Filomat 32:13 (2018), 4639–4654
https://doi.org/10.2298/FIL1813639Z

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Strong Consistency Rate of Estimators in Heteroscedastic Errors-in-variables Model for Negative Association Samples

Jing-jing Zhang[a], Ting Wang[a]

[a]*College of Science, University of Shanghai for Science and Technology, Shanghai 200093, PR China*

Abstract. This article is concerned with the estimating problem of heteroscedastic partially linear errors-in-variables (EV) models. We derive the strong consistency rate for estimators of the slope parameter and the nonparametric component in the case of known error variance with negative association (NA) random errors. Meanwhile, when the error variance is unknown, the strong consistency rate for the estimators of the slope parameter and the nonparametric component as well as variance function are considered for NA samples. In general, we concluded that the strong consistency rate for all estimators can achieve $o(n^{-1/4})$.

1. Introduction

Consider the following heteroscedastic partially linear EV model

$$\begin{cases} y_i = \xi_i\beta + g(t_i) + \epsilon_i, \\ x_i = \xi_i + \mu_i, \end{cases} \tag{1}$$

where $\epsilon_i = \sigma_i e_i$, $\sigma_i^2 = f(u_i)$, (ξ_i, t_i, u_i) are design points, (t_i, x_i, y_i) are observed samples, ξ_i are the potential variables cannot be observed, y_i are the response variables, and x_i are observed with measurement errors μ_i, $E\mu_i = 0$, while e_i are random errors with $Ee_i = 0$. $\beta \in \mathcal{R}$ is an unknown parameter that needs to be estimated. $h(\cdot)$ is a function defined on close interval $[0, 1]$ satisfying

$$\xi_i = h(t_i) + v_i. \tag{2}$$

where v_i are also nonrandom design points.

Model (1) and its special cases have been widely studied by many authors. Firstly, when the ξ_i can be accurately observed, $\sigma_i^2 = \sigma^2$, and the errors e_i are independent identically distribution(i.i.d). the model reduces to the general partially linear regression model, which was put forward by Engle et al. (1986). And then, when $g(t) \equiv 0$, $\sigma_i^2 = f(u_i)$, the model becomes into heteroscedastic linear model, which was extensively studied by Carroll (1982), Robinson (1987) and Carroll and Härdle (1989). In addition, when $g(t) \neq 0$ and the errors ϵ_i are i.i.d, the model (1) degenerates into partially linear EV model, which can be seen in Cui and Li (1998), Wang (1999), Liang et al. (1999) and so on. In recent years, semi-parametric EV models have

2010 *Mathematics Subject Classification*. Primary 62J12; Secondary 62E20.
Keywords. Partially linear EV model; Strong Consistency; Heteroscedastic; NA sequence; Least-squares estimator (LSE)
Received: 07 January 2018; Accepted: 01 June 2018
Communicated by Biljana Popović
Research supported by the National Natural Science Foundation of China (11701368).
Email addresses: jenny617@usst.edu.cn (Jing-jing Zhang), 15800980898@qq.com (Ting Wang)

been widely concerned. The EV models are widely applied in economy, biology and forestry. Early results of EV model can be seen in Fuller (1987), Cheng and Van Ness (1999) and Carrol (1995).

However, the independence assumption for the errors e_i in model (1) is not always appropriate in applications, especially for sequentially collected economic data, which often exhibit evident dependencies in the errors. So, for practical application, people need to weaken the restriction on independence. As we all known, the NA sequence is a weak dependent sequence, and it has extensive application in multivariate statistical analysis and systems reliability. When ξ_i is fully observed, the model (1) becomes semi-parametric model with NA samples, which has been studied by Baek and Liang (2006) for its strong consistency. However, few literature involves in the partially linear EV model for NA samples. Therefore, our paper is dedicated to this problem. We studied the strong consistency for the estimators of β, $f(\cdot)$, and $g(\cdot)$.

A finite family of random variables $\{X_i, 1 \le i \le n\}$ is said to be NA random variables if for every pair of disjoint subsets A and B of {1,2,...,n},we have

$$\mathrm{Cov}(f_1(X_i, i \in A), f_2(X_j, j \in B)) \le 0$$

whenever f_1 and f_2 are coordinatewise increasing function and such that the covariance exists. An infinite family of random variables is NA if every finite subfamily is NA.

The NA sequence was introduced by Alam and Saxena (1981); Then Joag-Dev and Proschan(1983) discovered the the character of multivariate distribution of NA sequence; Liang (2000) discovered complete convergence; Joag-Dev and Proschan(1983) discovered fundamental properties; Roussas derived asymptotic normality of the kernel estimate with a probability density function. NA sequence not only has been applied in the multivariate statistical analysis, reliability theory, seepage theory, but also in the oceans, weather, environment, risk analysis and time series analysis. In a word, the NA sequence has attracted considerable attention of scholars home and abroad recently. Therefore, this paper assumes that the error is NA sequence, which has certain theoretical significance and practical value.

The paper is organized as follows. In Section 2, we list some assumptions. The main results are given in Section 3. A simulation study is presented in section 4. Some preliminary lemmas are stated in Section 5. Proofs of the main results are provided in Sections 6.

2. Assumptions

(A0) Let $\{e_i, 1 \le i \le n\}$ be a sequence of NA random variables with mean zero, and let $\{\mu_i,\ 1 \le i \le n\}$ be a sequence of independent random variables with mean zero. $\{e_i, 1 \le i \le n\}$ is independent with $\{\mu_i, 1 \le i \le n\}$. Assume that $Ee_i^2 = 1$, $\sup_i E|e_i|^p < \infty$, for some $p > 4$, $\sup_i E|\mu_i|^p < \infty$, for some $p > 4$, $E\mu_i^2 = \Xi_\mu^2 > 0$ is known.

(A1) Let $\{v_i, 1 \le i \le n\}$ in condition (2) be a sequence satisfying

- (i) $\lim_{n\to\infty} n^{-1} \sum_{i=1}^n v_i^2 = \Sigma_0 \ \ (0 < \Sigma_0 < \infty)$;
- (ii) $\lim_{n\to\infty} \sup_n(\sqrt{n}\log n)^{-1} \cdot \max_{1\le m\le n} |\sum_{i=1}^m v_{j_i}| < \infty$, where $\{j_1, j_2, ..., j_n\}$ be a permutation of $(1, 2, ..., n)$ such that $V_{j1} \ge V_{j2} \ge ... \ge V_{jn}$.

(A2) (i) $0 < m_0 \le \min_{1\le i\le n} f(u_i) \le \max_{1\le i\le n} f(u_i) \le M_0 < \infty$;

- (ii) $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$ are continuous functions and satisfy the first-order Lipschitz condition on [0,1].

(A3) The probability weight functions $W_{nj}(t_i)$ are weight functions defined on $[0, 1]$ and satisfy

- (i) $\max_{1\le j\le n} \sum_{i=1}^n W_{nj}(t_i) = O(1)$;
- (ii) $\max_{1\le i\le n} \sum_{j=1}^n W_{nj}(t_i) I(|t_i - t_j| > n^{-1/4}) = o(n^{-1/4})$;
- (iii) $\max_{1\le i,j\le n} W_{nj}(t_i) = o(n^{-1/2}\log^{-1} n)$;

(A4) Let $\hat{W}_{ni}(\cdot)$ $(1 \le i \le n)$ be weight functions defined on [0,1]. Condition (A3) is satisfied replacing t_i and W_{ni} by u_i and $\hat{W}_{ni}$, respectively.

Remark 2.1. *Conditions (A0)-(A4) are standard regularity conditions and used commonly in the literature, see Härdle et al.(2000), Gao et al.(1994) and Chen et al.(1988);*

Remark 2.2. *Under some mild conditions, the following two weight functions satisfy hypothesis (A3):*

$$W_{ni}^{(1)}(t) = \frac{1}{h}\int_{s_i-1}^{s_i} K\Big(\frac{t-s}{h_n}\Big)ds,$$

$$W_{ni}^{(2)}(t) = K\Big(\frac{t-t_i}{h_n}\Big)\Big[\sum_{j=1}^{n} K\Big(\frac{t-t_i}{h_n}\Big)\Big]^{-1}.$$

where $s_i = (t_i + t_{i-1})/2,\ i = 1, 2, ..., n-1, s_0 = 0, s_n = 1, K(\cdot)$ is the Parzen-Rosenblatt kernel function, which we can see in Parzen(1962) or Härdle et al.(2000), and the h_n are bandwidth parameters.

3. Main Results

For model (1), we want to seek the estimators of β and $g(\cdot)$. Firstly, when the errors are homoscedastic and the ξ_i can be observed, we can apply the least squares estimation LSE method to estimate the parameter β. On the one hand, we assume the parameter β is known, and then to estimate $g(\cdot)$; for each given β, we have $g(t_i) = E(y_i - x_i\beta), 1 \le i \le n$. Therefore, based on the (x_i, t_i, y_i), we can define the estimator of $g(\cdot)$, that is $g_n^*(t, \beta) = \sum_{i=1}^{n} W_{ni}(t)(y_i - x_i\beta)$. Then, based on the model (1), we can also define the LSE of β by following formula:

$$\sum_{i=1}^{n}\Big[y_i - x_i\beta - g_n^*(t_i, \beta)\Big]^2 - \Xi_\mu^2\beta^2 = min!$$

On the other hand, under this condition of partially linear EV model, Liang et al.(1999) improved the LSE on the basis of the usual partially linear model, and employ the estimator of parameter β, as follow

$$\hat{\beta}_L = \Big[\sum_{i=1}^{n}(\tilde{x}_i^2 - \Xi_\mu^2)\Big]^{-1}\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i. \tag{3}$$

where $\tilde{x}_i = x_i - \sum_{j=1}^{n} W_{nj}(t_i)x_j$, $\tilde{y}_i = y_i - \sum_{j=1}^{n} W_{nj}(t_i)y_j$.

Secondly, when the errors are heteroscedastic, we consider two different cases according to $f(\cdot)$. If $\sigma_i^2 = f(u_i)$ are known, then the $\hat{\beta}_L$ is modified to be the weighted least-squares estimator (WLSE)

$$\hat{\beta}_{W_1} = \Big[\sum_{i=1}^{n}\sigma_i^{-2}(\tilde{x}_i^2 - \Xi_\mu^2)\Big]^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\tilde{x}_i\tilde{y}_i. \tag{4}$$

In fact, the $\sigma_i^2 = f(u_i)$ are unknown and must be estimated. In the case, we have $E[y_i - \xi_i\beta - g(t_i)]^2 = f(u_i)$, from $Ee_i^2 = 1$, Therefore, the estimator of $f(u_i)$ can be defined by

$$\hat{f}_n(u_i) = \sum_{j=1}^{n}\hat{W}_{nj}(u_i)(\tilde{y}_j - \tilde{x}_j\hat{\beta}_L)^2 - \Xi_\mu^2\hat{\beta}_L^2. \tag{5}$$

For convenience, we assume that $\min_{1\le i\le n}\hat{f}_n(u_i) > 0$. Then we can define a nonparametric estimator of σ_i^2, $\hat{\sigma}_{ni}^2 = \hat{f}_n(u_i)$. In consequence, when the errors are heteroscedastic and unknown, the WLSE of β is

$$\hat{\beta}_{W_2} = \Big[\sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}(\tilde{x}_i^2 - \Xi_\mu^2)\Big]^{-1}\sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{x}_i\tilde{y}_i. \tag{6}$$

Meanwhile, using $\hat\beta_L$, $\hat\beta_{W_1}$, $\hat\beta_{W_2}$, we can define three estimators for $g(\cdot)$:

$$\hat g_L(t) = \sum_{i=1}^n W_{ni}(t)(y_i - x_i\hat\beta_L), \tag{7}$$

$$\hat g_{W_1}(t) = \sum_{i=1}^n W_{ni}(t)(y_i - x_i\hat\beta_{W_1}), \tag{8}$$

$$\hat g_{W_2}(t) = \sum_{i=1}^n W_{ni}(t)(y_i - x_i\hat\beta_{W_2}). \tag{9}$$

In this paper, we provide some notions and a definition that will be used in the process of proof.

$$\begin{aligned}
&\tilde h_i = h(t_i) - \sum_{j=1}^n W_{nj}(t_i)h(t_j), \quad \tilde v_i = v_i - \sum_{j=1}^n W_{nj}(t_i)v_j, \quad \tilde g_i = g(t_i) - \sum_{j=1}^n W_{nj}(t_i)g(t_j),\\
&\tilde\epsilon_i = \epsilon_i - \sum_{j=1}^n W_{nj}(t_i)\epsilon_j, \quad \tilde\mu_i = \mu_i - \sum_{j=1}^n W_{nj}(t_i)\mu_j, \quad \tilde\xi_i = \xi_i - \sum_{j=1}^n W_{nj}(t_i)\xi_j,\\
&\eta_i = \epsilon_i - \mu_i\beta, \quad S_n^2 = \sum_{i=1}^n \tilde\xi_i^2, \quad T_n^2 = \sum_{i=1}^n \sigma_i^{-2}\tilde\xi_i^2,\\
&S_{1n}^2 = \sum_{i=1}^n(\tilde x_i^2 - \Xi_\mu^2), \quad T_{1n}^2 = \sum_{i=1}^n \sigma_i^{-2}(\tilde x_i^2 - \Xi_\mu^2), \quad U_{1n}^2 = \sum_{i=1}^n \hat\sigma_{ni}^{-2}(\tilde x_i^2 - \Xi_\mu^2).
\end{aligned} \tag{10}$$

Definition 3.1. *Let* $\left\{X_t, t = 0, \pm1, \pm2, \cdots\right\}$ *be a strictly stationary time series. For* $n = 1, 2, \cdots$, *define*

$$\rho(n) = \sup_{X\in L^2(F_{-\infty}^0), Y\in L^2(F_n^\infty)} \left|\mathrm{Corr}(X, Y)\right|$$

where F_i^j denotes the σ-algebra generated by $\left\{X_t, i \le t \le j\right\}$, and $L^2(F_i^j)$ consists of F_i^j-measurable random variables with finite second moment.

When $f(\cdot)$ is known, we give the strong consistency rate for LSE and WLSE of β and $g(\cdot)$.

Theorem 3.2. *Suppose that (A0)-(A3) are satisfied. Then*

(a) $\hat\beta_L - \beta = o(n^{-\frac14})$ *a.s.*

(b) $\hat\beta_{W_1} - \beta = o(n^{-\frac14})$ *a.s.*

Theorem 3.3. *Suppose that (A0)-(A3) are satisfied. For* $\forall\ t \in [0, 1]$, *we have*

(a) $\hat g_L(t) - E\hat g_L(t) = o(n^{-\frac14})$ *a.s.*

(b) $\hat g_{W_1}(t) - E\hat g_{W_1}(t) = o(n^{-\frac14})$ *a.s.*

When $f(\cdot)$ is unknown, we give the strong consistency rate for LSE and WLSE of β , $g(\cdot)$ and $f(\cdot)$.

Theorem 3.4. *Suppose that (A0)-(A4) are satisfied, where* $p > 6$ *in (A0). For* $\forall\ u \in [0, 1]$, *we have*

$$\hat f_n(u) - E\hat f_n(u) = o(n^{-\frac14}) \text{ a.s.}$$

Theorem 3.5. *Suppose that (A0)-(A4) are satisfied, where $p > 6$ in (A0). Then*

$$\hat{\beta}_{W_2} - \beta = o(n^{-\frac{1}{4}})\ a.s.$$

Theorem 3.6. *Suppose that (A0)-(A4) are satisfied, where $p > 6$ in (A0). For $\forall\ t \in [0,1]$, we have*

$$\hat{g}_{W_2}(t) - E\hat{g}_{W_2}(t) = o(n^{-\frac{1}{4}})\ a.s.$$

4. Simulation Study

In this section, we carry out a simulation to study the finite sample performance of the proposed estimators. In particular:

(i) we compare the performance of the estimators among $\hat{\beta}_L$, $\hat{\beta}_{W_1}$ and $\hat{\beta}_{W_2}$ by their mean squared errors (MSE), also, we compare the performance of the estimators among $\hat{g}_L(\cdot)$, $\hat{g}_{W_1}(\cdot)$ and $\hat{g}_{W_2}(\cdot)$ by their global mean squared errors (GMSE);

(ii) we give the boxplots for the estimators of β and $g(\cdot)$.

Observations are generated from

$$\begin{cases} y_i = \xi_i\beta + g(t_i) + \epsilon_i, \\ x_i = \xi_i + \mu_i, \quad i = 1, 2, \cdots, n, \end{cases}$$

where $\beta = 1$, $g(t) = \sin(2\pi t)$, $\sigma_i^2 = f(u_i)$, $f(u) = [1 + 0.5\cos(2\pi u)]^2$, $t_i = (i - 0.5)/n$ and $u_i = (i - 1)/n$, $\xi_i = t_i^2 + v_i$ with $v_i = \sin(i)/(n^{1/3})$. $\{\mu_i,\ 1 \le i \le n\}$ is an i.i.d. $N(0, 0.2^2)$ sequence. Following Joag-Dev and Proschan (1983), we know that $\{e_i,\ 1 \le i \le n\}$ is a multivariate normal distribution with $E(e_1, \cdots, e_n) = (0, \cdots, 0)$, $\mathrm{Cov}(e_i, e_j) = -4^{-(j-i)-1}$ $for\ i \ne j$ and $\mathrm{Var}(e_i) = 0.5^2$ $for\ 1 \le i \le n$. For the proposed estimators, the weight functions are taken as

$$W_{ni}(t) = \frac{K((t - t_i)/h_n)}{\sum_{j=1}^n K((t - t_j)/h_n)}, \qquad \hat{W}_{ni}(u) = \frac{K((u - u_i)/b_n)}{\sum_{j=1}^n K((u - u_j)/b_n)}.$$

where $K(\cdot)$ is a Gaussian kernel function, h_n and b_n are two bandwidth sequences.

4.1. The MSE for estimators of β, $g(\cdot)$ and $f(\cdot)$

In this subsection, we generate the observed data with sample sizes $n = 100$, 300 and 500 from the model above. The MSE of the estimators for β based on $M = 500$ replications are defined as

$$\mathrm{MSE}(\hat{\beta}) = \frac{1}{M}\sum_{l=1}^{M}\left[\hat{\beta}(l) - \beta_0\right]^2.$$

The GMSE of the estimators for $g(\cdot)$ is defined as

$$\mathrm{GMSE}(\hat{g}) = \frac{1}{Mn}\sum_{l=1}^{M}\sum_{k=1}^{n}\left[\hat{g}(t_k, l) - g(t_k)\right]^2.$$

We compute the MSE or GMSE for each estimators based on $M = 500$ replications and a grid of bandwidths h_n and b_n from $0.01 - 0.99$. Choose the optimal bandwidths to minimize the MSE or GMSE. The optimal bandwidths are chosen to minimize the MSE or GMSE. The smaller the MSE and GMSE are, the closer the estimators will be to the true values and the better the effects of the estimators will be. The minimum MSE or GMSE and the corresponding optimal bandwidths for the estimators are reported in Tables 1-2.

From Tables 1-2, it can be seen that: (i) for every fixed n, the $\hat{\beta}_{W1}$ have smaller MSE than that of the $\hat{\beta}_{W2}$; (ii)for every fixed n, the $\hat{g}_{W1}$ have smaller GMSE than that of the $\hat{g}_{W1}$. The estimated value β_{W1} is closer to the true value own to known $f(\cdot)$. (iii)the MSE or GMSE of all estimators decrease as the increasing of sample size n. So, our estimates are better.

Table 1: The MSE for the estimators of β and corresponding optimal bandwidths

n	MSE $\hat{\beta}_L$	h_1	MSE $\hat{\beta}_{W_1}$	h_1	MSE $\hat{\beta}_{W_2}$	h_1	h_2
100	0.1701	0.3600	0.0152	0.3900	0.0421	0.3800	0.1400
300	0.0840	0.3500	0.0046	0.3900	0.0167	0.3800	0.1500
500	0.0501	0.3600	0.0029	0.3900	0.0070	0.3800	0.1400

Table 2: The GMSE for the estimators of $g(\cdot)$ and $f(\cdot)$ and corresponding optimal bandwidths

n	GMSE $\hat{g}_L$	h_1	GMSE $\hat{g}_{W_1}$	h_1	GMSE $\hat{g}_{W_2}$	h_1	h_2	GMSE $\hat{f}_n$	h_1	h_2
100	0.0943	0.3500	0.0637	0.4000	0.0739	0.3800	0.1300	0.6155	0.4000	0.1300
300	0.0878	0.3500	0.0668	0.3900	0.0649	0.3800	0.1300	0.5837	0.3900	0.1200
500	0.0806	0.3500	0.0654	0.3900	0.0613	0.3900	0.1400	0.5652	0.4000	0.1600

4.2. Boxplots

In this subsection, we give the boxplots for the estimators. Under the condition that $f(\cdot)$ is known or unknown, we consider all estimators of β and $g(\cdot)$ taking the optimal bandwidths. In Figure 1, we give the boxplots for $\hat{\beta}_L$, $\hat{\beta}_{W_1}$ and $\hat{\beta}_{W_2}$ with $n = 100, 300$ and 500, respectively. In Figure 2, we provide the boxplots for the MSE of $\hat{g}_L(\cdot)$, $\hat{g}_{W_1}(\cdot)$ and $\hat{g}_{W_2}(\cdot)$ with $n = 100, 300$ and 500, respectively.

From Figures 1-2, one can see that: (i) the estimators $\hat{\beta}_{W_1}$ and $\hat{\beta}_{W_2}$ has better performance than $\hat{\beta}_L$; (ii)the estimators $\hat{g}_{W_1}$ and $\hat{g}_{W_2}$ has better performance than $\hat{g}_L$; (iii) for every estimator, the MSE of the estimators decrease as the increasing of sample size n. So, our estimates are better.

5. Preliminary Lemmas

In the sequel, let $c, c_1, \cdots$ and $C, C_1, \cdots$ be some finite positive constants, whose values are unimportant and may change. $a_n = O(b_n)$ means $|a_n| \le C|b_n|$, while $a_n = o(b_n)$ means $a_n/b_n \to 0$. $a^+ = \max(0, a)$, $a^- = \max(0, -a)$. And let $\{e_i,\ 1 \le i \le n\}$ be a stationary NA sequence with zero mean. Now, we introduce several lemmas, which will be used in the proof of the main results.

Lemma 5.1 (Baek and Liang (2006), Lemma 3.1). *Let $\alpha > 2$. Assume that $\{a_{ni}, 1 \le i \le n, n \ge 1\}$ is a triangular array of real numbers with $\max_{1\le i\le n} |a_{ni}| = O(n^{-1/2})$ and $\sum_{i=1}^n a_{ni}^2 = o(n^{-2/\alpha}(\log n)^{-1})$. If $\sup_i E|e_i|^p < \infty$ for some*

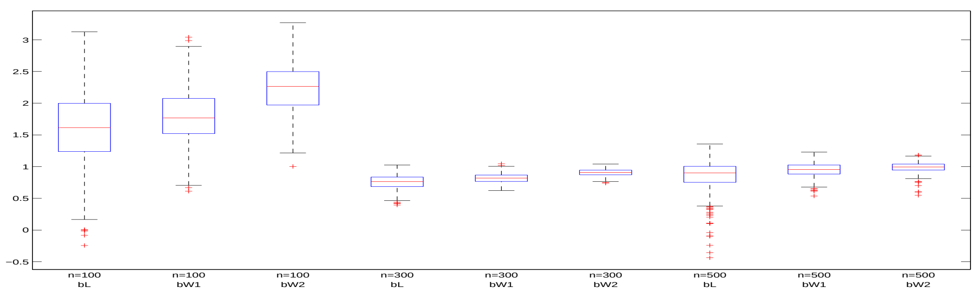

Figure 1: The boxplots for $\hat{\beta}_L$, $\hat{\beta}_{W_1}$ and $\hat{\beta}_{W_2}$ with N=500, n=100,300 and 500, respectively.

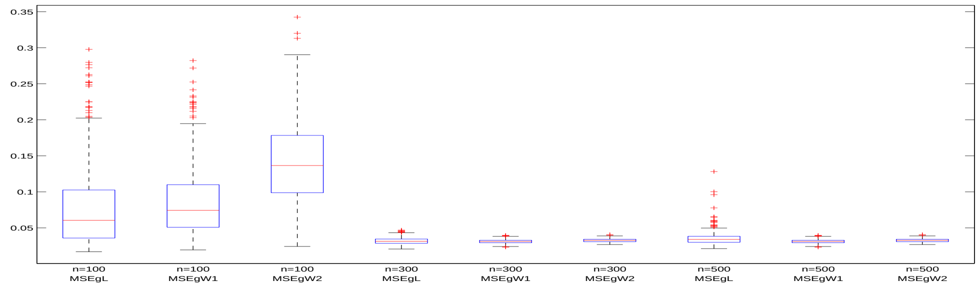

Figure 2: The boxplots of MSE for $\hat{g}_L$, $\hat{g}_{W_1}$ and $\hat{g}_{W_2}$ with N=500, n=100,300 and 500, respectively.

$p > 2\alpha/(\alpha-1)$. Then

$$\sum_{i=1}^{n} a_{ni}e_i = o(n^{-1/\alpha})\ a.s.$$

Remark 5.2. *In Lemma 5.1, it is quite clear that $p > 2$ as $\alpha \to \infty$ and $\sum_{i=1}^{n} a_{ni}e_i = o(1)$ a.s.; and $p > 4$ when $\alpha > 4$ and $\sum_{i=1}^{n} a_{ni}e_i = o(n^{-1/4})$ a.s. In addition, if all of the "o" is changed into "O", the conclusion is also right.*

Lemma 5.3 (Härdle et al. (2000), Lemma A.3). *Let $V_1, \cdots, V_n$ be independent random variables with $EV_i = 0$, finite variances and $\sup_{1\le j\le n} E|V_j|^r \le C < \infty$ $(r > 2)$. Assume that $\{a_{ki}, k, i = 1, \cdots, n\}$ is a sequence of real numbers such that $\sup_{1\le i,k\le n} |a_{ki}| = O(n^{-p_1})$ for some $0 < p_1 < 1$ and $\sum_{j=1}^{n} a_{ji} = O(n^{p_2})$ for $p_2 \ge \max(0, 2/r - p_1)$. Then*

$$\max_{1\le i\le n}\left|\sum_{k=1}^{n} a_{ki}V_k\right| = O(n^{-s}\log n)\ a.s.\ \text{for}\ s = (p_1 - p_2)/2.$$

Lemma 5.4 (Liu and Gan(2003)). *Assume a_n is a array of positive real numbers, and $\sum_{n=1}^{\infty} \sigma_n^2/a_n^2 < \infty$, where $\sigma_n^2 = \mathrm{Var}(e_n)$. If $0 < a_n \uparrow \infty$. Then*

$$\sum_{i=1}^{n} \frac{e_i}{a_n} = o(1)\ \ a.s.$$

Lemma 5.5 (Xu Bing(2002)). *Assume ϵ_i be a sequence of strong mixing, and $E\epsilon_i = 0$, when $p > 2$, $\sup_{i\ge1} E\left|\epsilon_i\right|^p < \infty$. And suppose that $\sum_{n=1}^{\infty}\left(\sum_{i=1}^{n} a_{ni}^2 \log n\right)^{p/2} < \infty$ and $\sum_{n=1}^{\infty} \alpha(n)^{(p-2)/p} < \infty$. Then*

$$\sum_{i=1}^{\infty} a_{ni}\epsilon_i = o(1)\ a.s.$$

where $\alpha(n)$ is the mixing coefficient, $\{a_{ni}, i = 1, 2, ...,\}$ are real sequence.

Following the proof line of Lemma 4.7 in Zhang and Liang (2011), one can verify the following Lemma 5.6.

Lemma 5.6. *(a) Under (A0) and (A3), we have $S_{1n}^2 \to S_n^2$ a.s.*

- *(b) Under (A1), (A2) and (A3), one can imply that $n^{-1}\sum_{i=1}^{n}\tilde{\xi}_i^2 \to \Sigma_0$, $\max_{1\le i\le n}|\tilde{\xi}_i| = o(n^{-1/2})$ and $S_n^{-2}\sum_{i=1}^{n}|\tilde{\xi}_i| \le C$;*
- *(c) Using (A1), (A2) and (A3), imply that $C_1 \le n^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\tilde{\xi}_i^2 \le C_2$ and $T_n^{-2}\sum_{i=1}^{n}|\sigma_i^{-2}\tilde{\xi}_i| \le C$;*
- *(d) Let $\tilde{A}_i = A(t_i) - \sum_{j=1}^{n} W_{nj}(t_i)A(t_j)$, where $A(\cdot) = f(\cdot)$, $g(\cdot)$ or $h(\cdot)$. Then (A2)(ii) and (A3)(ii) imply that $\max_{1\le i\le n}|\tilde{A}_i| = o(n^{-1/4})$.*

6. Proof of Main Results

In the sequel, we use the Abel Inequality (Härdle et al. (2000), page 183). Let $A_1, A_2, \cdots, A_n$; $B_1, B_2, \cdots, B_n$ ($B_1 \geq B_2 \geq \cdots \geq B_n \geq 0$) be two sequences of real numbers, and $S_k = \sum_{i=1}^{k} A_i$, $M_1 = \min_{1\leq k\leq n} S_k$, $M_2 = \max_{1\leq k\leq n} S_k$. Then, $B_1 M_1 \leq \sum_{k=1}^{n} A_k B_k \leq B_1 M_2$. Let $E_i, F_i (1 \leq i \leq n)$ to be arbitrary real numbers and $(j_1, j_2, \cdots, j_n)$ be a permutation of $(1, \cdots, n)$ such that $F_{j_1} \geq F_{j_2} \geq \cdots \geq F_{j_n}$. Then from the above equation, we have

$$\begin{aligned}\Big|\sum_{i=1}^{n} E_i F_i\Big| &= \Big|\sum_{i=1}^{n} E_{j_i} F_{j_i}\Big| \leq \Big|\sum_{i=1}^{n} E_{j_i}(F_{j_i} - F_{j_n})\Big| + \Big|\sum_{i=1}^{n} E_{j_i} F_{j_n}\Big| \\ &\leq C \max_{1\leq i\leq n} |F_i| \max_{1\leq m\leq n} \Big|\sum_{i=1}^{m} E_{j_i}\Big|.\end{aligned}$$

Proof of Theorem 3.2. We prove only *(a)*, as the proof of *(b)* is analogous. From $\big(3\big)$ and $\big(10\big)$, write that

$$\begin{aligned}\hat{\beta}_L - \beta &= S_{1n}^{-2}\Big[\sum_{i=1}^{n}(\tilde{\xi}_i + \tilde{\mu}_i)(\tilde{y}_i - \tilde{\xi}_i\beta - \tilde{\mu}_i\beta) + n\Xi_\mu^2\beta\Big] \\ &= S_{1n}^{-2}\Big\{\sum_{i=1}^{n}\Big[(\tilde{\xi}_i + \tilde{\mu}_i)(\tilde{\epsilon}_i - \tilde{\mu}_i\beta) + \Xi_\mu^2\beta\Big] + \sum_{i=1}^{n}\tilde{\xi}_i\tilde{g}_i + \sum_{i=1}^{n}\tilde{\mu}_i\tilde{g}_i\Big\} \\ &= S_{1n}^{-2}\Big\{\sum_{i=1}^{n}\tilde{\xi}_i(\epsilon_i - \mu_i\beta) + \sum_{i=1}^{n}\mu_i\epsilon_i - \sum_{i=1}^{n}(\mu^2 - \Xi_\mu^2)\beta + \sum_{i=1}^{n}\tilde{\xi}_i\tilde{g}_i \\ &\quad + \sum_{i=1}^{n}\tilde{\mu}_i\tilde{g}_i + \sum_{i=1}^{n}\sum_{j=1}^{n}W_{nj}(t_i)\tilde{\xi}_i\mu_j\beta - \sum_{i=1}^{n}\sum_{j=1}^{n}W_{nj}(t_i)\tilde{\xi}_i\epsilon_j \\ &\quad - \sum_{i=1}^{n}\sum_{j=1}^{n}W_{nj}(t_i)\epsilon_i\mu_j - \sum_{i=1}^{n}\sum_{j=1}^{n}W_{nj}(t_i)\mu_i\epsilon_j + 2\sum_{i=1}^{n}\sum_{j=1}^{n}W_{nj}(t_i)\mu_i\mu_j\beta \\ &\quad + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}W_{nj}(t_i)W_{nk}(t_i)\mu_j\epsilon_k - \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}W_{nj}(t_i)W_{nk}(t_i)\mu_j\mu_k\beta\Big\} \\ &:= S_{1n}^{-2}\sum_{k=1}^{12}A_{kn}. \qquad (11)\end{aligned}$$

Therefore, to prove $\hat{\beta}_L - \beta = o(n^{-1/4})$ *a.s.* we need to verify $S_{1n}^{-2}A_{kn} = o(n^{-1/4})$ *a.s. for* $k = 1, 2, \cdots, 12$. Using Lemma 5.6(a)(b), we only need to verify that $n^{-1}A_{kn} = o(n^{-1/4})$ *a.s.*

Step 1. Here, we prove that $S_n^{-2}A_{1n} = o(n^{-1/4})$ *a.s.*

Firstly, from (A0), we find out $\{\eta_i = \epsilon_i - \mu_i\beta, i \geq 1\}$ are sequences of NA random variables with $E\eta_i = 0$, $\sup_i E\big|\eta_i\big|^p < C\sup_i E\big|e_i\big|^p + C\sup_i E\big|\mu_i\big|^p < \infty$, for some $p > 4$. Since

$$S_n^{-2}A_{1n} = \sum_{i=1}^{n}\frac{\tilde{\xi}_i\eta_i}{S_n^2} := \sum_{i=1}^{n}B_{in}\eta_i\,.$$

$$\max_{1\leq i\leq n}\big|B_{in}\big| \leq \max_{1\leq i\leq n}\frac{\big|\tilde{\xi}_i\big|}{S_n^2} = o(n^{-\frac{3}{2}}), \qquad \sum_{i=1}^{n}B_{in}^2 = \sum_{i=1}^{n}\frac{\xi_i^2}{S_n^4} = O(n^{-1}).$$

We have $S_n^{-2}A_{1n} = o(n^{-1/4})$ *a.s.* by (A0), Lemma 5.6, Lemma 5.1 and Remark 5.2.

Step 2. We prove that $S_n^{-2}A_{2n} = o(n^{-1/4})$ *a.s.*

Since $\{\mu_i,\ i = 1, 2, ..., n\}$ is a sequence of independent random variables, $\{\epsilon_i,\ i = 1, 2, ..., n\}$ are sequences of NA random errors, independent of $\{\mu_i,\ i \geq 1\}$, and $E\mu_i = E\epsilon_i = 0$. So $\mathrm{Cov}(\mu_i\epsilon_i,\ \mu_j\epsilon_j) = 0$, then $Corr(\mu_i\epsilon_i,\ \mu_j\epsilon_j) = 0$. From definition 3.1, $\{\mu_i\epsilon_i,\ i \geq 1\}$ are sequences of ρ-mixing random variables, and the mixing coefficients $\rho(n) = 0$. In this situation, we can know ρ-mixing is also a sequence of strong mixing from Fan and Yao (2003), and we have $0 \leq \alpha(n) \leq \rho(n)/4 = 0$. Therefore, $\{\mu_i\epsilon_i,\ i \leq 1\}$ is a sequences of strong mixing random variables with the mixing coefficients $\alpha(n) = 0$. So, in Lemma 5.5, let $p = 4 + \delta$ for some $\delta > 0$ *and* $a_n = n^{-3/4}$. Then we have $\sum_{n=1}^{\infty} \left(\sum_{i=1}^{\infty} (n^{-3/4})^2 \log n \right)^{4+\delta/2} < \infty$, and $\sum_{n=1}^{\infty} \alpha(n)^{(p-2)/p} < \infty$, So

$$n^{-\frac{3}{4}} \sum_{i=1}^{n} \mu_i\epsilon_i = o(1)\ \ a.s. \qquad \frac{1}{n} \sum_{i=1}^{n} \mu_i\epsilon_i = o(n^{-\frac{1}{4}})\ \ a.s. \tag{12}$$

Thus, $S_n^{-2}A_{2n} \leq \frac{C}{n} \left| \sum_{i=1}^{n} \mu_i\epsilon_i \right| = o(n^{-1/4})$ *a.s.*

Step 3. We prove that $S_n^{-2}A_{kn} = o(n^{-1/4})$ *a.s.* for $k = 3, 4, 5, 6, 10, 12$.

By applying (A0), (A3) and Lemma 5.1 taking $\alpha = 4$ and Lemma 5.3, one can verity that

$$\sum_{i=1}^{n} \left(\zeta_i - E\zeta_i \right) = O(n^{\frac{1}{2}} \log n)\ a.s. \quad \max_{1\leq i\leq n} \left| \sum_{j=1}^{n} W_{nj}(t_i)\mu_j \right| = o(n^{-\frac{1}{4}})\ a.s. \tag{13}$$

where $\zeta_i = \left|\mu_i\right|,\ \mu_i^2\ or\ \mu_i$.

So, From Lemma 5.6(b)(d) and $\left(13\right)$, one can achieve that

$$\begin{aligned}
S_n^{-2}A_{3n} &\leq \frac{C}{n} \sum_{i=1}^{n} \left(\mu_i^2 - \Xi_\mu^2 \right)\beta = O(n^{-\frac{1}{2}} \log n)\ \ a.s. \\
S_n^{-2}A_{4n} &\leq \frac{C}{n} \Big| \sum_{i=1}^{n} \tilde{\xi}_i \tilde{g}_i \Big| \leq \frac{C}{n} \Big[\max_{1\leq i\leq n} \sum_{i=1}^{n} \left|\tilde{\xi}_i\right| \max_{1\leq i\leq n} \left|\tilde{g}_i\right| \Big] = o(n^{-\frac{1}{4}})\ \ a.s. \\
S_n^{-2}A_{5n} &\leq \frac{C}{n} \Big| \sum_{i=1}^{n} \tilde{\mu}_i \tilde{g}_i \Big| \leq \frac{C}{n} \Big[\sum_{i=1}^{n} \left|\mu_i\right| \left|\tilde{g}_i\right| + \sum_{i=1}^{n} \left|\tilde{g}_i\right| \Big| \sum_{j=1}^{n} W_{nj}(t_i)\mu_j \Big| \Big] \\
&\leq \frac{C}{n} \Big[\max_{1\leq i\leq n} \Big(\sum_{i=1}^{n} \left(\left|\mu_i\right| - E\left|\mu_i\right| \right) + \sum_{i=1}^{n} E\left|\mu_i\right| \Big) \max_{i\leq 1\leq n} \left|\tilde{g}_i\right| \\
&+ \sum_{i\leq 1\leq n} \left|\tilde{g}_i\right| \max_{1\leq i\leq n} \Big| \sum_{j=1}^{n} W_{nj}(t_i)\mu_j \Big| \Big] = o(n^{-\frac{1}{4}})\ \ a.s.
\end{aligned}$$

In the same way, from Lemma 5.6(b)(d) and $\left(13\right)$, $S_n^{-2}A_{kn} = o(n^{-1/4})$ *a.s.* for $k = 6, 10, 12$.

Step 4. Here, we prove that $S_n^{-2}A_{kn} = o(n^{-1/4})$ *a.s.* for $k = 7, 8, 9, 11$.

Firstly, the $\{\epsilon_i, i = 1, 2, ..., n\}$ is a stationary NA sequence with zero mean, let $\epsilon_i^+ = (|\epsilon_i| + \epsilon_i)/2$ and $\epsilon_i^- = (|\epsilon_i| - \epsilon_i)/2$. It's easy to know that the sequences $\{\epsilon_i^+, i = 1, 2, ...n\}$ and $\{\epsilon_i^-, i = 1, 2, ...n\}$ are all NA sequence satisfying $|\epsilon_i| = \epsilon_i^+ + \epsilon_i^-$. From Lemma 5.4 and (A0), one can get

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{n} \left|\epsilon_i\right| &= \frac{1}{n} \sum_{i=1}^{n} \epsilon_i^+ + \frac{1}{n} \sum_{i=1}^{n} \epsilon_i^- \\
&= \frac{1}{n} \sum_{i=1}^{n} \left(\epsilon_i^+ - E\epsilon_i^+ \right) + \frac{1}{n} \sum_{i=1}^{n} \left(\epsilon_i^- - E\epsilon_i^- \right) + \frac{1}{n} \sum_{i=1}^{n} E\left|\epsilon_i\right| \\
&= O(1)\ \ a.s.
\end{aligned} \tag{14}$$

Hence, by applying (A0) and (A3), Lemma 5.1 taking $\alpha = 4$, one can obtain that

$$\max_{1\le i\le n}\Big|\sum_{j=1}^{n} W_{nj}(t_i)\epsilon_j\Big| = o(n^{-\frac{1}{4}})\ a.s. \tag{15}$$

From Lemma 5.6(b), $\big(13\big)$, $\big(14\big)$ and $\big(15\big)$, we have

$$\begin{aligned} S_n^{-2}A_{7n} &\le \frac{C}{n}\sum_{i=1}^{n}\big|\tilde{\xi}_i\big|\Big|\sum_{j=1}^{n} W_{nj}(t_i)\epsilon_j\Big| = o(n^{-\frac{1}{4}})\ a.s.\\ S_n^{-2}A_{8n} &\le \frac{C}{n}\sum_{i=1}^{n}\big|\epsilon_i\big|\Big|\sum_{j=1}^{n} W_{nj}(t_i)\mu_j\Big| = o(n^{-\frac{1}{4}})\ a.s. \end{aligned}$$

The proof of $S_n^{-2}A_{kn} = o(n^{-1/4})\ a.s.$ for $k = 9, 11.$ is analogous. Thus, the proof of Theorem 3.2 is completed. ■

Proof of Theorem 3.3. We prove *(a)*, the proof for *(b)* is similar. From $\big(7\big)$, note that

$$\begin{aligned} \hat{g}_L(t) - E\hat{g}_L(t) &= \sum_{i=1}^{n} W_{ni}(t)\Big[y_i - x_i\hat{\beta}_L - Ey_i + E(x_i\hat{\beta}_L)\Big]\\ &= \sum_{i=1}^{n} W_{ni}(t)(\epsilon_i - \mu_i\beta) + \sum_{i=1}^{n} W_{ni}(t)\xi_i\Big(\beta - \hat{\beta}_L\Big)\\ &\quad - \sum_{i=1}^{n} W_{ni}(t)\xi_i E\Big(\beta - \hat{\beta}_L\Big) + \sum_{i=1}^{n} W_{ni}(t)\mu_i\Big(\beta - \hat{\beta}_L\Big)\\ &\quad - \sum_{i=1}^{n} W_{ni}(t)E\Big[\mu_i(\beta - \hat{\beta}_L)\Big]\\ &:= F_{1n}(t) + F_{2n}(t) - F_{3n}(t) + F_{4n}(t) - F_{5n}(t). \end{aligned}$$

Therefore, we only need to prove that $F_{kn}(t) = o(n^{-1/4})\ a.s.\ for\ k = 1, 2, \cdots, 5$

Firstly, We prove that $\big|\beta - \hat{\beta}_L\big| = o(n^{-1/4})\ a.s.$ and $E(\hat{\beta}_L - \beta)^2 = O(n^{-1})$. By the proof of Theorem 3.2 we have

$$\Big|\beta - \hat{\beta}_L\Big| = o(n^{-\frac{1}{4}})\ a.s. \tag{16}$$

Meanwhile, we take the same notations of A_{kn} for $k = 1, 2, \cdots, 12$ as in the proof of Theorem 3.2. Observe that

$$E\Big[S_{1n}^2(\hat{\beta}_L - \beta)\Big]^2 = E\Big[\sum_{k=1}^{12} A_{kn}\Big]^2 \le C\cdot\sum_{k=1}^{12} E(A_{kn})^2.$$

Noticing that $\{\epsilon_i - \mu_i\beta, 1 \le i \le n\}$ is a sequence of NA variables and $\{\epsilon_i\mu_i, 1 \le i \le n\}$ are sequences of α-mixing variables. Applying (A0), (A2), Lemma 5.6(a)(b) and $\big(13\big)$, one can achieve that

$$\begin{aligned}
\sup_n n^{-1}E(A_{1n})^2 &\leq \sup_n \frac{C}{n}E\Big(\sum_{i=1}^n \xi_i^2\epsilon_i^2 + \sum_{i=1}^n \xi_i^2\mu_i^2\beta^2 - 2\sum_{i=1}^n \xi_i^2\epsilon_i\mu_i\beta\Big) \\
&\leq \sup_n \frac{C}{n}\Big(\sum_{i=1}^n \xi_i^2 E\epsilon_i^2 + \sum_{i=1}^n \xi_i^2 E\mu_i^2\beta^2 + 2E\sum_{i=1}^n \Big|\xi_i^2\epsilon_i\mu_i\beta\Big|\Big) \\
&\leq \sup_n \frac{C}{n}\Bigg(O(n) + O(n) + \sqrt{\sum_{i=1}^n \xi_i^2 E\epsilon_i^2 \cdot \sum_{i=1}^n \xi_i^2 E\mu_i^2\beta^2}\Bigg) \ < \ \infty \\
\sup_n n^{-1}E(A_{2n})^2 &= \sup_n \frac{1}{n}E\Big(\sum_{i=1}^n \epsilon_i\mu_i\Big)^2 \ \leq \ \sup_n \frac{C}{n}\Big(\sum_{i=1}^n \sigma_i^2\Xi_\mu^2\Big) \ < \ \infty \\
\sup_n n^{-1}E(A_{3n})^2 &= \sup_n \frac{1}{n}\sum_{i=1}^n E\Big(\mu_i^2 - \Xi_\mu^2\Big)\beta^2 \ \leq \ \sup_n \frac{C}{n}n \ < \ \infty \\
\sup_n n^{-1}E(A_{8n})^2 &= \sup_n \frac{1}{n}E\Big\{\sum_{j=1}^n \Big[\sum_{i=1}^n W_{nj}(t_i)\epsilon_i\Big]\mu_j\Big\}^2 = \sup_n \frac{C}{n}\Big\{\sum_{j=1}^n E\Big[\sum_{i=1}^n W_{nj}(t_i)\epsilon_i\Big]^2 E\mu_j^2\Big\} \\
&\leq \sup_n \frac{C}{n}\Big[\sum_{j=1}^n \sum_{i=1}^n W_{nj}^2(t_i)E\Big(\sum_{i=1}^n \epsilon_i^2\Big)\Xi_\mu^2\Big] \leq C \cdot \sum_{j=1}^n \sum_{i=1}^n W_{nj}^2(t_i) \\
&\leq C \max_{1\leq j\leq n}\sum_{i=1}^n W_{nj}(t_i) \max_{1\leq i\leq n}\sum_{j=1}^n W_{nj}(t_i) \ < \ \infty
\end{aligned}$$

Similarly, one can deduce that $\sup_n n^{-1}E(A_{kn})^2 < \infty$ for $k = 4, 5, \cdots, 7, 9, \cdots, 12$. by (A0), (A2), (A3), Lemma 5.6(a)(b)(d), $\big(13\big)$ and $\big(15\big)$. Therefore, from Lemma 5.6(a)(b)(d), one can deduce that

$$E(\hat{\beta}_L - \beta)^2 = O(n^{-1}) \qquad E|\hat{\beta}_L - \beta| = O(n^{-1/2}) \tag{17}$$

So, from (A0), (A2), (A3), Lemma $\big(5.6\big)$, $\big(13\big)$, $\big(15\big)$, $\big(16\big)$ and $\big(17\big)$, one can get

$$\begin{aligned}
\Big|F_{1n}(t)\Big| &= \sum_{i=1}^n W_{ni}(t)\Big(\epsilon_i - \mu_i\beta\Big) = o(n^{-\frac{1}{4}}) \ \ a.s. \\
\Big|F_{2n}(t)\Big| &\leq \Big|\beta - \hat{\beta}_L\Big| \cdot \max_{1\leq i\leq n}\Big|\xi_i\Big| \cdot \sum_{i=1}^n W_{ni}(t) = o(n^{-\frac{3}{4}}) \ \ a.s. \\
\Big|F_{3n}(t)\Big| &\leq E\Big|\beta - \hat{\beta}_L\Big| \cdot \max_{1\leq i\leq n}\Big|\xi_i\Big| \cdot \sum_{i=1}^n W_{ni}(t) = o(n^{-1}) \\
\Big|F_{4n}(t)\Big| &\leq \Big|\beta - \hat{\beta}_L\Big| \cdot \max_{1\leq i\leq n}\Big|\sum_{i=1}^n W_{ni}(t)\mu_i\Big| = o(n^{-\frac{1}{2}}) \ \ a.s. \\
\Big|F_{5n}(5)\Big| &\leq \max_{1\leq i\leq n}\sum_{i=1}^n W_{ni}(t) \cdot \sqrt{E\mu_i^2 E\Big(\beta - \hat{\beta}_L\Big)^2} = O(n^{-\frac{1}{2}})
\end{aligned}$$

Thus, the proof of Theorem 3.3 is completed. ■

Proof of Theorem 3.4. From (5), write that

$$
\begin{aligned}
\hat{f}_n(u) - E\hat{f}_n(u) &= \sum_{i=1}^{n} \hat{W}_{ni}(u)\Big[(\tilde{y}_i - \tilde{x}_i\hat{\beta}_L)^2 - E(\tilde{y}_i - \tilde{x}_i\hat{\beta}_L)^2\Big] + \Xi_\mu^2(E\hat{\beta}_L^2 - \hat{\beta}_L^2) \\
&= \sum_{i=1}^{n} \hat{W}_{ni}(u)\Big\{\Big[(\tilde{\epsilon}_i - \tilde{\mu}_i\beta) + \tilde{\xi}_i(\beta - \hat{\beta}_L) + \tilde{\mu}_i(\beta - \hat{\beta}_L) + \tilde{g}_i\Big]^2 \\
&\quad -E\Big[(\tilde{\epsilon}_i - \tilde{\mu}_i\beta) + \tilde{\xi}_i(\beta - \hat{\beta}_L) + \tilde{\mu}_i(\beta - \hat{\beta}_L) + \tilde{g}_i\Big]^2\Big\} + \Xi_\mu^2(E\hat{\beta}_L^2 - \hat{\beta}_L^2) \\
&= \sum_{i=1}^{n} \hat{W}_{ni}(u)\Big[(\tilde{\epsilon}_i - \tilde{\mu}_i\beta)^2 - E(\tilde{\epsilon}_i - \tilde{\mu}_i\beta)^2\Big] + \sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i^2(\beta - \hat{\beta}_L)^2 \\
&\quad - \sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i^2 E(\beta - \hat{\beta}_L)^2 + \sum_{i=1}^{n} \hat{W}_{ni}(u)(\tilde{\mu}_i^2 - \Xi_\mu^2)(\beta - \hat{\beta}_L)^2 \\
&\quad - \sum_{i=1}^{n} \hat{W}_{ni}(u)E\Big[(\tilde{\mu}_i^2 - \Xi_\mu^2)(\beta - \hat{\beta}_L)^2\Big] + 2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i\tilde{g}_i(\beta - \hat{\beta}_L) \\
&\quad -2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i\tilde{g}_i E(\beta - \hat{\beta}_L) + 2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i(\tilde{\epsilon}_i - \tilde{\mu}_i\beta)(\beta - \hat{\beta}_L) \\
&\quad -2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i E\Big[(\tilde{\epsilon}_i - \tilde{\mu}_i\beta)(\beta - \hat{\beta}_L)\Big] + 2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i\tilde{\mu}_i(\beta - \hat{\beta}_L)^2 \\
&\quad -2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{\xi}_i E\Big[\tilde{\mu}_i(\beta - \hat{\beta}_L)^2\Big] + 2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{g}_i(\tilde{\epsilon}_i - \tilde{\mu}_i\beta) \\
&\quad +2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{g}_i\tilde{\mu}_i(\beta - \hat{\beta}_L) - 2\sum_{i=1}^{n} \hat{W}_{ni}(u)\tilde{g}_i E\Big[\tilde{\mu}_i(\beta - \hat{\beta}_L)\Big] \\
&\quad +2\sum_{i=1}^{n} \hat{W}_{ni}(u)\Big[\tilde{\epsilon}_i\tilde{\mu}_i - (\tilde{\mu}_i^2 - \Xi_\mu^2)\beta\Big](\beta - \hat{\beta}_L) \\
&\quad -2\sum_{i=1}^{n} \hat{W}_{ni}(u)E\Big\{\Big[\tilde{\epsilon}_i\tilde{\mu}_i - (\tilde{\mu}_i^2 - \Xi_\mu^2)\beta\Big](\beta - \hat{\beta}_L)\Big\} \\
&:= \sum_{k=1}^{16} I_{kn}(u).
\end{aligned}
$$

$$
\begin{aligned}
I_{1n}(u) &= \sum_{i=1}^{n} \hat{W}_{ni}(u)\Big[\Big(\epsilon_i - \mu_i\beta\Big)^2 - E\Big(\epsilon_i - \mu_i\beta\Big)^2\Big] + \sum_{i=1}^{n} \hat{W}_{ni}(u)\Big[\sum_{j=1}^{n} W_{nj}(t_i)(\epsilon_j - \mu_j\beta)\Big]^2 \\
&\quad - \sum_{i=1}^{n} \hat{W}_{ni}(u)E\Big[\sum_{j=1}^{n} W_{nj}(t_i)(\epsilon_j - \mu_j\beta)\Big]^2 \\
&\quad -2\sum_{i=1}^{n}\sum_{j=1}^{n} \hat{W}_{ni}(u)W_{nj}(t_i)(\epsilon_i - \mu_i\beta)(\epsilon_j - \mu_j\beta) \\
&\quad +2\sum_{i=1}^{n}\sum_{j=1}^{n} \hat{W}_{ni}(u)W_{nj}(t_i)E\Big[(\epsilon_i - \mu_i\beta)(\epsilon_j - \mu_j\beta)\Big] \\
&:= J_{1n}(u) + J_{2n}(u) - J_{3n}(u) - 2J_{4n}(u) + 2J_{5n}(u).
\end{aligned}
$$

Therefore, we only need to verify that $J_{tn}(u) = o(n^{-1/4})$ *a.s.* $I_{kn}(u) = o(n^{-1/4})$ *a.s.* for $t = 1, 2, \cdots, 5, k =$

$1, 2, \cdots, 16$.

Noticing that $\eta_i = \big(\epsilon_i - \mu_i\beta\big)$ and $\{\eta_i, 1 \le i \le n\}$ is a sequence of NA random variables with $E\eta_i = 0$. Then

$$\begin{aligned}\Big[\big(\epsilon_i - \mu_i\beta\big)^2 - E\big(\epsilon_i - \mu_i\beta\big)^2\Big] &= \Big[\eta_i^2 - E(\eta_i)^2\Big]\\ &= \Big\{\Big[\big(\eta_i^+\big)^2 - E\big(\eta_i^+\big)^2\Big] + \Big[\big(\eta_i^-\big)^2 - E\big(\eta_i^-\big)^2\Big]\Big\} := \varphi_i\end{aligned}$$

Therefore $\{\varphi_i,\ 1 \le i \le n\}$ is a sequence of NA random variables with mean zero and $\sup_i E|\varphi_i|^3 < \infty$. Thus, from (A0), (A3), (A4), Lemma 5.1, Lemma 5.6, $\big(15\big)$, we have

$$\begin{aligned}J_{1n}(u) &\le \max_{1\le i\le n}\Big|\sum_{i=1}^n \hat{W}_{ni}(u)\varphi_i\Big| = o(n^{-\frac{1}{4}}) \quad a.s.\\ J_{2n}(u) &\le \max_{1\le i\le n}\Big|\sum_{j=1}^n W_{nj}(t_i)(\epsilon_j - \mu_j\beta)\Big|^2 \cdot \sum_{i=1}^n \hat{W}_{ni}(u) = o(n^{-\frac{1}{2}}) \quad a.s.\\ J_{4n}(u) &\le \max_{1\le i\le n}\Big|\sum_{j=1}^n W_{nj}(t_i)(\epsilon_j - \mu_j\beta)\Big| \max_{1\le i\le n}\sum_{j=1}^n \Big|\hat{W}_{ni}(u)(\epsilon_i - \mu_i\beta)\Big| = o(n^{-\frac{1}{4}}) \quad a.s.\end{aligned}$$

Similarly, one can prove that $J_{tn}(u) = o(n^{-1/4})$ $a.s.$ for $t = 3, 5$. Meanwhile, From (A0), (A3), (A4), Lemma 5.6, $\big(13\big)$, $\big(15\big)$, $\big(16\big)$ we can deduce that

$$\begin{aligned}I_{2n}(u) &\le \big(\beta - \hat{\beta}_L\big)^2 \max_{1\le i\le n}\big|\tilde{\xi}_i\big|^2 \sum_{i=1}^n \Big|\hat{W}_{ni}(u)\Big| = o(n^{-\frac{3}{4}}) \quad a.s.\\ I_{4n}(u) &= \sum_{i=1}^n \hat{W}_{ni}(u)\Big\{(\mu_i^2 - \Xi_\mu^2)(\beta - \hat{\beta}_L)^2 - 2\mu_i \sum_{j=1}^n W_{nj}(t_i)\mu_j(\beta - \hat{\beta}_L)^2 + \Big[\sum_{j=1}^n W_{nj}(t_i)\mu_j\Big]^2(\beta - \hat{\beta}_L)^2\Big\}\\ &:= I_{41n}(u) - 2I_{42n}(u) + I_{43n}(u).\\ I_{41n}(u) &\le \sum_{i=1}^n |\mu_i^2 - \Xi_\mu^2| \cdot (\beta - \hat{\beta}_L)^2 \cdot \max_{1\le i\le n}\hat{W}_{ni}(u) = o(n^{-\frac{1}{2}}) \quad a.s.\\ I_{42n}(u) &= o(n^{-\frac{1}{2}}) \cdot \sum_{i=1}^n |\mu_i| \cdot \max_{1\le i\le n}\Big|\sum_{j=1}^n W_{nj}(t_i)\mu_j\Big| \cdot \max_{1\le i\le n}\hat{W}_{ni}(u) = o(n^{-\frac{1}{4}}) \quad a.s.\\ I_{43n}(u) &= o(n^{-\frac{1}{2}}) \cdot \sum_{i=1}^n \hat{W}_{ni}(u) \cdot \max_{1\le i\le n}\Big|\sum_{j=1}^n W_{nj}(t_i)\mu_j\Big|^2 = o(n^{-1}) \quad a.s.\\ I_{6n}(u) &\le \big|\beta - \hat{\beta}_L\big| \max_{1\le i\le n}\big|\tilde{g}_i\big| \max_{1\le i\le n}\big|\tilde{\xi}_i\big| \sum_{i=1}^n \hat{W}_{ni}(u) = o(n^{-1}) \quad a.s.\\ I_{8n}(u) &\le \big|\beta - \hat{\beta}_L\big| \max_{1\le i\le n}\big|\tilde{\xi}_i\big| \sum_{i=1}^n \hat{W}_{ni}(u)\Big|\big(\epsilon_i - \mu_i\beta\big) - \sum_{j=1}^n W_{nj}(t_i)(\epsilon_j - \mu_j\beta)\Big|\\ &\le o(n^{-\frac{3}{4}})\Big[\sum_{i=1}^n \hat{W}_{ni}(u)\Big(\big|\epsilon_i - \mu_i\beta\big| - E\big|\epsilon_i - \mu_i\beta\big| + E\big|\epsilon_i - \mu_i\beta\big|\Big)\\ &+ \max_{1\le i\le n}\sum_{i=1}^n \hat{W}_{ni}(u)\Big|\sum_{j=1}^n W_{nj}(t_i)(\epsilon_j - \mu_j\beta)\Big|\Big] = o(n^{-\frac{3}{4}}) \quad a.s.\end{aligned}$$

$$\begin{aligned} I_{10n}(u) &\le \left|\beta-\hat{\beta}_L\right|^2 \max_{1\le i\le n}\left|\tilde{\xi}_i\right| \sum_{i=1}^{n} \hat{W}_{ni}(u)\left|\mu_i-\sum_{j=1}^{n} W_{nj}(t_i)\mu_j\right| \\ &\le o(n^{-1})\Big[\sum_{i=1}^{n}\hat{W}_{ni}(u)\Big(\left|\mu_i\right|-E\left|\mu_i\right|+E\left|\mu_i\right|\Big) \\ &+ \max_{1\le i\le n}\sum_{i=1}^{n}\hat{W}_{ni}(u)\Big|\sum_{j=1}^{n}W_{nj}(t_i)\mu_i\Big|\Big] = o(n^{-\frac{3}{4}}) \quad a.s. \end{aligned}$$

In the same way, from (A0), (A3), (A4), Lemma 5.6, $\big(12\big)$, $\big(13\big)$, $\big(15\big)$, $\big(16\big)$ and $\big(17\big)$, one can get $I_{kn}(u) = o(n^{-\frac{1}{4}})$ *a.s.* for $k = 3, 5, 7, 9, 11, \cdots, 16$. Thus, the proof of Theorem 3.4 is completed. ■

Proof of Theorem 3.5. According to Theorem 3.2*(b)*, it suffices to show that $\hat{\beta}_{W_2} - \hat{\beta}_{W_1} = o(n^{-\frac{1}{4}})$ *a.s.* From $\big(4\big)$, $\big(6\big)$ and $\big(10\big)$ we have

$$\begin{aligned} \hat{\beta}_{W_1}-\hat{\beta}_{W_2} &= \frac{1}{T_{1n}^2}\sum_{i=1}^{n}\sigma_i^{-2}\tilde{x}_i(\tilde{x}_i\beta-\tilde{\mu}_i\beta+\tilde{g}_i+\tilde{\epsilon}_i)-\frac{1}{U_{1n}^2}\sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{x}_i(\tilde{x}_i\beta-\tilde{\mu}_i\beta+\tilde{g}_i+\tilde{\epsilon}_i) \\ &= \Big(\frac{1}{T_{1n}^2}-\frac{1}{U_{1n}^2}\Big)\sum_{i=1}^{n}\sigma_i^{-2}\Big[(\tilde{\xi}_i+\tilde{\mu}_i)(\tilde{g}_i+\tilde{\epsilon}_i-\tilde{\mu}_i\beta)+\Xi_\mu^2\beta\Big] \\ &\quad+\frac{1}{U_{1n}^2}\sum_{i=1}^{n}(\sigma_i^{-2}-\hat{\sigma}_{ni}^{-2})\Big[(\tilde{\xi}_i+\tilde{\mu}_i)(\tilde{g}_i+\tilde{\epsilon}_i-\tilde{\mu}_i\beta)+\Xi_\mu^2\beta\Big] \\ &= \frac{U_{1n}^2-T_{1n}^2}{U_{1n}^2T_{1n}^2}\sum_{i=1}^{n}\sigma_i^{-2}\Big[\tilde{\xi}_i\tilde{g}_i+\tilde{\xi}_i\tilde{\epsilon}_i-\tilde{\xi}_i\tilde{\mu}_i\beta+\tilde{\mu}_i\tilde{g}_i+\tilde{\mu}_i\tilde{\epsilon}_i-(\tilde{\mu}_i^2-\Xi_\mu^2)\beta\Big] \\ &\quad+\frac{1}{U_{1n}^2}\sum_{i=1}^{n}(\sigma_i^{-2}-\hat{\sigma}_{ni}^{-2})\Big[\tilde{\xi}_i\tilde{g}_i+\tilde{\xi}_i\tilde{\epsilon}_i-\tilde{\xi}_i\tilde{\mu}_i\beta+\tilde{\mu}_i\tilde{g}_i+\tilde{\mu}_i\tilde{\epsilon}_i-(\tilde{\mu}_i^2-\Xi_\mu^2)\beta\Big] \\ &:= \sum_{k=1}^{12}G_{kn}. \end{aligned}$$

Therefore, we only need to verify that $G_{kn} = o(n^{-\frac{1}{4}})$ *a.s.* for $k = 1, 2, \cdots, 12$

Step 1. we prove that

$$\max_{1\le i\le n}\left|\hat{f}_n(u_i)-f(u_i)\right| = o(n^{-\frac{1}{4}}) \; a.s. \tag{18}$$

Observe that $\left|\hat{f}_n(u_i) - f(u_i)\right| \le \left|\hat{f}_n(u_i) - E\hat{f}_n(u_i)\right| + \left|E\hat{f}_n(u_i) - f(u_i)\right|$. From Theorem 3.3 we know that $\left|\hat{f}_n(u_i) - E\hat{f}_n(u_i)\right| = o(n^{-1/4})$ *a.s.*. From the proof of the Theorem 3.4, (A0),(A3), and Lemma 5.6, one can deduce that $\left|E\hat{f}_n(u_i) - f(u_i)\right| = o(n^{-1/4})$

Step 2. Then, we prove that

$$\max_{1\le i\le n}\left|\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right| = o(n^{-\frac{1}{4}}) \; a.s. \;\; \frac{\left|U_{1n}^2-T_{1n}^2\right|}{\left|U_{1n}^2T_{1n}^2\right|} = o(n^{-\frac{5}{4}}) \;\; a.s. \;\; and \tag{19}$$

When n is large enough, we known from (A2) and $\big(18\big)$ that,

$$0 < m_1 \le \min_{1\le i\le n}\hat{f}_n(u_i) \le \max_{1\le i\le n}\hat{f}_n(u_i) \le M_1 < \infty \;\; a.s. \tag{20}$$

From $\big(18\big)$ and $\big(20\big)$

$$\max_{1\le i\le n}\left|\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right| = \max_{1\le i\le n}\left|\frac{f(u_i)-\hat{f}_{ni}(u_i)}{\hat{f}_{ni}(u_i)f(u_i)}\right| = o(n^{-\frac{1}{4}})$$

Note that from (A0), (A2), (A3) Lemma 5.6 and $\big(13\big)$, $\big(18\big)$, $\big(20\big)$, we have

$$\begin{aligned}
n^{-1}\left|U_{1n}^2-T_{1n}^2\right| &\le \max_{1\le i\le n}\left|\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right|\cdot\frac{1}{n}\sum_{i=1}^{n}\left|\tilde{x}_i^2-\Xi_\mu^2\right| \\
&= \max_{1\le i\le n}\frac{\left|\hat{f}_n(u_i)-f(u_i)\right|}{\hat{f}_n(u_i)f(u_i)}\cdot\frac{1}{n}\sum_{i=1}^{n}\left|\tilde{x}_i^2-\Xi_\mu^2\right| \\
&= o(n^{-\frac{5}{4}})\cdot\Big\{\sum_{i=1}^{n}\tilde{\xi}_i^2+\sum_{i=1}^{n}\left|\mu_i^2-\Xi_\mu^2\right|+\sum_{i=1}^{n}\Big[\sum_{j=1}^{n}W_{nj}(t_i)\mu_j\Big]^2+2\Big(\sum_{i=1}^{n}\tilde{\xi}_i^2\cdot\sum_{i=1}^{n}\mu_i^2\Big)^{1/2} \\
&\quad +2\sum_{i=1}^{n}|\tilde{\xi}_i|\cdot\Big|\sum_{j=1}^{n}W_{nj}(t_i)\mu_j\Big|+2\sum_{i=1}^{n}|\mu_i|\cdot\Big|\sum_{j=1}^{n}W_{nj}(t_i)\mu_j\Big|\Big\} = o(n^{-\frac{1}{4}})
\end{aligned}$$

On the other hand, from $\big(10\big)$, (A0), (A1), (A2), (A3), Lemma 5.6 and $\big(13\big)$, one can deduce that $n^{-1}|T_{1n}^2-T_n^2| = o(1)$ *a.s.* easily. Therefore, we get $C_3 n \le T_{1n}^2 \le C_4 n$ a.s. and $C_5 n \le U_{1n}^2 \le C_6 n$ *a.s.* So, $\big(19\big)$ is satisfied.

Step 3. Finally, we prove $G_{kn} = o(n^{-\frac{1}{4}})$ *a.s.* for $k = 1, 2, \cdots, 12$.

From Lemma 5.6, $\big(14\big)$ and $\big(19\big)$, one can deduce that

$$\begin{aligned}
G_{1n} &\le \max_{1\le i\le n}\frac{\left|U_{1n}^2-T_{1n}^2\right|}{\left|U_{1n}^2T_{1n}^2\right|}\sum_{i=1}^{n}\left|\tilde{\xi}_i\right|\max_{1\le i\le n}\left|\tilde{g}_i\right| = o(n^{-1/2})\ a.s. \\
G_{2n} &\le \max_{1\le i\le n}\frac{\left|U_{1n}^2-T_{1n}^2\right|}{\left|U_{1n}^2T_{1n}^2\right|}\sum_{i=1}^{n}\left|\tilde{\epsilon}_i\right|\max_{1\le i\le n}\left|\tilde{\xi}_i\right| = o(n^{-3/4})\ a.s.
\end{aligned}$$

In the same way, from (A0), (A2), Lemma 5.6, and $\big(12\big)$, $\big(13\big)$ and $\big(19\big)$, one can deduce that $G_{kn} = o(n^{-1/4})$ *a.s.* for $k = 3, 4, \cdots, 12$. Thus, the proof of Theorem 3.5 is completed. ■

Proof of Theorem 3.6. The proof of Theorem 3.6 is similar to the proof of Theorem 3.3 ■

References

[1] Alam, K., Saxena, K.M.L., Positive dependence in multivariate distributions, Communications in Statistics-Theory and Methods A10 (1981) 1183–1196.

[2] Baek, J. I., Liang, H. Y., Asymptotic of estimators in semi-parametric model under NA samples, Journal of Statist. Plann. Infer. 136 (2006) 3362–3382.

[3] Carroll, R.J., Adapting for heteroscedasticity in linear models, Ann. Statist 10(4) (1982) 1224–1233.

[4] Carroll, R.J., Härdle, W., Second order effects in semiparametric weighted least squares regression, Statistics 2 (1989) 179–186.

[5] Cui,H.J., Li,R.C., On parameter estimation for semi-linear errors-in-variables models, Journal of Multivariate Analysis 64(1) (1998) 1–24.

[6] Cheng, C. L., Van Ness, J. W., Statistical regression with measurement error, Arnold, London, 1999.

[7] Carroll, R.J., Ruppert, D. and Stefanski, L.A., Measurement Error in Nonlinear Models, Chapman Hall, London, 1995.

[8] Chen, H., Convergence rates for parametric components in a partly linear model, Ann. Statist 16 (1988) 136–146.

[9] Engle, R.F., Granger, C.W.J., Rice,J. and Weiss, A., Semiparametric estimation of the relation between weather and electricity sales, Journal of the American Statistical Association 81 (1986) 310–320.

[10] Fuller, W.A., Measurement Error Models, Wiley, New York, 1987.

[11] Fan, J.Q., Yao,Q.W., Nonlinear Time Series: Nonparametric and Parametric Methods, Springer-Verlag, New York, 2003.

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

[13] H.Y. Liang., Complete convergence for weighted sums of negatively associated random variables, Statist. Probab. Lett. 48 (2000) 317–325.
[14] Härdle, W., Liang, H. and Gao, J.T., Partial Linear Models, Physica-Verlag, Heidelberg, 2000.
[15] Joag-Dev, K., Proschan, F., Negative association of random variables with applications, Ann. Statist 11 (1983) 286–295.
[16] Liang,H., Hardle,W., and Carrol,R.J. Estimation in a semiparametric partially linear errosr-in-variables model, The Annals of Statistics 27(5) (1999) 1519–1935.
[17] Liu,J.J., Gan,S.X. and Chen,P.Y., The Hjeck-Rnyi inequality for the NA random variables and its application, Statistics and Probability letters 43(1) (1999) 99–105.
[18] Parzen, E., On the estimation of a probability density function and mode, Ann. Math. Statist 33 (1962) 1065–1076.
[19] Robinson, P.M., Asymptotically efficiency estimation in the presence of heteroscedasticity of unknown form, Econometrika 55 (1987) 875–891.
[20] Roussas, G.G., Asymptotic normality of the kernel estimate of a probability density function under association, Statist. Probab. Lett. 50 (2000) 1–12.
[21] Wang,Q.H., Estimation of partial linear errors-in-variables models with validation data, Journal of Multivariate Analysis 69(1) (1999) 30–64.
[22] Xu Bing., The convergence of the weighted sum for strong mixing dependent variable and its application, Journal of mathematics 45(2) (2002) 1025–1034.
[23] Zhang, J. J., Liang, H. Y., Berry-Esseen type bounds in heteroscedastic semi-parametric model, Journal of Statist. Plann. Infer. 141 (2011) 3447–3462.
[24] Zhang, J. J., Liang, H. Y., Asymptotic normality of estimators in heteroscedastic errors-in-variables model, Adv. Stat. Anal. 98 (2013) 165–195.