Supplementary Materials for "Generalizing Quantile Regression for Counting Processes with Applications to Recurrent Events"

December 3, 2014

SUPPLEMENT A: JUSTIFICATIONS FOR ALTERNATIVE MODEL REPRESENTATIONS

We present the justifications separately for the nonrecurrent event setting and the recurrent events setting.

Nonrecurrent event setting: Let $f_{\boldsymbol{Z}}^{T}(\cdot)$ denote the conditional density function of T given $\boldsymbol{Z}$, $S_{\boldsymbol{Z}}^{T}(\cdot)$ and $\Lambda_{\boldsymbol{Z}}^{T}(\cdot)$ denote the survival function and cumulative hazard function of T given $\boldsymbol{Z}$ respectively, and $S_{\boldsymbol{Z}}^{C}(\cdot)$ denote the survival function of C given $\boldsymbol{Z}$. Define $\mu^{nr}(t) = E\{N^{nr}(t)\}$. We assume the following regularity conditions:

(D0) T and C are independent given $\boldsymbol{Z}$;

(D1) $\boldsymbol{\beta}_0^{nr}(\tau)$ is continuously differentiable;

(D2) $f_{\boldsymbol{Z}}^{T}(e^{\boldsymbol{X}^{\mathsf{T}}\boldsymbol{\beta}_0^{nr}(\tau)})S_{\boldsymbol{Z}}^{C}(e^{\boldsymbol{X}^{\mathsf{T}}\boldsymbol{\beta}_0^{nr}(\tau)}) > 0$ for $\tau \in (0, \tau_U]$ and $\boldsymbol{Z} \in \mathcal{Z}$, where τ_U is some constant $\in (0, 1)$.

Proposition A1. (i) Under conditions (D0) and (D1), model (1) implies model (3); (ii) Under conditions (D0), (D1), and (D2), model (3) implies model (1) with $\tau \in (0, 1)$ replaced by $\tau \in (0, \tau_U]$, i.e.

$$Q_T(\tau|\boldsymbol{Z}) = \exp\{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)\}, \quad \tau \in (0, \tau_U]. \tag{A.1}$$

Proof of Proposition A1: Peng and Huang (2008) proved that (3) holds under model (1) and assumptions, (D0) and (D1). Thus we only need to show that (3) implies (A.1) under conditions (D0)–(D2).

First, we take derivative with respect to τ on both sides of equation (3). Under the random censoring assumption (D0) that T and C are independent given $\boldsymbol{Z}$, for $\tau \in (0, 1)$, we get

$$f_{\boldsymbol{Z}}^T(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) S_{\boldsymbol{Z}}^C(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)} \boldsymbol{X}^\top d\boldsymbol{\beta}_0^{nr}(\tau) = S_{\boldsymbol{Z}}^T(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) S_{\boldsymbol{Z}}^C(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) \frac{1}{1-\tau} d\tau.$$

Given the assumption (D2), this implies that for $\tau \in (0, \tau_U]$,

$$d\Lambda_{\boldsymbol{Z}}^T(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) = \frac{1}{1-\tau} d\tau. \tag{A.2}$$

By letting τ approaching 0, equation (3) implies that $\lim_{\tau \to 0} \mu_{\boldsymbol{Z}}^{nr}(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) = 0$. Given (D2), $\lim_{\tau \to 0} \mu_{\boldsymbol{Z}}^{nr}(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) = 0$ implies $e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(0)} = 0$ and hence $\Lambda_{\boldsymbol{Z}}^T(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(0)}) = 0$ for $\boldsymbol{Z} \in \mathcal{Z}$. Therefore, the differential equation (A.2) gives $\Lambda_{\boldsymbol{Z}}^T(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) = -\log(1-\tau)$ for $\tau \in (0, \tau_U]$, which means $S_{\boldsymbol{Z}}^T(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)}) = 1 - \tau$ and thus $Q_T(\tau|\boldsymbol{Z}) = \exp\{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau)\}$ for $\tau \in (0, \tau_U]$. This proves that (3) implies (A.1) under conditions (D0)–(D2).

Remark A1: By Proposition A1, the quantile regression model (1) implies the counting process model (3) as long as the standard random assumption is satisfied and regression quantiles in $\boldsymbol{\beta}_0(\tau)$ are smooth in τ. Showing the reverse relationship however requires an additional assumption (D2). In (D2), the density positiveness at regression quantiles is commonly adopted in quantile regression literature for establishing estimation consistency, while the other requirement on $S_{\boldsymbol{Z}}^C(\cdot)$ is necessitated by the identifiability of model (1) in the presence of censoring. For example, when the censoring variable C is always smaller than conditional quantiles at a certain quantile level, say τ^*, (i.e. $S_{\boldsymbol{Z}}^C(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0^{nr}(\tau^*)}) = 0$), $\boldsymbol{\beta}_0^{nr}(\tau)$

in model (1) would not be identifiable for $\tau \in (\tau^*, 1)$. In this case, model (1) cannot be implied by model (3), which is formulated based on the fully observable counting process $N_i(t)$ and at-risk process $Y_i(t)$. Such a view of condition (D2) may lead to an alternative interpretation of Proposition A1(ii). That is, under conditions (D0) and (D1), model (3) implies an identifiable version of model (1), model (A.1) with τ_U satisfying condition (D2).

Recurrent events setting: Let $F_{\boldsymbol{Z}}^W(x) = E\{Y(x)|\boldsymbol{Z}\}$. We assume the following regularity conditions:

(E0) (L, R) is independent of $\tilde{N}(\cdot)$ given $\boldsymbol{Z}$;

(E1) $\boldsymbol{\beta}_0(u)$ is continuously differentiable;

(E2) $\dot{\mu}_{\boldsymbol{Z}}(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)}) F_{\boldsymbol{Z}}^W(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)}) > 0$ for $\boldsymbol{Z} \in \mathcal{Z}$ and $u \in (0, U]$.

Proposition A2. (i) Under conditions (E0) and (E1), model (7) implies model (6); (ii) under conditions (E0)–(E2), model (6) implies (7).

Proof of Proposition A2: Suppose (7) holds. Under the random observation window assumption (E0), we see that $E\{N(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)})|L, R, \boldsymbol{Z}\} = \mu_{\boldsymbol{Z}}(R \wedge e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)}) - \mu_{\boldsymbol{Z}}(L)$. On the other hand, by the smoothness of $\boldsymbol{\beta}_0(\cdot)$ stated in (E1) and the definition of $\tau_{\boldsymbol{Z}}(u)$, we have $L < e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(s)} \le R$ is equivalent to $\mu_{\boldsymbol{Z}}(L) < G(s) \le \mu_{\boldsymbol{Z}}(R)$. As a result,

$$E\Big\{\int_0^u Y(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(s)}) g(s) ds | L, R, \boldsymbol{Z}\Big\} = E\Big\{\int_0^\infty I(\mu_{\boldsymbol{Z}}(L) < G(s) \le \mu_{\boldsymbol{Z}}(R) \wedge G(u)) dG(s) | L, R, \boldsymbol{Z}\Big\}$$

Since $G(u) = \mu_{\boldsymbol{Z}}(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)})$ for $u \in (0, U]$ under model (7), it follows from the above equation that for $u \in (0, U]$,

$$E\Big\{\int_0^u Y(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(s)}) g(s) ds | L, R, \boldsymbol{Z}\Big\} = \mu_{\boldsymbol{Z}}(R \wedge e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)}) - \mu_{\boldsymbol{Z}}(L).$$

Consequently, $E\{N(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)})|L, R, \boldsymbol{Z}\} = E\big\{\int_0^u Y(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(s)}) g(s) ds | L, R, \boldsymbol{Z}\big\}$ for $u \in (0, U]$, and therefore (6) is satisfied.

Suppose (6) holds. Under the random observation window assumption in (E0), we get from taking derivative with respect to u on both sides of equation (6) that

$$\dot{\mu}_{\boldsymbol{Z}}(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)}) F_{\boldsymbol{Z}}^W(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)}) e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)} \boldsymbol{X}^\top d\boldsymbol{\beta}_0(u) = F_{\boldsymbol{Z}}^W(e^{\boldsymbol{X}^\top \boldsymbol{\beta}_0(u)}) g(u) du.$$

It then follows from (E2) that $d\mu_{\boldsymbol{Z}}(e^{\boldsymbol{X}^\top\boldsymbol{\beta}_0(u)}) = g(u)du$ for $u \in (0, U]$.

By equation (6) and condition (E2), we have $e^{\boldsymbol{X}^\top\boldsymbol{\beta}_0(0)} = 0$ for $\boldsymbol{Z} \in \mathcal{Z}$. This, combined with $d\mu_{\boldsymbol{Z}}(e^{\boldsymbol{X}^\top\boldsymbol{\beta}_0(u)}) = g(u)du$ for $u \in (0, U]$, implies that $\mu_{\boldsymbol{Z}}(e^{\boldsymbol{X}^\top\boldsymbol{\beta}_0(u)}) = G(u)$ for $u \in (0, U]$. This means that (7) is satisfied. The proof of Proposition A2 is completed.

Remark A2: Like in the nonrecurrent event setting, condition (E2) is attached to the identifiability aspect of model (7) with recurrent events data under window observation. The result in proposition A2(ii) justifies the use of model (7) to interpret the counting process model (6) when (E2) is met.

SUPPLEMENT B: PROOFS OF THEOREM 1 AND THEOREM 2

Define $\mathcal{B}(d) = \{\boldsymbol{b} \in R^{p+1} : \inf_{u\in(0,U]} \|\boldsymbol{A}(\boldsymbol{b}) - \boldsymbol{A}(\boldsymbol{\beta}_0(u))\| \le d\}$. Here $\|\cdot\|$ denote the Euclidean norm. Without loss of generality, we assume $S_{L(n)}$ is a equally spaced grid, and thus $L(n) = U/a_n$, where $a_n = \|S_{L(n)}\|$. Define $\boldsymbol{\alpha}_0(u) = \boldsymbol{A}(\boldsymbol{\beta}_0(u))$, $\hat{\boldsymbol{\alpha}}(u) = \boldsymbol{A}(\hat{\boldsymbol{\beta}}(u))$, and $\mathcal{A}(d) = \{\boldsymbol{A}(\boldsymbol{b}) : \boldsymbol{b} \in \mathcal{B}(d)\}$.

Proof of Theorem 1. First, note that for any $\boldsymbol{b}_1$, $\boldsymbol{b}_2 \in \mathcal{B}(d_0)$, $(\boldsymbol{b}_1 - \boldsymbol{b}_2)\{\boldsymbol{A}(\boldsymbol{b}_1) - \boldsymbol{A}(\boldsymbol{b}_2)\} = E[(\boldsymbol{X}^\top\boldsymbol{b}_1 - \boldsymbol{X}^\top\boldsymbol{b}_2)\{N(\exp(\boldsymbol{X}^\top\boldsymbol{b}_1)) - N(\exp(\boldsymbol{X}^\top\boldsymbol{b}_2))\}] \ge 0$. Under condition C3, the equality, $(\boldsymbol{b}_1 - \boldsymbol{b}_2)\{\boldsymbol{A}(\boldsymbol{b}_1) - \boldsymbol{A}(\boldsymbol{b}_2)\} = 0$, holds only when $\boldsymbol{b}_1 = \boldsymbol{b}_2$. This implies that $\boldsymbol{A}(\cdot)$ is a one-to-one map from $\mathcal{B}(d_0)$ to $\mathcal{A}(d_0)$. Therefore, there exists an inverse function, denoted by $\boldsymbol{\kappa}(\cdot)$, from $\mathcal{A}(d_0)$ to $\mathcal{B}(d_0)$, such that $\boldsymbol{\kappa}(\boldsymbol{A}(\boldsymbol{b})) = \boldsymbol{b}$ for any $\boldsymbol{b} \in \mathcal{B}(d_0)$.

By the definition of $\hat{\boldsymbol{\beta}}(u)$, we have $\max_{k=1,\ldots,L(n)} \|\boldsymbol{\nu}_{n,k}\| = O_p(n^{-1})$, where $\boldsymbol{\nu}_{n,k} = n^{-1}\sum_{i=1}^n \boldsymbol{X}_i N_i(\exp\{\boldsymbol{X}_i^T\hat{\boldsymbol{\beta}}(u)\}) - n^{-1}\sum_{i=1}^n \int_0^u \boldsymbol{X}_i I(L_i \le \exp\{\boldsymbol{X}_i^T\hat{\boldsymbol{\beta}}(s)\} \le R_i)g(s)ds$. Define $\boldsymbol{v}_n(\boldsymbol{b}) = n^{-1}\sum_{i=1}^n \boldsymbol{X}_i N_i(\exp(\boldsymbol{X}_i^\top\boldsymbol{b})) - \boldsymbol{A}(\boldsymbol{b})$ and $\tilde{\boldsymbol{v}}_n(\boldsymbol{b}) = n^{-1}\sum_{i=1}^n \boldsymbol{X}_i I(L_i \le \exp(\boldsymbol{X}_i^\top\boldsymbol{b}) \le R_i) - \tilde{\boldsymbol{A}}(\boldsymbol{b})$. Consider the two function classes, $\mathcal{G}_1 = \{\boldsymbol{X}\sum_{j=1}^\infty I(L \le T^{(j)} \le \exp(\boldsymbol{X}^\top\boldsymbol{b}) \wedge R) : \boldsymbol{b} \in R^{p+1}\}$ and $\mathcal{G}_2 = \{\boldsymbol{X}I(L \le \exp(\boldsymbol{X}^\top\boldsymbol{b}) \le R) : \boldsymbol{b} \in R^{p+1}\}$. Under condition C1, both $\mathcal{G}_1$ and $\mathcal{G}_2$ are Glivenko-Cantelli (van der Vaart and Wellner 2000) because the class of uniformly bounded monotone function functions in the real line is Donsker and thus Glivenko-Cantelli.

Therefore, $\sup_{\boldsymbol{b}} \|\boldsymbol{v}_n(\boldsymbol{b})\| \to_{a.s.} 0$ and $\sup_{\boldsymbol{b}} \|\tilde{\boldsymbol{v}}_n(\boldsymbol{b})\| \to_{a.s.} 0$. These imply $r_n \equiv \sup_{1\le k\le L(n)} \| - \boldsymbol{v}_n(\hat{\boldsymbol{\beta}}(u_k)) + \int_0^{u_k} \tilde{\boldsymbol{v}}_n(\hat{\boldsymbol{\beta}}(s))g(s)ds\| \to_{a.s.} 0$.

Under conditions C1–C3, following the same arguments as in Peng and Huang (2008), we can show that, for $k = 1, \ldots, L(n)$, $\sup_{u\in[u_{k-1},u_k)} \|\hat{\boldsymbol{\alpha}}(u) - \boldsymbol{\alpha}_0(u)\|$ is bounded above almost surely by $\varepsilon_k \equiv (1 + C_3 a_n)^{k-1}(r_n + \varepsilon_0 C_3 a_n + C_1 n^{-1} + C_2 a_n)$, where C_l $(l = 1, 2, 3)$ are some positive constants. Given $\lim_{n\to\infty} a_n = 0$ and $L(n) = U/a_n$, we have $\lim_{n\to\infty}(1+C_3 a_n)^{L(n)} = e^{C_3 U}$. This, coupled with $r_n \to_{a.s.} 0$, then implies that

$$\sup_{u\in(0,U]} \|\hat{\boldsymbol{\alpha}}(u) - \boldsymbol{\alpha}_0(u)\| \to_p 0. \tag{B.1}$$

With applications of Taylor expansion of $\boldsymbol{\kappa}(\hat{\boldsymbol{\alpha}}(u))$ around $\boldsymbol{\alpha}_0(u)$, it follows under condition C4 that $\sup_{u\in[v,U]} \|\hat{\boldsymbol{\beta}}(u) - \boldsymbol{\beta}_0(u)\| \to_p 0$.

Proof of Theorem 2. Following the proof of Lemma B.1. in Peng and Huang (2008), we can show that, given $\sup_{u\in(0,U]} \|\hat{\boldsymbol{\alpha}}(u) - \boldsymbol{\alpha}_0(u)\| \to_p 0$,

$$\begin{aligned}\sup_{u\in(0,U]} \Big\| n^{-1/2} \sum_{i=1}^{n} \boldsymbol{X}_i \{\tilde{N}_i(\exp(\boldsymbol{X}_i^T \hat{\boldsymbol{\beta}}(u))) - \tilde{N}_i(\exp(\boldsymbol{X}_i^T \boldsymbol{\beta}_0(u)))\} \\ - n^{-1/2}\{\mathbf{A}(\hat{\boldsymbol{\beta}}(u)) - \mathbf{A}(\boldsymbol{\beta}_0(u))\}\Big\| \xrightarrow{p} 0.\end{aligned} \tag{B.2}$$

and

$$\begin{aligned}\sup_{u\in(0,U]} \Big\| n^{-1/2} \sum_{i=1}^{n} \boldsymbol{X}_i \{I(L_i \le \exp(\boldsymbol{X}_i^T \hat{\boldsymbol{\beta}}(u)) \le R_i) - I(L_i \le \exp(\boldsymbol{X}_i^T \boldsymbol{\beta}_0(u)) \le R_i)\} \\ - n^{-1/2}\{\tilde{\mathbf{A}}(\hat{\boldsymbol{\beta}}(u)) - \tilde{\mathbf{A}}(\boldsymbol{\beta}_0(u))\}\Big\| \xrightarrow{p} 0.\end{aligned} \tag{B.3}$$

Let $o_I(a_n)$ denote a term that is $o_p(a_n)$ uniformly in $u \in I$. Given $n^{1/2}\|S_{L(n)}\| \to 0$, by the definition of $\boldsymbol{S}_n(\cdot)$, we can show that $n^{1/2}\mathbf{S}_n(\hat{\boldsymbol{\beta}}, u) = o_{(0,U]}(1)$. Then (B.1), (B.2) and (B.3) imply that

$$
\begin{aligned}
& - n^{1/2}\mathbf{S}_n(\boldsymbol{\beta}_0, u) \\
=& n^{1/2}\{\mathbf{A}(\hat{\boldsymbol{\beta}}(u)) - \mathbf{A}(\boldsymbol{\beta}_0(u))\} - \int_0^u n^{1/2}\{\tilde{\mathbf{A}}(\hat{\boldsymbol{\beta}}(s)) - \tilde{\mathbf{A}}(\boldsymbol{\beta}_0(s))\}g(s)ds + o_{(0,U]}(1) \\
=& n^{1/2}\{\mathbf{A}(\hat{\boldsymbol{\beta}}(u)) - \mathbf{A}(\boldsymbol{\beta}_0(u))\} \\
& - \int_0^u \{\mathbf{J}(\boldsymbol{\beta}_0(s))\mathbf{B}(\boldsymbol{\beta}_0(s))^{-1}g(s) + o_{(0,U]}(1)\} \cdot n^{1/2}\{\mathbf{A}(\hat{\boldsymbol{\beta}}(s)) - \mathbf{A}(\boldsymbol{\beta}_0(s))\}ds + o_{(0,U]}(1)
\end{aligned}
$$

Using product integration theory (Andersen, Borgan, Gill, and Keiding 1998), we get

$$
n^{1/2}(\mathbf{A}(\hat{\boldsymbol{\beta}}(u)) - \mathbf{A}(\boldsymbol{\beta}_0(u))) = \boldsymbol{\phi}\{-n^{1/2}\mathbf{S}_n(\boldsymbol{\beta}_0(u), u)\} + o_{(0,U]}(1), \tag{B.4}
$$

where $\boldsymbol{\phi}(\cdot)$ is a linear operator on $\mathcal{F}$, defined as

$$
\boldsymbol{\phi}(\mathbf{w})(u) = \int_0^u \mathcal{I}(s,u)d\mathbf{w}(s), \tag{B.5}
$$

where $\mathcal{F} = \{\mathbf{w} : [0,U] \to R^{p+1}, \mathbf{w}$ is left-continuous with right limit, $\mathbf{w}(0) = 0\}$, and $\mathcal{I}(s,t) = (\boldsymbol{\pi}_{u\in(s,t]}[\mathbf{I}_{p+1} + (\mathbf{B}(\boldsymbol{\beta}_0(u))^{-1})^{\intercal}\mathbf{J}(\boldsymbol{\beta}_0(u))^{\intercal}g(u)du])^{\intercal}$. Using Taylor expansions and the uniform consistency of $\hat{\boldsymbol{\beta}}(u)$ stated in Theorem 1, we then have

$$
n^{1/2}\{\hat{\boldsymbol{\beta}}(u) - \boldsymbol{\beta}_0(u)\} = n^{-1/2}\sum_{i=1}^n \boldsymbol{\zeta}_i(u) + o_{[v,U]}(1),
$$

where $\boldsymbol{\zeta}_i(u) = \boldsymbol{B}(\boldsymbol{\beta}_0(u))^{-1}\boldsymbol{\phi}(\boldsymbol{\xi}_i)$ with $\boldsymbol{\xi}_i(u) = \boldsymbol{X}_i\{N_i(\exp\{\boldsymbol{X}_i^T\boldsymbol{\beta}_0(u)\}) - \int_0^u I(L_i \leq \exp\{\boldsymbol{X}_i^T\boldsymbol{\beta}_0(s)\} \leq R_i)g(s)ds\}$ $(i = 1,\dots,n)$. By Donsker theorem, $n^{1/2}\{\hat{\boldsymbol{\beta}}(u) - \boldsymbol{\beta}_0(u)\}$ converges to a mean zero Gaussian process with the covariance matrix

$$
\boldsymbol{\Sigma}(s,t) = E\{\boldsymbol{\zeta}_1(s)\boldsymbol{\zeta}_1(t)\}. \tag{B.6}
$$

This completes the proof of Theorem 2.

SUPPLEMENT C: JUSTIFICATION FOR THE PROPOSED COVARIANCE ESTIMATION

Given the closed form for the asymptotic covariance in (B.6), the key step to justify the proposed covariance estimation is to prove that $n^{-1/2}\boldsymbol{E}_n(u)\boldsymbol{D}_n^{-1}(u)$ and $n^{-1/2}\tilde{\boldsymbol{E}}_n(u)\boldsymbol{D}_n^{-1}(u)$ provide uniformly consistent estimates for $\boldsymbol{B}(\boldsymbol{\beta}_0(u))$ and $\boldsymbol{J}(\boldsymbol{\beta}_0(u))$ respectively.

Proposition C1: Under the regularity conditions in Theorem 2, we have

$$\lim_{n\to\infty}\sup_{u\in[v,U]}\|\boldsymbol{B}(\boldsymbol{\beta}_0(u)) - n^{-1/2}\boldsymbol{E}_n(u)\boldsymbol{D}_n^{-1}(u)\| = 0,\ \ \lim_{n\to\infty}\sup_{u\in[v,U]}\|\boldsymbol{J}(\boldsymbol{\beta}_0(u)) - n^{-1/2}\tilde{\boldsymbol{E}}_n(u)\boldsymbol{D}_n^{-1}(u)\| = 0.$$

Proof of Proposition C1: Let $\boldsymbol{R}(\boldsymbol{b}) = E[\{\boldsymbol{X}N(e^{\boldsymbol{X}^\top\boldsymbol{b}})\}^{\otimes 2}]$. By Glivenko-Cantelli Theorem, $\lim_{n\to\infty}\sup_{u\in(0,U]}\|\boldsymbol{\Omega}_n(u) - \boldsymbol{R}(\hat{\boldsymbol{\beta}}(u))\| = 0$. Following the lines for (B.1), we can show $\lim_{n\to\infty}\sup_{u\in(0,U]}\|\boldsymbol{R}(\hat{\boldsymbol{\beta}}(u)) - \boldsymbol{R}(\boldsymbol{\beta}_0(u)\| = 0$. Thus, we have $\sup_{u\in(0,U]}\|\boldsymbol{\Omega}_n(u)\| = O_p(1)$ and then $\sup_{u\in(0,U]}\|\boldsymbol{E}_n(u)\| = O_p(1)$. Here $\|\cdot\|$ with a matrix argument means the entrywise Euclidean norm. As a result,

$$\lim_{n\to\infty}\sup_{u\in(0,U]}\|n^{-1/2}\{\boldsymbol{L}_n(\boldsymbol{b}_{n,j}(u)) - \boldsymbol{L}_n(\hat{\boldsymbol{\beta}}(u))\}\| = \lim_{n\to\infty}\sup_{u\in(0,U]}\|n^{-1/2}\boldsymbol{e}_{n,j}(u)\| = 0 \qquad \text{(C.1)}$$

Applications of Glivenko-Cantelli Theorem further give $\lim_{n\to\infty}\sup_{u\in(0,U]}\|n^{-1/2}\boldsymbol{L}_n(\boldsymbol{b}_{n,j}(u)) - \boldsymbol{A}(\boldsymbol{b}_{n,j}(u))\| = 0$ and $\lim_{n\to\infty}\sup_{u\in(0,U]}\|n^{-1/2}\boldsymbol{L}_n(\hat{\boldsymbol{\beta}}(u)) - \boldsymbol{A}(\hat{\boldsymbol{\beta}}(u))\| = 0$. These, coupled with (B.1) and (C.1), imply

$$\sup_{u\in(0,U]}\|\boldsymbol{A}(\boldsymbol{b}_{n,j}(u)) - \boldsymbol{A}(\boldsymbol{\beta}_0(u))\| \to_p 0. \qquad \text{(C.2)}$$

Therefore, we can get (B.2) with $\hat{\boldsymbol{\beta}}(u)$ replaced by $\boldsymbol{b}_{n,j}(u)$. Given the uniform consistency result in Theorem 1, it then follows that

$$\boldsymbol{L}_n(\boldsymbol{b}_{n,j}(u)) - \boldsymbol{L}_n(\hat{\boldsymbol{\beta}}(u)) = \{\boldsymbol{B}(\boldsymbol{\beta}_0(u)) + \epsilon_{n,j}(u)\}\cdot n^{1/2}\{\boldsymbol{b}_{n,j}(u) - \hat{\boldsymbol{\beta}}(u)\},\ \ j = 1,\ldots,p+1. \qquad \text{(C.3)}$$

with $\sup_{u\in[v,U]}\|\epsilon_{n,j}(u)\| \to_{a.s.} 0$. Therefore, $\lim_{n\to\infty}\sup_{u\in(v,U]}\|\boldsymbol{B}(\boldsymbol{\beta}_0(u)) - n^{-1/2}\boldsymbol{E}_n(u)\boldsymbol{D}_n^{-1}(u)\| = 0$.

By mimicking the proof for Theorem 4 of Huang and Peng (2009), given $\sup_{u\in[v,U]}\|\hat{\boldsymbol{\beta}}(u) - \boldsymbol{\beta}_0(u)\| \to_p 0$ and $\sup_{u\in[v,U]}\|\boldsymbol{b}_{n,j}(u) - \boldsymbol{\beta}_0(u)\| \to_p 0$ (implied by (C.2)), we have

$$\sup_{u\in[v,U]}\|\tilde{\boldsymbol{L}}_n(\boldsymbol{b}_n(u)) - \tilde{\boldsymbol{L}}(\boldsymbol{\beta}_0(u)) - n^{1/2}\{\tilde{\boldsymbol{A}}(\boldsymbol{b}_n(u)) - \tilde{\boldsymbol{A}}(\boldsymbol{\beta}_0(u))\}\| \to_p 0,$$

where $\boldsymbol{b}_n(u)$ can be either $\hat{\boldsymbol{\beta}}(u)$ or $\boldsymbol{b}_{n,j}(u)$. Therefore, we get from applying Taylor expansion that

$$\tilde{\boldsymbol{L}}_n(\boldsymbol{b}_{n,j}(u)) - \tilde{\boldsymbol{L}}_n(\hat{\boldsymbol{\beta}}(u)) = \{\boldsymbol{J}(\boldsymbol{\beta}_0(u)) + \epsilon^*_{n,j}(u)\}\cdot n^{1/2}\{\boldsymbol{b}_{n,j}(u) - \hat{\boldsymbol{\beta}}(u)\},\ \ j = 1,\ldots,p+1,$$

where $\sup_{u\in[v,U]} \|\epsilon^*_{n,j}(u)\| \to_{a.s.} 0$. This shows that $\lim_{n\to\infty} \sup_{u\in[v,U]} \|\boldsymbol{J}(\boldsymbol{\beta}_0(u)) - n^{-1/2}\tilde{\boldsymbol{E}}_n(u)$ $\boldsymbol{D}_n^{-1}(u)\| = 0$.

REFERENCES

Andersen, P. K., Borgan, e., Gill, R. D., and Keiding, N. (1998), *Statistical Models Based on Counting Processes*, New York: Springer-Verlag., 2nd ed.

Huang, Y. and Peng, L. (2009), "Accelerated Recurrence Time Models," *Scandinavian Journal of Statistics*, 36, 636–648.

Peng, L. and Huang, Y. (2008), "Survival Analysis with Quantile Regression Models," *Journal of American Statistical Association*, 103, 637–649.

van der Vaart, A. W. and Wellner, J. (2000), *Weal Convergence and Emprical Processes with Applications to Statistics*, Springer-Verlag, 2nd ed.