Research Article

Statistical Inference for the Heteroscedastic Partially Linear Varying-Coefficient Errors-in-Variables Model with Missing Censoring Indicators

Yuye Zou ^{1,2} and Chengxin Wu ^{3,4}

¹Key Laboratory of Advanced Theory and Application in Statistics and Data Science - MOE, School of Statistics, East China Normal University, Shanghai 200062, China

²College of Economics and Management, Shanghai Maritime University, Shanghai 201306, China

³School of Mathematics, Hefei University of Technology, Hefei 230009, China

⁴School of Mathematics and Statistics, Huangshan University, Huangshan 245041, China

Correspondence should be addressed to Chengxin Wu; hsuchengxinwu@foxmail.com

Received 16 April 2021; Accepted 26 May 2021; Published 7 June 2021

Academic Editor: Chris Goodrich

Copyright © 2021 Yuye Zou and Chengxin Wu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we focus on heteroscedastic partially linear varying-coefficient errors-in-variables models under right-censored data with censoring indicators missing at random. Based on regression calibration, imputation, and inverse probability weighted methods, we define a class of modified profile least square estimators of the parameter and local linear estimators of the coefficient function, which are applied to constructing estimators of the error variance function. In order to improve the estimation accuracy and take into account the heteroscedastic error, reweighted estimators of the parameter and coefficient function are developed. At the same time, we apply the empirical likelihood method to construct confidence regions and maximum empirical likelihood estimators of the parameter. Under appropriate assumptions, the asymptotic normality of the proposed estimators is studied. The strong uniform convergence rate for the estimators of the error variance function is considered. Also, the asymptotic chi-squared distribution of the empirical log-likelihood ratio statistics is proved. A simulation study is conducted to evaluate the finite sample performance of the proposed estimators. Meanwhile, one real data example is provided to illustrate our methods.

1. Introduction

In regression analysis, for a long period of time, the flexible and refined statistical regression models are widely applied in theoretical study and practical application. The main results related to parameter regression models and non-parameter regression models are rather mature. Recently, semiparameter regression models can reduce the high risk of misspecification related to parameter regression models and avoid the “curse of dimensionality” for nonparametric regression models. Thanks to their advantage, semiparametric regression models enjoy consideration attention from statisticians. Semiparametric regression models have various forms. Specially, partially linear varying-coefficient errors-

in-variables (PLVCEV) model, as a typical example, was introduced by You and Chen [1] and has the following form:

$$\begin{cases} V = X^\top \beta + Z^\top \alpha(U) + \varepsilon, \\ \xi = X + e, \end{cases} \quad (1)$$

where V is the response variable, $X \in \mathbb{R}^p$, $Z \in \mathbb{R}^q$ are the covariates, $\beta = (\beta_1, \dots, \beta_p)^\top$ is a vector of p -dimensional unknown parameter, $\alpha(\cdot) = (\alpha_1(\cdot), \dots, \alpha_q(\cdot))^\top$ is an unknown q -dimensional vector of coefficient function, and ε is the random error. The measurement error e is independent of (X, Z, U) with mean zero and covariance matrix Σ_e . In order to identify the model, Σ_e is assumed to be known.

As one general and flexible semiparametric model, model (1) includes a variety of models of interest. When X is observed exactly, model (1) boils down to be PLVC model [2, 3]. When $Z = 1$, $q = 1$ and X is observed exactly, model (1) reduces to partially linear regression model [4]. When X is observed exactly, and $\alpha(\cdot)$ is a constant vector, model (1) becomes a linear regression model. When $q = 1$ and $Z = 1$, model (1) reduces to partially linear EV model [5]. For model (1), You and Chen [1] proposed estimators of parametric and nonparametric components and showed their asymptotic properties. Liu and Liang [6] constructed the asymptotical normality of jackknife estimator for error variance and standard chi-square distribution of jackknife empirical log-likelihood statistic. Fan et al. [7] established penalized profile least squares estimation of parameter and nonparameter in the model.

The literature mentioned above assumed that the random errors are homoscedastic, which means that the random error ε is independent of (X, Z, U) . However, in many practical application fields, the error variance function may change with the variables. Heteroscedastic error models have attracted much attention of many scholars. For example, You et al. [8] considered the estimation of parametric and nonparametric parts for partially linear regression models with heteroscedastic errors. Fan et al. [9] constructed confidence regions of parameter for heteroscedastic PLVCEV model based on empirical likelihood method. Shen et al. [10] discussed estimation and inference for PLVC model with heteroscedastic errors. Xu and Duan [11] extended the results of Shen et al. [10] to efficient estimation for PLVCEV model with heteroscedastic errors.

The above related works assumed that the responses are observed completely. However, in many practical fields, especially in biomedical studies and survival analysis, the response cannot be completely observed due to censored variables. Huang and Huang [12, 13] discussed the constructed confidence regions of the parameters for varying-coefficient single-index model and partially linear single-index EV model by empirical likelihood method under censored data, respectively. The aforementioned results require that the censoring indicators be always observed. However, the censoring indicators may not be observed completely. For example, the death of individual is attributable to the cause of interest that may require information that is not gathered or lost due to various reasons [14]. In this paper, we assume that the censoring indicators are missing at random (MAR), which is common and reasonable in statistical analysis with missing data [15]. There are a lot of works related to missing censoring indicators. For example, Wang and Dinse [16] and Li and Wang [17] proposed weighted least square estimators of unknown parameter and proved their asymptotical normality for linear regression model. Shen and Liang [18] discussed the estimation and variable selection for PLVC quantile regression model. Wang et al. [19] considered composite quantile regression for linear regression model. However, there is no literature focusing on the estimation and confidence regions of heteroscedastic errors model with right-censored data when the censoring indicators are MAR.

In this paper, we consider modified profile least square (PLS) estimators of the unknown parameter and local linear estimators of the coefficient function. Besides the point estimation, we are also interested in interval estimation in terms of empirical likelihood (EL) method, which, first introduced by Owen [20], is a very effective method for constructing confidence regions, which enjoys a lot of nice properties over the normal approximation-based methods and bootstrap approach. Thanks to its advantage, there are a lot of literature-related EL methods to refer to. For instance, Fan et al. [21] considered penalized EL for high-dimensional PLCVEV model. Wang and Drton [22] established estimation for linear structural equation models with dependent errors based on EL method. Fan et al. [23] discussed weighted EL for heteroscedastic varying-coefficient partially nonlinear model with missing data. Zou et al. [24] considered EL inference for partially linear single-index EV model with missing censoring indicators.

It is worth pointing out that it is innovative and interesting in studying the PLVCEV model with heteroscedastic errors under censoring indicators MAR. Thus, we consider estimation and confidence regions based on modified profiled LS method and EL inference, respectively. The main aims of this paper include the following aspects: (1) define a class of modified PLS estimators of the parameter and local linear estimators of coefficient function based on regression calibration, imputation, and inverse probability weighted approaches, and prove the asymptotical normality of the proposed estimators; (2) construct reweighted estimators of the parameter and coefficient function based on estimators of the error variance function, and establish the asymptotic properties of the proposed estimators; (3) develop the asymptotic standard chi-squared distribution of the empirical log-likelihood ratio functions, construct the confidence regions for the parameter, and propose the asymptotic distribution of the corresponding maximum EL estimators. Finally, a simulation study and a real data analysis are conducted to demonstrate the finite sample performance of the proposed procedures.

The rest of this paper is organized as follows. In Section 2, we construct modified PLE estimators of the parameter and local linear estimators of the coefficient function. In Section 3, we proposed empirical log-likelihood ratio statistics and maximum EL estimators. The main results are shown in Section 4. Section 5 presents simulation and real data analysis. In Section 6, we show some conclusions. The proofs of the main results are shown in Appendix.

2. Methodology

Suppose that $\{(V_i, X_i, Z_i, U_i), i = 1, \dots, n\}$ is a sample from model (1), that is,

$$\begin{cases} V_i = X_i^\top \beta + Z_i^\top \alpha(U_i) + \varepsilon_i, \\ \xi_i = X_i + e_i, \end{cases} \quad (2)$$

where the model error ε_i satisfies $E(\varepsilon_i | X_i, Z_i, U_i) = 0$ and $\text{Var}(\varepsilon_i | X_i, Z_i, U_i) = \sigma^2(U_i)$, which is an unknown function of U_i representing heteroscedastic error. In the practical

application, the response V_i may be right censored by various reasons. Let C_i be censoring time with distribution function (df) $G(\cdot)$. One can only observe $Y_i = \min(V_i, C_i)$ with df H and censoring indicator $\delta_i = I(V_i \leq C_i)$. Define the missing indicator to be ς_i , which is 0 if δ_i is missing; otherwise, it is 1. Throughout this article, we assume that V_i is independent of C_i , and δ_i is MAR, which implies that ς_i and δ_i are conditional independent given $T_i = (Y_i, X_i, Z_i, U_i)$, i.e.,

$$P(\delta_i = 1 | Y_i, X_i, Z_i, U_i, \varsigma_i) = P(\delta_i = 1 | Y_i, X_i, Z_i, U_i) := m(T_i). \quad (3)$$

$$\alpha_j(u) = \alpha_j(u_0) + \alpha'_j(u_0)(u - u_0) := a_j + b_j(u - u_0), \quad j = 1, \dots, q, \quad (4)$$

where $\alpha'_j(u) = \partial \alpha_j(u) / \partial u$. Then, (a_j, b_j) can be estimated by minimizing the following objective function:

$$\sum_{i=1}^n \frac{\delta_i}{1 - G(Y_i)} \left\{ Y_i - X_i^\top \beta - \sum_{j=1}^n [a_j + b_j(U_i - u_0)] Z_{ij} \right\}^2 K\left(\frac{U_i - u}{h_{1n}}\right), \quad (5)$$

where $K(\cdot)$ is a kernel function, and $0 < h_{1n} \rightarrow 0$ is a bandwidth sequence. Due to the missing indicators, some δ_i cannot be observed. Therefore, model (2) cannot be applied

2.1. Modified Profile Least Squares Estimation. The local linear regression technique is employed to estimate the coefficient function $\alpha(\cdot)$. If $\alpha(\cdot)$ has twice continuous derivative at point u_0 , for u in a small neighborhood of u_0 , one can approximate $\alpha(\cdot)$ by the following expansion with Taylor expansion:

directly. One can replace δ_i with its conditional expectation $m(T_i) = E[\delta_i | T_i]$. Thus, (a_j, b_j) can be defined as the minimizer of

$$\sum_{i=1}^n \frac{m(T_i)}{1 - G(Y_i)} \left\{ Y_i - X_i^\top \beta - \sum_{j=1}^n [a_j + b_j(U_i - u_0)] Z_{ij} \right\}^2 K\left(\frac{U_i - u}{h_{1n}}\right). \quad (6)$$

However, in practical fields, function $m(\cdot)$ is usually unknown. One can use parametric and nonparametric methods to estimate $m(\cdot)$. However, when the covariates are high-dimensional, nonparametric estimation may cause “the curse of dimensionality.” Hence, throughout this paper, we assume that $m(\cdot)$ follows a parametric model $m(T) = m(T, \theta)$, where θ is an unknown parameter vector. Following Wang and Dinse [16], the estimator $\hat{\theta}_n$ of θ can be obtained by maximizing the following likelihood function:

$$\prod_{i=1}^n m(T_i, \theta)^{\varsigma_i \delta_i} (1 - m(T_i, \theta))^{\varsigma_i (1 - \delta_i)}. \quad (7)$$

Let $\delta_i^c = m(T_i, \theta)$. Since $E[\delta_i^c] = E[m(T_i, \theta)] = E[E(\delta_i | T_i)] = E[\delta_i]$, we replace δ_i^c with its estimator $\hat{\delta}_i^c = m(T_i, \hat{\theta}_n)$. Hence, (a_j, b_j) can be estimated by minimizing the following objective function:

$$\sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} \left\{ Y_i - X_i^\top \beta - \sum_{j=1}^n [a_j + b_j(U_i - u_0)] Z_{ij} \right\}^2 K\left(\frac{U_i - u}{h_{1n}}\right), \quad (8)$$

where $\hat{G}_n(Y_i)$ is the estimator of $G(\cdot)$, which is defined by

$$\hat{G}_n(y) = 1 - \prod_{i: Y_i \leq y} \left(\frac{n - R_i}{n - R_i + 1} \right)^{1 - \hat{\mu}_n(y)},$$

$$R_i = \sum_{j=1}^n I(Y_j \leq Y_i), \quad (9)$$

$$\hat{\mu}_n(y) = \frac{\sum_{i=1}^n \delta_i \zeta_i L((y - Y_i)/a_n)}{\sum_{i=1}^n \zeta_i L((y - Y_i)/a_n)},$$

which is the Nadaraya-Watson estimator of $\mu(y) = E[\delta_i | Y_i = y]$ with the kernel function $L(\cdot)$ and bandwidth sequence $0 < a_n \rightarrow 0$.

For notational simplicity, let $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, $\mathbf{X} = (X_1, \dots, X_n)^\top$, $\mathbf{Z} = (Z_1, \dots, Z_n)^\top$,

$$D_u = \begin{pmatrix} Z_1^\top & \frac{U_1 - u}{h_{1n}} Z_1^\top \\ \vdots & \vdots \\ Z_n^\top & \frac{U_n - u}{h_{1n}} Z_n^\top \end{pmatrix},$$

$$M = \begin{pmatrix} Z_1^\top \alpha(U_1) \\ \vdots \\ Z_n^\top \alpha(U_n) \end{pmatrix},$$

$$\tilde{Y}_i = Y_i - S_i \mathbf{Y}, \quad (10)$$

$$\tilde{X}_i = X_i - S_i \mathbf{X},$$

$$\tilde{\xi}_i = \xi_i - S_i \xi,$$

$$\xi = (\xi_1, \dots, \xi_n),$$

$$\omega_u = \text{diag} \left(K \left(\frac{U_1 - u}{h_{1n}} \right), \dots, K \left(\frac{U_n - u}{h_{1n}} \right) \right),$$

$$S_i = (Z_i^\top 0) (D_u^\top \omega_u \hat{\Delta}_c D_u)^{-1} D_u^\top \hat{\Delta}_c \omega_u,$$

$$\hat{\Delta}_c = \text{diag} \left(\frac{\hat{\delta}_1^c}{(1 - \hat{G}_n(Y_1))}, \dots, \frac{\hat{\delta}_n^c}{(1 - \hat{G}_n(Y_n))} \right).$$

If β is known, one can obtain the local linear estimator of coefficient function by

$$\tilde{\alpha}_c(u) = (I_q 0) (D_u^\top \omega_u \hat{\Delta}_c D_u)^{-1} D_u^\top \hat{\Delta}_c \omega_u (\mathbf{Y} - \mathbf{X}\beta). \quad (11)$$

Substituting (11) into the original model (8) and eliminating bias produced by the measurement error, we get the following modified PLS estimator of the parameter based on regression calibration method,

$$\tilde{\beta}_c = \left\{ \sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) \right\}^{-1} \sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} \tilde{\xi}_i \tilde{Y}_i. \quad (12)$$

Then, the local linear regression estimator of $\alpha(u)$ is defined as follows:

$$\tilde{\alpha}_c(u) = (I_q 0) (D_u^\top \omega_u \hat{\Delta}_c D_u)^{-1} D_u^\top \hat{\Delta}_c \omega_u (\mathbf{Y} - \xi \tilde{\beta}_c). \quad (13)$$

Let $\delta_i^m = \zeta_i \delta_i + (1 - \zeta_i) m(T_i, \theta)$. Since $E[\delta_i^m] = E[\zeta_i \delta_i + (1 - \zeta_i) m(T_i, \theta)] = E[\delta_i]$ under the missing mechanism, we can impute δ_i with $\hat{\delta}_i^m = \zeta_i \delta_i + (1 - \zeta_i) m(T_i, \hat{\theta}_n)$ in expression (6). Hence, (a_j, b_j) can be estimated by minimizing the following objective function:

$$\sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} \left\{ Y_i - X_i^\top \beta - \sum_{j=1}^n [a_j + b_j (U_i - u_0)] Z_{ij} \right\}^2 K \left(\frac{U_i - u}{h_{1n}} \right). \quad (14)$$

If β is known, one can obtain the local linear estimator of coefficient function by

$$\tilde{\alpha}_m(u) = (I_q 0) (D_u^\top \omega_u \hat{\Delta}_m D_u)^{-1} D_u^\top \hat{\Delta}_m \omega_u (\mathbf{Y} - \mathbf{X}\beta), \quad (15)$$

where $\hat{\Delta}_m = \text{diag}(\hat{\delta}_1^m / (1 - \hat{G}_n(Y_1)), \dots, \hat{\delta}_n^m / (1 - \hat{G}_n(Y_n)))$. Substituting (21) into the original model and eliminating bias produced by the measurement error, hence, we obtain the following modified PLS estimator of β based on imputation method:

$$\tilde{\beta}_m = \left\{ \sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) \right\}^{-1} \sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} \tilde{\xi}_i \tilde{Y}_i. \quad (16)$$

Thus, the local linear regression estimator of $\alpha(\cdot)$ is defined as follows:

$$\tilde{\alpha}_m(u) = (I_q 0) (D_u^\top \omega_u \hat{\Delta}_m D_u)^{-1} D_u^\top \hat{\Delta}_m \omega_u (\mathbf{Y} - \xi \tilde{\beta}_m). \quad (17)$$

Let $\delta_i^w = (\zeta_i \delta_i / \pi(Y_i)) + (1 - (\zeta_i / \pi(Y_i))) m(T_i, \theta)$. Note that

$E[\delta_i^w] = E[(\zeta_i \delta_i / \pi(Y_i)) + (1 - (\zeta_i / \pi(Y_i))) m(T_i, \theta)] = E[\delta]$ under MAR assumption. Hence, we substitute δ_i with $\hat{\delta}_i^w = (\zeta_i \delta_i / \hat{\pi}_n(Y_i)) + (1 - (\zeta_i / \hat{\pi}_n(Y_i))) m(T_i, \hat{\theta}_n)$, where $\hat{\pi}_n(y) = (\sum_{i=1}^n \zeta_i \Omega((y - Y_i)/b_n) / \sum_{i=1}^n \Omega((y - Y_i)/b_n))$ is a nonparametric estimator of $\pi(y) = E[\zeta | Y = y]$ with kernel function $\Omega(\cdot)$ and bandwidth sequence $0 < b_n \rightarrow 0$. Hence, (a_j, b_j) can be estimated by minimizing the following objective function:

$$\sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} \left\{ Y_i - X_i^\top \beta - \sum_{j=1}^n [a_j + b_j (U_i - u_0)] Z_{ij} \right\}^2 K \left(\frac{U_i - u}{h_{1n}} \right). \quad (18)$$

If β is known, one can obtain the local linear estimator of coefficient function by

$$\tilde{\alpha}_w(u) = (I_q 0) (D_u^\top \omega_u \hat{\Delta}_w D_u)^{-1} D_u^\top \hat{\Delta}_w \omega_u (\mathbf{Y} - \mathbf{X}\beta), \quad (19)$$

where $\hat{\Delta}_w = \text{diag}(\hat{\delta}_1^w / (1 - \hat{G}_n(Y_1)), \dots, \hat{\delta}_n^w / (1 - \hat{G}_n(Y_n)))$. Substituting (19) into the original model and eliminating bias produced by the measurement error. Hence, we can get the following modified PLS estimator of β based on inverse probability weighted method:

$$\tilde{\beta}_w = \left\{ \sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) \right\}^{-1} \sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} \tilde{\xi}_i \tilde{Y}_i. \quad (20)$$

Hence, the local linear regression estimator of $\alpha(\cdot)$ is defined as follows:

$$\sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} \left\{ \left[(Y_i - \xi_i^\top \tilde{\beta}_c - Z_i^\top \tilde{\alpha}_c(U_i))^2 - \tilde{\beta}_c^\top \Sigma_e \tilde{\beta}_c \right] - \mu_1 - \nu_1 (U_i - u) \right\}^2 K\left(\frac{U_i - u}{h_{2n}}\right), \quad (22)$$

the local linear regression estimator of $\sigma^2(u)$ based on regression calibration method is defined by

$$\hat{\sigma}_c^2(u) = \sum_{i=1}^n W_{ni}^c \left\{ [Y_i - \xi_i^\top \tilde{\beta}_c - Z_i^\top \tilde{\alpha}_c(U_i)]^2 - \tilde{\beta}_c^\top \Sigma_e \tilde{\beta}_c \right\}, \quad (23)$$

$$\tilde{\alpha}_w(u) = (I_q \ 0) (D_u^\top \omega_u \hat{\Delta}_w D_u)^{-1} D_u^\top \hat{\Delta}_w \omega_u (Y - \xi \tilde{\beta}_w). \quad (21)$$

2.2. Estimation for Error Variance. In order to improve the estimation of parametric and nonparametric parts, we construct local linear estimators of the error variance function $\sigma^2(\cdot)$ in this subsection. Note that $\sigma^2(u) = E\{\delta_i^c / (1 - G(Y_i)) [(Y_i - \xi_i^\top \beta - Z_i \alpha(U_i))^2 - \beta^\top \Sigma_e \beta] | U_i\}$. By minimizing the following object function with respect to μ_1 ,

where the weight function $W_{ni}^c(\cdot)$ is defined by

$$W_{ni}^c(u) = \frac{(nh_{2n})^{-1} (\hat{\delta}_i^c / (1 - \hat{G}_n(Y_i))) K((U_i - u)/h_{2n}) (A_{n2}^c(u) - (U_i - u)A_{n1}^c(u))}{A_{n0}^c(u)A_{n2}^c(u) - [A_{n1}^c(u)]^2}, \quad (24)$$

with

$$A_{nj}^c = \frac{1}{nh_{2n}} \sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} K\left(\frac{U_i - u}{h_{2n}}\right) (U_i - u)^j, \quad \text{for } j = 0, 1, 2. \quad (25)$$

Note that

$$\sigma^2(u) = E\left\{ \frac{\delta_i^m}{1 - G(Y_i)} \left[(Y_i - \xi_i^\top \beta - Z_i \alpha(U_i))^2 - \beta^\top \Sigma_e \beta \right] | U_i \right\}. \quad (26)$$

By minimizing the following object function with respect to μ_2 ,

$$\sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} \left\{ \left[(Y_i - \xi_i^\top \tilde{\beta}_m - Z_i^\top \tilde{\alpha}_m(U_i))^2 - \tilde{\beta}_m^\top \Sigma_e \tilde{\beta}_m \right] - \mu_2 - \nu_2 (U_i - u) \right\}^2 K\left(\frac{U_i - u}{h_{2n}}\right), \quad (27)$$

the local linear regression estimator of $\sigma_m^2(u)$ based on imputation method is defined by

$$\hat{\sigma}_m^2(u) = \sum_{i=1}^n W_{ni}^m \left\{ [Y_i - \xi_i^\top \tilde{\beta}_m - Z_i^\top \tilde{\alpha}_m(U_i)]^2 - \tilde{\beta}_m^\top \Sigma_e \tilde{\beta}_m \right\}, \quad (28)$$

where the weight function $W_{ni}^m(\cdot)$ is defined by

$$W_{ni}^m(u) = \frac{(nh_{2n})^{-1}(\hat{\delta}_i^m / (1 - \hat{G}_n(Y_i))) K((U_i - u)/h_{2n})(A_{n2}^m(u) - (U_i - u)A_{n1}^m(u))}{A_{n0}^m(u)A_{n2}^m(u) - [A_{n1}^m(u)]^2}, \quad (29)$$

with

$$A_{nj}^m = \frac{1}{nh_{2n}} \sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} K\left(\frac{U_i - u}{h_{2n}}\right) (U_i - u)^j, \quad \text{for } j = 0, 1, 2. \quad (30)$$

Note that

$$\sigma^2(u) = E \left\{ \frac{\delta_i^w}{1 - G(Y_i)} \left[(Y_i - \xi_i^\top \beta - Z_i \alpha(U_i))^2 - \beta^\top \Sigma_e \beta \right] | U_i \right\}. \quad (31)$$

By minimizing the following object function with respect to μ_3 ,

$$\sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} \left\{ \left[(Y_i - \xi_i^\top \tilde{\beta}_w - Z_i^\top \tilde{\alpha}_w(U_i))^2 - \tilde{\beta}_w^\top \Sigma_e \tilde{\beta}_w \right] - \mu_3 - \nu_3(U_i - u) \right\}^2 K\left(\frac{U_i - u}{h_{2n}}\right), \quad (32)$$

the local linear regression estimator of $\sigma^2(u)$ based on inverse probability weighted method is defined by

$$\hat{\sigma}_w^2(u) = \sum_{i=1}^n W_{ni}^w \left\{ \left[Y_i - \xi_i^\top \tilde{\beta}_w - Z_i^\top \tilde{\alpha}_w(U_i) \right]^2 - \tilde{\beta}_w^\top \Sigma_e \tilde{\beta}_w \right\}, \quad (33)$$

where the weight function $W_{ni}^w(\cdot)$ is defined by

$$W_{ni}^w(u) = \frac{(nh_{2n})^{-1}(\hat{\delta}_i^w / (1 - \hat{G}_n(Y_i))) K((U_i - u)/h_{2n})(A_{n2}^w(u) - (U_i - u)A_{n1}^w(u))}{A_{n0}^w(u)A_{n2}^w(u) - [A_{n1}^w(u)]^2}, \quad (34)$$

with

$$A_{nj}^w = \frac{1}{nh_{2n}} \sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} K\left(\frac{U_i - u}{h_{2n}}\right) (U_i - u)^j, \quad \text{for } j = 0, 1, 2. \quad (35)$$

2.3. Reweighted Estimation. In this subsection, we construct the reweighted estimations of the parametric and non-parametric parts based on the error variance estimator $\hat{\sigma}_c^2(u)$ given in (23). By minimizing the following object function,

$$\sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_c^{-2}(U_i) \left\{ \left[Y_i - \xi_i^\top \beta - \sum_{j=1}^n (a_j + b_j(U_i - u_0)) Z_{ij} \right]^2 - \beta^\top \Sigma_e \beta \right\} K\left(\frac{U_i - u}{h_{3n}}\right), \quad (36)$$

then, we get the following reweighted estimator of β based on the regression calibration method:

$$\hat{\beta}_c = \left\{ \sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_c^{-2}(U_i) (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) \right\}^{-1} \cdot \sum_{i=1}^n \frac{\hat{\delta}_i^c}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_c^{-2}(U_i) \tilde{\xi}_i \tilde{Y}_i. \quad (37)$$

Furthermore, the reweighted estimator of the coefficient function $\alpha(u)$ is defined by

$$\hat{\alpha}_c(u) = (I_q 0) (D_u^\top \omega_u \hat{\Delta}_c D_u)^{-1} D_u^\top \hat{\Delta}_c \omega_u (\mathbf{Y} - \xi \hat{\beta}_c). \quad (38)$$

Similarly, based on the error variance estimator $\hat{\sigma}_m^2(u)$ given in (28) and minimizing the following object function

$$\sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_m^{-2}(U_i) \left\{ \left[Y_i - \xi_i^\top \beta - \sum_{j=1}^n (a_j + b_j(U_i - u_0)) Z_{ij} \right]^2 - \beta^\top \Sigma_e \beta \right\} K\left(\frac{U_i - u}{h_{3n}}\right), \quad (39)$$

then, we get the reweighted estimator of β based on the imputation method:

$$\hat{\beta}_m = \left\{ \sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_m^{-2}(U_i) (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) \right\}^{-1} \sum_{i=1}^n \frac{\hat{\delta}_i^m}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_m^{-2}(U_i) \tilde{\xi}_i \tilde{Y}_i. \quad (40)$$

Hence, the reweighted estimator of the coefficient function $\alpha(u)$ is defined by

$$\hat{\alpha}_m(u) = (I_q \ 0) (D_u^\top \omega_u \hat{\Delta}_m D_u)^{-1} D_u^\top \hat{\Delta}_m \omega_u (\mathbf{Y} - \xi \hat{\beta}_m). \quad (41)$$

From the error variance estimator $\hat{\sigma}_w^2(u)$ given in (33) and minimizing the following object function

$$\sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_w^{-2}(U_i) \left\{ \left[Y_i - \xi_i^\top \beta - \sum_{j=1}^n (a_j + b_j(U_i - u_0)) Z_{ij} \right]^2 - \beta^\top \Sigma_e \beta \right\} K\left(\frac{U_i - u}{h_{3n}}\right), \quad (42)$$

then, we get the reweighted estimator of β based on the inverse probability weighted method:

$$\hat{\beta}_w = \left\{ \sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_w^{-2}(U_i) (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) \right\}^{-1} \sum_{i=1}^n \frac{\hat{\delta}_i^w}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_w^{-2}(U_i) \tilde{\xi}_i \tilde{Y}_i. \quad (43)$$

Thus, the reweighted estimator of the coefficient function $\alpha(u)$ is defined by

$$\hat{\alpha}_w(u) = (I_q \ 0) (D_u^\top \omega_u \hat{\Delta}_w D_u)^{-1} D_u^\top \hat{\Delta}_w \omega_u (\mathbf{Y} - \xi \hat{\beta}_w). \quad (44)$$

3. Empirical Likelihood

The confidence regions of the parameter can be constructed by the asymptotic distribution of Theorems 1 and 4. However, the estimation of asymptotic covariance is quite complicated. In this section, we shall employ the EL method to construct confidence regions for β , which avoids to estimate the complicated covariance.

3.1. Regression Calibration Empirical Likelihood. We introduce the following auxiliary random vector based on regression calibration method:

$$\hat{\eta}_{i,c}(\beta) = \frac{m(T_i, \hat{\theta}_n)}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_c^{-2}(U_i) \left[\tilde{\xi}_i (\tilde{Y}_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta \right]. \quad (45)$$

Thus, we define the empirical log-likelihood ratio function as follows:

$$\hat{\mathcal{L}}_c(\beta) = \sup \left\{ \prod_{i=1}^n (np_i) : p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \hat{\eta}_{i,c}(\beta) = 0 \right\}. \quad (46)$$

The optimal value of p_i satisfying (46) is given by $p_i = (1/n) \cdot (1/(1 + \lambda^\top \hat{\eta}_{i,c}(\beta)))$, where $\lambda = \lambda_c(\beta)$ is the solution to the equation $\sum_{i=1}^n (\hat{\eta}_{i,c}(\beta)/(1 + \lambda^\top \hat{\eta}_{i,c}(\beta))) = 0$. By the Lagrange multiplier method, the corresponding empirical log-likelihood ratio function is represented as

$$\hat{\mathcal{L}}_c(\beta) = 2 \sum_{i=1}^n \log \{ 1 + \lambda^\top \hat{\eta}_{i,c}(\beta) \}. \quad (47)$$

By maximizing $\{-\hat{\mathcal{L}}_c(\beta)\}$, we can obtain a maximum EL estimator $\hat{\beta}_{\text{cel}}$ of β with regression calibration method.

3.2. Imputation Empirical Likelihood. We introduce the following auxiliary random vector based on imputation method:

$$\hat{\eta}_{i,m}(\beta) = \frac{\varsigma_i \delta_i + (1 - \varsigma_i) m(T_i, \hat{\theta}_n)}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_m^{-2}(U_i) \left[\tilde{\xi}_i (\tilde{Y}_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta \right]. \quad (48)$$

Hence, we define the empirical log-likelihood ratio function as follows:

$$\hat{L}_m(\beta) = \sup \left\{ \prod_{i=1}^n (np_i) : p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \hat{\eta}_{i,m}(\beta) = 0 \right\}. \quad (49)$$

The optimal value of p_i satisfying (49) is given by $p_i = (1/n) \cdot (1/(1 + \lambda^\top \hat{\eta}_{i,m}(\beta)))$, where $\lambda = \lambda_m(\beta)$ is the solution to the equation $\sum_{i=1}^n (\hat{\eta}_{i,m}(\beta)/(1 + \lambda^\top \hat{\eta}_{i,m}(\beta))) = 0$. By the Lagrange multiplier method, the corresponding empirical log-likelihood ratio function is

$$\hat{l}_m(\beta) = 2 \sum_{i=1}^n \log \{1 + \lambda^\top \hat{\eta}_{i,m}(\beta)\}. \quad (50)$$

By maximizing $\{-\hat{l}_m(\beta)\}$, we can obtain a maximum EL estimator $\hat{\beta}_{\text{mel}}$ of β with imputation method.

3.3. Inverse Probability Weighted Empirical Likelihood. We introduce the following auxiliary random vector based on inverse probability weighted method:

$$\begin{aligned} \hat{\eta}_{i,w}(\beta) &= \frac{(\zeta_i \delta_i / \hat{\pi}_n(Y_i)) + (1 - (\zeta_i / \hat{\pi}_n(Y_i)))m(T_i, \hat{\theta}_n)}{1 - \hat{G}_n(Y_i)} \hat{\sigma}_w^{-2} \\ &\cdot (U_i) [\tilde{\xi}_i (\tilde{Y}_i - \tilde{\zeta}_i^\top \beta) + \Sigma_e \beta]. \end{aligned} \quad (51)$$

Then, we define the empirical log-likelihood ratio function as follows:

$$\hat{L}_w(\beta) = \sup \left\{ \prod_{i=1}^n (np_i) : p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \hat{\eta}_{i,w}(\beta) = 0 \right\}. \quad (52)$$

The optimal value of p_i satisfying (52) is given by $p_i = (1/n) \cdot (1/(1 + \lambda^\top \hat{\eta}_{i,w}(\beta)))$, where $\lambda = \lambda_w(\beta)$ is the solution to the equation $\sum_{i=1}^n (\hat{\eta}_{i,w}(\beta)/(1 + \lambda^\top \hat{\eta}_{i,w}(\beta))) = 0$. By the Lagrange multiplier method, the corresponding empirical log-likelihood ratio function is represented as

$$\hat{l}_w(\beta) = 2 \sum_{i=1}^n \log \{1 + \lambda^\top \hat{\eta}_{i,w}(\beta)\}. \quad (53)$$

By maximizing $\{-\hat{l}_w(\beta)\}$, we can obtain a maximum EL estimator $\hat{\beta}_{\text{wel}}$ of β with inverse probability weighted method.

4. Main Results

For convenience and simplicity, we use $C_0, C_1, \dots$ and $c_0, c_1, \dots$ generically to represent any positive constants, which may take different values for each appearance. Let $\Gamma(u) = E[ZZ^\top | U = u]$, $\Phi(u) = E[ZX^\top | U = u]$,

$\Psi = X - \Phi^{-1}(U)\Gamma^{-1}(U)Z$, $\Theta = [X + e - \Phi(U)\Gamma^{-1}(U)\Phi^\top(U)](\varepsilon - e\beta)$ and $M^{\otimes 2} = MM^\top$. Denote

$$\begin{aligned} \Sigma_{11} &= E \left\{ \frac{m(T, \theta)}{1 - G(Y)} \Psi \Psi^\top \right\}, \\ \Sigma_{12} &= E \left\{ \frac{m(T_i, \theta)}{1 - G(Y_i)} \Theta + \Sigma_e \beta \right\}^{\otimes 2}, \\ \Sigma_{21} &= E \left\{ \frac{\zeta \delta + (1 - \zeta)m(T, \theta)}{1 - G(Y)} \Psi \Psi^\top \right\}, \\ \Sigma_{22} &= E \left\{ \frac{\zeta \delta + (1 - \xi)m(T, \theta)}{1 - G(Y)} \Theta + \Sigma_e \beta \right\}^{\otimes 2}, \\ \Sigma_{31} &= E \left\{ \frac{\zeta \delta + (1 - \xi)m(T, \theta)}{1 - G(Y)} \Psi \Psi^\top \right\}, \\ \Sigma_{32} &= E \left\{ \frac{(\zeta/\pi(Y))\delta + (1 - (\zeta/\pi(Y)))m(T, \theta)}{1 - G(Y)} \Theta + \Sigma_e \beta \right\}^{\otimes 2}, \\ \Sigma_{r,11} &= E \left\{ \frac{m(T, \theta)}{1 - G(Y)} \sigma^{-2}(U) \Psi \Psi^\top \right\}, \\ \Sigma_{r,12} &= E \left\{ \frac{m(T, \theta)}{1 - G(Y)} \sigma^{-2}(U) \Theta + \Sigma_e \beta \right\}^{\otimes 2}, \\ \Sigma_{r,21} &= E \left\{ \frac{\zeta \delta + (1 - \zeta)m(T, \theta)}{1 - G(Y)} \sigma^{-2}(U) \Psi \Psi^\top \right\}, \\ \Sigma_{r,22} &= E \left\{ \frac{\zeta \delta + (1 - \zeta)m(T, \theta)}{1 - G(Y)} \sigma^{-2}(U) \Theta + \Sigma_e \beta \right\}^{\otimes 2}, \\ \Sigma_{r,31} &= E \left\{ \frac{\zeta \delta + (1 - \zeta)m(T, \theta)}{1 - G(Y)} \sigma^{-2}(U) \Psi \Psi^\top \right\}, \\ \Sigma_{r,32} &= E \left\{ \frac{(\zeta/\pi(Y))\delta + (1 - (\zeta/\pi(Y)))m(T, \theta)}{1 - G(Y)} \sigma^{-2}(U) \Theta + \Sigma_e \beta \right\}^{\otimes 2}, \\ \Lambda_1 &= v_0 E \left\{ \frac{m(T, \theta)}{1 - G(Y)} g^{-1}(U) \Gamma^{-1}(U) (\varepsilon - e\beta)^2 \right\}, \\ \Lambda_2 &= v_0 E \left\{ \frac{\zeta \delta + (1 - \zeta)m(T_i, \theta)}{1 - G(Y_i)} g^{-1}(U) \Gamma^{-1}(U) (\varepsilon - e\beta)^2 \right\}, \\ \Lambda_3 &= v_0 E \left\{ \frac{(\zeta/\pi(Y))\delta + (1 - (\zeta/\pi(Y)))m(T, \theta)}{1 - G(Y_i)} g^{-1} \right. \\ &\quad \left. \cdot (U) \Gamma^{-1}(U) (\varepsilon - e\beta)^2 \right\}. \end{aligned} \quad (54)$$

In order to prove the main results, we give a set of assumptions that are stated in the following theorems:

(C1) The random variable U has bounded support $\mathcal{U}$ and its density function $g(\cdot)$ is Lipschitz continuous and away from zero on its support.

(C2) There is $s > 2$ such that $E(\|X_1\|^{2s}|U_1) < \infty$ a.s., $E(\|Z_1\|^{2s}|U_1) < \infty$ a.s., $E(|\varepsilon_1|^{2s}|X_1, Z_1) < \infty$ a.s. $\Gamma(u)$ and $\Phi(u)$ are nonsingular matrixes.

(C3) $\{\alpha_j(\cdot), j = 1, \dots, q\}$ has continuous second derivatives in $U \in \mathcal{U}$.

(C4) The variance function $\sigma^2(\cdot)$ with uniform boundedness has continuous second-order derivation and is bounded away from zero.

(C5) The kernel $K(\cdot)$ as a symmetric density function has compact support $[-1, 1]$, which is Lipschitz continuous, and satisfies $\int_{\mathbb{R}} t^j K(t) dt = \kappa_j < \infty$, $\int_{\mathbb{R}} t^j K^2(t) dt = \nu_j < \infty$, $j = 0, 1, 2$.

(C6) Denote $a_H = \inf\{t: H(t) = 1\}$ and $a_G = \inf\{t: G(t) = 1\}$. Let $a_H < a_G$ and $G(\cdot)$ is continuous. $H(t)$ is continuous for $t > a_H$.

(C7) The kernel functions $\Omega(\cdot)$ and $L(\cdot)$ are bounded with bounded compact supports, and $na_n \rightarrow \infty$, $na_n^2 \rightarrow 0$, $nb_n \rightarrow \infty$ and $nb_n^2 \rightarrow 0$.

(C8) $\pi(\cdot)$ and $\mu(\cdot)$ have bounded derivatives of order 1, and there exists $c > 0$ such that $\inf_x \pi(x) \geq c > 0$.

(C9) $I(\theta_0)$ is a positive definite. $\nabla m(T_i, \theta_0)$ is continuous at θ_0 .

(C10) The bandwidth h_{in} satisfies $h_{in} \rightarrow 0$, $nh_{in} \rightarrow \infty$, $nh_{in}^8 \rightarrow 0$ and $nh_{in}^2 / (\log n)^2 \rightarrow \infty$ for $i = 1, 2, 3$.

Remark 1.

(a) Assumptions (C1) and (C2) are used to establish the asymptotic normality and the oracle property of the estimators. Assumptions (C3) and (C4) are common conditions for varying-coefficients models with heteroscedastic error. Assumption (C5) requires that the kernel function is a proper density with finite second moment, which is required to derive the asymptotic variance of estimators. Assumption (C6) implies that $1 - G(Y_i)$ is bounded away from zero. Assumptions (C7)–(C9) are needed for the properties of $\hat{G}_n(y)$ and $m(T_i, \hat{\theta}_n)$. Assumption (C10) underlines the relationship bandwidth with sample size n , which implies the optimal bandwidth in nonparametric estimation.

(b) From the Taylor expansion and conclusion in Li and Wang [17], one can get

$$\begin{aligned} \sqrt{n}(\hat{\theta}_n - \theta) &= I^{-1}(\theta_0) \frac{1}{\sqrt{n}} \\ &\cdot \sum_{i=1}^n \frac{\xi_i (\delta_i - m(T_i, \theta_0)) \nabla m(T_i, \theta_0)}{m(T_i, \theta_0) \{1 - m(T_i, \theta_0)\}} + o_p(1), \end{aligned} \quad (55)$$

which, together with assumption (C9), gives $\hat{\theta}_n - \theta = O_p(n^{-1/2})$.

The asymptotic properties of the proposed estimators are shown in the following theorems.

Theorem 1. Suppose that assumptions (C1)–(C10) are satisfied; then, we have

$$\sqrt{n}(\tilde{\beta}_n - \beta) \xrightarrow{D} N(0, \Sigma_{j1}^{-1} \Sigma_{j2} \Sigma_{j1}^{-1}), \quad (56)$$

where $\tilde{\beta}_n$ is taken to be $\tilde{\beta}_c(u)$, $\tilde{\beta}_m(u)$ and $\tilde{\beta}_w(u)$. $j = 1, 2, 3$ correspond to $\tilde{\beta}_c$, $\tilde{\beta}_m$ and $\tilde{\beta}_w$, respectively.

Theorem 2. Suppose that assumptions (C1)–(C10) are satisfied; then, we have

$$\sqrt{nh_{1n}} \left(\tilde{\alpha}_n(u) - \alpha(u) - \frac{1}{2} h_n^2 \kappa^2 \alpha''(u) \right) \xrightarrow{D} N(0, \Lambda_j), \quad (57)$$

where $\tilde{\alpha}_n(u)$ is taken to be $\tilde{\alpha}_c(u)$, $\tilde{\alpha}_m(u)$ and $\tilde{\alpha}_w(u)$. $j = 1, 2, 3$ correspond to $\tilde{\alpha}_c(u)$, $\tilde{\alpha}_m(u)$ and $\tilde{\alpha}_w(u)$, respectively.

Theorem 3. Suppose that assumptions (C1)–(C10) are satisfied; let $\rho_n = (\log n / nh_{2n})^{1/2} + h_{2n}^2$; then, we have

$$\sup_u |\tilde{\sigma}_n^2(u) - \sigma^2(u)| = O_p(\rho_n), \quad (58)$$

where $\tilde{\sigma}_n^2(u)$ is taken to be one of $\tilde{\sigma}_c^2(u)$, $\tilde{\sigma}_m^2(u)$ and $\tilde{\sigma}_w^2(u)$.

Theorem 4. Suppose that assumptions (C1)–(C10) are satisfied; then, we have

$$\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{D} N(0, \Sigma_{r,j1}^{-1} \Sigma_{r,j2} \Sigma_{r,j1}^{-1}), \quad (59)$$

where $\hat{\beta}_n$ is taken to be one of $\hat{\beta}_c(u)$, $\hat{\beta}_m(u)$ and $\hat{\beta}_w(u)$. $j = 1, 2, 3$ correspond to $\hat{\beta}_c(u)$, $\hat{\beta}_m(u)$ and $\hat{\beta}_w(u)$, respectively.

Theorem 5. Suppose that assumptions (C1)–(C10) are satisfied; then, we have

$$\sqrt{nh_{1n}} \left(\hat{\alpha}_n(u) - \alpha(u) - \frac{1}{2} h_n^2 \kappa^2 \alpha''(u) \right) \xrightarrow{D} N(0, \Lambda_j), \quad (60)$$

where $\hat{\alpha}_n(u)$ is taken to be one of $\hat{\alpha}_c(u)$, $\hat{\alpha}_m(u)$ and $\hat{\alpha}_w(u)$. $j = 1, 2, 3$ correspond to $\hat{\alpha}_c(u)$, $\hat{\alpha}_m(u)$ and $\hat{\alpha}_w(u)$, respectively.

Theorem 6. Suppose that assumptions (C1)–(C10) are satisfied; if β is the true value, then we have

$$\hat{l}_n(\beta) \xrightarrow{D} \chi_1^2, \quad (61)$$

where $\hat{l}_n(\beta)$ denotes one of $\hat{l}_c(\beta)$, $\hat{l}_m(\beta)$, and $\hat{l}_w(\beta)$. χ_1^2 is a standard chi-squared random variable with 1 degree of freedom.

Theorem 7. Suppose that assumptions (C1)–(C10) are satisfied; then, we have

$$\sqrt{n}(\hat{\beta}_{el} - \beta) \xrightarrow{D} N(0, \Sigma_{r,j1}^{-1} \Sigma_{r,j2} \Sigma_{r,j1}^{-1}), \quad (62)$$

where $\hat{\beta}_{el}$ denotes one of $\hat{\beta}_{cel}$, $\hat{\beta}_{mel}$ and $\hat{\beta}_{wel}$. $j = 1, 2, 3$ correspond to $\hat{\beta}_{cel}$, $\hat{\beta}_{mel}$ and $\hat{\beta}_{wel}$, respectively.

TABLE 1: Choosing for v and $a = (a_1, a_2, a_3)$ and corresponding CR and MR in model (63).

n	CR (%)	MR (%)	v	$a = (a_1, a_2, a_3)$	$\theta = (\theta_1, \theta_2, \theta_3)$
100	10	10	9.25	$(-6.0, -1.1, -1.1)$	$(-2.546, -0.595, -0.215)$
		40	9.25	$(-1.3, -0.9, -0.8)$	$(-2.551, -0.351, -0.311)$
	40	10	0.79	$(-6.1, -1.3, -1.0)$	$(-0.152, -0.222, -1.650)$
		40	0.79	$(-1.5, -0.9, -0.7)$	$(-0.144, -0.213, -1.550)$
400	10	10	9.25	$(-6.0, -1.1, -1.1)$	$(-2.567, -0.601, -0.221)$
		40	9.25	$(-1.3, -0.9, -0.8)$	$(-2.541, -0.355, -0.302)$
	40	10	0.79	$(-6.1, -1.3, -1.0)$	$(-0.151, -0.213, -1.620)$
		40	0.79	$(-1.5, -0.9, -0.7)$	$(-0.145, -0.214, -1.550)$

FIGURE 1: The QQ-plots of $\tilde{\alpha}_w(\cdot)$ based on Σ_{e1} (a) and Σ_{e2} (b) with heteroscedasticity coefficient $\gamma = 1$ (c) and $\gamma = 4$ (d) under CR = 10% and MR = 10%.

Remark 2.

- (a) From Theorems 1 and 4, the asymptotic variance of the reweighted estimator $\hat{\beta}_n$ is not greater than that of the modified profile LS estimator $\hat{\beta}_n$; that is, $\Sigma_{j1}^{-1} \Sigma_{j2} \Sigma_{j1}^{-1} - \Sigma_{r,j1}^{-1} \Sigma_{r,j2} \Sigma_{r,j1}^{-1}$ is a positive semidefinite matrix. The asymptotic variance of the reweighted estimator $\hat{\beta}_m$ is smaller than that of $\hat{\beta}_w$, and is larger than that of $\hat{\beta}_c$,

which indicates that $\hat{\beta}_c$ performs the best, and $\hat{\beta}_w$ performs the worst. The modified PLS estimators $\hat{\beta}_c, \hat{\beta}_m, \tilde{\beta}_w$ enjoy the same conclusion.

- (b) From Theorems 2 and 5, the local polynomial estimator $\tilde{\alpha}_n(\cdot)$ and reweighted estimator $\tilde{\alpha}_n(\cdot)$ have the same asymptotic distribution, which reflects the characteristic of the local regression in nonparametric models.

FIGURE 2: The QQ-plots of $\tilde{\alpha}_w(\cdot)$ based on Σ_{e1} (a) and Σ_{e2} (b) with heteroscedasticity coefficient $\gamma = 1$ (c) and $\gamma = 4$ (d) under CR = 40% and MR = 40%.

- (c) From Theorem 6, the $100(1 - \tau)\%$ EL confidence region for β can be established as $I_\tau = \{\beta: \hat{l}_n(\beta) \leq \chi_1^2(1 - \tau)\}$, where $\chi_1^2(1 - \tau)$ is the upper $(1 - \tau)$ -quantile of distribution of χ_1^2 .

5. Simulation

In this subsection, we carry out some numerical simulation to investigate the finite sample behavior of the proposed estimators. We compare the performance of the estimators based on the regression calibration method (CA), imputation method (IM) and inverse probability weighted method (IPW), and their corresponding reweighted estimators (R-CA, R-IM, R-IPW). Besides, we conduct a comparison of the EL method with the normal approximation (NA) approach in terms of coverage probabilities (CP) and average interval lengths (AL) under different settings. At the same time, we give a real data analysis. The kernel functions are taken

as $K(u) = (3/4)(1 - u^2)I(|u| \leq 1)$, $L(u) = (15/16)(1 - u^2)^2I(|u| \leq 1)$ and $\Omega(u) = (1/2)I(|u| \leq 1)$. The bandwidths h_{1n} , h_{2n} and h_{3n} have taken the same values by leave-one-sample-out cross-validation. The following simulation is based on 500 replications. The sample size n is chosen to be 100 and 400, repeatedly.

5.1. Simulation Experiments. The data are generated from the following the PLVCEV model:

$$\begin{cases} T_i = X_{i1}^\top \beta_1 + X_{i2}^\top \beta_2 + Z_i^\top \alpha(U_i) + \varepsilon_i, \\ \xi_{i1} = X_{i1} + e_{i1}, \xi_{i2} = X_{i2} + e_{i2}, \quad i = 1, \dots, n, \end{cases} \quad (63)$$

where $\beta_1 = 1$, $\beta_2 = 2$, $\alpha(u) = \sin(2\pi u)$, the covariates X_{i1} and X_{i2} are from $N(0, 1)$ and pairwise covariance $\text{cov}(X_{ij}, X_{ik}) = 0.5^{|j-k|}$. Z_i is from $U(-1, 1)$. U_i is from $U(0, 1)$, the model error $\varepsilon_i \sim N(0, 1)$. The error variance function is taken as $\sigma^2(u) = 0.25 + [\gamma(1 + \sin(2\pi u))]^2$ for

FIGURE 3: The curves of $\sigma^2(\cdot)$ (solid line), $\hat{\sigma}_c^2(\cdot)$ (dotted line), $\hat{\sigma}_m^2(\cdot)$ (dashed line), and $\hat{\sigma}_w^2(\cdot)$ (dash-dotted line) based on Σ_{e1} (a) and Σ_{e2} (b) with heteroscedasticity coefficient $\gamma = 1$ (c) and $\gamma = 4$ (d) under CR=10% and MR=10%.

$\gamma = 1$ and 4. The measurement error $e_i \sim N(0, \Sigma_e)$. To represent different levels of measurement errors, we take Σ_e be $\Sigma_{e1} = \text{diag}(0.5, 0.5)$ and $\Sigma_{e2} = \text{diag}(0.25, 0.25)$ in the simulations, respectively. Let the censoring time C_i be from $\exp(v)$, where v is adapted to get different censoring ratios (CR). For the missing mechanism, we take $P(c_i | X_i, Z_i, U_i, Y_i) = 1/[1 + \exp(-a_1 - a_2(X_{i1} + Z_i + U_i) - a_3 Y_i)]$. We choose different v and $a = (a_1, a_2, a_3)$ to get different CR and missing ratios (MR). Suppose that $m(T_i, \theta)$ follows a logistic model, that is, $m(T_i, \theta) = 1/[1 + \exp(-\theta_1 - \theta_2 X_{i2} - \theta_3 Y_i)]$. Then, the parameter $\theta = (\theta_1, \theta_2, \theta_3)$ is estimated by maximum likelihood method. See Table 1 for details.

In the first simulation, we study the finite sample performance of the proposed modified PLS estimators and reweighted estimators of β based on mean squared error (MSE) defined as

$$\text{MSE}(\hat{\beta}_n) = \sqrt{(\hat{\beta}_n - \beta)^\top (\hat{\beta}_n - \beta)}, \quad (64)$$

and the global mean square error (GMSE) of $\hat{\alpha}_n(u)$ defined as

$$\text{GMSE} = \frac{1}{n_{\text{grid}} M} \sum_{m=1}^M \sum_{j=1}^{n_{\text{grid}}} \{\hat{\alpha}_n(u_j) - \alpha(u_j)\}^2, \quad (65)$$

where $\{u_j: j = 1, 2, \dots, n_{\text{grid}}\}$ is a sequence of grid points. In addition, we plot QQ-plots of the reweighted estimator $\hat{\alpha}_w(u)$ for $\alpha(u)$ under different settings in Figures 1 and 2. In the second simulations, we plot the curves of the proposed estimators $\hat{\sigma}_c^2(u)$, $\hat{\sigma}_m^2(u)$, and $\hat{\sigma}_w^2(u)$ under different settings in Figures 3 and 4. In the third simulations, we consider CP and AL of the confidence regions for β based on the EL method (CPE, ALE) and NA method (CPN, ALN) with nominal level 0.95 under different settings in Table 2.

From Tables 2–4 and Figures 1–4, it can be seen that

- (1) In Tables 3 and 4, the MSE and GMSE of reweighted estimators are smaller than those of modified PLS estimators under the same setting. The results of IM estimators are smaller than those of IPW estimators, and bigger than those of RC estimators. The results increase as measurement error, heteroscedasticity error, and CR and/or MR increase. The results decrease as the sample size increases. The results above

FIGURE 4: The curves of $\sigma^2(\cdot)$ (solid line), $\hat{\sigma}_c^2(\cdot)$ (dotted line), $\hat{\sigma}_m^2(\cdot)$ (dashed line), and $\hat{\sigma}_w^2(\cdot)$ (dash-dotted line) based on Σ_{e1} (a) and Σ_{e2} (b) with heteroscedasticity coefficient $\gamma = 1$ (c) and $\gamma = 4$ (d) under CR=40% and MR=40%.

imply that the reweighted estimators perform better than the modified PLS estimators. The RC method performs best, and IPW method performs worst, which confirms the theoretical results.

- (2) In Table 2, the CP of reweighted estimators is larger than that of the modified PLS estimators. The CP of the RC method is the smallest, and that of the IPW method is the biggest under the same settings. The CP decreases as heteroscedasticity error and CR and/or MR increase. The results increase as sample size increases. The CP based on the EL method is smaller than that of the NA method. The AL performs in the opposite way.
- (3) In Figures 1 and 2, the fit is better as decreasing the heteroscedasticity error and measurement error. The fit is worse as increasing CR and/or MR.
- (4) In Figures 3 and 4, the proposed estimators of error variance perform better as decreasing measurement error, heteroscedasticity error, and CR and/or MR. The estimator $\hat{\sigma}_c^2(\cdot)$ performs the best, and $\hat{\sigma}_w^2(\cdot)$ performs the worst under the same settings.

5.2. A Real Data Analysis. In real data analysis, we illustrate the methodology via an application to a dataset from a breast cancer clinical trial [25]. This clinical trial was conducted by the

Eastern Cooperative Oncology Group, whose target was evaluating tamoxifen as a treatment for stage II breast cancer among elderly women, who are older than 65. There are 169 elderly women participating in the trial, and we focus on 79 women who died by the end of the trial. But, unfortunately, the cause of death is incomplete. Among them, 44 women died from breast cancer, 17 died from other known causes, and 18 died from unknown causes. Let the censoring indicator δ show whether the death was caused by breast cancer, and let the missing indicator ζ show whether the cause of death was known. The dataset contains four covariates: whether the patients accepted the treatment (1, tamoxifen; 0, placebo), denoted as X_1 ; whether the estrogen receptor status was positive (1, yes; 0, unknown), denoted as X_2 ; whether there were four or more axillary lymph positive nodes (1, yes; 0, no), denoted as Z ; and whether the primary tumor is 3 cm or larger (1, yes; 0, no), denoted as U . Then, we employ the following model to fit the data,

$$Y = X_1^\top \beta_1 + X_2^\top \beta_2 + Z^\top \alpha(U) + \varepsilon, \quad (66)$$

where Y is the logarithm of the time to death due to breast cancer, which is censored, and the censoring indicator is MAR. The heteroscedastic error follows the form of $\text{Var}(\varepsilon|X, Z, U) = \sigma^2(U_i)$. For the purpose of comparison, we compute both mean squared prediction error (MSE) and mean

TABLE 2: CPE and CPN of β with confidence 95% and their corresponding ALE and ALN with Σ_{e1} .

n	CR (%)	MR (%)	Method	$\gamma = 1$				$\gamma = 4$			
				CPE	CPN	ALE	ALN	CPE	CPN	ALE	ALN
100	10	10	IPW	0.8310	0.5980	1.3577	0.5314	0.7270	0.5880	4.8436	1.9884
			IM	0.8900	0.6890	1.1943	0.5288	0.8660	0.6450	2.4760	1.5248
			RC	0.8930	0.6970	1.1923	0.5140	0.8690	0.6610	2.4609	1.5125
			R-IPW	0.8810	0.6430	0.8465	0.4914	0.8250	0.6120	1.0170	0.9900
			R-IM	0.9270	0.7240	0.7931	0.5090	0.8780	0.6600	0.9548	0.9740
			R-RC	0.9270	0.7340	0.7922	0.5067	0.8780	0.6650	0.9494	0.9235
	10	40	IPW	0.8410	0.5970	1.3613	0.6403	0.7580	0.5930	4.6717	1.9853
			IM	0.8900	0.6790	1.2079	0.5190	0.8580	0.6460	2.5634	1.5672
			RC	0.8920	0.6880	1.2026	0.5080	0.8660	0.6540	2.5301	1.5410
			R-IPW	0.8780	0.6280	0.8720	0.5063	0.8180	0.5990	1.0302	1.0481
			R-IM	0.9120	0.7020	0.8015	0.5007	0.8720	0.6600	0.9638	0.9703
			R-RC	0.9220	0.7150	0.7939	0.5006	0.8750	0.6680	0.9619	0.9256
	40	10	IPW	0.6700	0.4830	1.9149	0.6595	0.7060	0.5010	5.5129	1.7238
			IM	0.8870	0.6620	1.3387	0.6135	0.8300	0.5960	2.4391	1.4932
			RC	0.8960	0.6880	1.2997	0.5689	0.8360	0.6080	2.2743	1.4751
			R-IPW	0.7990	0.4810	1.1088	0.5471	0.7400	0.5470	1.5044	1.3151
			R-IM	0.9130	0.6830	0.8672	0.5353	0.8480	0.6100	1.1333	1.1330
			R-RC	0.9240	0.7080	0.8652	0.5005	0.8680	0.6330	1.0950	0.9522
	40	40	IPW	0.6900	0.4600	1.9381	0.6855	0.7070	0.5160	5.4336	1.6817
			IM	0.8620	0.6170	1.3993	0.6386	0.8330	0.5800	2.6316	1.5084
			RC	0.8830	0.6730	1.3304	0.5554	0.8380	0.6030	2.3778	1.4733
			R-IPW	0.7730	0.4690	1.1409	0.5522	0.7390	0.5480	1.4717	1.1889
			R-IM	0.8750	0.6400	0.8883	0.5355	0.8350	0.6130	1.1907	1.0976
			R-RC	0.8940	0.7000	0.8799	0.5050	0.8520	0.6320	1.1444	0.9509
400	10	10	IPW	0.7330	0.4170	0.7414	0.2631	0.6710	0.4580	4.9652	1.6687
			IM	0.9310	0.6520	0.6370	0.2560	0.8720	0.5440	1.9004	0.9526
			RC	0.9340	0.6640	0.6356	0.2541	0.8750	0.5590	1.8559	0.9386
			R-IPW	0.8320	0.4670	0.5251	0.2610	0.8190	0.5800	0.5444	0.4033
			R-IM	0.9390	0.6660	0.4485	0.2457	0.9330	0.6070	0.5092	0.3914
			R-RC	0.9400	0.6860	0.4413	0.2456	0.9360	0.6280	0.5025	0.3845
	10	40	IPW	0.7220	0.4040	0.7502	0.2656	0.6560	0.4490	4.8817	1.6279
			IM	0.9230	0.5920	0.6470	0.2588	0.8390	0.5430	2.1294	1.0194
			RC	0.9290	0.6530	0.6455	0.2468	0.8470	0.5850	2.0601	0.9871
			R-IPW	0.7970	0.4690	0.5518	0.3613	0.8040	0.5850	0.5493	0.3999
			R-IM	0.9410	0.6220	0.4450	0.2590	0.9240	0.6280	0.5108	0.3921
			R-RC	0.9560	0.6650	0.4415	0.2540	0.9270	0.6300	0.5102	0.3858
	40	10	IPW	0.3650	0.1730	1.0703	0.2953	0.5960	0.3530	6.6031	1.7275
			IM	0.8970	0.5410	0.7440	0.2768	0.8050	0.4640	2.0105	0.9612
			RC	0.9240	0.6210	0.7250	0.2759	0.8330	0.4930	1.8325	0.9396
			R-IPW	0.5270	0.3020	0.7678	0.2946	0.6120	0.5270	0.7763	0.4621
			R-IM	0.9180	0.5870	0.4991	0.2762	0.9170	0.5640	0.6400	0.4559
			R-RC	0.9450	0.6360	0.4895	0.2743	0.9240	0.5820	0.6284	0.4370
	40	40	IPW	0.3640	0.1640	1.1850	0.3123	0.5750	0.3610	7.0990	1.7615
			IM	0.8400	0.4730	0.8492	0.3005	0.7650	0.4310	2.4979	1.0689
			RC	0.9030	0.6070	0.8058	0.2996	0.7810	0.4820	2.1900	1.0364
			R-IPW	0.5410	0.2840	0.7799	0.2920	0.6010	0.5350	0.8070	0.4799
			R-IM	0.8490	0.5120	0.5752	0.2918	0.8810	0.5410	0.6692	0.4606
			R-RC	0.9140	0.6180	0.5599	0.2840	0.9060	0.5480	0.6614	0.4472

TABLE 3: The MSE of the estimators for β_1 and β_2 and GMSE of the estimators for $\alpha(u)$ with Σ_{e1} .

n	CR (%)	MR (%)	Method	$\gamma = 1$		$\gamma = 4$		$\gamma = 1$	$\gamma = 4$
				β_1	β_2	β_1	β_2	GMSE	GMSE
100	10	10	IPW	0.0696	0.0593	0.2406	0.2580	0.1797	0.3782
			IM	0.0488	0.0480	0.1237	0.1111	0.1670	0.2327
			RC	0.0432	0.0444	0.1153	0.1109	0.1665	0.2314
			R-IPW	0.0383	0.0407	0.3621	0.1619	0.1771	0.3097
			R-IM	0.0381	0.0372	0.0915	0.0854	0.1665	0.2299
			R-RC	0.0374	0.0367	0.0896	0.0838	0.1663	0.2293
	10	40	IPW	0.0617	0.0610	0.2864	0.2578	0.1802	0.3654
			IM	0.0475	0.0455	0.1814	0.1608	0.1695	0.2738
			RC	0.0424	0.0430	0.1742	0.1540	0.1646	0.2601
			R-IPW	0.0483	0.0399	0.1030	0.0961	0.1771	0.3338
			R-IM	0.0369	0.0365	0.0888	0.0900	0.1686	0.2654
			R-RC	0.0362	0.0358	0.0673	0.0791	0.1644	0.2531
	40	10	IPW	0.2205	0.2195	0.3168	0.3354	0.2859	0.4654
			IM	0.1918	0.1744	0.2467	0.2675	0.2553	0.4144
			RC	0.1286	0.1299	0.2456	0.2466	0.2505	0.3727
			R-IPW	0.1732	0.1670	0.2911	0.2214	0.2809	0.4626
			R-IM	0.1164	0.1138	0.2197	0.2236	0.2490	0.4124
			R-RC	0.1108	0.1098	0.2097	0.2107	0.2361	0.3651
	40	40	IPW	0.2237	0.2144	0.3245	0.3149	0.2926	0.4709
			IM	0.1759	0.2130	0.2893	0.2910	0.2708	0.3585
			RC	0.1740	0.1789	0.2611	0.2893	0.2659	0.3460
			R-IPW	0.1797	0.1696	0.2312	0.2418	0.2838	0.4656
			R-IM	0.1557	0.1574	0.2227	0.2281	0.2640	0.3473
			R-RC	0.1110	0.1279	0.2080	0.2087	0.2502	0.3440
400	10	10	IPW	0.0246	0.0270	0.2749	0.2756	0.1288	0.3326
			IM	0.0112	0.0125	0.0482	0.0526	0.1141	0.1503
			RC	0.0100	0.0099	0.0462	0.0513	0.1123	0.1417
			R-IPW	0.0139	0.0137	0.0293	0.0316	0.1265	0.2653
			R-IM	0.0092	0.0095	0.0288	0.0315	0.1136	0.1475
			R-RC	0.0090	0.0089	0.0282	0.0283	0.1121	0.1379
	10	40	IPW	0.0359	0.0359	0.3323	0.3580	0.1387	0.3695
			IM	0.0195	0.0186	0.1435	0.1406	0.1219	0.1909
			RC	0.0130	0.0185	0.1177	0.1030	0.1166	0.1777
			R-IPW	0.0128	0.0131	0.0401	0.0417	0.1338	0.2959
			R-IM	0.0110	0.0107	0.0365	0.0414	0.1206	0.1826
			R-RC	0.0101	0.0104	0.0310	0.0300	0.1159	0.1690
	40	10	IPW	0.1715	0.1684	0.2430	0.2599	0.2604	0.3664
			IM	0.0728	0.0708	0.1437	0.1432	0.1697	0.2188
			RC	0.0505	0.0466	0.1168	0.1116	0.1487	0.2157
			R-IPW	0.1336	0.1325	0.1596	0.1543	0.2562	0.3543
			R-IM	0.0702	0.0657	0.1292	0.1279	0.1690	0.2174
			R-RC	0.0468	0.0417	0.1065	0.1012	0.1482	0.2158
	40	40	IPW	0.1719	0.1639	0.2680	0.2706	0.2672	0.3895
			IM	0.1259	0.1193	0.2220	0.2037	0.2180	0.2578
			RC	0.0602	0.0692	0.1398	0.1340	0.1527	0.2230
			R-IPW	0.1360	0.1339	0.1820	0.1729	0.2648	0.3752
			R-IM	0.1208	0.1169	0.1536	0.1521	0.2161	0.2515
			R-RC	0.0492	0.0506	0.1161	0.1131	0.1523	0.2219

TABLE 4: The MSE of the estimators for β_1 and β_2 and GMSE of the estimators for $\alpha(u)$ with Σ_{e2} .

n	CR (%)	MR (%)	Method	$\gamma = 1$		$\gamma = 4$		$\gamma = 1$	$\gamma = 4$
				β_1	β_2	β_1	β_2	GMSE	GMSE
100	10	10	IPW	0.0650	0.0536	0.3470	0.2934	0.1798	0.3396
			IM	0.0534	0.0491	0.1332	0.1354	0.1725	0.2407
			RC	0.0529	0.0480	0.1301	0.1330	0.1723	0.2383
			R-IPW	0.0506	0.0493	0.1064	0.1116	0.1793	0.3138
			R-IM	0.0493	0.0460	0.1038	0.1082	0.1724	0.2404
			R-RC	0.0484	0.0452	0.0832	0.0926	0.1723	0.2379
	10	40	IPW	0.0633	0.0728	0.3404	0.3511	0.2519	0.3550
			IM	0.0527	0.0623	0.1804	0.1941	0.1715	0.2726
			RC	0.0511	0.0613	0.1642	0.1609	0.1699	0.2599
			R-IPW	0.9262	0.7587	0.1046	0.1067	0.1811	0.3335
			R-IM	0.0445	0.0459	0.0964	0.0986	0.1713	0.2670
			R-RC	0.0418	0.0434	0.0776	0.0844	0.1696	0.2572
	40	10	IPW	0.2238	0.2235	0.3408	0.3284	0.2943	0.3486
			IM	0.2232	0.2211	0.2761	0.2756	0.2752	0.3010
			RC	0.2058	0.2008	0.2639	0.2613	0.2591	0.2633
			R-IPW	0.1727	0.1876	0.3045	0.3014	0.2868	0.3441
			R-IM	0.1384	0.1359	0.2683	0.2630	0.2576	0.2743
			R-RC	0.1356	0.1411	0.2581	0.2525	0.2465	0.2613
	40	40	IPW	0.2371	0.2341	0.3404	0.3509	0.2885	0.3601
			IM	0.2282	0.2266	0.3097	0.3084	0.2735	0.3565
			RC	0.2132	0.2117	0.2959	0.2931	0.2734	0.3037
			R-IPW	0.1860	0.1871	0.3082	0.3289	0.2709	0.3564
			R-IM	0.1589	0.1592	0.2790	0.2770	0.2690	0.3548
			R-RC	0.1573	0.1524	0.2642	0.2609	0.2565	0.2968
400	10	10	IPW	0.0225	0.0218	0.3016	0.2508	0.1286	0.3290
			IM	0.0122	0.0133	0.0538	0.0597	0.1130	0.1439
			RC	0.0119	0.0126	0.0414	0.0521	0.1123	0.1364
			R-IPW	0.0134	0.0136	0.0293	0.0415	0.1261	0.2696
			R-IM	0.0101	0.0101	0.0317	0.0407	0.1128	0.1409
			R-RC	0.0101	0.0099	0.0304	0.0349	0.1119	0.1352
	10	40	IPW	0.0351	0.0302	0.3338	0.3939	0.1319	0.3946
			IM	0.0269	0.0206	0.1780	0.1377	0.1210	0.2414
			RC	0.0277	0.0215	0.1319	0.1242	0.1169	0.2034
			R-IPW	0.0152	0.0140	0.0496	0.0465	0.1298	0.3259
			R-IM	0.0149	0.0131	0.0480	0.0452	0.1193	0.2226
			R-RC	0.0143	0.0126	0.0318	0.0322	0.1151	0.1861
	40	10	IPW	0.1742	0.1734	0.3018	0.2901	0.2627	0.3830
			IM	0.0755	0.0729	0.1688	0.1726	0.1723	0.2229
			RC	0.0565	0.0524	0.1375	0.1394	0.1508	0.2186
			R-IPW	0.1293	0.1307	0.1607	0.1607	0.2581	0.3701
			R-IM	0.0710	0.0685	0.1356	0.1405	0.1720	0.2207
			R-RC	0.0521	0.0478	0.1114	0.1168	0.1504	0.2183
	40	40	IPW	0.1787	0.1743	0.2817	0.2786	0.2681	0.4099
			IM	0.1285	0.1268	0.2170	0.2305	0.2127	0.2755
			RC	0.0863	0.0841	0.1528	0.1715	0.1546	0.2333
			R-IPW	0.1282	0.1261	0.1751	0.1750	0.2647	0.2353
			R-IM	0.1203	0.1149	0.1494	0.1492	0.2126	0.2667
			R-RC	0.0580	0.0566	0.1195	0.1270	0.1542	0.3959

TABLE 5: The estimators, MSEs, and MADs for the breast cancer dataset.

Method	Estimates		MSE	MAD
	β_1	β_2		
RC	0.1015	6.7787	1.0572	0.6311
IM	0.5063	5.3196	1.1787	1.9918
IPW	0.5329	5.0336	1.7934	2.2677
R-RC	0.1087	7.7605	1.3351	0.6693
R-IM	0.5994	6.1303	1.1869	2.5233
R-IPW	0.6465	5.8090	2.2385	2.8778

FIGURE 5: The curves of $\hat{\alpha}_n(\cdot)$ for the breast cancer dataset based on RC, IM, and IPW methods.

absolute deviation (MAD) of the predictions, which are defined as follows: $MSE = (1/n) \sum_{i=1}^{169} (Y_i - \hat{Y}_i)^2$, and $MAD = (1/n) \sum_{i=1}^{169} |Y_i - \hat{Y}_i|$, where $\hat{Y}_i$ is the fitted value of Y_i . The values of MSE and MAD based on different methods are given in Table 5. In addition, the estimated curves of $\alpha(u)$ based on RC, IM, and IPW methods are reported in Figure 5.

From Table 5 and Figure 5, it can be seen that, (1) in Table 5, the estimators of β_1 are positive, which indicates that the breast-cancer deaths may live longer if they received the treatment. Among these estimators, the MSE and MAD based on the R-RC method are smallest, which confirms the conclusions in Theorems 1 and 4. (2) In Figure 5, the primary tumor size of patients is mainly from 0 to 3 cm. The survival time decreases obviously as the tumor size increases.

6. Conclusion

In this paper, we consider the estimation and confidence regions based on modified PLS method and EL inference for

PLVCEV model with heteroscedastic errors under censoring indicators MAR, respectively. Asymptotic properties of the proposed estimators are established, and the confidence regions of parameter are constructed. In addition, a simulation study and real data analysis are conducted to illustrate our proposed method.

Xu and Duan [11] established efficient estimation for varying-coefficient heteroscedastic partially linear model with additive errors, but their results are confined in responses observed completely. It is an innovative and challenging topic to study the statistical inference for heteroscedastic PLVCEV model under right-censored data with censoring indicators MAR.

An interesting problem is whether we can extend the estimation method and incomplete data to functional regression models.

Appendix

Proof of Main Results

Lemma A.1. Let $D_1, D_2, \dots, D_n$ be independent and identically distributed (i.i.d.) random variables. If $E|D_1|^s$ is bounded for $s > 1$, then $\max_{1 \leq i \leq n} |D_i| = o(n^{1/s})$ a.s.

Proof. Lemma A.1 can be verified as Lemma 3 in Owen [20]. $\square$

Lemma A.2. Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. random vectors, where Y_i and X_i are scalar random variables. Assume further that $E|Y|^s < \infty$ and $\sup_u \int |y|^s f(u, y) dy < \infty$, where $f(\cdot, \cdot)$ denotes the joint density of (X, Y) . Let $K(\cdot)$ be a bounded positive function with a bounded support and satisfying a Lipschitz condition. Then, $\sup_x |(\sum_{i=1}^n K_h(X_i - x) [Y_i - E(Y_i)]) / \sum_{i=1}^n K_h(X_i - x)| = O_p[(\ln(1/h)/nh)^{1/2}]$ provided that $n^{2\varepsilon-1} \rightarrow \infty$ for some $\varepsilon < 1 - s^{-1}$.

Proof. Lemma A.2 comes from the basic corollary in Mack and Silverman [26]. $\square$

Lemma A.3. Suppose that assumptions (C1)–(C10) hold; then, as $n \rightarrow \infty$, it holds that

$$\sup_{u \in \mathcal{U}} \left| \frac{1}{nh_{1n}} \sum_{i=1}^n K\left(\frac{U_i - u}{h_{1n}}\right) \left(\frac{U_i - u}{h_{1n}}\right)^\top Z_{ij_1} Z_{ij_2}^\top \right| = g(u) \Gamma_{j_1 j_2}(u) \kappa_l + O\left\{h_{1n}^2 + \frac{\log n}{nh_{1n}}\right\} \text{ a.s.}, \quad (\text{A.1})$$

$$\sup_{u \in \mathcal{U}} \left| \frac{1}{nh_{1n}} \sum_{i=1}^n K\left(\frac{U_i - u}{h_{1n}}\right) \left(\frac{U_i - u}{h_{1n}}\right)^\top Z_{ij} \varepsilon_i \right| = \left(\frac{\log n}{nh_{1n}}\right)^{1/2} \text{ a.s.},$$

where $j, j_1, j_2 = 1, \dots, q$, $l = 0, 1, 2, 4$, $\Gamma_{j_1 j_2}(u)$ is the (j_1, j_2) -th element of $\Gamma(u)$ and $\gamma_n = \{\log n / (nh_{1n})\}^{1/2} + h_{1n}^2$.

Proof. The proof of Lemma A.3 is similar to that of Lemma A.2 in Xia and Li [27]. $\square$

Lemma A.4. Under the assumptions (C6)–(C8), then we have

$$\begin{aligned} \widehat{G}_n(Y_i) - G(Y_i) &= (1 - G(Y_i)) \frac{1}{n} \sum_{j=1}^n \zeta(Y_j, \delta_j, \varsigma_j; Y_i) \\ &\quad + R_n(Y_i), \end{aligned} \quad (\text{A.2})$$

where

$$\begin{aligned} \zeta(Y_j, \delta_j, \varsigma_j; Y_i) &= -\frac{(\varsigma_j - \pi(Y_j))(\delta_j - \mu(Y_j))}{\pi(Y_j)(1 - H(Y_j))} I(Y_j \leq Y_i) + \int_0^{Y_j \wedge Y_i} \frac{d\tilde{H}_1(s)}{(1 - H(s))^2} + \frac{I(Y_j Y_i, \delta_j = 0)}{1 - H(Y_j)}, \\ R_n(Y_i) &= -\frac{e^{-\Xi^*(Y_i)}}{2} (\Xi_n(Y_i) - \Xi(Y_i))^2 + e^{-\Xi^{**}(Y_i)} (-\ln(1 - \widehat{G}_n(Y_i)) - \Xi_n(Y_i)) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{[1 - \widehat{\mu}_n(Y_j)] [H_n(Y_j) - H(Y_j)]}{[1 - H_n(Y_j)] [1 - H_n(Y_j)]} I(Y_j < Y_i) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{[\widehat{\mu}_n(Y_j) - \mu(Y_j)] [H_n(Y_j) - H(Y_j)]}{[1 - H_n(Y_j)] [1 - H(Y_j)]} I(Y_j < Y_i) \\ &\quad + \int_0^{Y_i} \frac{\mu(x) [H_n(x) - H(x)]}{[1 - H(x)]^2} d[H_n(x) - H(x)] + \int_0^{Y_i} \frac{\mu(x) [H_n(x) - H(x)]^2}{[1 - H_n(x)] [1 - H(x)]^2} dH_n(x) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{I(Y_j < Y_i)}{1 - H(Y_j)} \frac{1}{nb_n} \sum_{j=1}^n \frac{\varsigma_j (\mu(Y_i) - \mu(Y_j))}{V^*(Y_j)} \Omega\left(\frac{Y_j - X_j}{b_n}\right) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{I(Y_j < Y_i)}{1 - H(Y_j)} \frac{1}{nb_n} \sum_{j=1}^n \frac{\varsigma_j (\delta_j - \mu(Y_j)) (V_n^*(Y_j) - V(X_i))}{V_n^*(Y_j) V^*(Y_j)} \Omega\left(\frac{Y_j - X_j}{b_n}\right) \\ &\quad + \frac{1}{n} \sum_{l=1}^n \xi_l (\delta_l - \mu(X_l)) \left\{ \frac{1}{nb_n} \sum_{i=1}^n \frac{I(Y_j < t)}{(1 - H(Y_j)) V^*(Y_j)} \Omega\left(\frac{Y_j - Y_l}{b_n}\right) \right. \\ &\quad \left. - E_{X_j} \left[\frac{1}{nb_n} \sum_{i=1}^n \frac{I(Y_j < Y_i)}{(1 - H(Y_j)) V^*(Y_j)} \Omega\left(\frac{Y_j - Y_l}{b_n}\right) \right] \right\} \\ &\quad + \frac{1}{n} \sum_{l=1}^n \varsigma_l (\delta_l - \mu(Y_l)) I(Y_l < t) \left[\int \frac{(b_n u)^r}{r!} P_1^{(r)}(Y_l) + \frac{(b_n u)^{r+1}}{(r+1)!} P_1^{(r+1)}(Y_l^*) \right] \Omega(u) du \\ &\quad + \frac{1}{n} \sum_{l=1}^n \varsigma_l (\delta_l - \mu(Y_l)) \int \frac{I(Y_l + b_n u < Y_i) - I(Y_l < Y_i)}{\pi(X_j + b_n u) (1 - H(Y_l + b_n u))} \Omega(u) du, \end{aligned} \quad (\text{A.3})$$

with $\tilde{H}_1(t) = P(Y > t, \delta = 1)$, $H_n = (1/n) \sum_{i=1}^n I(Y \leq t)$, $\Xi(Y_i) = \int_0^{Y_i} (dG(t)/(1 - G(t)))$, $\Xi_n(Y_i) = \int_0^{Y_i} ((1 - \hat{\mu}_n(t))/(1 - H_n(t))) dH_n(t)$, $V_n^*(Y_j) = (1/na_n) \sum_{i=1}^n \xi_i L((Y_j - Y_i)/a_n)$, $V^*(Y_j) = \pi(Y_j)h(Y_j)$ and $P_1(x) = [\pi(x)(1 - H(x))]^{-1}$. $\Xi^*(Y_i)$ is between $\Xi_n(Y_i)$ and $\Xi(Y_i)$. $\Xi^{**}(Y_i)$ is between $-\ln(1 - \hat{G}_n(Y_i))$ and $\Xi_n(Y_i)$. Y_j^* is between Y_j and $Y_j + b_n u$.

Proof. Following the proof of Theorem 1 in Wang and Ng [28], one can get the proof of Lemma A.4. To save space, here, we omit the details. $\square$

Lemma A.5. Suppose that assumptions (C1)–(C10) hold; then, as $n \rightarrow \infty$, we have $\sup_{u \in \mathcal{U}} \|\tilde{\alpha}_n(u) - \alpha(u)\| = O_p(\gamma_n)$ and $\sup_{u \in \mathcal{U}} \|\tilde{\alpha}_n(u) - \alpha(u)\| = O_p(\gamma_n)$.

Proof. We only prove the results about $\tilde{\alpha}_c(u)$. The results related on $\tilde{\alpha}_m(u)$ and $\tilde{\alpha}_w(u)$ can be proved similarly. By the definition of $\tilde{\alpha}_c(u)$ defined as (13) in Section 2, we can write

$$\begin{aligned} \tilde{\alpha}_c(u) - \alpha(u) &= (I_q 0) \{D_u^\top \omega_u \hat{\Delta}_c D_u\}^{-1} D_u^\top \hat{\Delta}_c \omega_u (M - Z\alpha(u)) \\ &\quad + (I_q 0) \{D_u^\top \omega_u \hat{\Delta}_c D_u\}^{-1} D_u^\top \hat{\Delta}_c \omega_u \varepsilon \\ &\quad - (I_q 0) \{D_u^\top \omega_u \hat{\Delta}_c D_u\}^{-1} D_u^\top \hat{\Delta}_c \omega_u e^\top \beta \\ &\quad + (I_q 0) \{D_u^\top \omega_u \hat{\Delta}_c D_u\}^{-1} D_u^\top \hat{\Delta}_c \omega_u (\mathbf{X}(\tilde{\beta}_c - \beta) \\ &\quad + e^\top (\tilde{\beta}_c - \beta)) \\ &:= I_1 + I_2 + I_3 + I_4. \end{aligned} \quad (\text{A.4})$$

Standard arguments yield that $I_1 = 1_{q \times 1} O_p(h_{1n}^2)$. Following (A.9) in Shen et al. [10], we have $I_2 = 1_{q \times 1} O_p(\rho_n)$. Similarly, $I_3 = 1_{q \times 1} O_p(\rho_n)$. Note that $\|\tilde{\beta}_c - \beta\| = O_p(n^{-1/2})$, it is easy to prove that $I_4 = 1_{q \times 1} O_p(n^{-1/2})$. Therefore, the proof of $\sup_{u \in \mathcal{U}} \|\tilde{\alpha}_c(u) - \alpha(u)\| = O_p(\gamma_n)$ is finished. Using similar arguments, the verification of $\sup_{u \in \mathcal{U}} \|\tilde{\alpha}_n(u) - \alpha(u)\| = O_p(\gamma_n)$ can be done. $\square$

Lemma A.6. Suppose that assumptions (C1)–(C10) hold, and β is the true value; then, as $n \rightarrow \infty$, we have

$$\begin{aligned} (a) \quad & \frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \xrightarrow{D} N(0, \Sigma_{j1}), \frac{1}{n} \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \hat{\eta}_{i,n}^\top(\beta) \Sigma_{j1}, \frac{1}{n} \sum_{i=1}^n \frac{\partial \hat{\eta}_{i,n}(\beta)}{\partial \beta} - \Sigma_{j2}, \\ (b) \quad & \max_{1 \leq i \leq n} \|\hat{\eta}_{i,n}(\beta)\| = o_p(n^{1/2}), \frac{1}{n} \sum_{i=1}^n \|\hat{\eta}_{i,n}(\beta)\|^3 = o_p(n^{1/2}), \|\lambda\| = O_p(n^{-1/2}), \end{aligned} \quad (\text{A.5})$$

where $\hat{\eta}_{i,n}(\beta)$ is taken to be one of $\hat{\eta}_{i,c}(\beta)$, $\hat{\eta}_{i,m}(\beta)$ and $\hat{\eta}_{i,w}(\beta)$. The values $j = 1, 2, 3$ correspond to $\hat{\eta}_{i,c}(\beta)$, $\hat{\eta}_{i,m}(\beta)$ and $\hat{\eta}_{i,w}(\beta)$, respectively.

Proof. We only prove the result related to $\hat{\eta}_{i,c}(\beta)$. The proofs of $\hat{\eta}_{i,m}(\beta)$ and $\hat{\eta}_{i,w}(\beta)$ can be confirmed similarly. The verification of (b) can be done by Lemma 3 given in Owen [29]. Here, we omit the details. Consider

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\eta}_{i,c}(\beta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \frac{m(T_i, \theta_0)}{1 - G(Y_i)} + \frac{m(T_i, \hat{\theta}_n) - m(T_i, \theta_0)}{1 - G(Y_i)} + \frac{m(T_i, \theta_0)(\hat{G}_n(Y_i) - G(Y_i))}{(1 - \hat{G}_n(Y_i))(1 - G(Y_i))} \right. \\ &\quad \left. + \frac{((m(T_i, \hat{\theta}_n) - m(T_i, \theta_0))(\hat{G}_n(Y_i) - G(Y_i)))}{(1 - G(Y_i))(1 - \hat{G}_n(Y_i))} \right\} \hat{\sigma}_c^{-2}(U_i) [\tilde{\xi}_i(Y_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta] \\ &:= D_1 + D_2 + D_3 + D_4. \end{aligned} \quad (\text{A.6})$$

From Theorem 3, we have

$$\begin{aligned} D_1 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \sigma^{-2}(U_i) \tilde{\xi}_i(M_i - S_i M) + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \sigma^{-2}(U_i) \tilde{\xi}_i(\varepsilon_i - S_i \varepsilon) \\ &\quad - \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \sigma^{-2}(U_i) \tilde{\xi}_i(e_i - S_i e) \beta + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \sigma^{-2}(U_i) \Sigma_e \beta + o_p(1) \\ &:= D_{11} + D_{12} + D_{13} + D_{14} + o_p(1). \end{aligned} \quad (\text{A.7})$$

Since $M_i - S_i M = Z_i^\top \alpha(U_i) O_p(\gamma_n)$. Under assumption (C10), we have $D_{11} = o_p(1)$. Under assumption (C5), we have $S_i \varepsilon = Z_i^\top 1_q O_p(\log n/nh_{1n})$. On applying assumption (C10), one can get

$$D_{12} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} [\xi_i - \Phi(U_i) \Gamma^{-1}(U_i) \Phi^\top(U_i)] \varepsilon_i + o_p(1). \quad (\text{A.8})$$

Applying $S_i e = Z_i \Gamma^{-1}(U_i) E[Z_i e_i^\top | U_i] (1 + O_p(\gamma_n)) = 0$, we have

$$D_{13} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} [\xi_i - \Phi(U_i) \Gamma^{-1}(U_i) \Phi^\top(U_i)] e_i \beta + o_p(1). \quad (\text{A.9})$$

Recalling Remark 1 and from the results in Theorem 3, it is easy to prove

$$\begin{aligned} D_2 &= \frac{1}{n\sqrt{n}} \Gamma^{-1}(\theta_0) \sum_{i=1}^n \sum_{j=1}^n \frac{\nabla^\top m(T_i, \theta_0)}{1 - G(Y_i)} \cdot \frac{\xi_j (\delta_j - m(T_j, \theta_0)) \nabla m(T_j, \theta_0)}{m(T_j, \theta_0) \{1 - m(T_j, \theta_0)\}} \sigma^{-2}(U_i) \\ &\quad \times [\tilde{\xi}_i (M_i - S_i M) + \tilde{\xi}_i (\varepsilon_i - S_i \varepsilon) + \tilde{\xi}_i (e_i - S_i e) \beta + \Sigma_e \beta] + o_p(1) \\ &:= D_{21} + D_{22} + D_{23} + D_{24} + o_p(1). \end{aligned} \quad (\text{A.10})$$

Under the missing mechanism and similar to the proof of D_1 , it is easy to prove that $D_{2l} = o_p(1)$ for $l = 1, 2, 3, 4$. Hence, we have $D_2 = o_p(1)$. Consider

$$\begin{aligned} D_3 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0) (\hat{G}_n(Y_i) - G(Y_i))}{(1 - G(Y_i))^2} \sigma^{-2}(U_i) [\tilde{\xi}_i (Y_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta] \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0) (\hat{G}_n(Y_i) - G(Y_i))}{(1 - G(Y_i))^2} [\hat{\sigma}_c^{-2}(U_i) - \sigma^{-2}(U_i)] [\tilde{\xi}_i (Y_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta] \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0) (\hat{G}_n(Y_i) - G(Y_i))^2}{(1 - \hat{G}_n(Y_i))^2 (1 - G(Y_i))} \hat{\sigma}_c^{-2}(U_i) [\tilde{\xi}_i (Y_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta] \\ &:= D_{31} + D_{32} + D_{33}. \end{aligned} \quad (\text{A.11})$$

From Lemma A.4, we have

$$\begin{aligned} D_{31} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{(1 - G(Y_i))} \frac{1}{n} \sum_{j=1}^n \zeta(Y_j, \delta_j, \varsigma_j; Y_i) \sigma^{-2}(U_i) \tilde{\xi}_i (M_i - S_i M) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{(1 - G(Y_i))} \frac{1}{n} \sum_{j=1}^n \zeta(Y_j, \delta_j, \varsigma_j; Y_i) \sigma^{-2}(U_i) \tilde{\xi}_i (\varepsilon_i - S_i \varepsilon) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{(1 - G(Y_i))} \frac{1}{n} \sum_{j=1}^n \zeta(Y_j, \delta_j, \varsigma_j; Y_i) \sigma^{-2}(U_i) \tilde{\xi}_i (e_i - S_i e) \beta \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{(1 - G(Y_i))} \frac{1}{n} \sum_{j=1}^n \zeta(Y_j, \delta_j, \varsigma_j; Y_i) \sigma^{-2}(U_i) \Sigma_e \beta \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{(1 - G(Y_i))^2} R_n(Y_i) \sigma^{-2}(U_i) [\tilde{\xi}_i (Y_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta] \\ &:= D_{311} + D_{312} + D_{313} + D_{314} + D_{315}. \end{aligned} \quad (\text{A.12})$$

It is easy to prove $M_i - S_i M = Z_i^\top \alpha(U_i) O_p(\gamma_n)$, and we have $D_{311} = o_p(1)$ under assumption (C10).

$$\begin{aligned} E[D_{312}]^2 &= \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n E \left\{ \frac{m(T_i, \theta_0)}{(1 - G(Y_i))} \zeta(Y_j, \delta_j, \varsigma_j; Y_i) \sigma^{-2}(U_i) \tilde{\xi}_i(\varepsilon_i - S_i \varepsilon) \right\}^2 \\ &\leq \frac{C}{n} E \left\{ \frac{m(T_1, \theta_0)}{(1 - G(Y_1))} \zeta(Y_2, \delta_2, \varsigma_2; Y_1) \sigma^{-2}(U_1) \tilde{\xi}_1(\varepsilon_1 - S_1 \varepsilon) \right\}^2 = O_p(n^{-1}). \end{aligned} \quad (\text{A.13})$$

Hence, $D_{312} = o_p(1)$. Similarly, we have $D_{313} = o_p(1)$ and $D_{314} = o_p(1)$. Compared with D_{311} , D_{315} is far smaller than D_{311} , and then $D_{315} = o_p(1)$. From Theorem 3, we have $D_{32} = o_p(1)$. Similarly, $D_{33} = o_p(1)$. Hence, we have $D_3 = o_p(1)$. Note that $m(T_i, \hat{\theta}_n) - m(T_i, \theta_0) = \nabla^\top m(T_i, \theta_0)(\hat{\theta}_n - \theta_0)(1 + o(1))$. From assumption (C9), we get

$$\max_{1 \leq i \leq n} |m(T_i, \hat{\theta}_n) - m(T_i, \theta_0)| = O_p(n^{-1/2}). \quad (\text{A.14})$$

Thus, it can be checked that $D_4 = o_p(1)$. Hence, combining with D_{14} and collecting the results above, one can obtain

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\eta}_{i,c}(\beta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \left\{ [\xi_i - \Phi(U_i) \Gamma^{-1} \right. \\ &\quad \cdot (U_i) \Phi^\top(U_i)] (\varepsilon_i - e_i \beta) + \Sigma_e \beta \left. \right\} + o_p(1). \end{aligned} \quad (\text{A.15})$$

By the central limit theorem, we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\eta}_{i,c}(\beta) \xrightarrow{D} N(0, \Sigma_{11}). \quad (\text{A.16})$$

It follows from the law of large numbers that

$$\frac{1}{n} \sum_{i=1}^n \hat{\eta}_{i,c}(\beta) \hat{\eta}_{i,c}^\top(\beta) \Sigma_{11}. \quad (\text{A.17})$$

Consider the partial derivative of $\hat{\eta}_{i,c}(\beta)$ with respect to β , and then

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \frac{\partial \hat{\eta}_{i,c}(\beta)}{\partial \beta} &= -\frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \sigma^{-2}(u) [X_i - Z_i^\top \Gamma^{-1} \\ &\quad \cdot (U_i) \Phi(U_i)]^\top [X_i - Z_i^\top \Gamma^{-1} (U_i) \Phi(U_i)] + o_p(1). \end{aligned} \quad (\text{A.18})$$

By the law of large numbers, we have

$$\frac{1}{n} \sum_{i=1}^n \frac{\partial \hat{\eta}_{i,c}(\beta)}{\partial \beta} = \Sigma_{12}. \quad (\text{A.19})$$

Therefore, the verification of Lemma A.6 is finished. $\square$

Proof. of Theorem 1. We prove Theorem 1 for $\tilde{\beta}_c$. The verifications related to $\tilde{\beta}_m$ and $\tilde{\beta}_w$ can be obtained similarly. Recalling the definition of $\tilde{\beta}_c$ in Section 2, one can write $\sqrt{n}(\tilde{\beta}_c - \beta) = \Sigma_c^{-1} A_c$, where

$$\Sigma_c = \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \hat{\theta}_n)}{1 - \hat{G}_n(Y_i)} \left(\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e \right), \quad (\text{A.20})$$

$$A_c = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \hat{\theta}_n)}{1 - \hat{G}_n(Y_i)} \left[\tilde{\xi}_i (Y_i - \tilde{\xi}_i^\top \beta) + \Sigma_e \beta \right]. \quad (\text{A.21})$$

Replace $\hat{\theta}_n$ and $\hat{G}_n(Y_i)$ with their true values, and then Σ_c can be rewritten as

$$\begin{aligned} \Sigma_c &= \frac{1}{n} \sum_{i=1}^n \left\{ \frac{m(T_i, \theta_0)}{1 - G(Y_i)} + \frac{m(T_i, \hat{\theta}_n) - m(T_i, \theta_0)}{1 - G(Y_i)} + \frac{m(T_i, \theta_0)(\hat{G}_n(Y_i) - G(Y_i))}{(1 - \hat{G}_n(Y_i))(1 - G(Y_i))} \right. \\ &\quad \left. + \frac{(m(T_i, \hat{\theta}_n) - m(T_i, \theta_0))(\hat{G}_n(Y_i) - G(Y_i))}{(1 - G(Y_i))(1 - \hat{G}_n(Y_i))} \right\} \left(\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e \right) \\ &:= \Sigma_{c1} + \Sigma_{c2} + \Sigma_{c3} + \Sigma_{c4}. \end{aligned} \quad (\text{A.22})$$

Recalling the definition of $\tilde{\xi}_i$ in Section 2, then we have

$$\begin{aligned}\Sigma_{c1} &= \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} (X_i - S_i \mathbf{X})^\top (X_i - S_i \mathbf{X}) + \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} (e_i - S_i e)^\top (X_i - S_i \mathbf{X}) \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} (X_i - S_i \mathbf{X})^\top (e_i - S_i e) + \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} [(e_i - S_i e)^\top (e_i - S_i e) - \Sigma_e] \\ &:= \Sigma_{c11} + \Sigma_{c12} + \Sigma_{c13} + \Sigma_{c14}.\end{aligned}\tag{A.23}$$

Standard computations yield that $S_i \mathbf{X} = Z_i \Gamma^{-1}(U_i) \Phi(U_i) (1 + O_p(\gamma_n))$; then, it is easy to calculate

$$\Sigma_{c11} = \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} [X_i - Z_i^\top \Gamma^{-1}(U_i) \Phi(U_i)]^\top [X_i - Z_i^\top \Gamma^{-1}(U_i) \Phi(U_i)] + o_p(1).\tag{A.24}$$

Note that $S_i e = Z_i \Gamma^{-1}(U_i) E[Z_i e_i^\top | U_i] (1 + O_p(\gamma_n)) = 0$. Hence, we have $\Sigma_{c12} = o_p(1)$, $\Sigma_{c13} = o_p(1)$ and

$\Sigma_{c14} = (1/n) \sum_{i=1}^n (m(T_i, \theta_0)/1 - G(Y_i)) (e_i e_i^\top - \Sigma_e) = 0$. Thus, it can be concluded that

$$\Sigma_{c1} = \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} [X_i - Z_i^\top \Gamma^{-1}(U_i) \Phi(U_i)]^\top [X_i - Z_i^\top \Gamma^{-1}(U_i) \Phi(U_i)] + o_p(1).\tag{A.25}$$

Note that $\max_{1 \leq i \leq n} |m(T_i, \hat{\theta}_n) - m(T_i, \theta_0)| = O_p(n^{-1/2})$, which indicates that $\Sigma_{c2} = o_p(1)$. Consider

$$\Sigma_{c3} = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{m(T_i, \theta_0)(\hat{G}_n(Y_i) - G(Y_i))}{(1 - G(Y_i))^2} + \frac{m(T_i, \theta_0)(\hat{G}_n(Y_i) - G(Y_i))^2}{(1 - G(Y_i))^2(1 - \hat{G}_n(Y_i))} \right\} (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) := \Sigma_{c31} + \Sigma_{c32}.\tag{A.26}$$

In terms of the expansion of $\hat{G}_n(Y_i) - G(Y_i)$ in Lemma A.4, one can get

$$\Sigma_{c31} = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{1}{n} \sum_{j=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \zeta(\tilde{Y}_j, \delta_j, \xi_j; \tilde{Y}_i) + R_n(\tilde{Y}_i) \right\} (\tilde{\xi}_i \tilde{\xi}_i^\top - \Sigma_e) := \Sigma_{c311} + \Sigma_{c312}.\tag{A.27}$$

Note that $E[\zeta(Y_j, \delta_j, \xi_j; Y_i) | Y_i] = 0$. Based on the independence of $\{Y_j, \delta_j, \xi_j\}$, then we have $E[\Sigma_{c311}]^2 = O_p(n^{-1})$. Hence, it holds that $\Sigma_{c311} = o_p(n^{-1/2})$. Compared with Σ_{c311} , Σ_{c312} is far smaller than Σ_{c311} , and then

$\Sigma_{c312} = o_p(1)$. Similarly, $\Sigma_{c31} = o_p(1)$ and $\Sigma_{c32} = o_p(1)$. Hence, we have $\Sigma_{c3} = o_p(1)$. Following (A.20) in Lemma A.4, it is easy to prove $\Sigma_{c4} = o_p(1)$. Collecting the results above, we have

$$\Sigma_c = \frac{1}{n} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} [X_i - Z_i^\top \Gamma^{-1}(U_i) \Phi(U_i)]^\top [X_i - Z_i^\top \Gamma^{-1}(U_i) \Phi(U_i)] + o_p(1).\tag{A.28}$$

Analogous to the arguments as the proof of Lemma A.6 (a), then one can obtain

$$A_c = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \{ [\xi_i - \Phi(U_i) \Gamma^{-1}(U_i) \Phi^\top(U_i)] (\varepsilon_i - e_i \beta) + \Sigma_e \beta \} + o_p(1). \quad (\text{A.29})$$

Collecting the results above, the proof of Theorem 1 is completed. $\square$

Proof. of Theorem 2. We only prove the results about $\tilde{\alpha}_c(u)$. The results related on $\tilde{\alpha}_m(u)$ and $\tilde{\alpha}_w(u)$ can be proved similarly. By the definition of $\tilde{\alpha}_c(u)$, we can write

$$\begin{aligned} \tilde{\alpha}_c(u) - \alpha(u) &= (I_q \ 0) \{ D_u^\top \omega_u \Delta D_u \}^{-1} D_u^\top \Delta \omega_u (M - \mathbf{Z} \alpha(u)) \\ &+ (I_q \ 0) \{ D_u^\top \omega_u \Delta D_u \}^{-1} D_u^\top \Delta \omega_u (\varepsilon - e^\top \beta) \\ &+ (I_q \ 0) \{ D_u^\top \omega_u \Delta D_u \}^{-1} D_u^\top \Delta \omega_u (\mathbf{X}(\beta - \tilde{\beta}_c) + e^\top (\beta - \tilde{\beta}_c)) \\ &+ (I_q \ 0) \{ D_u^\top \omega_u \Delta D_u \}^{-1} D_u^\top (\hat{\Delta}_c - \Delta) \omega_u (M + \varepsilon - e^\top \tilde{\beta}_c + \mathbf{X}(\tilde{\beta}_c - \beta)) + O_p(n^{-1/2}) \\ &:= B_1 + B_2 + B_3 + B_4. \end{aligned} \quad (\text{A.30})$$

By Taylor expansion, it can be checked that

$$M = \begin{pmatrix} Z_1^\top \alpha(u) + (U_1 - u) Z_1^\top \alpha'(u) + \frac{1}{2} (U_1 - u)^2 Z_1^\top \alpha''(u) \\ \vdots \\ Z_n^\top \alpha(u) + (U_n - u) Z_n^\top \alpha'(u) + \frac{1}{2} (U_n - u)^2 Z_n^\top \alpha''(u) \end{pmatrix} + o_p(h_{1n}^2). \quad (\text{A.31})$$

Hence, from assumption (C5), we have

$$B_1 = \frac{1}{2} h_{1n}^2 \kappa_2 \alpha''(u) + o_p(h_{1n}^2). \quad (\text{A.32})$$

One direct simplification implies

$$B_2 = g^{-1}(u) \Gamma^{-1}(u) \frac{1}{n h_{1n}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} Z_i K\left(\frac{U_i - u}{h_{1n}}\right) (\varepsilon_i - e_i^\top \beta). \quad (\text{A.33})$$

Theorem 1 implies that $\|\tilde{\beta}_c - \beta\| = O_p(n^{-1/2})$. It is easy to verify that $B_3 = O_p(n^{-1/2})$ and $B_4 = O_p(n^{-1/2+1/2s})$. Hence, we have

$$\begin{aligned} &\sqrt{n h_{1n}} \left(\tilde{\alpha}_c(u) - \alpha(u) - \frac{1}{2} h_{1n}^2 \kappa_2 \alpha''(u) \right) \\ &= g^{-1}(u) \Gamma^{-1}(u) \frac{1}{\sqrt{n h_{1n}}} \sum_{i=1}^n \frac{m(T_i, \theta_0)}{1 - G(Y_i)} \end{aligned} \quad (\text{A.34})$$

$$Z_i K\left(\frac{U_i - u}{h_{1n}}\right) (\varepsilon_i - e_i^\top \beta) + o_p(1).$$

By Slutsky's theorem, we finish the proof of Theorem 2. $\square$

Proof. of Theorem 3. We only prove the results about $\hat{\sigma}_c^2(\cdot)$. The results related on $\hat{\sigma}_m^2(\cdot)$ and $\hat{\sigma}_w^2(\cdot)$ can be proved similarly. To save space, here, we omit the details. Denote $M_i = Z_i \alpha(U_i)$ and $\tilde{M}_i = Z_i \tilde{\alpha}_c(U_i)$. Note that

$$\begin{aligned}
\hat{\sigma}_c^2(u) - \sigma^2(u) &= \left\{ \sum_{i=1}^n \varepsilon_i^\top W_{ni}^c(u) \varepsilon_i - \sigma^2(u) \right\} - 2 \sum_{i=1}^n [(M_i - \tilde{M}_i)]^\top W_{ni}^c(u) \varepsilon_i \\
&\quad - 2 \sum_{i=1}^n [X_i(\tilde{\beta}_c - \beta)]^\top W_{ni}^c(u) \varepsilon_i + \sum_{i=1}^n [X_i(\tilde{\beta}_c - \beta)]^\top W_{ni}^c(u) [X_i(\tilde{\beta}_c - \beta)] \\
&\quad + \sum_{i=1}^n (M_i - \tilde{M}_i)^\top W_{ni}^c(u) [X_i(\tilde{\beta}_c - \beta)] + \sum_{i=1}^n (M_i - \tilde{M}_i)^\top W_{ni}^c(u) (M_i - \tilde{M}_i) \\
&\quad + \left\{ \sum_{i=1}^n (e\tilde{\beta}_c)^\top W_{ni}^c(u) (e\tilde{\beta}_c) - \tilde{\beta}_c^\top \Sigma_e \tilde{\beta}_c \right\} := \sum_{l=1}^9 Q_l. \tag{A.35}
\end{aligned}$$

Following Lemma 4.1 in Chiou [30], it is easy to prove that $\sup_l |\sum_{i=1}^n \varepsilon_i^\top W_{ni}(u) \varepsilon_i - \sigma^2(u)| = O_p(\rho_n)$. From Theorem 3.5 in You and Chen [1], we have $\sup_{u \in \mathcal{U}} \|\tilde{\alpha}_c(u) - \alpha(u)\| = O_p(\gamma_n)$, which indicates that $M_i - \tilde{M}_i = O_p(\rho_n)$. Based on the fact that $\|\tilde{\beta}_c - \beta\| = O_p(n^{-1/2})$, it is easy to prove $Q_i = O_p(\rho_n)$ for $i = 2, 3, \dots, 8$. In terms of Lemma A.2, one can obtain $\sup_u |Q_9| = O_p(\rho_n)$, which completes the proof of Theorem 3. $\square$

Proof. of Theorem 4. Based on Theorem 3, similar to the arguments, the proof of Theorems 1 and 4 can be verified easily. $\square$

Proof. of Theorem 5. Theorem 1 implies that $\|\hat{\beta}_n - \beta\| = O_p(n^{-1/2})$ and $\sup_{u \in \mathcal{U}} \|\hat{\alpha}_n(u) - \alpha(u)\| = O_p(\gamma_n)$. Analogous to the proof of Theorem 3, it is easy to verify Theorem 5. $\square$

Proof. of Theorem 6. Applying the Taylor expansion for the empirical log-likelihood ratio function, then we have

$$\hat{l}_n(\beta) = 2 \sum_{i=1}^n \left\{ \lambda^\top \hat{\eta}_{i,n}(\beta) - \frac{1}{2} [\lambda^\top \hat{\eta}_{i,n}(\beta)]^2 \right\} + o_p(1). \tag{A.36}$$

Note that λ is the solution of equation $\sum_{i=1}^n (\hat{\eta}_{i,n}(\beta) / (1 + \lambda^\top \hat{\eta}_{i,n}(\beta))) = 0$, and then one can get

$$\begin{aligned}
0 &= \sum_{i=1}^n \frac{\hat{\eta}_{i,n}(\beta)}{1 + \lambda^\top \hat{\eta}_{i,n}(\beta)} = \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) - \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \\
&\quad \cdot \sum_{i=1}^n \hat{\eta}_{i,n}^\top(\beta) \lambda + \sum_{i=1}^n \frac{\hat{\eta}_{i,n}(\beta) \{\lambda^\top \hat{\eta}_{i,n}(\beta)\}}{1 + \lambda^\top \hat{\eta}_{i,n}(\beta)}. \tag{A.37}
\end{aligned}$$

Similar to Owen [20], we derive that $\|\lambda\| = O_p(n^{-1/2})$. Hence, we have

$$\lambda = \left\{ \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \hat{\eta}_{i,n}^\top(\beta) \right\}^{-1} \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) + o_p(n^{-1/2}). \tag{A.38}$$

Hence, from (A.36) and (A.38), it can be checked that

$$\hat{l}_n(\beta) = \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \right\}^\top \left\{ \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \hat{\eta}_{i,n}^\top(\beta) \right\}^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \right\} + o_p(1), \tag{A.39}$$

in which, together with Lemma A.6, the proof of Theorem 6 is finished. $\square$

Proof. of Theorem 7. For convenience, let

$$\begin{aligned}
\Lambda_{1n}(\beta, \lambda) &= \frac{1}{n} \sum_{i=1}^n \frac{\hat{\eta}_{i,n}(\beta)}{1 + \lambda^\top \hat{\eta}_{i,n}(\beta)}, \\
\Lambda_{2n}(\beta, \lambda) &= \frac{1}{n} \sum_{i=1}^n \frac{1}{1 + \lambda^\top \hat{\eta}_{i,n}(\beta)} \left(\frac{\partial \hat{\eta}_{i,n}(\beta)}{\partial \beta^\top} \right) \lambda. \tag{A.40}
\end{aligned}$$

Note that β^* and $\lambda^* = \lambda(\beta^*)$ satisfy $\Lambda_{1n}(\beta^*, \lambda^*) = 0$ and $\Lambda_{2n}(\beta^*, \lambda^*) = 0$. By expanding $\Lambda_{jn}(\beta^*, \lambda^*) = 0$ at $(\beta^\top, 0)^\top$ for $j = 1, 2$, it can be shown that

$$0 = \Lambda_{jn}(\beta^*, \lambda^*) = \Lambda_{jn}(\beta, 0) + \frac{\partial \Lambda_{jn}(\beta, 0)}{\partial \beta} (\beta^* - \beta) + \frac{\partial \Lambda_{jn}(\beta, 0)}{\partial \lambda} \lambda^* + o_p(\|\beta^* - \beta\|). \tag{A.41}$$

where $\|\beta_n\| = \|\beta^* - \beta\| + \|\lambda^*\|$. Hence,

$$\begin{pmatrix} \lambda^* \\ \beta^* - \beta \end{pmatrix} = \begin{pmatrix} -\frac{1}{n} \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) \hat{\eta}_{i,n}^\top(\beta) & \frac{1}{n} \sum_{i=1}^n \frac{\partial \hat{\eta}_{i,n}(\beta)}{\partial \beta} \\ \frac{1}{n} \sum_{i=1}^n \frac{\partial \hat{\eta}_{i,n}(\beta)}{\partial \beta} & 0 \end{pmatrix}^{-1} \begin{pmatrix} -\frac{1}{n} \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) + o_p(\|\cdot\|_n) \\ o_p(\|\cdot\|_n) \end{pmatrix}. \quad (\text{A.42})$$

From Lemma A.6, we have $(1/n) \sum_{i=1}^n \hat{\eta}_{i,n}(\beta) = O_p(n^{-1/2})$, $\|\eta_n\| = O_p(n^{-1/2})$ and $(1/n) \sum_{i=1}^n (\partial \hat{\eta}_{i,n}(\beta) / \partial \beta) = -\Sigma_{12}(1 + o_p(1))$ which, together with Lemma A.6, shows the proof of Theorem 7. $\square$

Data Availability

The simulation study is based on the Monte Carlo simulation to study the finite sample performance of the proposed estimators. The real dataset is in [25]. Clinical trial E1178 conducted by the Eastern Cooperative Oncology Group compared tamoxifen therapy and placebo in elderly (≥ 65) women with stage II breast cancer.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the China Postdoctoral Science Foundation (2019M651422), the National Natural Science Foundation of China (11831008 and 11971171), the National Social Science Foundation Key Program (17ZDA091), the 111 Project of China (B14019), and the Natural Science Foundation of Shanghai (17ZR1409000).

References

- [1] J. You and G. Chen, "Estimation of a semiparametric varying-coefficient partially linear errors-in-variables model," *Journal of Multivariate Analysis*, vol. 97, no. 2, pp. 324–341, 2006.
- [2] X. He, X. Feng, X. Tong, and X. Zhao, "Semiparametric partially linear varying coefficient models with panel count data," *Lifetime Data Analysis*, vol. 23, no. 3, pp. 439–466, 2017.
- [3] B. Kai, R. Li, and H. Zou, "New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models," *Annals of Statistics*, vol. 39, no. 1, pp. 305–332, 2011.
- [4] R. F. Engle, C. W. J. Granger, J. Rice, and A. Weiss, "Semiparametric estimates of the relation between weather and electricity sales," *Journal of the American Statistical Association*, vol. 81, no. 394, pp. 310–320, 1986.
- [5] H. Liang, W. Härdle, and R. J. Carroll, "Estimation in a semiparametric partially linear errors-in-variables model," *The Annals of Statistics*, vol. 27, pp. 1519–1535, 1999.
- [6] A. A. Liu and H. Y. Liang, "Jackknife empirical likelihood of error variance in partially linear varying-coefficient errors-in-variables models," *Statistical Papers*, vol. 58, no. 1, pp. 1–28, 2017.
- [7] G.-l. Fan, H.-y. Liang, and L.-x. Zhu, "Penalized profile least squares-based statistical inference for varying coefficient partially linear errors-in-variables models," *Science China Mathematics*, vol. 61, no. 9, pp. 1677–1694, 2018.
- [8] J. You, G. Chen, and Y. Zhou, "Statistical inference of partially linear regression models with heteroscedastic errors," *Journal of Multivariate Analysis*, vol. 98, no. 8, pp. 1539–1557, 2007.
- [9] G.-L. Fan, H.-Y. Liang, and J.-F. Wang, "Statistical inference for partially time-varying coefficient errors-in-variables models," *Journal of Statistical Planning and Inference*, vol. 143, no. 3, pp. 505–519, 2013.
- [10] S.-L. Shen, J.-L. Cui, C.-L. Mei, and C.-W. Wang, "Estimation and inference of semi-varying coefficient models with heteroscedastic errors," *Journal of Multivariate Analysis*, vol. 124, pp. 70–93, 2014.
- [11] H. Xu and X. Duan, "Efficient estimation for partially linear varying-coefficient errors-in-variables models with heteroscedastic errors," *Instrumentation Mesure Métrologie*, vol. 18, no. 2, pp. 295–314, 2018.
- [12] Z. Huang, "Empirical likelihood for single-index varying-coefficient models with right-censored data," *Journal of the Korean Statistical Society*, vol. 39, no. 4, pp. 533–544, 2010.
- [13] Z. Huang, "Empirical likelihood for a partially linear single-index measurement error model with right-censored data," *Communications in Statistics - Theory and Methods*, vol. 40, no. 6, pp. 1015–1029, 2011.
- [14] Q. Wang and J. Shen, "Estimation and confidence bands of a conditional survival function with censoring indicators missing at random," *Journal of Multivariate Analysis*, vol. 99, no. 5, pp. 928–948, 2008.
- [15] R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*, John Wiley & Sons, New York, NY, USA, 1987.
- [16] Q. Wang and G. E. Dinse, "Linear regression analysis of survival data with missing censoring indicators," *Lifetime Data Analysis*, vol. 17, no. 2, pp. 256–279, 2011.
- [17] X. Li and Q. Wang, "The weighted least square based estimators with censoring indicators missing at random," *Journal of Statistical Planning and Inference*, vol. 142, no. 11, pp. 2913–2925, 2012.
- [18] Y. Shen and H.-Y. Liang, "Quantile regression for partially linear varying-coefficient model with censoring indicators missing at random," *Computational Statistics & Data Analysis*, vol. 117, pp. 1–18, 2018.
- [19] J. F. Wang, W. J. Jiang, F. Y. Xu, and W. X. Fu, "Weighted composite quantile regression with censoring indicators missing at random," *Communications in Statistics-Theory and Methods*, vol. 51, 2019.
- [20] A. B. Owen, "Empirical likelihood ratio confidence regions," *The Annals of Statistics*, vol. 18, no. 1, pp. 90–120, 1990.
- [21] G.-L. Fan, H.-Y. Liang, and Y. Shen, "Penalized empirical likelihood for high-dimensional partially linear varying coefficient model with measurement errors," *Journal of Multivariate Analysis*, vol. 147, pp. 183–201, 2016.
- [22] Y. S. Wang and M. Drton, "Empirical likelihood for linear structural equation models with dependent errors," *Stat*, vol. 6, no. 1, pp. 434–447, 2017.

- [23] G. L. Fan, L. L. Wang, and H. X. Xu, "Weighted empirical likelihood for heteroscedastic varying coefficient partially non-linear models with missing data," *Stat*, vol. 10, no. 1, 2021.
- [24] Y. Zou, G. Fan, and R. Zhang, "Empirical likelihood and variable selection for partially linear single-index EV models with missing censoring indicators," *Journal of the Korean Statistical Society*, vol. 50, no. 1, pp. 134–162, 2021.
- [25] F. J. Cummings, R. Gray, T. E. Davis et al., "Tamoxifen versus placebo: Double-blind adjuvant trial in elderly women with stage ii breast cancer," *NCI Monographs*, vol. 1, pp. 119–123, 1986.
- [26] Y. P. Mack and B. W. Silverman, "Weak and strong uniform consistency of kernel regression estimates," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 61, no. 3, pp. 405–415, 1982.
- [27] Y. C. Xia and W. K. Li, "On the estimation and testing of functional-coefficient linear models," *Statistica Sinica*, vol. 9, pp. 735–757, 1999.
- [28] Q. Wang and K. W. Ng, "Asymptotically efficient product-limit estimators with censoring indicators missing at random," *Statistic Sinica*, vol. 18, no. 2, pp. 749–768, 2008.
- [29] A. B. Owen, "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, vol. 75, no. 2, pp. 237–249, 1988.
- [30] J. M. Chiou and H. G. Müller, "Nonparametric quasi-likelihood," *The Annals of Statistics*, vol. 27, no. 1, pp. 36–64, 1999.