Web-based Supplementary Materials for "Assessing quantile prediction with censored quantile regression models," by Ruosha Li and Limin Peng

Supplementary Materials A: Asymptotic Properties of $\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}})$ and $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$

Define $\boldsymbol{\mu}(\tau, \boldsymbol{b}) = E[\boldsymbol{Z}'\{I(T \leqslant \boldsymbol{Z}\boldsymbol{b}) - \tau\}]$ and let $\widetilde{\boldsymbol{\beta}}(\tau)$ be the solution to $\boldsymbol{\mu}(\tau, \boldsymbol{b}) = 0$. We require the following mild regularity conditions, which are essentailly very similar to those required in Zhou (2006) and Peng and Fine (2009).

C1 The covariates space $\mathcal{Z}$ is bounded.

C2 $\Pr(C > u) > 0$, where C is the log censoring time.

C3 (i) Define $F(t|\boldsymbol{Z}) = \Pr(T \leqslant t|\boldsymbol{Z})$ and $f(t|\boldsymbol{Z}) = dF(t|\boldsymbol{Z})/dt$. The density function $f(t|\boldsymbol{Z})$ is bounded from above uniformly for t and $\boldsymbol{Z} \in \mathcal{Z}$. (ii) For a large constant $M_b > 0$, define $\mathcal{B} = \{\boldsymbol{b} : ||\boldsymbol{b}|| \leqslant M_b \text{ and } \max_{\mathcal{Z}} \boldsymbol{Z}\boldsymbol{b} \leqslant u\}$. For $\tau \in [\tau_L, \tau_U]$, $\widetilde{\boldsymbol{\beta}}(\tau) \in \mathcal{B}$ and is uniformly bounded and Lipschitz continuous in τ.

C4 Let $\boldsymbol{J}(\tau, \boldsymbol{b}) = \partial \boldsymbol{\mu}(\tau, \boldsymbol{b})/\partial \boldsymbol{b}$, $\inf_{\tau \in [\tau_L, \tau_U]} \text{eigmin}\{\boldsymbol{J}(\tau, \widetilde{\boldsymbol{\beta}})\} > 0$, where eigmin$(\cdot)$ denotes the minimum eigenvalue of a matrix.

Important notations used in the manuscript include:

- $L(\tau, \boldsymbol{\beta}) = E\rho_\tau\{Y^{0u} - \boldsymbol{Z}^0\boldsymbol{\beta}(\tau)\}$ is the expected prediction loss for a certain $\boldsymbol{\beta}(\tau)$.
- $L(\tau) = L(\tau, \widetilde{\boldsymbol{\beta}})$, where $\widetilde{\boldsymbol{\beta}}(\tau)$ satisfies $\boldsymbol{\mu}\{\tau, \widetilde{\boldsymbol{\beta}}(\tau)\} = 0$.
- ζ_τ is the true unconditional τ_{th} quantile of T^0, and $L_0(\tau) = E\rho_\tau(T^{0u} - \zeta_\tau)$.
- $R^1(\tau) \equiv R^1(\tau, \widetilde{\boldsymbol{\beta}}) = 1 - \dfrac{L(\tau)}{L_0(\tau)}$.
- $\hat{L}_n(\tau, \boldsymbol{\beta}) = n^{-1}\sum_{i=1}^n \Delta_i/\hat{G}(Y_i^u)\rho_\tau\{Y_i^u - \boldsymbol{Z}_i\boldsymbol{\beta}(\tau)\}$ is the empirical loss function evaluated at a certain $\boldsymbol{\beta}(\tau)$. Also write $\hat{L}_n^G(\tau, \boldsymbol{\beta}) = n^{-1}\sum_{i=1}^n \Delta_i/G(Y_i^u)\rho_\tau\{Y_i^u - \boldsymbol{Z}_i\boldsymbol{\beta}(\tau)\}$.
- $\hat{R}_n^1(\tau, \widehat{\boldsymbol{\beta}}) = [\hat{L}_n(\tau, \hat{\zeta}_\tau) - \hat{L}_n(\tau, \widehat{\boldsymbol{\beta}})]/\hat{L}_n(\tau, \hat{\zeta}_\tau)$, where $\hat{\zeta}_\tau$ represents Zhou (2006)'s estimator in an intercept-only model.
- $\boldsymbol{S}_n(\tau, \boldsymbol{b}) = n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i' \dfrac{\Delta_i}{\hat{G}(Y_i^u)}\{I(Y_i^u \leqslant \boldsymbol{Z}\boldsymbol{b}) - \tau\}$, then $\widehat{\boldsymbol{\beta}}(\tau)$ is obtained by solving $\boldsymbol{S}_n(\tau, \boldsymbol{\beta}) = 0$.

- Let $\{\omega_i\}_{i=1}^n$ be i.i.d. unit Exponential random variables. The perturbed loss function is $\hat{L}_n^*(\tau,\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^n \omega_i \frac{\Delta_i}{G^*(Y_i^u)}\rho_\tau(Y_i^u - \boldsymbol{Z}_i\boldsymbol{\beta})$, where $G^*(\cdot)$ is a perturbed Kaplan-Meier estimator. $\widehat{\boldsymbol{\beta}}^*$ is the minimizer of this perturbed loss function.

- In the CV-type estimator $\hat{L}_{CV}(\tau,\widehat{\boldsymbol{\beta}}) = \frac{1}{K}\sum_{k=1}^K \hat{L}^k\{\tau,\widehat{\boldsymbol{\beta}}_{(-k)}\}$, $\widehat{\boldsymbol{\beta}}_{(-k)}$ is the estimator based on all the observations that satisfy $V_i \neq k$, and $\hat{L}^k\{\tau,\boldsymbol{\beta}\} = \frac{K}{n}\sum_{i=1}^n \frac{\Delta_i I(V_i = k)}{\hat{G}(Y_i^u)}\rho_\tau\{Y_i^u - \boldsymbol{Z}_i\boldsymbol{\beta}(\tau)\}$.

A.1 Proof of Theorem 1

Write $\boldsymbol{S}_n^G(\tau,\boldsymbol{b}) = n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i'\Delta_i/G(Y_i^u)\big\{I(Y_i^u \leqslant \boldsymbol{Z}_i\boldsymbol{b}) - \tau\big\}$. We see that

$$E\boldsymbol{S}_n^G(\tau,\boldsymbol{b}) = E[\boldsymbol{Z}'\{I(T_i^u \leqslant \boldsymbol{Z}_i\boldsymbol{b}) - \tau\}] = \boldsymbol{\mu}(\tau,\boldsymbol{b}), \quad \forall \boldsymbol{b} \in \mathcal{B}.$$

Let $\mathcal{F} = \{\boldsymbol{Z}_i'\Delta_i/G(Y_i^u)\big\{I(Y_i^u \leqslant \boldsymbol{Z}_i\boldsymbol{b}) - \tau\big\} : \boldsymbol{Z}_i \in \mathcal{Z}, \boldsymbol{b} \in \mathcal{B}, \tau \in [\tau_L,\tau_U]\}$. Under **C1-C2**, $\boldsymbol{Z}_i$ and $G^{-1}(Y_i^u)$ are both bounded from above. Thus, the functional class $\mathcal{F}$ can be shown to be Donsker. This is because that the class of indicator functions is Donsker, and that Donsker's property is preserved under Lipschitz transformations (Van Der Vaart and Wellner, 2000). Since Donsker's property implies Glivenko-cantelli, we get

$$\sup_{\tau,\boldsymbol{b}\in\mathcal{B}} ||\boldsymbol{S}_n^G(\tau,\boldsymbol{b}) - \boldsymbol{\mu}(\tau,\boldsymbol{b})|| = o_p(1), \tag{A.1}$$

Write $N_i^G(t) = I(Y_i \leqslant t, \delta_i = 0)$, $Y_i^G(t) = I(Y_i \geqslant t)$, $y^G(t) = EY_i^G(t)$ and $M_i^G(t) = N_i^G(t) - \int_0^t Y_i^G(s)d\Lambda^G(s)$, where $\Lambda^G(\cdot)$ denotes the cumulative hazard of C. Let $g_i(t) = -G(t)\int_0^t y^G(s)^{-1}dM_i^G(s)$, we have $n^{1/2}\{\hat{G}(t) - G(t)\} = n^{-1/2}\sum_{i=1}^n g_i(t) + o_p^{t\in[0,u]}(1)$ for the Kaplan-Meier estimator $\hat{G}(\cdot)$ (Pepe, 1991), where $o_p^{t\in[0,u]}(1)$ represents a term that converges to 0 in probability uniformly for $t \in [0,u]$. Moreover, $\mathcal{F}_G = \{g_i(t) : t \in [0,u]\}$ can be shown to be a Donsker's class. It follows that

$$n^{1/2}\{\hat{G}^{-1}(t) - G^{-1}(t)\} = -n^{-1/2}\sum_{i=1}^n \frac{g_i(t)}{G^2(t)} + o_p^{t\in[0,u]}(1), \tag{A.2}$$

where $\mathcal{F}_{G1} = \{-g_i(t)/G^2(t) : t \in [0,u]\}$ remains Donsker, due to the fact that $G^{-2}(t)$ is bounded from above uniformly for $t \in [0,u]$. When combined with an application of the Glivenko-cantelli theorem on $\mathcal{F}_{G1}$, this fact further implies that $\sup_{t\in[0,u]} ||\hat{G}^{-1}(t) - G^{-1}(t)|| = o_p(1)$ and that

$$\sup_{\tau,\boldsymbol{b}\in\mathcal{B}} ||\boldsymbol{S}_n(\tau,\boldsymbol{b}) - \boldsymbol{S}_n^G(\tau,\boldsymbol{b})|| = o_p(1). \tag{A.3}$$

Noting $\boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}}) = 0$ and $\boldsymbol{S}_n(\tau, \widehat{\boldsymbol{\beta}}) = o_p^\tau(1)$, with $o_p^\tau(1)$ denoting a term that converges to 0 in probability uniformly in $\tau \in [\tau_L, \tau_U]$, we consider the following manipulation,

$$\boldsymbol{\mu}(\tau, \widehat{\boldsymbol{\beta}}) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}}) = \boldsymbol{\mu}(\tau, \widehat{\boldsymbol{\beta}}) - \boldsymbol{S}_n^G(\tau, \widehat{\boldsymbol{\beta}}) + \boldsymbol{S}_n^G(\tau, \widehat{\boldsymbol{\beta}}) - \boldsymbol{S}_n(\tau, \widehat{\boldsymbol{\beta}}) + o_p^\tau(1). \tag{A.4}$$

The combination of (A.1), (A.3) and (A.4) implies $\sup_\tau ||\boldsymbol{\mu}(\tau, \widehat{\boldsymbol{\beta}}) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}})|| = o_p(1)$. Since eigmin$\{\boldsymbol{J}(\tau, \widetilde{\boldsymbol{\beta}})\}$ is bounded from below uniformly in τ, as required in **C4**, we can use Taylor expansion and get

$$\sup_{\tau \in [\tau_L, \tau_U]} ||\widehat{\boldsymbol{\beta}}(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)|| = o_p(1). \tag{A.5}$$

To show the consistency of $\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}})$, we decompose $\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau)$ as follows:

$$\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau) = \{\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}^G(\tau, \widehat{\boldsymbol{\beta}})\} + \{\hat{L}^G(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widehat{\boldsymbol{\beta}})\} + \{L(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}})\}, \tag{A.6}$$

where $L(\tau, \widetilde{\boldsymbol{\beta}}) = L(\tau)$ by definition. From the uniform consistency of $\hat{G}(\cdot)$ and the uniform boundedness of $\rho_\tau\{Y_i^u - \boldsymbol{Z}_i \boldsymbol{b}\}$ for $\boldsymbol{b} \in \mathcal{B}$, we have

$$\sup_{\tau, \boldsymbol{b} \in \mathcal{B}} |\hat{L}_n(\tau, \boldsymbol{b}) - \hat{L}^G(\tau, \boldsymbol{b})| = o_p(1).$$

Write $\mathcal{F}_L = \{\Delta_i \rho_\tau(Y_i^u - \boldsymbol{Z}_i \boldsymbol{b})/G(Y_i^u) : \boldsymbol{Z} \in \mathcal{Z}, \boldsymbol{b} \in \mathcal{B}, \tau \in [\tau_L, \tau_U]\}$. Under **C1-C4**, we can follow the arguments above and apply the Glivenko-cantelli to see

$$\sup_{\tau, \boldsymbol{b}} |\hat{L}^G(\tau, \boldsymbol{b}) - L(\tau, \boldsymbol{b})| = o_p(1).$$

Furthermore, a Taylor expansion on $L(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}})$, coupled with (A.5) and that $\frac{\partial}{\partial \boldsymbol{b}} L(\tau, \boldsymbol{b})\big|_{\boldsymbol{b} = \widetilde{\boldsymbol{\beta}}(\tau)} = 0$, implies that $L(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}}) = o_p^\tau(1)$. Finally, we can combine these facts with (A.6) to get

$$\sup_{\tau} |\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau)| = o_p(1). \tag{A.7}$$

Following the same argument, $\hat{L}_n(\tau, \hat{\zeta}_\tau)$ in $\hat{R}_n^1(\tau, \widehat{\boldsymbol{\beta}})$ is also uniformly consistent to $L_0(\tau)$. When combined with (A.7), this further ensures the uniform consistency of $\hat{R}_n^1(\tau, \widehat{\boldsymbol{\beta}})$ for $\tau \in [\tau_L, \tau_U]$.

A.2 Proof of Theorem 2

Since $\boldsymbol{S}_n(\tau,\widehat{\boldsymbol{\beta}}) = o_p(n^{-1/2})$ and $\boldsymbol{\mu}(\tau,\widetilde{\boldsymbol{\beta}}) = 0$, we re-write $-\sqrt{n}\{\boldsymbol{S}_n(\tau,\widehat{\boldsymbol{\beta}}) - \boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}})\}$ as

$$-\sqrt{n}\{\boldsymbol{S}_n(\tau,\widehat{\boldsymbol{\beta}}) - \boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}})\} = \sqrt{n}\{\boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}}) - \boldsymbol{S}_n^G(\tau,\widetilde{\boldsymbol{\beta}})\} + \sqrt{n}\{\boldsymbol{S}_n^G(\tau,\widetilde{\boldsymbol{\beta}}) - \boldsymbol{\mu}(\tau,\widetilde{\boldsymbol{\beta}})\} + o_p^\tau(1). \quad \text{(A.8)}$$

For the first term on the right-handside, we can utilize (A.2) to get

$$\begin{aligned}
\sqrt{n}\{\boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}}) - \boldsymbol{S}_n^G(\tau,\widetilde{\boldsymbol{\beta}})\} &= -n^{-1/2}\sum_{j=1}^n \boldsymbol{Z}_j'\Delta_j I\big[\{Y_j^u \leqslant \boldsymbol{Z}_j\widetilde{\boldsymbol{\beta}}(\tau)\} - \tau\big]\cdot\frac{1}{n}\sum_{i=1}^n \frac{g_i(Y_j^u)}{G^2(Y_j^u)} + o_p^\tau(1) \\
&= -n^{-1/2}\sum_{i=1}^n \frac{1}{n}\sum_{j=1}^n \boldsymbol{Z}_j'\Delta_j\big[I\{Y_j^u \leqslant \boldsymbol{Z}_j\widetilde{\boldsymbol{\beta}}(\tau)\} - \tau\big]\frac{g_i(Y_j^u)}{G^2(Y_j^u)} + o_p^\tau(1). \quad \text{(A.9)}
\end{aligned}$$

Define

$$\boldsymbol{\eta}_{1i}(\tau,\boldsymbol{\beta}) = -E\Big(\boldsymbol{Z}'\Delta\big[I\{Y^u \leqslant \boldsymbol{Z}\boldsymbol{\beta}(\tau)\} - \tau\big]g_i(Y^u)/G^2(Y^u)|\boldsymbol{D}_i\Big),$$

with $\boldsymbol{D}_i = (Y_i, \delta_i, \bar{\boldsymbol{Z}}_i)$ denoting the data from the i_{th} observation. It follows from an application of the Glivenko-cantelli theorem that

$$\sqrt{n}\{\boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}}) - \boldsymbol{S}_n^G(\tau,\widetilde{\boldsymbol{\beta}})\} = n^{-1/2}\sum_{i=1}^n \boldsymbol{\eta}_{1i}(\tau,\widetilde{\boldsymbol{\beta}}) + o_p^\tau(1). \quad \text{(A.10)}$$

Note that $\boldsymbol{\eta}_{1i}(\tau,\widetilde{\boldsymbol{\beta}})$ accounts for the uncertainty in $\boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}})$ due to the estimated inverse weights $\hat{G}(Y_i^u)$. Now write $\boldsymbol{\eta}_{2i}(\tau,\boldsymbol{\beta}) = \boldsymbol{Z}_i'\Delta_i/G(Y_i^u)\big[I\{Y_i^u \leqslant \boldsymbol{Z}_i\boldsymbol{\beta}(\tau)\} - \tau\big] - \boldsymbol{\mu}(\tau,\boldsymbol{\beta})$ and $\boldsymbol{\eta}_i(\tau,\boldsymbol{\beta}) = \boldsymbol{\eta}_{2i}(\tau,\boldsymbol{\beta}) + \boldsymbol{\eta}_{2i}(\tau,\boldsymbol{\beta})$, we can combine (A.8) and (A.10) to see

$$\sqrt{n}\boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}}) = n^{-1/2}\sum_{i=1}^n \boldsymbol{\eta}_i(\tau,\widetilde{\boldsymbol{\beta}}) + o_p^\tau(1). \quad \text{(A.11)}$$

The influence functions, $\boldsymbol{\eta}_i(\tau,\widetilde{\boldsymbol{\beta}})$, can be consistently estimated by plugging in their estimated counterpart. Furthermore, we can follow the lines of Peng and Fine (2009) to get

$$\sqrt{n}\{\boldsymbol{S}_n\{\tau,\widehat{\boldsymbol{\beta}}\} - \boldsymbol{S}_n(\tau,\widetilde{\boldsymbol{\beta}})\} = \sqrt{n}[\boldsymbol{\mu}(\tau,\widehat{\boldsymbol{\beta}}) - \boldsymbol{\mu}(\tau,\widetilde{\boldsymbol{\beta}})] + o_p^\tau(1), \quad \text{(A.12)}$$

using (A.5) and condition **C3**. We can then use (A.11) to see $\sqrt{n}\{\boldsymbol{\mu}(\tau,\widehat{\boldsymbol{\beta}})-\boldsymbol{\mu}(\tau,\widetilde{\boldsymbol{\beta}})\} = -n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_i(\tau,\widetilde{\boldsymbol{\beta}})+$ $o_p^\tau(1)$, and moreover use Taylor expansion and condition **C4** to get

$$\sqrt{n}\{\widehat{\boldsymbol{\beta}}(\tau)-\widetilde{\boldsymbol{\beta}}(\tau)\} = -n^{-1/2}\sum_{i=1}^{n}\boldsymbol{J}^{-1}(\tau,\widetilde{\boldsymbol{\beta}})\boldsymbol{\eta}_i(\tau,\widetilde{\boldsymbol{\beta}}) + o_p^\tau(1). \tag{A.13}$$

Therefore, we see that $\sqrt{n}\{\widehat{\boldsymbol{\beta}}(\tau)-\widetilde{\boldsymbol{\beta}}(\tau)\}$ still converges weakly to a centered Gaussian process when the working quantile regression model is misspecified, despite that $\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}(\tau)$ may no longer correspond to the true conditional quantile.

Next, we utilize (A.6) again to study the distributional properties of $\hat{L}_n(\tau,\widehat{\boldsymbol{\beta}})$. Define

$$\pi_{1i}(\tau,\boldsymbol{b}) = -E\big[\frac{\Delta g_i(Y^u)}{G^2(Y^u)}\rho_\tau\{Y^u-\boldsymbol{Z}\widetilde{\boldsymbol{\beta}}(\tau)\}|\ \mathbf{D}_i\big],\ \pi_{2i}(\tau,\boldsymbol{b}) = \Delta_i\rho_\tau(Y_i^u-\boldsymbol{Z}_i\boldsymbol{b})/G(Y_i^u) - L(\tau,\boldsymbol{b}), \tag{A.14}$$

and $\pi_i(\tau,\boldsymbol{b}) \equiv \pi_{1i}(\tau,\boldsymbol{b}) + \pi_{2i}(\tau,\boldsymbol{b})$. With similar arguments as those used in (A.9), and noting that $\widetilde{\boldsymbol{\beta}}(\tau)$ is the minimizer of $L(\tau,\boldsymbol{b})$, we can show that the third term in (A.6) is asymptotically neglegible and that

$$\sqrt{n}\{\hat{L}_n(\tau,\widehat{\boldsymbol{\beta}}) - L(\tau)\} = n^{-1/2}\sum_{i=1}^{n}\pi_i(\tau,\widetilde{\boldsymbol{\beta}}) + o_p^\tau(1), \tag{A.15}$$

where $E\pi_i(\tau,\widetilde{\boldsymbol{\beta}}) = 0$ and $\{\pi_i(\tau,\boldsymbol{b}) : \tau\in[\tau_L,\tau_U], \boldsymbol{b}\in\mathcal{B}\}$ forms a Donsker's class. Thus we see that $\sqrt{n}\{\hat{L}_n(\tau,\widehat{\boldsymbol{\beta}}) - L(\tau)\}$ converges weakly to a zero-mean Gaussian process.

To derive the asymptotic distribution of $\hat{R}_n^1(\tau,\widehat{\boldsymbol{\beta}})$, we first write $\pi_{i0}(\tau)$ as the influence functions for $\sqrt{n}\{\hat{L}_n(\tau,\hat{\zeta}_\tau) - L_0(\tau)\}$. Using the continuous mapping theorem and the functional delta method, it is easy to see that $\sqrt{n}\{\hat{R}_n^1(\tau,\widehat{\boldsymbol{\beta}}) - R_n^1(\tau,\widetilde{\boldsymbol{\beta}})\}$ also converge weakly to a zero-mean Gaussian process, and the influence function equals

$$r_i(\tau,\widetilde{\boldsymbol{\beta}}) = -\frac{\pi_i(\tau,\widetilde{\boldsymbol{\beta}})}{L_0(\tau)} + \frac{\pi_{i0}(\tau)\times L(\tau,\widetilde{\boldsymbol{\beta}})}{L_0(\tau)^2}. \tag{A.16}$$

This completes the proof of Theorem 2.

A.3 Asymptotic Results for $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$

Recall that $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}}) = K^{-1} \sum_{k=1}^{K} \hat{L}^k(\tau, \widehat{\boldsymbol{\beta}}_{(-k)})$, where

$$\hat{L}^k(\tau, \boldsymbol{\beta}) = \frac{K}{n} \sum_{i=1}^{n} \frac{\Delta_i I(V_i = k)}{\hat{G}(Y_i^u)} \rho_\tau \{Y_i^u - \boldsymbol{Z}_i \boldsymbol{\beta}(\tau)\},\ k = 1, 2, ..., K,$$

and K is a small fixed integer such that $K = O(1)$ and $n/K = O(n)$. Following Sections A.1–A.2, for $k = 1, 2, ..., K$, $\widehat{\boldsymbol{\beta}}_{(-k)}(\tau)$ are all uniformly consistent estimators for $\widetilde{\boldsymbol{\beta}}(\tau)$ for $\tau \in [\tau_L, \tau_U]$, and that $(\frac{nK-n}{K})^{1/2}\{\widehat{\boldsymbol{\beta}}_{(-k)}(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)\}$ converges weakly to a zero-mean Gaussian process. The following arguments are similar to those in Appendix 3 of Tian et al. (2007). Specifically, define

$$\hat{L}^{kG}(\tau, \boldsymbol{\beta}) = \frac{K}{n} \sum_{i=1}^{n} \frac{\Delta_i I(V_i = k)}{G(Y_i^u)} \rho_\tau \{Y_i^u - \boldsymbol{Z}_i \boldsymbol{\beta}(\tau)\}. \tag{A.17}$$

Also write $\hat{L}_{CV}^G(\tau, \widehat{\boldsymbol{\beta}}) = K^{-1} \sum_{k=1}^{K} \hat{L}^{kG}(\tau, \widehat{\boldsymbol{\beta}}_{(-k)})$, and $L_{CV}(\tau, \widehat{\boldsymbol{\beta}}) = K^{-1} \sum_{k=1}^{K} L(\tau, \widehat{\boldsymbol{\beta}}_{(-k)})$. We write

$$\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}}) = \hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}_{CV}^G(\tau, \widehat{\boldsymbol{\beta}}) + \hat{L}_{CV}^G(\tau, \widehat{\boldsymbol{\beta}}) - L_{CV}(\tau, \widehat{\boldsymbol{\beta}}) + L_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}}). \tag{A.18}$$

We first provide a sketch proof for the uniform consistency of $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$ to $L(\tau)$. Because $\partial L(\tau, \boldsymbol{b})\big|_{\boldsymbol{b}=\widetilde{\boldsymbol{\beta}}(\tau)} = 0$, the uniform consistency of $\widehat{\boldsymbol{\beta}}_{(-k)}(\tau)$ implies that $\sup_{\tau \in [\tau_L, \tau_U]} |L_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}})| = o_p(1)$. Next, we use the consistency of $\widehat{\boldsymbol{\beta}}_{(-k)}(\tau)$ and the fact that $\sum_{k=1}^{K} I(V_i = k) = 1$ to derive

$$\begin{aligned} \hat{L}_{CV}^G(\tau, \widehat{\boldsymbol{\beta}}) - L_{CV}(\tau, \widehat{\boldsymbol{\beta}}) &= \hat{L}_{CV}^G(\tau, \widetilde{\boldsymbol{\beta}}) - L_{CV}(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1) \\ &= n^{-1} \sum_{i=1}^{n} \frac{\Delta_i}{G(Y_i^u)} \rho_\tau \{Y_i^u - \boldsymbol{Z}_i \widetilde{\boldsymbol{\beta}}(\tau)\} - L(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1), \end{aligned}$$

where the right-hand-side is $o_p(1)$ according to the Glivenko-cantelli theorem. Moreover, we have

$$\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}_{CV}^G(\tau, \widehat{\boldsymbol{\beta}}) = n^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n} \{\hat{G}^{-1}(Y_i^u) - G^{-1}(Y_i^u)\} \Delta_i I(V_i = k) \rho_\tau \{Y_i^u - \boldsymbol{Z}_i \widehat{\boldsymbol{\beta}}_{(-k)}(\tau)\},$$

which is also $o_p(1)$ according to the uniform consistency of $\hat{G}^{-1}(\cdot)$ to $G^{-1}(\cdot)$. These arguments, when combined with (A.18), gives the uniform consistency of $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$ to $L(\tau, \widetilde{\boldsymbol{\beta}})$.

To derive the asymptotic distribution of $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$, we use the consistency $\widehat{\boldsymbol{\beta}}_{(-1)}(\tau), ..., \widehat{\boldsymbol{\beta}}_{(-K)}(\tau)$

and $\sum_{k=1}^{K} I(V_i = k) = 1$ again to derive

$$
\begin{aligned}
& \sqrt{n}\{\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}^G_{CV}(\tau, \widehat{\boldsymbol{\beta}})\} + \sqrt{n}\{\hat{L}^G_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - L_{CV}(\tau, \widehat{\boldsymbol{\beta}})\} \\
= \;& n^{-1/2} \sum_{k=1}^{K} \sum_{i=1}^{n} \Delta_i I(V_i = k) \rho_\tau\{Y_i^u - \boldsymbol{Z}_i \widetilde{\boldsymbol{\beta}}(\tau)\}\{\hat{G}^{-1}(Y_i^u) - G^{-1}(Y_i^u)\} \\
& \quad + n^{-1/2} \sum_{k=1}^{K} \sum_{i=1}^{n} I(V_i = k) \pi_{2i}\{\tau, \widehat{\boldsymbol{\beta}}_{(-k)}\} + o_p^\tau(1) \\
= \;& n^{-1/2} \sum_{i=1}^{n} \pi_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p^\tau(1). \qquad \text{(A.19)}
\end{aligned}
$$

Further, it follows from the fact that $\frac{\partial}{\partial \boldsymbol{b}} L(\tau, \boldsymbol{b})\big|_{\boldsymbol{b} = \widetilde{\boldsymbol{\beta}}(\tau)} = 0$ and the weak convergence of $(\frac{nK-n}{K})^{1/2}\{\widehat{\boldsymbol{\beta}}_{(-k)}(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)\}$ that $\sqrt{n}\{L_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}})\} = \sqrt{n}/K \sum_{k=1}^{K}\{L(\tau, \widehat{\boldsymbol{\beta}}_{(-k)}) - L(\tau, \widetilde{\boldsymbol{\beta}})\}$ is asymptotically of smaller order. The combination of these arguments gives

$$
\sqrt{n}\{\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}})\} = n^{-1/2} \sum_{i=1}^{n} \pi_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p^\tau(1),
$$

suggesting that $\sqrt{n}\{\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}})\}$ is asymptotically equivalent to $\sqrt{n}\{\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}}) - L(\tau, \widetilde{\boldsymbol{\beta}})\}$.

A.4 Justification for the Proposed Perturbation Method

Noticing that $E(\omega_i) = Var(\omega_i) = 1$, we first show that

$$
\sup_{\tau \in [\tau_L, \tau_U]} ||\widehat{\boldsymbol{\beta}}^*(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)|| = o_p(1).
$$

To see this, one can use the following algebraic manipulation and get

$$
\boldsymbol{\mu}(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}}) = \boldsymbol{\mu}(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{S}_n^{G*}(\tau, \widehat{\boldsymbol{\beta}}^*) + \boldsymbol{S}_n^{G*}(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{S}_n^{*}(\tau, \widehat{\boldsymbol{\beta}}^*) + o_p(1), \qquad \text{(A.20)}
$$

where $\boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}}) = 0$, $\boldsymbol{S}_n^{*}(\tau, \widehat{\boldsymbol{\beta}}^*) = o_p(1)$, and $\boldsymbol{S}_n^{*G}(\tau, \boldsymbol{b}) = n^{-1} \sum_{i=1}^{n} \omega_i \boldsymbol{Z}_i' \Delta_i / G(Y_i^u)\big\{I(Y_i^u \leqslant \boldsymbol{Z}_i \boldsymbol{b}) - \tau\big\}$. Following the proof in Section A.1, it is not hard to verify that $\sup_{\tau \in [\tau_L, \tau_U]} ||\boldsymbol{\mu}(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}})|| = o_p(1)$, which further entails the consistency of $\widehat{\boldsymbol{\beta}}^*(\tau)$ to $\widetilde{\boldsymbol{\beta}}(\tau)$.

The following Lemma states the distributional properties.

Lemma 1. *For i.i.d. random variables* $\{\omega_i\}_{i=1}^n$*, where* $\omega_i > 0$ *and* $E(\omega_i) = Var\ (\omega_i) = 1$*, we have*

$$\sqrt{n}\{\widehat{\boldsymbol{\beta}}^*(\tau) - \widehat{\boldsymbol{\beta}}(\tau)\} = -n^{-1/2}\sum_{i=1}^{n}(\omega_i - 1)\boldsymbol{J}^{-1}(\tau, \widetilde{\boldsymbol{\beta}})\boldsymbol{\eta}_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1). \tag{A.21}$$

and

$$\sqrt{n}\{\hat{L}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \hat{L}_n(\tau, \widehat{\boldsymbol{\beta}})\} = n^{-1/2}\sum_{i=1}^{n}(\omega_i - 1)\pi_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1). \tag{A.22}$$

Proof. Write $N^*(t) = n^{-1}\sum_{i=1}^n \omega_i N_i(t)$ and $Y^*(t) = n^{-1}\sum_{i=1}^n \omega_i Y_i(t)$, and let $G^*(t) = \mathcal{P}_{s\in[0,t]}\{1 - dN^*(t)/Y^*(t)\}$, where $\mathcal{P}$ is the product-integral operator. An application of the Duhamel's equation (Andersen, 1993) and empirical process techniques gives

$$\sqrt{n}\{G^*(t) - G(t)\} = -n^{-1/2}G(t)\sum_{i=1}^{n}\omega_i\int_0^t y^G(s)dM_i^G(s) = n^{-1/2}\sum_{i=1}^{n}\omega_i g_i(t) + o_p^{t\in[0,u]}(1), \tag{A.23}$$

where $g_i(t)$ is the influence function of $\hat{G}(t)$. Next, we obtain the following decomposition using $\boldsymbol{S}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) = o_p^\tau(n^{-1/2})$ and $\boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}}) = 0$:

$$\begin{aligned} -\sqrt{n}\{\boldsymbol{S}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{S}_n^*(\tau, \widetilde{\boldsymbol{\beta}})\} &= \sqrt{n}\{\boldsymbol{S}_n^*(\tau, \widetilde{\boldsymbol{\beta}}) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}})\} + o_p^\tau(1). \\ &= \sqrt{n}\{\boldsymbol{S}_n^*(\tau, \widetilde{\boldsymbol{\beta}}) - \boldsymbol{S}_n^{*G}(\tau, \widetilde{\boldsymbol{\beta}})\} + \sqrt{n}\{\boldsymbol{S}_n^{*G}(\tau, \widetilde{\boldsymbol{\beta}}) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}})\} + o_p^\tau(1), \end{aligned}$$

where $\boldsymbol{S}_n^*(\tau, \boldsymbol{b}) = n^{-1}\sum_{i=1}^n \omega_i \boldsymbol{Z}_i' \dfrac{\Delta_i}{G^*(Y_i^u)}\big\{I(Y_i^u \leqslant \boldsymbol{Z}_i\boldsymbol{b}) - \tau\big\}$. First, we see that

$$\sqrt{n}\{\boldsymbol{S}_n^{*G}(\tau, \widetilde{\boldsymbol{\beta}}) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}})\} = n^{-1/2}\sum_{i=1}^{n}\omega_i\boldsymbol{\eta}_{2i}(\tau, \widetilde{\boldsymbol{\beta}}). \tag{A.24}$$

Moreover, writing $\Phi_i(\tau, \boldsymbol{b}) = I\{I(Y_i^u \leqslant \boldsymbol{Z}_i\boldsymbol{b}) - \tau\}$ and using (A.23), we get

$$\begin{aligned} & \sqrt{n}\{\boldsymbol{S}_n^*(\tau, \widetilde{\boldsymbol{\beta}}) - \boldsymbol{S}_n^{*G}(\tau, \widetilde{\boldsymbol{\beta}})\} \\ = \; & n^{-1/2}\sum_{j=1}^{n} \omega_j \boldsymbol{Z}_j' \Delta_j \Phi_j(\tau, \widetilde{\boldsymbol{\beta}})\{G^{*-1}(Y_j^u) - G^{-1}(Y_j^u)\} \\ = \; & -n^{-1/2}\sum_{j=1}^{n} \omega_j \boldsymbol{Z}_j' \Delta_j \Phi_j(\tau, \widetilde{\boldsymbol{\beta}}) \frac{1}{n}\sum_{i=1}^{n} \frac{\omega_i g_i(Y_j^u)}{G^2(Y_j^u)} + o_p^{\tau}(1) \\ = \; & n^{-1/2}\sum_{i=1}^{n} \omega_i \frac{1}{n}\sum_{j=1}^{n} \frac{-\omega_j \boldsymbol{Z}_j' \Delta_j \Phi_j(\tau, \widetilde{\boldsymbol{\beta}}) g_i(Y_j^u)}{G^2(Y_j^u)} + o_p^{\tau}(1) \\ = \; & n^{-1/2}\sum_{i=1}^{n} \omega_i \boldsymbol{\eta}_{1i}(\tau, \widetilde{\boldsymbol{\beta}}) + o_p^{\tau}(1). \end{aligned} \tag{A.25}$$

where $\boldsymbol{\eta}_{1i}(\tau, \boldsymbol{b})$ and $\boldsymbol{\eta}_{2i}(\tau, \boldsymbol{b})$ were defined in Section A.2. Hence, we have

$$-\sqrt{n}\{\boldsymbol{S}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{S}_n^*(\tau, \widetilde{\boldsymbol{\beta}})\} = n^{-1/2}\sum_{i=1}^{n} \omega_i \boldsymbol{\eta}_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p^{\tau}(1).$$

Next, it can be shown that $\sqrt{n}\{\boldsymbol{S}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{S}_n^*(\tau, \widetilde{\boldsymbol{\beta}})\} = \sqrt{n}\{\boldsymbol{\mu}(\tau, \widehat{\boldsymbol{\beta}}^*) - \boldsymbol{\mu}(\tau, \widetilde{\boldsymbol{\beta}})\} + o_p^{\tau}(n^{-1/2})$, following similar arguments to those in Peng and Huang (2008), lemma B.1 and Huang (2010), Lemma 2. We then use Taylor expansion to get

$$\sqrt{n}\{\widehat{\boldsymbol{\beta}}^*(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)\} = -n^{-1/2}\sum_{i=1}^{n} \omega_i \boldsymbol{J}^{-1}(\tau, \widetilde{\boldsymbol{\beta}})\boldsymbol{\eta}_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p^{\tau}(1), \tag{A.26}$$

and furthermore

$$\sqrt{n}\{\widehat{\boldsymbol{\beta}}^*(\tau) - \widehat{\boldsymbol{\beta}}(\tau)\} = -n^{-1/2}\sum_{i=1}^{n} (\omega_i - 1)\boldsymbol{J}^{-1}(\tau, \widetilde{\boldsymbol{\beta}})\boldsymbol{\eta}_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p^{\tau}(1) \tag{A.27}$$

when combined with (A.13).

We now examine the distribution of $\sqrt{n}\{\hat{L}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) - L(\tau, \widetilde{\boldsymbol{\beta}})\}$, by decomposing it as

$$\sqrt{n}\{\hat{L}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \hat{L}_n^{*G}(\tau, \widehat{\boldsymbol{\beta}}^*)\} + \sqrt{n}\{\hat{L}_n^{*G}(\tau, \widehat{\boldsymbol{\beta}}^*) - L(\tau, \widehat{\boldsymbol{\beta}}^*)\} + \sqrt{n}\{L(\tau, \widehat{\boldsymbol{\beta}}^*) - L(\tau, \widetilde{\boldsymbol{\beta}})\}, \tag{A.28}$$

where $\hat{L}_n^{*G}(\tau, \boldsymbol{\beta}) = n^{-1}\sum_{i=1}^{n} \omega_i \dfrac{\Delta_i}{G(Y_i^u)} \rho_\tau(Y_i^u - \boldsymbol{Z}_i\boldsymbol{\beta})$. The third term is asymptotically negligible, because $\widetilde{\boldsymbol{\beta}}(\tau)$ is the minimizer of $L(\tau, \boldsymbol{b})$ and that $\widehat{\boldsymbol{\beta}}^*(\tau)$ is uniformly consistent to $\widetilde{\boldsymbol{\beta}}(\tau)$. The second term, by using the uniform consistency of $\widehat{\boldsymbol{\beta}}^*(\tau)$ to $\widetilde{\boldsymbol{\beta}}(\tau)$, equals $n^{-1/2}\sum_{i=1}^{n} \omega_i \pi_{2i}(\tau, \widetilde{\boldsymbol{\beta}}) + o_p^{\tau}(1)$. Finally, the first term equals $n^{-1/2}\sum_{i=1}^{n}\omega_i\pi_{1i}(\tau,\widetilde{\boldsymbol{\beta}})+o_p^\tau(1)$ by (A.23) and algebraic manipulations similar to those in (A.25). The mathematical forms of $\pi_{1i}(\cdot)$ and $\pi_{2i}(\cdot)$ are in (A.14). Thus we see that

$$\sqrt{n}\{\hat{L}_n^*(\tau,\widehat{\boldsymbol{\beta}}^*)-L(\tau,\widetilde{\boldsymbol{\beta}})\}=n^{-1/2}\sum_{i=1}^{n}\omega_i\pi_i(\tau,\widetilde{\boldsymbol{\beta}})+o_p^\tau(1),$$

which further implies that

$$\sqrt{n}\{\hat{L}_n^*(\tau,\widehat{\boldsymbol{\beta}}^*)-\hat{L}_n(\tau,\widehat{\boldsymbol{\beta}})\}=n^{-1/2}\sum_{i=1}^{n}(\omega_i-1)\pi_i(\tau,\widetilde{\boldsymbol{\beta}})+o_p^\tau(1).$$

Since $E(\omega_i)=Var(\omega_i)=1$, this result justifies the use of the perturbed $\sqrt{n}\{\hat{L}_n^*(\tau,\widehat{\boldsymbol{\beta}}^*)-\hat{L}_n(\tau,\widehat{\boldsymbol{\beta}})\}$ for approximating the asymptotic distribution of $\sqrt{n}\{\hat{L}_n(\tau,\widehat{\boldsymbol{\beta}})-L(\tau)\}$.

Supplementary Materials B: Justifications for the Resampling-based Hypothesis Testing Procedures

B.1 Proof of Theorem 3

Write $\boldsymbol{J}_n(\tau)=n^{-1}\sum_{i=1}^{n}[\boldsymbol{Z}_i^{\otimes 2}f\{\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}(\tau)|\boldsymbol{Z}_i\}]$, with $E\{\boldsymbol{J}_n(\tau)\}=\boldsymbol{J}(\tau)$. Similar to Chen et al. (2008) and Rao and Zhao (1992), we define $\boldsymbol{W}_{in}(\tau)=n^{-1/2}\boldsymbol{J}_n(\tau)^{-1/2}\boldsymbol{Z}_i'$. Under condition **C1**, we see that $\sup_i\boldsymbol{W}_{in}(\tau)=O_p^\tau(n^{-1/2})$. Define $\psi_\tau(u)=\tau-I(u\leqslant 0)$. Temporarily suppressing the τ-index in $\widetilde{\boldsymbol{\beta}}(\tau)$, $\psi_\tau(u)$ and $\boldsymbol{W}_{in}(\tau)$, we follow Lemma 2.2 in Rao and Zhao (1992) and define

$$\begin{aligned}
f_n(\tau,\boldsymbol{\gamma}) &= \sum_{i=1}^{n}\omega_i\frac{\Delta_i}{G^*(Y_i^u)}\{\rho_\tau(Y_i^u-\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}-\boldsymbol{W}_{in}'\gamma)-\rho_\tau(Y_i^u-\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})+\boldsymbol{W}_{in}'\gamma\psi(Y_i^u-\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\};\\
f_n^G(\tau,\boldsymbol{\gamma}) &= \sum_{i=1}^{n}\omega_i\frac{\Delta_i}{G(Y_i^u)}\{\rho_\tau(Y_i^u-\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}-\boldsymbol{W}_{in}'\gamma)-\rho_\tau(Y_i^u-\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})+\boldsymbol{W}_{in}'\gamma\psi(Y_i^u-\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\}.
\end{aligned}$$

Lemma 2. *$f_n(\tau,\boldsymbol{\gamma})=f_n^G(\tau,\boldsymbol{\gamma})+o_p(1)$ when $||\boldsymbol{\gamma}||\leqslant c$, where c is any given positive constant. Also, $\sup_{||\boldsymbol{\gamma}||\leqslant c}|f_n^G(\tau,\boldsymbol{\gamma})-\frac{1}{2}\boldsymbol{\gamma}^T\boldsymbol{\gamma}|=o_p(1)$.*

Proof. Since $\rho_\tau(x+t)-\rho_\tau(x) \geqslant \psi(x)\times t$ for any x and t, we have

$$\begin{aligned}
& |f_n(\tau,\boldsymbol{\gamma}) - f_n^G(\tau,\boldsymbol{\gamma})| \\
= & |\sum_{i=1}^n \omega_i \Delta_i \{\hat{G}^{*-1}(Y_i^u) - G^{-1}(Y_i^u)\}\{\rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} - \boldsymbol{W'}_{in}\boldsymbol{\gamma}) - \rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}) + \boldsymbol{W'}_{in}\boldsymbol{\gamma}\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\}| \\
\leqslant & \sup_{t\leqslant u}|\frac{G(t)-G^*(t)}{G^*(t)}| \times \sum_{i=1}^n \omega_i \frac{\Delta_i}{G(Y_i^u)}\{\rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} - \boldsymbol{W'}_{in}\boldsymbol{\gamma}) - \rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}) + \boldsymbol{W'}_{in}\boldsymbol{\gamma}\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\}. \\
= & \sup_{t\leqslant u}|\frac{G(t)-G^*(t)}{G^*(t)}| \times f_n^G(\tau,\boldsymbol{\gamma}) = o_p(1)\times f_n^G(\tau,\boldsymbol{\gamma}).
\end{aligned}$$

Also note that

$$\begin{aligned}
& \rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} - \boldsymbol{W'}_{in}\boldsymbol{\gamma}) - \rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}) + \boldsymbol{W'}_{in}\boldsymbol{\gamma}\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}) \\
= & \int_0^{-\boldsymbol{W'}_{in}\boldsymbol{\gamma}} \{\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} - s) - \psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\} ds.
\end{aligned}$$

Under condition **C3**, we follow Equation (2.5) in Rao and Zhao (1992) and get

$$\begin{aligned}
& E\big[f_n^G(\tau,\boldsymbol{\gamma})\big|\{\boldsymbol{Z}_i\}_{i=1}^n\big] \\
= & \sum_{i=1}^n E\big[\rho_\tau(T_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} - \boldsymbol{W'}_{in}\boldsymbol{\gamma}) - \rho_\tau(T_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}) + \boldsymbol{W'}_{in}\boldsymbol{\gamma}\psi(T_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\big|\{\boldsymbol{Z}_i\}_{i=1}^n\big] \\
= & \sum_{i=1}^n E\big[\int_0^{-\boldsymbol{W'}_{in}\boldsymbol{\gamma}} \{\psi(T_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} + u) - \psi(T_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\} du\big|\{\boldsymbol{Z}_i\}_{i=1}^n\big] \\
= & \sum_{i=1}^n \int_0^{-\boldsymbol{W'}_{in}\boldsymbol{\gamma}} f(\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}|\boldsymbol{Z}_i) u\{1+o(1)\} du \\
= & \sum_{i=1}^n \frac{1}{2}(\boldsymbol{W'}_{in}\boldsymbol{\gamma})^2 f(\boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}|\boldsymbol{Z}_i) + o(1) \to \frac{1}{2}\boldsymbol{\gamma'\gamma}, \qquad \text{(B.1)}
\end{aligned}$$

where the last step follows from the definition of $\boldsymbol{W}_{in}(\cdot)$ and $\boldsymbol{J}_n(\cdot)$. This result further implies that $E\{f_n^G(\tau,\boldsymbol{\gamma})\} \to \frac{1}{2}\boldsymbol{\gamma'\gamma}$. Similarly, we can show that $\mathrm{Var}\{f_n^G(\tau,\boldsymbol{\gamma})\} \to 0$. It then follows that $f_n^G(\tau,\boldsymbol{\gamma}) \xrightarrow{p} \frac{1}{2}\boldsymbol{\gamma'\gamma}$ for any given $\boldsymbol{\gamma}$ and τ. An application of Lemma 2.1 in Rao and Zhao (1992) completes the proof of Lemma 2. □

An immediate implication of Lemma 2 is that

$$\sum_{i=1}^n \omega_i \frac{\Delta_i}{G^*(Y_i^u)}\{\rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} - \boldsymbol{W'}_{in}\boldsymbol{\gamma}) - \rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}) + \boldsymbol{W'}_{in}\boldsymbol{\gamma}\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\} = \frac{1}{2}\boldsymbol{\gamma'\gamma} + o_p(1). \quad \text{(B.2)}$$

Using very similar techniques, we can also derive that

$$\sum_{i=1}^{n} \frac{\Delta_i}{\hat{G}(Y_i^u)} \{\rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}} - \boldsymbol{W}'_{in}\boldsymbol{\gamma}) - \rho_\tau(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}}) + \boldsymbol{W}'_{in}\boldsymbol{\gamma}\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\} = \frac{1}{2}\boldsymbol{\gamma}'\boldsymbol{\gamma} + o_p(1). \quad \text{(B.3)}$$

We consider $\boldsymbol{\gamma} = \sqrt{n}\boldsymbol{J}_n^{1/2}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})$. Note that $\boldsymbol{W}'_{in}\boldsymbol{\gamma} = n^{-1/2}\boldsymbol{Z}_i\boldsymbol{J}_n^{-1/2}\boldsymbol{\gamma} = \boldsymbol{Z}_i(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})$, we plug this $\boldsymbol{\gamma}$ into (B.3) and get

$$n\{\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}_n(\tau, \widetilde{\boldsymbol{\beta}})\} = -\{\sum_{i=1}^{n} \frac{\Delta_i}{\hat{G}(Y_i^u)} \boldsymbol{Z}_i\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) + \frac{n}{2}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}_n(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) + o_p(1), \quad \text{(B.4)}$$

where $-\{\sum_{i=1}^{n} \frac{\Delta_i}{\hat{G}(Y_i^u)} \boldsymbol{Z}'_i\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\} = n\boldsymbol{S}_n(\tau, \widetilde{\boldsymbol{\beta}})$. From the proof of Theorem 2, we know that

$$\sqrt{n}\boldsymbol{S}_n(\tau, \widetilde{\boldsymbol{\beta}}) = n^{-1/2}\sum_{i=1}^{n} \boldsymbol{\eta}_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1) = -\boldsymbol{J}(\tau)\sqrt{n}\{\widehat{\boldsymbol{\beta}}(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)\} + o_p(1).$$

Plugging this result into (B.4) yields

$$\begin{aligned} n\{\hat{L}_n(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}_n(\tau, \widetilde{\boldsymbol{\beta}})\} &= -n(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) + \frac{n}{2}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}_n(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) + o_p(1) \\ &= -\frac{n}{2}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) + o_p(1). \quad \text{(B.5)} \end{aligned}$$

where $\boldsymbol{J}_n$ converges in probability to $\boldsymbol{J}$.

Similarly, we plug $\boldsymbol{\gamma}_1 = \sqrt{n}\boldsymbol{J}_n^{1/2}(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}})$ and $\boldsymbol{\gamma}_2 = \sqrt{n}\boldsymbol{J}_n^{1/2}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})$ respectively into (B.2) and get

$$\begin{aligned} n\{\hat{L}_n^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \hat{L}_n^*(\tau, \widetilde{\boldsymbol{\beta}})\} &= -\{\sum_{i=1}^{n} \omega_i \frac{\Delta_i}{G^*(Y_i^u)} \boldsymbol{Z}_i\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\}\{\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}}\} + \frac{n}{2}(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}_n(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}}) + o_p(1); \\ n\{\hat{L}_n^*(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}_n^*(\tau, \widetilde{\boldsymbol{\beta}})\} &= -\{\sum_{i=1}^{n} \omega_i \frac{\Delta_i}{G^*(Y_i^u)} \boldsymbol{Z}_i\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\}\{\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}\} + \frac{n}{2}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}_n(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) + o_p(1). \\ &, \quad \text{(B.6)} \end{aligned}$$

where

$$-n^{-1/2}\{\sum_{i=1}^{n} \omega_i \frac{\Delta_i}{G^*(Y_i^u)} \boldsymbol{Z}'_i\psi(Y_i^u - \boldsymbol{Z}_i\widetilde{\boldsymbol{\beta}})\} = \sqrt{n}\boldsymbol{S}_n^*(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1) = -\sqrt{n}\boldsymbol{J}(\tau)\{\widehat{\boldsymbol{\beta}}^*(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)\} + o_p(1).$$

Combine this with (B.6), we can further derive

$$
\begin{aligned}
& n\{\hat{L}^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \hat{L}^*(\tau, \widehat{\boldsymbol{\beta}})\} \\
= \quad & n\{\hat{L}^*(\tau, \widehat{\boldsymbol{\beta}}^*) - \hat{L}^*(\tau, \widetilde{\boldsymbol{\beta}})\} - n\{\hat{L}^*(\tau, \widehat{\boldsymbol{\beta}}) - \hat{L}^*(\tau, \widetilde{\boldsymbol{\beta}})\} \\
= \quad & -n(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}}) + \frac{n}{2}(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}}) + n(\widehat{\boldsymbol{\beta}}^* - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) - \frac{n}{2}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}) + o_p(1) \\
= \quad & -\frac{n}{2}(\widehat{\boldsymbol{\beta}}^* - \widehat{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}}^* - \widehat{\boldsymbol{\beta}}) + o_p(1), \qquad \text{(B.7)}
\end{aligned}
$$

utilizing the uniform consistency of $\boldsymbol{J}_n(\tau)$ to $\boldsymbol{J}(\tau)$. This completes the proof of Theorem 3.

B.2 Justifications for the Hypothesis Testing Procedures in the Nested Case

By (A.11) and (A.13), we see that $\sqrt{n}\boldsymbol{S}_n(\tau, \widetilde{\boldsymbol{\beta}}) = n^{-1/2}\sum_{i=1}^n \boldsymbol{\eta}_i(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1)$, where $\boldsymbol{\eta}_i(\tau, \widetilde{\boldsymbol{\beta}})$ are i.i.d. with mean 0. Furthermore,

$$
n\{\widehat{\boldsymbol{\beta}}(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)\}'\boldsymbol{J}(\tau)\{\widehat{\boldsymbol{\beta}}(\tau) - \widetilde{\boldsymbol{\beta}}(\tau)\} = n\boldsymbol{S}_n(\tau, \widetilde{\boldsymbol{\beta}})'\boldsymbol{J}(\tau)^{-1}\boldsymbol{S}_n(\tau, \widetilde{\boldsymbol{\beta}}) + o_p(1). \qquad \text{(B.8)}
$$

Now, suppose that $\boldsymbol{Z}_B = (\boldsymbol{Z}_A, \dot{\boldsymbol{Z}})$ without loss of generality, where $\dot{\boldsymbol{Z}}$ is of length $r_B - r_A$. Temporarily suppressing the τ index, an application of Theorem 3 under H_0 gives

$$
\begin{aligned}
n\mathcal{T}_{AB}(\tau) \quad &= \quad -\frac{n}{2}[(\widehat{\boldsymbol{\beta}}_A - \widetilde{\boldsymbol{\beta}}_A)'\boldsymbol{J}_A(\widehat{\boldsymbol{\beta}}_A - \widetilde{\boldsymbol{\beta}}_A) - (\widehat{\boldsymbol{\beta}}_B - \widetilde{\boldsymbol{\beta}}_B)'\boldsymbol{J}_B(\widehat{\boldsymbol{\beta}}_B - \widetilde{\boldsymbol{\beta}}_B)] + o_p(1) \qquad \text{(B.9)} \\
n\mathcal{T}_{AB}^{n*}(\tau) \quad &= \quad -\frac{n}{2}[(\widehat{\boldsymbol{\beta}}_A^* - \widehat{\boldsymbol{\beta}}_A)'\boldsymbol{J}_A(\widehat{\boldsymbol{\beta}}_A^* - \widehat{\boldsymbol{\beta}}_A) - (\widehat{\boldsymbol{\beta}}_B^* - \widehat{\boldsymbol{\beta}}_B)'\boldsymbol{J}_B(\widehat{\boldsymbol{\beta}}_B^* - \widehat{\boldsymbol{\beta}}_B)] + o_p(1). \qquad \text{(B.10)}
\end{aligned}
$$

Let $\boldsymbol{K}$ be a $r_B \times r_A$ matrix that equals the first r_A^{th} column of an $r_B \times r_B$ identity matrix. Under H_0, we have $\boldsymbol{Z}_A = \boldsymbol{Z}_B\boldsymbol{K}$, $\boldsymbol{S}_{nA}(\widetilde{\boldsymbol{\beta}}_A) = \boldsymbol{K}'\boldsymbol{S}_{nB}(\widetilde{\boldsymbol{\beta}}_B)$ and $\boldsymbol{J}_A = \boldsymbol{K}'\boldsymbol{J}_B\boldsymbol{K}$. Combining these with (B.8) and (B.9), we obtain

$$
\begin{aligned}
& n\mathcal{T}_{AB} = -\frac{n}{2}\boldsymbol{S}_{nA}(\tau, \widetilde{\boldsymbol{\beta}}_A)'\boldsymbol{J}_A^{-1}\boldsymbol{S}_{nA}(\tau, \widetilde{\boldsymbol{\beta}}_A) + \frac{n}{2}\boldsymbol{S}_{nB}(\tau, \widetilde{\boldsymbol{\beta}}_B)'\boldsymbol{J}_B^{-1}\boldsymbol{S}_{nB}(\tau, \widetilde{\boldsymbol{\beta}}_B) + o_p(1) \\
= \quad & \frac{1}{2}\{n^{-1/2}\sum_{i=1}^n \boldsymbol{\eta}_{iB}(\widetilde{\boldsymbol{\beta}}_B)\}'\{\boldsymbol{J}_B^{-1} - \boldsymbol{K}(\boldsymbol{K}'\boldsymbol{J}_B\boldsymbol{K})^{-1}\boldsymbol{K}'\}\{n^{-1/2}\sum_{i=1}^n \boldsymbol{\eta}_{iB}(\widetilde{\boldsymbol{\beta}}_B)\} + o_p(1).
\end{aligned}
$$

Similarly, we can derive that

$$
n(\widehat{\boldsymbol{\beta}}^* - \widehat{\boldsymbol{\beta}})'\boldsymbol{J}(\widehat{\boldsymbol{\beta}}^* - \widehat{\boldsymbol{\beta}}) = \{n^{-1/2}\sum_{i=1}^n(\omega_i - 1)\boldsymbol{\eta}_i(\widetilde{\boldsymbol{\beta}})\}'\boldsymbol{J}^{-1}\{n^{-1/2}\sum_{i=1}^n(\omega_i - 1)\boldsymbol{\eta}_i(\widetilde{\boldsymbol{\beta}})\} + o_p(1),
$$

and $\boldsymbol{\eta}_{iA}(\widetilde{\boldsymbol{\beta}}_A) = \boldsymbol{K}'\boldsymbol{\eta}_{iB}(\widetilde{\boldsymbol{\beta}}_B)$ under H_0, which further gives

$$n\mathcal{T}_{AB}^{n*} = \frac{1}{2}\{n^{-1/2}\sum_{i=1}^{n}(\omega_i - 1)\boldsymbol{\eta}_{iB}(\widetilde{\boldsymbol{\beta}}_B)\}'\{\boldsymbol{J}_B^{-1} - \boldsymbol{K}(\boldsymbol{K}'\boldsymbol{J}_B\boldsymbol{K})^{-1}\boldsymbol{K}'\}\{n^{-1/2}\sum_{i=1}^{n}(\omega_i - 1)\boldsymbol{\eta}_{iB}(\widetilde{\boldsymbol{\beta}}_B)\} + o_p(1).$$

Following the arguments of Chen et al. (2008), proof of theorem 3, we see that the conditional distribution of $n\mathcal{T}_{AB}^{n*}$ given the observed data is equivalent to the unconditional distribution of $n\mathcal{T}_{AB}$, which then justifies the validity of the perturbation scheme under the nested case.

Supplementary Materials C: Additional Numerical Results

The numerical performance of $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$ and $\hat{R}_{CV}^1(\tau, \widehat{\boldsymbol{\beta}})$ under Models A and B were summarized in Table C.1. To obtain these results, we implemented the random split multiple times for the same dataset and then calculate the average of the resulting $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$ and $\hat{R}_{CV}^1(\tau, \widehat{\boldsymbol{\beta}})$. This improves the numerical performance slightly, with no change to the asymptotic behavior. We observe that the estimators are also asymptotically unbiased. The empirical standard errors agrees with the estimated standard errors based on influence functions. It is clear that the performance of the CV-type estimators also improves with the sample size. Overall, the performances of CV-type estimators are quite similar to those of the plug-in estimators.

Table C.1: Simulations: summary statistics for $\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$ at τ = 0.1, 0.3, 0.5, 0.6 under model A and model B.

n		**Model A**					**Model B**				
	τ	**TRUE**	**EB** $\times 10^3$	**ESE** $\times 10^3$	**ASE** $\times 10^3$	**C95** (%)	TRUE	**EB** $\times 10^3$	**ESE** $\times 10^3$	**ASE** $\times 10^3$	**C95** (%)
					$\hat{L}_{CV}(\tau, \widehat{\boldsymbol{\beta}})$						
200	0.1	0.117	0	12	11	91.4	0.129	0	12	11	91.5
	0.3	0.231	0	22	20	92.2	0.255	0	23	21	91.7
	0.5	0.263	0	27	24	90.9	0.291	1	28	24	91.2
	0.6	0.253	1	26	23	90.8	0.281	0	27	24	90.4
400	0.1	0.117	0	10	8	92.7	0.129	0	9	8	92.6
	0.3	0.231	0	16	15	93.1	0.255	0	16	15	92.9
	0.5	0.263	0	18	17	92.3	0.291	1	19	17	92.7
	0.6	0.253	1	18	17	92.4	0.281	0	19	17	92.3
600	0.1	0.117	0	7	6	93.8	0.129	0	7	7	94.0
	0.3	0.231	0	12	12	94.3	0.255	1	12	12	94.5
	0.5	0.263	0	15	14	93.9	0.291	1	15	14	93.6
	0.6	0.253	1	14	14	93.5	0.281	1	15	14	93.3
					$\hat{R}^1_{CV}(\tau, \widehat{\boldsymbol{\beta}})$						
200	0.1	0.478	−3	54	51	93.5	0.425	−3	55	52	93.8
	0.3	0.473	−3	50	47	93.0	0.417	−2	52	48	93.0
	0.5	0.472	−3	53	49	92.8	0.415	−3	55	50	92.4
	0.6	0.473	−3	56	51	92.4	0.416	−4	58	52	92.1
400	0.1	0.478	0	37	36	94.1	0.425	−1	39	37	93.9
	0.3	0.473	0	34	33	94.0	0.417	0	35	34	93.9
	0.5	0.472	0	37	35	93.5	0.415	0	37	35	93.8
	0.6	0.473	0	38	36	93.4	0.416	−1	39	37	93.6
600	0.1	0.478	0	29	29	94.1	0.425	−1	30	30	94.3
	0.3	0.473	0	27	27	94.5	0.417	0	28	28	94.4
	0.5	0.472	0	29	28	94.3	0.415	0	30	29	93.8
	0.6	0.473	0	30	30	94.2	0.416	−1	31	30	93.8

Table C.2 represents sensitivity analysis results for $\hat{R}_n^1(\tau, \widehat{\boldsymbol{\beta}})$, where the distribution of C is misspecified and $n = 400$. The performance of the proposed method deteriorates slightly but is still acceptable.

Table C.2: Simulations: summary statistics for $\hat{R}_n^1(\tau, \widehat{\boldsymbol{\beta}})$ at $\tau = 0.1, 0.3, 0.5, 0.6$ under model A and model B, where the censoring distribution depends on covariates but is estimated using the Kaplan-Meier method.

n		**Model A**					**Model B**				
	τ	**TRUE**	**EB** $\times 10^3$	**ESE** $\times 10^3$	**ASE** $\times 10^3$	**C95** (%)	**TRUE**	**EB** $\times 10^3$	**ESE** $\times 10^3$	**ASE** $\times 10^3$	**C95** (%)
400	0.1	0.478	−6	38	37	94.4	0.425	−5	39	38	94.0
	0.3	0.473	−7	36	35	93.5	0.417	−5	37	35	92.7
	0.5	0.472	−5	39	37	92.7	0.415	−3	41	37	92.4
	0.6	0.473	−5	42	39	92.4	0.416	−3	44	40	92.0

For two *non-nested* models, we can build a Wald-type confidence interval for the difference between their corresponding $R^1(\tau)$'s, using the estimated $\hat{R}_n^1(\tau, \widehat{\boldsymbol{\beta}})$'s and the influence functions in (A.16). We examined the performance of the confidence intervals using simulations. In Table C.3 below, the rows labeled "Diff" present the difference in $R^1(\tau)$ between models A vs. B and models B vs. C, respectively. The rows labeled "C95" display the corresponding empirical coverage rates using 95% Wald-type confidence intervals. We observe that the empirical coverage rates are all very close to the nominal level.

Table C.3: Simulation results: empirical coverage rates of 95% confidence intervals (C95) for the difference in $R^1(\tau)$ between two non-nested working models, where $n = 400$ and $E(\sum_{i=1}^n \delta_i) = 288$.

		$\boldsymbol{\tau}$ **0.1**	**0.3**	**0.5**	**0.6**
A vs. B	Diff	0.052	0.057	0.057	0.056
	C95	0.955	0.951	0.951	0.950
B vs. C	Diff	0.062	−0.036	−0.057	−0.052
	C95	0.945	0.948	0.952	0.952

Table C.4: Simulation results: empirical rejection rates (ERR) based on $\mathcal{R}_{AB}$, where $[\tau_L, \tau_U] = [0.1, 0.6]$.

	n=200		n=400		n=600	
	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$
(i) A vs. B	0.823	0.890	0.979	0.992	0.998	0.999
(ii) A vs. C	0.976	0.994	1.000	1.000	1.000	1.000
(iii) E vs. A	**0.061**	**0.119**	**0.062**	**0.120**	**0.060**	**0.113**
(iv) B vs. C	0.206	0.302	0.345	0.462	0.493	0.610
(v) B vs. D	0.855	0.923	0.999	1.000	1.000	1.000

Table C.4 summarizes the performance of the overall hypothesis testing procedure, where the censoring distribution is correctly specified. Table C.5 presents sensitivity results for hypothesis testing when the censoring distribution is mis-specified by the Kaplan-Meier method, where $n = 400$. The results in Table C.5 are quite comparable to those in Table 2 in the manuscript, suggesting that the proposed hypothesis testing procedure is insensitive to moderate mis-specification of the censoring distribution.

Table C.5: Simulation results when the distribution of C is mis-specified: empirical rejection rates (ERR) based on $\mathcal{T}_{AB}(\tau)$ and perturbations, where the bolded cells are empirical sizes, and the remaining cells correspond to empirical power. The significance level $\alpha = 0.05$ and $E(\sum_{i=1}^{n} \delta_i) = 288$.

	τ				$\tau \in [0.1, 0.6)$
	0.1	**0.3**	**0.5**	**0.6**	
A vs. B	0.850	0.936	0.915	0.877	0.958
A vs. C	1.000	0.821	**0.058**	0.315	1.000

Finally, Figure C.1 below presents the estimated regression coefficients for the dialysis dataset.

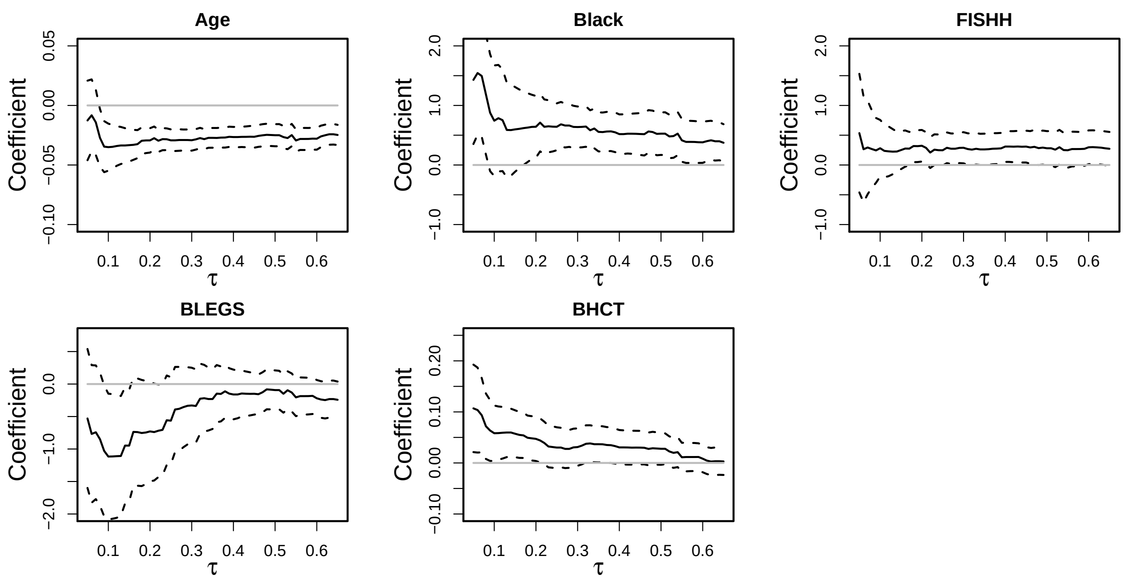

Figure C.1: Analysis of the dialysis data: estimated regression coefficients and pointwise 95% confidence intervals under Model A.

References

Andersen, P. K. (1993). *Statistical models based on counting processes.* Springer Verlag.

Chen, K., Ying, Z., Zhang, H., and Zhao, L. (2008). Analysis of least absolute deviation. *Biometrika* **95,** 107–122.

Huang, Y. (2010). Quantile calculus and censored regression. *The Annals of Statistics* **38,** 1607–1637.

Peng, L. and Fine, J. (2009). Competing risks quantile regression. *Journal of the American Statistical Association* **104,** 1440–1453.

Peng, L. and Huang, Y. (2008). Survival analysis with quantile regression models. *Journal of the American Statistical Association* **103,** 637–649.

Pepe, M. S. (1991). Inference for events with dependent risks in multiple endpoint studies. *Journal of the American Statistical Association* **86,** 770–778.

Rao, C. R. and Zhao, L. (1992). Approximation to the distribution of m-estimates in linear models by randomly weighted bootstrap. *Sankhyā: The Indian Journal of Statistics, Series A* **54,** 323–331.

Tian, L., Cai, T., Goetghebeur, E., and Wei, L. (2007). Model evaluation based on the sampling distribution of estimated absolute prediction error. *Biometrika* **94,** 297–311.

Van Der Vaart, A. and Wellner, J. (2000). *Weak convergence and empirical processes.* Springer Verlag, New York.

Zhou, L. (2006). A simple censored median regression estimator. *Statistica Sinica* **16,** 1043–1058.