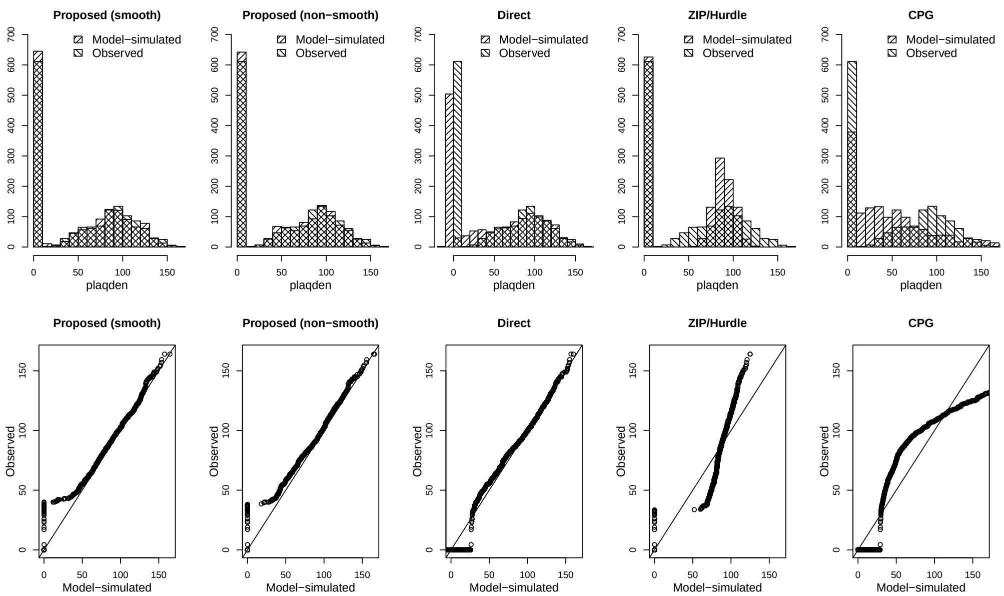
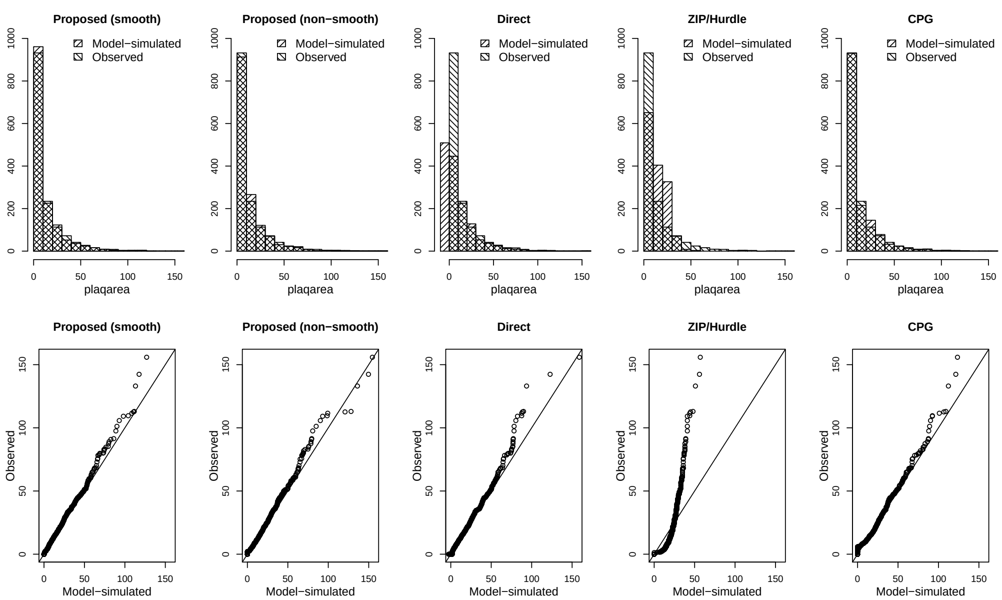


Figure S2: Comparison of the proposed method (without smoothing) with the direct quantile regression, ZIP / hurdle and CPG models in estimating the AQE of increasing systolic blood pressure from 120 mmHg to 170 mmHg on all quantiles of echodensity.



(a) Model fitness for the echodensity data.



(b) Model fitness for the plaque area data.

Figure S3: Model fitness for the NOMAS data.

Table S1: Summary of RIMSE(%), RIBias^2(%), and RIVar(%) of the estimated conditional quantile functions of echodensity around the change point of $Q_Y(\tau|\boldsymbol{X})$ (from zero to positive).

	Proposed (smooth)			Proposed (non-smooth)			Direct		
(gender, systolic)	RIMSE	RIBias^2	RIVar	RIMSE	RIBias^2	RIVar	RIMSE	RIBias^2	RIVar
$(0, 130.78)$	0.30	0.03	0.27	0.33	0.03	0.31	0.44	0.29	0.15
$(0, 139.88)$	0.20	0.01	0.19	0.24	0.01	0.22	0.26	0.13	0.13
$(0, 150.00)$	0.16	0.01	0.15	0.19	0.01	0.18	0.17	0.06	0.11
$(0, 160.12)$	0.18	0.01	0.17	0.22	0.01	0.21	0.16	0.03	0.13
$(0, 169.22)$	0.21	0.01	0.20	0.27	0.01	0.25	0.24	0.10	0.14
$(1, 130.78)$	0.28	0.02	0.26	0.32	0.02	0.29	0.30	0.13	0.17
$(1, 139.88)$	0.20	0.01	0.19	0.23	0.01	0.22	0.22	0.10	0.13
$(1, 150.00)$	0.17	0.01	0.16	0.19	0.01	0.18	0.14	0.03	0.11
$(1, 160.12)$	0.16	0.01	0.15	0.19	0.01	0.18	0.17	0.05	0.12
$(1, 169.22)$	0.16	0.01	0.16	0.21	0.01	0.20	0.33	0.21	0.12
	ZIP			Hurdle			CPG		
(gender, systolic)	RIMSE	RIBias^2	RIVar	RIMSE	RIBias^2	RIVar	RIMSE	RIBias^2	RIVar
$(0, 130.78)$	1.44	0.69	0.75	1.44	0.69	0.75	0.74	0.72	0.02
$(0, 139.88)$	1.48	0.85	0.62	1.48	0.85	0.62	0.72	0.71	0.02
$(0, 150.00)$	1.55	0.99	0.56	1.55	0.99	0.56	0.68	0.67	0.01
$(0, 160.12)$	1.82	1.15	0.68	1.82	1.15	0.68	0.71	0.69	0.02
$(0, 169.22)$	2.00	1.19	0.80	2.00	1.19	0.80	0.66	0.62	0.04
$(1, 130.78)$	1.91	0.98	0.93	1.91	0.98	0.93	0.82	0.81	0.02
$(1, 139.88)$	2.10	1.27	0.83	2.10	1.27	0.83	0.87	0.85	0.02
$(1, 150.00)$	2.39	1.60	0.80	2.39	1.60	0.80	0.88	0.84	0.04
$(1, 160.12)$	2.38	1.60	0.78	2.38	1.60	0.78	0.70	0.65	0.05
$(1, 169.22)$	2.20	1.45	0.75	2.20	1.45	0.75	0.52	0.47	0.05

Table S2: Comparison of the correctly predicted rate (%) for the zero outcomes, coverage of the 95% prediction upper bound (%), and its average length for positive outcomes in predicting echodensity or plaque area by the proposed quantile methods, direct quantile regression, ZIP/hurdle and CPG models.

	Measure	Proposed (smooth)	Proposed (non-smooth)	Direct	ZIP/Hurdle	CPG
plaqden = 0	Correctly predicted rate (%)	45.86	47.15	1.50	45.86	13.09
plaqden > 0	Coverage of 95% prediction upper bound (%)	96.23	95.48	95.62	82.19	98.42
	Length of prediction bound	128.26	127.10	127.16	102.68	145.73

	Measure	Proposed (smooth)	Proposed (non-smooth)	Direct	ZIP/Hurdle	CPG
plaqarea = 0	Correctly predicted rate (%)	45.02	47.12	6.76	45.02	92.85
plaqarea > 0	Coverage of 95% prediction upper bound (%)	93.70	93.70	93.63	84.25	94.45
	Length of prediction bound	46.92	45.71	45.71	25.12	44.85