Strong Consistency of Estimators under Missing Responses

Linran Zhang, Jingjing Zhang*

College of Science, University of Shanghai for Science and Technology, Shanghai, China
Email: zlr18231178176@foxmail.com, *jenny617@usst.edu.cn

Abstract

In this article, we focus on the semi-parametric error-in-variables model with missing responses: $y_i = \xi_i \beta + g(t_i) + \epsilon_i$, $x_i = \xi_i + \mu_i$, where y_i are the response variables missing at random, (ξ_i, t_i) are design points, ξ_i are the potential variables observed with measurement errors μ_i, the unknown slope parameter β and nonparametric component $g(\cdot)$ need to be estimated. Here we choose two different approaches to estimate β and $g(\cdot)$. Under appropriate conditions, we study the strong consistency for the proposed estimators.

Keywords

Semi-Parametric Model, Error-in-Variables, Missing Responses, Strong Consistency

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 y_i are the response variables, (ξ_i, t_i) are design points, ξ_i are the potential variables observed with measurement errors μ_i, $E\mu_i = 0$, ϵ_i are random errors with $E\epsilon_i = 0$. $\beta \in \mathcal{R}$ is an unknown parameter that needs to be estimated. $g(\cdot)$ is a unknown function defined on close interval $[0,1]$, $h(\cdot)$ is a known function defined on $[0,1]$ satisfying

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

where v_i are also design points.

Model (1.1) and its special forms have gained much attention in recent years. When $\mu_i \equiv 0$, ξ_i are observed exactly, the model (1.1) reduces to the general semi-parametric model, which was first introduced by Engle *et al.* [1]. However, in many applications, there are often covariates measurement errors. So the EV models are somewhat more practical than the ordinary regression model. In addition, when y_i are complete observed and $g(\cdot) \equiv 0$, the model (1.1) reduces to the usual linear EV model, which has been studied by Liu and Chen [2], Miao *et al.* [3], Miao and Liu [4], Fan *et al.* [5] and so on. For complete data, the model (1) itself has also been studied by many authors: See Cui and Li [6], Liang *et al.* [7], Zhou *et al.* [8] and so on. In recent years, the semi-parametric EV models have been widely concerned.

On the other hand, we often encounter incomplete data in the practical application of the models. In particular, some response variables may be missing, by design or by happenstance. For example, the responses y_i may be very expensive to measure and only part of y_i are available. Actually, missing of responses is very common in opinion polls, social-economic investigations, market research surveys and so on. Therefore, we focus our attention on the case that missing data occur only in the response variables. When ξ_i can fully be observed, the model (1.1) reduces to the usual semi-parametric model which has been studied by many scholars in the literature: See Wang *et al.* [9], Wang and Sun [10], Bianco *et al.* [11].

To deal with missing data, one method is to impute a plausible value for each missing datum and then analyze the results as if they are complete. In regression problems, common imputation approaches include linear regression imputation by Healy and Westmacott [12], nonparametric kernel regression imputation by Cheng [13], semi-parametric regression imputation by Wang *et al.* [9], Wang and Sun [10], among others. We here extend the methods to the estimation of β and $g(\cdot)$ under the semi-parametric EV model (1.1). We obtain two approaches to estimate β and $g(\cdot)$ with missing responses and study the strong consistency for the estimators.

In this paper, suppose we obtain a random sample of incomplete data $\{(y_i, \delta_i, x_i, t_i)\}$ from the model (1.1), where $\delta_i = 0$ if y_i is missing, otherwise $\delta_i = 1$. Throughout this paper, we assume that y_i is missing at random. The assumption implies that δ_i and y_i are independent. That is, $P(\delta_i = 1 \mid y_i) = P(\delta_i = 1)$. This assumption is a common assumption for statistical analysis with missing data and is reasonable in many practical situations.

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

2. Assumptions

In this section, we list some assumptions which will be used in the main results. Here $a_n = O(b_n)$ means $|a_n| \leq C|b_n|$ for every $n \geq 1$, $a_n = o(b_n)$ means

$a_n / b_n \to 0$ as $n \to \infty$, while a.s. is stand for almost sure.

(A0) Let $\{\epsilon_i, 1 \le i \le n\}$, $\{\mu_i, 1 \le i \le n\}$ and $\{\delta_i, 1 \le i \le n\}$ be sequences of independent random variables satisfying

i) $E\epsilon_i = 0$, $E\mu_i = 0$, $E\epsilon_i^2 = 1$, $E\mu_i^2 = \Xi_\mu^2 > 0$ is known.

ii) $\sup_i E|\epsilon_i|^p < \infty$, $\sup_i E|\mu_i|^p < \infty$ for some $p > 4$.

iii) $\{\epsilon_i, 1 \le i \le n\}$, $\{\mu_i, 1 \le i \le n\}$, $\{\delta_i, 1 \le i \le n\}$ are independent of each other.

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

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. $(0 < \Sigma_0, \Sigma_1 < \infty)$.

ii) $\lim_{n \to \infty} \sup_n \left(\sqrt{n} \log n\right)^{-1} \cdot \max_{1 \le m \le n} \left|\sum_{i=1}^m v_{j_i}\right| < \infty$, where $\{j_1, j_2, \cdots, j_n\}$ is a permutation of $(1, 2, \cdots, n)$.

iii) $\max_{1 \le i \le n} |v_i| = O(1)$.

(A2) $g(\cdot)$ and $h(\cdot)$ are continuous functions satisfying the first-order Lipschitz condition on the close interval $[0,1]$.

(A3) Let $W_{nj}^c(t_i)(1 \le i, j \le n)$ be weight functions defined on [0, 1] and satisfy

i) $\max_{1 \le j \le n} \sum_{i=1}^n \delta_j W_{nj}^c(t_i) = O(1)$ a.s.

ii) $\max_{1 \le i \le n} \sum_{j=1}^n \delta_j W_{nj}^c(t_i) I\left(|t_i - t_j| > a \cdot n^{-1/4}\right) = o(n^{-1/4})$ a.s. for any $a > 0$.

iii) $\max_{1 \le i, j \le n} W_{nj}^c(t_i) = o(n^{-1/2} \log^{-1} n)$ a.s.

(A4) The probability weight functions $W_{nj}(t_i)(1 \le i, j \le n)$ are defined on $[0,1]$ and satisfy

i) $\max_{1 \le j \le n} \sum_{i=1}^n W_{nj}(t_i) = O(1)$.

ii) $\max_{1 \le i \le n} \sum_{j=1}^n W_{nj}(t_i) I\left(|t_i - t_j| > a \cdot n^{-1/4}\right) = o(n^{-1/4})$, for any $a > 0$.

iii) $\max_{1 \le i, j \le n} W_{nj}(t_i) = o(n^{-1/2} \log^{-1} n)$.

Remark 2.1. *Conditions* (A0)-(A4) *are standard regularity conditions and used commonly in the literature, see Härdle et al.* [14], *Gao et al.* [15] *and Chen* [16].

3. Main Results

For model (1.1), we want to seek the estimators of β and $g(\cdot)$. The most natural idea is to delete all the missing data. Therefore, one can get 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 squares estimation method to estimate the parameter β. If the parameter β is known, using the complete data $\{(\delta_i y_i, \delta_i x_i, \delta_i t_i), 1 \le i \le n\}$, we can define the estimator of $g(\cdot)$ to be

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

where $W_{nj}^c(t)$ are weight functions satisfying (A3). On the other hand, under the condition of the semi-parametric EV model, Liang *et al.* [7] improved the least squares estimator (LSE) on the basis of the usual partially linear model, and employ the estimator of parameter β to minimize the following formula:

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

Therefore, we can achieve the modified LSE of β as follow:

$$\hat{\beta}_c = \left[\sum_{i=1}^{n} \left(\delta_i \tilde{x}_i^{c2} - \delta_i \Xi_\mu^2 \right) \right]^{-1} \sum_{i=1}^{n} \delta_i \tilde{x}_i^c \tilde{y}_i^c, \qquad (3.1)$$

where $\tilde{x}_i^c = x_i - \sum_{j=1}^{n} \delta_j W_{nj}^c(t_i) x_j$, $\tilde{y}_i^c = y_i - \sum_{j=1}^{n} \delta_j W_{nj}^c(t_i) y_j$. We substitute (3.1) into $g_n^*(t, \beta)$, then

$$\hat{g}_n^c(t) = \sum_{j=1}^{n} \delta_j W_{nj}^c(t) \left(y_j - x_j \hat{\beta}_c \right). \qquad (3.2)$$

Apparently, the estimators $\hat{\beta}_c$ and $\hat{g}_n^c(t)$ 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 [10], and let

$$U_i^{[I]} = \delta_i y_i + (1 - \delta_i) \left[x_i \hat{\beta}_c + \hat{g}_n^c(t_i) \right]. \qquad (3.3)$$

Therefore, Using complete data $\left\{ \left(U_i^{[I]}, x_i, t_i \right), 1 \le i \le n \right\}$, similar to (3.1)-(3.2), one can get another estimators for β and $g(\cdot)$, that is

$$\hat{\beta}_I = \left[\sum_{i=1}^{n} \left(\tilde{x}_i^2 - \delta_i \Xi_\mu^2 \right) \right]^{-1} \sum_{i=1}^{n} \tilde{x}_i \tilde{U}_i^{[I]}, \qquad (3.4)$$

$$\hat{g}_n^{[I]}(t) = \sum_{j=1}^{n} W_{nj}(t) \left(U_j^{[I]} - x_j \hat{\beta}_I \right). \qquad (3.5)$$

where $\tilde{U}_i^{[I]} = U_i^{[I]} - \sum_{j=1}^{n} W_{nj}(t_i) U_j^{[I]}$, $\tilde{x}_i = x_i - \sum_{j=1}^{n} W_{nj}(t_i) x_j$, $W_{nj}(t)$ are weight functions satisfying (A4).

Based on the estimators for β and $g(\cdot)$, we have the following results.

Theorem 3.1 *Suppose that* $(A0)$-$(A3)$ *are satisfied. For every* $t \in [0,1]$, *we have*

a) $\hat{\beta}_c \to \beta$ a.s.

b) $\hat{g}_n^c(t) \to g(t)$ a.s.

Theorem 3.2 *Suppose that* $(A0)$-$(A4)$ *are satisfied. For every* $t \in [0,1]$, *we have*

a) $\hat{\beta}_I \to \beta$ a.s.

b) $\hat{g}_n^{[I]}(t) \to g(t)$ a.s.

4. Preliminary Lemmas

In the sequel, let $C, C_1, \cdots$ 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 4.1 (Baek ang Liang [17], Lemma 3.1) *Let* $\alpha > 2$, $e_1, \cdots, e_n$ *be independent random variables with* $Ee_i = 0$. *Assume that* $\{a_{ni}, 1 \le i \le n\}$ *is a triangular array of numbers with* $\max_{1 \le i \le n} |a_{ni}| = O(n^{-1/2})$ *and* $\sum_{i=1}^{n} a_{ni}^2 = o(n^{-2/\alpha} \log^{-1} n)$. *If* $\sup_i E|e_i|^p < \infty$ *for some* $p > 2\alpha/(\alpha - 1)$. *Then*

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

Lemma 4.2 (Hardle *et al.* [14], 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 numbers such that* $\sup_{1 \le i, k \le n} |a_{ki}| = O\left(n^{-p_1}\right)$ *for some* $0 < p_1 < 1$ *and* $\sum_{j=1}^{n} a_{ji} = O\left(n^{p_2}\right)$ *for* $p_2 \ge \max\left(0, 2/r - p_1\right)$ *. Then*

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

Lemma 4.3

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 $\max_{1 \le i \le n} |\tilde{A}_i| = o\left(n^{-1/4}\right)$ and $\max_{1 \le i \le n} |\tilde{A}_i^c| = o\left(n^{-1/4}\right)$ a.s.

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

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

Lemma 4.4 *Suppose that* $(A0)$-$(A4)$ *are satisfied. Then one can deduce that*

$$\max_{1 \le i \le n} \left| \hat{g}_n^c(t_i) - g(t_i) \right| = o\left(n^{-\frac{1}{4}}\right) \text{ a.s.}$$

One can easily get Lemma 4.3 by (A0)-(A4). The proof Lemma 4.4 is analogous to the proof of Theorem 3.1(*b*).

5. Proof of Main Results

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

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

$$\tilde{g}_i^c = g(t_i) - \sum_{j=1}^{n} \delta_j W_{nj}^c(t_i) g(t_j), \quad \tilde{\epsilon}_i^c = \epsilon_i - \sum_{j=1}^{n} \delta_j W_{nj}^c(t_i) \epsilon_j,$$

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

$$\tilde{g}_i = g(t_i) - \sum_{j=1}^{n} W_{nj}(t_i) g(t_j), \quad \tilde{\epsilon}_i = \epsilon_i - \sum_{j=1}^{n} W_{nj}(t_i) \epsilon_j, \quad \eta_i = \epsilon_i - \mu_i \beta,$$

$$B_{1n}^2 = \sum_{i=1}^{n} \delta_i \tilde{\xi}_i^2, \quad S_n^2 = \sum_{i=1}^{n} \tilde{\xi}_i^2, \quad S_{1n}^2 = \sum_{i=1}^{n} \left(\delta_i \tilde{x}_i^2 - \delta_i \Xi_\mu^2\right), \quad S_{2n}^2 = \sum_{i=1}^{n} \left(\tilde{x}_i^2 - \Xi_\mu^2\right).$$

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

$$\hat{\beta}_c - \beta = S_{1n}^{-2} \left[\sum_{i=1}^{n} \delta_i \left(\tilde{\xi}_i^c + \tilde{\mu}_i^c\right)\left(\tilde{y}_i^c - \tilde{\xi}_i^c \beta - \tilde{\mu}_i^c \beta\right) + \sum_{i=1}^{n} \delta_i \Xi_\mu^2 \beta \right]$$

$$= S_{1n}^{-2} \left\{ \sum_{i=1}^{n} \left[\delta_i \left(\tilde{\xi}_i^c + \tilde{\mu}_i^c\right)\left(\tilde{\epsilon}_i^c - \tilde{\mu}_i^c \beta\right) + \delta_i \Xi_\mu^2 \beta \right] + \sum_{i=1}^{n} \delta_i \tilde{\xi}_i^c \tilde{g}_i^c + \sum_{i=1}^{n} \delta_i \tilde{\mu}_i^c \tilde{g}_i^c \right\}$$

$$
\begin{aligned}
= S_{1n}^{-2} \Bigg\{ & \sum_{i=1}^{n} \delta_i \tilde{\xi}_i^c (\epsilon_i - \mu_i \beta) + \sum_{i=1}^{n} \delta_i \mu_i \epsilon_i - \sum_{i=1}^{n} \delta_i \left(\mu^2 - \Xi_\mu^2 \right) \beta + \sum_{i=1}^{n} \delta_i \tilde{\xi}_i^c \tilde{g}_i^c \\
& + \sum_{i=1}^{n} \delta_i \tilde{\mu}_i^c \tilde{g}_i^c - \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \delta_j W_{nj}^c (t_i) \tilde{\xi}_i^c \epsilon_j - \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \delta_j W_{nj}^c (t_i) \epsilon_i \mu_j \\
& - \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \delta_j W_{nj}^c (t_i) \mu_i \epsilon_j + \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \delta_j W_{nj}^c (t_i) \tilde{\xi}_i^c \mu_j \beta \\
& + 2\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \delta_j W_{nj}^c (t_i) \mu_i \mu_j \beta + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \delta_i \delta_j \delta_k W_{nj}^c (t_i) W_{nk}^c (t_i) \mu_j \epsilon_k \\
& - \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \delta_i \delta_j \delta_k W_{nj}^c (t_i) W_{nk}^c (t_i) \mu_j \mu_k \beta \Bigg\} \\
:= & \; S_{1n}^{-2} \sum_{k=1}^{12} A_{kn}.
\end{aligned}
\tag{5.1}
$$

Thus, to prove $\hat{\beta}_c \to \beta$ a.s., we only need to verify that $S_{1n}^{-2} \leq C n^{-1}$ a.s. and $n^{-1} A_{kn} = o(1)$ a.s. for $k = 1,2,\cdots,12$.

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

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

By Lemma 4.3(a), we have $n^{-1} B_{1n} \to \Sigma_1$ a.s. Hence, it suffices to verify that $B_{kn} = o(B_{1n}) = o(n)$ a.s. for $k = 2,3,\cdots,6$. Applying (A0), taking $r = p/2 > 2$, $p_1 = 1/2$, $p_2 = 1/2$ in Lemma 4.2, we can verity that

$$
\sum_{i=1}^{n} (\zeta_i - E\zeta_i) = n^{\frac{1}{2}} \cdot \sum_{i=1}^{n} n^{-\frac{1}{2}} (\zeta_i - E\zeta_i) = O\left(n^{\frac{1}{2}} \log n \right) \text{ a.s.}
\tag{5.2}
$$

where $\{\zeta_i\}$ is a sequence of independent random variables satisfying $E\zeta_i = 0$ and $\sup_{1 \leq i \leq n} E|\zeta_i|^{p/2} < \infty$. Therefore, we obtain $B_{2n} = O(n^{1/2} \log n) = o(n)$ a.s. from (A0) and (5.2). On the other hand, taking $\alpha = 4, p > 4$ in Lemma 4.1, we have

$$
\max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} \delta_j W_{nj}^c (t_i) \zeta_j \right| = o\left(n^{-\frac{1}{4}} \right) \text{ a.s.,} \quad \max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} W_{nj}^c (t_i) \zeta_j \right| = o\left(n^{-\frac{1}{4}} \right) \text{ a.s.}
\tag{5.3}
$$

where $\{\zeta_i\}$ is a sequence of independent random variables satisfying $E\zeta_i = 0$ and $\sup_{1 \leq i \leq n} E|\zeta_i|^{p} < \infty$. By (A0) and Lemma 4.3, taking $r = 4$, $p_1 = 1/4$, $p_2 = 3/4$ in Lemma 4.2, one can also deduce that

$$
|B_{4n}| = 2n^{\frac{1}{4}} \cdot \left| \sum_{i=1}^{n} n^{-\frac{1}{4}} \delta_i \tilde{\xi}_i^c \mu_i \right| = O\left(n^{\frac{1}{2}} \log n \right) = o(n) \text{ a.s.}
\tag{5.4}
$$

Note that, from Lemma 4.3(a), (5.2) and (5.3), we have

$$\left|B_{3n}\right| \leq \sum_{i=1}^{n}\left|\delta_i\right|\cdot\max_{1\leq i\leq n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c\left(t_i\right)\mu_j\right|^2 = o\left(n^{\frac{1}{2}}\right)\text{a.s.} \tag{5.5}$$

$$\left|B_{5n}\right| \leq 2\sum_{i=1}^{n}\left|\delta_i\tilde{\xi}_i^c\right|\cdot\max_{1\leq i\leq n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c\left(t_i\right)\mu_j\right| = o\left(n^{\frac{3}{4}}\right)\text{a.s.} \tag{5.6}$$

$$\left|B_{6n}\right| \leq 2\left[\sum_{i=1}^{n}\left(\left|\delta_i\mu_i\right|-E\left|\delta_i\mu_i\right|\right)+\sum_{i=1}^{n}E\left|\delta_i\mu_i\right|\right]\cdot\max_{1\leq i\leq n}\left|\sum_{j=1}^{n}\delta_j W_{nj}^c\left(t_i\right)\mu_j\right| = o\left(n^{\frac{3}{4}}\right)\text{a.s.} \tag{5.7}$$

Therefore, for (5.2)-(5.7), one can deduce that
$S_{1n}^2 = B_{1n}+o\left(n\right)=B_{1n}+o\left(B_{1n}\right)\text{a.s.}$, which yields that

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

Therefore, by the Lemma 4.3(b), we can get that $S_{1n}^{-2}\leq Cn^{-1}$ a.s.

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

$$n^{-1}A_{1n}=n^{-1}\sum_{i=1}^{n}\delta_i\tilde{\xi}_i^c\eta_i=O\left(n^{-\frac{1}{2}}\log n\right)\text{a.s.}$$

Meanwhile, from (A0)-(A3), Lemma 4.3, (5.2) and (5.3), one can achieve that

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

$$S_n^{-2}A_{3n}\leq\frac{C}{n}\left|\sum_{i=1}^{n}\delta_i\left(\mu_i^2-\Xi_\mu^2\right)\beta\right|=o\left(1\right)\text{a.s.}$$

$$S_n^{-2}A_{4n}\leq\frac{C}{n}\left|\sum_{i=1}^{n}\delta_i\tilde{\xi}_i^c\tilde{g}_i^c\right|\frac{C}{n}\left[\max_{1\leq i\leq n}\sum_{i=1}^{n}\left|\delta_i\tilde{\xi}_i^c\right|\max_{1\leq i\leq n}\left|\tilde{g}_i^c\right|\right]=o\left(1\right)\text{a.s.}$$

$$S_n^{-2}A_{5n}\leq\frac{C}{n}\left|\sum_{i=1}^{n}\delta_i\tilde{\mu}_i^c\tilde{g}_i^c\right|\frac{C}{n}\left[\sum_{i=1}^{n}\left|\mu_i\right|\cdot\left|\delta_i\tilde{g}_i^c\right|+\sum_{i=1}^{n}\left|\delta_i\tilde{g}_i^c\right|\cdot\left|\sum_{j=1}^{n}\delta_j W_{nj}^c\left(t_i\right)\mu_j\right|\right]$$

$$\leq\frac{C}{n}\cdot\max_{1\leq i\leq n}\left(\sum_{i=1}^{n}\left(\left|\mu_i\right|-E\left|\mu_i\right|\right)+\sum_{i=1}^{n}E\left|\mu_i\right|\right)\cdot\max_{1\leq i\leq n}\left|\delta_i\tilde{g}_i^c\right|+o\left(n^{-\frac{1}{2}}\right)=o\left(1\right)\text{a.s.}$$

The proof of $n^{-1}A_{kn}=o\left(1\right)$ a.s. for $k=6,\cdots,12$ is analogous. Thus, the proof of Theorem 3.1(a) is completed. ∎

Proof of Theorem 3.1(b). From (3.2), for every $t\in[0,1]$, one can write that

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

$$=\sum_{j=1}^{n}W_{nj}^c\left(t\right)\delta_j\left[\xi_j\beta+g\left(t_j\right)+\epsilon_j-\left(\xi_j+\mu_j\right)\hat{\beta}_c\right]-g\left(t\right)$$

$$=\sum_{j=1}^{n}W_{nj}^c\left(t\right)\delta_j\xi_j\left(\beta-\hat{\beta}_c\right)+\sum_{j=1}^{n}W_{nj}^c\left(t_i\right)\delta_j\left[g\left(t_j\right)-g\left(t\right)\right]$$

$$+\sum_{j=1}^{n}W_{nj}^c\left(t_i\right)\delta_j\epsilon_j+\sum_{j=1}^{n}W_{nj}^c\left(t_i\right)\delta_j\mu_j\beta+\sum_{j=1}^{n}W_{nj}^c\left(t_i\right)\delta_j\mu_j\left(\hat{\beta}_c-\beta\right)$$

$$:=F_{1n}\left(t\right)+F_{2n}\left(t\right)+F_{3n}\left(t\right)+F_{4n}\left(t\right)+F_{5n}\left(t\right).$$

Therefore, we only need to prove that $F_{kn}(t) \to 0$ a.s. for every $t \in [0,1]$ and $k = 1, 2, \cdots, 5$. From (A0)-(A3), Theorem 3.1(*a*), Lemma 4.3, (2) and (3), for every $t \in [0,1]$ and any $a > 0$, one can get

$$F_{1n}(t) \le |\beta - \hat{\beta}_c| \cdot \max_{1 \le j \le n} \left| h(t_j) + v_j \right| \cdot \sum_{j=1}^{n} \delta_j W_{nj}^c(t) \right| = o\left(n^{-\frac{1}{4}}\right) \text{ a.s.}$$

$$F_{2n}(t) \le \sum_{j=1}^{n} \delta_j W_{nj}^c(t) \cdot \left[g(t_j) - g(t) \right] \cdot I\left(|t_j - t| > a \cdot n^{-\frac{1}{4}} \right)$$

$$+ \sum_{j=1}^{n} \delta_j W_{nj}^c(t) \cdot \left[g(t_j) - g(t) \right] \cdot I\left(|t_j - t| < a \cdot n^{-\frac{1}{4}} \right)$$

$$\le C \cdot a \cdot n^{-\frac{1}{4}} = o(1) \text{ a.s.}$$

$$F_{3n}(t) \le \left| \sum_{j=1}^{n} W_{nj}^c(t) \delta_j \epsilon_j \right| = o(1) \text{ a.s.}$$

$$F_{4n}(t) \le \left| \sum_{j=1}^{n} W_{nj}^c(t) \delta_j \mu_j \beta \right| = o(1) \text{ a.s.}$$

$$F_{5n}(t) \le |\beta - \hat{\beta}_c| \cdot \left| \sum_{j=1}^{n} W_{nj}^c(t) \delta_j \mu_j \right| = o(1) \text{ a.s.}$$

Thus, the proof of Theorem 3.1(*b*) is completed. ■

Proof of Theorem 3.2(*a*). From (3.3)-(3.4), write that

$$\hat{\beta}_I - \beta$$

$$= S_{2n}^{-2} \left\{ \sum_{i=1}^{n} \left(\tilde{\xi}_i + \tilde{\mu}_i \right) \left[\delta_i \left(y_i - \xi_i \beta - \mu_i \beta \right) + \left(1 - \delta_i \right) \left(\xi_i + \mu_i \right) \left(\hat{\beta}_c - \beta \right) \right. \right.$$

$$+ \left(1 - \delta_i \right) \hat{g}_n^c(t_i) - \sum_{j=1}^{n} W_{nj}(t_i) \left(\delta_j \left(y_j - \xi_j \beta - \mu_j \beta \right) \right.$$

$$\left. + \left(1 - \delta_j \right) \left(\xi_j + \mu_j \right) \left(\hat{\beta}_c - \beta \right) + \left(1 - \delta_j \right) \hat{g}_n^c(t_j) \right) \right] + \sum_{i=1}^{n} \delta_i \Xi_{\mu}^2 \beta \right\}$$

$$= S_{2n}^{-2} \left\{ \sum_{i=1}^{n} \left(\tilde{\xi}_i + \tilde{\mu}_i \right) \left[\delta_i \left(\epsilon_i - \mu_i \beta \right) + \delta_i \left(g(t_i) - \hat{g}_n^c(t_i) \right) \right. \right.$$

$$+ \left(1 - \delta_i \right) \left(\xi_i + \mu_i \right) \left(\hat{\beta}_c - \beta \right) + \hat{g}_n^c(t_i) - \sum_{j=1}^{n} W_{nj}(t_i) \left(\delta_j \left(\epsilon_j - \mu_j \beta \right) \right.$$

$$\left. + \delta_j \left(g(t_j) - \hat{g}_n^c(t_j) \right) + \left(1 - \delta_j \right) \left(\xi_j + \mu_j \right) \left(\hat{\beta}_c - \beta \right) + \hat{g}_n^c(t_j) \right) \right] + \sum_{i=1}^{n} \delta_i \Xi_{\mu}^2 \beta \right\}$$

$$= S_{2n}^{-2} \left\{ \sum_{i=1}^{n} \delta_i \tilde{\xi}_i \left(\epsilon_i - \mu_i \beta \right) + \sum_{i=1}^{n} \delta_i \tilde{\xi}_i \left(g(t_i) - \hat{g}_n^c(t_i) \right) \right.$$

$$+ \sum_{i=1}^{n} \tilde{\xi}_i \left(\xi_i + \mu_i \right) \left(1 - \delta_i \right) \left(\hat{\beta}_c - \beta \right) - \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_j W_{nj}(t_i) \tilde{\xi}_i \epsilon_j$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_j W_{nj}(t_i) \tilde{\xi}_i \mu_j \beta - \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_j W_{nj}(t_i) \tilde{\xi}_i \left(g(t_j) - \hat{g}_n^c(t_j) \right)$$

$$- \sum_{i=1}^{n} \sum_{j=1}^{n} W_{nj}(t_i) \tilde{\xi}_i \left(1 - \delta_j \right) \left(\xi_j + \mu_j \right) \left(\hat{\beta}_c - \beta \right) + \sum_{i=1}^{n} \tilde{\xi}_i \tilde{g}_i^c + \sum_{i=1}^{n} \delta_i \mu_i \epsilon_i$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_i W_{nj}(t_i)\epsilon_i\mu_j - \sum_{i=1}^{n}\delta_i\left(\mu_i^2 - \Xi_\mu^2\right)\beta + \sum_{i=1}^{n}\sum_{j=1}^{n}\delta_i W_{nj}(t_i)\mu_i\mu_j\beta$$

$$+\sum_{i=1}^{n}\delta_i\tilde{\mu}_i\left(g(t_i) - \hat{g}_n^c(t_i)\right) + \sum_{i=1}^{n}\tilde{\mu}_i(1-\delta_i)(\xi_i+\mu_i)\left(\hat{\beta}_c - \beta\right)$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_j W_{nj}(t_i)\mu_i\epsilon_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\delta_j W_{nj}(t_i)W_{nk}(t_i)\mu_k\epsilon_j$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_j W_{nj}(t_i)\mu_i\mu_j\beta - \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\delta_j W_{nj}(t_i)W_{nk}(t_i)\mu_j\mu_k\beta$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_j W_{nj}(t_i)\tilde{\mu}_i\left(g(t_j) - \hat{g}_n^c(t_j)\right)$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n}W_{nj}(t_i)(1-\delta_j)\tilde{\mu}_i(\xi_j+\mu_j)\left(\hat{\beta}_c - \beta\right) + \sum_{i=1}^{n}\tilde{\mu}_i\tilde{g}_i^c$$

$$:= S_{2n}^{-2}\sum_{k=1}^{21}D_{kn}.$$

Using a similar approach as step 1 in the proof of Theorem 3.1(a), one can get $S_{2n}^{-2} \leq Cn^{-1}$ a.s.

Therefore, we only need to verify that $n^{-1}D_{kn} = o(1)$ a.s. for $k = 1, 2, \cdots, 21$. From (A0)-(A4), Lemmas 4.2-4.4, Theorem 3.1(a), (5.2)-(5.4), we have

$$n^{-1}D_{1n} = n^{-1}\sum_{i=1}^{n}\delta_i\tilde{\xi}_i\left(\epsilon_i - \mu_i\beta\right) = n^{-1}\cdot O\left(n^{\frac{1}{2}}\log n\right) = o(1)\,\text{a.s.}$$

$$n^{-1}D_{2n} \leq n^{-1}\left[\sum_{i=1}^{n}\left|\delta_i\tilde{\xi}_i\right|\cdot\max_{1\leq i\leq n}\left|g(t_i) - \hat{g}_n^c(t_i)\right|\right] = o(1)\,\text{a.s.}$$

$$n^{-1}D_{3n} = n^{-1}\sum_{i=1}^{n}\tilde{\xi}_i^2(1-\delta_i)\left(\hat{\beta}_c - \beta\right) + n^{-1}\sum_{i=1}^{n}\tilde{\xi}_i(1-\delta_i)\sum_{j=1}^{n}W_{nj}(t_i)\xi_j\left(\hat{\beta}_c - \beta\right)$$

$$+ n^{-1}\sum_{i=1}^{n}\tilde{\xi}_i\mu_i(1-\delta_i)\left(\hat{\beta}_c - \beta\right) = o(1)\,\text{a.s.}$$

$$n^{-1}D_{4n} \leq n^{-1}\sum_{i=1}^{n}\left|\tilde{\xi}_i\right|\cdot\max_{1\leq i\leq n}\left|\sum_{j=1}^{n}\delta_j W_{nj}(t_i)\epsilon_j\right| = o(1)\,\text{a.s.}$$

In the same way, from (A0)-(A4), Lemmas 4.2-4.4, (5.2) and (5.3), one can similarly deduce that $n^{-1}D_{kn} = o\left(n^{-1/4}\right)$ a.s. for $k = 5, 6, \cdots, 21$. Thus, the proof of Theorem 3.2(a) is completed. ∎

Proof of Theorem 3.2(b). From (3.4), write that

$$\hat{g}_n^{[I]}(t) - g(t)$$

$$= \sum_{j=1}^{n}W_{nj}(t)\left\{\delta_j y_j + (1-\delta_j)\left[(\xi_j+\mu_j)\hat{\beta}_c + \hat{g}_n^c(t_j)\right] - (\xi_i+\mu_j)\hat{\beta}_I\right\} - g(t)$$

$$= \sum_{j=1}^{n}W_{nj}(t)\left\{\delta_j\left[\xi_j\beta + g(t_j) + \epsilon_j\right] + (1-\delta_j)\left[(\xi_j+\mu_j)\hat{\beta}_c + \hat{g}_n^c(t_j)\right]\right.$$

$$\left. - (\xi_j+\mu_j)\hat{\beta}_I\right\} - g(t)$$

$$= \sum_{j=1}^{n}W_{nj}(t)\left\{\delta_j\xi_j\beta + \delta_j g(t_j) + \delta_j\epsilon_j + \xi_j\hat{\beta}_c + \mu_j\hat{\beta}_c - \delta_j\xi_j\hat{\beta}_c - \delta_j\mu_j\hat{\beta}_c\right.$$

$$\left. + \hat{g}_n^c(t_j) - \delta_j\hat{g}_n^c(t_j) - \xi_j\hat{\beta}_I - \mu_j\hat{\beta}_I\right\} - g(t)$$

$$= \sum_{j=1}^{n} W_{nj}(t) \delta_j \xi_j \left(\beta - \hat{\beta}_c \right) + \sum_{j=1}^{n} W_{nj}(t) \delta_j \left[g(t_j) - \hat{g}_n^c(t_j) \right] + \sum_{j=1}^{n} W_{nj}(t) \delta_j \epsilon_j$$

$$+ \sum_{j=1}^{n} W_{nj}(t) \xi_j \left(\hat{\beta}_c - \beta \right) + \sum_{j=1}^{n} W_{nj}(t) \mu_j \left(\hat{\beta}_c - \beta \right) - \sum_{j=1}^{n} W_{nj}(t) \delta_j \mu_j \hat{\beta}_c$$

$$+ \sum_{j=1}^{n} W_{nj}(t) \left[\hat{g}_n^c(t_j) - g(t_j) \right] + \sum_{j=1}^{n} W_{nj}(t) \left[g(t_j) - g(t) \right]$$

$$+ \sum_{j=1}^{n} W_{nj}(t) \xi_j \left(\beta - \hat{\beta}_I \right) + \sum_{j=1}^{n} W_{nj}(t) \mu_j \left(\beta - \hat{\beta}_I \right)$$

$$:= \sum_{k=1}^{10} G_{kn}.$$

Therefore, we only need to prove that $G_{kn}(t) \to 0$ a.s. for every $t \in [0,1]$ and $k = 1, 2, \cdots, 10$. From (A0)-(A4), Lemma 4.3-4.4, (5.2), (5.3), one can get

$$G_{1n}(t) \le \left| \beta - \hat{\beta}_c \right| \cdot \max_{1 \le j \le n} \left| \delta_j \xi_j \right| \cdot \sum_{j=1}^{n} W_{nj}(t)$$

$$\le \left| \beta - \hat{\beta}_c \right| \cdot \max_{1 \le j \le n} \left| \xi_j \right| \cdot \sum_{j=1}^{n} W_{nj}(t) = o(1) \text{ a.s.}$$

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

Meanwhile, the proof of $G_{kn}(t) \to 0$ a.s. for every $t \in [0,1]$ and $k = 3, \cdots, 10$ is analogous. Thus, the proof of Theorem 3.2(*b*) is completed. ∎

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Engle, R.F., Granger, C.W.J., Rice, J. and Weiss, A. (1986) Semiparametric Estimation of the Relation between Weather and Electricity Sales. *Journal of the American Statistical Association*, **81**, 310-320.
https://doi.org/10.1080/01621459.1986.10478274

[2] Liu, J.X. and Chen, X.R. (2005) Consistency of LS Estimator in Simple Linear EV Regression Models. *Acta Mathematica Scientia*, *Series B*, **25**, 50-58.
https://doi.org/10.1016/S0252-9602(17)30260-6

[3] Miao, Y., Yang, G. and Shen, L. (2007) The Central Limit Theorem for LS Estimator in Simple Linear EV Regression Models. *Communications in Statistics—Theory and Methods*, **36**, 2263-2272. https://doi.org/10.1080/03610920701215266

[4] Miao, Y. and Liu, W. (2009) Moderate Deviations for LS Estimator in Simple Linear EV Regression Model. *Journal of Statistical Planning and Inference*, **139**, 3122-3131.

https://doi.org/10.1016/j.jspi.2009.02.021

[5] Fan, G.L., Liang, H.Y., Wang, J.F. and Xu, H.X. (2010) Asymptotic Properties for LS Estimators in EV Regression Model with Dependent Errors. *Advances in Statistical Analysis*, **94**, 89-103. https://doi.org/10.1007/s10182-010-0124-3

[6] Cui, H.J. and Li, R.C. (1998) On Parameter Estimation for Semi-Linear Errors-in-Variables Models. *Journal of Multivariate Analysis*, **64**, 1-24. https://doi.org/10.1006/jmva.1997.1712

[7] Liang, H., Härdle, W. and Carrol, R.J. (1999) Estimation in a Semiparametric Partially Linear Errosr-in-Variables Model. *The Annals of Statistics*, **27**, 1519-1935.

[8] Zhou, H.B., You, J.H. and Zhou, B. (2010) Statistical Inference for Fixed-Effects Partially Linear Regression Models with Errors in Variables. *Statistical Papers*, **51**, 629-650. https://doi.org/10.1007/s00362-008-0150-3

[9] Wang, Q., Linton, O. and Härdle, W. (2004) Semiparametric Regression Analysis with Missing Response at Random. *Journal of the American Statistical Association*, **99**, 334-345. https://doi.org/10.1198/016214504000000449

[10] Wang, Q. and Sun, Z. (2007) Estimation in Partially Linear Models with Missing Responses at Random. *Journal of Multivariate Analysis*, **98**, 1470-1493. https://doi.org/10.1016/j.jmva.2006.10.003

[11] Bianco, A., Boente, G., Gonzlez-Manteiga, W. and Prez-Gonzlez, A. (2010) Estimation of the Marginal Location under a Partially Linear Model with Missing Responses. *Computational Statistics & Data Analysis*, **54**, 546-564. https://doi.org/10.1016/j.csda.2009.09.028

[12] Healy, M.J.R. and Westmacott, M. (1956) Missing Values in Experiments Analysis on Automatic Computers. *Journal of Applied Statistics*, **5**, 203-206. https://doi.org/10.2307/2985421

[13] Cheng, P.E. (1994) Nonparametric Estimation of Mean Functionals with Data Missing at Random. *Journal of the American Statistical Association*, **89**, 81-87. https://doi.org/10.1080/01621459.1994.10476448

[14] Härdle, W., Liang, H. and Gao, J.T. (2000) Partial Linear Models. Physica-Verlag, Heidelberg. https://doi.org/10.1007/978-3-642-57700-0

[15] Gao, J.T., Chen, X.R. and Zhao, L.C. (1994) Asymptotic Normality of a Class of Estimators in Partial Linear Models. *Acta Mathematica Sinica*, **37**, 256-268.

[16] Chen, H. (1988) Convergence Rates for Parametric Components in a Partly Linear Model. *The Annals of Statistics*, **16**, 136-146. https://doi.org/10.1214/aos/1176350695

[17] Baek, J.I. and Liang, H.Y. (2006) Asymptotic of Estimators in Semi-Parametric Model under NA Samples. *Journal of Statistical Planning and Inference*, **136**, 3362-3382. https://doi.org/10.1016/j.jspi.2005.01.008