Supplementary to "Copula and composite quantile regression based estimating equations for longitudinal data"

Kangning Wang[1]*and Wen Shan[1]
[1]School of Statistics, Shandong Technology and Business University, No. 191, Binhai Middle Road, Laishan District, Yantai 264005, China

To establish the asymptotic properties of the proposed estimators, the following regularity conditions are needed in this paper.

(C1) There exists a unique solution $\boldsymbol{\theta}_0 \in \Theta$ to the equation $E_{\boldsymbol{u}}[\frac{\partial \log[c(u_1,\cdots,u_m;\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}}] = \mathbf{0}$, where $\boldsymbol{u} = (u_1, \cdots, u_m)$, and Θ is a convex and compact set.

(C2) The integral $\int_{\boldsymbol{u}} \sup_{\boldsymbol{\theta}\in\Theta} \|\frac{\partial \log[c(u_1,\cdots,u_m;\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}}\| d\boldsymbol{u}$ exists.

(C3) The partial derivative $sup_{1/(\kappa_n+1)\leq \boldsymbol{u}\leq \kappa_n/(\kappa_n+1)} \|\frac{\partial \log[c(u_1,\cdots,u_m;\boldsymbol{\theta}_0)]}{\partial \boldsymbol{u}}\| \leq Q\kappa_n^a$, where Q and a are some positive constants.

(C4) The conditional density function of T, $f_{\boldsymbol{X},\boldsymbol{Z}}(t)$, satisfies that $0 < c_1 < f_{\boldsymbol{X},\boldsymbol{Z}}(T) < c_2 < \infty$ uniformly in $\boldsymbol{X}$, $\boldsymbol{Z}$ and T for some positive constants c_1 and c_2. Furthermore, there is a constant M such that the conditional densities satisfy the Lipschitz condition $|f_j(s|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij}) - f_j(t|\boldsymbol{X}_{ij}, \boldsymbol{Z}_{ij})| \leq M|s-t|$, for all i and j.

(C5) The covariates $\boldsymbol{X}_{ij}$ and $\boldsymbol{Z}_{ij}$ satisfy that $\max_{1\leq i\leq n, 1\leq j\leq m} \|\boldsymbol{X}_{ij}\| = O_p(n^v)$ and $\max_{1\leq i\leq n, 1\leq j\leq m} \|\boldsymbol{Z}_{ij}\| = O_p(n^v)$ for some positive constant $v < \frac{1}{6}$.

*Corresponding author. E-mail addresses: wkn1986@126.com (Kangning Wang), shanwen2019@126.com (Wen Shan). The research was supported by NNSF project (11901356), wealth management project (2019ZBKY047) of Shandong Technology and Business University. The online version of this article contains supplementary material.

(C6) $\alpha_{0k}(t) \in \mathcal{H}_r$ for some $r > 1/2$. Where $\mathcal{H}_r$ is the collection of all functions on $[0,1]$ whose mth order derivative satisfies the Hölder condition of order ν with $r \equiv m + \nu$.

(C7) For any positive definite matrix $\boldsymbol{\Upsilon}_{i(k)}$, $\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\tilde{\boldsymbol{X}}_{i(k)}^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{\Upsilon}_{i(k)}\boldsymbol{\Lambda}_{i(k)}\tilde{\boldsymbol{X}}_{i(k)}$ converges to a positive definite matrix.

(C8) Matrix $\boldsymbol{\Omega}$ is positive definite and $\|\boldsymbol{\Omega}\| = O\left(\frac{1}{n}\right)$.

Conditions (C1) and (C4), together with (C5), are commonly used in quantile regression literature. Conditions (C2) and (C3) specify conditions on the copula function, which are satisfied by many commonly used copulas including Gaussian copula, t-copula and so on. Condition (C6) states the smoothness condition on the coefficient functions, which describes a requirement on the best convergence rate that the coefficient functions can be estimated. Conditions (C7)-(C8) are used to represent the asymptotic covariance matrix and obtain the optimal convergence rate.

Lemma 1. Under conditions (C1)-(C5), if $\kappa_n^{2a+1}/n^{1-2v} \to 0$ and $\kappa_n \to \infty$ as $n \to \infty$, then $\hat{\boldsymbol{\theta}} \to_p \boldsymbol{\theta}_0$, where $\to_p$ denotes convergence in probability.

Proof of Lemma 1. The proof is similar with the proof of Lemma 1 in Wang et al. (2018), we omit the details for saving space.

Proof of Theorem 1. Let

$$\bar{\boldsymbol{U}}(\boldsymbol{\zeta}) = \sum_{i=1}^{n}\sum_{k=1}^{K}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{P}_{i(k)}(\boldsymbol{\zeta})$$

with $\boldsymbol{P}_{i(k)}(\boldsymbol{\zeta}) = \left(\tau_k - \Pr(Y_{i1} - \boldsymbol{D}_{i1}\boldsymbol{\zeta} < \hat{b}_k), \cdots, \tau_k - \Pr(Y_{im} - \boldsymbol{D}_{im}\boldsymbol{\zeta} < \hat{b}_k)\right)^T$. By Lemma 5.1 and Hendricks and Koenker (1992), we can obtain that $\boldsymbol{V}_{i(k)}(\hat{\boldsymbol{\theta}}) \to_p \boldsymbol{V}_{i(k)}(\boldsymbol{\theta}_0)$ and $\hat{\boldsymbol{\Lambda}}_{i(k)} \to_p \boldsymbol{\Lambda}_{i(k)}$. Hence, we can directly consider the estimation equation $\hat{\boldsymbol{U}}(\boldsymbol{\zeta}) = \sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}) = \boldsymbol{0}$. By some direct calculation, we can get that

$$\begin{aligned}
&\frac{1}{n}\left[\hat{\boldsymbol{U}}(\boldsymbol{\zeta})-\bar{\boldsymbol{U}}(\boldsymbol{\zeta})\right] \\
&=\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\left[\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta})-\boldsymbol{P}_{i(k)}(\boldsymbol{\zeta})\right] \\
&=\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\begin{bmatrix}\Pr(Y_{i1}-\boldsymbol{D}_{i1}\boldsymbol{\zeta}<\hat{b}_k)-I(Y_{i1}-\boldsymbol{D}_{i1}\boldsymbol{\zeta}<\hat{b}_k)\\ \vdots \\ \Pr(Y_{im}-\boldsymbol{D}_{im}\boldsymbol{\zeta}<\hat{b}_k)-I(Y_{im}-\boldsymbol{D}_{im}\boldsymbol{\zeta}<\hat{b}_k)\end{bmatrix} \\
&=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{K}\boldsymbol{a}_{ij(k)}\left[\Pr(Y_{ij}-\boldsymbol{D}_{ij}\boldsymbol{\zeta}<\hat{b}_k)-I(Y_{ij}-\boldsymbol{D}_{ij}\boldsymbol{\zeta}<\hat{b}_k)\right],
\end{aligned}$$

where $\boldsymbol{a}_{ij(k)}$ is a vector with $\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)=(\boldsymbol{a}_{i1(k)},\cdots,\boldsymbol{a}_{im(k)})$. Under condition (C5) and from the law of large numbers (Pollard 1990), we have that

$$\sup_{\boldsymbol{\zeta}}\left|\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{K}\boldsymbol{a}_{ij(k)}\left[\Pr(Y_{ij}-\boldsymbol{D}_{ij}\boldsymbol{\zeta}<\hat{b}_k)-I(Y_{ij}-\boldsymbol{D}_{ij}\boldsymbol{\zeta}<\hat{b}_k)\right]\right|=o\left(\frac{1}{\sqrt{n}}\right).$$

Therefore, $\sup_{\boldsymbol{\zeta}}\left\|\frac{1}{n}\left[\bar{\boldsymbol{U}}(\boldsymbol{\zeta})-\hat{\boldsymbol{U}}(\boldsymbol{\zeta})\right]\right\|=o\left(\frac{1}{\sqrt{n}}\right)$. Hence $\hat{\boldsymbol{\zeta}}^o\to_p\boldsymbol{\zeta}_0$ as $n\to\infty$. For any $\boldsymbol{\zeta}$ satisfying $\|\boldsymbol{\zeta}-\boldsymbol{\zeta}_0\|=O(n^{-1/3})$,

$$\begin{aligned}
&\hat{\boldsymbol{U}}(\boldsymbol{\zeta})-\hat{\boldsymbol{U}}(\boldsymbol{\zeta}_0) \\
&=\sum_{i=1}^{n}\sum_{k=1}^{K}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\left[\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta})-\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}_0)\right] \\
&=\left\{\sum_{i=1}^{n}\sum_{k=1}^{K}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{P}_{i(k)}(\boldsymbol{\zeta})+\sum_{i=1}^{n}\sum_{k=1}^{K}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\left[\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta})-\hat{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta}_0)-\boldsymbol{P}_{i(k)}(\boldsymbol{\zeta})\right]\right\} \\
&=\sum_{i=1}^{n}\sum_{k=1}^{K}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{P}_{i(k)}(\boldsymbol{\zeta}) \\
&\quad+\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{K}\boldsymbol{a}_{ij(k)}\left[\Pr(Y_{ij}-\boldsymbol{D}_{ij}\boldsymbol{\zeta}<\hat{b}_k)-I(Y_{ij}-\boldsymbol{D}_{ij}\boldsymbol{\zeta}<\hat{b}_k)+I(Y_{ij}-\boldsymbol{D}_{ij}\boldsymbol{\zeta}_0<\hat{b}_k)-\tau_k\right].
\end{aligned}$$

According to Lemma 3 of Jung (1996), we have that

$$\sup \left| \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{K} \boldsymbol{a}_{ij(k)} \left[\Pr(Y_{ij} - \boldsymbol{D}_{ij}\boldsymbol{\zeta} < \hat{b}_k) - I(Y_{ij} - \boldsymbol{D}_{ij}\boldsymbol{\zeta} < \hat{b}_k) + I(Y_{ij} - \boldsymbol{D}_{ij}\boldsymbol{\zeta}_0 < \hat{b}_k) - \tau_k \right] \right|$$
$$= o_p(\sqrt{n}).$$

Therefore, $\hat{\boldsymbol{U}}(\boldsymbol{\zeta}) - \hat{\boldsymbol{U}}(\boldsymbol{\zeta}_0) = \bar{\boldsymbol{U}}(\boldsymbol{\zeta}) + o_p(\sqrt{n})$. By Taylor's expansion of $\bar{\boldsymbol{U}}(\boldsymbol{\zeta})$ together with $\bar{\boldsymbol{U}}(\boldsymbol{\zeta}_0) = \boldsymbol{0}$, we can obtain $\hat{\boldsymbol{U}}_0(\boldsymbol{\zeta}) - \hat{\boldsymbol{U}}_0(\boldsymbol{\zeta}_0) = \boldsymbol{D}(\boldsymbol{\zeta}_0)(\boldsymbol{\zeta} - \boldsymbol{\zeta}_0) + o_p(\sqrt{n})$, where

$$\boldsymbol{D}(\boldsymbol{\zeta}_0) = \frac{\partial \bar{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0} = -\sum_{i=1}^{n} \sum_{k=1}^{K} \boldsymbol{D}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{D}_i.$$

Note that $\hat{\boldsymbol{\zeta}}^o$ is in the $n^{-1/3}$ neighborhood of $\boldsymbol{\zeta}_0$ and $\hat{\boldsymbol{U}}_0(\hat{\boldsymbol{\zeta}}^o) = \boldsymbol{0}$, we have $\hat{\boldsymbol{\zeta}}^o - \boldsymbol{\zeta}_0 = -\boldsymbol{D}^{-1}(\boldsymbol{\zeta}_0)\hat{\boldsymbol{U}}(\boldsymbol{\zeta}_0) + o_p(\sqrt{n})$. To obtain the closed form expression of $\hat{\boldsymbol{\beta}}^o$, we write the inverse of $\boldsymbol{D}(\boldsymbol{\zeta}_0)$ as the following block form

$$\begin{aligned}
&\boldsymbol{D}^{-1}(\boldsymbol{\zeta}_0) \\
&= \begin{bmatrix} \sum_{i=1}^{n} \sum_{k=1}^{K} \boldsymbol{X}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{X}_i & \sum_{i=1}^{n} \sum_{k=1}^{K} \boldsymbol{X}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{B}_i \\ \sum_{i=1}^{n} \sum_{k=1}^{K} \boldsymbol{B}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{X}_i & \sum_{i=1}^{n} \sum_{k=1}^{K} \boldsymbol{B}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{B}_i \end{bmatrix}^{-1} \\
&= \begin{bmatrix} \boldsymbol{D}_{\boldsymbol{XX}} & \boldsymbol{D}_{\boldsymbol{XB}} \\ \boldsymbol{D}_{\boldsymbol{BX}} & \boldsymbol{D}_{\boldsymbol{BB}} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{D}^{11} & \boldsymbol{D}^{12} \\ \boldsymbol{D}^{21} & \boldsymbol{D}^{22} \end{bmatrix},
\end{aligned}$$

where $\boldsymbol{D}^{11} = \left(\boldsymbol{D}_{\boldsymbol{XX}} - \boldsymbol{D}_{\boldsymbol{XB}} \boldsymbol{D}_{\boldsymbol{BB}}^{-1} \boldsymbol{D}_{\boldsymbol{BX}}\right)^{-1}$, $\boldsymbol{D}^{22} = \left(\boldsymbol{D}_{\boldsymbol{BB}} - \boldsymbol{D}_{\boldsymbol{BX}} \boldsymbol{D}_{\boldsymbol{XX}}^{-1} \boldsymbol{D}_{\boldsymbol{XB}}\right)^{-1}$, $\boldsymbol{D}^{12} = -\boldsymbol{D}^{11} \boldsymbol{D}_{\boldsymbol{XB}} \boldsymbol{D}_{\boldsymbol{BB}}^{-1}$ and $\boldsymbol{D}^{21} = -\boldsymbol{D}^{22} \boldsymbol{D}_{\boldsymbol{BX}} \boldsymbol{D}_{\boldsymbol{XX}}^{-1}$. Furthermore, let $\hat{\boldsymbol{U}}(\boldsymbol{\zeta}_0) = (U^1(\boldsymbol{\zeta}_0)^T, U^2(\boldsymbol{\zeta}_0)^T)^T$, $U^1(\boldsymbol{\zeta}_0) = \sum_{i=1}^{n} \sum_{k=1}^{K} \boldsymbol{X}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}_0)$ and $U^2(\boldsymbol{\zeta}_0) = \sum_{i=1}^{n} \sum_{k=1}^{K} \boldsymbol{B}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}_0)$. Then

$$\begin{aligned}
\hat{\boldsymbol{\beta}}^o - \boldsymbol{\beta}_0 &= -\left[\boldsymbol{D}^{11} U^1(\boldsymbol{\zeta}_0) + \boldsymbol{D}^{12} U^2(\boldsymbol{\zeta}_0)\right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \boldsymbol{D}^{11} \sum_{i=1}^{n} \sum_{k=1}^{K} \left[\boldsymbol{X}_i - \boldsymbol{B}_i \boldsymbol{D}_{\boldsymbol{BB}}^{-1} \boldsymbol{D}_{\boldsymbol{BX}}\right]^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}_0) + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Thus,

$$\sqrt{n}(\hat{\boldsymbol{\beta}}^o - \boldsymbol{\beta}_0) = (n\boldsymbol{D}^{11}) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{k=1}^{K} \left[\boldsymbol{X}_i - \boldsymbol{B}_i \boldsymbol{D}_{\boldsymbol{BB}}^{-1} \boldsymbol{D}_{\boldsymbol{BX}}\right]^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{\Lambda}_{i(k)} \boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}_0) + o_p(1).$$

Because $\boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}_0)$ are independent random variables with mean zero, and

$$\operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{k=1}^{K}\left[\boldsymbol{X}_i-\boldsymbol{\Pi}_i\boldsymbol{D}_{\boldsymbol{BB}}^{-1}\boldsymbol{D}_{\boldsymbol{BX}}\right]^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Lambda}_{i(k)}\boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}_0)\right)=\boldsymbol{\Xi}.$$

The multivariate central limit theorem implies that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{k=1}^{K}\left[\boldsymbol{X}_i-\boldsymbol{\Pi}_i\boldsymbol{D}_{\boldsymbol{BB}}^{-1}\boldsymbol{D}_{\boldsymbol{BX}}\right]^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Lambda}_{i(k)}\boldsymbol{S}_{i(k)}(\boldsymbol{\zeta}_0)\rightarrow_d N(\boldsymbol{0},\boldsymbol{\Xi}),$$

furthermore, $n\boldsymbol{D}^{11}=\left[\frac{1}{n}\left(\boldsymbol{D}_{\boldsymbol{XX}}-\boldsymbol{D}_{\boldsymbol{XB}}\boldsymbol{D}_{\boldsymbol{XB}}^{-1}\boldsymbol{D}_{\boldsymbol{BX}}\right)\right]^{-1}\rightarrow_p\boldsymbol{\Sigma}$, by the law of large numbers. Then, by using Slutskys theorem, it follows that $\sqrt{n}(\hat{\boldsymbol{\beta}}^o-\boldsymbol{\beta}_0)\rightarrow_d N(\boldsymbol{0},\boldsymbol{\Sigma}^{-1}\boldsymbol{\Xi}\boldsymbol{\Sigma}^{-1})$. This complete the proof of part (a). Furthermore, by using the same arguments of proving the part (a), we can get $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\boldsymbol{B}^T(T_{ij})\left(\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_{0k}\right)\right]=O_p\left(\frac{K_n}{n}\right)$. The triangular inequality implies that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}(\hat{\alpha}_k^o(T_{ij})-\alpha_{0k}(T_{ij}))^2\\
&\leq\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\boldsymbol{B}^T(T_{ij})\left(\hat{\boldsymbol{\gamma}}_k-\boldsymbol{\gamma}_{0k}\right)\right]^2+\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\boldsymbol{B}^T(T_{ij})\boldsymbol{\gamma}_{0k}-\alpha_{0k}(T_{ij})\right]^2\\
&=O_p\left(\frac{K_n}{n}+K_n^{-2r}\right).
\end{aligned}$$

The proof of (b) is completed.

Similar with the proof of (a) and (b), in order to prove (c) and (d), we can directly consider the estimation equation $\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})=\sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\tilde{\boldsymbol{S}}_{i(k)}(\boldsymbol{\zeta})=\boldsymbol{0}$. We first prove that $\frac{1}{\sqrt{n}}[\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})-\hat{\boldsymbol{U}}(\boldsymbol{\zeta})]=o_p(1)$. Let

$$\varsigma_{(k)}\left(\boldsymbol{\zeta}\right)=\begin{pmatrix}\boldsymbol{\Xi}^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\\ K_n^{-1/2}\boldsymbol{H}_{(k)}(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0)+K_n^{1/2}\boldsymbol{H}_{(k)}^{-1}\boldsymbol{B}^T\boldsymbol{\Lambda}_{(k)}\boldsymbol{X}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\end{pmatrix},$$

where $\boldsymbol{H}_{(k)}^2=K_n\boldsymbol{B}^T\boldsymbol{\Lambda}_{(k)}\boldsymbol{B}$. Furthermore, we standardize $\widetilde{\boldsymbol{X}}_{ij}=\boldsymbol{\Xi}^{1/2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_{ij}^T$ and $\widetilde{\boldsymbol{B}}_{ij(k)}=K_n^{1/2}\boldsymbol{H}_{(k)}^{-1}\boldsymbol{B}_{ij}$. Note that $\tilde{S}_{ij(k)}(\boldsymbol{\zeta})-\hat{S}_{ij(k)}(\boldsymbol{\zeta})=I(-\Delta_{ij(k)})\Phi(-|\Delta_{ij(k)}|)$, where $\Delta_{ij(k)} = (\epsilon_{ij} + u_{ij(k)})/r_{ij}$ with $u_{ij(k)} = -(\boldsymbol{\varsigma}_{(k)} (\boldsymbol{\zeta})^T (\widetilde{\boldsymbol{X}}_{ij}^T, \widetilde{\boldsymbol{B}}_{ij(k)}^T)^T + R_{nij})$ and $R_{nij} = \boldsymbol{B}_{ij}^T \boldsymbol{\gamma}_0 - \sum_{l=1}^q Z_{ij,l} \alpha_{0l}(T_{ij})$. We can obtain

$$\frac{1}{\sqrt{n}} \left[\tilde{\boldsymbol{U}}(\boldsymbol{\zeta}) - \hat{\boldsymbol{U}}(\boldsymbol{\zeta}) \right] = \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{k=1}^K \boldsymbol{D}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \begin{bmatrix} I(-\Delta_{i1(k)})\Phi(-|\Delta_{i1(k)}|) \\ \vdots \\ I(-\Delta_{im(k)})\Phi(-|\Delta_{im(k)}|) \end{bmatrix}$$
$$= \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K \boldsymbol{a}_{ij(k)} I(-\Delta_{ij(k)})\Phi(-|\Delta_{ij(k)}|).$$

Note that

$$\begin{aligned} & E\left(\tilde{S}_{ij(k)}(\boldsymbol{\zeta}) - S_{ij(k)}(\boldsymbol{\zeta})\right) \\ =& \int_{-\infty}^{\infty} I(-\Delta_{ij(k)})\Phi(-|\Delta_{ij(k)}|) f_{ij}(\epsilon) d\epsilon \\ =& \int_{-\infty}^{\infty} \Phi(-|\epsilon_{ij} + u_{ij(k)}|/r_{ij})\{2I(\epsilon_{ij} + u_{ij(k)} < 0) - 1\} f_{ij}(\epsilon) d\epsilon \\ =& r_{ij} \int_{-\infty}^{\infty} \Phi(-|t|)\{2I(t<0) - 1\} \left[f_{ij}(0) + f_{ij}'(\eta(t))(r_{ij}t - u_{ij(k)}) \right] dt, \end{aligned}$$

where $\eta(t)$ is between 0 and $r_{ij}t - u_{ij(k)}$. Because $\int_{-\infty}^{\infty} \Phi(-|t|)\{2I(t<0)-1\}dt = 0$, and by condition (C4), there exists a constant C such that $\sup_{i,j} |f_{ij}^T(\eta(t))| \le C$. Then note that $\int_{-\infty}^{\infty} \Phi(-|t|)|t|dt = 1/2$, we have that

$$\left| E\left(\tilde{S}_{ij(k)}(\boldsymbol{\zeta}) - S_{ij(k)}(\boldsymbol{\zeta})\right) \right| \le r_{ij}^2 \int_{-\infty}^{\infty} \Phi(-|t|)|t||f_{ij}'(\eta(t))|dt \le \frac{Cr_{ij}^2}{2}.$$

Under conditions (C7) and (C8), as $n \to \infty$, we can obtain

$$\left\| \frac{1}{\sqrt{n}} E\left[\tilde{\boldsymbol{U}}(\boldsymbol{\zeta}) - \hat{\boldsymbol{U}}(\boldsymbol{\zeta}) \right] \right\| = \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K \sup_{i,j,k} \|\boldsymbol{a}_{ij(k)}\| \frac{Cr_{ij}^2}{2} = o(1).$$

Note that

$$\tilde{\boldsymbol{U}}(\boldsymbol{\zeta}) - \hat{\boldsymbol{U}}(\boldsymbol{\zeta}) = \sum_{k=1}^K \left\{ \sum_{i=1}^n \boldsymbol{D}_i^T \boldsymbol{\Lambda}_{i(k)} \boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0) \begin{bmatrix} I(-\Delta_{i1(k)})\Phi(-|\Delta_{i1(k)}|) \\ \vdots \\ I(-\Delta_{im(k)})\Phi(-|\Delta_{im(k)}|) \end{bmatrix} \right\} = \sum_{k=1}^K M_k.$$

In addition, by Cauchy-Schwartz inequality,

$$\begin{aligned}
\frac{1}{n}\text{var}\left[M_k\right] &= \frac{1}{n}\sum_{i=1}^{n}\text{var}\left[\sum_{j=1}^{m}\boldsymbol{a}_{ij(k)}I(-\Delta_{ij(k)})\Phi(-|\Delta_{ij(k)}|)\right] \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\boldsymbol{a}_{ij(k)}^{T}\boldsymbol{a}_{ij(k)}\text{var}\left[\tilde{S}_{ij(k)}(\boldsymbol{\zeta})-S_{ij(k)}(\boldsymbol{\zeta})\right] \\
&+ \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{j^*=1,j^*\neq j}^{m}\boldsymbol{a}_{ij(k)}^{T}\boldsymbol{a}_{ij^*(k)}\left\{\text{var}\left[\tilde{S}_{ij(k)}(\boldsymbol{\zeta})-S_{ij(k)}(\boldsymbol{\zeta})\right]\text{var}\left[\tilde{S}_{ij^*(k)}(\boldsymbol{\zeta})-S_{ij^*(k)}(\boldsymbol{\zeta})\right]\right\}^{1/2}.
\end{aligned}$$

For $j = 1, \cdots, m$,

$$\begin{aligned}
\text{var}\left[\tilde{S}_{ij(k)}(\boldsymbol{\zeta})-S_{ij(k)}(\boldsymbol{\zeta})\right] &\leq E\left[\tilde{S}_{ij(k)}(\boldsymbol{\zeta})-S_{ij(k)}(\boldsymbol{\zeta})\right]^2 \\
&= \int_{-\infty}^{\infty}\left\{I(-\Delta_{ij(k)})\Phi(-|\Delta_{ij(k)}|)\right\}^2 f_{ij}(\epsilon)d\epsilon \\
&= r_{ij}\int_{-\infty}^{\infty}\Phi^2(-|t|)f_{ij}(r_{ij}t-u_{ij(k)})dt \\
&= r_{ij}\int_{|t|>c}\Phi^2(-|t|)f_{ij}(r_{ij}t-u_{ij(k)})dt \\
&\quad + r_{ij}\int_{|t|\leq c}\Phi^2(-|t|)f_{ij}(r_{ij}t-u_{ij(k)})dt \\
&\leq \Phi^2(-c)+r_{ij}cf_{ij}(\eta),
\end{aligned}$$

where η is a positive value, and η lies between $(-r_{ij}c - u_{ij(k)}, r_{ij}c - u_{ij(k)})$. Let $c = n^{1/3}$, under condition (C8), since $r_{ij} = O(n^{-1/2})$, then $r_{ij}c = O(n^{-1/6})$. Note that $\Phi^2(-c) \to 0$ and $r_{ij}cf_{ij}(\eta) \to 0$, as $n \to \infty$. By conditions (C4) and (C6), it is easy to get that $\frac{1}{n}\text{var}\left[M_k\right] = o(1)$. Therefore, by Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\frac{1}{n}\text{var}\left\{\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})-\hat{\boldsymbol{U}}(\boldsymbol{\zeta})\right\} &= \frac{1}{n}\text{var}\left\{\sum_{k=1}^{K}M_k\right\} \\
&\leq \frac{1}{n}\sum_{k=1}^{K}\text{var}\left\{M_k\right\}+\frac{1}{n}\sum_{k=1}^{K}\sum_{k'=1,\neq k}^{K}\sqrt{\text{var}\left\{M_k\right\}}\sqrt{\text{var}\left\{M_k'\right\}} \\
&= o(1).
\end{aligned}$$

Therefore, we have $\frac{1}{\sqrt{n}}\left[\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})-\hat{\boldsymbol{U}}(\boldsymbol{\zeta})\right]\to 0$ as $n\to\infty$ for any $\boldsymbol{\zeta}$.
Furthermore, note that $\sup_{\boldsymbol{\zeta}}\|\frac{1}{n}\left(\bar{\boldsymbol{U}}(\boldsymbol{\zeta})-\hat{\boldsymbol{U}}(\boldsymbol{\zeta})\right)\| = o\left(\frac{1}{\sqrt{n}}\right)$, thus we have that $\sup_{\boldsymbol{\zeta}}\|\frac{1}{n}\left(\bar{\boldsymbol{U}}(\boldsymbol{\zeta})-\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})\right)\| = o\left(\frac{1}{\sqrt{n}}\right)$. Because that $\boldsymbol{\zeta}_0$ is the unique solution of equation of $\bar{\boldsymbol{U}}(\boldsymbol{\zeta})=\mathbf{0}$. This together with the definition of $\hat{\boldsymbol{\zeta}}^s$ implies that $\hat{\boldsymbol{\zeta}}^s\to_p \boldsymbol{\zeta}_0$ as $n\to\infty$. In order to prove the asymptotic normality, we first prove that $\frac{1}{n}\{\frac{\partial\tilde{\boldsymbol{U}}_0(\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0}-\boldsymbol{D}(\boldsymbol{\zeta}_0)\}\to_p 0$. Note that

$$
\begin{aligned}
E\left[\frac{\partial\tilde{\boldsymbol{U}}_0(\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0}\right]-\boldsymbol{D}(\boldsymbol{\zeta}_0) &= -\sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{D}_i^T\boldsymbol{\Lambda}_{i(k)}\boldsymbol{V}_{i(k)}^{-1}(\boldsymbol{\theta}_0)\left\{E\left[\boldsymbol{\Phi}_{i(k)}\right]-\boldsymbol{\Lambda}_{i(k)}\right\}\boldsymbol{D}_i \\
&= -\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{m}\boldsymbol{a}_{ij(k)}\left\{\frac{1}{r_{ij}}E\phi\left(\frac{\epsilon_{ij}+u_{ij(k)}}{r_{ij}}\right)-f_{ij}(0)\right\}\boldsymbol{D}_{ij},
\end{aligned}
$$

and because that

$$
\begin{aligned}
&\left|\frac{1}{r_{ij}}E\phi\left(\frac{\epsilon_{ij}+u_{ij(k)}}{r_{ij}}\right)-f_{ij}(0)\right| \\
=&\left|\frac{1}{r_{ij}}\int_{-\infty}^{+\infty}\phi\left(\frac{\epsilon+u_{ij(k)}}{r_{ij}}\right)f_{ij}(\epsilon)d\epsilon-f_{ij}(0)\right| \\
=&\left|\int_{-\infty}^{+\infty}\phi(t)\left\{f_{ij(k)}(0)+[r_{ij(k)}t-u_{ij(k)}]f_{ij}(\eta(t))\right\}dt-f_{ij}(0)\right| \\
=&\left|\int_{-\infty}^{+\infty}\phi(t)[r_{ij}t-u_{ij(k)}]f_{ij}(\eta(t))dt\right| \\
\le& r_{ij}\int_{-\infty}^{+\infty}\left|\phi(t)tf_{ij}(\eta(t))\right|dt+\int_{-\infty}^{+\infty}\left|\phi(t)u_{ij(k)}f_{ij}(\eta(t))\right|dt,
\end{aligned}
$$

where $\eta(t)$ lies between 0 and $r_{ij}t-u_{ij(k)}$. By condition (C4), $f_{ij}(\cdot)$ is uniformly bounded, hence there exists a constant C satisfying $|f_{ij}^T(\eta(t))|\le C$, and by condition (C8), we have $\left|\frac{1}{r_{ij}}E\phi\left(\frac{\epsilon_{ij}+u_{ij(k)}}{r_{ij}}\right)-f_{ij}(0)\right|\to 0$. By the strong law of large number, we know that $\frac{1}{n}\left[\frac{\partial\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0}\right]\to E\left[\frac{1}{n}\frac{\partial\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0}\right]$. Using the triangle inequality, we have

$$
\begin{aligned}
&\left|\frac{1}{n}\left\{\frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0} - \boldsymbol{D}(\boldsymbol{\zeta}_0)\right\}\right| \\
&\leq \left|\frac{1}{n}\left\{\frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0} - E\left[\frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0}\right]\right\}\right| + \left|\frac{1}{n}\left\{E\left[\frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0}\right] - \boldsymbol{D}(\boldsymbol{\zeta}_0)\right\}\right| \\
&= o(1),
\end{aligned}
$$

which implies that $\frac{1}{n}\left\{\frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0} - \boldsymbol{D}(\boldsymbol{\zeta}_0)\right\} \to_p 0$. By Taylor series expansion of $\tilde{\boldsymbol{U}}(\boldsymbol{\zeta})$ around $\boldsymbol{\zeta}_0$, we have $\tilde{\boldsymbol{U}}(\boldsymbol{\zeta}) = \tilde{\boldsymbol{U}}(\boldsymbol{\zeta}_0) + \frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}^*}(\boldsymbol{\zeta} - \boldsymbol{\zeta}_0)$, where $\boldsymbol{\zeta}^*$ lies between $\boldsymbol{\zeta}$ and $\boldsymbol{\zeta}_0$. Because $\tilde{\boldsymbol{U}}(\hat{\boldsymbol{\zeta}}^s) = \mathbf{0}$ and $\hat{\boldsymbol{\zeta}}^s \to \boldsymbol{\zeta}_0$, we therefore obtain $\boldsymbol{\zeta}^* \to \boldsymbol{\zeta}_0$ and $\frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}^*} \to \frac{\partial \tilde{\boldsymbol{U}}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_0}$. Then by the same arguments used in the proof of (a) and (b), we can complete the proof of (c) and (d).

References

Hendricks, W., Koenker, R. (1992). Hierarchical spline models for conditional quantiles and the demand for electricity. *Journal of the American Statistical Association*, 87, 58-68.

Jung, S. (1996). Quasi-likelihood for median regression models. *Journal of the American Statistical Association*, 91, 251-257.

Pollard, D. (1990). *Empirical processes: Theories and applications*. Hayward, CA: Institute of Mathematical Statistics.

Wang, H., Feng, X., Dong, C. (2018). Copula-based quantile regression for longitudinal data. *Statistica Sinica*, doi: 10.5705/ss.202016.0135.