Supplementary Materials for "Inference for High Dimensional Censored Quantile Regression"

July 14, 2021

Appendix A: Comparisons with Ad Hoc Approaches

We perform additional simulations to compare our method with various ad hoc approaches under the same settings as in Example 2. The ad hoc approaches include: the "OneSplit" approach which conducts variable selection and CQR estimation based on a single data split; the "NoSplit" approach which conducts both variable selection and CQR estimation on the full data; the "Half" approach which conducts Fused-HDCQR but performs selection and estimation on the same half of data. The results, as reported in WebTable 1, show that our proposed "Fused" method has the smallest biases and standard errors (SEs); "OneSplit" has smaller biases than "NoSplit" but with larger SEs; "NoSplit" has larger biases than "Fused" and "OneSplit," but smaller biases than 'Half"; and the "Half" approach produces the largest biases under all of the configurations.

Appendix B: Proofs of the Main Results

We prove our main results. All of the lemmas mentioned in the proofs are to be formally stated and proved in Appendix C.

Under Condition (A1), we have $E\left[\int_0^{\tau_0} 1\left\{\log X_i \geq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(u)\right\} dH(u)\right] = \tau_0$ (Zheng et al., 2018). Therefore, following the discussions underneath Assumption 3.1 in Zheng et al. (2018),

$$\sum_{i=1}^n \mathbf{Z}_i \left(\sum_{r=0}^{k-1} \int_{\tau_r}^{\tau_{r+1}} 1\left\{\log X_i \geq \mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)\right\} dH(u) + \int_0^{\tau_0} 1\left\{\log X_i \geq \mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(0)\right\} dH(u)\right),$$

the grid approximation of $\sum_{i=1}^n \mathbf{Z}_i \int_0^{\tau} 1\left\{\log X_i \geq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(u)\right\} dH(u)$, can be shown to be asymptotically equivalent to

$$\sum_{i=1}^n \mathbf{Z}_i \left(\sum_{r=0}^{k-1} \int_{\tau_r}^{\tau_{r+1}} 1\left\{\log X_i \geq \mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)\right\} dH(u) + \tau_0\right).$$

Let $\psi_{\tau_0}(\mathbf{h}) := 2\sum_{i=1}^n \Delta_i \rho_{\tau_0}(\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i) + 2\tau_0 \sum_{i=1}^n (1-\Delta_i)(\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i)$. Denote by

$$\begin{aligned}\acute{L}_k(\mathbf{h}) = &\sum_{i=1}^n \Delta_i \left|\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i\right| + \mathbf{h}^{\mathrm{T}} \sum_{i=1}^n \Delta_i \mathbf{Z}_i \\ &-2\mathbf{h}^{\mathrm{T}} \sum_{i=1}^n \mathbf{Z}_i \left(\sum_{r=0}^{k-1} \int_{\tau_r}^{\tau_{r+1}} 1\left\{\log X_i \geq \mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)\right\} dH(u) + \tau_0\right). \qquad (1)\end{aligned}$$

In particular, $\acute{L}_0(\mathbf{h}) = \sum_{i=1}^n \Delta_i \left|\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i\right| + \mathbf{h}^{\mathrm{T}} \sum_{i=1}^n \Delta_i \mathbf{Z}_i - 2\mathbf{h}^{\mathrm{T}} \sum_{i=1}^n \mathbf{Z}_i \tau_0$.

Given $\boldsymbol{\delta} \in \mathbf{R}^p$, let $S(\boldsymbol{\delta}) = \{j : \delta_j \neq 0\}$. Given $S \subset \{1, \ldots, p\}$ such that $S^* \subseteq S$ and $|S| \leq K_1 n^{c_1}$, we denote by $\acute{\Omega}_{S,k}(a,b)$, $k = 0, \ldots, m$ the event that for all $0 \leq r \leq k$,

$$\inf_{\|(E[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}])^{1/2}\boldsymbol{\delta}\| = \sqrt{\lambda_{\min}} a\nu_{r,n}(b), S(\boldsymbol{\delta})\subseteq S} \acute{L}_r(\boldsymbol{\beta}^*(\tau_r) + \boldsymbol{\delta}) - \acute{L}_r(\boldsymbol{\beta}^*(\tau_r)) > 0,$$

where $\{\nu_{k,n}(b), k = 0, \ldots, m\}$ is a sequence satisfying $\nu_{0,n}(b) = \nu_{0,n} = \sqrt{K_1 n^{c_1-1}\log n}$ and $\nu_{k+1,n}(b) = \nu_{k,n}(1+b\epsilon_n)$ for some constant $b > 0$; $\nu_{k,n}$ increases with k and $\nu_{m,n} = \nu_{0,n}(1+b\epsilon_n)^m \leq \nu_{0,n}(1+bc_0 n^{-1})^{n/c_0} \leq e^b(K_1 n^{c_1-1}\log n)^{1/2}$. Event $\acute{\Omega}_{S,k}(a,b)$ and the convexity of $\acute{L}_r(\mathbf{h})$ together ensure the uniform consistency of $\acute{\boldsymbol{\beta}}(\tau)$ for all $\nu \leq \tau \leq \tau_k$.

With Propositions 1 and 2 below, we show that $\acute{\Omega}_{S,k}(a,b)$ holds with probability going to 1. In Proposition 1, we prove that there exists a constant ζ_1, such that $\mathrm{P}(\acute{\Omega}_{S,0}(\zeta_1, 0)) > 1 - 16K_1 n^{c_1-4}$. Thus, $\acute{\boldsymbol{\beta}}(\tau_0)$ is consistent with rate $\left(K_1 n^{c_1-1}\log n\right)^{1/2}$, which establishes the baseline result for induction. Then in Proposition 2, we show that there exists a constant ζ_2, such that given that event $\acute{\Omega}_{S,k-1}(\zeta_1,\zeta_2)$ holds, event $\acute{\Omega}_{S,k}(\zeta_1,\zeta_2)$ holds with probability at least $1 - 4(5k+8)K_1 n^{c_1-4}$.

Propositions 1 and 2 will lead to the estimation consistency of $\acute{\boldsymbol{\beta}}(\tau)$ over $[\tau_0, \tau_U]$. Specifically, $\nu_{m,n}(\zeta_2) \leq \exp(\zeta_2)(K_1 n^{c_1-1}\log n)^{1/2}$ and $\mathrm{P}(\acute{\Omega}_{S,m}(\zeta_1,\zeta_2)) \geq 1 - \sum_{k=0}^m 4(5k+8)K_1 n^{c_1-4} \geq 1 - 20c_0^{-2}K_1 n^{c_1-2}$ imply the estimation consistency of $\acute{\boldsymbol{\beta}}(\tau_k)$ at $k = 1, \ldots, m$, and consequently the uniform consistency of $\acute{\beta}_j(\tau)$, $j = 1, .., p$ over $[\tau_0, \tau_U]$, as shown in Theorem 1. Then utilizing some empirical process techniques, we can establish the weak convergence of $\acute{\beta}_j$ for any $j \in S$ in Theorem 2.

Proposition 1. *Under Conditions (A1) – (A7), one can find a sufficiently large constant* ζ_1 *such that event* $\acute{\Omega}_{S,0}(\zeta_1, 0)$ *holds with probability at least* $1 - 16K_1 n^{c_1-4}$, *where* $S \subset \{1, \ldots, p\}$ *such that* $S^* \subseteq S$ *and* $|S| \leq K_1 n^{c_1}$.

Proof of Proposition 1. Consider $\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))$. It can be written as

$$E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))\right] + \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau)) - E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))\right].$$

By Lemma 2, uniformly for $\boldsymbol{\delta}$ that satisfies $\|\boldsymbol{\delta}\|_0 \leq K_1 n^{c_1}$, $\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2, t \leq \kappa/\sqrt{\lambda_{\min}}$,

$$n^{-1}E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta})\right] - n^{-1}E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))\right] \geq \underline{g}t^2 - 2At^3/(3c_2). \tag{2}$$

It follows that

$$\sup_{\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2, S(\boldsymbol{\delta})\subseteq S} \left|\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0)) - E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))\right]\right| \leq 2\sqrt{n}\mathcal{A}_0(t),$$

where

$$\mathcal{A}_0(t) := \sup_{\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} \leq t^2, S(\boldsymbol{\delta})\subseteq S} \left|\mathbb{G}_n\left[\bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0))\right]\right|,$$

and $\bar{\rho}_{\tau_0,i}(\mathbf{h}) := \Delta_i \rho_{\tau_0}(\log X_i - \mathbf{Z}_i^{\mathrm{T}}\mathbf{h}) + \tau_0(1 - \Delta_i)(\log X_i - \mathbf{Z}_i^{\mathrm{T}}\mathbf{h})$. According to Lemma 3,

$$\mathrm{P}\left(\mathcal{A}_0(t) \geq 24\sqrt{2}\left(\lambda_{\min}^{-1}K_1 n^{c_1}\log n\right)^{1/2} t\right) \leq 16K_1 n^{c_1-4}. \tag{3}$$

By (2) and (3), we obtain that

$$\inf_{E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2, S(\boldsymbol{\delta})\subseteq S} n^{-1}\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta}) - \psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))\right] \geq t\left\{\underline{g}t - \frac{2}{3c_2}At^2 - 48\sqrt{2}\left(\lambda_{\min}^{-1}K_1 n^{c_1-1}\log n\right)^{1/2}\right\}$$

with probability at least $1-16K_1n^{c_1-4}$. Therefore, there exists a sufficiently large constant ζ_1, such that

$$\inf_{\boldsymbol{\delta}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta}=\lambda_{\min}\zeta_1^2K_1n^{c_1-1}\log n,S(\boldsymbol{\delta})\subseteq S}\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0)+\boldsymbol{\delta})-\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))>0,$$

with probability at least $1-16K_1n^{c_1-4}$.

Since $\psi_{\tau_0}(\mathbf{h})$ is convex with respect to $\mathbf{h}$, we have with probability at least $1-16K_1n^{c_1-4}$,

$$\left\|(E[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}])^{1/2}\left(\acute{\boldsymbol{\beta}}(\tau_0)-\boldsymbol{\beta}^*(\tau_0)\right)\right\|\le\sqrt{\lambda_{\min}}\zeta_1\left(K_1n^{c_1-1}\log n\right)^{1/2}.$$

By Condition (A6), $\left\|\acute{\boldsymbol{\beta}}(\tau_0)-\boldsymbol{\beta}^*(\tau_0)\right\|\le\zeta_1\left(K_1n^{c_1-1}\log n\right)^{1/2}$. □

Proposition 2. *Suppose Conditions (A1)–(A7) hold and ζ_1 is a constant from Proposition 1, there exists a universal constant ζ_2 such that under event $\acute{\Omega}_{S,k-1}(\zeta_1,\zeta_2)$, $1\le k\le m$, event $\acute{\Omega}_k(\zeta_1,\zeta_2)$ holds with probability at least $1-4(5k+8)K_1n^{c_1-4}$, where $S\subset\{1,\dots,p\}$ such that $S^*\subseteq S$ and $|S|\le K_1n^{c_1}$.*

Proof of Proposition 2. We note that

$$\begin{aligned}
&\acute{L}_k(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta})-\acute{L}_k(\boldsymbol{\beta}^*(\tau_k))\\
&=E\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta})-\acute{L}_k(\boldsymbol{\beta}^*(\tau_k))\right]\\
&\quad+n^{1/2}\mathbb{G}_n\left[\Delta_i\left(\left|\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}\right)\right|-\left|\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)\right|\right)+\sum_{i=1}^n\Delta_i\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}\right]\\
&\quad-2n^{1/2}\mathbb{G}_n\left[\boldsymbol{\delta}^{\mathrm{T}}\sum_{i=1}^n\mathbf{Z}_i\left(\sum_{r=0}^{k-1}\int_{\tau_r}^{\tau_{r+1}}1\left\{\log X_i\ge\mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)\right\}dH(u)+\tau_0\right)\right].
\end{aligned}$$

We choose some constant ζ_2 such that

$$\zeta_2>\frac{22\bar{f}\lambda_{\max}^{1/2}}{\underline{g}\lambda_{\min}^{1/2}(1-\tau_U)}+\frac{128\sqrt{2}}{\underline{g}\lambda_{\min}(1-\tau_U)\zeta_1}+\frac{2L\epsilon_nn^{1/2-c_1}}{\underline{g}\lambda_{\min}^{1/2}(1-\tau_U)K_1\log n},$$

where $\underline{g}$ is defined in Condition (A3) and $\lambda_{\min}$ are defined in Condition (A6). It can be seen that the choice of ζ_2 does not depend on n as the last three terms go to zero as n increases. Then we show under event $\acute{\Omega}_{k-1}(\zeta_1,\zeta_2)$, event $\acute{\Omega}_k(\zeta_1,\zeta_2)$ holds with large probability.

We follow the similar arguments used in Proposition 1. By Lemmas 5, 6 and 7, we have under $\acute{\Omega}_{S,k-1}(\zeta_1,\zeta_2)$,

$$\begin{aligned}
&\inf_{\boldsymbol{\delta}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta}=t^2,\|\boldsymbol{\delta}\|_0\le K_1n^{c_1}}n^{-1}\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta})-\acute{L}_k(\boldsymbol{\beta}^*(\tau_k))\right]\\
&\ge t\Bigg\{\underline{g}t-\frac{2}{3c_2}At^2-2\frac{\epsilon_n}{1-\tau_U}\sum_{r=0}^{k-1}\left(\bar{f}\zeta_1\lambda_{\max}^{1/2}\nu_{r,n}(\zeta_2)+L\epsilon_n\right)\\
&\quad-(40+64\tau_0)\sqrt{2}\left(\lambda_{\min}^{-1}K_1n^{c_1-1}\log n\right)^{1/2}-128\sqrt{2}\sum_{r=0}^{k-1}\frac{\epsilon_n}{1-\tau_U}\left(\lambda_{\min}^{-1}K_1n^{c_1-1}\log n\right)^{1/2}\\
&\quad-20\frac{\epsilon_n}{1-\tau_U}\sum_{r=0}^{k-1}\bar{f}\zeta_1\lambda_{\max}^{1/2}\nu_{r,n}(\zeta_2)\Bigg\}
\end{aligned}$$

with probability at least $1-4(5k+8)K_1n^{c_1-4}$. Thus, under $\acute{\Omega}_{S,k-1}(\zeta_1,\zeta_2)$. we have

$$\begin{aligned}
0 \le\ & \underline{g}\zeta_1\sqrt{\lambda_{\min}}\nu_{k-1,n}(\zeta_2) - \frac{2}{3c_2}A\lambda_{\min}(\zeta_1\nu_{k-1,n}(\zeta_2))^2 \\
& -\frac{\epsilon_n}{1-\tau_U}\sum_{r=0}^{k-2}\left(22\bar{f}\zeta_1\sqrt{\lambda_{\min}}\nu_{r,n}(\zeta_2)+2L\epsilon_n\right) \\
& -(40+64\tau_0)\sqrt{2}\left(\lambda_{\min}^{-1}K_1n^{c_1-1}\log n\right)^{1/2} - 128\sqrt{2}\sum_{r=0}^{k-2}\frac{\epsilon_n}{1-\tau_U}\left(\lambda_{\min}^{-1}K_1n^{c_1-1}\log n\right)^{1/2}
\end{aligned} \tag{4}$$

Let $\nu_{k,n}(\zeta_2)=(1+\zeta_2\epsilon)\nu_{k-1,n}(\zeta_2)$. Simple algebra yields that

$$\begin{aligned}
& \underline{g}\zeta_1\sqrt{\lambda_{\min}}\nu_{k-1,n}(\zeta_2) - \frac{2}{3c_2}A\lambda_{\min}(\zeta_1(1+\zeta_2\epsilon_n)\nu_{k-1,n}(\zeta_2))^2 \\
& -\frac{\epsilon_n}{1-\tau_U}\sum_{r=0}^{k-2}\left(22\bar{f}\zeta_1\lambda_{\max}^{1/2}\nu_{r,n}(\zeta_2)+2L\epsilon_n\right) \\
& -(40+64\tau_0)\sqrt{2}\left(\lambda_{\min}^{-1}K_1n^{c_1-1}\log n\right)^{1/2} - 128\sqrt{2}\sum_{r=0}^{k-2}\frac{\epsilon_n}{1-\tau_U}\left(\lambda_{\min}^{-1}K_1n^{c_1-1}\log n\right)^{1/2} \\
& +\underline{g}\zeta_1\zeta_2\sqrt{\lambda_{\min}}\epsilon_n\nu_{k-1,n}(\zeta_2) - \frac{\epsilon_n}{1-\tau_U}\left(22\bar{f}\zeta_1\lambda_{\max}^{1/2}\nu_{k-1,n}(\zeta_2)+2L\epsilon_n\right) \\
& -128\sqrt{2}\frac{\epsilon_n}{1-\tau_U}\left(\lambda_{\min}^{-1}K_1n^{c_1-1}\log n\right)^{1/2} > 0,
\end{aligned}$$

by our choice of ζ_2. Again, since (1) is convex with respect to $\mathbf{h}$, under $\acute{\Omega}_{k-1}(\zeta_1,\zeta_2)$, we have with probability at least $1-8(2k+3)K_1n^{c_1-4}$,

$$\left\|(E[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}])^{1/2}\left(\acute{\boldsymbol{\beta}}(\tau_k)-\boldsymbol{\beta}^*(\tau_k)\right)\right\| \le \sqrt{\lambda_{\min}}\zeta_1\nu_{k,n}(\zeta_2).$$

By condition (A6), we have $\left\|\acute{\boldsymbol{\beta}}(\tau_k)-\boldsymbol{\beta}^*(\tau_k)\right\| \le \zeta_1\nu_{k,n}(\zeta_2)$. □

Proof of Theorem 1. By Propositions 1 and 2, we have

$$\begin{aligned}
\mathrm{P}(\acute{\Omega}_{S,m}(\zeta_1,\zeta_2)) &\ge 1-\sum_{k=0}^{m}4(5k+8)K_1n^{c_1-4} \ge 1-\sum_{k=0}^{(c_0n^{-1})^{-1}}4(5k+8)K_1n^{c_1-4} \\
&\ge 1-10c_0^{-2}K_1n^{c_1-2}-42c_0^{-1}K_1n^{c_1-3}-32K_1n^{c_1-4} \ge 1-20c_0^{-2}K_1n^{c_1-2},
\end{aligned}$$

when n is sufficiently large. Thus, we have with probability at least $1-20c_0^{-2}K_1n^{c_1-2}$,

$$\begin{aligned}
& \sup_{\tau_0\le\tau\le\tau_U}\|\acute{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}^*(\tau)\| \\
\le\ & \max_{k=0,\ldots,m-1}\left\{\sup_{\tau_k\le\tau<\tau_{k+1}}\|\acute{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}^*(\tau)\|, \|\acute{\boldsymbol{\beta}}(\tau_m)-\boldsymbol{\beta}^*(\tau_m)\|\right\} \\
\le\ & \max_{k=0,\ldots,m-1}\|\acute{\boldsymbol{\beta}}(\tau_k)-\boldsymbol{\beta}^*(\tau_k)\| + \sup_{\tau_k\le\tau<\tau_{k+1}}\|\boldsymbol{\beta}^*(\tau)-\boldsymbol{\beta}^*(\tau_k)\|, \zeta_1\nu_{m_n,n}(\zeta_2) \\
\le\ & \max\left\{\max_{k=0,\ldots,m-1}\zeta_1\nu_{k,n}(\zeta_2)+Lc_0n^{-1}(K_1n^{c_1})^{1/2}, \zeta_1\nu_{m_n,n}(\zeta_2)\right\} \\
\le\ & \zeta_1\nu_{m_n,n}(\zeta_2) \le \zeta_1e^{\zeta_2}(K_1n^{c_1-1}\log n)^{1/2}.
\end{aligned} \tag{5}$$

Thus $\acute{\boldsymbol{\beta}}(\tau)$ is uniformly consistent to $\boldsymbol{\beta}^*(\tau)$ with the convergence rate $(n^{c_1-1}\log n)^{1/2}$ across $\tau\in[\tau_0,\tau_U]$. □

Proof of Theorem 2. For a set $S \subset \{1,\dots,p\}$ satisfying $S^* \subseteq S$ and $|S| \le K_1 n^{c_1}$, $0 \le c_1 < 1/3$ and $K_1 \le 1$, let $\acute{\boldsymbol{\beta}}_S(\tau)$ be the estimator from fitting CQR $Q_Y(\tau|Z_S) = \mathbf{Z}_S^{\mathrm{T}}\boldsymbol{\beta}_S(\tau)$, and $\forall j \in S$, the j-th entry $\acute{\beta}_j(\tau)$ is the coefficient for variable Z_j. Further denote $\theta_{iS}(\tau) = \mathbf{Z}_{iS}^{\mathrm{T}}\boldsymbol{\beta}_S(\tau)$ and $\theta_{iS}^*(\tau) = \mathbf{Z}_{iS}^{\mathrm{T}}\boldsymbol{\beta}_S^*(\tau)$ for subject i. Then $\acute{\boldsymbol{\beta}}_S(\tau)$ is the solution to the following estimating equation as in Peng and Huang (2008),

$$n^{1/2}\mathbf{U}_n(\boldsymbol{\beta}_S,\tau) = 0,$$

where

$$\mathbf{U}_n(\boldsymbol{\beta}_S,\tau) = n^{-1}\sum_{i=1}^{n}\mathbf{Z}_{iS}\left(N_i\left(\theta_{iS}(\tau)\right) - \int_0^\tau I[\log X_i \ge \theta_{iS}(u)]dH(u)\right).$$

Let $\mathbf{u}(\boldsymbol{\beta}_S,\tau) = E[\mathbf{U}_n(\boldsymbol{\beta}_S,\tau)]$. By the Martingale property, $\mathbf{u}(\boldsymbol{\beta}_S^*,\tau) = 0$, $\tau \in (0,1)$. For a vector $\mathbf{b} = (b_1, b_2, .., b_{|S|})^{\mathrm{T}}$ of length $|S|$, we define $\boldsymbol{\mu}_S(\mathbf{b}) = E[\mathbf{Z}_S N(\mathbf{Z}_S^{\mathrm{T}}\mathbf{b})] = E[\mathbf{Z}_S G(\mathbf{Z}_S^{\mathrm{T}}\mathbf{b}|\mathbf{Z}_S)]$, $\mathbf{B}_S(\mathbf{b}) = E[\mathbf{Z}_S\mathbf{Z}_S^{\mathrm{T}} g\left(\mathbf{Z}_S^{\mathrm{T}}\mathbf{b}|\mathbf{Z}_S\right)]$, and $\mathbf{J}_S(\mathbf{b}) = -E[\mathbf{Z}_S\mathbf{Z}_S^{\mathrm{T}} f\left(\mathbf{Z}_S^{\mathrm{T}}\mathbf{b}|\mathbf{Z}_S\right)]$. For $d > 0$, define

$\mathcal{B}(d) = \{\mathbf{b} \in \mathbf{R}^{|S|} : \inf_{\tau\in(0,\tau_U]} \|\boldsymbol{\mu}_S(\mathbf{b}) - \boldsymbol{\mu}_S[\boldsymbol{\beta}_S^*(\tau)]\| \le d\}$, and $\mathcal{A}(d) = \{\boldsymbol{\mu}_S(\mathbf{b}) : \mathbf{b} \in \mathcal{B}(d)\}$.

By the restricted eigenvalue condition (A6), together with $|S| \le K_1 n^{c_1}$, $\boldsymbol{\mu}_S$ is a one-to-one map from $\mathcal{B}(d_0)$ to $\mathcal{A}(d_0)$ for some $d_0 > 0$. By Conditions (A3) and (A6), the inverse of $\mathbf{B}_S(\boldsymbol{\beta}_S^*(\tau))$ exists and we use $\mathbf{B}_S^{-1}(\boldsymbol{\beta}_S^*(\tau))$ to denote the inverse. Furthermore, let $\mathbf{e}_j = (1\{i = j\})_{i=1,..,|S|}$, be the unit vector of which the jth element is 1. Then $\acute{\beta}_j(\tau) = \mathbf{e}_j^{\mathrm{T}}\acute{\boldsymbol{\beta}}_S(\tau)$ and $\mathbf{e}_j^{\mathrm{T}} = [\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\mathbf{B}_S(\boldsymbol{\beta}_S^*(\tau))$, where $\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))$ is the jth column of $\mathbf{B}_S^{-1}(\boldsymbol{\beta}_S^*(\tau))$. By the Taylor Expansion, $\boldsymbol{\mu}_S\{\acute{\boldsymbol{\beta}}_S(\tau)\} - \boldsymbol{\mu}_S\{\boldsymbol{\beta}_S^*(\tau)\} = \mathbf{B}_S(\boldsymbol{\beta}_S^*(\tau))\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big) + \big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)/2$ for some $\mathbf{b}_\tau \in \mathcal{B}(d_0)$ between $\boldsymbol{\beta}_S^*(\tau)$ and $\acute{\boldsymbol{\beta}}_S(\tau)$, where $\nabla^2\boldsymbol{\mu}_S(\mathbf{b})$ is the second derivative of $\boldsymbol{\mu}_S$. The jth element of $\nabla^2\boldsymbol{\mu}_S(\mathbf{b})$ is $E\left[g'(\mathbf{Z}_S^{\mathrm{T}}\mathbf{b}|\mathbf{Z}_S)Z_{Sj}\mathbf{Z}_S\mathbf{Z}_S^{\mathrm{T}}\right]$, where Z_{Sj} is the jth element of $\mathbf{Z}_S$.

$$\begin{aligned}
&\acute{\beta}_j(\tau) - \beta_j^*(\tau) = [\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\mathbf{B}_S(\boldsymbol{\beta}_S^*(\tau))\left(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\right)\\
= \quad &[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\left(\boldsymbol{\mu}_S\{\acute{\boldsymbol{\beta}}_S(\tau)\} - \boldsymbol{\mu}_S\{\boldsymbol{\beta}_S^*(\tau)\}\right)\\
&-[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)/2. \qquad (6)
\end{aligned}$$

We first consider $[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)/2$.

$$\begin{aligned}
&|[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)|\\
&= |\mathbf{e}_j^{\mathrm{T}}\mathbf{B}_S^{-1}(\boldsymbol{\beta}_S^*(\tau))\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)|\\
&\le \|\mathbf{B}_S^{-1}(\boldsymbol{\beta}_S^*(\tau))\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)\|_2\\
&\le \underline{g}^{-1}\lambda_{\min}^{-1}\left\|\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)\right\|_2\\
&\le \underline{g}^{-1}\lambda_{\min}^{-1}\left(K_1 n^{c_1}\right)^{1/2}\max_{1\le j\le|S|} E\left[g'(\mathbf{Z}_S^{\mathrm{T}}\mathbf{b}|\mathbf{Z}_S)Z_{Sj}\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\mathbf{Z}_S\mathbf{Z}_S^{\mathrm{T}}\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)\right]\\
&\le \underline{g}^{-1}\lambda_{\min}^{-1}A\left(K_1 n^{c_1}\right)^{1/2} E\left[\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\mathbf{Z}_S\mathbf{Z}_S^{\mathrm{T}}\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)\right]\\
&= \underline{g}^{-1}\lambda_{\min}^{-1}\lambda_{\max}AK_1^{3/2}O_p(n^{3c_1/2-1}\log n) = O_p(n^{3c_1/2-1}\log n),
\end{aligned}$$

where the first inequality follows from $\|\mathbf{e}_j\|_2 = 1$, the second inequality follows from the definition of $\mathbf{B}_S(\boldsymbol{\beta}_S^*(\tau))$ and Conditions (A3) and (A6), the third inequality is trivial, the fourth inequality follows from Conditions (A2) and (A3), and the last equality follows from Condition (A6) and Theorem 1. Then

$$\begin{aligned}
&n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)^{\mathrm{T}}\nabla^2\boldsymbol{\mu}_S(\mathbf{b}_\tau)\big(\acute{\boldsymbol{\beta}}_S(\tau) - \boldsymbol{\beta}_S^*(\tau)\big)/2\\
&= O_p(n^{3c_1/2-1/2}\log n) = o_p(1), \qquad (7)
\end{aligned}$$

by Condition (A4).

We next consider $[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\left(\boldsymbol{\mu}_S\{\acute{\boldsymbol{\beta}}_S(\tau)\} - \boldsymbol{\mu}_S\{\boldsymbol{\beta}_S^*(\tau)\}\right)$. We modify the decomposition in Appendix C of Peng and Huang (2008) by multiplying both sides by $[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}$,

$$
\begin{aligned}
-n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\mathbf{U}_n(\boldsymbol{\beta}_S^*,\tau) =& n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}[\boldsymbol{\mu}_S\{\acute{\boldsymbol{\beta}}_S(\tau)\} - \boldsymbol{\mu}_S\{\boldsymbol{\beta}_S^*(\tau)\}] \\
& - \int_0^\tau [\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\left[\mathbf{J}_S\left(\boldsymbol{\beta}_S^*(u)\right)\mathbf{B}_S\left(\boldsymbol{\beta}_S^*(u)\right)^{-1} + o_{(0,\tau_U](1)}\right] \\
& \times n^{1/2}\left[\boldsymbol{\mu}_S\{\acute{\boldsymbol{\beta}}_S(\tau)\} - \boldsymbol{\mu}_S\{\boldsymbol{\beta}_S^*(\tau)\}\right]dH(u) + o_{(0,\tau_U]}(1).
\end{aligned}
$$

View the equation as a stochastic differential equation for $n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\left[\boldsymbol{\mu}_S\{\acute{\boldsymbol{\beta}}_S(\tau)\} - \boldsymbol{\mu}_S\{\boldsymbol{\beta}_S^*(\tau)\}\right]$. We use the production integration theory (Andersen et al. (2012) II.6) and obtain

$$
\begin{aligned}
& n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\left[\boldsymbol{\mu}_S\{\acute{\boldsymbol{\beta}}_S(\tau)\} - \boldsymbol{\mu}_S\{\boldsymbol{\beta}_S^*(\tau)\}\right] \\
=& \boldsymbol{\phi}_j\left[-n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\mathbf{U}_n(\boldsymbol{\beta}_S^*,\tau)\right] + o_{(0,\tau_U]}(1),
\end{aligned} \tag{8}
$$

where $\boldsymbol{\phi}_j$ is a map from $\mathcal{G}$ to $\mathcal{G}$ such that for $\mathbf{g} \in \mathcal{G}$, $\boldsymbol{\phi}_j(\mathbf{g})(\tau) = \int_0^\tau \mathcal{I}_j(s,\tau)d\mathbf{g}(s)$, with

$$
\mathcal{I}_j(s,t) = \prod_{u\in[s,t]}\left[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(t))\right]^{\mathrm{T}}\left[\mathbf{I}_{|S|} + \mathbf{J}_S\{\boldsymbol{\beta}_S^*(u)\}\mathbf{B}_S\{\boldsymbol{\beta}_S^*(u)\}^{-1}dH(u)\right] \quad \text{and}
$$

$$
\mathcal{G} = \{\mathbf{g} : [0,\tau_U] \to \mathbf{R}, \mathbf{g} \text{ is left-continuous with right limit, } \mathbf{g}(0) = 0\},
$$

where $\mathbf{I}_l$ is a $l \times l$ identity matrix.

Next we show the convergence of $-n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\mathbf{U}_n(\boldsymbol{\beta}_S^*,\tau)$. Since $\mathbf{U}_n$ is of dimension $|S|$, which increases with n, we apply the results in Section 2.11.3 of Van Der Vaart and Wellner (2000). We write the class

$$
\mathcal{F}_n = \left\{f_{n,\tau} = [\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\mathbf{Z}_{iS}\left(N_i\left(\theta_{iS}^*(\tau)\right) - \int_0^\tau 1\left\{\log X_i \geq \theta_{iS}^*(u)\right\}dH(u)\right) : \tau \in [\nu,\tau_U]\right\}.
$$

Since $N_i\left(\theta_{iS}^*(\tau)\right) - \int_0^\tau 1\left\{\log X_i \geq \theta_{iS}^*(u)\right\}dH(u)$ is uniformly bounded by some K_3 over $\tau \in [\nu,\tau_U]$, we choose $F_n = \sup_{\tau\in[\nu,\tau_U]} K_3\underline{g}^{-1}\lambda_{\min}^{-1}\|\mathbf{Z}_S\|$. One can check that

$$
\begin{aligned}
& \mathrm{P}^*F_n^2 = O(1), \\
& \mathrm{P}^*F_n^2\{F_n > \eta\sqrt{n}\} \to 0, \quad \forall \eta > 0, \text{ and} \\
& \sup_{|\tau-\tilde{\tau}|<\delta_n} \mathrm{P}\left(f_{n,\tau} - f_{n,\tilde{\tau}}\right)^2 \to 0, \quad \forall \delta_n \downarrow 0,
\end{aligned}
$$

where P^* is the outer probability. By Conditions (A3) and (A6), $f_{n,\tau}$ is Lipschitz. By Lemma 2.7.11 of Van Der Vaart and Wellner (2000), $N_{[]}(\epsilon\|F_n\|_{P,2}, \mathcal{F}_n, L_2(P)) \leq N(\epsilon/2, [0,1], L_1) \leq 2/\epsilon$. We refer to Page 83 in Van Der Vaart and Wellner (2000) for the definitions of the bracketing number $N_{[]}()$ and covering number $N()$. Let $u = \log(2/\epsilon)$. Then as $\delta_n \to 0$,

$$
\int_0^{\delta_n}(\log(2/\epsilon))^{1/2}d\epsilon = \int_{\log(2/\delta_n)}^\infty 2u^{1/2}e^{-u}du \to 0.
$$

By Theorem 2.11.23 of Van Der Vaart and Wellner (2000), $-n^{1/2}[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(\tau))]^{\mathrm{T}}\mathbf{U}_n(\boldsymbol{\beta}_S^*,\tau)$ is tight in $\tau \in [\nu,\tau_U]$, and converges in distribution to a tight Gaussian process $\mathbf{G}_S(\tau)$ with mean zero and covariance $\boldsymbol{\Sigma}(s,t)$, where $\boldsymbol{\Sigma}(s,t) = E\{[\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(s))]^{\mathrm{T}}\imath_{iS}(s)\imath_{iS}(t)^{\mathrm{T}}\mathbf{B}_{jS}^{-1}(\boldsymbol{\beta}_S^*(t))\}$, and

$$
\imath_{iS}(\tau) = \mathbf{Z}_{iS}\left(N_i\left(\theta_{iS}^*(\tau)\right) - \int_0^\tau 1\left\{\log X_i \geq \theta_{iS}^*(\tau)\right\}du\right).
$$

Last, because $\boldsymbol{\phi}_j$ is a linear operator, $\boldsymbol{\phi}_j\{\mathbf{G}_S(\tau)\}, \tau \in [\nu,\tau_U]$ is Gaussian as well (Römisch, 2014). This coupled with (6) and (7), yields that $\sqrt{n}\left(\acute{\beta}_j(\tau) - \beta_j^*(\tau)\right), \tau \in [\nu,\tau_U]$ converges weakly to a mean zero Gaussian process denoted as $\boldsymbol{\phi}_j\{\mathbf{G}_S(\tau)\}$.

□

Now we are equipped to prove Theorem 3.

Proof of Theorem 3. We first introduce the oracle estimators of $\beta_j^*(\tau)$'s assuming the true active set S^* is known. For each $j \in \{1, \dots, p\}$, once again $S_{+j}^* = \{j\} \cup S^*$, and note that $S_{+j}^* = S^*$ if $j \in S^*$. Let $\check{\boldsymbol{\beta}}_{S_{+j}^*}(\tau)$ be the oracle estimator by fitting the following CQR on the full data,

$$Q_Y(\tau|\mathbf{Z}_{S_{+j}^*}) = \mathbf{Z}_{S_{+j}^*}^{\mathrm{T}} \boldsymbol{\beta}_{S_{+j}^*}(\tau).$$

Then the oracle estimator for $\beta_j^*(\tau)$ is $\check{\beta}_j(\tau) = \left(\check{\boldsymbol{\beta}}_{S_{+j}^*}(\tau)\right)_j$, the entry corresponding to the coefficient for variable Z_j. Analogically, let $\check{\beta}_j^b(\tau)$ denote the oracle estimator fitted on the b-th sub-sample D_2^b in the Fused-HDCQR procedure.

The objective can be decomposed as below,

$$\begin{aligned}
&\sqrt{n}\left(\widehat{\beta}_j(\tau) - \beta_j^*(\tau)\right) \\
=&\sqrt{n}\left(\check{\beta}_j(\tau) - \beta_j^*(\tau)\right) + \sqrt{n}\left(\widehat{\beta}_j(\tau) - \check{\beta}_j(\tau)\right) \\
=&\sqrt{n}\left(\check{\beta}_j(\tau) - \beta_j^*(\tau)\right) + \sqrt{n}\left(\frac{1}{B}\sum_{b=1}^{B}\widetilde{\beta}_j^b(\tau) - \check{\beta}_j(\tau)\right) \\
=&\underbrace{\sqrt{n}\left(\check{\beta}_j(\tau) - \beta_j^*(\tau)\right)}_{\mathrm{I}} + \underbrace{\sqrt{n}\left(\frac{1}{B}\sum_{b=1}^{B}\check{\beta}_j^b(\tau) - \check{\beta}_j(\tau)\right)}_{\mathrm{II}} + \underbrace{\sqrt{n}\left(\frac{1}{B}\sum_{b=1}^{B}\left\{\widetilde{\beta}_j^b(\tau) - \check{\beta}_j^b(\tau)\right\}\right)}_{\mathrm{III}}.
\end{aligned}$$

We will study the asymptotic behavior of the three terms separately. As the first two terms do not involve the selections $\widehat{S}^b$'s, they deal with the oracle estimators and the true active set S^*.

- $\mathrm{I} = \sqrt{n}\left(\check{\beta}_j(\tau) - \beta_j^*(\tau)\right)$ converges weakly to a mean zero Gaussian process;
- $\mathrm{II} = \sqrt{n}\left(\frac{1}{B}\sum_{b=1}^{B}\check{\beta}_j^b(\tau) - \check{\beta}_j(\tau)\right) = o_p(1)$, uniformly in $\tau \in [\nu, \tau_U]$;
- $\mathrm{III} = \sqrt{n}\left(\frac{1}{B}\sum_{b=1}^{B}\left\{\widetilde{\beta}_j^b(\tau) - \check{\beta}_j^b(\tau)\right\}\right) = o_p(1)$, uniformly in $\tau \in [\nu, \tau_U]$.

By Slutsky's theorem for random processes (Theorem 18.10 in Van der Vaart (2000)), if the above statements all hold, we would conclude that $\sqrt{n}\left(\widehat{\beta}_j(\tau) - \beta_j^*(\tau)\right), \tau \in [\nu, \tau_U]$ converges weakly to a mean zero Gaussian process.

a) Let $S = S_{+j}^*$ for each $j \in \{1, \dots, p\}$, and by Theorem 2, $\mathrm{I} = \sqrt{n}\left(\check{\beta}_j(\tau) - \beta_j^*(\tau)\right), \tau \in [\nu, \tau_U]$ converges weakly to a mean zero Gaussian process $\boldsymbol{\phi}_j\{\mathbf{G}_S(\tau)\}$, where $\boldsymbol{\phi}_j(\cdot), \mathbf{G}_S(\cdot)$ are defined in the proof of Theorem 2. Denote its covariance as $\sigma_j^*(s,t)$, which is uniformly bounded for $s, t \in [\nu, \tau_U]$.

b) To show $\mathrm{II} = o_p(1)$, we first denote $\xi_{b,n}(\tau) = \sqrt{n}\left(\check{\beta}_j^b(\tau) - \check{\beta}_j(\tau)\right)$, then $\mathrm{II} = \left(\sum_{b=1}^{B}\xi_{b,n}(\tau)\right)/B$. Since D_2^b's are random sub-samples, $\xi_{b,n}(\tau)$'s are i.i.d. conditional on data. Using a similar argument as in Appendix C of Peng and Huang (2008), the conditional distribution of $\sqrt{n}\left(\check{\beta}_j^b(\tau) - \check{\beta}_j(\tau)\right)$ given the observed data is asymptotically the same as the unconditional distribution of $\mathrm{I} = \sqrt{n}\left(\check{\beta}_j(\tau) - \beta_j^*(\tau)\right)$, which is mean zero Gaussian from part a). Thus $E(\xi_{b,n}(\tau)|D^{(n)}) \to E(\mathrm{I}) \to 0$ and $\mathrm{Var}\,(\xi_{b,n}(\tau)|D^{(n)}) \to \sigma_j^*(\tau,\tau) \doteq \sigma_j^2(\tau)$, as $n \to \infty$. Denote $\sigma_j^2 = \sup_{\tau\in[\nu,\tau_U]} \sigma_j^2(\tau) < \infty$, then $E(\mathrm{II}|D^{(n)}) \to 0$ uniform in $\tau \in [\nu, \tau_U]$, and

$$\mathrm{Var}\,(\mathrm{II}|D^{(n)}) = \frac{1}{B^2}\sum_{b=1}^{B} \mathrm{Var}\,(\xi_{b,n}(\tau)|D^{(n)}) \le \frac{2\sigma_j^2(\tau)}{B} \le \frac{2\sigma_j^2}{B},\ \tau \in [\nu, \tau_U].$$

Now for any $\delta, \zeta > 0$, there exist $N_0, B_0 > 0$ such that, for any $\tau \in [\nu, \tau_U]$, when$n > N_0, B > B_0$,

$$\begin{aligned}\mathrm{P}(|\mathrm{II}| \ge \delta) &\le \int_{D^{(n)}\in\Omega_n} \mathrm{P}\left(|\mathrm{II}| \ge \delta \,\middle|\, D^{(n)}\right) \mathrm{dP}(D^{(n)}) \\ &\le \int_{\Omega_n} \mathrm{P}\left(|\mathrm{II} - E(\mathrm{II})| \ge \delta/2 \,\middle|\, D^{(n)}\right) \mathrm{dP}(D^{(n)}) \\ &\le \int_{\Omega_n} \frac{\mathrm{Var}\,\left(\mathrm{II}\,\middle|\, D^{(n)}\right)}{\delta^2/4} \mathrm{dP}(D^{(n)}) \le \frac{2\sigma_j^2}{B_0\delta^2/4}\int_{\Omega_n} \mathrm{dP}(D^{(n)}) \le \zeta.\end{aligned}$$

Thus, $\mathrm{II} = o_p(1)$ uniformly in $\tau \in [\nu, \tau_U]$.

c) Each subsample D_1^b can be regarded as a random sample of $\lceil n/2 \rceil$ i.i.d. observations from the population distribution for which assumption (A4) holds, that is $|\widehat{S}^b| \le K_1 n^{c_1}$ and $\mathrm{P}\left(S^* \subset \widehat{S}^b\right) \ge 1 - K_2(p \vee n)^{-1}$. Notice that whenever $S^* \subset \widehat{S}^b$, the estimators based on the respective selections both converge to the truth by Theorem 2, i.e. $\sqrt{n}\left(\widetilde{\beta}_j^b(\tau) - \check{\beta}_j^b(\tau)\right) \to 0, \tau \in [\nu, \tau_U]$. Define $\eta_b(\tau) = 1\left\{S^* \not\subset \widehat{S}^b\right\}\sqrt{n}\left\{\widetilde{\beta}_j^b(\tau) - \check{\beta}_j^b(\tau)\right\}$, while omitting subscripts j in η for simplicity, then $\mathrm{III} = \left(\sum_{b=1}^{B} \eta_b(\tau)\right)/B$.

By Lemma 1, there exists $M_0 > 0$ such that $\sup_{\tau\in[\nu,\tau_U]}\left|\widetilde{\beta}_j^b(\tau) - \check{\beta}_j^b(\tau)\right| \le 2M_0$ for any $\widehat{S}^b$ with $|\widehat{S}^b| \le K_1 n^{c_1}$. Therefore, by (A4),

$$E(\eta_b(\tau)) \le \mathrm{P}\left(S^* \not\subset \widehat{S}^b\right)\sqrt{n} \sup_{b\in[B],\tau\in[\nu,\tau_U]}\left|\widetilde{\beta}_j^b(\tau) - \check{\beta}_j^b(\tau)\right| \le 2M_0\sqrt{n}K_2(p\vee n)^{-1-c_2} \to 0;$$

$$\mathrm{Var}\,(\eta_b(\tau)) \le \mathrm{P}\left(S^* \not\subset \widehat{S}^b\right) n \sup_{b\in[B],\tau\in[\nu,\tau_U]}\left(\widetilde{\beta}_j^b(\tau) - \check{\beta}_j^b(\tau)\right)^2 \le 4M_0^2 n K_2(p\vee n)^{-1-c_2} \to 0.$$

Although $\eta_b(\tau)$'s are dependent, we further have

$$E(\mathrm{III}) = E\left\{\left(\sum_{b=1}^{B}\eta_b(\tau)\right)/B\right\} \le 2M_0\sqrt{n}K_2(p\vee n)^{-1-c_2} \to 0;$$

$$\mathrm{Var}\,(\mathrm{III}) = \frac{1}{B^2}\sum_{b=1}^{B}\sum_{b'=1}^{B}\mathrm{Cov}\,(\eta_b(\tau), \eta_{b'}(\tau)) \le 4M_0^2 n K_2(p\vee n)^{-1-c_2} \to 0.$$

Thus $\mathrm{III} = o_p(1)$ uniformly in $\tau \in [\nu, \tau_U]$ by definition, as $\forall\delta, \zeta > 0$, $\exists N_0 > 0$ such that $\forall\tau \in [\nu, \tau_U], n > N_0$,

$$\mathrm{P}(|\mathrm{III}| \ge \delta) \le \mathrm{P}\left(|\mathrm{III} - E(\mathrm{III})| \ge \delta/2\right) \le \frac{\mathrm{Var}\,(\mathrm{III})}{\delta^2/4} \le \frac{16M_0^2K_2}{\delta^2}(p\vee n)^{-c_2} \le \zeta.$$

□

Appendix C: Lemmas and Proofs

We present the lemmas used in the proofs of the theorems and propositions and their proofs.

Lemma 1. ***(Bounds of coefficients)*** *Under assumptions (A1) – (A3), (A5) – (A7), for any* $S \subset \{1,\dots,p\}$ *with* $|S| \le K_1 n^{c_1}$, $0 \le c_1 < 1/3$ *and* $K_1 \le 1$, *there exists a constant* $M_0 > 0$, *such that* $\sup_{j\in S,\tau\in[\nu,\tau_U]} |\acute{\beta}_j(\tau)| < M_0$ *almost surely.*

Proof of Lemma 1. From Peng and Huang (2008), $\acute{\boldsymbol{\beta}}_S(\tau)$ is sequentially estimated for $\tau_k \in \Gamma_m$, $k = 0, 1, \ldots, m$ by solving the following minimization problem of an L_1-type convex objective function for $\mathbf{h}$ at k,

$$\begin{aligned} n^{-1}\acute{L}_k(\mathbf{h}) =& n^{-1}\sum_{i=1}^{n}\Delta_i\left|\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i\right| + n^{-1}\mathbf{h}^{\mathrm{T}}\sum_{i=1}^{n}(-\Delta_i\mathbf{Z}_i) \\ & - 2n^{-1}\mathbf{h}^{\mathrm{T}}\sum_{i=1}^{n}\mathbf{Z}_i\left(\sum_{r=0}^{k-1}\int_{\tau_r}^{\tau_{r+1}} I[\log X_i \geq \mathbf{Z}_i^{\mathrm{T}}\acute{\beta}_j(\tau_r)]dH(u) + \tau_0\right), \end{aligned}$$

and $n^{-1}\acute{L}_0(\mathbf{h}) = n^{-1}\sum_{i=1}^{n}\Delta_i\left|\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i\right| + n^{-1}\mathbf{h}^{\mathrm{T}}\sum_{i=1}^{n}\Delta_i\mathbf{Z}_i - 2n^{-1}\mathbf{h}^{\mathrm{T}}\sum_{i=1}^{n}\mathbf{Z}_i\tau_0$.

Since $\acute{\boldsymbol{\beta}}_S(\tau)$ is defined as a right-continuous function on the grid Γ_m, to show the boundedness of $\acute{\beta}_j(\tau)$'s, we only need to show it at the grid points τ_k's. We first prove $n^{-1}\acute{L}_k(\mathbf{h})$ is a coercive function in $\mathbf{h}$, that is $n^{-1}\acute{L}_k(\mathbf{h}) \to \infty$ whenever $\|\mathbf{h}\| \to \infty$.

Since $n^{-1}\acute{L}_k(\mathbf{h}) \geq n^{-1}\sum_{i=1}^{n}\Delta_i\left|\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i\right|$, where the right hand of the inequality does not depend on τ or k, it is sufficient to show $L(\mathbf{h}) = n^{-1}\sum_{i=1}^{n}\Delta_i\left|\log X_i - \mathbf{h}^{\mathrm{T}}\mathbf{Z}_i\right|$ is coercive. By Proposition 12.3.1 in Lange (2004), a sufficient and necessary condition is that $L(\mathbf{h})$ is coercive along all nontrivial rays $\{\mathbf{h} : \mathbf{h} = t\mathbf{v}, t \geq 0\}$. The condition is met because $\forall \mathbf{v} \in \mathbf{R}^{|S|}$, $L(t\mathbf{v}) = n^{-1}\sum_{i=1}^{n}\Delta_i\left|\log X_i - t\mathbf{v}^{\mathrm{T}}\mathbf{Z}_i\right|$ is an absolute value function in t, and thus goes to infinity as $t \to \infty$. Now let $L_0 = L_k(\mathbf{0})$, which does not depend on k and is bounded almost surely by Condition (A2), then the set $\{\mathbf{h} : n^{-1}\acute{L}_k(\mathbf{h}) \leq L_0\}$ is compact and contains the minimizer $\acute{\boldsymbol{\beta}}_S(\tau_k)$. Thus there exists a uniform bound $M_0 > 0$ depending on L_0, such that $\sup_{j\in S,\tau\in[\nu,\tau_U]}|\acute{\beta}_j(\tau)| < M_0$. □

Lemma 2. *Under Conditions (A1), (A3), and (A6), given $0 < t \leq \kappa/\sqrt{\lambda_{\min}}$, we have*

$$n^{-1}E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta})\right] - n^{-1}E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))\right] \geq \underline{g}t^2 - 2At^3/(3c_2),$$

uniformly for $\boldsymbol{\delta}$ that satisfies $\|\boldsymbol{\delta}\|_0 \leq K_1 n^{c_1}$ and $\boldsymbol{\delta}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2$.

Proof of Lemma 2. By Condition (A1), we have

$$\begin{aligned} E\left[\tau_0 - \Delta_i 1\left\{\log X_i \leq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0)\right\}|\mathbf{Z}_i\right] &= \tau_0 - E\left[1\left\{T_i \leq C_i\right\}1\left\{\log X_i \leq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0)\right\}\right] \\ = \tau_0 - E\left[1\left\{\log T_i \leq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0)\right\}\right] &+ E\left[1\left\{T_i > C_i\right\}1\left\{\log T_i \leq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0)\right\}\right] = 0. \end{aligned} \quad (9)$$

Given any $\boldsymbol{\delta} \in A_{\tau_0}$ satisfying $\boldsymbol{\delta}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2$, by the identity from Knight (1998) that for any $u \neq 0$,

$$|u - v| - |u| = -v\left[1 - 2\cdot 1\{u < 0\}\right] + 2\int_0^v \left[1\{u \leq t\} - 1\{u \leq 0\}\right]dt,$$

we have

$$\begin{aligned} & n^{-1}E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0) + \boldsymbol{\delta})\right] - n^{-1}E\left[\psi_{\tau_0}(\boldsymbol{\beta}^*(\tau_0))\right] \\ =& -2E[\Delta_i\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}\left(\tau_0 - 1\{\log X_i \leq \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0)\}\right) + \tau_0(1 - \Delta_i)\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}] \\ & + 2E\left[\Delta_i\int_0^{\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}} 1\left\{\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0) \leq u\right\} - 1\left\{\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0) \leq 0\right\}du\right] \\ =& 2E\left[\int_0^{\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}} \Delta_i\left(1\left\{\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0) \leq u\right\} - 1\left\{\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0) \leq 0\right\}\right)du\right] \\ \geq& 2E\left[\int_0^{|\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}|}\left[g\left(\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0)\middle|\mathbf{Z}_i\right)u - Au^2\right]du\right] \geq \underline{g}t^2 - 2At^3/(3c_2), \end{aligned}$$

where the second equation follows from (9), the first inequality follows from the law of iterated expectations, mean value expansion, and Condition (A3), and the second inequality follows from Condition (A6). The above inequality holds uniformly for $\boldsymbol{\delta}$ that satisfies $\|\boldsymbol{\delta}\|_0 \leq K_1 n^{c_1}$ and $\boldsymbol{\delta}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2$. □

Lemma 3. *Let* $\bar{\rho}_{\tau_0,i}(\mathbf{h}) := \Delta_i \rho_{\tau_0}(\log X_i - \mathbf{Z}_i^{\mathrm{T}}\mathbf{h}) + \tau_0(1-\Delta_i)(\log X_i - \mathbf{Z}_i^{\mathrm{T}}\mathbf{h})$ *and*

$$\mathcal{A}_0(t) := \sup_{\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} \le t^2, \|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}} \left|\mathbb{G}_n\left[\bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0)+\boldsymbol{\delta}) - \bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0))\right]\right|,$$

under Conditions (A1)-(A7), we have for any $C_1 > t$,

$$\mathrm{P}\left(\mathcal{A}_0(t) \ge 12C_1\right) \le 16K_1 n^{c_1} \exp\left(-\frac{C_1^2}{2K_1 n^{c_1} t^2/\lambda_{\min}}\right).$$

Proof of Lemma 3. For any $\boldsymbol{\delta}$ such that $\|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}$ and $\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} \le t^2$,

$\mathrm{Var}\left(\mathbb{G}_n\left[\bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0)+\boldsymbol{\delta}) - \bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0))\right]\right) \le E[(\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta})^2] \le t^2$. Applying Lemma 2.3.7 from Van Der Vaart and Wellner (2000) yields that, for each $M > 2t$,

$$Pr\left(\mathcal{A}_0(t) \ge M\right) \le \frac{2Pr\left(\mathcal{A}_0^0(t) \ge M/4\right)}{1 - 4t^2/M^2},$$

where

$$\mathcal{A}_0^0(t) := \sup_{\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} \le t^2, \|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}} \left|\mathbb{G}_n\left[V_i\left\{\bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0)+\boldsymbol{\delta}) - \bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0))\right\}\right]\right|$$

is a symmetrized version of $\mathcal{A}_0(t)$, and V_i's are Rademacher random variables. Since $\bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0)+\boldsymbol{\delta}) - \bar{\rho}_{\tau_0,i}(\boldsymbol{\beta}^*(\tau_0)) = -\tau_0 \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta} + D_i(\tau_0,\boldsymbol{\delta})$, where $D_i(\tau_0,\boldsymbol{\delta}) := \Delta_i(\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0) - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta})_- - \Delta_i(\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_0))_-$ and u_- denotes $u1\{u<0\}$, we have $\mathcal{A}_0^0(t) \le \mathcal{B}_0^0(t) + \mathcal{C}_0^0(t)$, where

$$\mathcal{B}_0^0(t) := \sup_{\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} \le t^2, \|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}} \left|\mathbb{G}_n\left[V_i \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}\right]\right| \quad \text{and} \quad \mathcal{C}_0^0(t) := \sup_{\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} \le t^2, \|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}} \left|\mathbb{G}_n\left[V_i D_i(\tau_0,\boldsymbol{\delta})\right]\right|.$$

First, we consider $\mathcal{B}_0^0(t)$. Recall that $Z_{j,i}$ denotes the jth element of the vector $\mathbf{Z}_i$. Since $\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} \le t^2$ and $\|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}$, then $\|\boldsymbol{\delta}\| \le t/\sqrt{\lambda_{\min}}$, and $\|\boldsymbol{\delta}\|_1 \le \sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}$,

$$\begin{aligned} E\left[\exp(\theta\mathcal{B}_0^0(t))\right] &\le E\left[\exp\left(\sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}\theta \max_{j\in S(\boldsymbol{\delta})}\left|\mathbb{G}_n\left[V_i Z_{j,i}\right]\right|\right)\right] \\ \le& \sum_{j\in S(\boldsymbol{\delta})} E\left[\exp\left(\sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}\theta\left|\mathbb{G}_n\left[V_i Z_{j,i}\right]\right|\right)\right] \qquad (10) \\ \le& \; 2\sum_{j\in S(\boldsymbol{\delta})} E\left[\exp\left(\sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}\theta\mathbb{G}_n\left[V_i Z_{j,i}\right]\right)\right] \le 2K_1 n^{c_1}\exp\left[\left(\sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}\theta\right)^2/2\right], \end{aligned}$$

where the first two inequalities are elementary, the third inequality follows from the fact that $E[\exp(|W|)] \le E[\exp(W) + \exp(-W)] \le 2E[\exp(W)]$ for any symmetric random variable W, and the last inequality follows from $E\left[\exp(uV_i)\right] \le \exp(u^2/2)$. Then for any $C_1 > 0$,

$$\begin{aligned} &Pr\left(\mathcal{B}_0^0(t) \ge C_1\right) \le \min_{\theta\ge 0}\exp(-\theta C_1 E\left[\exp(\theta\mathcal{B}_0^0(t))\right] \\ &\le \min_{\theta\ge 0}\exp(-\theta C_1) 2K_1 n^{c_1}\exp\left[\left(\sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}\theta\right)^2/2\right] = 2K_1 n^{c_1}\exp\left(-\frac{C_1^2}{2K_1 n^{c_1} t^2/\lambda_{\min}}\right), \end{aligned}$$

where the first inequality follows from the Markov inequality, the second inequality follows from (10), and the minimum is achieved at $\theta = C_1\left(K_1 n^{c_1} t^2/\lambda_{\min}\right)^{-1}$. Next, we consider $\mathcal{C}_0^0(t)$. We have

$$\begin{aligned} &E\left[\exp(\theta\mathcal{C}_0^0(t))\right] \le E\left[\exp\left(\theta \sup_{\|\boldsymbol{\delta}\|_1 \le \sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}, \|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}}\left|\mathbb{G}_n\left[V_i D_i(\tau_0,\boldsymbol{\delta})\right]\right|\right)\right] \\ &\le E\left[\exp\left(2\theta \sup_{\|\boldsymbol{\delta}\|_1 \le \sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}, \|\boldsymbol{\delta}\|_0 \le K_1 n^{c_1}}\left|\mathbb{G}_n\left[\Delta_i V_i \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}\right]\right|\right)\right] \\ &\le 2K_1 n^{c_1}\exp\left[\left(2\sqrt{K_1} n^{c_1/2} t/\sqrt{\lambda_{\min}}\theta\right)^2/2\right], \end{aligned}$$

where the second inequality follows from Theorem 4.12 in Ledoux and Talagrand (1991), and the contractive property that $|D_i(\tau_0,\boldsymbol{\delta}_1)-D_i(\tau_0,\boldsymbol{\delta}_2)|\le|\Delta_i\mathbf{Z}_i^{\mathrm{T}}(\boldsymbol{\delta}_1-\boldsymbol{\delta}_2)|$, and the rest inequalities follow exactly as for $\mathcal{B}_0^0(t)$. By Markov inequality again,

$$Pr\left(\mathcal{C}_0^0(t)\ge 2C_1\right)\le 2K_1n^{c_1}\exp\left(-\frac{C_1^2}{2K_1n^{c_1}t^2/\lambda_{\min}}\right).$$

If we choose $C_1>t$, we obtain

$$Pr\left(\mathcal{A}_0(t)\ge 12C_1\right)\le\frac{2Pr\left(\mathcal{A}_0^0(t)\ge 3C_1\right)}{1-\frac{4t^2}{144C_1^2}}\le 16K_1n^{c_1}\exp\left(-\frac{C_1^2}{2K_1n^{c_1}t^2/\lambda_{\min}}\right).$$

□

Lemma 4. *Given $0\le k\le m-1$ and $S\subset\{1,\dots,p\}$ such that $|S|\le K_1n^{c_1}$, under conditions (A1) – (A7), if $\|\boldsymbol{\beta}-\boldsymbol{\beta}^*(\tau_k)\|\le\zeta_1\nu_{k,n}(\zeta_2)$ and $S(\boldsymbol{\beta})\subseteq S$, then for sufficiently large n,*

$$Pr\Bigg(\sup_{\|\boldsymbol{\beta}-\boldsymbol{\beta}^*(\tau_k)\|\le\zeta_1\nu_{k,n}(\zeta_2),S(\boldsymbol{\beta})\subseteq S}\Bigg|\mathbb{E}_n\Big[Z_{j,i}\Big(1\Big\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)>0\Big\}$$
$$-1\Big\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_r)>0\Big\}\Big)\Big]\Bigg|>8\bar f\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2)\Bigg)\le 4\exp(-4K_1n^{c_1}\log n).$$

Proof of Lemma 4. We cover the ball $\|\boldsymbol{\delta}\|\le\zeta_1\nu_{k,n}(\zeta_2)$ and $S(\boldsymbol{\delta})\subseteq S$ with cubes $C=\{C(\boldsymbol{\delta}_l)\}$, where $C(\boldsymbol{\delta}_l)$ is a cube containing $\boldsymbol{\delta}_l$ with sides of length $\zeta_1\nu_{k,n}(\zeta_2)n^{-2}$ so that $N:=|C|=(4n^2)^{K_1n^{c_1}}$, $\|\boldsymbol{\delta}_l\|\le\zeta_1\nu_{k,n}(\zeta_2)$ and for $\boldsymbol{\delta}\in C(\boldsymbol{\delta}_l)$, $\|\boldsymbol{\delta}-\boldsymbol{\delta}_l\|\le\zeta_1\nu_{k,n}(\zeta_2)(K_1n^{c_1-4})^{1/2}=:\zeta_{k,n}$.

Let $T_{n,k}(\boldsymbol{\delta}):=\mathbb{E}_n\left[Z_{j,i}1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}\right)>0\right\}\right]$.

$$|T_{n,k}(\boldsymbol{\delta})-T_{n,k}(0)|\le\max_{1\le l\le N}|T_{n,k}(\boldsymbol{\delta}_l)-T_{n,k}(\mathbf{0})|+\max_{1\le l\le N}\sup_{\boldsymbol{\delta}\in C(\boldsymbol{\delta}_l)}|T_{n,k}(\boldsymbol{\delta})-T_{n,k}(\boldsymbol{\delta}_l)|$$
$$\le\max_{1\le l\le N}|T_{n,k}(\boldsymbol{\delta}_l)-T_{n,k}(\mathbf{0})|$$
$$+\max_{1\le l\le N}\Bigg|n^{-1}\sum_{i=1}^n|Z_{j,i}|\,1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)+(K_1n^{c_1})^{1/2}\zeta_{k,n}>0\right\}$$
$$-E\left[|Z_{j,i}|\,1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)+(K_1n^{c_1})^{1/2}\zeta_{k,n}>0\right\}\right]$$
$$-n^{-1}\sum_{i=1}^n|Z_{j,i}|\,1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}$$
$$+E\left[|Z_{j,i}|\,1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}\right]\Bigg|$$
$$+\max_{1\le l\le N}E\Big[|Z_{j,i}|\,1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)+(K_1n^{c_1})^{1/2}\zeta_{k,n}>0\right\}$$
$$-|Z_{j,i}|\,1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}\Big]$$
$$:=I_1+I_2+I_3.$$

We consider I_3 first. Noting that $|\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}_l|\le(K_1n^{c_1})^{1/2}\zeta_1\nu_{k,n}(\zeta_2)$, we obtain

$$\begin{aligned}I_3&\le[\bar f+A(K_1n^{c_1})^{1/2}\left(\zeta_1\nu_{k,n}(\zeta_2)+\zeta_{k,n}\right)](K_1n^{c_1})^{1/2}\zeta_{k,n}\\&\le 2\bar f(K_1n^{c_1})^{1/2}\zeta_{k,n}=2\bar f\zeta_1\nu_{k,n}(\zeta_2)K_1n^{c_1-2},\end{aligned}\tag{11}$$

where the first inequality follows from the conditional expectation and from Conditions (A2) and (A3), and the last two inequalities are trivial for sufficiently large n.

We next consider I_1.

$$\begin{aligned}|T_{n,k}(\boldsymbol{\delta}_l)-T_{n,k}(0)| &= n^{-1/2}\left|\mathbb{G}_n\left[Z_{j,i}\Big\{1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}-1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)>0\right\}\Big\}\right]\right|\\ &\quad+\left|E\left[Z_{j,i}\Big\{1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}-1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)>0\right\}\Big\}\right]\right|\\ &=: I_{1,1}+I_{1,2}.\end{aligned}$$

Conditional on $\mathbf{Z}_i$,

$$\begin{aligned}I_{1,2} &\leq E\left[|Z_{j,i}|\left\{\bar{f}+A(K_1n^{c_1})^{1/2}\zeta_1\nu_{k,n}(\zeta_2)\right\}|\boldsymbol{\delta}_l^{\mathrm{T}}\mathbf{Z}_i|\right]\leq 2\bar{f}E\left[|\boldsymbol{\delta}_l^{\mathrm{T}}\mathbf{Z}_i|\right]\\ &\leq 2\bar{f}\left(E\left[\boldsymbol{\delta}_l^{\mathrm{T}}\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}_l\right]\right)^{1/2}\leq 2\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2).\end{aligned}\tag{12}$$

where the first inequality follows from Condition (A3), the second inequality follows from Condition (A2), the third inequality is trivial, and the last one follows from Condition (A6).

Similarly, we can show that

$$E\left[\left(Z_{j,i}\Big\{1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}-1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)>0\right\}\Big\}\right)^2\right]\leq 2\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2).$$

As $Z_{j,i}\Big\{1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}-1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)>0\right\}\Big\}$ is bounded under Condition (A2), by Bernstein's inequality, we have

$$\begin{aligned}&Pr\left(n^{-1/2}\left|\mathbb{G}_n\left[Z_{j,i}\Big\{1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}-1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)>0\right\}\Big\}\right]\right|>t\right)\\ &\leq 2\exp\left(-\frac{1}{2}\frac{(nt)^2}{2\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2)n+nt/3}\right)=2\exp\left(-\frac{1}{2}\frac{nt^2}{2\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2)+t/3}\right).\end{aligned}$$

Choose $t=8\left(\bar{f}\lambda_{\max}^{1/2}\zeta_1e^{\zeta_2}K_1^{3/2}\right)^{1/2}n^{\frac{3(c_1-1)}{4}}\log^{3/4}n$. It follows that $t=o(n^{\frac{c_1-1}{2}}\log^{1/2}n)$. Thus, $t=o(\zeta_1\nu_{k,n}(\zeta_2))$. Then

$$\begin{aligned}&Pr\left(\sup_l n^{-1/2}\left|\mathbb{G}_n\left[Z_{j,i}\Big\{1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\left(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta}_l\right)>0\right\}-1\left\{\log X_i-\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)>0\right\}\Big\}\right]\right|>t\right)\\ &\leq 2(4n^2)^{K_1n^{c_1}}\exp\left(-\frac{1}{2}\frac{n\times 64\left(\bar{f}\lambda_{\max}^{1/2}\zeta_1e^{\zeta_2}K_1^{3/2}\right)n^{\frac{3(c_1-1)}{2}}\log^{3/2}n}{4\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2)}\right)\\ &\leq 2\exp\left(-\frac{8n\times\zeta_1e^{\zeta_2}K_1^{3/2}n^{\frac{3(c_1-1)}{2}}\log^{3/2}n}{\zeta_1e^{\zeta_2}(K_1n^{c_1-1}\log n)^{1/2}}-K_1n^{c_1}(2\log n+\log 4)\right)\leq 2\exp(-4K_1n^{c_1}\log n).\end{aligned}\tag{13}$$

Combining (12), (13), and $n^{\frac{3(c_1-1)}{4}}\log^{3/4}n=o(n^{\frac{c_1-1}{2}}\log^{1/2}n)$ together yields that

$$Pr\left(I_1>4\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2)\right)\leq 2\exp(-4K_1n^{c_1}\log n).\tag{14}$$

Now we consider I_2. By the same arguments used for $I_{1,1}$, we can show that

$$Pr\left(I_2>8n^{c_1-1}\right)\leq 2\exp(-4K_1n^{c_1}\log n).\tag{15}$$

By (11), (14), (15), and $n^{c_1-1}=o(n^{\frac{c_1-1}{2}})=o(\zeta_1\nu_{k,n}(\zeta_2))$, we obtain that

$$Pr\left(\sup_{\|\boldsymbol{\delta}\|\leq\zeta_1\nu_{k,n}(\zeta_2),S(\boldsymbol{\delta})\subseteq S}|T_{n,k}(\boldsymbol{\delta})-T_{n,k}(0)|\geq 8\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{k,n}(\zeta_2)\right)\leq 4\exp(-4K_1n^{c_1}\log n).$$

□

Lemma 5. *Given $1 \le k \le m$ and a $S \subset \{1,\dots,p\}$ such that $S^* \subseteq S$ and $|S| \le K_1 n^{c_1}$, under event $\acute{\Omega}_{S,k-1}(\zeta_1,\zeta_2)$, if $t \le \kappa/\sqrt{\lambda_{\min}}$*

$$\begin{aligned} & n^{-1}E\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta})\right] - n^{-1}E\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k))\right] \\ & \ge \underline{g}t^2 - \frac{2}{3c_2}At^3 - 2\frac{\epsilon_n}{1-\tau_U}\sum_{r=0}^{k-1}\left(\bar{f}\zeta_1\sqrt{\lambda_{\max}}\nu_{r,n}(\zeta_2)+L\epsilon_n\right)t \end{aligned}$$

uniformly for $\boldsymbol{\delta}$ satisfying $\boldsymbol{\delta}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2$ and $S(\boldsymbol{\delta}) \subseteq S$.

Proof of Lemma 5. By the identity used in Lemma 2, we have

$$\begin{aligned} & n^{-1}E\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta})\right] - n^{-1}E\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k))\right] \\ &= 2E\left[\Delta_i\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}1\left\{\log X_i < \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)\right\} - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}\left(\sum_{r=0}^{k-1}\int_{\tau_r}^{\tau_{r+1}} 1\left\{\log X_i \ge \mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)\right\}dH(\tau)+\tau_0\right)\right] \\ &\quad + 2E\left[\Delta_i\int_0^{\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}}\left(1\left\{\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k) < t\right\} - 1\left\{\log X_i < \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)\right\}\right)dt\right] \\ &:= I_1 + I_2. \end{aligned}$$

We consider I_1 first. By the Martingale equality, I_1 can be written as

$$2E\left[\mathbf{Z}^{\mathrm{T}}\boldsymbol{\delta}\sum_{r=0}^{k-1}\int_{\tau_r}^{\tau_{r+1}} 1\left\{\log X \ge \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(u)\right\} - 1\left\{\log X \ge \mathbf{Z}_i^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)\right\}dH(u)\right].$$

Under $\acute{\Omega}_{S,k-1}(\zeta_1,\zeta_2)$, we obtain

$$\begin{aligned} |I_1| &\le 2\sup_{\boldsymbol{\delta}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta}=t^2,\|\boldsymbol{\delta}\|_0\le K_1n^{c_1}}\left|E\left[\mathbf{Z}^{\mathrm{T}}\boldsymbol{\delta}\sum_{r=0}^{k-1}\int_{\tau_r}^{\tau_{r+1}}E\left[1\left\{\log X \ge \mathbf{Z}^{\mathrm{T}}\boldsymbol{\beta}^*(u)\right\} - \right.\right.\right. \\ &\qquad\qquad \left.\left.\left. 1\left\{\log X \ge \mathbf{Z}^{\mathrm{T}}\acute{\boldsymbol{\beta}}(\tau_r)\right\}|\mathbf{Z}\right]dH(u)\right]\right| \\ &\le 2\sum_{r=0}^{k-1}\left(2\bar{f}\zeta_1\sqrt{\lambda_{\max}}\nu_{r,n}(\zeta_2)+L\epsilon_n\right)\left(H(\tau_{r+1})-H(\tau_r)\right)\left(E\left[\boldsymbol{\delta}^{\mathrm{T}}\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}\right]\right)^{1/2} \\ &\le 2\frac{\epsilon_n}{1-\tau_U}\sum_{r=0}^{k-1}\left(2\bar{f}\zeta_1\sqrt{\lambda_{\max}}\nu_{r,n}(\zeta_2)+L\epsilon_n\right)t, \end{aligned} \tag{16}$$

where the first inequality follows the proof of (11).

We now evaluate I_2. By Condition (A6),

$$\begin{aligned} & 2E\left[\int_0^{\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}}\Delta_i\left(1\left\{\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k) < t\right\} - 1\left\{\log X_i < \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)\right\}\right)dt\right] \\ \ge\ & 2E\left[\int_0^{\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta}} g\left(\exp\left\{\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau_k)\right\}|\mathbf{Z}_i\right)u - Au^2du\right] \ge \underline{g}t^2 - 2At^3/(3c_2). \end{aligned} \tag{17}$$

Inequalities (16) and (17) together imply that

$$\begin{aligned} & n^{-1}E\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k)+\boldsymbol{\delta})\right] - n^{-1}E\left[\acute{L}_k(\boldsymbol{\beta}^*(\tau_k))\right] \\ & \ge \underline{g}t^2 - \frac{2}{3c_2}At^3 - 2\frac{\epsilon_n}{1-\tau_U}\sum_{r=0}^{k-1}\left(2\bar{f}\zeta_1\sqrt{\lambda_{\max}}\nu_{r,n}(\zeta_2)+L\epsilon_n\right)t \end{aligned}$$

uniformly for $\boldsymbol{\delta}$ satisfying $\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right] \boldsymbol{\delta} = t^2$ and $\|\boldsymbol{\delta}\|_0 \leq K_1 n^{c_1}$.

□

Lemma 6. *Given $1 \leq k \leq m$, let*

$$\mathcal{A}_k(t) := \sup_{\boldsymbol{\delta}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\delta} = t^2, \|\boldsymbol{\delta}\|_0 \leq K_1 n^{c_1}} \left| \mathbb{G}_n \left[\Delta_i \left(\left| \log X_i - \mathbf{Z}_i^{\mathrm{T}}(\boldsymbol{\beta}^*(\tau) + \boldsymbol{\delta}) \right| \right.\right.\right.$$
$$\left.\left.\left. - \left| \log X_i - \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}^*(\tau) \right| \right) + \Delta_i \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\delta} \right] \right|,$$

under conditions (C1)-(C6), we have for $C_1 > t$,

$$Pr\left(\mathcal{A}_k(t) \geq 20 C_1\right) \leq 16 K_1 n^{c_1} \exp\left(-\frac{C_1^2}{2K_1 n^{c_1} t^2/\lambda_{\min}}\right).$$

Proof of Lemma 6. Since $\Delta_i \left[\left|\log X_i - \mathbf{Z}_i^{\mathrm{T}}(\boldsymbol{\beta}^*(\tau) + \boldsymbol{\delta})\right| - \left|\log X_i - \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\beta}^*(\tau)\right| \right] = -\Delta_i \mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\delta} + 2\Delta_i D_i(\tau, \boldsymbol{\delta})$, where D_i is defined in Lemma 3, The proof follows the exactly same arguments used in Lemma 3. Therefore, we omit the details here. □

Lemma 7. *Given $1 \leq k \leq m$ and $S \subset \{1, \ldots, p\}$ such that $S^* \subseteq S$ and $|S| \leq K_1 n^{c_1}$, under conditions (A1) – (A7), for sufficiently large n, let $\mathcal{C}_k(t) :=$*

$$\sup_{\substack{\boldsymbol{\xi}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\xi} \leq t^2, S(\boldsymbol{\xi}) \subseteq S, \\ \|\boldsymbol{\beta}(\tau_r) - \boldsymbol{\beta}^*(\tau_r)\| \leq \zeta_1 \nu_{r,n}(\zeta_2), S(\boldsymbol{\beta}(\tau_r)) \subseteq S, \forall r \leq k-1}} \left| \mathbb{G}_n \left[\boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \left(\sum_{r=0}^{k-1} \int_{\tau_r}^{\tau_{r+1}} 1\left\{\log X_i > \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}(\tau_r)\right\} dH(u) + \tau_0 \right) \right] \right|,$$

we have for $C_1 > t$,

$$Pr\left(\mathcal{C}_k(t) > \sum_{r=0}^{k-1} \left[32 \frac{\epsilon_n}{1 - \tau_U} C_1 + 10 \frac{\epsilon_n}{1 - \tau_U} \bar{f} \lambda_{\max}^{1/2} \zeta_1 \nu_{r,n}(\zeta_2) \sqrt{n} t \right] + 16 \tau_0 C_1 \right)$$
$$\leq 4(5k + 4) K_1 n^{c_1} \exp\left(-\frac{C_1^2}{2K_1 n^{c_1} t^2/\lambda_{\min}}\right).$$

Proof of Lemma 7. We first consider

$$\mathcal{C}_{k,r}(t) := \sup_{\substack{\boldsymbol{\xi}^{\mathrm{T}} E\left[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\xi} \leq t^2, S(\boldsymbol{\xi}) \subseteq S, \\ \|\boldsymbol{\beta}(\tau_r) - \boldsymbol{\beta}^*(\tau_r)\| \leq \zeta_1 \nu_{r,n}(\zeta_2), S(\boldsymbol{\beta}(\tau_r)) \subseteq S}} \left| \mathbb{G}_n \left[\boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \int_{\tau_r}^{\tau_{r+1}} 1\left\{\log X_i > \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}(\tau_r)\right\} dH(u) \right] \right|$$

for $0 \le r \le k-1$. We can verify

$$\begin{aligned}
\mathcal{C}_{k,r}(t) \le & \sup_{\substack{\boldsymbol{\xi}^{\mathrm{T}} E[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}]\boldsymbol{\xi} \le t^2, S(\boldsymbol{\xi}) \subseteq S, \\ \|\boldsymbol{\beta}(\tau_r) - \boldsymbol{\beta}^*(\tau_r)\| \le \zeta_1 \nu_{r,n}(\zeta_2), S(\boldsymbol{\beta}(\tau_r)) \subseteq S}} \sqrt{n} \left| \mathbb{E}_n \left[\boldsymbol{\delta}^{\mathrm{T}} \mathbf{Z}_i \int_{\tau_r}^{\tau_{r+1}} \left[1\left\{ \log X_i > \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}(\tau_r) \right\} \right. \right. \right. \\
& \left. \left. \left. - 1\left\{ \log X_i > \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}^*(\tau_r) \right\} \right] dH(u) \right] \right| \\
& + \sup_{\substack{\boldsymbol{\xi}^{\mathrm{T}} E[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}]\boldsymbol{\xi} \le t^2, S(\boldsymbol{\xi}) \subseteq S, \\ \|\boldsymbol{\beta}(\tau_r) - \boldsymbol{\beta}^*(\tau_r)\| \le \zeta_1 \nu_{r,n}(\zeta_2), S(\boldsymbol{\beta}(\tau_r)) \subseteq S}} \sqrt{n} \left| E \left[\boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \int_{\tau_r}^{\tau_{r+1}} \left[1\left\{ \log X_i > \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}(\tau_r) \right\} \right. \right. \right. \\
& \left. \left. \left. - 1\left\{ \log X_i > \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}^*(\tau_r) \right\} \right] dH(u) \right] \right| \\
& + \sup_{\boldsymbol{\xi}^{\mathrm{T}} E[\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}]\boldsymbol{\xi} \le t^2, S(\boldsymbol{\xi}) \subseteq S} \left| \mathbb{G}_n \left[\boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \int_{\tau_r}^{\tau_{r+1}} 1\left\{ \log X_i > \mathbf{Z}_i^{\mathrm{T}} \boldsymbol{\beta}^*(\tau_r) \right\} dH(u) \right] \right| \\
=: & \mathcal{D}_{k,r,1}(t) + \mathcal{D}_{k,r,2}(t) + \mathcal{D}_{k,r,3}(t).
\end{aligned}$$

Now we consider $\mathcal{D}_{k,r,2}(t)$ first. Following the same arguments used for the term I_1 in the proof of Lemma 5, we obtain that

$$\mathcal{D}_{k,r,2}(t) \le \frac{2\epsilon_n}{1-\tau_U} \bar{f} \zeta_1 \sqrt{\lambda_{\max}} \nu_{r,n}(\zeta_2) \sqrt{n} t. \tag{18}$$

Next, we consider $\mathcal{D}_{k,r,3}(t)$. Applying the same arguments used in Lemma 3 yields, for any $C_1 > t$,

$$Pr\left(\mathcal{D}_{k,r,3}(t) \ge 16 \frac{\epsilon_n}{1-\tau_U} C_1 \right) \le 8K_1 n^{c_1} \exp\left(-\frac{C_1^2}{2K_1 n^{c_1} t^2 / \lambda_{\min}} \right). \tag{19}$$

Next, we evaluate $\mathcal{D}_{k,r,1}(t)$. Let $T_{n,r}(\boldsymbol{\delta}, \boldsymbol{\xi}) := \mathbb{E}_n \left[\boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i 1\left\{ \log X_i - \mathbf{Z}_i^{\mathrm{T}} \left(\boldsymbol{\beta}^*(\tau_k) + \boldsymbol{\delta} \right) > 0 \right\} \right]$.

$$\begin{aligned}
|T_{n,r}(\boldsymbol{\xi}, \boldsymbol{\delta}) - T_{n,r}(\boldsymbol{\xi}, 0)| \le & \max_{1 \le l \le N} |T_{n,r}(\boldsymbol{\xi}, \boldsymbol{\delta}_l) - T_{n,r}(\boldsymbol{\xi}, 0)| + \max_{1 \le l \le N} \sup_{\boldsymbol{\delta} \in C(\boldsymbol{\delta}_l)} |T_{n,r}(\boldsymbol{\xi}, \boldsymbol{\delta}) - T_{n,r}(\boldsymbol{\xi}, \boldsymbol{\delta}_l)| \\
\le & \max_{1 \le l \le N} |T_{n,r}(\boldsymbol{\xi}, \boldsymbol{\delta}_l) - T_{n,r}(\boldsymbol{\xi}, 0)| \\
& + \max_{1 \le l \le N} \left| n^{-1} \sum_{i=1}^{n} \left| \boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \right| 1\left\{ \log X_i - \mathbf{Z}_i^{\mathrm{T}} \left(\boldsymbol{\beta}^*(\tau_k) + \boldsymbol{\delta}_l \right) + (K_1 n^{c_1})^{1/2} \zeta_{k,n} > 0 \right\} \right. \\
& - E\left[\left| \boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \right| 1\left\{ \log X_i - \mathbf{Z}_i^{\mathrm{T}} \left(\boldsymbol{\beta}^*(\tau_k) + \boldsymbol{\delta}_l \right) + (K_1 n^{c_1})^{1/2} \zeta_{k,n} > 0 \right\} \right] \\
& - n^{-1} \sum_{i=1}^{n} \left| \boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \right| 1\left\{ \log X_i - \mathbf{Z}_i^{\mathrm{T}} \left(\boldsymbol{\beta}^*(\tau_k) + \boldsymbol{\delta}_l \right) > 0 \right\} \\
& \left. + E\left[\left| \boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \right| 1\left\{ \log X_i - \mathbf{Z}_i^{\mathrm{T}} \left(\boldsymbol{\beta}^*(\tau_k) + \boldsymbol{\delta}_l \right) > 0 \right\} \right] \right| \\
& + \max_{1 \le l \le N} E\left[\left| \boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \right| 1\left\{ \log X_i - \mathbf{Z}_i^{\mathrm{T}} \left(\boldsymbol{\beta}^*(\tau_k) + \boldsymbol{\delta}_l \right) + (K_1 n^{c_1})^{1/2} \zeta_{k,n} > 0 \right\} \right. \\
& \left. - \left| \boldsymbol{\xi}^{\mathrm{T}} \mathbf{Z}_i \right| 1\left\{ \log X_i - \mathbf{Z}_i^{\mathrm{T}} \left(\boldsymbol{\beta}^*(\tau_k) + \boldsymbol{\delta}_l \right) > 0 \right\} \right],
\end{aligned}$$

where $\boldsymbol{\delta}_l$'s are defined in Lemma 4.

Following the arguments used for $\mathcal{D}_{k,r,2}(t)$, $\mathcal{D}_{k,r,3}(t)$ and Lemma 4, we obtain

$$Pr\left(\mathcal{D}_{k,r,1}(t) > 16 \frac{\epsilon_n}{1-\tau_U} C_1 + \frac{8\epsilon_n}{1-\tau_U} \bar{f} \lambda_{\max}^{1/2} \zeta_1 \nu_{r,n}(\zeta_2) \sqrt{n} t \right) \le 12 K_1 n^{c_1} \exp\left(-\frac{C_1^2}{2K_1 n^{c_1} t^2 / \lambda_{\min}} \right). \tag{20}$$

Combining (18), (19) and (20) yields

$$Pr\left(\mathcal{C}_{k,r}(t) > 32\frac{\epsilon_n}{1-\tau_U}C_1 + \frac{10\epsilon_n}{1-\tau_U}\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{r,n}(\zeta_2)\sqrt{n}t\right) \leq 20K_1n^{c_1}\exp\left(-\frac{C_1^2}{2K_1n^{c_1}t^2/\lambda_{\min}}\right). \quad (21)$$

We now consider $\mathcal{C}_{k,\tau_0}(t) := \tau_0 \sup_{\boldsymbol{\xi}^{\mathrm{T}}E\left[\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}\right]\boldsymbol{\xi}\leq t^2, S(\boldsymbol{\xi})\subseteq S}\left|\mathbb{G}_n\left[\boldsymbol{\delta}^{\mathrm{T}}\mathbf{Z}_i\right]\right|$. From the proof of Lemma 3, we have

$$Pr\left(\mathcal{C}_{k,\tau_0}(t) > 16\tau_0C_1\right) \leq 16K_1n^{c_1}\exp\left(-\frac{C_1^2}{2K_1n^{c_1}t^2/\lambda_{\min}}\right). \quad (22)$$

By (21) and (22), we obtain

$$\begin{aligned}&Pr\left(\mathcal{C}_k(t) > \sum_{r=0}^{k-1}\left[32\frac{\epsilon_n}{1-\tau_U}C_1 + 10\frac{\epsilon_n}{1-\tau_U}\bar{f}\lambda_{\max}^{1/2}\zeta_1\nu_{r,n}(\zeta_2)\sqrt{n}t\right] + 16\tau_0C_1\right)\\ &\leq 4(5k+4)K_1n^{c_1}\exp\left(-\frac{C_1^2}{2K_1n^{c_1}t^2/\lambda_{\min}}\right).\end{aligned}$$

□

References

Andersen, P. K., O. Borgan, R. D. Gill, and N. Keiding (2012). *Statistical models based on counting processes.* Springer Science & Business Media.

Knight, K. (1998). Limiting distributions for l_1 regression estimators under general conditions. *The Annals of Statistics 26*, 755–770.

Lange, K. (2004). *Optimization* (2 ed.). Springer.

Ledoux, M. and M. Talagrand (1991). *Probability in Banach Space: Isoperimetry and Processes.*, Volume 23. Springer Science & Business Media.

Peng, L. and Y. Huang (2008). Survival analysis with quantile regression models. *Journal of the American Statistical Association 103*(482), 637–649.

Römisch, W. (2014). Delta method, infinite dimensional. *Wiley StatsRef: Statistics Reference Online*.

Van Der Vaart, A. and J. Wellner (2000). *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer, New York.

Van der Vaart, A. W. (2000). *Asymptotic Statistics*, Volume 3. Cambridge University Press.

Zheng, Q., L. Peng, and X. He (2018). High dimensional censored quantile regression. *The Annals of Statistics 46*(1), 308–343.

WebTable 1: Comparisons with three ad-hoc approaches based on 100 simulated datasets. The results are summarized at $\tau = 0.75$ for $\beta_4(\tau)$.

	Bias				SE				Selection frequency			
	Fused	OneSplit	NoSplit	Half	Fused	OneSplit	NoSplit	Half	Fused	OneSplit	NoSplit	Half
					$n = 300, p = 1000$							
β_4	-0.14	-0.18	-0.16	-0.54	0.34	0.61	0.38	0.32	0.73	0.72	0.83	0.71
β_{21}	-0.00	0.02	-0.06	-0.41	0.13	0.26	0.15	0.27	0.69	0.63	0.95	0.64
β_{41}	0.02	-0.00	-0.08	-0.20	0.14	0.26	0.16	0.14	0.99	0.99	1.00	0.99
β_{61}	-0.01	-0.05	-0.15	-0.29	0.15	0.26	0.17	0.14	1.00	1.00	1.00	1.00
					$n = 700, p = 1000$							
β_4	-0.03	-0.05	-0.07	-0.24	0.23	0.31	0.25	0.23	0.89	0.92	0.99	0.90
β_{21}	0.01	0.02	-0.02	-0.07	0.09	0.15	0.10	0.10	0.99	0.98	1.00	0.99
β_{41}	0.01	0.01	-0.05	-0.11	0.09	0.17	0.10	0.09	1.00	1.00	1.00	1.00
β_{61}	-0.00	-0.01	-0.07	-0.15	0.10	0.17	0.13	0.11	1.00	1.00	1.00	1.00