The Strong Consistency for the Estimators of Longitudinal Data in Semiparametric Regression Model with $\tilde{\rho}$-Mixing Errors

Ping Xiao, Xinsheng Liu *

State Key Laboratory of Mechanics and Control of Mechanical Structures, Institute of Nano Science, and College of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

* Corresponding Author Email: xsliu@nuaa.edu.cn

Abstract. Consider the following semiparametric regression model for longitudinal data with $\tilde{\rho}$-mixing errors: $Y_{ij} = X_{ij}^T\beta + g(t_{ij}) + e_{ij}, i = 1, \dots, n \ \ j = 1, \dots, m_i$, where, the response variable Y_{ij} and the covariate vector $X_{ij}\epsilon R^p$ taken from the i-th subject at time t_{ij}, $X_{ij} = (x_{ij1}, \dots, x_{ijp})^T$. $\{e_{ij}\}$ is the $\tilde{\rho}$-mixing random variables, $Ee_{ij} = 0$. We establish a strong consistency for the least squares estimator of the parametric β and the estimator of the non-parametric function $g(\cdot)$ under some mild conditions.

Keywords: $\tilde{\rho}$-mixing random variables, semiparametric regression model, longitudinal data, Strong consistency, convergence rate.

1. Introduction

Longitudinal data refers to the data obtained by observing the same group of subjects several times at different times. Longitudinal data incorporate cross-sectional data and time series data. Therefore, the structure of longitudinal data is more complex and the data dimension is higher. Longitudinal data are a common type of data in biological and economic applications. Generally, the observations of different individuals are independent of each other. However, different observations of the same individual are often correlated.

Consider the following partial linear regression model for longitudinal data:

$$Y_{ij} = X_{ij}^T\beta + g(t_{ij}) + e_{ij} \qquad i = 1, \dots, n \ \ j = 1, \dots, m_i, \tag{1}$$

Where, $(X_{ij}, t_{ij}) \in R^p \times R$ are known fixed design points. Without loss of generality, we assume $t_{ij}\epsilon[0,1]$. Assume that observations from different observing individuals are independent of each other, but observations from the same individual are interdependent. $g(\cdot)$ is the unknown smooth function, $\beta$is the unknown p-dimensional parameter. The observations are obtained from n subjects and for the i-th subject with m_i observations over time. We assume that n can be sufficiently large and $\{m_i\}$ is a sequence of positive integers.

Model (1) combines the features of parametric and nonparametric models, so it has stronger interpretability. Considering that the random errors are independent and identically distributed in model (1), Gao and Chen, et al. [1] estimated $g(\cdot)$ by kernel method and Nearest-neighbor method, and estimated the parameter β by least weighted square method, and proved the asymptotic normality of the estimator. Hu and Wu [2] obtained the strong convergence rate of the estimator by using the wavelet estimation when the random errors were NA random sequences. In addition, some scholars have studied the partial linear regression model for longitudinal data. Tian and Xue [3] studied the estimation of the unknown parameter β and the unknown function $g(\cdot)$ in the model (1) with independent and identically distributed errors. Zhou and Lin [4, 5] gave the estimator by weighted least squares method and proved the strong consistency of the estimator under the random errors as martingale difference, $\emptyset - \text{mixing}$ and $\rho - \text{mixing}$ sequence.

In this paper, it is assumed that there are $\tilde{\rho}$ −mixing correlations among different observations of the same individual, that is, we consider the error in model (1) as $\tilde{\rho}$ −mixing.

$\{X_n, n \in N\}$ be a sequence of random variables defined on the fixed probability space$(\Omega, \mathcal{A}, \boldsymbol{P})$, $\sigma -$ fields $\mathcal{F}\mathrm{s} = \sigma(X_i, i \in S \subset N)$, $F\,_1^k 1k = \sigma(X_i, i \leq k)$, $F\,_{k+n}^{+\infty} = \sigma(X_i, i \geq k+n)$, $S_j = \sum_{i=1}^{j} X_i$, $||X||_p = (E|X|^p)^{\frac{1}{p}}$. F, R are fixed $\sigma -$ fileds in A. Let

$$\rho(\mathcal{F}, \mathcal{R}) = \sup\{\left|\mathrm{corr}(X, Y)\right| : X \in L_2\,(\mathcal{F}), Y \in L_2(\mathcal{R})\} \tag{2}$$

Where $\mathrm{corr}(X, Y)$ represents the correlation coefficient of random variables X, Y. For $n \geq 0$, Let

$$\rho(n) = \sup_{k \in N} \rho\,((\mathcal{F}_1^k, \mathcal{F}_{k+n}^{+\infty}), \tag{3}$$

$$\tilde{\rho}(\rho) = \sup\,\{\rho(\mathcal{F}_S, \mathcal{R}_T)\} : S, T \text{ are finite subsets of } \mathbf{N}, \text{ s.t. } \mathrm{dist}(S, T) \geq n\}, \tag{4}$$

Where, $\mathrm{dist}(S, T)$ represents the distance between S and T.

$$0 \leq \tilde{\rho}(n+1) \leq \tilde{\rho}(n) \leq 1, \tilde{\rho}(0) = 1. \tag{5}$$

Definition 2.1 [8]: Let $\{X_n, n \geq 1\}$ be a sequence of random variables. $\{X_n, n \geq 1\}$ is said to be $\rho -$ mixing, if $\rho(n) \to 0, n \to \infty$; $\{X_n, n \geq 1\}$ is called $\tilde{\rho} -$ mixing, or weakly dependent, if there exists $n_0 \geq 1$, s.t. $\tilde{\rho}(n_0) < 1$, then.

The $\rho -$ mixing sequence was first proposed by Kolmogorov and Rozanov [6] in 1960, which aroused the research of many scholars and obtained a series of achievements. The $\tilde{\rho} -$ mixing sequence was first proposed by Bradley (1990) [7], which is a generalization of the $\rho -$ mixing sequence. Wang, et al [8] pointed out that the condition of $\tilde{\rho} -$ mixing sequence is weaker than that of $\rho -$ mixing sequence, so $\tilde{\rho} -$ mixing is a more extensive dependent sequence, and the study has important theoretical significance and practical value. Many scholars have studied $\tilde{\rho} -$ mixing sequence and obtained a series of research results, such as: Sung [9, 10] and Wu, et al [11] studied the complete convergence of $\tilde{\rho} -$ mixing random variables. Wu [12, 13]studied the strong convergence of $\tilde{\rho} -$ mixing sequence, etc.

Wang, et al [8] studied the strong convergence, r-th ($r > 2$) mean consistency and complete consistency for the estimators of semiparametric regression model based on weakly dependent errors. On the basis of this literature, this paper considers a partial linear regression model for longitudinal data, and proves the strong consistency of the estimator and the convergence rate.

2. Main Result

To obtain the estimators of the unknown parameter β and the unknown function $g(\cdot)$ in model (1), we first assume that the parameter β is known, then model (1) can be regarded as a nonparametric regression model $Y_{ij} - X_{ij}^T\beta = g(t_{ij}) + e_{ij}$ $\quad i = 1, \dots, n$ $\; j = 1, \dots, m_i$. Using the method of probability weight function, we can get a preliminary estimate of $g(\cdot)$ as:

$$\hat{g}_0(t, \beta) = \sum_{i=1}^{n}\sum_{j=1}^{m_i} W_{nij}(t)(Y_{ij} - X_{ij}^T\beta), \tag{6}$$

Where $W_{nij}(t), t \in [0,1]$ is probability weight function. Next we estimate β by weighted least squares

$$\hat{\beta}_n = \underset{\beta}{\mathrm{argmin}} \sum_{i=1}^{n}\sum_{j=1}^{m_i}(Y_{ij} - X_{ij}^T\beta - \hat{g}_0(t_{ij}, \beta))^2. \tag{7}$$

For convenience, let $\tilde{X}_{ij} = X_{ij} - \sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}(t_{ij})\, X_{kl}$, $\tilde{Y}_{ij} = Y_{ij} - \sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}(t_{ij})\, Y_{kl}$. The estimator of parameter β is

$$\hat{\beta}_n = (\sum_{i=1}^{n}\sum_{j=1}^{m_i} \tilde{X}_{ij}\, \tilde{X}_{ij}^{\;T})^{-1} \sum_{i=1}^{n}\sum_{j=1}^{m_i} \tilde{X}_{ij}\, Y_{ij}. \tag{8}$$

Substitute Equation (8) into Equation (6), we get

$$\hat{g}_n(t,\beta) = \hat{g}_0\left(t,\hat{\beta}_n\right) = \sum_{i=1}^{n}\sum_{j=1}^{m_i} W_{nij}(t)(Y_{ij} - X_{ij}^T\hat{\beta}_n). \tag{9}$$

Denote $S_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{m_i} \widetilde{X}_{ij}\,\widetilde{X}_{ij}^{\,T}, \tilde{g}(t) = g(t) - \sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}(t)\,g(t_{kl})$.

To obtain the results of this paper, we give the following appropriate conditions:

A1 (i) $\{e_{ij}, 1 \le j \le m_i\}$ is a $\tilde{\rho} -$ mixing sequence, $Ee_{ij} = 0, i = 1, \dots n$.

A2 (i) $\max\limits_{1\le i\le n} m_i = o\left(n^{\delta}\right)\ 0 < \delta < \frac{r-2}{2r}, r > 2$.

(ii) $\lim\limits_{n\to\infty} \frac{S_n^2}{N(n)} = \sum$, where, $\sum$ is a positive definite matrix, $N(n) = \sum_{i=1}^{n} m_i$

(iii) $g(\cdot)$ satisfies the first-order Lipschitz condition.

A3 (i) $\sum_{i=1}^{n}\sum_{j=1}^{m_i} W_{nij}(t) = 1,\ t \in [0,1]$.

(ii) $\sup\limits_{0\le t\le 1} \max\limits_{\substack{1\le i\le n,\\ 1\le j\le m_i}} W_{nij}(t) = O\left(n^{-\frac{1}{2}}\right)$.

(iii) $\sup\limits_{0\le t\le 1} \sum_{i=1}^{n}\sum_{j=1}^{m_i} W_{nij}(t)\, I\left(\left|t_{ij} - t\right| > a_n\right) = O(b_n)$ where $b_n = n^{-\frac{1}{3}}\log n$.

(iv) $\sup\limits_{0\le t\le 1} \left\|\sum_{i=1}^{n}\sum_{j=1}^{m_i} W_{nij}(t)X_{ij}\right\| = O(1)$.

(v) $\forall s, t \in [0,1]$, uniformly for $\max\limits_{1\le i\le n, 1\le j\le m_i} m_i \quad |W_{nij}(t) - W_{nij}(s)| \le C|t - s|$.

Remark 1: When many scholars study longitudinal data, they often limit the number of repeated observations of each individual to a finite number of times, that is, restrict $\{m_i\}$ to be a bounded sequence of positive integers. The constraint on $\{m_i\}$ in A1 (ii) is weaker, and m_i can also be sufficiently large.

Theorem 2.1 Let $\left\{e_{ij}, 1 \le j \le m_i\right\}$ be a $\tilde{\rho} -$ mixing sequence, $Ee_{ij} = 0, i = 1, \dots n$. $\exists \alpha > 3$ s.t.

$$\max_{\substack{1\le i\le n,\\ 1\le j\le m_i}} E(|e_{ij}|^{\alpha}) \le C \quad \text{a.s.}$$

Suppose that conditions (A1)-(A3) hold, we have

$$\left\|\hat{\beta}_n - \beta\right\| = O\left(n^{-\frac{1}{3}}\log n\right) \tag{10}$$

Theorem 2.2 Under the same assumptions in Theorem 2.1, we have

$$\sup_{0\le t\le 1} \left|\hat{g}_n(t) - g(t)\right| = O\left(n^{-\frac{1}{3}}\log n\right). \tag{11}$$

3. Several technical lemma

Lemma 3.1 Suppose A1(i) holds. Let $\alpha > 1, 0 < r < \alpha$ and

$$e'_{ij} = e_{ij} I\left(\left|e_{ij}\right| \le \varepsilon i^{\frac{1}{r}} m_i\right),$$

$$e''_{ij} = e_{ij} - e'_{ij} = e_{ij} I\left(e_{ij} > \varepsilon i^{\frac{1}{r}} m_i\right) + e_{ij} I\left(e_{ij} < -\varepsilon i^{\frac{1}{r}} m_i\right),$$

$\forall \varepsilon > 0$, if $\max\limits_{\substack{1\le i\le n\\ 1\le j\le m_i}} E|e_{ij}|^{\alpha} \le C$, a.s., we have

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left|e''_{ij}\right| < \infty \ \text{a.s.}.$$

Proof: it can be obtained by simple calculation referring to Lemma 3.1 in Ref.[5].

Lemma 3.2 Let $\left\{e_{ij}, 1 \le j \le m_i\right\}$ be a $\tilde{\rho} -$ mixing sequence, there is a constant $M > 0$, s.t. $|e_{ij}| \le M$, for $0 \le x \le \frac{2}{M}$, we have

$$\mathrm{E}\left(\exp\left(\mathrm{x}\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{e_{ij}}\right)\right) \le \exp\left(\frac{1}{2}(1+\mathrm{xM})\mathrm{x}^2\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{Ee_{ij}}^2\right).$$

Proof: First prove that $\mathrm{E}(\exp\,(\mathrm{xe_{ij}})) \le \exp\,(\frac{1}{2}(1+\mathrm{xM})\mathrm{x}^2\mathrm{Ee_{ij}}^2)$ is established

$$\mathrm{E}\big(\exp(\mathrm{xe_{ij}})\big) = \sum_{\mathrm{k}=0}^{\infty}\frac{\mathrm{E}((\mathrm{xe_{ij}})^{\mathrm{k}})}{\mathrm{k}!} \le 1 + \mathrm{E}\big(\mathrm{x}^2\mathrm{e_{ij}}^2\big)\left(\frac{1}{2!}+\frac{\mathrm{xM}}{3!}+\frac{\mathrm{x}^2\mathrm{M}^2}{4!}+\cdots\right)$$
$$\le 1 + \frac{\mathrm{x}^2\mathrm{E}\big(\mathrm{e_{ij}}^2\big)}{\mathrm{x}^2\mathrm{M}^2}\left(\frac{(\mathrm{xM})^2}{2!}+\frac{(\mathrm{xM})^3}{3!}+\frac{(\mathrm{xM})^4}{4!}+\cdots\right)$$
$$= 1 + \frac{\mathrm{E}\big(\mathrm{e_{ij}}^2\big)}{\mathrm{M}^2}(\mathrm{e^{xM}}-1-\mathrm{xM}) \le \exp\,(\mathrm{E}\big(\mathrm{e_{ij}}^2\big)\frac{\mathrm{e^{xM}}-1-\mathrm{xM}}{\mathrm{M}^2})$$
$$\le \exp\left(\frac{1}{2}(1+\mathrm{xM})\mathrm{x}^2\mathrm{Ee_{ij}}^2\right) \tag{12}$$

The last inequality holds for $|\mathrm{xM}| < 2$. According to the definition of $\tilde{\rho}-$mixing sequence, $\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{e_{ij}}$ is also a $\tilde{\rho}-$mixing sequence, so Lemma 3.2 is proved. ■

Lemma 3.3 Suppose $\{\mathrm{e_{ij}}, 1 \le \mathrm{j} \le \mathrm{m_i}\}$ be a $\tilde{\rho}-$mixing sequence, and there is a constant $\alpha > 3$, $\mathrm{s.t.}\ \sup_{\substack{\mathrm{i}\ge 1,\\ 1\le \mathrm{j}\le \mathrm{m_i}}} \mathrm{E}|\mathrm{e_{ij}}|^{\alpha} \le \mathrm{C} < \infty$. Suppose $\mathrm{a_{nij}(t)}$ satisfy

(i) $\exists \mathrm{C_1} > 0, \mathrm{C_2} > 0, \mathrm{s.t.}$ for $\forall \mathrm{t} \in [0,1]$ and sufficiently large n, have

$$\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{a_{nij}(t)} \le \mathrm{C_1},\ \max_{\substack{1\le \mathrm{i}\le \mathrm{n}\\ 1\le \mathrm{j}\le \mathrm{m_i}}} \mathrm{a_{nij}(t)n^{\frac{1}{2}}} \le \mathrm{C_2}$$

(ii) $\exists$ constant $\mathrm{C_3}\ \mathrm{s.t.} \forall \mathrm{s,t} \in [0,1]$ has $\max_{\substack{1\le \mathrm{i}\le \mathrm{n}\\ 1\le \mathrm{j}\le \mathrm{m_i}}} |\mathrm{a_{nij}(t)} - \mathrm{a_{nij}(s)}| \le \mathrm{C_3}\ |\mathrm{s}-\mathrm{t}|$, then when n is sufficiently large, we have

$$\sup_{0\le \mathrm{t}\le 1}\left|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{a_{nij}(t)\,e_{ij}}\right| = \mathrm{O}(\mathrm{n}^{-\frac{1}{3}}\mathrm{logn})\ \mathrm{a.s.}\ \mathrm{n}\to\infty.$$

Lemma 3.3 is a generalization of Lemma 2 in the Ref. [14]. Compared with the original lemma, the lemma here is applicable to longitudinal data. Furthermore, this lemma weakens the conditions from $\max_{\substack{1\le \mathrm{i}\le \mathrm{n}\\ 1\le \mathrm{j}\le \mathrm{m_i}}} \mathrm{a_{nij}(t)n^{\frac{2}{3}}} \le \mathrm{C_2}$ to $\max_{\substack{1\le \mathrm{i}\le \mathrm{n}\\ 1\le \mathrm{j}\le \mathrm{m_i}}} \mathrm{a_{nij}(t)n^{\frac{1}{2}}} \le \mathrm{C_2}$. Lemma 3.3 is a generalization of the original lemma.

Proof: Denote $\mathrm{B_n(t)} = \sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{a_{nij}(t)\,e_{ij}}$, $\mathrm{B_n} = \sup_{0\le \mathrm{t}\le 1}|\mathrm{B_n(t)}|$. For $\forall$ n, $\exists\ \mathrm{O}(\mathrm{n}^2)$ small fields $\mathrm{D_{nj}}$ with radius n^2 covering $[0,1]$, let the center of $\mathrm{D_{nj}}$ is $\mathrm{S_{nj}}$, then for $\forall \mathrm{t} \in [0,1]$, there is some $\mathrm{S_{nj}} = \mathrm{S_n(t)}, \mathrm{s.t.}\, \mathrm{t} \in \mathrm{D_{nj}}$.

$$\mathrm{B_n} = \sup_{0\le \mathrm{t}\le 1}|\mathrm{B_n(t)}| = \sup_{0\le \mathrm{t}\le 1}|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{a_{nij}(t)\,e_{ij}}| \le \sup_{0\le \mathrm{t}\le 1}\left|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{a_{nij}(t)}\left(\mathrm{e_{ij}}-\mathrm{e'_{ij}}\right)\right|$$
$$+\sup_{0\le \mathrm{t}\le 1}\left|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}(\mathrm{a_{nij}(t)}-\mathrm{a_{nij}}\big(\mathrm{S_n(t)}\big))\mathrm{e'_{ij}}\right| + \sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}a_{nij}\big(S_n(t)\big)(\,e'_{ij}-Ee'_{ij})\right|$$
$$+\sup_{0\le \mathrm{t}\le 1}\left|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}(\mathrm{a_{nij}(t)}-\mathrm{a_{nij}}\big(\mathrm{S_n(t)}\big))\mathrm{Ee'_{ij}}\right| + \sup_{0\le \mathrm{t}\le 1}\left|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{a_{nij}(t)\,E}\big(\mathrm{e'_{ij}}-\mathrm{e_{ij}}\big)\right|$$
$$= \mathrm{B_{1n}}+\mathrm{B_{2n}}+\mathrm{B_{3n}}+\mathrm{B_{4n}}+\mathrm{B_{5n}} \tag{13}$$

Denote $\mathrm{e''_{ij}} = \mathrm{e_{ij}} - \mathrm{e'_{ij}}$. From Lemma 3.1 we know $\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}|\mathrm{e''_{ij}}| < \infty$ a.s.

$$\mathrm{B_{1n}} = \sup_{0\le \mathrm{t}\le 1}\left|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{a_{nij}(t)}\left(\mathrm{e''_{ij}}\right)\right|$$
$$\le \sup_{0\le \mathrm{t}\le 1}\max_{\substack{1\le \mathrm{i}\le \mathrm{n}\\ 1\le \mathrm{j}\le \mathrm{m_i}}}|\mathrm{a_{nij}(t)}|\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{m_i}}\mathrm{e''_{ij}} = \mathrm{O}\left(\mathrm{n}^{-\frac{1}{2}}\right)\ \mathrm{a.s.}. \tag{14}$$

$$\begin{aligned} B_{2n} &= \sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left(a_{nij}(t)-a_{nij}\left(S_n(t)\right)\right)e'_{ij}\right| \\ &\le \sup_{0\le t\le 1}\max_{i,j}\left|a_{nij}(t)-a_{nij}\left(S_n(t)\right)\right|\cdot\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}e'_{ij}\right| \\ &\le C_3\ \left|t-S_n(t)\right|\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}e'_{ij}\right| = O\left(n^{2\delta+\frac{1}{r}-2}\right)\ = O\left(n^{-\frac{2}{3}}\right)\ \text{a.s..} \end{aligned} \tag{15}$$

Similarly, we can get:

$$B_{4n} = O\left(n^{-\frac{2}{3}}\right)\ \text{a.s..} \tag{16}$$

$$\begin{aligned} B_{5n} &= \sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}a_{nij}(t)\,E\left(e'_{ij}\ -e_{ij}\right)\right| \\ &\le \sup_{0\le t\le 1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}|a_{nij}(t)|\cdot E|e_{ij}|I(|e_{ij}|>\varepsilon i^{\frac{1}{r}}m_i)) \\ &\le \sup_{0\le t\le 1}\max_{\substack{1\le i\le n\\1\le j\le m_i}}\left|a_{nij}(t)\right|\sum_{i=1}^{n}\sum_{j=1}^{m_i}\ \left(\varepsilon i^{\frac{1}{r}}m_i\right)^{-2}E|e_{ij}|^3 \\ &\le C_3 n^{-\frac{1}{2}}\varepsilon^{-2}n^{\frac{1}{3}}n^{-2\delta}\max_{\substack{1\le i\le n\\1\le j\le m_i}}\ E\left|e_{ij}\right|^3 \end{aligned}$$

Take $\varepsilon = n^{-\frac{1}{6}}$, then

$$B_{5n} = O\left(n^{-\frac{1}{2}}\right). \tag{17}$$

We next prove:

$$B_{3n} = \sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}a_{nij}\left(S_n(t)\right)(\ e'_{ij}-Ee'_{ij})\right| = O\left(n^{-\frac{1}{3}}\log n\right).$$

Denote $W_{ij}^{(n)} = n^{\frac{13}{12}}\max\limits_{\substack{1\le i\le n\\1\le j\le m_i}}\left|a_{nij}(u)\right|\cdot a_{nij}(u)\left(e'_{ij}-Ee'_{ij}\right),\ u\in D_n(S_n(t)),$

$$M = 2n^{\frac{13}{12}}\max_{\substack{1\le i\le n\\1\le j\le m_i}}\left|a_{nij}(u)\right|^2, x = C_2^{-1}n^{-\frac{3}{4}}(\max_{\substack{1\le i\le n\\1\le j\le m_i}}\left|a_{nij}(u)\right|)^{-1}, a_n = n^{\frac{3}{4}}\log n\max_{\substack{1\le i\le n\\1\le j\le m_i}}\left|a_{nij}(u)\right|.$$

Apparently $\{W_{ij}^{(n)}, 1\le j\le m_i\}$ is a $\tilde{\rho}-$mixing sequence. From Lemma 3.2,

$$\begin{aligned} &P(\sum_{i=1}^{n}\sum_{j=1}^{m_i}a_{nij}(u)(\ e'_{ij}-Ee'_{ij}) > C_5 n^{-\frac{1}{3}}\log n) \\ &= P\left(\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{ij}^{(n)} > C_5 a_n\right) \\ &\le \exp\,(-C_5 a_n x)E\exp(x\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{ij}^{(n)}) \\ &\le \exp(-C_5 a_n x)\cdot\exp\,(\frac{1}{2}(1+xM)x^2\sum_{i=1}^{n}\sum_{j=1}^{m_i}E{W_{ij}^{(n)}}^2) \\ &\le \exp\left(-C_5 a_n x + C_6 x^2 n^2\max_{\substack{1\le i\le n\\1\le j\le m_i}}\left|a_{nij}(u)\right|^2\sum_{i=1}^{n}\sum_{j=1}^{m_i}E{e_{ij}}^2(a_{nij}(u))^2\right) \\ &\le \exp\left(-C_5 C_2^{-1}\log n + C_6 C_2^{-1}C_1\max_{\substack{1\le i\le n\\1\le j\le m_i}}E{e_{ij}}^2\right) \\ &= \exp\left(C_6 C_2^{-1}C_1\max_{\substack{1\le i\le n\\1\le j\le m_i}}E{e_{ij}}^2\right)n^{-\frac{C_5}{C_2}}. \end{aligned} \tag{18}$$

Using symmetry, we get

$$P\left(\sum_{i=1}^{n}\sum_{j=1}^{m_i} a_{nij}(u)(e_{ij}' - Ee_{ij}'\right) < -C_5 n^{-\frac{1}{3}}\log n) \le \exp\left(C_6 C_2^{-1} C_1 \max_{\substack{1\le i\le n\\ 1\le j\le m_i}} Ee_{ij}{}^2 \right) n^{-\frac{C_5}{C_2}}. \qquad (19)$$

Combining equation (18) and equation (19), we can get

$$P\left(\sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i} a_{nij}\big(S_n(t)\big)(e_{ij}' - Ee_{ij}')\right| > C_5 n^{-\frac{1}{3}}\log n\right)$$
$$\le P\left(\bigcup_{j=1}^{n^3}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i} a_{nij}\big(S_n(t)\big)(e_{ij}' - Ee_{ij}')\right| > C_5 n^{-\frac{1}{3}}\log n\right)$$
$$\le \sum_{j=1}^{n^3} P\left(\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i} a_{nij}\big(S_n(t)\big)(e_{ij}' - Ee_{ij}')\right| > C_5 n^{-\frac{1}{3}}\log n\right)$$
$$\le Cn^3 n^{-\frac{C_5}{C_2}}. \qquad (20)$$

Choose an appropriately large C_5 , according to the Borel-Cantelli lemma, $B_{3n} = O\left(n^{-\frac{1}{3}}\log n\right)$ a. s..Together with equations (3.3) - (3.6), Lemma 3.3 is prove.■

Remark: If $\{a_{nij}(t), 1 \le i \le n, 1 \le j \le m_i\}$ is a constant sequence Lemma 3.3 still holds.

4. Proof of Theorems

4.1. Proof of Theorem 2.1

$$\hat{\beta}_n - \beta = \left(\sum_{i=1}^{n}\sum_{j=1}^{m_i} \tilde{X}_{ij}\tilde{X}_{ij}{}^{T}\right)^{-1} \sum_{i=1}^{n}\sum_{j=1}^{m_i} \tilde{X}_{ij}\,(\tilde{Y}_{ij} - \tilde{X}_{ij}^{T}\beta)$$
$$= S_n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{m_i} \tilde{X}_{ij}\left[e_{ij} - \sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}\,(t_{ij})e_{kl} + \tilde{g}(t_{ij})\right]$$
$$= \left(\frac{S_n^2}{N(n)}\right)^{-1}\Big[\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\tilde{X}_{ij}}{N(n)}e_{ij} - \sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\tilde{X}_{ij}}{N(n)}\left(\sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}\,(t_{ij})e_{kl}\right)$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\tilde{X}_{ij}}{N(n)}\tilde{g}(t_{ij})\Big]=D_{1n} + D_{2n} + D_{3n} \qquad (21)$$

From condition A2(ii),we know

$$\frac{1}{N(n)}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\|\tilde{X}_{ij}\| = O(1), \qquad (22)$$

$$\max_{\substack{1\le i\le n,\\ 1\le j\le m_i}}\|\tilde{X}_{ij}\| = o\left(N(n)^{\frac{1}{2}}\right). \qquad (23)$$

By Lemma 3.3 and (4.2)-(4.3), we get

$$\|D_{1n}\| = \left(\frac{S_n^2}{N(n)}\right)^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\tilde{X}_{ij}}{N(n)}e_{ij}$$
$$\le C\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\|\tilde{X}_{ij}\|}{N(n)}e_{ij}\right| = O\left(n^{-\frac{1}{3}}\log n\right). \qquad (24)$$

$$\|D_{2n}\| = -\left(\frac{S_n^2}{N(n)}\right)^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\tilde{X}_{ij}}{N(n)}\left(\sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}\,(t_{ij})e_{kl}\right)$$
$$\le C\max_{\substack{1\le i\le n,\\ 1\le j\le m_i}}\left\|\sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}\,(t_{ij})e_{kl}\right\|\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\|\tilde{X}_{ij}\|}{N(n)}$$
$$= O\left(n^{-\frac{1}{3}}\log n\right). \qquad (25)$$

Take $b_n = n^{-\frac{1}{3}}\log n$, from condition A2(iii) and A3(iii)

$$\sup_{0\le t\le 1}|\tilde{g}(t)| = \sup_{0\le t\le 1}\left|g(t) - \sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}\,(t)g(t_{kl})\right|$$
$$\le \sup_{0\le t\le 1}\left|\sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}\,(t)\big(g(t) - g(t_{kl})\big)I(|t - t_{kl}| > b_n)\right|$$

$$
\begin{aligned}
&+\sup_{0\le t\le 1}\left|\sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}(t)\big(g(t)-g(t_{kl})\big)I(|t-t_{kl}|\le b_n)\right| \\
&\le \sup_{0\le t\le 1}\left|\sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}(t)L|t-t_{kl}|\cdot I(|t-t_{kl}|>b_n)\right| \\
&+\sup_{0\le t\le 1}\left|\sum_{k=1}^{n}\sum_{l=1}^{m_i} W_{nkl}(t)L|t-t_{kl}|\cdot I(|t-t_{kl}|\le b_n)\right| \\
&= O\left(n^{-\frac{1}{3}}\log n\right). \qquad (26)
\end{aligned}
$$

$$
D_{3n}=\left(\frac{S_n^2}{N(n)}\right)^{-1}\left[\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\tilde{X}_{ij}}{N(n)}\tilde{g}(t_{ij})\right]\le C\max_{\substack{1\le i\le n,\\ 1\le j\le m_i}}\left|\tilde{g}(t_{ij})\right|\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\|\tilde{X}_{ij}\|}{N(n)}=O\left(n^{-\frac{1}{3}}\log n\right). \qquad (27)
$$

Theorem 1 is proved.■

4.2. Proof of Theorem 2.2

$$
\begin{aligned}
&\sup_{0\le t\le 1}|\hat{g}_n(t)-g(t)|=\sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t)\big(Y_{ij}-X_{ij}^T\hat{\beta}_n\big)-g(t)\right|=\sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t)X_{ij}^T\big(\beta-\hat{\beta}_n\big)+\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t)e_{ij}-\tilde{g}(t)\right|\le \\
&\sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t)X_{ij}^T\big(\beta-\hat{\beta}_n\big)\right|+\sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t)e_{ij}\right|+\sup_{t}|\tilde{g}(t)|=E_{1n}(t)+E_{2n}(t)+E_{3n}(t). \qquad (28)
\end{aligned}
$$

From Lemma 3.3 we get

$$
E_{1n}(t)=\sup_{0\le t\le 1}\left|\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t)X_{ij}^T\big(\beta-\hat{\beta}_n\big)\right|\le\sup_{0\le t\le 1}\left\|\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t)X_{ij}\right\|\cdot\left\|\beta-\hat{\beta}_n\right\|=O\left(n^{-\frac{1}{3}}\log n\right). \qquad (29)
$$

According to Lemma 2.5 and equation (26), we can get $E_{2n}(t)=O\left(n^{-\frac{1}{3}}\log n\right), E_{3n}(t)=O\left(n^{-\frac{1}{3}}\log n\right)$. Together with equation (29), equation (11) is established. ■

Acknowledgements

We are grateful to the anonymous reviewers and the editors for their valuable comments and constructive suggestions that greatly improved the paper. This work is supported by Graduate Research and Practice Innovation Program of NUAA (xcxjh20210801).

References

[1] Gao J T, Chen X R, Zhao L C. Asymptotic normality of a class of estimators in partial linear models [J]. Acta Math. Sinica, 1994, 37(2): 256-268.

[2] Hu H, Wu L. Convergence rates of wavelet estimators in semiparametric regression models under NA samples [J]. Chinese Annals of Mathematics, Series B, 2012, 33(4): 609-624.

[3] Ping T, Lin Y, Liugen X. Asymptotic properties of estimators in partially linear single-index model for longitudinal data [J]. Acta Mathematica Scientia, 2010, 30(3): 677-687.

[4] Zhou X C, Lin J G. Semiparametric regression estimation for longitudinal data in models with martingale difference error's structure [J]. Statistics, 2013, 47(3): 521-534.

[5] Zhou X, Lin J. Strong consistency of estimators in partially linear models for longitudinal data with mixing-dependent structure [J]. Journal of Inequalities and Applications, 2011, 2011(1): 1-18.

[6] Kolmogorov A N, Rozanov Y A. On strong mixing conditions for stationary Gaussian processes [J]. Theory of Probability & Its Applications, 1960, 5(2): 204-208.

[7] Bradley R C. Equivalent mixing conditions for random fields [J]. The Annals of Probability, 1993, 21(4): 1921-1926.

[8] Wang X, Deng X, Xia F, et al. The consistency for the estimators of semiparametric regression model based on weakly dependent errors [J]. Statistical Papers, 2017, 58(2): 303-318

[9] Sung S H. Complete convergence for weighted sums of-mixing random variables [J]. Discrete Dynamics in Nature and Society, 2010, 2010.

[10] Sung S H. On the strong convergence for weighted sums of random variables [J]. Statistical Papers, 2011, 52(2): 447-454.

[11] Wu Y, Wang C, Volodin A. Limiting behavior for arrays of rowwise ρ*-mixing random variables [J]. Lithuanian Mathematical Journal, 2012, 52(2): 214-221.

[12] Wu Q, Jiang Y. Some strong limit theorems for ρ-mixing sequences of random variables [J]. Statistics & Probability Letters, 2008, 78(8): 1017-1023.

[13] Wu Q, Jiang Y. Chover-type laws of the k-iterated logarithm for ρ-mixing sequences of random variables [J]. Journal of Mathematical Analysis and Applications, 2010, 366(2): 435-443.

[14] Zhu C H. Convergence Rate of the Estimate of the Regression Function under NA Sequence [J]. College Mathematics, 2008, 24(4): 69-75.