Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Strong consistency rates for the estimators in a heteroscedastic EV model with missing responses

Jing-Jing Zhang[1*] and Yun-Peng Xiao[1]

*Correspondence:
jenny62961700@163.com
[1]College of Science, University of Shanghai for Science and Technology, Shanghai, P.R. China

Abstract

This article is concerned with the semi-parametric error-in-variables (EV) model with missing responses: $y_i = \xi_i\beta + g(t_i) + \epsilon_i$, $x_i = \xi_i + \mu_i$, where $\epsilon_i = \sigma_i e_i$ is heteroscedastic, $f(u_i) = \sigma_i^2$, y_i are the response variables missing at random, the design points (ξ_i, t_i, u_i) are known and non-random, β is an unknown parameter, $g(\cdot)$ and $f(\cdot)$ are functions defined on closed interval $[0, 1]$, and the ξ_i are the potential variables observed with measurement errors μ_i, e_i are random errors. Under appropriate conditions, we study the strong consistent rates for the estimators of β, $g(\cdot)$ and $f(\cdot)$. Finite sample behavior of the estimators is investigated via simulations.

MSC: 62J12; 62N02; 62E99

Keywords: Semi-parametric error-in-variables model; Heteroscedastic; Missing responses; Strong consistent rate

1 Introduction

Consider the following semi-parametric error-in-variables (EV) model:

$$\begin{cases} y_i = \xi_i\beta + g(t_i) + \epsilon_i, \\ x_i = \xi_i + \mu_i, \end{cases} \tag{1.1}$$

where $\epsilon_i = \sigma_i e_i$, $\sigma_i^2 = f(u_i)$, y_i are the response variables, (ξ_i, t_i, u_i) are design points, ξ_i are the potential variables observed with measurement errors μ_i, $E\mu_i = 0$, e_i are random errors with $Ee_i = 0$ and $Ee_i^2 = 1$, β is an unknown parameter, $g(\cdot)$ and $f(\cdot)$ are functions defined on closed interval $[0, 1]$. In model (1.1), there exists a function $h(\cdot)$ defined on closed interval $[0, 1]$ satisfying

$$\xi_i = h(t_i) + \nu_i, \tag{1.2}$$

where ν_i are also design points.

Model (1.1) includes many special forms which were studied by many scholars in recent years for complete data. When $\mu_i \equiv 0$, it reduces to the general semi-parametric model, which was first introduced by Engle et al. [5] to study the effect of weather on electricity demand.

In many applications, however, there often exist covariate measurement errors. For example, it has been well documented in the literature that covariates such as blood pressure, urinary sodium chloride level, and exposure to pollutants are often subject to measurement errors, and covariate measurement errors may cause difficulties and complications in conducting statistical analysis. So EV models are somewhat more practical than ordinary regression models. When $g(\cdot) \equiv 0$, the model (1.1) reduces to the usual linear EV model. Fan et al. [6] discussed the strong consistency and asymptotic normality for the estimators of the ordinary least-square estimators. Further discussions can be found in Miao et al. [15], Miao and Liu [14], Liu and Chen [13] and so on. Model (1.1) has also been investigated by some authors. Cui and Li [4] derived the asymptotic distributions for the estimators of β, $g(\cdot)$ and error variance. Hu et al. [10] derived the asymptotic normality for the estimators of the parametric and nonparametric components; Zhou et al. [22] discussed the inference process for asymptotic distribution of estimators.

In the literature on the semi-parametric model and the EV model, scholars mainly study the situation of complete data, but in practical applications, they often encounter the situation of incomplete data. Incomplete data includes missing data, truncated data and censored data. Among them, missing data is the most common. In practical application, variables may be lost due to design or happenstance. For example, the responses y_i may be very expensive to measure and only part of the y_i are available. Another example is that the y_i represent the responses to a set of questions and some sampled individuals refuse to supply the desired information. Actually, missing of responses is very common in opinion polls, market research surveys, mail enquiries, social-economic investigations, and so on. Xu et al. [19] investigated the problem of testing nonparametric function and proposed two bias-corrected test statistics based on the quadratic conditional moment method. Yang and Xia [20] derived the asymptotically normal distribution for the restricted estimator of the parametric component. Wei and Mei [18] defined an empirical likelihood-based statistic and error-prone covariates and proved that its limiting distribution is a chi-square distribution.

Missing data includes completely random loss (MCAR), random missing (MAR) and nonrandom missing (NMAR). In this paper, we choose model (1.1) and assume that y_i is missing at random (MAR). Therefore, we obtain a random incomplete data sequence $(y_i, \delta_i, \xi_i, t_i)$. The MAR assumption implies that δ_i and y_i are conditionally independent given ξ_i and t_i. That is, $P(\delta_i = 1|y_i, \xi_i, t_i) = P(\delta_i = 1|\xi_i, t_i)$. The MAR is a common assumption for statistical analysis with missing data and is reasonable in many practical situations; see Little and Rubin [12].

When people face the loss of data, one method is to delete data with missing values. Another method is interpolation, which is based on the analysis of fully observed data and uses predicted or average values to interpolate the missing part of the data. In regression problems, commonly used imputation approaches include linear regression imputation by Healy and Westmacott [9], nonparametric kernel regression imputation by Cheng [3], semi-parametric regression imputation by Wang et al. [16], and by Wang and Sun [17]. This paper extends the methods to the estimation of β and $g(\cdot)$ under the semi-parametric EV model (1.1). We use two methods to estimate β and $g(\cdot)$ with missing responses and study the strong consistent rates for the estimators of β and $g(\cdot)$, according to $f(\cdot)$ being known or unknown.

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

2 Assumptions

In this section, we list some assumptions, which will be used in the theorems below.

(A0) Let $\{e_i, 1 \le i \le n\}$, $\{\mu_i, 1 \le i \le n\}$ and $\{\delta_i, 1 \le i \le n\}$ be independent random variables satisfying
 (i) $Ee_i = 0$, $E\mu_i = 0$, $Ee_i^2 = 1$, $E\mu_i^2 = \Xi_\mu^2 > 0$.
 (ii) $\sup_i E|e_i|^{\gamma_1} < \infty$, $\sup_i E|\mu_i|^{\gamma_2} < \infty$ from some $\gamma_1 > 8/3$, $\gamma_2 > 4$.
 (iii) $\{e_i, 1 \le i \le n\}$, $\{\mu_i, 1 \le i \le n\}$ and $\{\delta_i, 1 \le i \le n\}$ are independent of each other.

(A1) Let $\{v_i, 1 \le i \le n\}$ in (1.2) be a sequence satisfying $0 < \Sigma_i < \infty$ for $i = 0, 1, 2, 3$.
 (i) $\lim_{n\to\infty} n^{-1}\sum_{i=1}^n v_i^2 = \Sigma_0$, $\lim_{n\to\infty} n^{-1}\sum_{i=1}^n \delta_i v_i^2 = \Sigma_1$ *a.s.*
 (ii) $\lim_{n\to\infty} n^{-1}\sum_{i=1}^n \sigma_i^{-2} v_i^2 = \Sigma_2$, $\lim_{n\to\infty} n^{-1}\sum_{i=1}^n \sigma_i^{-2}\delta_i v_i^2 = \Sigma_3$ *a.s.*
 (iii) $\lim_{n\to\infty}\sup_n(\sqrt{n}\log n)^{-1}\max_{1\le m\le n}|\sum_{i=1}^m \delta_{j_i} v_{j_i}| < \infty$ *a.s.*, where $\{j_1, j_2, \dots, j_n\}$ is a permutation of $(1, 2, \dots, n)$.
 (iv) $\max_{1\le i\le n}|v_i| = O(n^{1/8})$.

(A2) (i) $0 < m \le \min_{1\le i\le n} f(u_i) \le \max_{1\le i\le n} f(u_i) \le M < \infty$.
 (ii) $f(\cdot)$, $g(\cdot)$, $h(\cdot)$ satisfy the first-order Lipschitz condition on a closed interval $[0, 1]$.

(A3) Let $W_{nj}^c(t)$ $(1 \le j \le n)$ be weight functions defined on $[0, 1]$ and satisfying
 (i) $\max_{1\le j\le n}\sum_{i=1}^n \delta_j W_{nj}^c(t_i) = O(1)$ a.s.
 (ii) $\sup_{t\in[0,1]}\sum_{j=1}^n \delta_j W_{nj}^c(t) I(|t - t_j| > a\cdot n^{-1/4}) = o(n^{-1/4})$ a.s. for any $a > 0$.
 (iii) $\sup_{t\in[0,1]}\max_{1\le j\le n} W_{nj}^c(t) = o(n^{-1/2}\log^{-1} n)$ a.s.

(A4) Let $W_{nj}(t)$ $(1 \le j \le n)$ be weight functions defined on $[0, 1]$ and satisfying
 (i) $\max_{1\le j\le n}\sum_{i=1}^n W_{nj}(t_i) = O(1)$.
 (ii) $\sup_{t\in[0,1]}\sum_{j=1}^n W_{nj}(t) I(|t - t_j| > a\cdot n^{-1/4}) = o(n^{-1/4})$ for any $a > 0$.
 (iii) $\sup_{t\in[0,1]}\max_{1\le j\le n} W_{nj}(t) = o(n^{-1/2}\log^{-1} n)$.

(A5) Let $\hat{W}_{nj}^c(u)$ $(1 \le j \le n)$ be weight functions defined on $[0, 1]$ and satisfying (A3).

Remark 2.1 Assumptions (A0)–(A5) are standard regularity conditions and they are used commonly in the literature; see Gao et al. [7], Härdle et al. [8] and Chen [2].

3 Main results

3.1 Estimation without considering heteroscedasticity

For model (1.1) without heteroscedasticity, firstly, one deletes all the missing data. Then one can get the model $\delta_i y_i = \delta_i \xi_i \beta + \delta_i g(t_i) + \delta_i \epsilon_i$. If ξ_i can be observed, we can apply the least-square estimation method to estimate the parameter β. If β is known, using the complete data $(\delta_i y_i, \delta_i x_i, \delta_i t_i)$, $1 \le i \le n$, the estimator of $g(\cdot)$, given β, is

$$g_n^*(t,\beta) = \sum_{j=1}^n W_{nj}^c(t)(\delta_j y_j - \delta_j x_j \beta). \tag{3.1}$$

Then under this condition of the semi-parametric EV model, Liang et al. [11] improved the least-square estimator (LSE) on the basis of the usual partially linear model, and employed the estimator of parameter β to minimize the following formula:

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

Therefore, one can get the LSE of β

$$\hat{\beta}_c = \left[\sum_{i=1}^{n} \delta_i \left(\tilde{x_i^c}^2 - \Xi_\mu^2 \right) \right]^{-1} \sum_{i=1}^{n} \delta_i \tilde{x_i^c} \tilde{y_i^c}, \tag{3.2}$$

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

Using $\hat{\beta}_c$, we define the following the estimators of $g(\cdot)$:

$$\hat{g}_n^c(t) = \sum_{j=1}^{n} \delta_j W_{nj}^c(t)(y_j - x_j\hat{\beta}_c). \tag{3.3}$$

Apparently, the estimators β_c and $g_n^c(t)$ are formed without taking all sample information into consideration. Hence, in order to make up for the missing data, we imply an imputation method from Wang and Sun [17], and let

$$U_i^I = \delta_i y_i + (1-\delta_i)\left[x_i\hat{\beta}_c + \hat{g}_n^c(t_i) \right]. \tag{3.4}$$

Therefore, using complete data $(U_i^I, x_i, t_i), 1 \le i \le n$, similar to (3.2) and (3.3), one can get other estimators for β and $g(\cdot)$, that is,

$$\hat{\beta}_I = \left[\sum_{i=1}^{n} \left(\tilde{x}_i^2 - \delta_i \Xi_\mu^2 \right) \right]^{-1} \sum_{i=1}^{n} \tilde{x}_i \tilde{U}_i^I. \tag{3.5}$$

Using $\hat{\beta}_I$, we define the following the estimators of $g(\cdot)$:

$$\hat{g}_n^I(t) = \sum_{j=1}^{n} W_{nj}(t)\left(U_j^I - x_j\hat{\beta}_I \right), \tag{3.6}$$

where $\tilde{U}_i^I = U_i^I - \sum_{j=1}^{n} W_{nj}(t_i)U_j^I$, $\tilde{x}_i = x_i - \sum_{j=1}^{n} W_{nj}(t_i)x_j$.

3.2 Estimation when $\sigma_i^2 = f(u_i)$ are known

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

$$\hat{\beta}_{W_1} = \left[\sum_{i=1}^{n} \sigma_i^{-2} \delta_i \left(\tilde{x_i^c}^2 - \Xi_\mu^2 \right) \right]^{-1} \sum_{i=1}^{n} \sigma_i^{-2} \delta_i \tilde{x_i^c} \tilde{y_i^c}. \tag{3.7}$$

Using $\hat{\beta}_{W_1}$, we define the following the estimators of $g(\cdot)$:

$$\hat{g}_n^{W_1}(t) = \sum_{j=1}^{n} \delta_j W_{nj}^c(t)(y_j - x_j\hat{\beta}_{W_1}). \tag{3.8}$$

Then similar to (3.4) one can make up for the missing data and let

$$U_i^{I_1} = \delta_i y_i + (1-\delta_i)\left[x_i\hat{\beta}_{W_1} + \hat{g}_n^{W_1}(t_i)\right]. \tag{3.9}$$

Therefore, using complete data $(U_i^{I_1}, x_i, t_i), 1 \le i \le n$, similar to (3.4)–(3.5), one can get other estimators for β and $g(\cdot)$, that is,

$$\hat{\beta}_{I_1} = \left[\sum_{i=1}^{n} \sigma_i^{-2}\left(\tilde{x}_i^2 - \delta_i \Xi_\mu^2\right)\right]^{-1} \sum_{i=1}^{n} \sigma_i^{-2}\tilde{x}_i \tilde{U}_i^{I_1}. \tag{3.10}$$

Using $\hat{\beta}_{I_1}$, we define the following the estimators of $g(\cdot)$:

$$\hat{g}_n^{I_1}(t) = \sum_{j=1}^{n} W_{nj}(t)\left(U_j^{I_1} - x_j\hat{\beta}_{I_1}\right), \tag{3.11}$$

where $\tilde{U}_i^{I_1} = U_i^{I_1} - \sum_{j=1}^{n} W_{nj}(t_i)U_j^{I_1}$, $\tilde{x}_i = x_i - \sum_{j=1}^{n} W_{nj}(t_i)x_j$.

Therefore, we have the following results.

Theorem 3.1 *Suppose that* (A0), (A1)(i), (ii), (iii), (A2)–(A3) *are satisfied. For every* $t \in [0,1]$, *we have*

(a) $\hat{\beta}_{W_1} - \beta = o(n^{-\frac{1}{4}})$ *a.s.*

(b) $\hat{g}_n^{W_1}(t) - g(t) = o(n^{-\frac{1}{4}})$ *a.s.*

Theorem 3.2 *Suppose that* (A0), (A1)(i), (ii), (iii), (A2)–(A4) *are satisfied. For every* $t \in [0,1]$, *we have*

(a) $\hat{\beta}_{I_1} - \beta = o(n^{-\frac{1}{4}})$ *a.s.*

(b) $\hat{g}_n^{I_1}(t) - g(t) = o(n^{-\frac{1}{4}})$ *a.s.*

3.3 Estimation when $\sigma_i^2 = f(u_i)$ are unknown

We address the case that the $\sigma_i^2 = f(u_i)$ are unknown and must be estimated. Note that, when $Ee_i^2 = 1$, we have $E[y_i - \xi_i\beta - g(t_i)]^2 = f(u_i)$. So, the estimator of $f(u_i)$ can be defined by

$$\hat{f}_n(u_i) = \sum_{j=1}^{n} \delta_j \hat{W}_{nj}^c(u_i)\left(\tilde{y}_j^c - \tilde{x}_j^c\hat{\beta}_c\right)^2 - \Xi_\mu^2\hat{\beta}_c^2. \tag{3.12}$$

For the sake of 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, $\hat{\sigma}_{ni}^2 = \hat{f}_n(u_i)$. Consequently, the WLSE of β is

$$\hat{\beta}_{W_2} = \left[\sum_{i=1}^{n} \hat{\sigma}_{ni}^{-2}\delta_i\left(\tilde{x}_i^{c2} - \Xi_\mu^2\right)\right]^{-1} \sum_{i=1}^{n} \hat{\sigma}_{ni}^{-2}\delta_i\tilde{x}_i^c\tilde{y}_i^c. \tag{3.13}$$

Using $\hat{\beta}_{W_2}$, we define the following estimators of $g(\cdot)$:

$$\hat{g}_n^{W_2}(t) = \sum_{j=1}^{n} \delta_j W_{nj}^c(t)(y_j - x_j\hat{\beta}_{W_2}). \tag{3.14}$$

Similarly, one can make up for the missing data and let

$$U_i^{I_2} = \delta_i y_i + (1-\delta_i)\big[x_i\hat{\beta}_{W_2} + \hat{g}_n^{W_2}(t_i)\big]. \tag{3.15}$$

Therefore, using complete data $(U_i^{I_2}, x_i, t_i), 1 \le i \le n$, one can get other estimators for β and $g(\cdot)$, that is,

$$\hat{\beta}_{I_2} = \left[\sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\big(\tilde{x}_i^2 - \delta_i \Xi_\mu^2\big)\right]^{-1}\sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{x}_i\tilde{U}_i^{I_2}. \tag{3.16}$$

Using $\hat{\beta}_{I_2}$, we define the following the estimators of $g(\cdot)$:

$$\hat{g}_n^{I_2}(t) = \sum_{j=1}^{n} W_{nj}(t)\big(U_j^{I_2} - x_j\hat{\beta}_{I_2}\big), \tag{3.17}$$

where $\tilde{U}_i^{I_2} = U_i^{I_2} - \sum_{j=1}^{n} W_{nj}(t_i)U_j^{I_2}$, $\hat{\sigma}_{ni}^2 = \hat{f}_n(u_i)$.

Therefore, we have the following results.

Theorem 3.3 *Suppose that* (A0)–(A3), (A5) *are satisfied with* $\gamma_1 > 16/3$ *and* $\gamma_2 > 16/3$ *in* (*A*0). *For every* $u \in [0,1]$, *we have* $\hat{f}_n(u) - f(u) = o(n^{-1/4})$ *a.s.*

Theorem 3.4 *Suppose that* (A0)–(A3), (A5) *are satisfied with* $\gamma_1 > 16/3$ *and* $\gamma_2 > 16/3$ *in* (A0). *For every* $t \in [0,1]$, *we have*

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

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

Theorem 3.5 *Suppose that* (A0)–(A5) *are satisfied with* $\gamma_1 > 16/3$ *and* $\gamma_2 > 16/3$ *in* (A0). *For every* $t \in [0,1]$, *we have*

(a) $\hat{\beta}_{I_2} - \beta = o(n^{-\frac{1}{4}})$ *a.s.*

(b) $\hat{g}_n^{I_2}(t) - g(t) = o(n^{-\frac{1}{4}})$ *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:

(1) we compare the performance of the estimators $\hat{\beta}_{W_1}$, $\hat{\beta}_{I_1}$, $\hat{\beta}_{W_2}$ and $\hat{\beta}_{I_2}$ by their mean squared errors (MSE), also, we compare the performance of the estimators $\hat{g}_n^{W_1}(\cdot)$, $\hat{g}_n^{I_1}(\cdot)$, $\hat{g}_n^{W_2}(\cdot)$ and $\hat{g}_n^{I_2}(\cdot)$ by their global mean squared errors (GMSE);

(2) we give the boxplots for the estimators of β and $g(t_{n/2})$;

(3) we give the fitting figure for the estimators of $g(\cdot)$.

Observations are generated from

$$\begin{cases} y_i = \xi_i\beta + g(t_i) + \epsilon_i, \\ x_i = \xi_i + \mu_i, \end{cases} \tag{4.1}$$

where $\beta = 1$, $g(t) = \sin(2\pi t)$, $\epsilon_i = \sigma_i e_i$, $\sigma_i^2 = f(u_i)$, $f(u) = [1 + 0.5\cos(2\pi u)]^2$, $t_i = (i-0.5)/n$, $u_i = (i-1)/n$ and $\xi_i = t_i^2 + \nu_i$. $\{\nu_i, 1 \le i \le n\}$ is an i.i.d. $N(0,1)$ sequence. $\{\mu_i, 1 \le i \le n\}$

Table 1 The MSE for the estimators of β

n	p	$\hat{\beta}_{W_1}$	$\hat{\beta}_{I_1}$	$\hat{\beta}_{W_2}$	$\hat{\beta}_{I_2}$
50	0.1	0.0129	0.0104	0.0123	0.0091
100	0.1	0.0053	0.0041	0.0047	0.0041
200	0.1	0.0024	0.0020	0.0029	0.0021
50	0.25	0.0166	0.0131	0.0189	0.0146
100	0.25	0.0086	0.0070	0.0096	0.0074
200	0.25	0.0045	0.0032	0.0047	0.0036
50	0.5	0.0201	0.0167	0.0226	0.0195
100	0.5	0.0132	0.0114	0.0158	0.0110
200	0.5	0.0058	0.0041	0.0050	0.0039

is an i.i.d. $N(0,0.2^2)$ sequence. $\{e_i, 1 \le i \le n\}$ is an i.i.d. $N(0,1)$ sequence. $\{\nu_i, 1 \le i \le n\}$, $\{\mu_i, 1 \le i \le n\}$ and $\{e_i, 1 \le i \le n\}$ are independent. d_i is a $B(1, 1-p)$ sequence where the missing probability $p = 0.1, 0.25, 0.5$. For the proposed estimators, the weight functions are taken as

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

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

4.1 Compare the estimators for β and $g(\cdot)$

Because otherwise there would be too much computation, we have to take a small sample size for convenience of the simulation. We generate the observed data with sample size n=50, 100 and 200 from the model above. The MSE of the estimators for β based on M = 100 replications are defined as

$$\mathrm{MSE}(\hat{\beta}) = \frac{1}{M}\sum_{l=1}^{M}\bigl[\hat{\beta}(l) - \beta\bigr]^2,$$

where $\hat{\beta}(l)$ is the lth estimator of β. The GMSE of the estimators for $g(\cdot)$ are defined as

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

where $\hat{g}(t_k, l)$ is the lth estimator of $g(t_k)$.

It is well known that an important issue is the selection of an appropriate bandwidth sequence. The common methods are grid point and cross-validation. Here we use the grid point method to select optimal bandwidths. The bandwidth sequences h_n, b_n, l_n are taken uniformly over 50 points with step length of 0.02 on the closed interval $[0,1]$. Then we calculate the MSE for the estimators of β and the GMSE for the estimators $g(\cdot)$ for each (h_n, b_n, l_n) and select optimal bandwidths to minimize the MSE for the estimators of β and the GMSE for the estimators $g(\cdot)$. The MSE or GMSE for the estimators are reported in Tables 1–2. On the other hand, we give the boxplots for the estimators of β and $g(t_{n/2})$ with $n = 50, 100, 200$ and $p = 0.25$.

From Tables 1–2 and Fig. 1, it can be seen that:

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

n	p	$\hat{g}_n^{W_1}(\cdot)$	$\hat{g}_n^{I_1}(\cdot)$	$\hat{g}_n^{W_2}(\cdot)$	$\hat{g}_n^{I_2}(\cdot)$	$\hat{f}_n(\cdot)$
50	0.1	0.1199	0.1134	0.1112	0.1028	0.2816
100	0.1	0.0646	0.0577	0.0557	0.0547	0.1563
200	0.1	0.0401	0.0397	0.0374	0.0346	0.0945
50	0.25	0.1504	0.1335	0.1314	0.1304	0.3413
100	0.25	0.0795	0.0785	0.0752	0.0705	0.2048
200	0.25	0.0510	0.0505	0.0461	0.0434	0.1011
50	0.5	0.1700	0.1580	0.1718	0.1525	0.3693
100	0.5	0.1134	0.1082	0.1027	0.1018	0.2305
200	0.5	0.0583	0.0568	0.0531	0.0508	0.1550

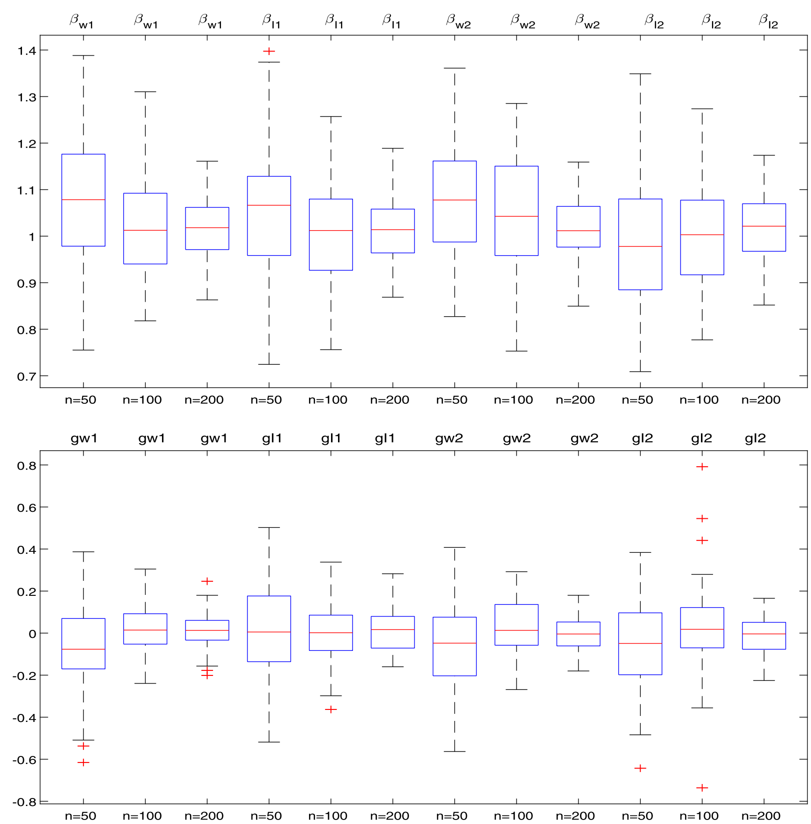


Figure 1 $p = 0.25$ and $n = 50, 100, 200$,The boxplots for the estimators of β and $g(\cdot)$

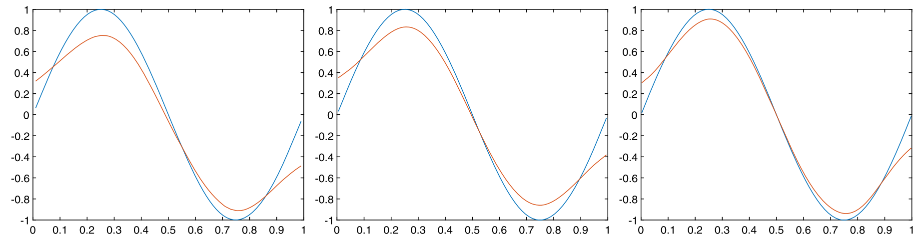
(i) For every fixed n and p, the MSE of $\hat{\beta}_{I_1}$ and $\hat{\beta}_{I_2}$ are smaller than that of $\hat{\beta}_{W_1}$ and $\hat{\beta}_{W_2}$, the GMSE of $\hat{g}_n^{I_1}(\cdot)$ and $\hat{g}_n^{I_2}(\cdot)$ are smaller than that of $\hat{g}_n^{W_1}(\cdot)$ and $\hat{g}_n^{W_2}(\cdot)$. It shows that the interpolation method is more effective than the delection method.

(ii) For every fixed n and p, the MSE of $\hat{\beta}_{W_2}$ and $\hat{\beta}_{I_2}$ are very close to that of $\hat{\beta}_{W_1}$ and $\hat{\beta}_{I_1}$, the GMSE of $\hat{g}_n^{W_2}(\cdot)$ and $\hat{g}_n^{I_2}(\cdot)$ are close to that of $\hat{g}_n^{W_1}(\cdot)$ and $\hat{g}_n^{I_2}(\cdot)$.

(iii) For every fixed n, the MSE for the estimators of β and the GMSE for the estimators of $g(\cdot)$ increase as the increasing of p.

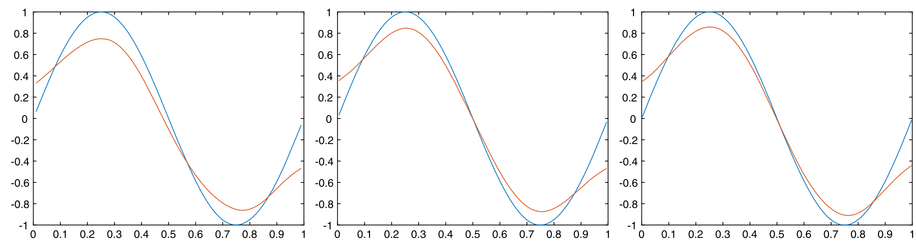
(iv) For every fixed p, the MSE for the estimators of β and the GMSE for the estimators of $g(\cdot)$ all decrease as the increasing of n.

(v) Fig. 1 shows that the variances of the estimators decrease on increasing of sample size n.



(a) n=50 (b) n=100 (c) n=200

Figure 2 $p = 0.25$, $n = 50$, 100 and 200, the fitting figure for $\hat{g}_n^{I_1}$



(a) n=50 (b) n=100 (c) n=200

Figure 3 $p = 0.25$, $n = 50$, 100 and 200, the fitting figure for $\hat{g}_n^{I_2}$

(vi) The simulation results are consistent with the theoretical results.

4.2 The fitting figure for the estimators of $g(\cdot)$

In this section, we give the fitting figure of $\hat{g}_n^{I_1}(\cdot)$ and $\hat{g}_n^{I_2}(\cdot)$ with $p = 0.25$. From Figs. 2–3, one can see that

(i) for every fixed n, the graph for the estimators of $g(\cdot)$ is very close to $g(\cdot)$;
(ii) for every fixed n, the graph of $\hat{g}_n^{I_1}(\cdot)$ is very close to $\hat{g}_n^{I_2}(\cdot)$;
(iii) the fitting effect is better on the increase of n;
(iv) when n reaches 200, the fitting effect is ideal;
(v) the simulation results are consistent with the theoretical results.

5 Preliminary lemmas

In the sequel, let $C, C_1, C_2, \ldots$ be some finite positive constants, whose values are unimportant and may change. Now, we introduce several lemmas, which will be used in the proof of the main results.

Lemma 5.1 (Baek and Liang [1], Lemma 3.1) *Let $\alpha > 2$, $e_1, \ldots, e_n$ be independent random variables with $Ee_i = 0$. Assume that $\{a_{ni}, 1 \le i \le n, n \ge 1\}$ is a triangular array of numbers with $\max_{1\le i\le n}|a_{ni}| = O(n^{-1/2})$ and $\sum_{i=1}^n a_{ni}^2 = o(n^{-2/\alpha}(logn)^{-1})$. If $\sup_i E|e_i|^\gamma < \infty$ for some $\gamma > 2\alpha/(\alpha-1)$, then*

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

Lemma 5.2 (Härdle et al. [8], Lemma A.3) *Let $V_1, \ldots, V_n$ be independent random variables with $EV_i = 0$, and $\sup_{1\le j\le n} E|V_j|^r \le C < \infty (r > 2)$. Assume that $\{a_{ki}, k, i = 1, \ldots, n\}$*

is a sequence of 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*

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

Following the proof line of Lemma 4.7 in Zhang and Liang [21], one can verify the following two lemmas.

Lemma 5.3

(a) *Let* $\tilde{A}_i = A(t_i) - \sum_{j=1}^{n} W_{nj}(t_i)A(t_j)$, *where* $A(\cdot) = g(\cdot)$ *or* $h(\cdot)$. *Let* $\tilde{A}_i^c = A(t_i) - \sum_{j=1}^{n} \delta_j W_{nj}^c(t_i)A(t_j)$, *where* $A(\cdot) = g(\cdot)$ *or* $h(\cdot)$. *Then* (A0)–(A4) *imply that* $\max_{1\le i\le n}|\tilde{A}_i| = o(n^{-1/4})$ *and* $\max_{1\le i\le n}|\tilde{A}_i^c| = o(n^{-1/4})$ *a.s.*

(b) (A0)–(A4) *imply that* $n^{-1}\sum_{i=1}^{n}\tilde{\xi}_i^2 \to \Sigma_0$, $\sum_{i=1}^{n}|\tilde{\xi}_i| \le C_1 n$, $n^{-1}\sum_{i=1}^{n}\delta_i(\tilde{\xi}_i^c)^2 \to \Sigma_1$ *a.s. and* $\sum_{i=1}^{n}|\delta_i\tilde{\xi}_i^c| \le C_2 n$ *a.s.*

(c) (A0)–(A4) *imply that* $n^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\tilde{\xi}_i^2 \to \Sigma_2$, $\sum_{i=1}^{n}|\sigma_i^{-2}\tilde{\xi}_i| \le C_3 n$, $n^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\delta_i(\tilde{\xi}_i^c)^2 \to \Sigma_3$ *a.s. and* $\sum_{i=1}^{n}|\sigma_i^{-2}\delta_i\tilde{\xi}_i^c| \le C_4 n$ *a.s.*

(d) (A0)–(A4) *imply that* $\max_{1\le i\le n}|\tilde{\xi}_i| = O(n^{1/8})$ *and* $\max_{1\le i\le n}|\tilde{\xi}_i^c| = O(n^{1/8})$ *a.s.*

(e) (A0)–(A4) *imply that* $\max_{1\le i\le n}|\sigma_i^{-2}\tilde{\xi}_i| = O(n^{1/8})$ *and* $\max_{1\le i\le n}|\sigma_i^{-2}\tilde{\xi}_i^c| = O(n^{1/8})$ *a.s.*

Lemma 5.4

(a) *Suppose that* (A0)–(A4) *are satisfied. Then one can deduce that*

$$\max_{1\le i\le n}\bigl|\hat{g}_n^{W_1}(t_i) - g(t_i)\bigr| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.}$$

(b) *Suppose that* (A0)–(A4) *are satisfied. Then one can deduce that*

$$\max_{1\le i\le n}\bigl|\hat{g}_n^{W_2}(t_i) - g(t_i)\bigr| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.}$$

One can easily get Lemma 5.3 by (A0)–(A4). The proof of Lemma 5.4 is analogous to the proof of Theorem 3.1(b) and Theorem 3.2(b).

6 Proof of main results

Now, we introduce some notations which will be used in the proofs below.

$$\tilde{\xi_i^c} = \xi_i - \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\xi_j, \qquad \tilde{\mu_i^c} = \mu_i - \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j,$$

$$\tilde{g_i^c} = g(t_i) - \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)g(t_j),$$

$$\tilde{\epsilon_i^c} = \epsilon_i - \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\epsilon_j, \qquad \tilde{\xi}_i = \xi_i - \sum_{j=1}^{n} W_{nj}(t_i)\xi_j, \qquad \tilde{\mu}_i = \mu_i - \sum_{j=1}^{n} W_{nj}(t_i)\mu_j,$$

$$\tilde{g}_i = g(t_i) - \sum_{j=1}^{n} W_{nj}(t_i)g(t_j), \qquad \tilde{\epsilon}_i = \epsilon_i - \sum_{j=1}^{n} W_{nj}(t_i)\epsilon_j, \qquad S_n^2 = \sum_{i=1}^{n}\sigma_i^{-2}\tilde{\xi}_i^2,$$

$$S_{1n}^2=\sum_{i=1}^n\sigma_i^{-2}\big(\delta_i\tilde{x}_i^{c^2}-\delta_i\Xi_\mu^2\big),\qquad S_{2n}^2=\sum_{i=1}^n\sigma_i^{-2}\big(\tilde{x}_i^2-\delta_i\Xi_\mu^2\big).$$

Proof of Theorem 3.1(*a*) From (3.7), one can write

$$\begin{aligned}
\hat{\beta}_{W_1}-\beta&=S_{1n}^{-2}\Bigg[\sum_{i=1}^n\sigma_i^{-2}\delta_i\big(\tilde{\xi}_i^c+\tilde{\mu}_i^c\big)\tilde{y}_i^c-\sum_{i=1}^n\sigma_i^{-2}\delta_i\big(\big(\tilde{\xi}_i^c+\tilde{\mu}_i^c\big)^2-\Xi_\mu^2\big)\beta\Bigg]\\
&=S_{1n}^{-2}\Bigg[\sum_{i=1}^n\sigma_i^{-2}\delta_i\big(\tilde{\xi}_i^c+\tilde{\mu}_i^c\big)\big(\tilde{g}_i^c+\tilde{\epsilon}_i^c-\tilde{\mu}_i^c\beta\big)+\sum_{i=1}^n\sigma_i^{-2}\delta_i\Xi_\mu^2\beta\Bigg]\\
&=S_{1n}^{-2}\Bigg\{\sum_{i=1}^n\sigma_i^{-2}\delta_i\tilde{\xi}_i^c(\epsilon_i-\mu_i\beta)+\sum_{i=1}^n\sigma_i^{-2}\delta_i\mu_i\epsilon_i-\sum_{i=1}^n\sigma_i^{-2}\delta_i\big(\mu_i^2-\Xi_\mu^2\big)\beta\\
&\quad+\sum_{i=1}^n\sigma_i^{-2}\delta_i\tilde{\xi}_i^c\tilde{g}_i^c+\sum_{i=1}^n\sigma_i^{-2}\delta_i\tilde{\mu}_i^c\tilde{g}_i^c-\sum_{i=1}^n\sum_{j=1}^n\sigma_i^{-2}\delta_i\delta_jW_{nj}^c(t_i)\tilde{\xi}_i^c\epsilon_j\\
&\quad-\sum_{i=1}^n\sum_{j=1}^n\sigma_i^{-2}\delta_i\delta_jW_{nj}^c(t_i)\epsilon_i\mu_j-\sum_{i=1}^n\sum_{j=1}^n\sigma_i^{-2}\delta_i\delta_jW_{nj}^c(t_i)\mu_i\epsilon_j\\
&\quad+\sum_{i=1}^n\sum_{j=1}^n\sigma_i^{-2}\delta_i\delta_jW_{nj}^c(t_i)\tilde{\xi}_i^c\mu_j\beta+2\sum_{i=1}^n\sum_{j=1}^n\sigma_i^{-2}\delta_i\delta_jW_{nj}^c(t_i)\mu_i\mu_j\beta\\
&\quad+\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sigma_i^{-2}\delta_i\delta_j\delta_kW_{nj}^c(t_i)W_{nk}^c(t_i)\mu_j\epsilon_k\\
&\quad-\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sigma_i^{-2}\delta_i\delta_j\delta_kW_{nj}^c(t_i)W_{nk}^c(t_i)\mu_j\mu_k\beta\Bigg\}\\
&:=S_{1n}^{-2}\sum_{k=1}^{12}A_{kn}.
\end{aligned}\tag{6.1}$$

Thus, to prove $\hat{\beta}_{W_1}-\beta=o(n^{-1/4})$ a.s., we only need to verify that $S_{1n}^{-2}\le Cn^{-1}$ a.s. and $n^{-1}A_{kn}=o(n^{-1/4})$ a.s. for $k=1,2,\ldots,12$.

Step 1. We prove $S_{1n}^{-2}\le Cn^{-1}$ a.s. Note that

$$\begin{aligned}
S_{1n}^2&=\sum_{i=1}^n\sigma_i^{-2}\delta_i\big(\tilde{x}_i^{c^2}-\Xi_\mu^2\big)=\sum_{i=1}^n\big[\sigma_i^{-2}\delta_i\big(\tilde{\xi}_i^c+\tilde{\mu}_i^c\big)^2-\sigma_i^{-2}\delta_i\Xi_\mu^2\big]\\
&=\sum_{i=1}^n\Bigg\{\sigma_i^{-2}\delta_i\Bigg[\tilde{\xi}_i^c+\mu_i-\sum_{j=1}^n\delta_jW_{n_j}^c(t_i)\mu_j\Bigg]^2-\sigma_i^{-2}\delta_i\Xi_\mu^2\Bigg\}\\
&=\sum_{i=1}^n\sigma_i^{-2}\delta_i\big(\tilde{\xi}_i^c\big)^2+\sum_{i=1}^n\sigma_i^{-2}\delta_i\big(\mu_i^2-\Xi_\mu^2\big)+\sum_{i=1}^n\sigma_i^{-2}\delta_i\Bigg[\sum_{j=1}^n\delta_jW_{nj}^c(t_i)\mu_j\Bigg]^2\\
&\quad+2\sum_{i=1}^n\sigma_i^{-2}\delta_i\tilde{\xi}_i^c\mu_i-2\sum_{i=1}^n\sigma_i^{-2}\delta_i\tilde{\xi}_i^c\sum_{j=1}^n\delta_jW_{nj}^c(t_i)\mu_j-2\sum_{i=1}^n\sigma_i^{-2}\delta_i\mu_i\sum_{j=1}^n\delta_jW_{nj}^c(t_i)\mu_j\\
&:=B_{1n}+B_{2n}+B_{3n}+B_{4n}+B_{5n}+B_{6n}.
\end{aligned}$$

By Lemma 5.3(c), we have $n^{-1}B_{1n} \to \Sigma_1$ a.s. Next, we verify that $B_{kn} = o(B_{1n}) = o(n)$ a.s. for $k = 2, 3, \ldots, 6$. Applying (A0), taking $r > 2$, $p_1 = 1/2$, $p_2 = 1/2$ in Lemma 5.2, we have

$$\sum_{i=1}^{n}(\zeta_i - E\zeta_i) = n^{\frac{1}{2}} \cdot \sum_{i=1}^{n} n^{-\frac{1}{2}}(\zeta_i - E\zeta_i) = O\bigl(n^{\frac{1}{2}} \log n\bigr) = o(n) \quad \text{a.s.,} \tag{6.2}$$

where ζ_i are independent random variables satisfying $E\zeta_i = 0$ and $\sup_{1\le i\le n} E|\zeta_i|^r < \infty$. Therefore, we obtain $B_{2n} = O(n^{1/2}\log n) = o(n)$ a.s. from (A0) and (6.2). On the other hand, taking $\alpha = 4$, $\gamma > 8/3$ in Lemma 5.1, we have

$$\max_{1\le i\le n}\left|\sum_{j=1}^{n} \delta_j W_{nj}^c(t_i)\zeta_j\right| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.,} \qquad \max_{1\le i\le n}\left|\sum_{j=1}^{n} W_{nj}(t_i)\zeta_j\right| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.,} \tag{6.3}$$

where the ζ_i are independent random variables satisfying $E\zeta_i = 0$ and $\sup_{1\le i\le n} E|\zeta_i|^\gamma < \infty$, for some $\gamma > 8/3$. It also holds if one replace $W_{nj}^c(t_i)$ with $\hat{W}_{nj}^c(u_i)$. Meanwhile, by (A0) and Lemma 5.2, taking $r = p > 2$, $p_1 = 1/4$, $p_2 = 3/4$ in Lemma 5.2, one can also deduce that

$$|B_{4n}| = 2n^{\frac{1}{4}} \cdot \left|\sum_{i=1}^{n} n^{-\frac{1}{4}}\sigma_i^{-2}\delta_i\tilde{\xi}_i^c\mu_i\right| = O\bigl(n^{\frac{1}{2}}\log n\bigr) = o(n) \quad \text{a.s.} \tag{6.4}$$

Note that, from Lemma 5.3, (6.2) and (6.3), we have

$$|B_{3n}| \le \sum_{i=1}^{n}\bigl|\sigma_i^{-2}\delta_i\bigr| \cdot \max_{1\le i\le n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j\right|^2 = o\bigl(n^{\frac{1}{2}}\bigr) \quad \text{a.s.,} \tag{6.5}$$

$$|B_{5n}| \le 2\sum_{i=1}^{n}\bigl|\sigma_i^{-2}\delta_i\tilde{\xi}_i^c\bigr| \cdot \max_{1\le i\le n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j\right| = o\bigl(n^{\frac{3}{4}}\bigr) \quad \text{a.s.,} \tag{6.6}$$

$$\begin{aligned}|B_{6n}| &\le 2\left[\sum_{i=1}^{n}\bigl(\bigl|\sigma_i^{-2}\delta_i\mu_i\bigr| - E\bigl|\sigma_i^{-2}\delta_i\mu_i\bigr|\bigr) + \sum_{i=1}^{n}E\bigl|\sigma_i^{-2}\delta_i\mu_i\bigr|\right] \cdot \max_{1\le i\le n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j\right| \\ &= o\bigl(n^{\frac{3}{4}}\bigr) \quad \text{a.s.}\end{aligned} \tag{6.7}$$

Therefore, for (6.2)–(6.7), one can deduce that $S_{1n}^2 = B_{1n} + o(n) = B_{1n} + o(B_{1n})$ a.s., which yields

$$\lim_{n\to\infty}\frac{B_{1n}}{S_{1n}^2} = \lim_{n\to\infty}\frac{B_{1n}}{B_{1n} + o(B_{1n})} = 1 \quad \text{a.s.}$$

Therefore, by Lemma 5.3(b), we get $S_{1n}^{-2} \le Cn^{-1}$ a.s.

Step 2. We verify that $n^{-1}A_{kn} = o(n^{-1/4})$ a.s. for $k = 1, 2, \ldots, 12$. From (A0), we find $\{\eta_i = \epsilon_i - \mu_i\beta, 1 \le i \le n\}$ are sequences of independent random variables with $E\eta_i = 0$, $\sup_i E|\eta_i|^p \le C\sup_i E|\epsilon_i|^p + C\sup_i E|\mu_i|^p < \infty$. Similar to (6.4), we deduce that

$$n^{-1}A_{1n} = n^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\delta_i\tilde{\xi}_i^c\eta_i = O\bigl(n^{-\frac{1}{2}}\log n\bigr) = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.} \tag{6.8}$$

Similar to the proofs of (6.2)–(6.3), one can easily deduce that

$$\frac{1}{n}\sum_{i=1}^{n}|\omega_i| = \frac{1}{n}\left[\sum_{i=1}^{n}\bigl(|\omega_i| - E|\omega_i|\bigr) + \sum_{i=1}^{n}E|\omega_i|\right] = O(1) \quad \text{a.s.,} \tag{6.9}$$

$$\sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)|\omega_i| = O(1) \quad \text{a.s.,} \qquad \sum_{i=1}^{n}W_{nj}(t_i)|\omega_i| = O(1) \quad \text{a.s.,} \tag{6.10}$$

$$\frac{1}{n}\sum_{i=1}^{n}\bigl|\tilde{\omega}_i^c\bigr| = O(1) \quad \text{a.s.,} \qquad \frac{1}{n}\sum_{i=1}^{n}|\tilde{\omega}_i| = O(1) \quad \text{a.s.,} \tag{6.11}$$

where $\tilde{\omega}_i = \omega_i - \sum_{j=1}^n W_{nj}(t_i)\omega_j$, $\tilde{\omega}_i^c = \omega_i - \sum_{j=1}^n \delta_j W_{nj}^c(t_i)\omega_j$, ω_i are independent random variables satisfying $E\omega_i = 0$ and $\sup_{1\le i\le n}E|\omega_i|^r < \infty$ for some $r > 2$.

Meanwhile, from (A0)–(A3), Lemma 5.3, (6.2)–(6.3), (6.9)–(6.11), one can achieve

$$n^{-1}A_{5n} \le n^{-1}\left[\sum_{i=1}^{n}\bigl|\sigma_i^{-2}\delta_i\mu_i\bigr|\cdot\bigl|\tilde{g}_i^c\bigr| + \sum_{i=1}^{n}\bigl|\sigma_i^{-2}\delta_i\tilde{g}_i^c\bigr|\cdot\left|\sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j\right|\right] = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.,}$$

$$n^{-1}A_{6n} \le n^{-1}\sum_{i=1}^{n}\bigl|\sigma_i^{-2}\delta_i\tilde{\xi}_i^c\bigr|\cdot\max_{1\le i\le n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\epsilon_j\right| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.,}$$

$$\begin{aligned} n^{-1}A_{11n} &\le n^{-1}\sum_{i=1}^{n}\bigl|\sigma_i^{-2}\delta_i\bigr|\cdot\max_{1\le i\le n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j\right|\cdot\max_{1\le i\le n}\left|\sum_{j=1}^{n}\delta_k W_{nk}^c(t_i)\epsilon_k\right| \\ &= o\bigl(n^{-\frac{1}{2}}\bigr) \quad \text{a.s.} \end{aligned}$$

One can similarly get $n^{-1}A_{in} = o(n^{-1/4})$ for $i = 2,3,4,7,8,9,10,12$. Thus, the proof of Theorem 3.1(a) is completed. □

Proof of Theorem 3.1(*b*) From (3.8), for every $t \in [0,1]$, one can write

$$\begin{aligned} \hat{g}_n^{W_1}(t) - g(t) &= \sum_{j=1}^{n}\delta_j W_{nj}^c(t)(y_j - x_j\hat{\beta}_{W_1}) - g(t) \\ &= \sum_{j=1}^{n}\delta_j W_{nj}^c(t)\bigl[\xi_j\beta + g(t_j) + \epsilon_j - (\xi_j + \mu_j)\hat{\beta}_{W_1}\bigr] - g(t) \\ &= \sum_{j=1}^{n}\delta_j W_{nj}^c(t)\xi_j(\beta - \hat{\beta}_{W_1}) + \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\delta_j\bigl[g(t_j) - g(t)\bigr] \\ &\quad + \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\epsilon_j + \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j\beta + \sum_{j=1}^{n}\delta_j W_{nj}^c(t_i)\mu_j(\hat{\beta}_{W_1} - \beta) \\ &:= F_{1n}(t) + F_{2n}(t) + F_{3n}(t) + F_{4n}(t) + F_{5n}(t). \end{aligned}$$

Therefore, we only need to prove that $F_{kn}(t) = o(n^{-1/4})$ a.s. for $k = 1,2,\ldots,5$. From (A0)–(A3), Theorem 3.1(a), Lemma 5.3, (6.2), (6.3), for every $t \in [0,1]$ and any $a > 0$, one can get

$$\begin{aligned}
F_{1n}(t) &\le |\beta - \hat{\beta}_{W_1}| \cdot \left|\sum_{j=1}^{n} \delta_j W_{nj}^{c}(t)\xi_j\right| \\
&\le o\big(n^{-\frac{1}{4}}\big) \cdot \left[\left|\sum_{j=1}^{n} \delta_j W_{nj}^{c}(t)h(t_j)\right| + \left|\sum_{j=1}^{n} \delta_j W_{nj}^{c}(t)v_j\right|\right] \\
&\le o\big(n^{-\frac{1}{4}}\big) \cdot \left[O(1) + C \max_{1\le j\le n}\big|W_{nj}^{c}(t)\big| \max_{1\le m\le n}\left|\sum_{k=1}^{m} \delta_{j_k} v_{j_k}\right|\right] = o\big(n^{-\frac{1}{4}}\big) \quad \text{a.s.}, \\
F_{2n}(t) &\le \sum_{j=1}^{n} \delta_j W_{nj}^{c}(t) \cdot \big[g(t_j) - g(t)\big] \cdot I\big(|t_j - t| > a \cdot n^{-\frac{1}{4}}\big) \\
&\quad + \sum_{j=1}^{n} \delta_j W_{nj}^{c}(t) \cdot \big[g(t_j) - g(t)\big] \cdot I\big(|t_j - t| < a \cdot n^{-\frac{1}{4}}\big) \\
&\le C \cdot a \cdot n^{-\frac{1}{4}} = o\big(n^{-\frac{1}{4}}\big) \quad \text{a.s.}
\end{aligned}$$

One can easily get $F_{kn}(t) = o(n^{-1/4})$ a.s. for $k = 3, 4, 5$ from (6.3) and Theorem 3.1(a). The proof of Theorem 3.1(b) is completed. □

Proof of Theorem 3.2(*a*) From (3.9)–(3.10), one can write

$$\begin{aligned}
\hat{\beta}_{I_1} - \beta &= S_{2n}^{-2}\left[\sum_{i=1}^{n} \sigma_i^{-2}(\tilde{\xi}_i + \tilde{\mu}_i)\tilde{U}_i^{I_1} - \sum_{i=1}^{n} \sigma_i^{-2}\big((\tilde{\xi}_i + \tilde{\mu}_i)^2 - \delta_i \Xi_{\mu}^2\big)\beta\right] \\
&= S_{2n}^{-2}\left\{\sum_{i=1}^{n} \sigma_i^{-2}(\tilde{\xi}_i + \tilde{\mu}_i)\left[\delta_i(\epsilon_i - \mu_i\beta) + \delta_i\big(g(t_i) - \hat{g}_n^{W_1}(t_i)\big)\right.\right. \\
&\quad + (1 - \delta_i)(\xi_i + \mu_i)(\hat{\beta}_{W_1} - \beta) + \hat{g}_n^{W_1}(t_i) - \sum_{j=1}^{n} W_{nj}(t_i)\big(\delta_j(\epsilon_j - \mu_j\beta) \\
&\quad \left.\left. + \delta_j\big(g(t_j) - \hat{g}_n^{W_1}(t_j)\big) + (1 - \delta_j)(\xi_j + \mu_j)(\hat{\beta}_{W_1} - \beta) + \hat{g}_n^{W_1}(t_j)\big)\right] + \sum_{i=1}^{n} \sigma_i^{-2}\delta_i \Xi_{\mu}^2\beta\right\} \\
&= S_{2n}^{-2}\{\sum_{i=1}^{n} \sigma_i^{-2}\delta_i\tilde{\xi}_i(\epsilon_i - \mu_i\beta) + \sum_{i=1}^{n} \sigma_i^{-2}\delta_i\tilde{\xi}_i\big[g(t_i) - \hat{g}_n^{W_1}(t_i)\big] \\
&\quad + \sum_{i=1}^{n} \sigma_i^{-2}\tilde{\xi}_i(\xi_i + \mu_i)(1 - \delta_i)(\hat{\beta}_{W_1} - \beta) - \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\delta_j W_{nj}(t_i)\tilde{\xi}_i(\epsilon_j - \mu_j\beta) \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\delta_j W_{nj}(t_i)\tilde{\xi}_i\big(g(t_j) - \hat{g}_n^{W_1}(t_j)\big) \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2} W_{nj}(t_i)\tilde{\xi}_i(1 - \delta_j)(\xi_j + \mu_j)(\hat{\beta}_{W_1} - \beta) + \sum_{i=1}^{n} \sigma_i^{-2}\tilde{\xi}_i\big(\hat{g}_n^{W_1}(t_i) - g(t_i)\big) \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\tilde{\xi}_i W_{nj}(t_i)\big(\hat{g}_n^{W_1}(t_j) - g(t_j)\big) + \sum_{i=1}^{n} \sigma_i^{-2}\tilde{\xi}_i\tilde{g}_i + \sum_{i=1}^{n} \sigma_i^{-2}\delta_i\mu_i\epsilon_i
\end{aligned}$$

$$
\begin{aligned}
&- \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\delta_i W_{nj}(t_i)\epsilon_i\mu_j - \sum_{i=1}^{n}\sigma_i^{-2}\delta_i\bigl(\mu_i^2 - \Xi_\mu^2\bigr)\beta \\
&+ \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\delta_i W_{nj}(t_i)\mu_i\mu_j\beta \\
&+ \sum_{i=1}^{n}\sigma_i^{-2}\delta_i\tilde{\mu}_i\bigl(g(t_i) - \hat{g}_n^{W_1}(t_i)\bigr) + \sum_{i=1}^{n}\sigma_i^{-2}\tilde{\mu}_i(1-\delta_i)(\xi_i + \mu_i)(\hat{\beta}_{W_1} - \beta) \\
&- \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\delta_j W_{nj}(t_i)\mu_i\epsilon_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sigma_i^{-2}\delta_j W_{nj}(t_i)W_{nk}(t_i)\mu_k\epsilon_j \\
&+ \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\delta_j W_{nj}(t_i)\mu_i\mu_j\beta - \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sigma_i^{-2}\delta_j W_{nj}(t_i)W_{nk}(t_i)\mu_j\mu_k\beta \\
&- \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\delta_j W_{nj}(t_i)\tilde{\mu}_i\bigl[g(t_j) - \hat{g}_n^{W_1}(t_j)\bigr] \\
&- \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2} W_{nj}(t_i)(1-\delta_j)\tilde{\mu}_i(\xi_j + \mu_j)(\hat{\beta}_{W_1} - \beta) + \sum_{i=1}^{n}\sigma_i^{-2}\tilde{\mu}_i\bigl(g_n^{W_1}(t_i) - g(t_i)\bigr) \\
&+ \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i^{-2}\tilde{\mu}_i W_{nj}(t_i)\bigl(\hat{g}_n^{W_1}(t_j) - g(t_j)\bigr) + \sum_{i=1}^{n}\sigma_i^{-2}\tilde{\mu}_i\tilde{g}_i \\
:={}& S_{2n}^{-2}\sum_{k=1}^{24} D_{kn}.
\end{aligned}
$$

Using a similar approach to step 1 in the proof of Theorem 3.1(a), one can get $S_{2n}^{-2} \le Cn^{-1}$ a.s. Therefore, we only need to verity that $n^{-1}D_{kn} = o(n^{-1/4})$ a.s. for $k = 1, 2, \dots, 24$. From (A0)–(A4), Lemmas 5.2–5.4, Theorem 3.1, (6.2)–(6.4), (6.9)–(6.11), one obtains

$$
\begin{aligned}
n^{-1}D_{1n} &= n^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\delta_i\tilde{\xi}_i(\epsilon_i - \mu_i\beta) = n^{-1}\cdot O\bigl(n^{\frac{1}{2}}\log n\bigr) = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.}, \\
n^{-1}D_{2n} &\le n^{-1}\max_{1\le i\le n}\bigl|g(t_i) - \hat{g}_n^{W_1}(t_i)\bigr| \cdot \sum_{i=1}^{n}\bigl|\sigma_i^{-2}\delta_i\tilde{\xi}_i\bigr| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.}, \\
n^{-1}D_{3n} &\le n^{-1}\left|\sum_{i=1}^{n}\sigma_i^{-2}(1-\delta_i)\tilde{\xi}_i^2\right| \cdot |\hat{\beta}_{W_1} - \beta| \\
&\quad + n^{-1}\sum_{i=1}^{n}\bigl|\sigma_i^{-2}\tilde{\xi}_i\bigr| \cdot |1-\delta_i| \cdot \max_{1\le i\le n}\left|\sum_{j=1}^{n} W_{nj}(t_i)\xi_j\right| \cdot |\hat{\beta}_{W_1} - \beta| \\
&\quad + n^{-1}\left|\sum_{i=1}^{n}\sigma_i^{-2}\tilde{\xi}_i(1-\delta_i)\mu_i\right| \cdot |\hat{\beta}_{W_1} - \beta| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.}, \\
n^{-1}D_{4n} &\le n^{-1}\max_{1\le i\le n}\left|\sum_{j=1}^{n} W_{nj}(t_i)\delta_j(\epsilon_j - \mu_j\beta)\right| \cdot \sum_{i=1}^{n}\bigl|\sigma_i^{-2}\tilde{\xi}_i\bigr| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.} \\
n^{-1}D_{5n} &\le n^{-1}\max_{1\le i\le n}\left|\sum_{i=1}^{n}\delta_j W_{nj}(t_i)\right| \cdot \max_{1\le i\le n}\bigl|g(t_i) - \hat{g}_n^{W_1}(t_i)\bigr| \cdot \sum_{i=1}^{n}\bigl|\sigma_i^{-2}\tilde{\xi}_i\bigr| = o\bigl(n^{-\frac{1}{4}}\bigr) \quad \text{a.s.}
\end{aligned}
$$

$$n^{-1}D_{6n} \leq n^{-1}\sum_{i=1}^{n}\left|\sigma_i^{-2}\tilde{\xi}_i\right| \cdot \max_{1\leq i\leq n}\left|\sum_{j=1}^{n}W_{nj}(t_i)(1-\delta_j)\xi_j\right| \cdot |\hat{\beta}_{W_1}-\beta|$$

$$+ n^{-1}\sum_{i=1}^{n}\left|\sigma_i^{-2}\tilde{\xi}_i\right| \cdot \max_{1\leq i\leq n}\left|\sum_{j=1}^{n}W_{nj}(t_i)(1-\delta_j)\mu_j\right| \cdot |\hat{\beta}_{W_1}-\beta| = o\left(n^{-\frac{1}{4}}\right) \quad \text{a.s.}$$

One can easily get $n^{-1}D_{kn} = o(n^{-1/4})$ a.s. for $k = 7, 8, \ldots, 24$. Thus, the proof of Theorem 3.2(a) is completed. □

Proof of Theorem 3.2(*b*) From (3.11), for every $t \in [0,1]$, one can write

$$\begin{aligned}
\hat{g}_n^{I_1}(t) - g(t) &= \sum_{j=1}^{n} W_{nj}(t)\left\{\delta_j y_j + (1-\delta_j)\left[(\xi_j+\mu_j)\hat{\beta}_{W_1} + \hat{g}_n^{W_1}(t_j)\right] - (\xi_j+\mu_j)\hat{\beta}_{I_1}\right\} - g(t) \\
&= \sum_{j=1}^{n} W_{nj}(t)(1-\delta_j)\xi_j(\hat{\beta}_{W_1}-\beta) + \sum_{j=1}^{n} W_{nj}(t)(1-\delta_j)\left[\hat{g}_n^{W_1}(t_j) - g(t_j)\right] \\
&\quad + \sum_{j=1}^{n} W_{nj}(t)\delta_j\epsilon_j + \sum_{j=1}^{n} W_{nj}(t)\mu_j(\hat{\beta}_{W_1}-\beta) - \sum_{j=1}^{n} W_{nj}(t)\delta_j\mu_j\hat{\beta}_{W_1} \\
&\quad + \sum_{j=1}^{n} W_{nj}(t)\left[g(t_j)-g(t)\right] + \sum_{j=1}^{n} W_{nj}(t)\xi_j(\beta-\hat{\beta}_{I_1}) + \sum_{j=1}^{n} W_{nj}(t)\mu_j(\beta-\hat{\beta}_{I_1}) \\
&:= \sum_{k=1}^{8} G_{kn}.
\end{aligned}$$

Therefore, we only need to prove that $G_{kn}(t) = o(n^{-1/4})$ a.s. for $k = 1, 2, \ldots, 8$. Similar to $F_{1n}(t) = o(n^{-1/4})$, we can get

$$G_{1n}(t) \leq \left|\sum_{j=1}^{n} W_{nj}(t)(1-\delta_j)\xi_j\right| \cdot |\hat{\beta}_{W_1}-\beta| = o\left(n^{-\frac{1}{4}}\right) \quad \text{a.s.}$$

Then from (A0)–(A4), Lemmas 5.3–5.4, (6.2) and (6.3), for every $t \in [0,1]$ and any $a > 0$, one can get

$$G_{2n}(t) \leq \left|\sum_{j=1}^{n} W_{nj}(t)(1-\delta_j)\right| \cdot \max_{1\leq j\leq n}\left|\hat{g}_n^{W_1}(t_j) - g(t_j)\right| = o\left(n^{-\frac{1}{4}}\right) \quad \text{a.s.}$$

$$\begin{aligned}
G_{5n}(t) &\leq \left|\sum_{j=1}^{n} W_{nj}(t)\delta_j\mu_j(\hat{\beta}_{W_1}-\beta)\right| + \left|\sum_{j=1}^{n} W_{nj}(t)\delta_j\mu_j\beta\right| \\
&\leq |\beta-\hat{\beta}_{W_1}| \cdot \left|\sum_{j=1}^{n} W_{nj}(t)\delta_j\mu_j\right| + \left|\sum_{j=1}^{n} W_{nj}(t)\mu_j\beta\right| = o\left(n^{-\frac{1}{4}}\right) \quad \text{a.s.,}
\end{aligned}$$

$$\begin{aligned}
G_{6n}(t) &\leq \sum_{j=1}^{n} W_{nj}(t) \cdot \left[g(t_j)-g(t)\right] \cdot I\left(|t_j - t| > a \cdot n^{-\frac{1}{4}}\right) \\
&\quad + \sum_{j=1}^{n} W_{nj}(t) \cdot \left[g(t_j)-g(t)\right] \cdot I\left(|t_j - t| < a \cdot n^{-\frac{1}{4}}\right) = o\left(n^{-\frac{1}{4}}\right) \quad \text{a.s.}
\end{aligned}$$

One can easily get $G_{kn}(t) = o(n^{-1/4})$ a.s. for $k = 3, 4, 7, 8$. Thus, the proof of Theorem 3.2(b) is completed. □

Proof of Theorem 3.3 From (3.12), one can write

$$\begin{aligned}
&\hat{f}_n(u) - f(u) \\
&\quad= \sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\bigl\{\bigl[\tilde{\xi}_i^c(\beta - \hat{\beta}_c) + \tilde{g}_i^c + \tilde{\epsilon}_i^c - \tilde{\mu}_i^c\hat{\beta}_c\bigr]^2 - \Xi_\mu^2\hat{\beta}_c^2 - f(u)\bigr\} \\
&\quad= \sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\bigl(\epsilon_i^2 - f(u)\bigr) + \sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{\xi}_i^{c2}(\beta - \hat{\beta}_c)^2 + \sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{g}_i^{c2} \\
&\qquad+ \sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\left(\sum_{j=1}^{n} \delta_j W_{nj}^c(t_i)\epsilon_j\right)^2 + \sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\bigl(\mu_i^2 - \Xi_\mu^2\bigr)\hat{\beta}_c^2 \\
&\qquad+ \sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\left(\sum_{j=1}^{n} \delta_j W_{nj}^c(t_i)\mu_j\hat{\beta}_c\right)^2 + 2\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{\xi}_i^c\tilde{g}_i^c(\beta - \hat{\beta}_c) \\
&\qquad+ 2\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{\xi}_i^c(\beta - \hat{\beta}_c)\epsilon_i - 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{\xi}_i^c\delta_j W_{nj}^c(t_i)\epsilon_j(\beta - \hat{\beta}_c) \\
&\qquad- 2\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{\xi}_i^c\mu_i(\beta - \hat{\beta}_c)\hat{\beta}_c + 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{\xi}_i^c\delta_j W_{nj}^c(t_i)\mu_j(\beta - \hat{\beta}_c)\hat{\beta}_c \\
&\qquad+ 2\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{g}_i^c\epsilon_i - 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{g}_i^c\delta_j W_{nj}^c(t_i)\epsilon_j - 2\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{g}_i^c\mu_i\hat{\beta}_c \\
&\qquad+ 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\tilde{g}_i^c\delta_j W_{nj}^c(t_i)\mu_j\hat{\beta}_c - 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\epsilon_i\delta_j W_{nj}^c(t_i)\epsilon_j \\
&\qquad- 2\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\epsilon_i\mu_i\hat{\beta}_c + 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\epsilon_i\delta_j W_{nj}^c(t_i)\mu_j\hat{\beta}_c \\
&\qquad+ 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\mu_i\delta_j W_{nj}^c(t_i)\mu_j\hat{\beta}_c - 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \hat{W}_{ni}^c(u)\mu_i\delta_j W_{nj}^c(t_i)\mu_j\hat{\beta}_c^2 \\
&\qquad- 2\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \delta_i \hat{W}_{ni}^c(u)\delta_j W_{nj}^c(t_i)\epsilon_j\delta_k W_{nk}^c(t_i)\mu_k\hat{\beta}_c \\
&\quad:= \sum_{k=1}^{21} U_{kn}(u),
\end{aligned}$$

$$\begin{aligned}
&U_{1n}(u) \\
&\quad\le \left|\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\bigl(\epsilon_i^2 - f(u)\bigr)\right| \\
&\quad\le \left|\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)f(u_i)\bigl(e_i^2 - 1\bigr)\right| + \left|\sum_{i=1}^{n} \delta_i \hat{W}_{ni}^c(u)\bigl(f(u_i) - f(u)\bigr)\right|
\end{aligned}$$

$$:= J_{11n}(u) + J_{12n}(u).$$

We get $E(e_i^2 - 1) = 0$ and $\sup_{1\le i\le n} E|e_i^2 - 1|^{\gamma_2/2} < \infty$, where $\gamma_2 > 16/3$. From (A0)–(A3), (A5), Theorem 3.1(a), Lemma 5.3, (6.2)–(6.3), we have $J_{1kn}(u) = o(n^{-1/4})$ a.s. for $k = 1, 2$. Therefore, we can get $U_{1n}(u) = o(n^{-1/4})$ a.s. By Lemma 5.2, taking $p_1 = 3/8$, $p_2 = 1/8$, $\gamma = 4$, we can get $U_{kn}(u) = o(n^{-1/4})$ a.s. for $k = 8, 10$. By Lemma 5.1, taking $\gamma > 4$, $\alpha = 2$, we can get $U_{kn}(u) = o(n^{-1/4})$ a.s. for $k = 12, 14$. Similarly, one can deduce that $U_{kn}(u) = o(n^{-1/4})$ a.s. for $k = 2, 3, \ldots, 21$. Thus, the proof of Theorem 3.3 is completed. □

Proof of Theorem 3.4(a) Let $T_{1n}^2 = \sum_{i=1}^n \hat{\sigma}_{ni}^{-2}\delta_i(\tilde{x}_i^{c2} - \Xi_\mu^2)$, similar to Theorem 3.1(a), one can write

$$\begin{aligned}
&\hat{\beta}_{W_2} - \beta \\
&= T_{1n}^{-2}\Bigg\{\sum_{i=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\tilde{\xi}_i^c(\epsilon_i - \mu_i\beta) + \sum_{i=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\mu_i\epsilon_i - \sum_{i=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\big(\mu_i^2 - \Xi_\mu^2\big)\beta \\
&\quad + \sum_{i=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\tilde{\xi}_i^c\tilde{g}_i^c + \sum_{i=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\tilde{\mu}_i^c\tilde{g}_i^c - \sum_{i=1}^n\sum_{j=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\delta_j W_{nj}^c(t_i)\tilde{\xi}_i^c\epsilon_j \\
&\quad - \sum_{i=1}^n\sum_{j=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\delta_j W_{nj}^c(t_i)\epsilon_i\mu_j - \sum_{i=1}^n\sum_{j=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\delta_j W_{nj}^c(t_i)\mu_i\epsilon_j \\
&\quad + \sum_{i=1}^n\sum_{j=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\delta_j W_{nj}^c(t_i)\tilde{\xi}_i^c\mu_j\beta + 2\sum_{i=1}^n\sum_{j=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\delta_j W_{nj}^c(t_i)\mu_i\mu_j\beta \\
&\quad + \sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\delta_j\delta_k W_{nj}^c(t_i)W_{nk}^c(t_i)\mu_j\epsilon_k \\
&\quad - \sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n \hat{\sigma}_{ni}^{-2}\delta_i\delta_j\delta_k W_{nj}^c(t_i)W_{nk}^c(t_i)\mu_j\mu_k\beta\Bigg\} \\
&:= T_{1n}^{-2}\sum_{k=1}^{12} G_{kn}.
\end{aligned}$$

Similar to step 1 in the proof of Theorem 3.1, we can get $T_{1n}^2 \le Cn^{-1}$ a.s. Therefore, we only need to verify that $n^{-1}G_{kn} = o(n^{-1/4})$ a.s. for $k = 1, 2, \ldots, 12$. From (A0)–(A5), Lemmas 5.2–5.4, Theorem 3.1(a), Theorem 3.3, (6.2)–(6.4), (6.9)–(6.11), one obtains

$$\begin{aligned}
n^{-1}G_{1n} &\le n^{-1}\Bigg|\sum_{i=1}^n\big(\hat{\sigma}_{ni}^{-2} - \sigma_i^{-2}\big)\delta_i\tilde{\xi}_i^c(\epsilon_i - \mu_i\beta)\Bigg| + n^{-1}\Bigg|\sum_{i=1}^n \sigma_i^{-2}\delta_i\tilde{\xi}_i^c(\epsilon_i - \mu_i\beta)\Bigg| \\
&\le \max_{1\le i\le n}\big|\hat{\sigma}_{ni}^{-2} - \sigma_i^{-2}\big| \cdot n^{-1}\sum_{i=1}^n\big|\delta_i\tilde{\xi}_i^c(\epsilon_i - \mu_i\beta)\big| + o\big(n^{-\frac{1}{4}}\big) = o\big(n^{-\frac{1}{4}}\big) \quad \text{a.s.}, \\
n^{-1}G_{2n} &\le n^{-1}\Bigg|\sum_{i=1}^n\big(\hat{\sigma}_{ni}^{-2} - \sigma_i^{-2}\big)\delta_i\mu_i\epsilon_i\Bigg| + n^{-1}\Bigg|\sum_{i=1}^n \sigma_i^{-2}\delta_i\mu_i\epsilon_i\Bigg| \\
&\le \max_{1\le i\le n}\big|\hat{\sigma}_{ni}^{-2} - \sigma_i^{-2}\big| \cdot n^{-1}\sum_{i=1}^n|\delta_i\mu_i\epsilon_i| + o\big(n^{-\frac{1}{4}}\big) = o\big(n^{-\frac{1}{4}}\big) \quad \text{a.s.},
\end{aligned}$$

$$
\begin{aligned}
n^{-1}G_{3n} &\le n^{-1}\left|\sum_{i=1}^{n}(\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2})\delta_i(\mu_i^2-\Xi_\mu^2)\beta\right| + n^{-1}\left|\sum_{i=1}^{n}\sigma_i^{-2}\delta_i(\mu_i^2-\Xi_\mu^2)\beta\right| \\
&\le \max_{1\le i\le n}\left|\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right|\cdot n^{-1}\sum_{i=1}^{n}\left|\delta_i(\mu_i^2-\Xi_\mu^2)\beta\right| + o(n^{-\frac{1}{4}}) = o(n^{-\frac{1}{4}}) \quad \text{a.s.}
\end{aligned}
$$

The proofs of $n^{-1}G_{kn} = o(n^{-1/4})$ a.s. for $k = 4, 5, \dots, 12$ are similar. Thus, the proof of Theorem 3.4(a) is completed. □

The proof of Theorem 3.4(b) is similar to the proof of Theorem 3.1(b).

Proof of Theorem 3.5(*a*) Let $T_{2n}^2 = \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}(\tilde{x}_i^2 - \delta_i\Xi_\mu^2)$, similar to Theorem 3.2(a), one can write

$$
\begin{aligned}
&\hat{\beta}_{I_2} - \beta \\
&= T_{2n}^{-2}\{\sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_i\tilde{\xi}_i(\epsilon_i-\mu_i\beta) + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_i\tilde{\xi}_i\big(g(t_i)-\hat{g}_n^{W_2}(t_i)\big) \\
&\quad + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\xi}_i(\xi_i+\mu_i)(1-\delta_i)(\hat{\beta}_{W_2}-\beta) - \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_j W_{nj}(t_i)\tilde{\xi}_i(\epsilon_j-\mu_j\beta) \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_j W_{nj}(t_i)\tilde{\xi}_i\big(g(t_j)-\hat{g}_n^{W_2}(t_j)\big) \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}W_{nj}(t_i)\tilde{\xi}_i(1-\delta_j)(\xi_j+\mu_j)(\hat{\beta}_{W_2}-\beta) + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\xi}_i\big(\hat{g}_n^{W_2}(t_i)-g(t_i)\big) \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\xi}_i W_{nj}(t_i)\big(\hat{g}_n^{W_2}(t_j)-g(t_j)\big) + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\xi}_i\tilde{g}_i^c + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_i\mu_i\epsilon_i \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_i W_{nj}(t_i)\epsilon_i\mu_j - \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_i\big(\mu_i^2-\Xi_\mu^2\big)\beta + \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_i W_{nj}(t_i)\mu_i\mu_j\beta \\
&\quad + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_i\tilde{\mu}_i\big(g(t_i)-\hat{g}_n^{W_2}(t_i)\big) + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\mu}_i(1-\delta_i)(\xi_i+\mu_i)(\hat{\beta}_{W_2}-\beta) \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_j W_{nj}(t_i)\mu_i\epsilon_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_j W_{nj}(t_i)W_{nk}(t_i)\mu_k\epsilon_j \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_j W_{nj}(t_i)\mu_i\mu_j\beta - \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_j W_{nj}(t_i)W_{nk}(t_i)\mu_j\mu_k\beta \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\delta_j W_{nj}(t_i)\tilde{\mu}_i\big[g(t_j)-\hat{g}_n^{W_2}(t_j)\big] \\
&\quad - \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}W_{nj}(t_i)(1-\delta_j)\tilde{\mu}_i(\xi_j+\mu_j)(\hat{\beta}_{W_2}-\beta) + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\mu}_i\big(g_n^{W_2}(t_i)-g(t_i)\big) \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\mu}_i W_{nj}(t_i)\big(\hat{g}_n^{W_2}(t_j)-g(t_j)\big) + \sum_{i=1}^{n}\hat{\sigma}_{ni}^{-2}\tilde{\mu}_i\tilde{g}_i^c
\end{aligned}
$$

$$:= T_{2n}^{-2}\sum_{k=1}^{24} H_{kn}.$$

Using a similar approach to step 1 in the proof of Theorem 3.4(a), we can get $T_{2n}^{-2} \le C_2 n^{-1}$ a.s. Then from (A0)–(A4), Lemmas 5.2–5.4, Theorem 3.4(a), Theorem 3.3, (6.2)–(6.4), (6.9)–(6.11) one obtains

$$\begin{aligned} n^{-1}H_{1n} &\le n^{-1}\left|\sum_{i=1}^{n}\left(\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right)\delta_i\tilde{\xi}_i(\epsilon_i-\mu_i\beta)\right| + n^{-1}\left|\sum_{i=1}^{n}\sigma_i^{-2}\delta_i\tilde{\xi}_i(\epsilon_i-\mu_i\beta)\right| \\ &\le \max_{1\le i\le n}\left|\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right|\cdot n^{-1}\sum_{i=1}^{n}\left|\delta_i\tilde{\xi}_i(\epsilon_i-\mu_i\beta)\right| + o\left(n^{-\frac{1}{4}}\right) = o\left(n^{-\frac{1}{4}}\right) \quad \text{a.s.}, \\ n^{-1}H_{2n} &\le n^{-1}\left|\sum_{i=1}^{n}\left(\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right)\delta_i\tilde{\xi}_i\left(g(t_i)-\hat{g}_n^{W_2}(t_i)\right)\right| + n^{-1}\left|\sum_{i=1}^{n}\sigma_i^{-2}\delta_i\tilde{\xi}_i\left(g(t_i)-\hat{g}_n^{W_2}(t_i)\right)\right| \\ &\le n^{-1}\cdot\max_{1\le i\le n}\left|\hat{\sigma}_{ni}^{-2}-\sigma_i^{-2}\right|\cdot\max_{1\le i\le n}\left|g(t_i)-\hat{g}_n^{W_2}(t_i)\right|\cdot\sum_{i=1}^{n}\left|\sigma_i^{-2}\tilde{\xi}_i\right| + o\left(n^{-\frac{1}{4}}\right) = o\left(n^{-\frac{1}{4}}\right) \\ &\quad \text{a.s.} \end{aligned}$$

The proofs of $n^{-1}H_{kn} = o(n^{-1/4})$ a.s. for $k = 3, 4, \ldots, 24$ are similar. Thus, the proof of Theorem 3.5(a) is completed. □

The proof of Theorem 3.5(b) is similar to the proof of Theorem 3.2(b).

Acknowledgements
The authors greatly appreciate the constructive comments and suggestions of the Editor and referee.

Funding
This research was supported by the National Natural Science Foundation of China (11701368).

Availability of data and materials
This paper did not use the actual data. Simulation data was produced by matlab software.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
JJZ gave the framework of the article and guided YPX to complete the theoretical proof of the article, YPX completed the simulation. All authors read and approved the final manuscript.

Publisher's Note

Received: 20 January 2020 Accepted: 11 May 2020 Published online: 24 May 2020

References

1. Baek, J.I., Liang, H.Y.: Asymptotic of estimators in semi-parametric model under NA samples. J. Stat. Plan. Inference **136**, 3362–3382 (2006)
2. Chen, H.: Convergence rates for parametric components in a partly linear model. Ann. Stat. **16**, 136–146 (1988)
3. Cheng, P.E.: Nonparametric estimation of mean functionals with data missing at random. J. Am. Stat. Assoc. **89**, 81–87 (1994)
4. Cui, H.J., Li, R.C.: On parameter estimation for semi-linear errors-in-variables models. J. Multivar. Anal. **64**(1), 1–24 (1998)
5. Engle, R.F., Granger, C.W.J., Rice, J., Weiss, A.: Semiparametric estimation of the relation between weather and electricity sales. J. Am. Stat. Assoc. **81**, 310–320 (1986)
6. Fan, G.L., Liang, H.Y., Wang, J.F., Xu, H.X.: Asymptotic properties for LS estimators in EV regression model with dependent errors. AStA Adv. Stat. Anal. **94**, 89–103 (2010)

7. Gao, J.T., Chen, X.R., Zhao, L.C.: Asymptotic normality of a class of estimators in partial linear models. Acta Math. Sin. **37**(2), 256–268 (1994)
8. Härdle, W., Liang, H., Gao, J.T.: Partial Linear Models. Physica-Verlag, Heidelberg (2000)
9. Healy, M.J.R., Westmacott, M.: Missing values in experiments analysis on automatic computers. Appl. Stat. **5**, 203–206 (1956)
10. Hu, X.M., Wang, Z.Z., Liu, F.: Zero finite-order serial correlation test in a semi-parametric varying-coefficient partially linear errors-in-variables model. Stat. Probab. Lett. **78**, 1560–1569 (2008)
11. Liang, H., Härdle, W., Carrol, R.J.: Estimation in a semiparametric partially linear errosr-in-variables model. Ann. Stat. **27**(5), 1519–1935 (1999)
12. Little, R.J., Rubin, D.B.: Statistical Analysis with Missing Data. Wiley, New York (1987)
13. Liu, J.X., Chen, X.R.: Consistency of LS estimator in simple linear EV regression models. Acta Math. Sci. Ser. B Engl. Ed. **25**, 50–58 (2005)
14. Miao, Y., Liu, W.: Moderate deviations for LS estimator in simple linear EV regression model. J. Stat. Plan. Inference **139**(9), 3122–3131 (2009)
15. Miao, Y., Yang, G., Shen, L.: The central limit theorem for LS estimator in simple linear EV regression models. Commun. Stat., Theory Methods **36**, 2263–2272 (2007)
16. Wang, Q., Linton, O., Härdle, W.: Semiparametric regression analysis with missing response at random. J. Am. Stat. Assoc. **99**(466), 334–345 (2004)
17. Wang, Q., Sun, Z.: Estimation in partially linear models with missing responses at random. J. Multivar. Anal. **98**, 1470–1493 (2007)
18. Wei, C.H., Mei, C.L.: Empirical likelihood for partially linear varying-coefficient models with missing response variables and error-prone covariates. J. Korean Stat. Soc. **41**, 97–103 (2012)
19. Xu, H.X., Fan, G.L., Chen, Z.L.: Hypothesis tests in partial linear errors-in-variables models with missing response. Stat. Probab. Lett. **126**, 219–229 (2017)
20. Yang, H., Xia, X.C.: Equivalence of two tests in varying coefficient partially linear errors in variable model with missing responses. J. Korean Stat. Soc. **43**, 79–90 (2014)
21. Zhang, J.J., Liang, H.Y.: Berry–Esseen type bounds in heteroscedastic semi-parametric model. J. Stat. Plan. Inference **141**, 3447–3462 (2011)
22. Zhou, H.B., You, J.H., Zhou, B.: Statistical inference for fixed-effects partially linear regression models with errors in variables. Stat. Pap. **51**, 629–650 (2010)