Yunming Mu[1] and Ying Wei[2]

[1] *Portland State University and* [2]*Columbia University*

Supplementary Material

This supplement contains technical details in proving Theorem 1.

Lemma S1. Denote $\theta = (\lambda, \beta, x, t, c)$ and let $\Theta = \Omega_\tau \bigotimes \mathcal{B} \bigotimes \mathcal{X} \bigotimes [0,1]^2$ be the product parameter space for θ. If Assumptions 2-8 hold and $h_n \to 0$ and $nh_n \to \infty$, the following expansions hold

$$\sup_{\lambda \in \Omega_\tau, x \in \mathcal{X}, t, c \in [0,1]} \| R_n(x, \lambda, \tilde{\beta}_n(\lambda; t)) - R_n(x, \lambda, \beta(\lambda, t); t) \| = o_p(1), \tag{S1.1}$$

and

$$\sup_{\theta \in \Theta} R_n(x, \lambda, \beta; t) = O_p(1). \tag{S1.2}$$

Proof.

We first establish the following uniform expansion

$$\sup_{(\lambda,\beta,x,t,c)\in\Theta} |R_n(x,\lambda,\beta;t) - R_n(x,\lambda,\beta(\lambda,t);t) - E[R_n(x,\lambda,\beta;t)] + E[R_n(x,\lambda,\beta(\lambda,t);t)]| = o_p(1). \tag{S1.3}$$

For a fixed $\theta = (\lambda, \beta, x, t, c)$, denote $Z_i = \{(X_{ij}, Y_{ij}), j = 1, \ldots . n_i\}$ and

$$\psi_i(z_i, \theta) = \sum_{j=1}^{n_i} \mathrm{I}\{t_{ij} \in I_{n,c}(t)\} \mathrm{I}\{x_{ij} \le x\} \left[\tau - \mathrm{I}\{y_{ij}^{(\lambda)} - x_{ij}^T \beta \le 0\} \right].$$

Define $u(z_i, \theta, d) = \sup_{|\theta_1 - \theta_2| \le d} |\psi_i(z_i, \theta_1) - \psi_i(z_i, \theta_2)|$, where $|\cdot|$ is taken to be the sup norm of vectors. With some work one can show

$$\begin{aligned} u^2(z_i, \theta, d) &= \{ \sup_{|\theta_1 - \theta| \le d} |\psi_i(z_i, \theta_1) - \psi_i(z_i, \theta_2)| \}^2 \\ &\le n_i \sum_{j=1}^{n_i} \left\{ \sup_{|\theta_1 - \theta| \le d} \left| \mathrm{I}\{t_{ij} \in I_{n,c_1}(t)\} \mathrm{I}\{x_{ij} \le x_1\} \left[\tau - \mathrm{I}\{y_{ij}^{(\lambda_1)} - x_{ij}^T \beta_1 \le 0\} \right] \right. \right. \\ &\quad \left. \left. - \mathrm{I}\{t_{ij} \in I_{n,c}(t)\} \mathrm{I}\{x_{ij} \le x\} \left[\tau - \mathrm{I}\{y_{ij}^{(\lambda)} - x_{ij}^T \beta \le 0\} \right] \right| \right\}^2 . \end{aligned} \tag{S1.4}$$

Observing (S1.4), (S1.3) can be proven similarly as Lemma 1 of Mu (2005) under Assumptions 2-8. Applying a Taylor expansion, it is easy to show, for any $\|\beta - \beta(\lambda, t)\| = o(1)$,

$$E[R_n(x, \lambda, \beta)] = E[R_n(x, \lambda, \beta(\lambda, t))](1 + o(1)).$$

Substitute this into (S1.3), we obtain

$$R_n(x, \lambda, \beta; t) - R_n(x, \lambda, \beta(\lambda, t); t) = o_p(1), \tag{S1.5}$$

where the remainder term in (S1.5) is uniform in (λ, x, t, c) and $\beta : ||\beta - \beta(\lambda, t)|| = o(1)$. Thus to verify (S1.1), it suffices to show $\|\tilde{\beta}_n(\lambda; t) - \beta(\lambda, t)\| = o_p(1)$. For this purpose, define $\beta_n^*(\lambda, t) = \text{argmin}_b E\{h_n^{-1}\text{I}\{T \in I_{n,c}(t)\} \cdot \rho_\tau(Y(T)^{(\lambda)} - X(T)^T b)\}$. Following similar argument of Lemma 4 of Mu (2005), we can demonstrate under Assumptions 2-8 that

$$\sup_{\lambda \in \Omega_\tau, t, c \in [0,1]} \left\| \tilde{\beta}_n(\lambda; t) - \beta_n^*(\lambda, t) \right\| = O_p((nh_n)^{-1/2}(\log(nh_n))^{1/2}).$$

Denote $\varphi(\lambda, b, t) = E\{\rho_\tau(Y(T)^{(\lambda)} - X(T)^T b)|T = t\}$ and recall $\beta(\lambda, t) = \text{argmin}_b \varphi(\lambda, b, t)$. Note that

$$E\{h_n^{-1}\text{I}\{T \in I_{n,c}(t)\} \cdot \rho_\tau(Y(T)^{(\lambda)} - X(T)^T b)\} = E\{h_n^{-1}\text{I}\{T \in I_{n,c}(t)\}\varphi(\lambda, b, T)\},$$

and by Taylor's expansion

$$\varphi(\lambda, b, t') = \varphi(\lambda, b, t) + O(h_n) \quad \text{whenever } |t' - t| \le h_n.$$

Hence

$$\begin{aligned} E\{h_n^{-1}\text{I}\{T \in I_{n,c}(t)\} \cdot \rho_\tau(Y(T)^{(\lambda)} - X(T)^T b)\} &= h_n^{-1}\varphi(\lambda, b, t)P(T \in I_{n,c}(t))(1 + O(h_n)) \\ &= \varphi(\lambda, b, t)g_T(t)(1 + O(h_n)). \end{aligned} \tag{S1.6}$$

It can easily be proved by some elementary arguments that the minimizer $\beta_n^*(\lambda, t)$ of $E\{\text{I}\{T \in I_{n,c}(t)\} \cdot \rho_\tau(Y(T)^{(\lambda)} - X(T)^T b)\}$ approaches the minimizer $\beta(\lambda, t)$ of the right-hand side of (S1.6):

$$\sup_{\lambda \in \Omega_\tau, t, c \in [0,1]} \|\tilde{\beta}_n(\lambda; t) - \beta(\lambda, t)\| = O(h_n). \tag{S1.7}$$

(S1.5) and (S1.7) proves (S1.1). To verify (S1.2), it suffices to prove $E\left\{\sup_{\theta\in\Theta} R_n(x,\lambda,\beta;t)\right\} = O(1)$. But $\sup_{\theta\in\Theta} R_n(x,\lambda,\beta;t) = \sup_{\theta\in\Theta}\left|E\left\{R_n(x,\lambda,\beta;t)\right\}\right| + o_p(1) = O_p(1)$ as a consequence of (S1.3). This completes the proof of Lemma S1.

Lemma S2. Under the same assumptions for Lemma S1 and Assumption 10, if $h_n \to 0$ and $nh_n \to \infty$, then (S1.21) and (S1.22) hold.

Proof. As an immediate result of (S1.3), the following expansion holds

$$R_n(x,\lambda,\beta;t) - E[R_n(x,\lambda,\beta(\lambda,t);t)] = o_p(1), \tag{S1.8}$$

where the remainder term in (S1.8) is uniform in (λ, x, t, c) and $\beta : ||\beta - \beta(\lambda,t)|| = o(1)$. As a consequence of (S1.7) we have

$$R_n(x,\lambda,\tilde{\beta}_n(\lambda;t);t) - E[R_n(x,\lambda,\beta(\lambda,t);t)] = o_p(1), \tag{S1.9}$$

Recall $\phi(x,\lambda,\beta,t_{ij}) = E\left\{\mathrm{I}\{x_{ij} \le x\}\left[\tau - F(x_{ij}^T(\beta - \beta(\lambda,t_{ij})); t_{ij}, x_{ij}, \lambda)\right] | t_{ij}\right\}$. Due to Assumption 10, $\phi(x,\lambda,\beta,t_{ij}) = \phi(\lambda,\beta,x,t) + O(h_n)$ for any $t_{ij} \in I_{n,c}(t)$. Thus

$$\begin{aligned}
E[R_n(x,\lambda,\beta;t)] &= \frac{1}{nh_n}\sum_{i,j} E\left\{\mathrm{I}\{t_{ij} \in I_{n,c}(t)\}\mathrm{I}\{x_{ij} \le x\}\left[\tau - \mathrm{I}\{y_{ij}^{(\lambda)} - x_{ij}^T\beta \le 0\}\right]\right\} \\
&= \frac{1}{nh_n}\sum_{i,j} E\left\{\mathrm{I}\{t_{ij} \in I_{n,c}(t)\}\phi(\lambda,\beta,x,t_{ij})\right\} \\
&= \frac{1}{nh_n}\sum_{i,j} E\left\{\mathrm{I}\{t_{ij} \in I_{n,c}c(t)\}\phi(\lambda,\beta,x,t) + O(h_n)\right\} \\
&= d_T(t)\phi(x,\lambda,\beta,t)(1 + O(h_n))) + O(h_n),
\end{aligned} \tag{S1.10}$$

where Assumption 5 has been applied in the last step of (S1.10). From (S1.9), we obtain

$$\begin{aligned}
V_n^0(\lambda;t) &= \frac{1}{nh_n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\mathrm{I}\{t_{ij} \in I_{n,c}(t)\}\cdot\left\{E\left[R_n(x_{ij},\lambda,\beta(\lambda;t);t)\right] + o_p(1)\right\}^2. \\
&= \frac{1}{nh_n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\mathrm{I}\{t_{ij} \in I_{n,c}(t)\}\cdot\left\{E\left[R_n(x_{ij},\lambda,\beta(\lambda;t);t)\right]\right\}^2 + o_p(1).
\end{aligned} \tag{S1.11}$$

We insert (S1.10) into (S1.11) to obtain

$$V_n^0(\lambda;t) = \frac{1}{nh_n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} \mathrm{I}\{T_{ij}\in I_{n,c}(t)\}\cdot d_T^2(t)\phi^2(X_{ij},\lambda,\beta(\lambda,t),t)(1+O(h_n)) + O_p(h_n) + o_p(1)$$

$$= \frac{d_T^2(t)}{nh_n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} \mathrm{I}\{T_{ij}\in I_{n,c}(t)\}\cdot \phi^2(X_{ij},\lambda,\beta(\lambda,t),t)(1+O(h_n)) + o_p(1) \qquad \text{(S1.12)}$$

Due to Assumption 10, by similar arguments of Lemma S1, we can demonstrate that

$$\frac{1}{nh_n}\sum_{i,j} \mathrm{I}\{T_{ij}\in I_{n,c}(t)\}\cdot \phi^2(X_{ij},\lambda,\beta(\lambda,t),t)$$
$$-\frac{1}{nh_n}\sum_{i,j} E\left\{\mathrm{I}\{T_{ij}\in I_{n,c}(t)\}\cdot \phi^2(X_{ij},\lambda,\beta(\lambda,t),t)\right\} = o_p(1), \qquad \text{(S1.13)}$$

where the remainder term is uniform in (λ, t). To avoid repetition, we skip a proof here. (S1.12) and (S1.13) implies

$$V_n^0(\lambda;t) = \frac{d_T^2(t)}{nh_n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\mathrm{I}\{T_{ij}\in I_{n,c}(t)\}\cdot \phi^2(X_{ij},\lambda,\beta(\lambda,t),t)\right\}(1+O(h_n))$$
$$+ o_p(h_n) + o_p(1). \qquad \text{(S1.14)}$$

Now consider

$$E\left\{\mathrm{I}\{T_{ij}\in I_{n,c}(t)\}\cdot \phi^2(X_{ij},\lambda,\beta(\lambda,t),t)\right\}$$
$$= E\left\{\mathrm{I}\{T_{ij}\in I_{n,c}(t)\}\cdot E\left[\phi^2(X_i(T),\lambda,\beta(\lambda,t),t)|T=t_{ij}\right]\right\} \qquad \text{(S1.15)}$$

Using a Taylor expansion, it is easy to show $E\left[\phi^2(X(T),\lambda,\beta(\lambda,t),t)|T=t_{ij}\right] = E\left[\phi^2(X(T),\lambda,\beta(\lambda,t),t)|T\right.$ $o_p(1)$. This, (S1.14) and (S1.15) implies

$$V_n^0(\lambda;t) = d_T^3(t)E\left[\phi^2(X(T),\lambda,\beta(\lambda,t),t)|T=t\right](1+O(h_n))$$

Denote $V_\tau(\lambda,t) = d_T^3(t)E\left[\phi^2(X(T),\lambda,\beta(\lambda,t),t)|T=t\right]$ and we have shown (S1.21). Let $X_1(t)$ and $X_2(t)$ denote two independent realizations from the process $X(t)$. Note that

$$V_\tau(\lambda,t) \;=\; d_T^3(t)E\left\{E^2\left[\mathrm{I}\{X_2(t)\le X_1(T)\}\left(\tau - F(0;t,X_2^T(t),\lambda)\right)|X_1(T)\right]|T=t\right\}.$$

First $V_\tau(\lambda,t)$ is continuous at all $(t,\lambda)\in[0,1]\bigotimes\Omega_\tau$ due to Assumption 9. Under Assumptions 3 and 4, the identifiability conditions of Mu & He (2007) are satisfied for $\lambda_\tau(t)$ at every

$t \in [0,1]$. Note that $V_\tau(\lambda, t) \geq 0$ at all t and λ, and $V(\lambda_\tau(t), t) = 0$ almost surely in t. Thus $\lambda_\tau(t)$ minimizes $V_\tau(\lambda, t)$ almost surely in $t \in [0,1]$. Now we demonstrate the uniqueness of $\lambda_\tau(t)$. Suppose that $\lambda^*(t) \neq \lambda_\tau(t)$ also minimizes $V_\tau(\lambda, t)$, then we must have $V_\tau(\lambda^*(t), t) = 0$ almost surely in t. As a consequence, $E\left\{E^2\left[\mathrm{I}\{X_2(t) \leq X_1(t)\}\left(\tau - F(0; t, X_2^T(t), \lambda^*(t))\right)|X_1(t)\right]\right\} = 0$ almost surely in $t \in [0,1]$. This further implies that $E\left[\mathrm{I}\{X_2(t) \leq x\}\left(\tau - F(0; t, X_2^T(t), \lambda^*(t))\right)\right] = 0$, or, $\tau = F(0; t, x, \lambda^*(t))$ almost surely in x and t. But this would contradict Assumption the identifiability of $\lambda_\tau(t)$. This proves the uniqueness of $\lambda_\tau(t)$ and completes the proof of Lemma S2.

Lemma S3. Let $Q_n(\theta, t)$ be a random real-valued function with a parameters $\theta \in \Theta \subseteq \mathcal{R}$ and $t \in \mathcal{T} \subseteq \mathcal{R}$, and $Q_n(\theta, t)$ converges to a non-stochastic function $Q(\theta, t)$ for each $t \in \mathcal{T}$. Denote $\theta_0(t) = \mathrm{argmin}_{\theta\in\Theta} Q(\theta, t)$ and $\hat{\theta}_n(t) = \mathrm{argmin}_{\theta\in\Theta} Q_n(\theta, t)$. Assume the following assumptions

C1. The parameter space $\Theta \bigotimes \mathcal{T}$ is a compact subset of $\mathcal{R}^2$.

C2. $Q(\theta, t)$ attains a unique global minimum at $\theta_0(t)$ for all $t \in \mathcal{T}$.

C3. $Q(\theta, t)$ is continuous at every $(\theta, t) \in \Theta \bigotimes \mathcal{T}$.

C4. $Q_n(\theta, t)$ converges in probability to $Q(\theta, t)$ uniformly in $\theta \in \Theta$ and in $t \in \mathcal{T}$ as $n \to \infty$.

Under assumptions C1-C4, we have

$$\sup_{t\in\mathcal{T}} |\hat{\theta}(t) - \theta_0(t)| = o_p(1)$$

Proof. For any $\delta > 0$, denote $\mathcal{N}_t = \{\theta; |\theta - \theta_0(t)| < \delta\}$, and $\mathcal{N}_t^c$ be the complement of $\mathcal{N}_t$, $\mathcal{N}_t^c = \mathcal{R} - \mathcal{N}_t$. Then $\Theta \cap \mathcal{N}_t^c$ is compact, so that $\min_{\theta\in\Theta\cap\mathcal{N}_t^c} Q(\theta, t)$ exists. The minimum of a continuous function always exist on a compact set. Denote $\varepsilon_\delta(t) = Q(\theta_0(t), t) - \min_{\theta\in\Theta\cap\mathcal{N}_t^c} Q(\theta, t)$. Assumption C2 implies that $\varepsilon_\delta(t) > 0$ for all $t \in \mathcal{T}$. Then Assumption C3 guarantees that there exists a constant $\epsilon_\delta > 0$ such that $\min_{t\in\mathcal{T}} \varepsilon_\delta(t) = \epsilon_\delta > 0$. Let E_n

be the event

$$|Q_n(\theta, t) - Q(\theta, t)| < \frac{1}{3}\epsilon_\delta \,\text{for all}\, \theta \in \Theta, t \in [0, 1]$$

Then

$$E_n \Rightarrow Q(\hat{\theta}(t), t) < Q_n(\hat{\theta}(t), t) + \frac{1}{3}\epsilon_\delta \tag{S1.16}$$

and

$$E_n \Rightarrow Q_n(\theta_0(t), t) < Q(\theta_0(t), t) + \frac{1}{3}\epsilon_\delta \tag{S1.17}$$

But

$$Q_n(\hat{\theta}(t), t) = \min_{\theta \in \Theta} Q_n(\theta, t) \le Q_n(\theta_0(t), t), \tag{S1.18}$$

and we can use (S1.18) to rewrite (S1.16) as,

$$E_n \Rightarrow Q(\hat{\theta}(t), t) < Q_n(\theta_0(t), t) + \frac{1}{3}\epsilon_\delta. \tag{S1.19}$$

Combine (S1.17) and (S1.19) to get

$$E_n \Rightarrow Q(\hat{\theta}(t), t) < Q(\theta_0(t), t) + \frac{2}{3}\epsilon_\delta.$$

This and our definition of ϵ_δ implies

$$E_n \Rightarrow \hat{\theta}(t) \in \mathcal{N}_t \,\text{for all}\, t \in [0, 1],$$

which in turn implies

$$E_n \Rightarrow \sup_{t \in [0,1]} |\hat{\theta}(t) - \theta_0(t)| < \delta,$$

so that $P(E_n) \le P(\sup_{t \in [0,1]} |\hat{\theta}(t) - \theta_0(t)| < \delta)$. However Assumption C4 implies that $\lim_{n\to\infty} P(E_n) = 1$, so that we have

$$1 \ge \lim_{n\to\infty} P(\sup_{t \in [0,1]} |\hat{\theta}(t) - \theta_0(t)| < \delta) \ge \lim_{n\to\infty} P(E_n) = 1$$

Proof of Theorem 1: We sketch the proof here. Major steps of the argument include

$$(i) \sup_{\lambda \in \Omega_\tau, t, c \in [0,1]} |V_n(\lambda; t) - V_n^0(\lambda; t)| = o_p(1), \tag{S1.20}$$

(ii)There exists a deterministic function $V_\tau(\lambda, t)$ such that

$$\sup_{\lambda \in \Omega_\tau, t, c \in [0,1]} |V_n^0(\lambda; t) - V_\tau(\lambda, t)]| = o_p(1), \tag{S1.21}$$

$(iii) V_\tau(\lambda, t)$ is uniquely minimized at $\lambda_\tau(t)$ for every $t \in [0, 1]$. (S1.22)

Note that (S1.20) is a direct result of Lemma S1. We have shown (S1.21) and (S1.22) in Lemma S2. By Lemma S3, $\tilde{\lambda}_n(t)$ converges in probability to $\lambda_\tau(t)$ uniformly in $t \in [0,1]$. Thus we have shown that if $h_n \to 0$ and $nh_n \to \infty$

$$\sup_{t\in[0,1]} |\tilde{\lambda}_n(t) - \lambda_\tau(t)| = o_p(1).$$

To verify the uniform consistency of $\hat{\lambda}_n(t)$, we notice that the smoothing spline estimator $\hat{\lambda}_n(t)$ is linear in the observations $\tilde{\lambda}_n(t_l)$, in the sense that there exists a weight function $H(s,t;\gamma)$ such that

$$\hat{\lambda}_n(t) = k_n^{-1} \sum_{l=1}^{k_n} H(t,t_l;\gamma) \cdot \tilde{\lambda}_n(t_l) = k_n^{-1} \sum_{l=1}^{k_n} H(t,t_l;\gamma) \cdot (\lambda_\tau(t_l) + o_p(1)).$$

Under suitable restrictions on the rate that γ converges to zero and the smoothness assumption of $\lambda_\tau(t)$, Lemma 6.1 of Nychika (1995) shows that $k_n^{-1}\sum_{l=1}^{k_n} H(t,t_l;\gamma)\cdot(\lambda_\tau(t_l)+o_p(1)) = \lambda_\tau(t_l) + o_p(1)$ uniformly in $t \in [0,1]$. The proof of (1) in Theorem 1 is complete.

Now we prove (2) in Theorem 1. The result follows from the continuity of quantiles as a set-valued solution and the consistency property of the coefficient functions assuming $\lambda_\tau(t)$ as known. Let $\Lambda_\lambda(y)$ denote the first derivative of $\Lambda(y) = \frac{y^\lambda - 1}{\lambda}$ with respect to λ, then $\Lambda_\lambda(y) = \frac{y^\lambda \ln y}{\lambda} - \frac{y^\lambda}{\lambda^2} + \frac{1}{\lambda^2}$. Due to the boundedness of X and the robustness property of quantiles, it suffices to consider y in a compact set, i.e., $c \le y \le C$. Note that $\Lambda_\lambda(y)$ is continuous and hence bounded on $\Omega_\tau \bigotimes [c,C]$. This implies that $y_i^{(\hat{\lambda}_n(t_i))} = y_i^{(\lambda_\tau(t_i))} + o_p(1)$, where $o_p(1)$ is independent of y_i and t_i. Let $\check{\beta}_{n,j}(t)$ denote the B-spline estimator of $\beta_{\tau,j}(t)$ assuming the true transformation function $\lambda_\tau(t)$ is given. With a slight modification of the arguments in Kim (2007), one can demonstrate the consistency of $\check{\beta}_{n,k}(\cdot)$ in the case of longitudinal data under our stated assumptions, i.e.,

$$\frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\check{\beta}_{n,k}(t_{ij}) - \beta_{\tau,k}(t_{ij}))^2 = o_p(1), \quad k = 1, \cdots, p.$$

To save space, we do not present a proof here. Then (2) in Theorem 1 is a consequence of the continuity of quantile estimator. We refer to Portnoy S. and Mizera I. (1998) for a discussion of continuity of LAD estimator as set-valued solutions on nonsingular designs.